

Proceedings of the Augmented Visual Display (AVID) Research Workshop

*Proceedings of a workshop held at
Ames Research Center
Moffett Field, California
March 10-12, 1993*

(NASA-CP-10128) PROCEEDINGS OF THE
WORKSHOP ON AUGMENTED VISUAL
DISPLAY (AVID) RESEARCH (NASA)
428 p

N94-25490
--THRU--
N94-25514
Unclass

G3/54 0205730

National Aeronautics and
Space Administration

Ames Research Center



NASA Conference Publication 10128

249912
14P35

Proceedings of the Augmented Visual Display (AVID) Research Workshop

Edited by
Mary K. Kaiser and Barbara T. Sweet
Ames Research Center
Moffett Field, California

Proceedings of a workshop held at
Ames Research Center
Moffett Field, California
March 10-12, 1993



National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California 94035-1000

December 1993

CONTENTS

	Page
INTRODUCTION.....	vii
WORKSHOP PROGRAM.....	ix
SUMMARY.....	1
I. SENSOR SYSTEMS	
Infrared Sensors and Systems for Enhanced Vision/Autonomous Landing Applications..... <i>J. Richard Kerr</i>	5 - 1
Synthetic Vision System Flight Test Results and Lessons Learned..... <i>Jeffrey Radke</i>	29 - 2
94 GHz MMW Imaging Radar System..... <i>Yair Alon and Lon Ulmer</i>	47 - 3
Passive Millimeter-Wave Imaging..... <i>Stephen K. Young, Roger A. Davidheiser, Bruce I. Hauss, Paul S. C. Lee, Michael Mussetto, Merit M. Shoucri, & Larry Yujiri</i>	61 - 4
Application of Aircraft Navigation Sensors to Enhanced Vision Systems..... <i>Barbara T. Sweet</i>	79 - 5
II. SENSOR MODELING	
The IRGen Infrared Data Base Modeler..... <i>Uri Bernstein</i>	99 - 6
The Radar Image Generation (RIG) Model..... <i>Anthony J. Stenger</i>	121 - 7
Advanced Radiometric & Interferometric Millimeter-Wave Scene Simulations..... <i>B. I. Hauss, P. J. Moffa, W. G. Steele, H. Agravante, R. Davidheiser, T. Samec, & S. K. Young</i>	133 - 8
Simulation of a Passive Millimeter Wave Sensor..... <i>William M. Kahlbaum</i>	163 - 9

Engineering Workstation: Sensor Modeling..... <i>M. Pavel & B. Sweet</i>	181-10
III. SENSOR FUSION	
Image Fusion..... <i>M. Pavel</i>	193-11
Enhanced Image Capture Through Fusion..... <i>Peter J. Burt, Keith Hanna, & Raymond J. Kolczynski</i>	207-12
Radar E-O Image Fusion..... <i>William F. O'Neil</i>	225-13
IV. IMAGE PROCESSING: COMPUTER VISION	
Expansion-Based Passive Ranging..... <i>Yair Barniv</i>	235-14
Multiple-Camera/Motion Stereoscopy for Range Estimation in Helicopter Flight..... <i>Phillip N. Smith, Banavar Sridhar, & Raymond E. Suorsa</i>	287-15
Validation of Vision-Based Range Estimation Algorithms Using Helicopter Flight Data..... <i>Phillip N. Smith</i>	293-16
A Model-Based Approach for Detection of Objects in Low Resolution Passive-Millimeter Wave Images..... <i>Yuan-Liang Tang, Sadashiva Devadiga, Rangachar Kasturi, & Randall L. Harris, Sr.</i>	313-17
Image Processing for Flight Crew Enhanced Situational Awareness..... <i>Barry Roberts</i>	329-18
V. IMAGE PROCESSING: HUMAN VISION	
DCT Quantization Matrices Visually Optimized for Individual Images..... <i>Andrew B. Watson</i>	363-19

Extracting Heading and Temporal Range from Optic Flow: Human Performance Issues.....	379 -20
<i>Mary K. Kaiser, John A. Perrone, Leland Stone, Martin S. Banks, & James A. Crowell</i>	
Optical Information in Landing Scenes.....	397 -21
<i>Walter W. Johnson</i>	
 VI. IMAGE EVALUATION AND METRICS	
Sensor Fusion Display Evaluation Using Information Integration Models in Enhanced/Synthetic Vision Applications.....	409 -22
<i>David C. Foyle</i>	
Quality Metrics for Sensor Images.....	421 -23
<i>Al Ahumada</i>	
Evaluation of Image Quality.....	423 -24
<i>M. Pavel</i>	
APPENDIX: AVID WORKSHOP ATTENDEES.....	433

INTRODUCTION

In March 1993, NASA Ames Research Center hosted a three-day workshop covering two of the major research domains of the Augmented Visual Display (AVID) Research Program. Researchers from industry, government laboratories, and universities were brought together to discuss common interests in the areas of sensor modeling and simulation, and image processing and evaluation. The workshop attendees represented a wide range of disciplines, from sensor engineering to aerospace human factors. The panel sessions were unified by the common goal of developing systems to enhance pilots' functional vision in low-visibility and constrained-visibility conditions.

The AVID Research Program is dedicated to the support of generic research which underpins NASA's focused programs which rely on development of advanced display technologies. These include (but are not limited to) the Terminal Area Productivity Program (whose low-visibility element seeks to enable all equipped airliners to land and taxi under Category IIIA conditions at Type I facilities), and the High Speed Research Program (which seeks to enable pilots to land and perform ground operations in the absence of forward-looking windows). Because of its generic nature, the research encompassed by the AVID Research Program will also contribute to display solutions in the rotorcraft and space domains. In addition to the topics discussed at this workshop, the AVID Program supports work in display requirements and formatting, and on the systems integration/integrity issues associated with advanced displays.

It is our expectation that the AVID Program will continue to serve as the common touchstone for human-centered research on advanced display. As was clearly demonstrated in this workshop, such a program is critical for keeping industry apprised of relevant advances in the research community and, in turn, informing researchers about critical concerns and constraints of the operational community.

Mary K. Kaiser
Barbara T. Sweet
June, 1993

PRECEDING PAGE BLANK NOT FILMED

Program of the Augmented Visual Display (AVID) Research Workshop
NASA Ames Research Center (Building 262, Room 100)
March 10 - 12, 1993

Wednesday, March 10

8:30 - 9:00	Sign-in, Coffee and Donuts
9:00 - 9:15	Welcome C. Thomas Snyder, Director of Aerospace Systems, NASA Ames Research Center Mary Kaiser, Workshop Chair, NASA Ames Research Ctr.
9:15 - 11:00	Sensor Systems Panel (B. Sweet, chair) J. Richard Kerr, FLIR Systems, Inc. Jeffrey Radke, Systems Research Center, Honeywell, Inc. Yair Alon, Lear Astronics Bruce Hauss, TRW Applied Technology Division Barbara Sweet, NASA Ames Research Center
11:00 - 11:30	IR/Active Radar Modeling Uri Bernstein, Technology Services Corporation
11:45 - 12:45	Lunch (Galileo Room, Ames Cafeteria)
1:00 - 1:30	IRGen Demonstration (Room 129)
1:30 - 3:15	Sensor Modeling Panel (R. Harris, chair) Randall Harris, NASA Langley Research Center Bruce Hauss William Kahlbaum, Lockheed Engineering and Sciences Misha Pavel, WAL-NASA Ames Research Center/NYU
3:15 - 3:30	Afternoon Break
3:30 - 5:15	Sensor Fusion Panel (M. Pavel, chair) Misha Pavel Keith Hanna, David Sarnoff Research Center Bill O'Neil, Westinghouse Philip Hontalas, NASA Ames Research Center
5:30 - 8:00	Drinks and Dinner, Carlos Murphys (seating at 6:15)

Thursday, March 11

8:30 - 8:45	Coffee and Donuts
8:45 - 11:50	Image Processing: Computer Vision Panel (B. Sridhar, chair) Banavar Sridhar, NASA Ames Research Center Yair Barniv, NASA Ames Research Center Phillip Smith, NASA Ames Research Center Raymond Suorsa, NASA Ames Research Center Rangachar Kasturi, Penn State University Barry Roberts, Systems Research Center, Honeywell
12:00 - 1:00	Lunch (Galileo Room, Ames Cafeteria)

PRECEDING PAGE BLANK NOT FILMED

Thursday, March 11 (continued)

- 1:15 - 3:00 Image Processing: Human Vision Panel (M. Kaiser, chair)
 A. Beau Watson, NASA Ames Research Center
 Jeffrey Mulligan, NASA Ames Research Center
 Mary Kaiser
 Walter Johnson, NASA Ames Research Center
- 3:00 - 3:15 Afternoon Break
- 3:15 - 5:00 Image Evaluation and Metrics Panel (D. Foyle, chair)
 Sidney Connor, Maryland Advanced Development Lab
 David Foyle, NASA Ames Research Center
 Albert Ahumada, NASA Ames Research Center
 Misha Pavel

Friday, March 12

- 8:30 - 9:00 Coffee and Donuts
- 9:00 - 10:00 Workshop Review/Group Discussion
 Commentator: Tony Lambregts, Boeing CAG
- 10:00 - 11:45 Laboratory and Facility Demonstrations (Room 129 and
 Building 257 - Man/Vehicle Systems Research Facility)
- 12:00 - 1:30 Lunch (Golden Wok Restaurant in Mountain View)

Proceedings of the Augmented Visual Display (AVID) Research Workshop

Mary K. Kaiser and Barbara T. Sweet, Editors

Ames Research Center

SUMMARY

The papers, abstracts, and presentations in this volume were presented at a three day workshop focused on sensor modeling and simulation, and image enhancement, processing, and fusion. The technical sessions emphasized how sensor technology can be used to create visual imagery adequate for aircraft control and operations. Participants from industry, government, and academic laboratories contributed to panels on Sensor Systems, Sensor Modeling, Sensor Fusion, Image Processing (Computer and Human Vision), and Image Evaluation and Metrics.

I. SENSOR SYSTEMS

PRECEDING PAGE BLANK NOT FILMED

PAGE 2 INTENTIONALLY BLANK

Infrared Sensors and Systems for Enhanced Vision/Autonomous Landing Applications

J. Richard Kerr
FLIR Systems, Inc.

ABSTRACT

Infrared Imaging Through Fog

There exists a large body of data spanning more than two decades, regarding the ability of infrared imagers to "see" through fog, i.e., in Category III weather conditions. Much of this data is anecdotal, highly specialized, and/or proprietary.

In order to determine the efficacy and cost effectiveness of these sensors under a variety of climatic/weather conditions, there is a need for systematic data spanning a significant range of slant-path scenarios. These data should include simultaneous video recordings at visible, midwave (3-5 micron), and longwave (8-12 micron) wavelengths, with airborne weather pods that include the capability of determining the fog droplet size distributions.

Existing data tend to show that infrared is more effective than would be expected from analysis and modeling. It is particularly more effective for inland (radiation) fog as compared to coastal (advection) fog, although both of these archetypes are oversimplifications. In addition, as would be expected from droplet size vs wavelength considerations, longwave outperforms midwave, in many cases by very substantial margins. Longwave also benefits from the higher level of available thermal energy at ambient temperatures.

Imager Technologies

The principal attraction of midwave sensors is that staring focal plane technology is available at attractive cost-performance levels. However, longwave technology such as that developed at FLIR Systems, Inc. (FSI), has achieved high performance in small, economical, reliable imagers utilizing serial-parallel scanning techniques.

In addition, FSI has developed dual-waveband systems particularly suited for enhanced vision flight testing. These systems include a substantial, embedded processing capability which can perform video-rate image enhancement and multisensor fusion. This is achieved with proprietary algorithms and includes such operations as real-time histograms, convolutions, and fast Fourier transforms.

RELEVANT ACTIVITIES AT FSI

LOCKHEED (C-130) HTTP

OTHER AIR FORCE ACTIVITIES

RNLN - P3

SITUATION AWARENESS FLIR

SAFIRE (Q22)

HIGH PERFORMANCE
3-AXIS INERTIAL STABILIZATION
FULL DIGITAL FEATURES
TURRET WILL ACCOMMODATE EVS SENSORS
- FLEXIBILITY, COMPATABILITY

RELEVANT ACTIVITIES AT FSI (CONT)

COMPACT FLIR DEVELOPMENT - APPLY TO LONGWAVE EVS
STARING ARRAY DEVELOPMENT - APPLY TO MIDWAVE EVS

DUALBAND IMAGING RADIOMETER DEVELOPMENT
- APPLY TO DUAL-WAVELENGTH EVS

EMBEDDED PROCESSOR

SENSOR FUSION CONCEPTS

KEY: ECONOMICAL!

IR SENSOR TECHNOLOGIES

TECHNOLOGY	WAVEBAND	RESOLUTION (HxV)	SENSITIVITY (NETD)	IMAGER SIZE (IN)
COMPACT SCANNING	EITHER	500x375	0.5°C (3-5) 0.2°C (8-2)	5.7x5.7x10.9
STARING ARRAY (Pt Si)	3-5 MICRONS	320x244 640x488	0.08°C	5x5x9 6x6x10
STARING & UNCOOLED ARRAYS	8-12 MICRONS	FUTURE		

BASIC FSI TECHNOLOGY

DETECTOR MINI-ARRAYS WITH TDI: "BEST OF SERIAL AND PARALLEL SCAN"

PROPERTIES OF SERIAL-SCAN FLIRS:

CHALLENGES --

Low dwell time per resolution element

High speed azimuth scanner

ADVANTAGES --

Few detectors = high sensitivity and uniformity

optimized front-end electronics

high yield (economical)

efficient cold shield

Easy channel balance/low fixed-pattern noise

Freedom from vertical aliasing

BASIC FSI TECHNOLOGY, (continued)

Video output from simple electronics

- no complex E-Mux
- no complex DSC

AC coupling artifacts and low-frequency noise minimized

"Fast" optics (low f#'s)

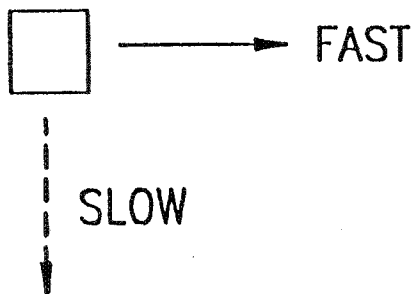
TDI vs SPRITE:

- freedom from charge carrier diffusion
- fast optics are permitted
- easier material fab (carrier lifetimes)
- less heat (resistance X bias current)

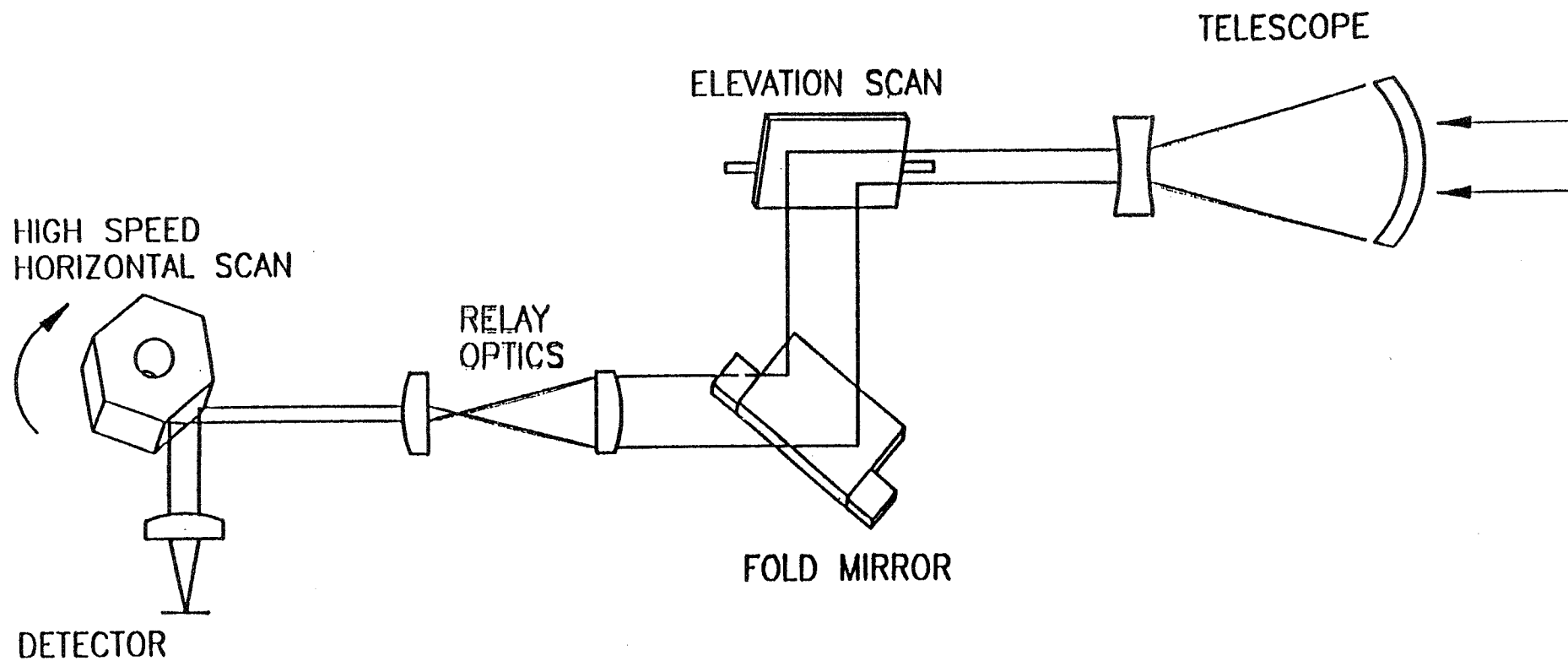
RELIABILITY

MAINTAINABILITY

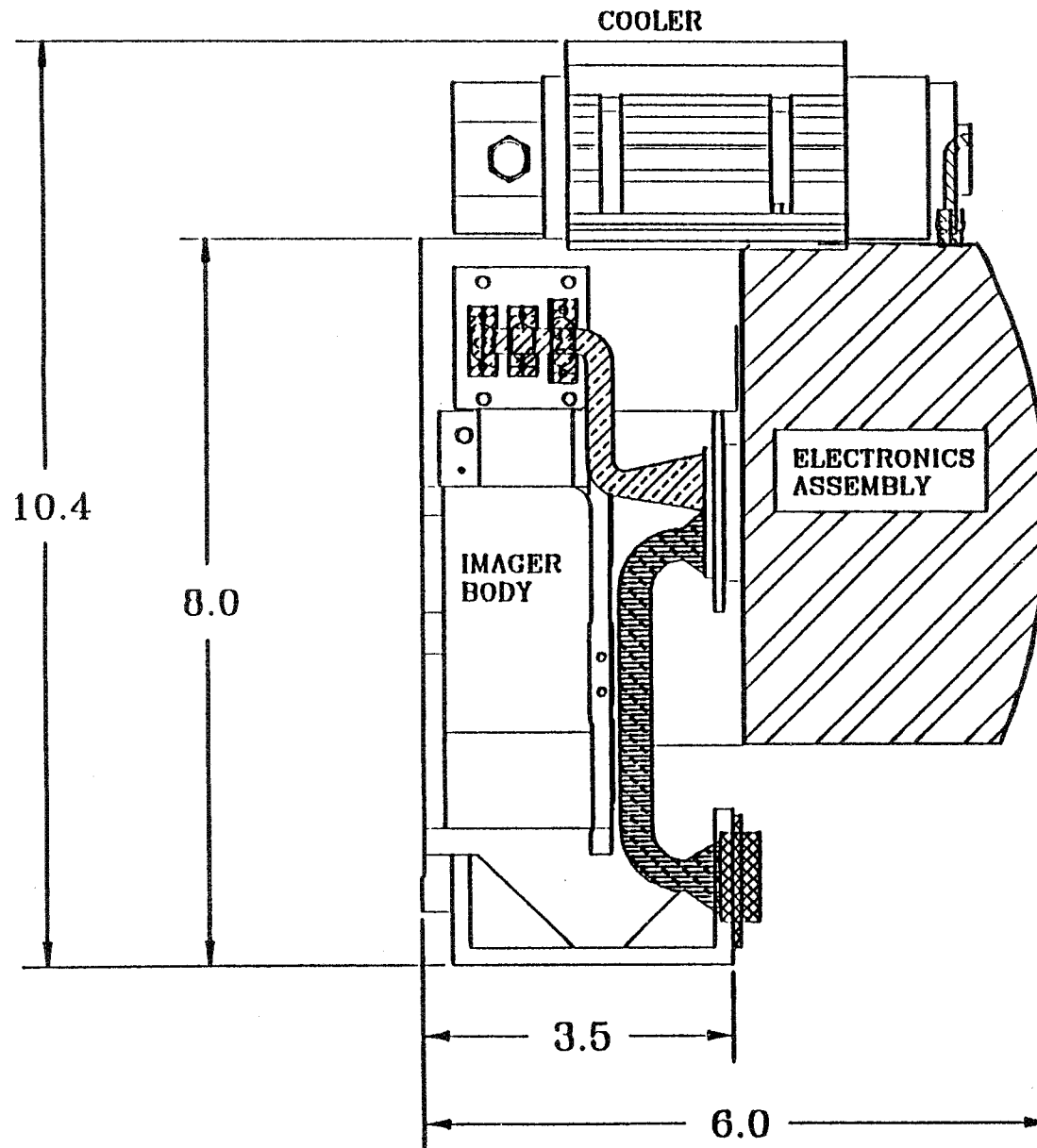
LIFE-CYCLE COST

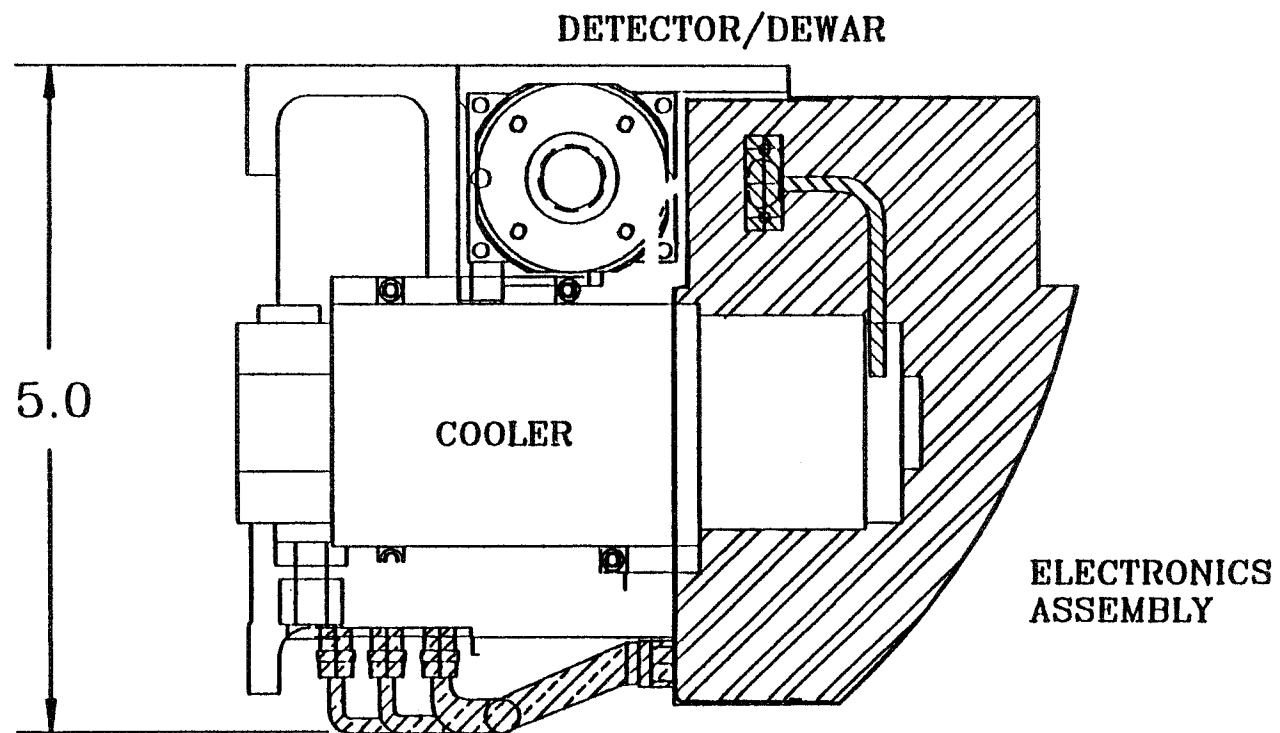


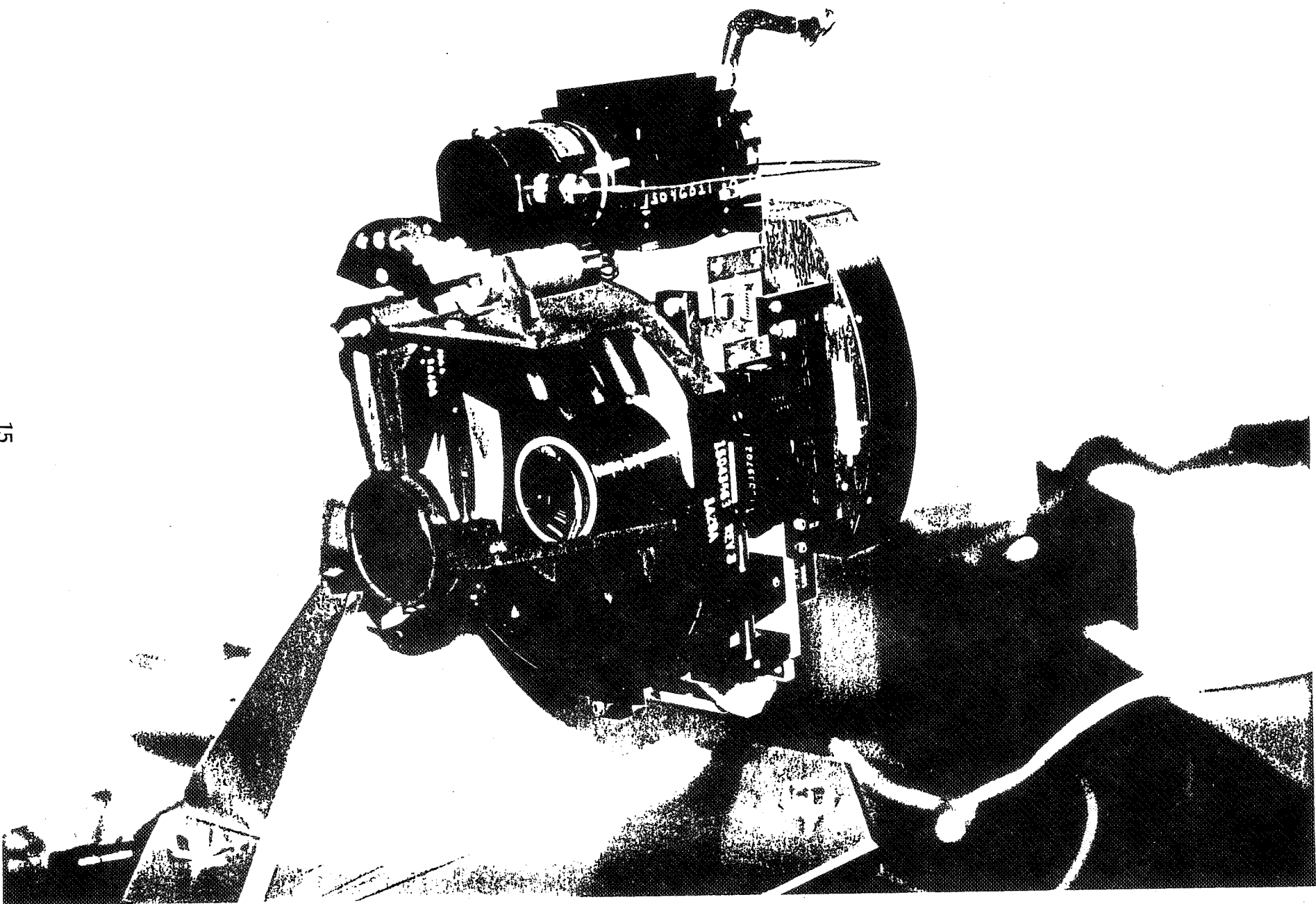
SERIAL SCAN



IMPLEMENTATION







OPTIONAL AND GROWTH FEATURES

SIMPLE PAN-TILT

SNAP-LOOK INTO TURNS

DIGITAL ZOOM

COMMON PROCESSING WITH RADAR (SENSOR FUSION)

INTEGRATION WITH GPS

"CAN INFRARED PROVIDE SEE-TO-LAND AT 2400 FEET IN CAT IIIA?"

***** BASIS OF DECISIONS**

- MEANINGFUL FLIGHT (APPROACH) AND FIELD DATA
- COMPLETE, OPEN VS ANECDOTAL, PROPRIETARY
- CHARACTERIZE PROPAGATION MEDIUM

SIMULATION

- EVS SENSOR AND HF ISSUES

RELATIVE ADVANTAGES OF WAVEBANDS

MIDWAVE (3 - 5 MICRONS)

STARING-ARRAY TECHNOLOGY MORE MATURE AT MIDWAVE COMPARED TO LONGWAVE

RUNWAY LIGHTS ARE "BEACONS" AT MIDWAVE

BETTER AT LONG RANGES IN HIGH HUMIDITY ATMOSPHERE

APPROPRIATE FOR TAXI AND TAKEOFF

- **LANDING GEAR MOUNT**
- **FOG CHARACTERISTICS NEAR GROUND**

RELATIVE ADVANTAGES OF WAVEBANDS (continued)

LONGWAVE (8 - 12 MICRONS)

BETTER PERFORMANCE IN LOW-THERMAL-CONTRAST AMBIENT CONDITIONS

BETTER PERFORMANCE IN MANY FOG SCENARIOS

- **SEE 1-3 TIMES VISUAL RANGE**
- **ALWAYS AS GOOD AS MIDWAVE**
- **CAN BE 100'S OF TIMES BETTER THAN MIDWAVE**

FLIR PERFORMANCE ISSUES (continued)

LONGWAVE (8-12 MICRONS) VS MIDWAVE (3-5 MICRONS) FLIRS

- **LONGWAVE CAN HAVE HUGE ADVANTAGE FOR INLAND FOG**
- **SHORTWAVE CAN HAVE SOME ADVANTAGE IN HUMID ATMOSPHERES**
- **MAY BE COMPARABLE FOR COASTAL FOG**
- **STARING ARRAY FLIRS NOT AVAILABLE AT LONGWAVE**

HIGHLY DESIRABLE: INTEGRATED DUAL-WAVEBAND FLIR AT LOW PRICE

PERFORMANCE SPECIFICATIONS DUAL-WAVEBAND SYSTEM

	<u>MIDWAVE</u>	<u>LONGWAVE</u>
FIELD OF VIEW	30 x 20 DEG	30 x 20 DEG
RESOLUTION	1.5 MRAD	1.5 MRAD
SENSOR HEAD ENVELOPE	10.5" H x 5" W x 6" L	
SENSOR HEAD WEIGHT	LESS THAN 12 POUNDS	
POWER REQUIREMENT	23V AT 4A MAX	
FIELD OF REGARD	FIXED, FORWARD (CONFORMAL WITH HUD) AZIMUTH PAN/SNAP-LOOK OPTION ELECTRONIC ZOOM OPTION	
SPLIT-STIRLING COOLER		
DUAL RS-170 OUTPUT		
ENVIRONMENTALLY QUALIFIED TO MIL-STD-810D		

ALTERNATIVE ("HYBRID") DUALBAND SYSTEM

SCANNING LONGWAVE / STARING MIDWAVE

COMMON VS CONTIGUOUS OPTICAL PATHS

DUAL OUTPUTS

RETAIN: COMPACTNESS, ECONOMY

- COMMON COOLER

SENSOR FUSION

IMAGE PROCESSOR SUMMARY

COMPONENTS

SENSOR INTERFACE

DISPLAY INTERFACE

CENTRAL PROCESSOR INTERFACE

STRUCTURE

VME BACKPLANE (32 ADDRESS/16 DATA)

CUSTOM/PARALLEL 16-BIT VIDEO BUS

DUAL EUROCARD STANDARD (6-U)

Provides for easy addition of standard processing alternatives

CENTRAL PROCESSOR PERIPHERALS

SCSI I/O (TAPE, HARD AND FLOPPY DISK, OPTICAL STORAGE)

RS-232 (OR 422 OR 485))

As a console for operator communication

Or in a protected host communication packet mode with error checking

ETHERNET

CENTRONICS

SENSOR FUSION

SENSOR INTERFACE

COMPONENTS

SCAN CONVERTER

INPUT PROCESSING ALU

SENSOR IDENTITY MODULE

FRAME MEMORY

SCAN CONVERTER

CONVERTS NON-STANDARD SENSOR SCANS INTO TELEVISION
SYNCHRONIZED RASTER FOR THE FRAME MEMORY

POSITION CONTROL

ZOOM

OUTPUT INTO VME ADDRESSABLE BIT-MAP VIDEO MEMORY

PAN/BORESITE

IMAGE REGISTRATION

VIDEO INPUT PROCESSOR

BRIGHTNESS CORRECTION

ALC

CONTRAST CORRECTION

AGC

REALTIME AVERAGING

NOISE REDUCTION

FAST FOURIER

(not supported in current products)

CAN PEAK DETECT/MASK SUBTRACT

SENSOR FUSION

SENSOR INTERFACE (CONTINUED)

SENSOR IDENTITY MODULE

SENSOR SYNCHRONIZATION

SENSOR CONTROL

MANCHESTER OR RS-232 COMMUNICATIONS ARE COMMON

ANALOG INPUT option

10 bit/w AGC

DIGITAL INPUT option

PARALLEL

SERIAL (TAXI - UP TO 10 BITS)

FRAME MEMORY

CONFIGURES TO SCANNER RESOLUTION/SENSITIVITY

256 X 512 X 8 UP TO 1024 X 1024 X 16

THREE PORTED MEMORY

VME READ/WRITE

SENSOR WRITE INPUT

TELEVISION RASTER SCANNED READ

(CONTROLLED BY DISPLAY TIMING)

SENSOR FUSION

DISPLAY PROCESSOR

COMPONENTS

8 BIT CONFIGURABLE OVERLAY

TMS 34020 GRAPHICS SUPPORT (TIGA)

HISTOGRAM PROCESSOR (KEYPLANE ADDRESSABLE AREA)

CONVOLUTION PROCESSOR (3 X 3)

FULL DISPLAY LEVEL ADDRESSING (REMAPPING IN RAM LUT)

OUTPUT SECTION

WINDOWS SUPPORT

INPUT SWITCHABLE BETWEEN 3 SENSORS

OVERLAY

1024 X 1024 X 8

(CAN BE CONFIGURED FOR LESS RESOLUTION)

SHARED CPU/VME MEMORY ADDRESSING (for DMA)

KEY-PLANE CONTROL OF WINDOW AREAS

GRAPHICS PROCESSOR SUPPORT IS VME ADDRESSABLE

HISTOGRAM AND CONVOLVER ARE REAL TIME

HISTOGRAM MEMORY IS SHARED WITH CPU for DMA ACCESS

SENSOR FUSION

DISPLAY SUPPORT (CONTINUED)

VIDEO OUTPUTS

RGB AND SYNC

NTSC (OR PAL)

S-VHS OUTPUT (Y/C)

PARALLEL DIGITAL PROCESSING

RS-170 (OR CCIR) BLACK AND WHITE OUTPUT
(TO 10 BITS + SYNC)

FULLY REMAPPABLE OUTPUT COLORS (TO 8 BITS EACH)

PROVISION FOR CGA/VGA AND S-VGA
(NOT IMPLEMENTED YET)

ON BOARD VIDEO MEMORY SUPPORT (not real time)

(QUADRANT DISPLAY (only one quadrant can be live at a time)

FULL 16 BIT TO 10 BIT BLACK AND WHITE OR 8 BIT COLOR
REMAPPING ABILITY

(for histogram image correction, gamma correction, etc)

ACCEPTS EXTERNAL SYNCHRONIZATION

Synthetic Vision System Flight Test Results and Lessons Learned

Jeffrey Radke
Honeywell Systems and Research Center

ABSTRACT

Honeywell Systems and Research Center developed and demonstrated an active 35 GHz Radar Imaging system as part of the FAA/USAF/Industry sponsored Synthetic Vision System Technology Demonstration (SVSTD) Program. The objectives of this presentation are to provide a general overview of flight test results, a system level perspective that encompasses the efforts of the SVSTD and Augmented Visual Display (AVID) programs, and more importantly, provide the AVID workshop participants with Honeywell's perspective on the lessons that were learned from the SVS flight tests.

One objective of the SVSTD program was to explore several known system issues concerning radar imaging technology. The program ultimately resolved some of these issues, left others open, and in fact created several new concerns. In some instances, the interested community has drawn improper conclusions from the program by globally attributing implementation specific issues to radar imaging technology in general. The motivation for this presentation is therefore to provide AVID researchers with a better understanding of the issues that truly remain open, and to identify the perceived issues that are either resolved or were specific to Honeywell's implementation.

CHART 1: Synthetic Vision System Flight Test

The SVSTD program was motivated by an existing "catch-22" situation, in which the avionics user community was unaware of the capabilities and benefits of an adverse weather (fog, rain, snow, haze) imaging system, while potential manufacturers of such a product did not perceive an existing marketplace. The program focused on demonstrating this technical capability, as well as on a first step toward resolution of the many issues associated with the system's certification.

A Gulfstream 2 was used as the flight test aircraft. Honeywell developed an active 35 GHz imaging radar and integrated it with the Gulfstream 2 avionics system. A scanning antenna and the radar transmit/receive unit were mounted behind the radome. A real-time display processing unit, housed within a single, ruggedized VME chassis, was mounted in the aircraft cabin. The Honeywell display processor provided

pilot-perspective radar video to a Head Up Display (HUD) mounted in the cockpit. The HUD electronics projected a holographic image onto the HUD combining glass, effectively overlaying the radar image on the pilot's real world scene.

The test aircraft was outfitted with a host of related sensors and instrumentation. In addition to Honeywell's 35 GHz radar, the Gulfstream 2 was equipped with a 3-5 micron-band forward looking infrared (FLIR) camera and a visible-band camera. Separate flight tests were briefly flown using a Lear 94 GHz radar imager in place of the 35 GHz radar. The aircraft cabin was equipped with recording equipment, allowing radar, FLIR, and visible-band imagery to be simultaneously recorded. In order to support accurate analysis of the performance of each sensor as a function of weather conditions, the aircraft was also equipped with wing-mounted pods that measured atmospheric liquid content (both water density and droplet size).

Hundreds of approaches were flown into more than 25 airports across the US, encountering a wide variety of weather conditions. The program executed a flight test matrix, involving both instrumented and non-precision approaches with several test pilots, under varying weather conditions. The Honeywell 35 GHz radar demonstrated clear pilot advantages in most situations. Pilot performance across the flight test matrix was well documented, but will not be addressed in detail within this presentation.

CHART 2: Autonomous Airplane Technology - System Concept

Honeywell envisions an overall system concept that is much broader in scope than the fundamental Synthetic Vision System previously described. Ultimately, an aircraft can achieve greater autonomy through the integration of advanced cockpit decision aids and display technology, high-precision navigation aids, forward visibility sensors, and hazard detection sensors. Honeywell is actively involved with Boeing in the development of an Enhanced Situational Awareness System (ESAS) that could potentially take advantage of such technology capabilities.

CHART 3: Autonomous Airplane Technology - System Functions

A strawman block diagram could potentially include display electronics, forward visibility sensors, navigation and landing aides, and advanced processor systems. High precision guidance and navigation can be achieved using one or more of several candidate navigation/landing aides. A digital terrain map registered with a radar altimeter can also be used for increased accuracy. A millimeter wave (radar) imager, a FLIR, and/or digitally stored imagery are potential sources of images that can be presented to the pilot on some type of display. These image sources could be used in several ways, including selection of

the sensor with the best image at some time instance, fusion of multiple sensor images, or registration of a digitally stored image to one or more of the sensors. Other variations upon these themes can be constructed.

CHART 4: Honeywell 35 GHz Radar Imaging System Hardware

The major components that were flight tested include a 34"x4"x8" electro-mechanically scanned antenna, a radar receiver/transmitter (R/T) unit, an R/T Controller unit, and the Display Processor. The antenna and RT unit were both mounted behind the aircraft radome. The R/T Controller and Display Processor were mounted in the aircraft cabin. The majority of processing was housed within the Display Processor, implemented primarily with commercially available hardware mounted within a ruggedized VME chassis.

CHART 5: Honeywell SVS Function Block Diagram

A custom RF Interface card within the VME chassis is responsible for controlling the radar and antenna, as well as digitizing range samples. All range samples are then passed through the display processing pipeline, implemented with TI TMS320C30 digital signal processors. The display processing pipeline is controlled by a system processor. The system processor is also responsible for communicating with avionics bus interface cards, as well as storing raw radar data for post-flight analysis.

CHART 6: SVS Image Beam Sharpening

The display processing pipeline contains hardware allocated for optional execution of image enhancement functions. Honeywell has developed several algorithms for image contrast enhancement, noise reduction, and beam sharpening. Although the image enhancement algorithm suite was not part of the SVSTD flight test baseline configuration, Honeywell's beam sharpening algorithm has shown promising results.

The beamsharpening algorithm operates across the image, attempting to improve azimuthal resolution. Azimuthal resolution is most critical, in that runway acquisition range is typically driven by the ability of the sensor to fully contain one beamwidth between the runway edges, and thus provide the necessary contrast between the runway and the surrounding terrain. Honeywell's beamsharpening algorithm can be executed with real-time, flight-worthy hardware, to produce approximately a 2.5:1 improvement in azimuthal resolution.

CHART 7: A Honeywell SVS Image

An example of a pilot's perspective radar image is shown to include the flight director and navigational symbology that is overlaid by the GEC HUD. One issue that was identified by the SVSTD program concerns the tendency of HUD symbology to obstruct the runway at far ranges, or hide obstacles on the runway from the pilot's view.

CHART 8: SVS Lessons Learned

Several issues were studied or brought about by the SVSTD program. This presentation addresses those that are more of a concern from a radar imaging perspective, and represent only Honeywell's point of view. Other issues, perhaps at a higher system level, were addressed by the SVS Certification Issues Study Team, as presented at their January 1993 conference in Williamsburg, VA. An attempt is made to classify the issues according to the radar subsystem from which they are derived. Some issues are truly introduced at the system level, while others that have been related to a particular subsystem are indeed a system issue.

Minimum Range is an issue that concerns the inability of the radar system to sense near range signal returns. This "blind spot" is necessary to allow time for the saturated radar receiver to "settle" after each 1 kW pulse is transmitted. The visual effect is an absence of image in the near range. The Honeywell configuration that was tested began sampling radar returns at 150 feet. As shown in Chart 9, a 75 foot minimum range is more tolerable, and can be achieved within the current implementation with only minor adjustments.

Resolution at 35 GHz was a concern. The program demonstrated that 35 GHz resolution is marginally acceptable. As discussed earlier, beamsharpening can be applied to the imagery to provide image resolution which would approach that inherent in a 94 GHz radar with equivalent antenna aperture. A beamsharpened 94 GHz image would offer excellent resolution. Similarly, a 10 GHz (X-band) system using beamsharpening would at best be marginally acceptable (about equivalent to 35 GHz without beamsharpening).

Intrusion Detection was an operational capability tested by the SVSTD program. Pilots could usually detect foreign obstacles on the runway after some exposure to a "normal" runway radar scene. The few occasions when the pilot failed to detect intrusions may be attributed to one or more problems. The tendency for overlaid HUD symbology to obstruct obstacles shown in the radar image was evident on some occasions. Additionally, the radar image itself contained secondary artifacts, that with further radar

development work may be resolved, but tended to cause problems for pilots in discerning obstacles from the artifacts.

Motion Compensation with a low scan rate antenna is an approach that may or may not be viable as an alternative to expensive high scan rate antennas. Honeywell did not study this approach, opting instead to use a relatively high scan rate antenna (>10 Hz). It is still an open issue as to whether a slow antenna with motion compensation will allow adequate pilot performance based on only the radar image.

Antenna Performance Requirements were fairly well determined by the flight test program, as well as previous research. Prior research had shown that frame rates in excess of 17 - 18 fps provided diminishing return in terms of pilot performance. The 10 Hz Honeywell system was marginally acceptable. The 30 degree antenna field of view (fov) was driven primarily by inherent limitations in the HUD. It was established that a 40 degree fov would be desirable, especially for high crab-angle approaches.

Achieving high scan rate and wide fov is very challenging for antenna designs. The approach taken by Malibu Research in developing Honeywell's antenna was effectively to piece two antennas side-by-side. One resulting effect was a dark line in the center of the image, caused by a gain imbalance between the two antenna halves. This imbalance may have been resolved with extensive antenna tuning, or with addition processing downstream. System designers should note this problem as an artifact of the Malibu antenna design, and not necessarily a characteristic of all radar imaging systems.

Antenna Pitch Stabilization was a debated requirement until flight testing proved its necessity. The Honeywell flight test configuration did not pitch stabilize the antenna. Since the antenna vertical beamwidth is relatively narrow, even slight changes in the aircraft pitch attitude tended to produce dynamic intensity variations across the runway scene. The most notable problem, however, was the inability to optimize the pitch angle for both approach and taxi. Nominally, a look-down angle of 3 degrees was optimal for approach on typical glidepaths. For ground operations, however, the antenna fixed at 3 degrees down was very inefficient since the scene ahead was nominal at 0 degrees. For purposes of the flight test, a compromise configuration was used (without pitch stabilization) as shown in Chart 10. Ultimately, the imaging radar should use a pitch stabilized antenna.

Antenna Sidelobe Suppression is critical to the radar imaging system implementation. The Malibu antenna implementation had fairly low sidelobes, however runway artifacts observed during flight testing may be attributed to the sidelobe returns. Although the sidelobe returns would be relatively low in amplitude, they would still tend to stand out against the extremely low runway returns onto which the sidelobe returns

would be mapped. It may be possible to remove sidelobe returns with additional signal processing, however this issue remains open.

Radome Effects were negligible for Honeywell's 35 GHz implementation. The development of radomes with high transmissivity at 94 GHz is still a problem, as witnessed by the 94 GHz Lear system tests. The difficulty at 94 GHz is in developing radome materials that are thin enough to allow 94 GHz transmission, yet strong enough to tolerate bird strikes and other stresses.

"Ground Rush" is a phenomena in which the motion in the radar image tends to convey increasing aircraft ground speed as altitude is decreased through the last few hundred feet. This effect is attributed to the fact that the Honeywell implementation used linear range samples (ie. one sample every 25 feet). Linear sampling produces too few samples per display pixel in the near range, and too many samples per display pixel in the far range. In the Honeywell implementation, this produced very blocky imagery in the near range. A more sophisticated approach would either use non-linear sampling, providing more samples in the near range, or would perform more processing intensive interpolation on near range pixels with a linear sampling approach.

Power vs Backscatter is a relationship that requires further study. The issue concerns the ability of a radar signal to penetrate weather. First instincts would suggest that more transmit power would result in better weather penetration. The reality is that at some point, the atmospheric backscatter begins to blind the radar, much like car headlights in fog. The point where this occurs can be theoretically derived, but was not verified by the flight test program.

Snow and Rain Performance was not adequately documented by the flight test program. More data needs to be collected and analyzed in this area. Of specific concern is the fact that radar cross sections from snow cover tend to vary widely depending upon several factors associated with the snow itself. This coupled with many potential runway states (snow covered, icy, freshly plowed, etc.) will not allow very accurate modelling or prediction of system performance in many situations.

Processing Latency: The processing latency, observed as the time from start of an aircraft maneuver until the radar image showed correlated effects, was about 0.4 seconds for Honeywell's prototype SVS system. Contrary to what some have purveyed, the system frame rate (> 10 Hz) is unaffected by processing latency. Latency through the image processing pipeline was actually only about 0.2 seconds. An implementation problem with the servicing of avionics bus interrupts accounted for the additional latency. Since aircraft orientation parameters were not being efficiently updated, image perspective was substantially (0.5 sec) lagging real world orientation changes (roll, pitch, yaw), even though the data

presented was relatively current. Display processing hardware used within the prototype primarily consisted of commercially available boards selected to enable rapid system development. Latency could be improved to about 0.2 seconds using this hardware, with minor changes to system control software. Ultimately, a more custom hardware approach would have substantial latency improvement.

Beamsharpening: Image enhancement that can be accomplished through antenna beam sharpening techniques is a well understood issue, and has been discussed in previous charts.

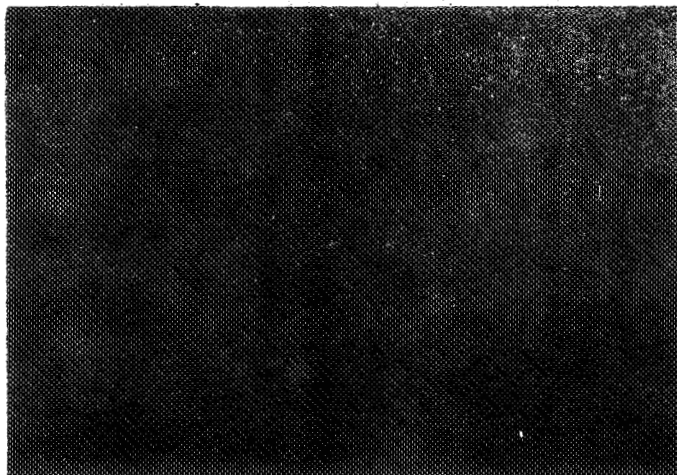
Image Enhancement: Other image enhancement techniques for noise reduction and contrast enhancement to the radar image are actively being developed at Honeywell. Image enhancement is a very open area of research if one begins to consider the potential impact of fusion with other image sources such as FLIR, terrain databases, or computer graphics.

Display Registration: Registration of the radar image on the HUD with the true world scene was a concern at the onset of the SVS flight test. Several techniques were used to accomplish radar image registration, resolving the issue. An interesting artifact of registering the radar scene to the real world relates to the fact that the radar has limited range. Since the radar doesn't "see" to the horizon, the radar horizon line in the image usually appears lower than the true world horizon if the remainder of the radar image is registered. This is at first misleading, however the pilots seemed to become comfortable with the artifact. Future implementations may wish to artificially extend the radar horizon if the image is to be displayed in original (not fused) format.

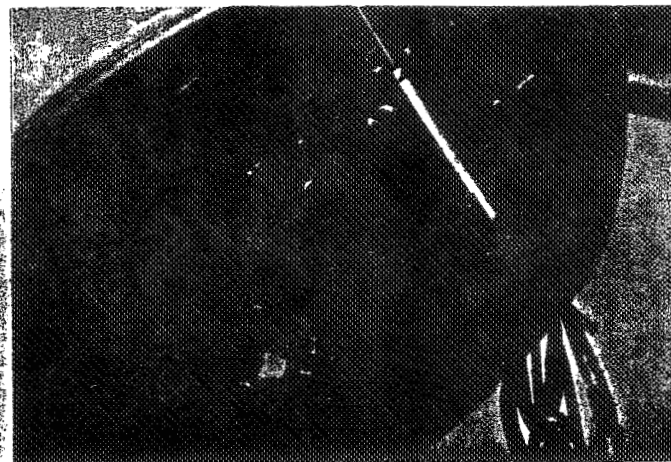
Taxi Display: Due to the fact that the radar has a limited vertical ranging angle, the resulting perspective transform image at low altitudes becomes very "short" vertically. This made taxi and ground operations very difficult for pilots during the flight test program. Some experimentation was performed in which the perspective altitude was artificially increased by 50 to 75 feet, giving more of a "god's eye" view while at low altitude or on the ground. Although this lead to a slightly, generally mis-registered image, the pilots found it was a much more useful than the true perspective during ground operations. An extension to this concept would be to present the radar "plan" view as an augmentation to the C-scope image.

Fusion: Clearly sensor fusion is an open area of research, and is one of the main topics for the AVID workshop.

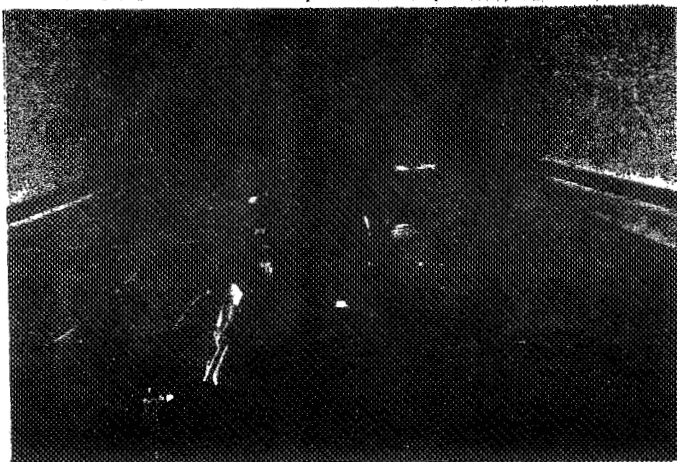
Synthetic Vision System Flight Test



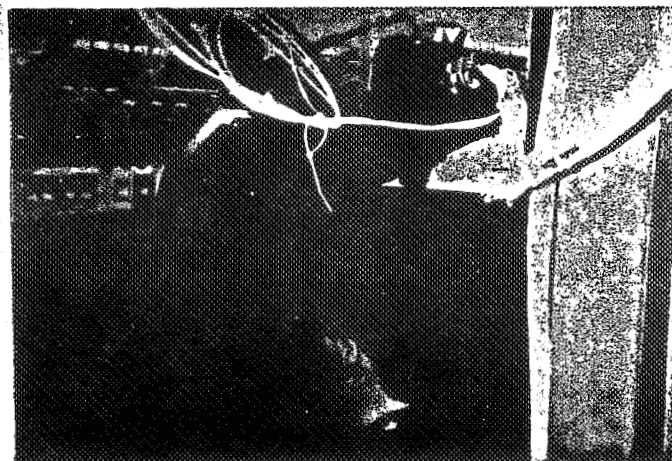
Test Aircraft



Antenna Mount



Flight Instrumentation



Cockpit/HUD

CHART 1

Autonomous Airplane Technology

System Concept

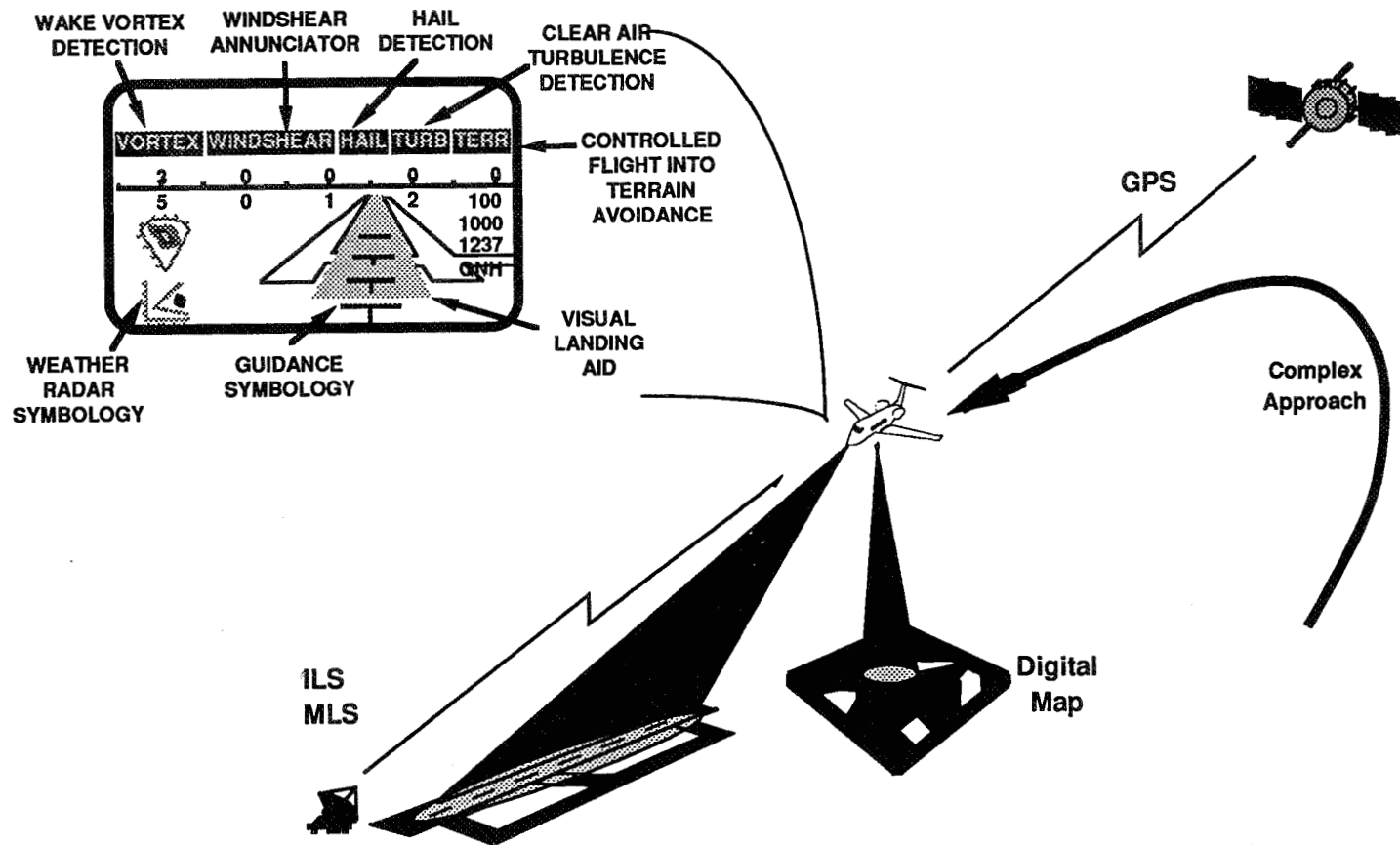


CHART 2

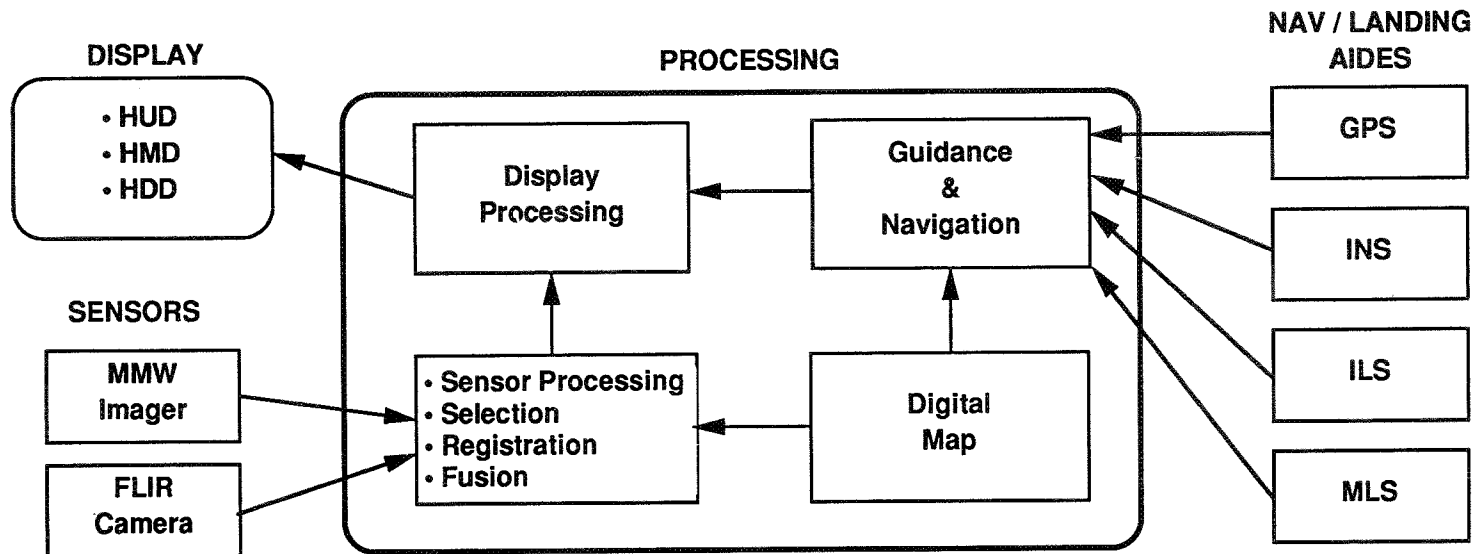


CHART 3

SVS Flight Demo Hardware

39

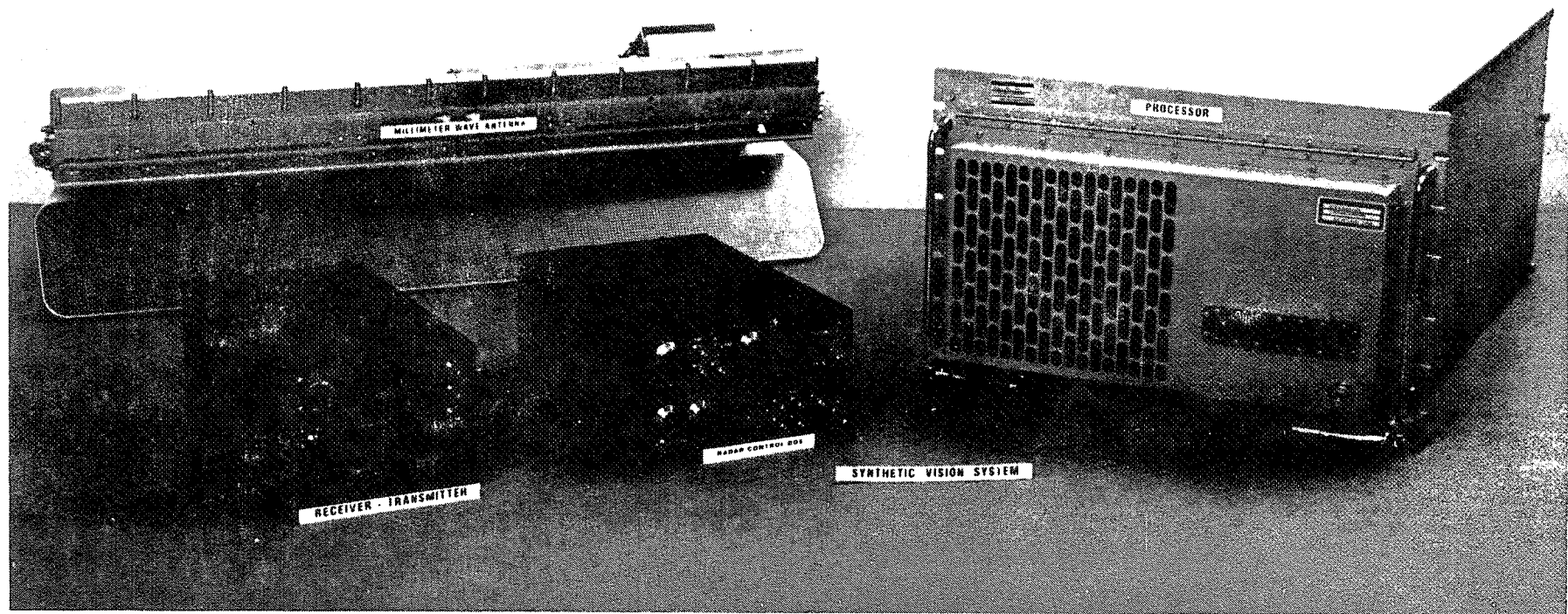


CHART 4

**SYNTHETIC VISION SYSTEM
FUNCTION BLOCK DIAGRAM**

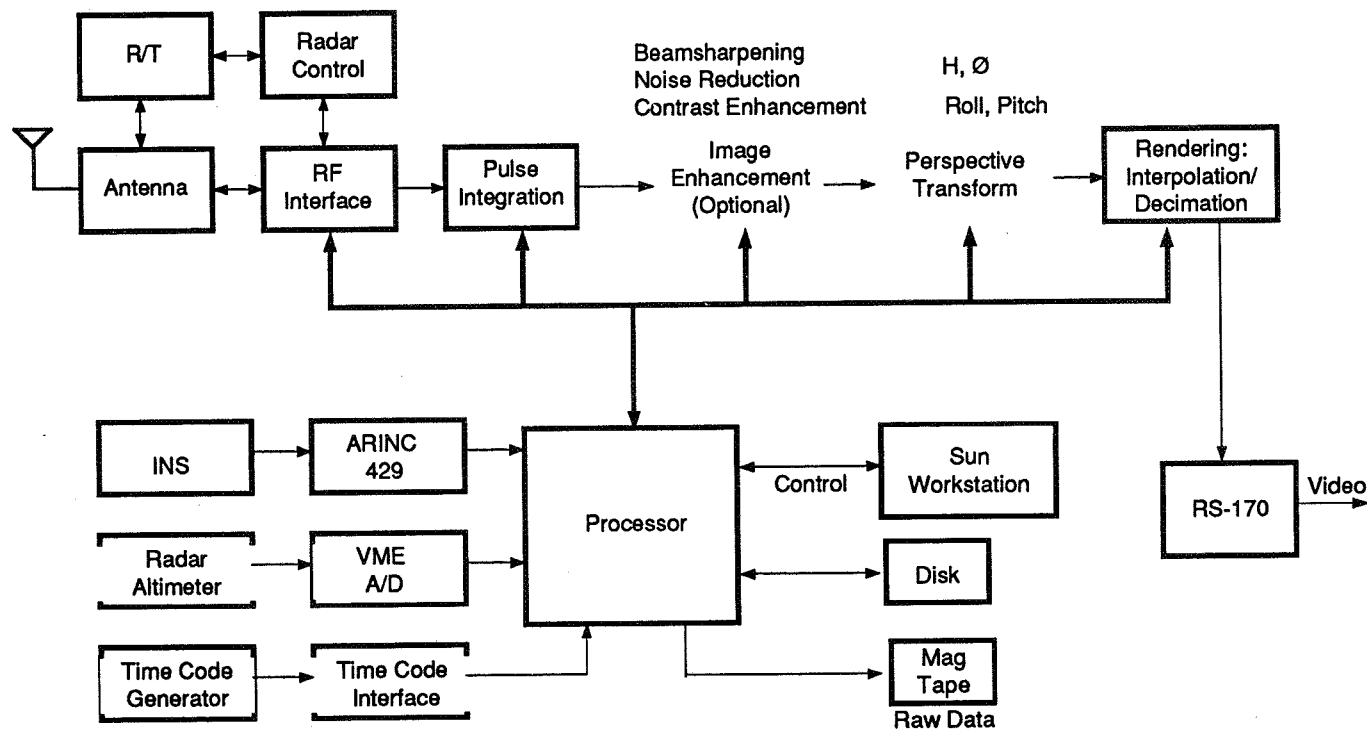
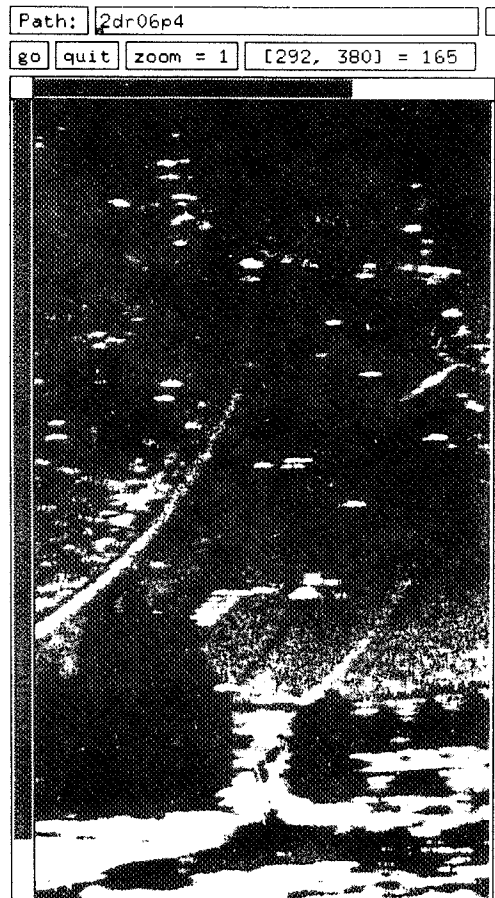


Chart 5

Synthetic Vision Radar Image



Unenhanced Image



Image After Beam Sharpening

CHART 6

APP 155 221:15:48:15 206

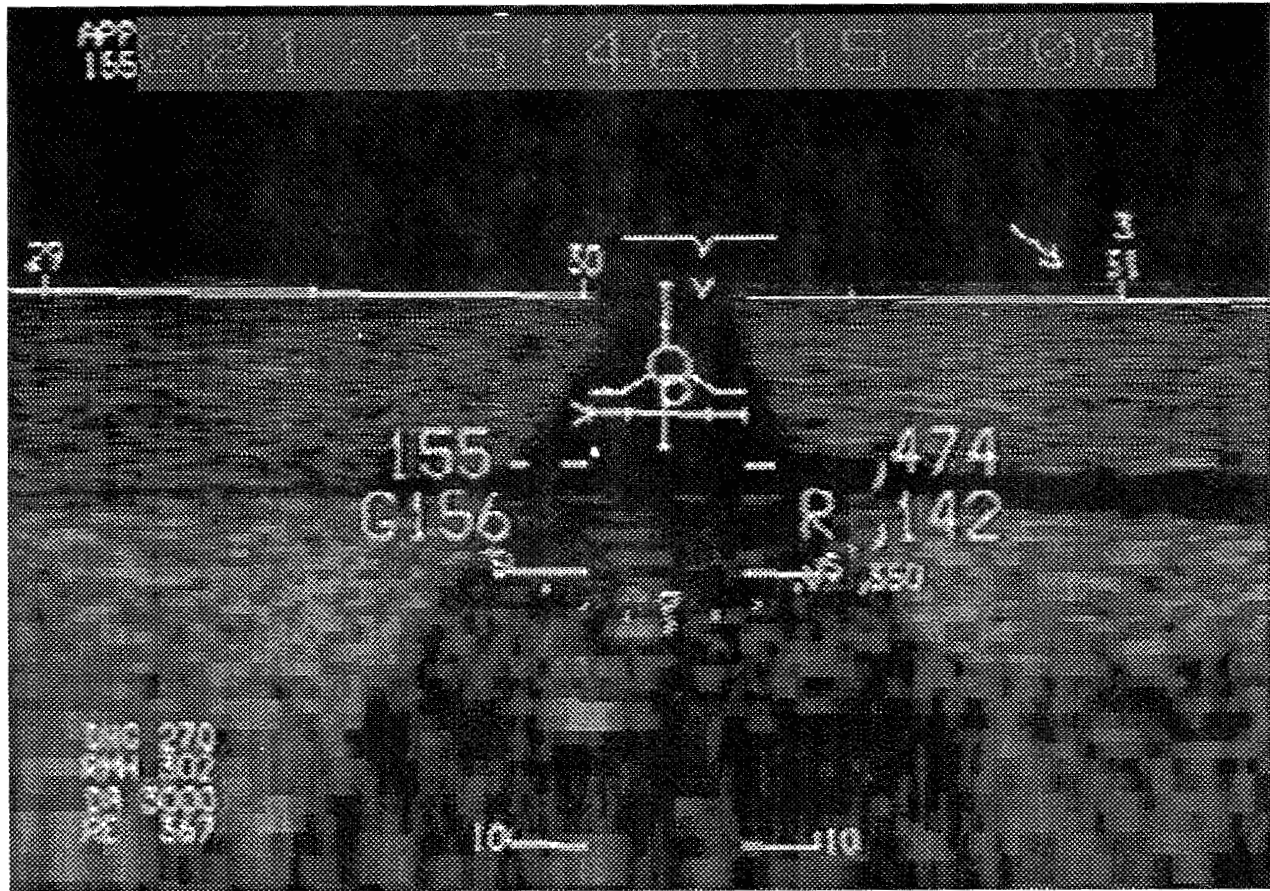


CHART 7

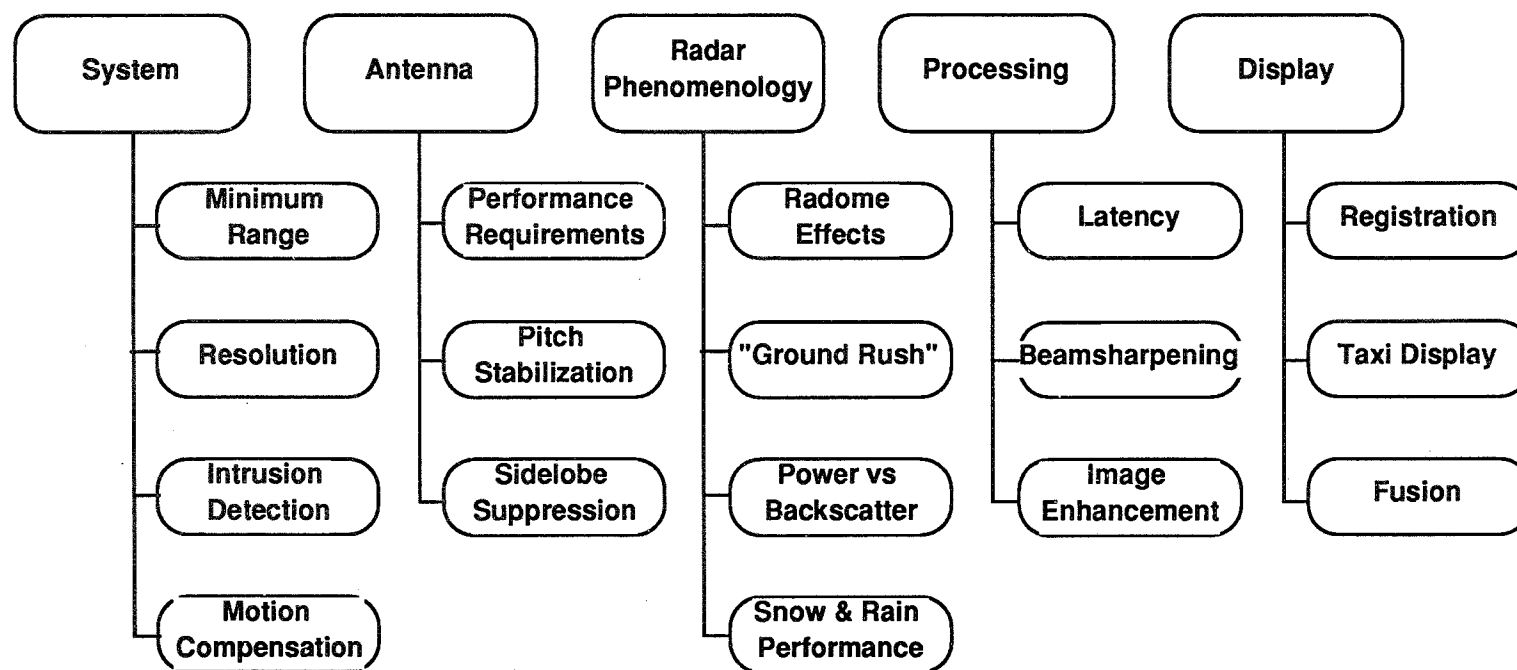
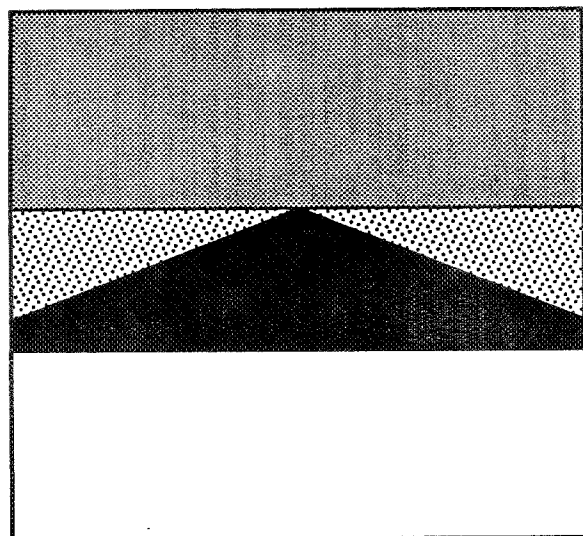


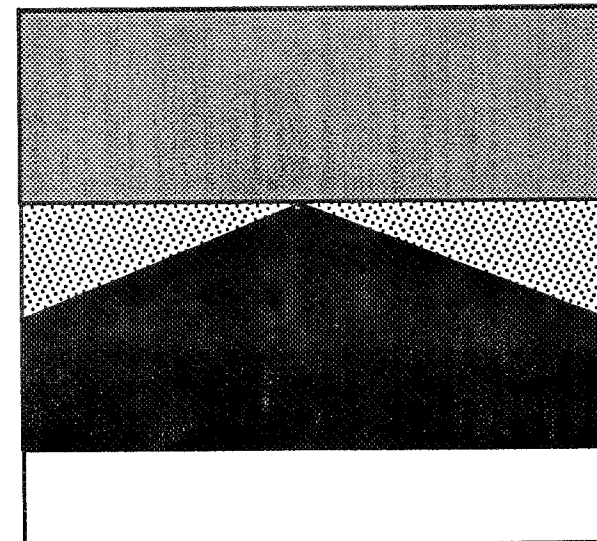
CHART 8

HUD View During Taxi with Minimum Radar Range of 150 ft and 75 ft



12,000 ft.

150 ft.



12,000 ft

75 ft.

Taxi strip width: 100 ft.
HUD FOV: $\pm 15^\circ$ Azimuth; 26° Elevation
Perspective View Elevation: 12 ft.

CHART 9

Flight Test Antenna Configuration

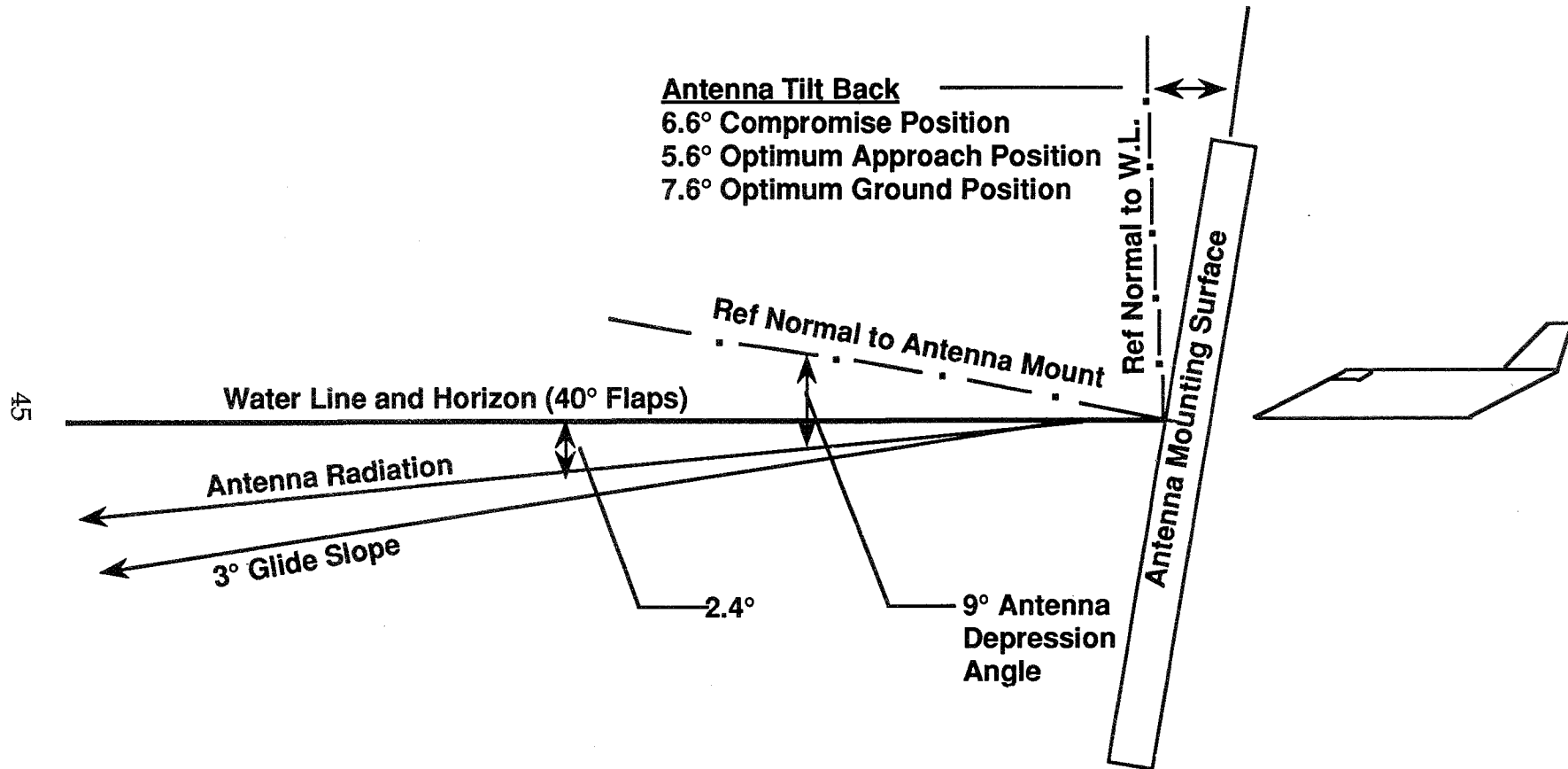


CHART 10

94 GHz MMW Imaging Radar System

Yair Alon and Lon Ulmer
Lear Astronics Corporation

23-32
205733
P. 14
249915

ABSTRACT

The 94 GHz MMW airborne radar system that provides a runway image in adverse weather conditions is now undergoing tests at Wright-Patterson Air Force Base (WPAFB). This system, which consists of a solid state FMCW transceiver, antenna and digital signal processor, has an update rate of 10 times per second, 0.35° azimuth resolution and up to 3.5 meter range resolution. The radar B scope (range versus azimuth) image, once converted to C scope (elevation versus azimuth), is compatible with the standard TV presentation and can be displayed on the Head Up Display (HUD) or Head Down Display (HDD) to aid the pilot during landing and takeoff in limited visibility conditions.

INTRODUCTION

The technology now exists to take the next step in all-weather landing capability. An Enhanced Vision System employing a weather penetrating sensor interfaced to a raster/stroke heads-up-display will give the pilot an out-the-window view of the runway which allows a "VFR" manually flown approach in CAT III weather conditions at facilities that have only CAT I quality precision or non-precision approach guidance. This provides several advantages over conventional autoland operations:

- Potentially autonomous CAT IIIa or IIIb operation on any runway
- Ground movement at any RVR
- Takeoff at 300 ft RVR at any facility
- Runway incursion detection
- Reduced approach spacing — "VFR operations"

The final system configuration is illustrated in Figure 1. It consists of a scanned antenna, solid state TX/RX, DSP, radar controller and HUD.

EVS TESTBED

An EVS testbed has been developed by Lear Astronics Corp. under a joint FAA/Air Force contract in order to evaluate quantitatively the performance of a 94 GHz FMCW imaging radar in real weather conditions.

The testbed depicted in Figure 2 is being evaluated in a stationary tower test at Wright-Patterson AFB starting in August of 1991, and will then be integrated into a Gulfstream II business class jet for flight testing in adverse weather conditions during 1992.

The testbed consists of a 94 GHz tilt-scanner antenna, a solid state transceiver, a radar interface unit, a digital signal processor, and an integral radar/video data recording system. The antenna with its drive electronics, the TX/RX, and the radar interface unit will mount in the radome of the GII, and the DSP and data recording equipment will be rack-mounted in the cabin.

OPERATIONAL RADAR REQUIREMENTS

The results of a trade-off study to establish radar performance requirements are summarized in Table I.

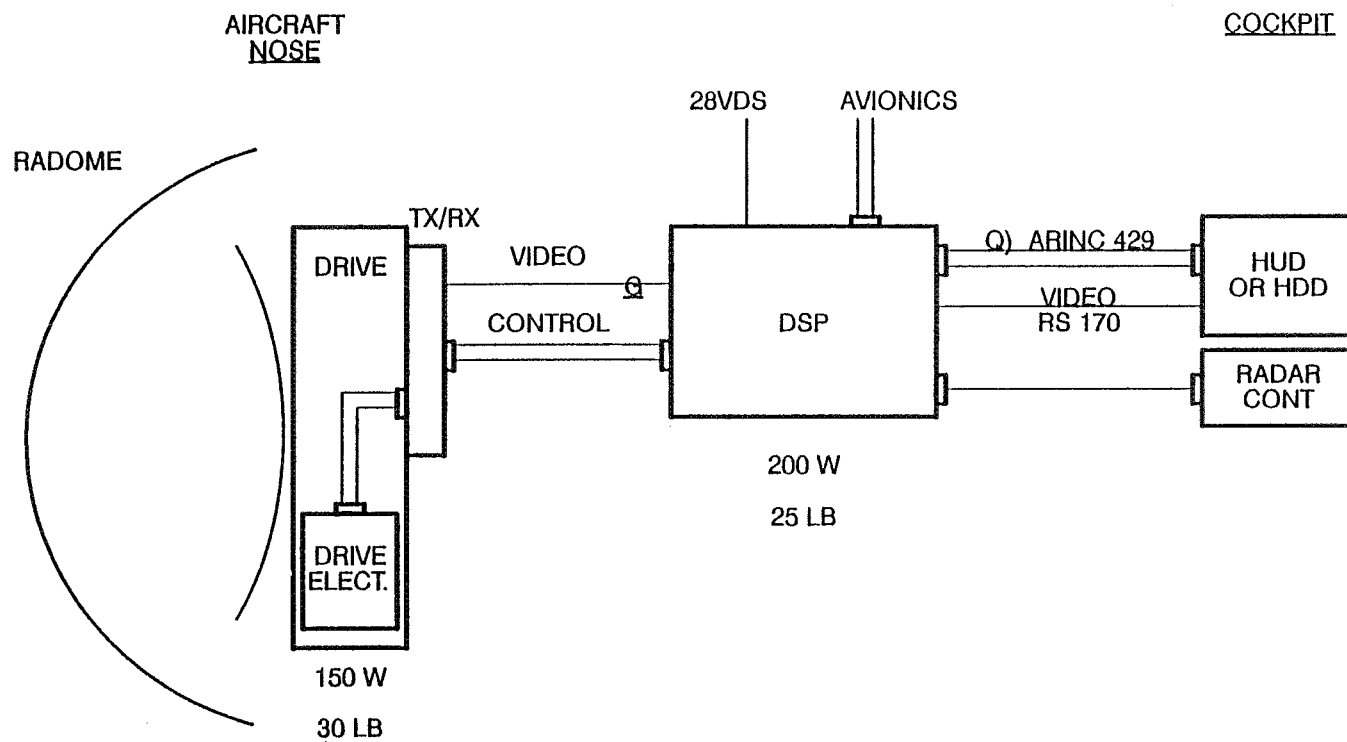


Figure 1. Final Production System Equipment Rack

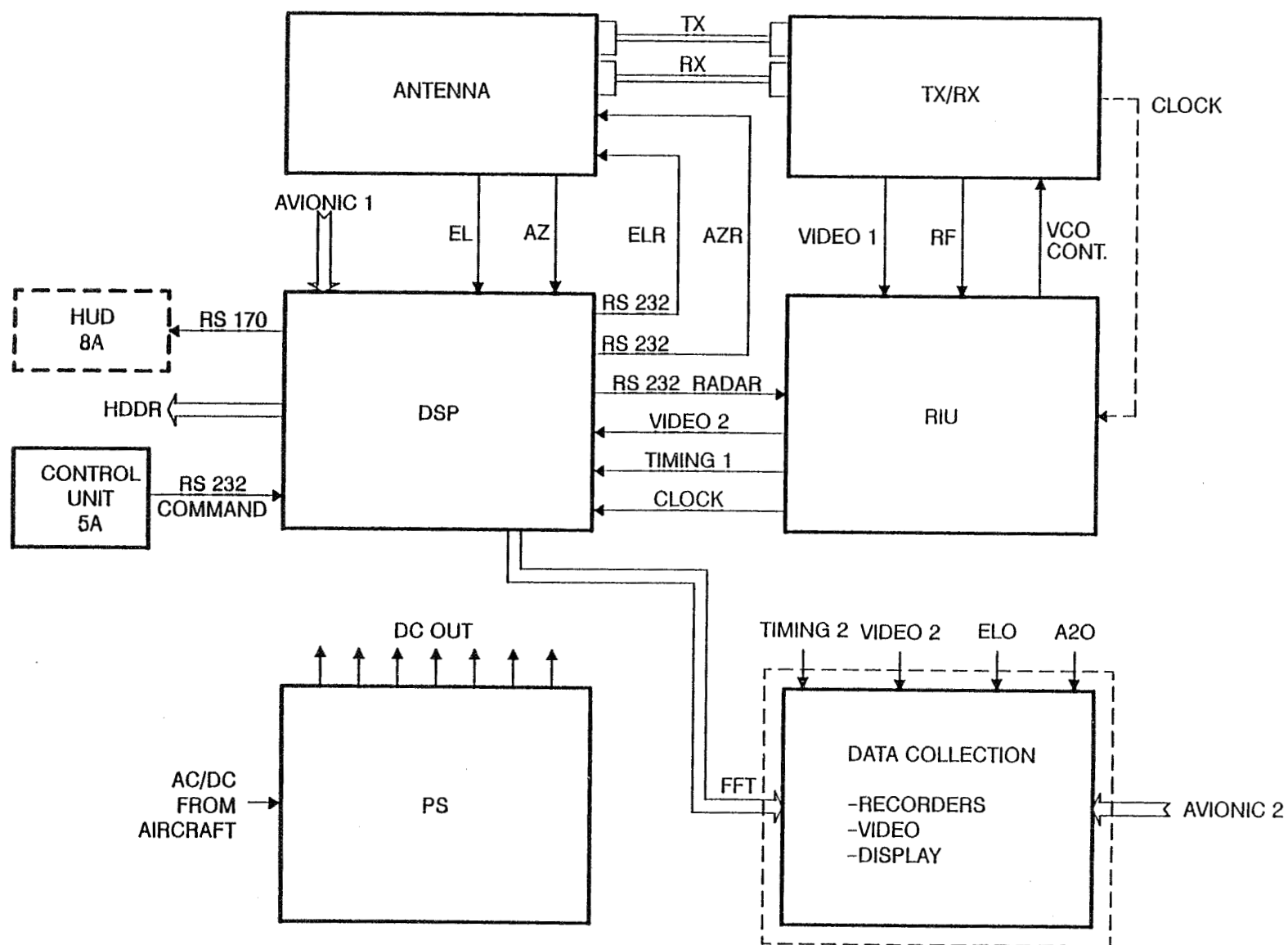


Figure 2. Basic Configuration

Table I. Operational Requirements Summary

RADAR OPERATIONAL SPECIFICATIONS

- Display — C scope (elevation versus azimuth)
- Maximum Processed Range — 6,000 meter, acquisition mode
3,000 meter, approach mode
1,500 meter, taxi mode
 - The acquisition mode is for runway detection and ground map display.
 - Approach mode is selected during the final phase of the landing process (range to runway threshold less than 2,000 meters).
 - Taxi mode is selected during aircraft taxi and takeoff.
- Mode Change — Automatic or manual
- Update Rate — 10 times per second
 - Radar antenna horizontal scan of 10 times per second (5 Hertz) is utilized.
- Scan Angle in Azimuth — ± 15 degrees
 - The horizontal scan covers the total HUD field of view, 30 degrees.
- Elevation Stabilization — ± 15 degrees
 - Adjustment to compensate for aircraft pitch changes to maintain optimum runway illumination.
- Elevation Rate — 30 degrees/second
- Azimuth Resolution — 0.35 degree (5.4 milliradian)
 - Two way antenna azimuth beamwidth.
- Range Resolution — 14 meters for 6,000 meter range, acquisition mode
7 meters for 3,000 meter range, approach mode
3.5 meters for 1,500 meter range, taxi mode
- Azimuth Accuracy — 0.3 degree
 - Azimuth pointing accuracy of 5 meters at 1,000 meters from runway.
- Elevation Accuracy — 0.3 degree
 - Accuracy is affected by altitude and roll data avionic input.

For nonprecision approaches and autonomous OPs the radar must allow the pilot to detect, acquire, and track the SVS scene prior to the Visual Descent Point (VDP), which requires a processed range of about 3000m. A horizontal scan rate of 10X/sec (5 Hz antenna rate) was selected to minimize scene latency. The capability of further extrapolating the scene, using the aircraft state vector to "smooth" the image and decrease scene flicker, has also been incorporated in the system and software design. (Eventually, it may be

desirable to slow the actual antenna scan to be compatible with X-band rate to as low as 1 Hz weather radars.)

The azimuth scan angle of ± 15 degrees was selected to make the scanned scene compatible with typical HUD azimuth fields-of-view. This permits the crew to "see" the runway on the HUD under required cross wind conditions. The antenna is pitch stabilized with a range of $\pm 15^\circ$ to maintain optimum runway illumination in all radar modes and flight path angles.

An azimuth resolution of 0.35 degree was selected as the minimum resolution required to provide the crew an adequate image. This number directly affects the antenna size, hence is an important design parameter that should be verified through simulation and flight test. If larger azimuth resolutions can be tolerated a smaller antenna can be used which would simplify the radome integration problem.

EVS SENSOR TECHNOLOGIES

A 94 GHz FMCW was selected from potential EVS sensors including FLIR, active 35 GHz radar, and a passive 94 GHz radiometer. It was felt that the 94 GHz radar was a mature technology that provided the best overall operational capabilities in low visibility when compared to the other sensors.

FLIR was eliminated as a technology due to poor performance in fog. The extinction coefficient of IR in fog is too large to meet the range requirements of an EVS sensor. The IR sensor may have an application in the taxi mode where the high resolution, TV-like image of the FLIR may be desirable for ground movement and where the visual range requirements are not so demanding.

The most decisive factor in choosing the 94 GHz active radar technology over 35 GHz radar is that the 94 GHz radar yields much better azimuth resolution for a given aperture. To achieve the required 0.35 degree azimuth resolution the 94 GHz radar allows a much smaller antenna size that easily fits into the form factor of existing radomes. The EVS testbed uses a 24 inch antenna. To obtain the same resolution from a 35 GHz radar would require a 64 inch antenna without using some advanced processing technology such as "super-resolution."

Although the 35 GHz radar provides better meteorological parameters, as can be seen in Table II, these only come into play at ranges beyond what is operationally required for EVS. For the ranges of interest, the 94 GHz penetrates the weather adequately, meets the azimuth resolution requirements with a workable size antenna, and is a mature technology at the required transmitted power levels (< 1 watt).

The Frequency Modulated Continuous Wave (FMCW) selected utilizes the change in frequency

to resolve target range. The transmitted signal is swept over a wide frequency range in linear form. The received signal, when mixed with a portion of the transmitter waveform, will produce a beat frequency proportional to the delay introduced by the target range. In the approach mode, the EVS transmitter sweeps 100 MHz within 1.8 msec; this is equivalent to 370.37 Hz for each 1 meter delay.

ANTENNA

The 24" x 8" Flat Parabolic Surfaces (FLAPS) scanning reflector antenna was developed by Malibu Research Associates for the EVS testbed (Figure 3).

This technology is designed such that a flat surface behaves electromagnetically as if it were a shaped reflector. A FLAPS surface is essentially a single large printed circuit board. The feed is fixed and only the lightweight reflector scans ± 7.5 degrees. The antenna produces a 2:1 scan enhancement, which gives a ± 15 degree field-of-view. The FLAPS surface focuses the beam, converts from linear to circular polarization, and forms the COSEC² elevation shaped beam.

The TX/RX is mounted integrally to the antenna assembly behind the reflector surfaces to minimize waveguide losses. The antenna is scanned at 5 Hz (10X through center), in azimuth, and can be pitch stabilized under computer control through a pitch gimbal that has ± 15 degree authority.

RADAR TRANSCEIVER

The 94 GHz solid state FMCW linearized transceiver developed by Marconi Defence Systems, depicted in Figure 4, consists of two LRUs, the RF unit (TX/RX) mounted directly on the antenna and the Radar Interface Unit (RIU) colocated with it. The radar transmitter uses a phase lock loop linearized VCO and an Injection Locked Oscillator (ILO) to produce the 400 mW output power. The received signal is downconverted by an MIC assembly to baseband and then amplified by a digitally gain controlled amplifier stage to produce the frequency/range related signal. The conversion from frequency to range is performed in the system Digital Signal Processor (DSP).

Table II. Meteorological Parameters			
PARAMETER	35 GHz	94 GHz	REMARKS
Attenuation dB/km One Way			
Clear Air	0.12	0.4	
Fog 0.2 gm/m	0.15	0.8	
Rain 5 mm/hr	1.1	4.0	
Rain 10 mm/hr	3	6.3	
Snow 2.5 mm/hr	0.3	1.46	Dry Snow
Backscatter, Circular Polarization Volumetric Clutter (m^2/M^3) $\times 10^{(-4)}$			
Fog 0.2 gm/m	—	—	
Rain 5 mm/hr	0.063	0.25	
Rain 10 mm/hr	0.19	0.44	
Reflectivity (dB).3 Degree Grazing Angle			
Grass (Dry)	-24	-18	
Concrete	< -35	< -30	
Snow (Dry)	-18	-13	
Snow (Wet)	-28	-18	

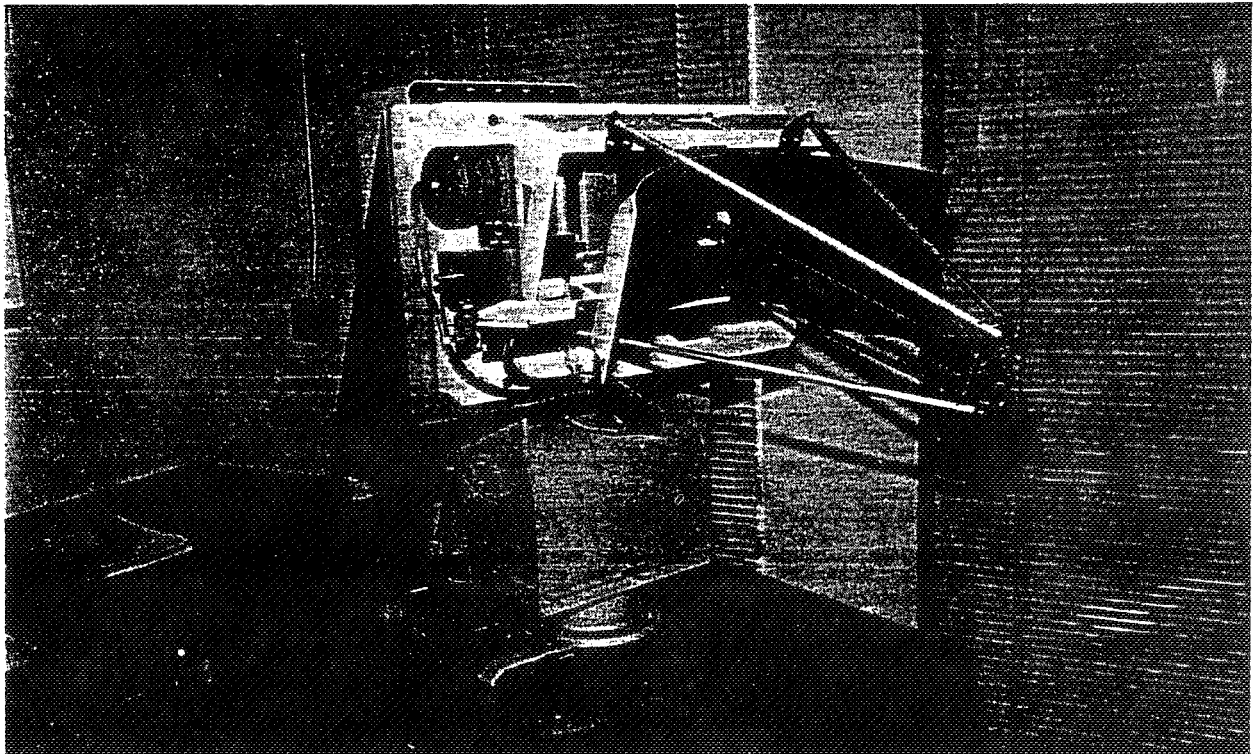


Figure 3. The 24" x 8" Flat Parabolic Surface (FLAPS) Scanning Reflector Antenna

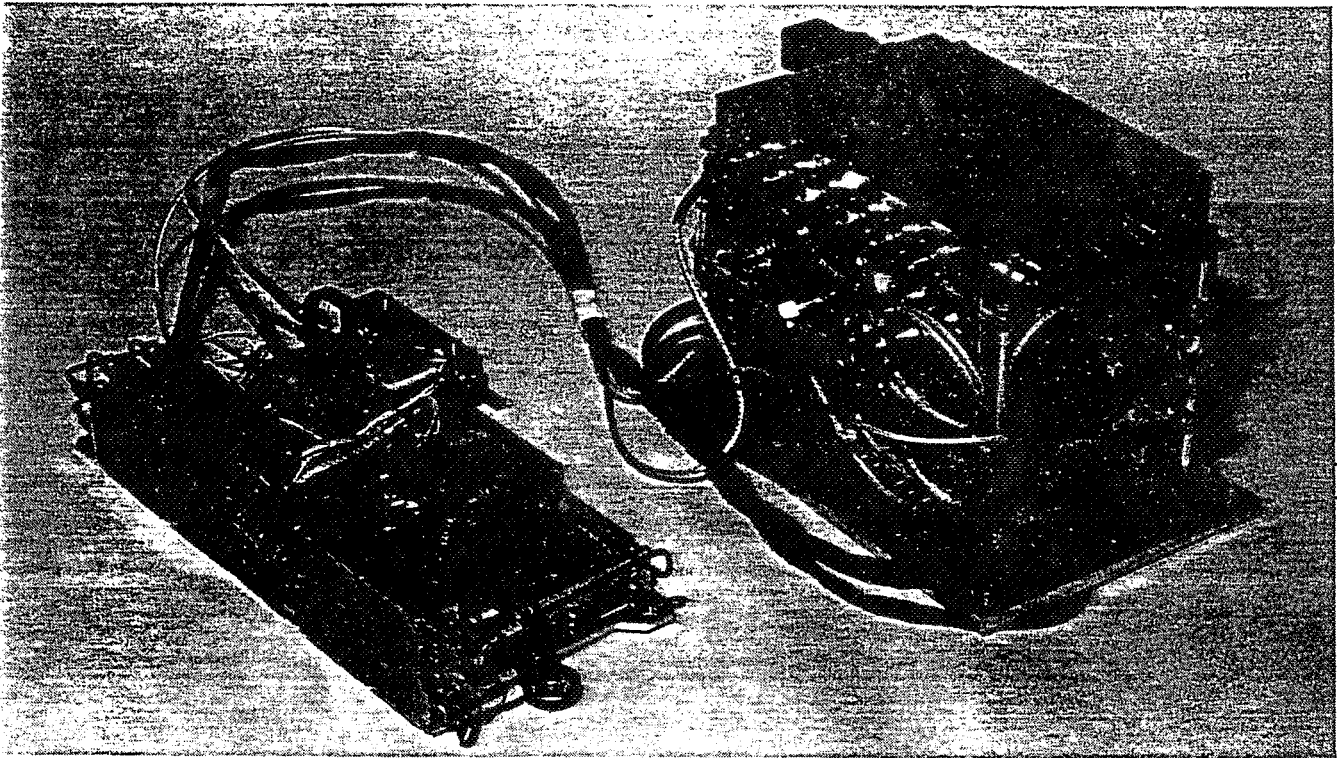


Figure 4. TX/RX Assembly

DIGITAL SIGNAL PROCESSING UNIT (DSPU)

The DSPU, depicted in Figure 5, consists of a fast (400 μ sec conversion time) FFT card, a scan converter, and six RISC architecture MIPS R3000 processor/memory card pairs in a single chassis.

The DSP's primary function is to process a radar return signal and convert it to a displayable picture of the runway scene. The radar return input is digitized and stepped through an FFT calculation, creating 256 range profiles per scene, each consisting of 512 range bins. Each range profile is processed individually to enhance the scene definition. Scenes are processed at a rate of 10 per second. The standard radar B scope (range versus azimuth) is converted, in real time, to C scope (elevation versus azimuth display).

After processing, the range profiles are collected in the scene memory space of the scan converter. Motion compensation of the scene for changes in aircraft attitude may be performed before data conversion to RS-170 output format.

Scene update to the display is at a rate of 30 per second.

The DSP functions include the following:

- Radar return digitization and FFT processing
- Range profile processing
- Scan conversion with motion compensation
- Command and control interface to operator console
- Command and data interface to radar unit
- Data interface to aircraft avionics
- Image enhancement (Level II software)

TEST RESULTS

Starting in May 1991, the radar system was tested in several locations and runway images were collected for evaluation. Since none of the locations has the required 3° glide slope, or the position toward the runway is to the side, the image evaluation is somewhat limited.

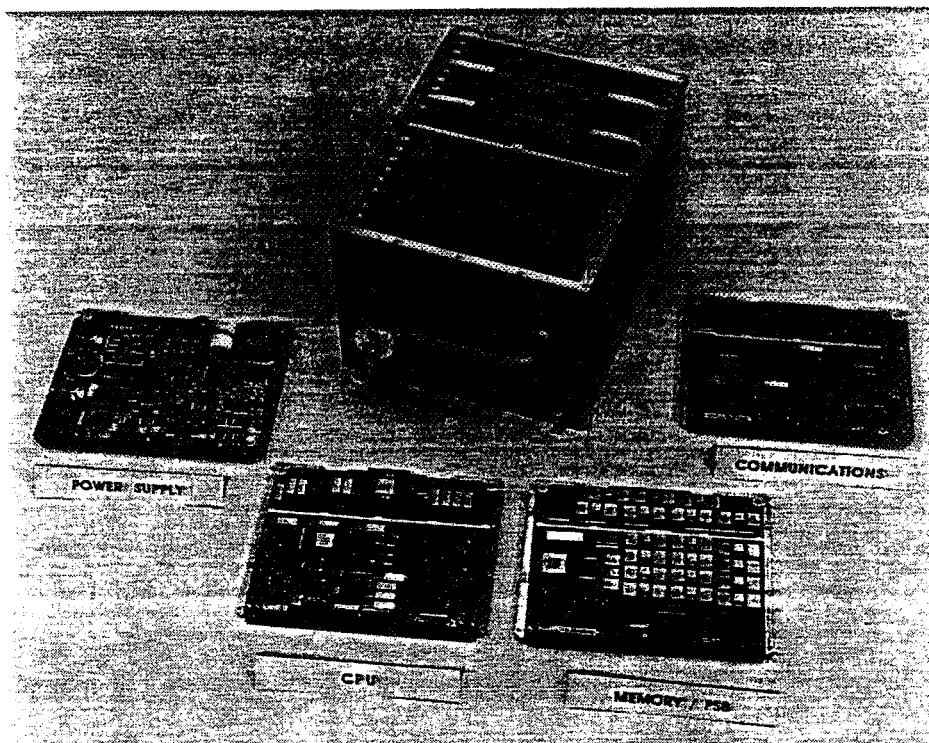


Figure 5. EVS DSPU

Figure 6 illustrates the runway detection from 90° to the side at a very shallow angle ($<1^\circ$). Runway detection prior to touchdown is presented in Figure 7. The runway at a distance of 2,000 to 3,000m is presented in Figure 8. The dark area in front of the runway is the result of the shadow caused by the tree line.

The effect of the DSPU image processing is illustrated in Figure 9. The raw B scope image presented in Figure 9A is converted to C scope (Figure 9B); the image is then smoothed (Figure 9C) and further processed (Figure 9D).

CONCLUSION

The 94 GHz MMW radar system, now being tested at WPAFB, provides a real time runway image up to a distance of 3 km. The runway can be easily discriminated from the grass surrounding it. Utilizing image processing techniques, the image quality can be further enhanced for a clear HUD runway presentation.

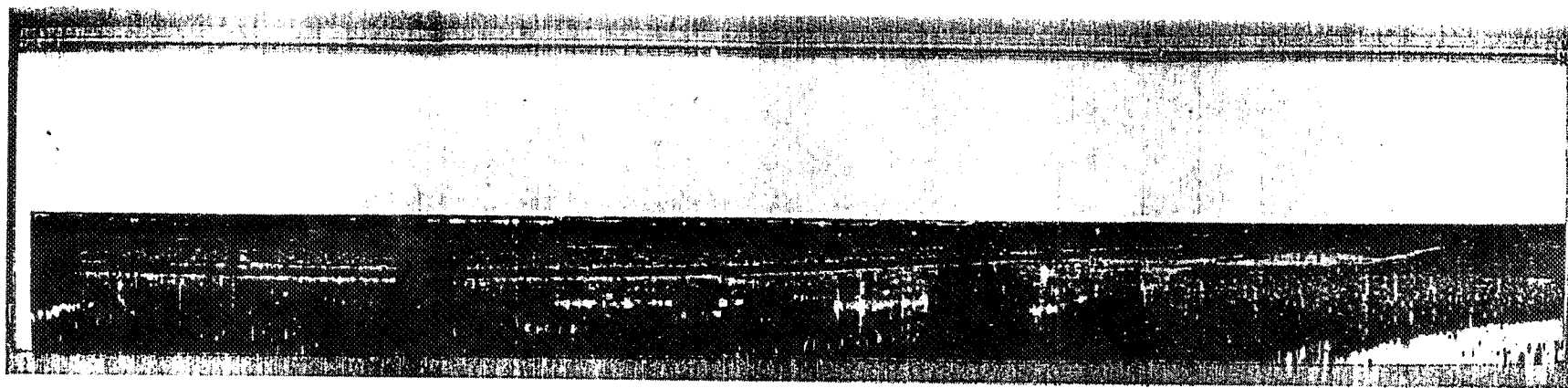


Figure 6. Runway at 90°

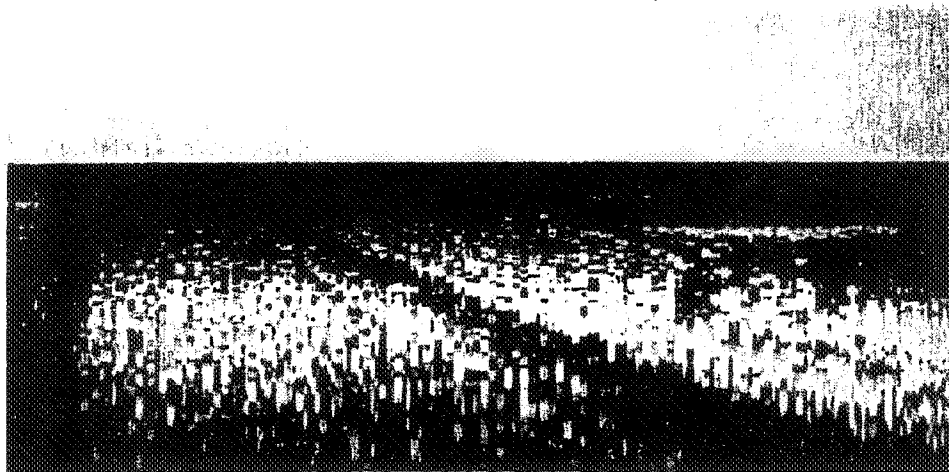
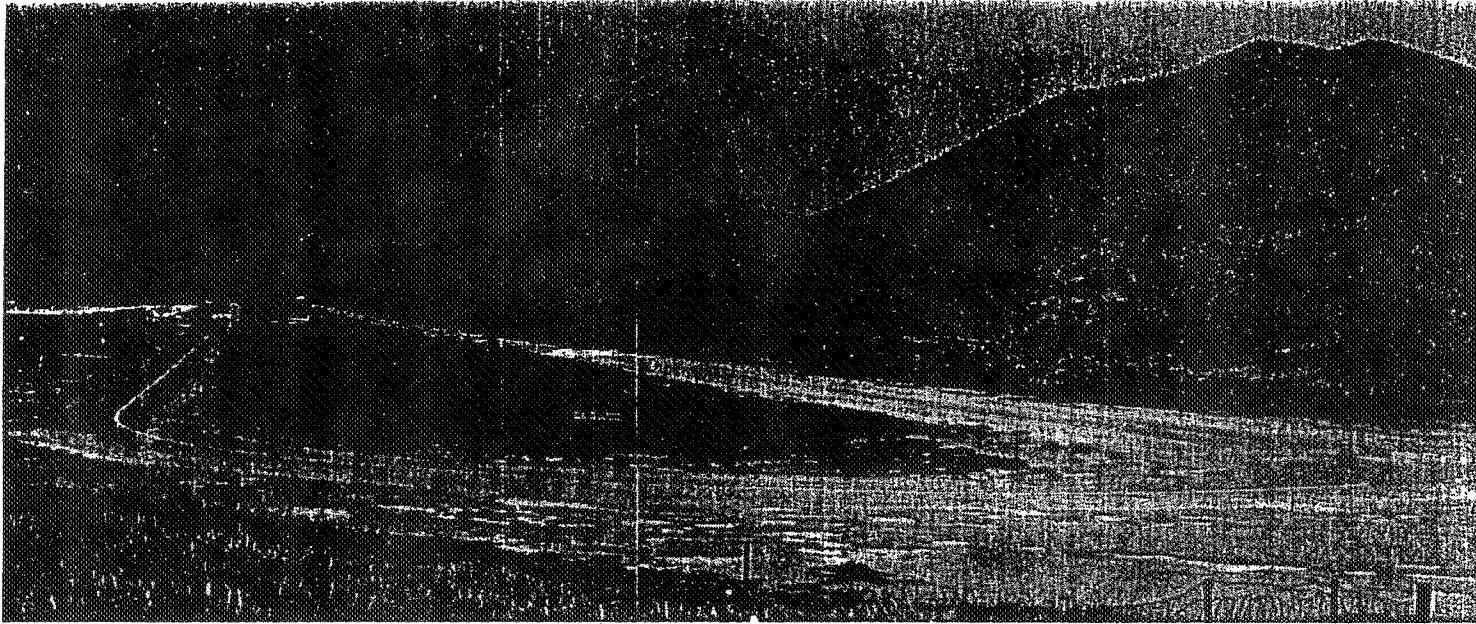


Figure 7. Runway at Short Distance

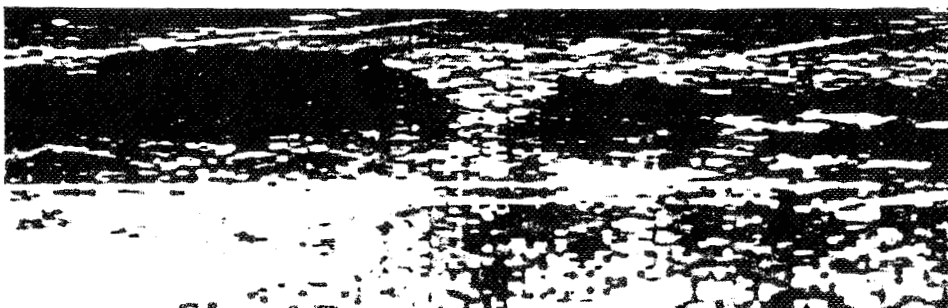


Figure 8. Runway Image, WPAFB



Figure 9A. B Scope Image



Figure 9B. C Scope Image

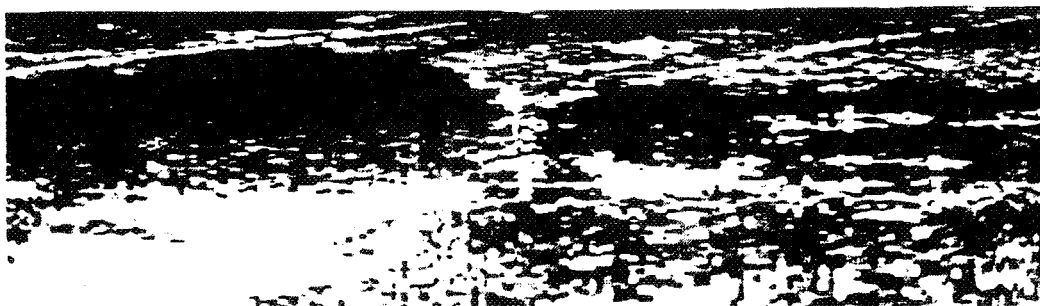


Figure 9C. Processed C Scope Image



Figure 9D. Cluster Process C Scope Image

When "the fog comes on little cat feet," we want to see what it's hiding. The millimeter-wave regime of the electromagnetic spectrum can show us—if we have the necessary vision.

Passive millimeter-wave imaging

by Stephen K. Young,
Roger A. Davidheiser,
Bruce Hauss,
Paul S. C. Lee,
Michael Mussetto,
Merit M. Shoucri,
and Larry Yujiri

54-32
205734
P-18
249916

The regime of the electromagnetic spectrum where it is possible for humans to see is that part where the sun's radiance peaks: the visible regime. In that regime, the human eye responds to different wavelengths of light scattered by objects by recognizing different colors. In the absence of sunlight, however, the natural emissions from Earth objects (at 300 Kelvin) are concentrated in the infra-red (IR) regime. Advances in IR-sensor technology in the last 40 years now make night vision possible. The exploitation of the millimeter wave regime follows a natural progression in the quest to expand our vision, for the great advantage of millimeter-wave radiation is that it can be used at night, in fog, and in other poor-visibility conditions that would normally limit our ability to see.

The millimeter-wave region of the electromagnetic spectrum lies between 30 and 300 GHz, with corresponding wavelengths of 10 and 1.0 mm. It is a region that has not been widely explored for passive imaging for three main reasons: weak natural emission, hardware limitations, and poor resolving power. Objects emit millimeter-wave radiation similar to IR and visible radiation, but that radiation is weak by comparison. The product of emissivity (e) and true physical temperature of an object equals its brightness (or radiometric) temperature. A perfect absorber has $e = 1$ and is known as a blackbody, as opposed to a perfect reflector, which has $e = 0$. The emissivity of an object (which is polarization-dependent) is a function of the dielectric properties of its constituents, its surface roughness, and the angle of observation. (A sample of the measured emissivities of diverse materials at various frequencies is given in the table on the next page.) The radiation intensity of a 300-Kelvin blackbody falls exponentially by about eight orders of magnitude from a peak value in the IR to the millimeter-wave regime at around 94 GHz (Figure 1). This large decrease in intensity is partially compensated for by the lower photon energy that occurs at millimeter-wave frequencies. However, this situation is dramatically reversed in fog and other inclement weather when one takes into account the signal attenuation by atmospheric constituents. Here the strength of the propagated signal peaks in the millimeter-wave region, as the figure shows.

The second reason, hardware limitations, is due to the low millimeter-wave power flux, but is not the problem it once was. Several recent technological advances have enabled the exploitation of millimeter waves. Receivers with mixer front-ends using Schottky-barrier diodes have demonstrated double-sideband noise figures of 6 to 10 dB over the 94- to 300-GHz regime,

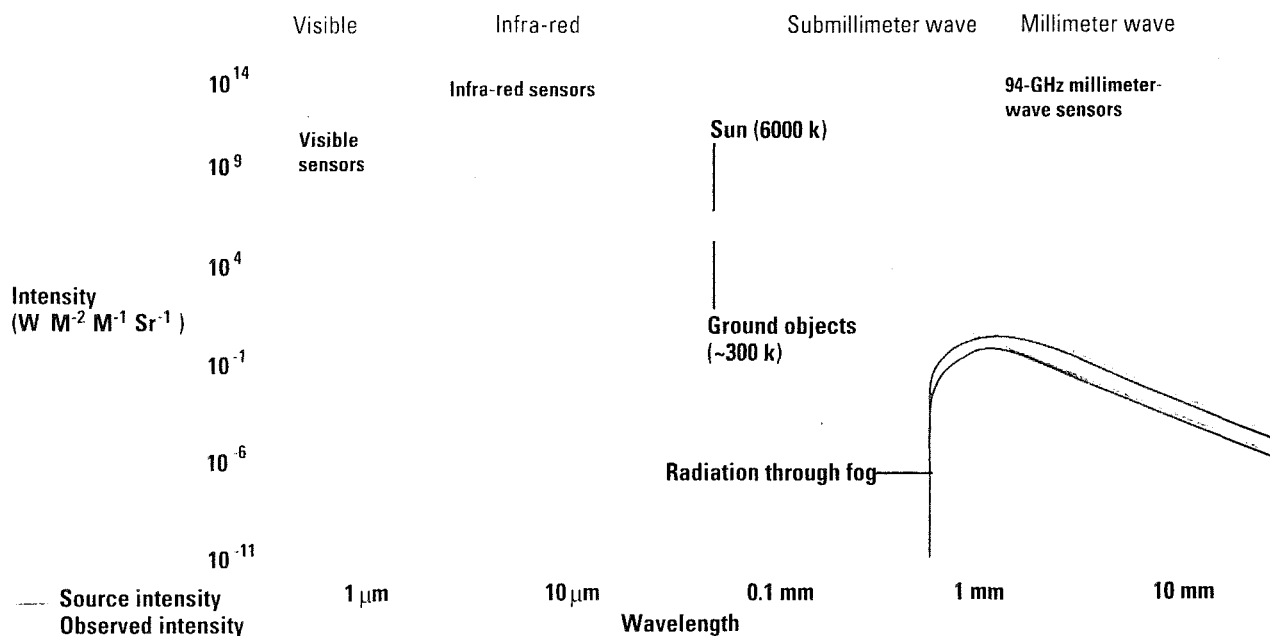
Effective emissivity for vertical look-down assuming specular reflection. The emissivity of an object (which is polarization-dependent) at a given frequency is a function of the dielectric properties of its constituents, its surface roughness, and the angle of observation.

Surface	Effective Emissivity		
	44 GHz	94 GHz	140 GHz
Bare metal	0.008	0.040	0.058
Painted metal	0.034	0.098	0.122
Painted metal under canvas	0.181	0.240	0.299
Painted metal under camouflage	0.222	0.389	0.463
Dry gravel	0.879	0.921	0.957
Dry asphalt	0.891	0.914	0.941
Dry concrete	0.861	0.905	0.946
Smooth water	0.472	0.588	0.662
Rough dirt	1.0	1.0	1.0
Hard-packed dirt	1.0	1.0	1.0

which is adequate for imaging, and high-electron-mobility transistors are demonstrating a 1.9 dB noise figure with greater than 7 dB associated gain at 94 GHz. In addition, supercooled Josephson junctions operating at helium temperatures have even better performance with quantum-efficient detection. Transmission lines and antenna technologies have also kept pace, partly because of the recent interest in radio astronomy applications. The advent of Millimeter Wave Monolithic Integrated Circuit technology has also greatly increased the regime's potential: direct detection and low-noise amplification are now a reality.

The third reason, limited imaging resolution at millimeter-wave frequencies, has traditionally restricted the regime's use to short-range applications. At 3-mm wavelength, and using diffraction-limited optics with a one-meter aperture, the angular resolution is approximately 4 milliradians compared to 12 microradians in the IR region (10-micron wavelength) and 0.7-microradian in the visible region (6,000 angstroms). At a 5-km range, this translates into a passive millimeter-wave spatial resolution of 20 meters, barely adequate for discerning such landmarks as roads and buildings. From a range of 1,000 km, typical of low-Earth-orbit satellite applications, the resolution is 4 km, which again borders on the utility limit for observing mesoscale meteorological phenomena. A typical cloud, for example, is 10 km in extent and

Figure 1. The effect of fog on blackbody radiation observed at a distance of 1 km from the source



the cloud scale of interest is on the order of 100 meters. Again, the situation is changing. With the advent of long-baseline interferometry, millimeter waves need no longer be relegated to coarse-scale applications—the correlation of radiometric signals from receivers separated spatially, the so-called ‘sparse-array’ configuration, has the net effect of increasing the receiving aperture, leading to improved resolution.

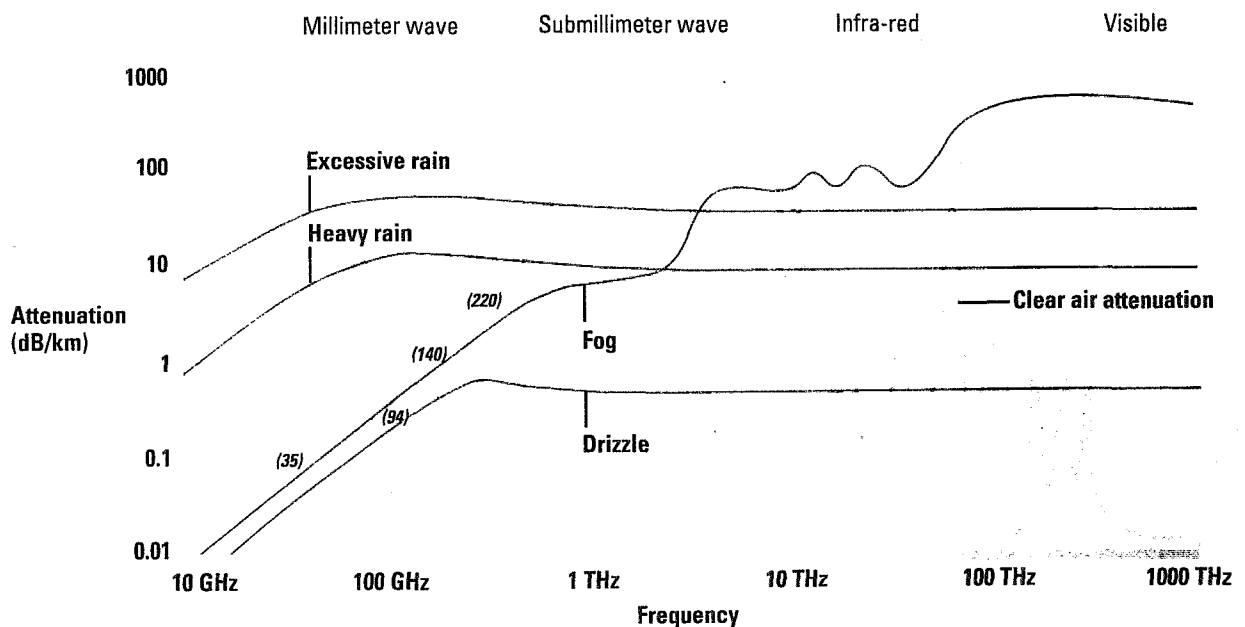
From the standpoint of technology, the time is ripe for millimeter-wave exploitation. At the Applied Technology Division, we have developed a strong phenomenology base for understanding millimeter waves through extensive field measurements and theoretical modeling. Current research in radiometry and interferometry includes such applications as oil-spill monitoring, atmospheric sensing, surveillance, and aircraft landing, as well as millimeter-wave component and subsystem development using superconducting electronics for quantum-efficient detection and low-noise operation.

We are developing millimeter-wave hardware systems. Our approach begins with identifying and defining the applications. System requirements are then specified based on mission needs using our end-to-end performance model. The model has been benchmarked against existing data bases and, where data is deficient, it is acquired via field measurements. The derived system requirements are then validated with the appropriate field measurements using our imaging testbeds and hardware breadboards. The result is a final system that satisfies all the requirements of the target mission.

Phenomenology

Atmospheric propagation. The usefulness of millimeter waves lies in the peculiarities of atmospheric attenuation phenomenologies over the prescribed frequency regime. Figure 2 shows the attenuation of electromagnetic signals in dB/km of propagation path-length from the microwave through the visible regime. This spans the frequency range from 10 GHz to 1,000 THz, with corresponding free-space wavelengths from 3 cm to 0.3 micron. Propagation of electromagnetic waves over this frequency range is subject to continuum

Figure 2. The attenuation of millimeter waves by atmospheric gases, rain, and fog.

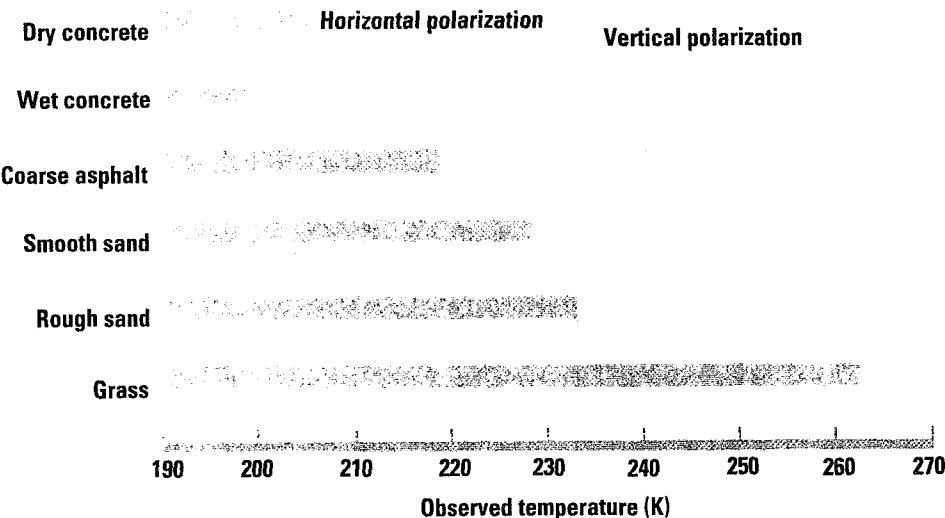


as well as resonant absorption by various atmospheric constituents, including water (in both vapor and droplet form), oxygen, nitrogen, carbon dioxide, ozone, etc. In clear weather, IR and visible radiation propagates with little attenuation. However, water content in the atmosphere in the form of fog, clouds, and rain causes significant absorption and scattering. Conversely, in the millimeter-wave regime, there are propagation windows at 35, 94, 140, and 220 GHz, where the attenuation is relatively modest in both clear air and fog. Even taking into account the much higher blackbody radiation at the IR and the visible, millimeter waves give the strongest radiometric signals in fog when propagated over distances of interest. It is this ability that makes millimeter waves the best candidate for imaging in adverse weather.

While an imaging system benefits from the propagation window in the millimeter-wave regime, an atmospheric-sensing system uses the various molecular absorption lines. For example, the oxygen resonance line around 60 GHz (or 120 GHz) enables temperature-sounding in the atmosphere. Radiometric observations at a number of frequency channels around the oxygen resonance from a satellite platform can be used to unfold the vertical atmospheric temperature profile because atmospheric layers at various altitudes are 'sensed' with different observing frequencies. Essentially, the observed 'brightness' temperature is the result of superposing the radiometric contributions from various layers of oxygen in the atmosphere, less the attenuation of the electromagnetic energy by intervening layers as it propagates toward the observer. Pressure-broadening of the oxygen resonance and the variation of density and pressure with altitude give rise to weighting functions for the various altitude layers in their contribution to the measured brightness temperature at a given frequency in the neighborhood of the oxygen resonance. Similarly, the water-vapor absorption line around 180 GHz allows retrieval of atmospheric moisture profiles. Finally, atmospheric ozone can be monitored by observing the ozone absorption lines around 110 GHz.

Data interpretation and modeling. Figure 3 shows typical 94-GHz scene signatures from various surfaces at grazing incidence angle plotted as a function of polarization. The observed radiometric temperature of a scene is based on the following factors: emissions from scene constituents, reflections of the downwelling sky radiation by the scene, upwelling atmospheric emissions between the scene and observer, and propagation of the electromagnetic energy from the scene to the observer.

Figure 3. Measured grazing angle scene signatures at 94 GHz as a function of polarization and types of surfaces.



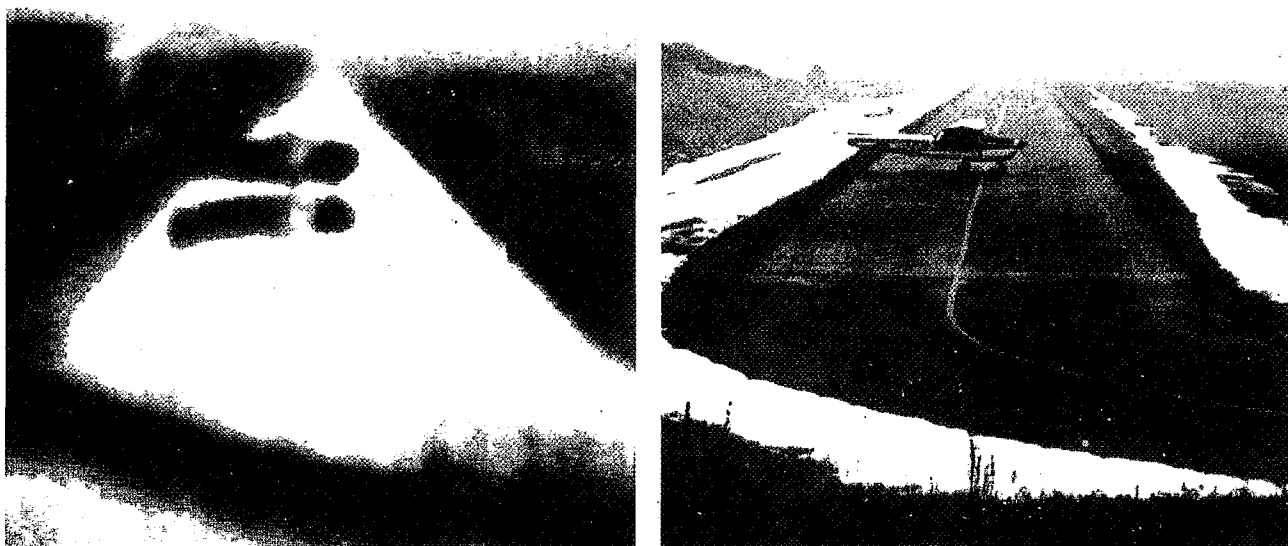


Figure 4. The airport scene on the left was acquired at 94 GHz with a TRW passive millimeter-wave field imaging system. The photo on the right is a visible image of the same scene.

The left-hand photo in Figure 4 shows a millimeter-wave image as measured by the TRW radiometric field imaging system; the right-hand photo shows a visible image of the same airport scene. In the radiometric image, the increasingly darker shades denote increasingly colder temperatures. Thus, the aircraft on the runway appears cold because parts of its metal surface, which is nearly perfectly reflecting, reflect the overhead sky, which is colder than the sky at the horizon. The asphalt runway, on the other hand, although also a good reflector at grazing incidence, reflects primarily the sky at the horizon, which is much hotter. The dirt adjacent to the runway is colder than the runway because the roughness of the dirt surface, although increasing its emissivity at grazing incidence, also mixes the reflections from various parts of the sky, effectively lowering the reflected sky temperature. One interesting feature that emerges from the image is the mirror image of the plane on the asphalt runway. This occurs because the asphalt runway, instead of reflecting the hot sky at the horizon, now sees a colder part of the sky overhead through reflections off the plane. Note that passive millimeter-wave images, unlike radar images, have a visual quality like IR and visible images.

We have developed a sophisticated end-to-end model with four components for the interpretation of millimeter-wave data and for the development of system requirements. The phenomenology model component includes models for the atmospheric propagation effects and meteorology; surface/terrain physics describing the mix of emission and scattering (based on bulk dielectric properties and surface/subsurface geometry) from scene constituents; ray-tracing algorithms for solution of the radiative transfer equation; and the use of combinatorial geometry for constructing complex scenes. Each aspect of the phenomenology model has been individually benchmarked against both measured data and other models in the literature. In addition, the phenomenology model as a whole has been benchmarked against the field imaging data that we have collected. The sensor model component includes the sensor optics, detector, and mechanical/electrical-effects models. It constructs realistic images as seen by the sensor, based on diffraction optics, and includes such effects as finite detector size and noise. The image-processing model component includes image-enhancement and image-restoration techniques. It takes as input raw data from the sensor and applies noise filtering, up-sampling (interpolation), temperature bandpass filtering, contrast

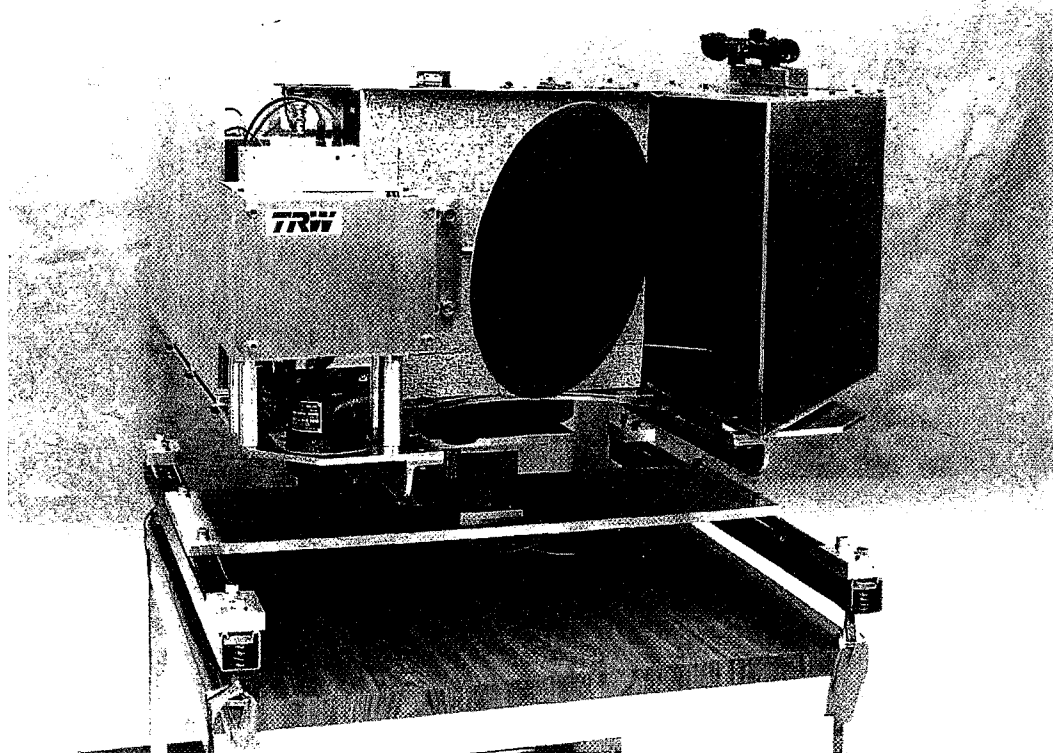


Figure 5. The TRW semiconductor-based multispectral radiometer has a 44-GHz detector channel and an integrated 94- and 140-GHz channel using a Gaussian optics lens antenna.

enhancement, and edge-sharpening techniques to enhance the resulting image. Computer-aided symbology can be superposed on the image to facilitate display and image interpretation. Finally, the display model component captures the enhanced images, frame-by-frame, on video tape for replay at the frame-rate for which the images were produced. Various flight symbolologies (heading, glide-slope, etc.) can also be incorporated in the images to simulate the complete scene a pilot might see on a heads-up display.

Laboratory and field imaging. We have developed multispectral radiometers to provide both ground- and flight-imaging capabilities. Flight and ground systems incorporating these radiometers have been built and used for technology demonstration and for acquisition of images under a variety of weather conditions. Advanced superconducting sensors and associated cryogenics have also been designed, fabricated, and demonstrated in a flight radiometer. For the exploration of high-resolution millimeter-wave imagery, a laboratory interferometer was built to assess sparse-array image collection with model scenes (see Technology Development, below).

The multispectral millimeter-wave imaging radiometers were developed using conventional semiconductor and superconducting detectors, low-noise signal-conditioning electronics, microwave optics for imaging, computerized scene scanning, data acquisition, and image processing and enhancement. Our 'workhorse' semiconductor-based radiometer, shown in Figure 5, consists of a 44-GHz detector channel and an integrated 94- and 140-GHz channel using a Gaussian optics lens antenna. Flight capability for millimeter-wave imaging has been demonstrated by acquiring flight radiometric images using a vibration-isolated, gyrostabilized platform that is mounted in a helicopter (Figure 6).

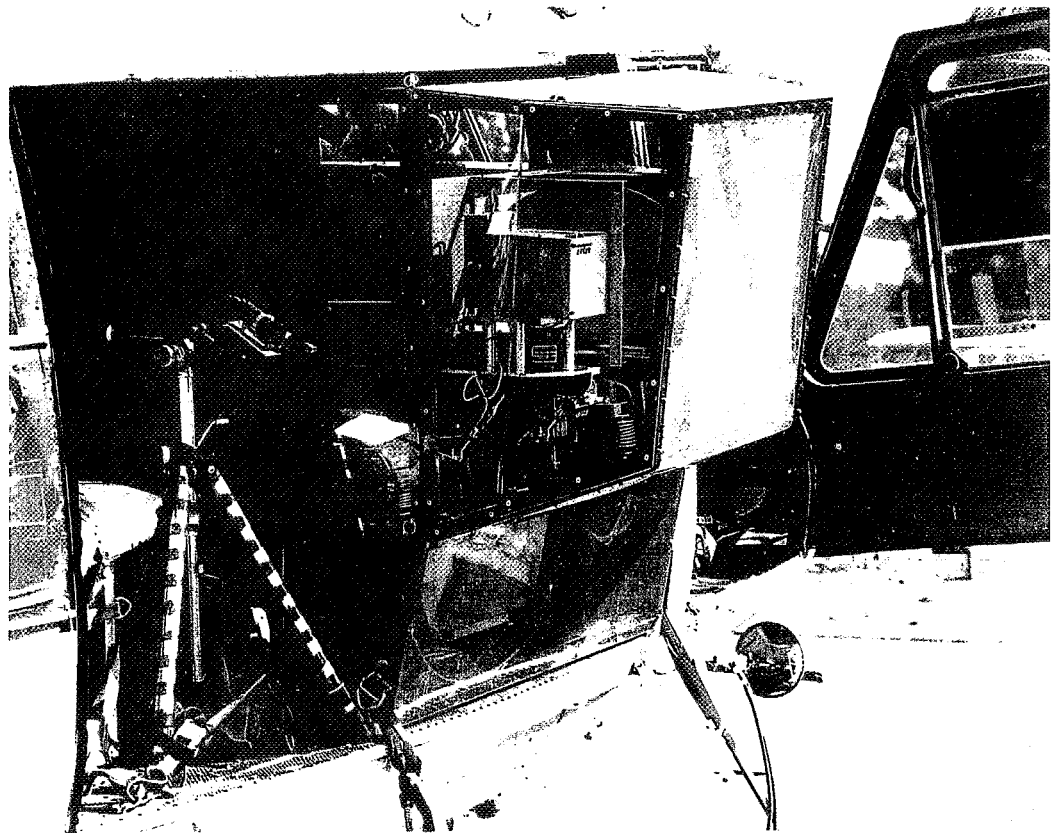


Figure 6. The vibration-isolated, gyrostabilized platform is mounted in a helicopter to give the radiometer flight capability

This instrument has successfully acquired images through clouds and at night, and has imaged special targets such as harbors, ships, boat wakes, refineries, airports, camouflaged vehicles, and oil spills (Figures 7 and 8). Buildings, ships, and rows of storage containers are visible in the harbor image. The oil-spill images were obtained during the Huntington Beach, CA, oil spill of February 1990. An oil layer on the water is highly visible because it acts like an optical coating with varying thicknesses and resultant reflectivities.

Advanced microstrip integrated-circuit superconducting millimeter-wave video detectors for single- and multiple-frequency operation have been designed, fabricated, and tested. Our superconductor-based radiometer uses a two-dewar cryogenic system for separate 35- and 94-GHz tunneling-junction millimeter-wave detectors. This radiometer, like our semiconductor-based instrument, has flight capability using our gyrostabilized platform.

Extensive ground tests with our millimeter-wave radiometers have been conducted. Multi-frequency (44-, 94-, and 140-GHz) studies of imaging phenomenology were performed by measuring the polarization and view-angle-dependent signatures of scene constituents. These include metal surfaces (bare and painted; under canvas, foliage, and camouflage), grass, water, asphalt, concrete, dirt, sand, gravel, and sky. Scenes of military interest, containing vehicles in mixed terrain, have been imaged with 3-meter resolution over several incidence angles from normal to near-grazing.

To support the development of an aircraft landing system for use during low visibility, we have conducted a series of runway imaging tests with the



Figure 7. The top photo is a passive millimeter-wave image of the Long Beach, CA, harbor at 94 GHz. The photo on the right is a visible image of the same scene.

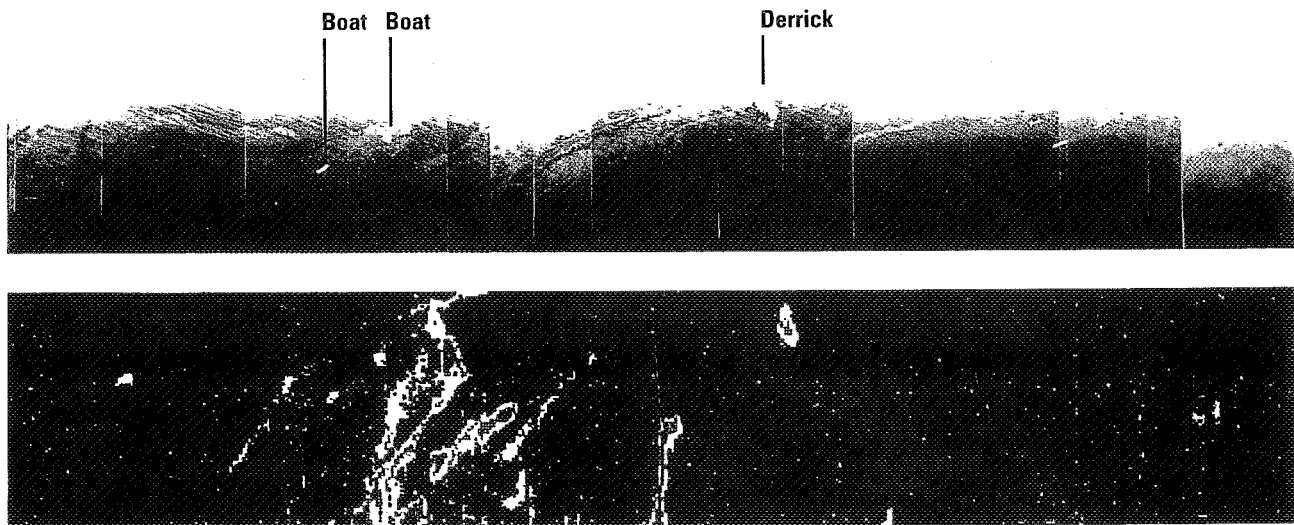
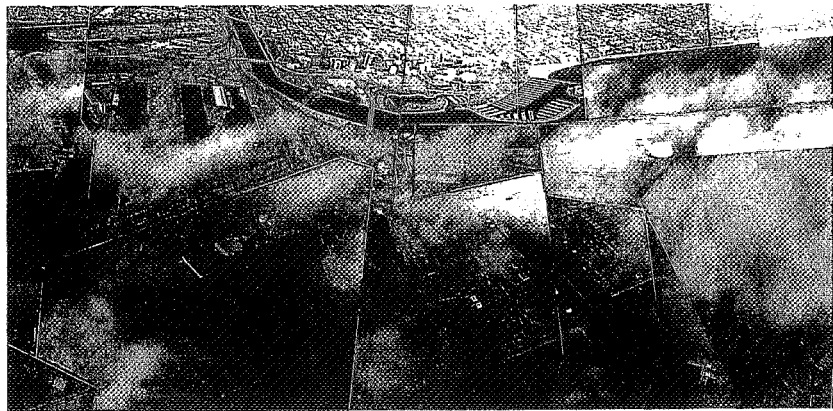


Figure 8. Passive millimeter-wave images of the February 1990 Huntington Beach, CA, oil spill.

94-GHz radiometer, using 4-, 2-, and 1-ft-diameter antennas in fog (Figure 9), rain, and with snow on the ground. These field data serve to validate and benchmark our phenomenology model and define requirements for the aircraft landing augmentation sensor. The airport scene shown in Figure 4 was obtained with this field imaging system. For a potential shipboard navigation system, we have demonstrated the system by imaging a ship (the *Queen Mary*) across a harbor channel (Figure 10).

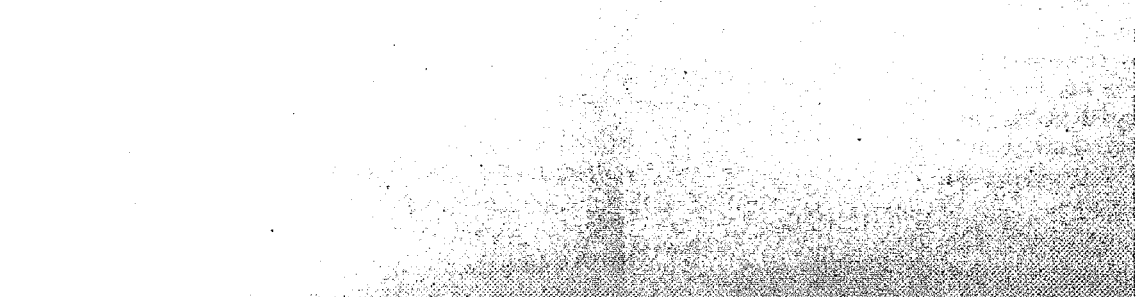
a.



b.



c.



d.



Figure 9. The results of 94-GHz radiometer runway imaging tests: a. and c. show visible images in clear and foggy weather; b. and d. show corresponding 94-GHz images.

Technology Development

The demand for high image resolution drives system development toward high-frequency systems. The millimeter-wave radiometric imaging system resolution is described by the 3-dB spot size of the receiver antenna given as $3\text{-dB spot} = 70^\circ \times \text{Wavelength/Size of optics}$. This equation expresses the fundamental relationship that millimeter-wave image resolution is inversely proportional to frequency and antenna size, and drives the trade-off involved in passive millimeter-wave imaging system development. The goal

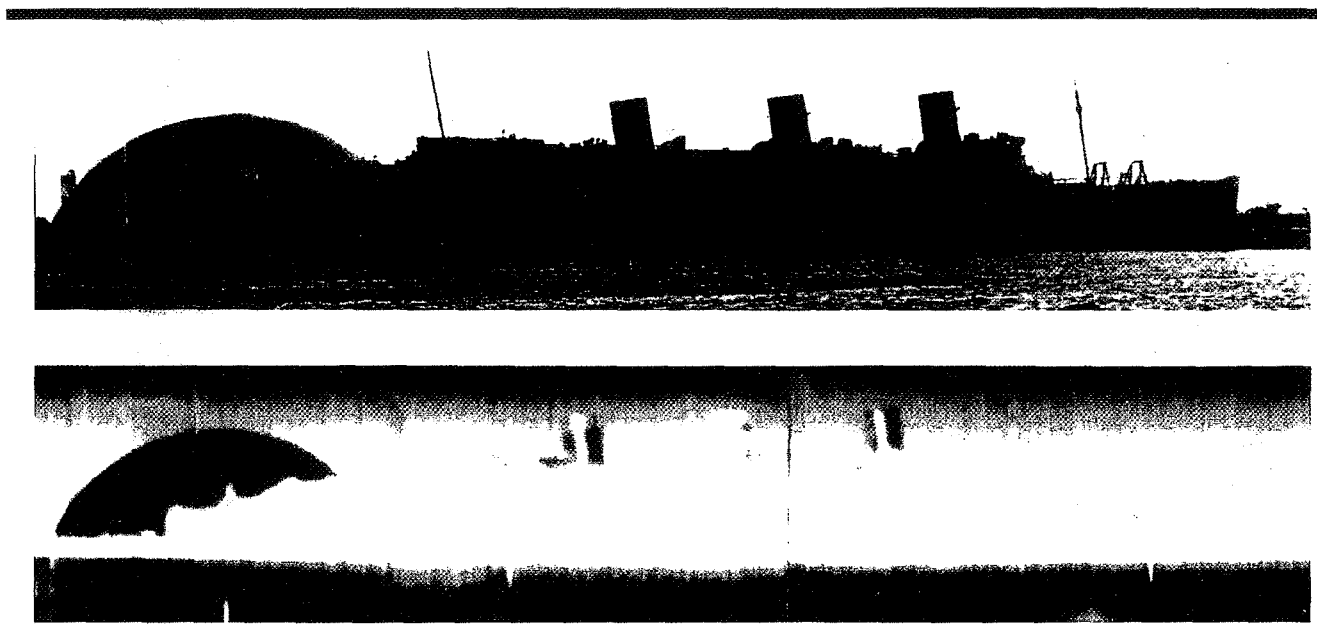


Figure 10. The *Queen Mary* radiometrically imaged at a range of 1,700 ft across the Long Beach, CA, harbor channel at 94 GHz. The dome that covers the *Spruce Goose* appears in the background, left.

is to develop ever-higher-frequency millimeter-wave hardware technology for finer image resolution with a given size optics, or to use higher-frequency hardware to maintain resolution while achieving the smaller and lighter system packaging that is crucial to many applications.

In step with this drive for higher-frequency millimeter-wave technology is the development of a practical system of utility within the bounds of hardware technology maturity and economics. Technology maturity includes sensitivity, compactness, and reliability; technology economics include system affordability, demand, and manufacturability.

Waveguide components and systems. The engineering of waveguide-type microwave component technologies is much better understood and in a more advanced stage of development than are its counterparts, the hybrid and the monolithic printed-circuit microwave components. As a result, development of passive millimeter-wave technologies usually begins with waveguide component building blocks that provide flexibility in design iterations and a much faster engineering process from design to breadboard. After the concept and system design are perfected, the breadboard is then turned into millimeter-wave hybrid systems or highly integrated, monolithic millimeter-wave prototype systems.

We have effectively used off-the-shelf millimeter-wave waveguide hardware to build field-measurement systems for phenomenology measurements, and have also produced numerous high-sensitivity waveguide components. Further development is under way in superconducting heterodyne mixers for higher signal detection sensitivity and in high-temperature superconducting millimeter-wave devices for simpler and more compact application systems.

The MMIC advantage. Innovation in millimeter-wave focal-plane array (FPA) design (see sidebar) using printed hybrid circuit technologies has led to the manufacture of 94-GHz millimeter-wave FPAs for passive imaging applications. Our 8-by-8-pixel passive millimeter-wave camera, built by Millitech Corp., has verified the design and maturity of the hybrid technol-

Imaging a two-dimensional scene with a single millimeter-wave detector is slow because of the large number of picture elements needed for a high-quality image and the per-picture element detector dwell-time needed to achieve the required sensitivity. When imaging a stationary scene, this slow process is acceptable; for a dynamic scene (from a moving vehicle, airborne platform, or satellite), time is simply not available because detector dwell-time will be very limited. A sensitive, high-density image can only be acquired with two-dimensional focal-plane arrays imaging in the video-frame mode or with line arrays imaging in the pushbroom mode.

Two-dimensional focal-plane arrays (FPAs) produce an image much like an everyday video camera that employs a visible FPA. Very high sensitivity in millimeter-wave imaging is achieved with each FPA element by staring at the scene of interest during the entire image acquisition time, instead of scanning through each picture element of the scene. The equation—Sensitivity (K) = Instrument noise temperature/(Bandwidth x Signal averaging time)^{1/2}—shows sensitivity is improved by a factor equal to the square root of the total number of focal-plane elements.

A line array acquires images in the pushbroom mode by mechanically scanning the line array in one dimension, or by mounting the line array on a moving platform and flying the platform over the scene of interest. For the same required number of picture elements, sensitivity is increased by a factor equal to the square root of the number of line-array elements over that which can be achieved by a single detector.

High-resolution images require high-density, tightly packed FPA elements mandated by the image-sampling theorem. The FPA element separation should be as close as possible to 0.5 wavelength. The challenge of implementing millimeter-wave imaging with FPAs resides in the hardware design: it must be a closely packed array of millimeter-wave receivers that is sensitive and is both RF and thermally stable.

In the past year, TRW funded Millitech Corp. to implement their patented breakthrough design that solves the close packaging and stability requirement for FPA fabrication at 94 GHz. This solution (Figure A) uses a combination of heterodyne receivers with an external, quasi-optically injected local oscillator and state-of-the-art, low-power, hybrid-component technology. Compact FPA thermal-loading issues are resolved by the separation of the local oscillator from the FPA assembly. The compact receiver design is made with millimeter-wave printed-circuit technology and with a design that has circuit elements extending from front to back. The receiver circuit component includes a printed antenna; a single-end microstrip mixer; a high-gain, low-noise IF ampli-

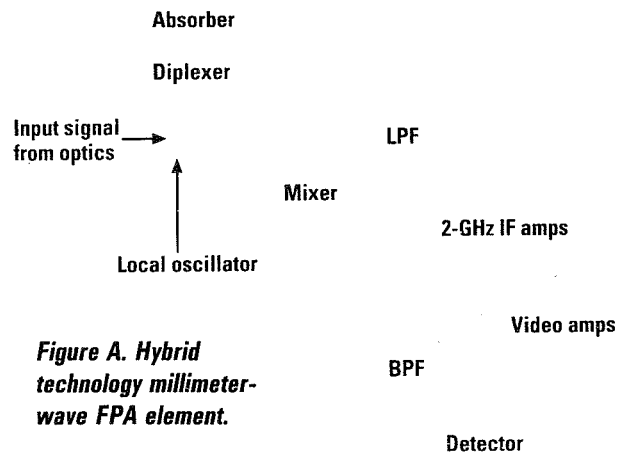


Figure A. Hybrid technology millimeter-wave FPA element.

fier chain with off-the-shelf MMICs; and signal conditioning and multiplexing circuits. The design also takes into account large-FPA manufacturability issues by integrating 8 FPA elements into a single subarray assembly with a single multiplexed signal output. It provides for automated assembly and ease of quality control, and forms the basic building block for large FPA assemblies.

With this 1-by-8-element subarray, a two-dimensional millimeter-wave FPA imaging system will then consist of a two-dimensional assembly of this subarray coupled to the local oscillator assembly; the imaging optics; and an image acquisition, analysis, and display system (Figure B). The TRW/Millitech team proved the maturity of this design and verified the technology's maturity for production-scale readiness. We built an 8-by-8-element, hybrid-technology, 94-GHz heterodyne detection FPA with a quasi-optical injected local oscillator. Provision for field-imaging demonstration was also implemented with 24-in.-diameter lens optics and an image acquisition and display system. The imaging quality of a large FPA was simulated by mosaic-image construction with the 8-by-8-element FPA. We are currently developing a 44- and 94-GHz pushbroom line-array imaging instrument. The line-array design will employ a side-by-side assembly of the 1-by-8-element subarray discussed above.

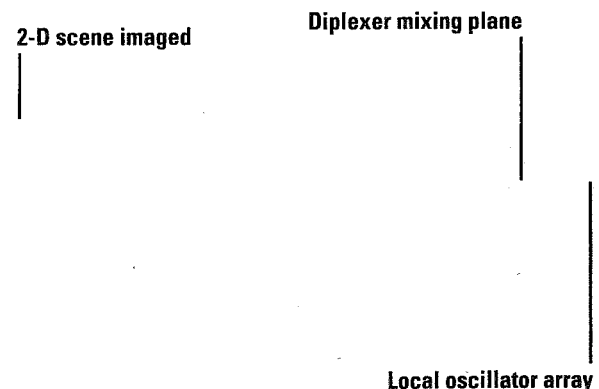


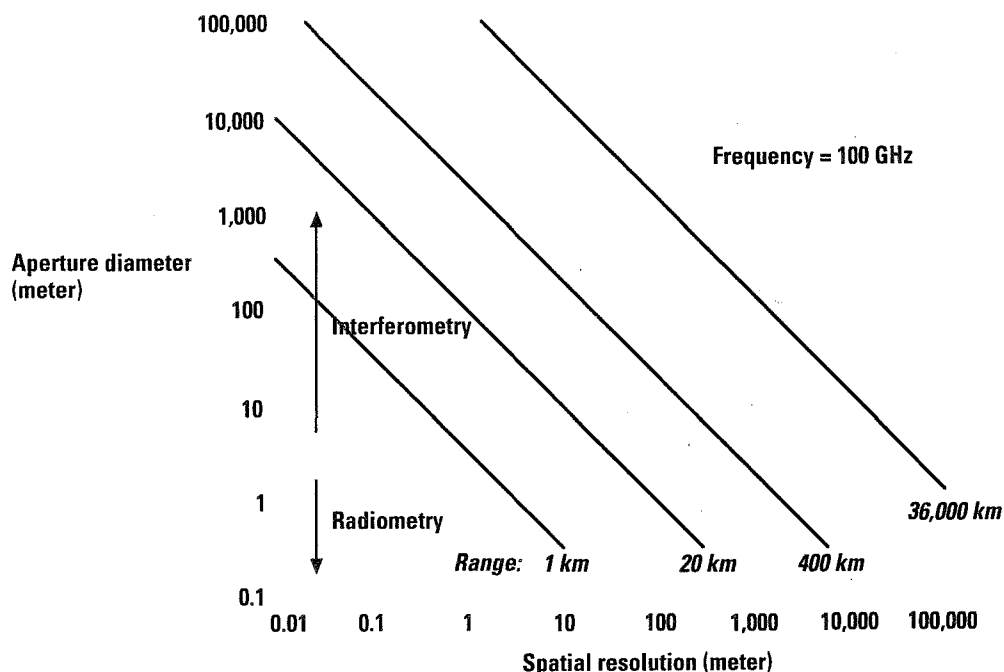
Figure B. Two-dimensional 94-GHz passive millimeter-wave imaging camera.

ogy involved. At the same time, our experience in the design and fabrication of the FPA revealed certain technology areas which, with improvement, would greatly enhance the reliability, manufacturability, and affordability of passive millimeter-wave FPAs. These areas include a highly integrated receiver circuit; an improved local oscillator injection diplexer design; improved local oscillator/receiver coupling to decrease local oscillator power requirements, which can result in a reduced thermal load and increased affordability; and system architecture improvement to decrease the number of functional blocks currently required, thereby improving manufacturability.

TRW's Microwave/Millimeter Wave Monolithic Integrated Circuit and IR&D programs recently produced advances in millimeter-wave amplifier and detector technology that can revolutionize passive millimeter-wave FPA design methodology and manufacturability. The recent millimeter-wave monolithic integrated circuit advances have led to the feasibility of simpler, lower-power-consumption, and more sensitive FPA designs. In the long run, as technology improvement leads to higher-yield monolithic-chip production, integration of many receiver millimeter-wave circuits into a single chip will further simplify the FPA circuit component count and the assembly process, and will result in a more economical final product. Ultimately, as the circuit reduces in size, a planar FPA design will become feasible and millimeter-wave FPA-on-a-chip will be a reality.

Figure 11. Apertures vary greatly with required resolution and range for low-altitude airborne, high-altitude airborne, low-Earth-orbit, and geosynchronous-Earth-orbit applications. Large aperture applications are enabled by interferometry.

Interferometry. A millimeter-wave laboratory interferometer was designed and built to demonstrate the concept of sparse-array interferometric imaging of terrestrial scenes and to evaluate the effectiveness of image-reconstruction schemes. Sparse-array aperture-synthesis techniques from radio astronomy permit the large apertures for high-resolution imaging (Figure 11). In aperture synthesis, also known as long- or very-long-baseline interferometry, the correlated output of antenna pairs sample the wavefront of scene emissions in an area known as the aperture. A two-dimensional inverse Fourier transform allows the scene image to be reconstructed with these samples. Image resolution is determined by the antenna spacing, rather than the physical size of



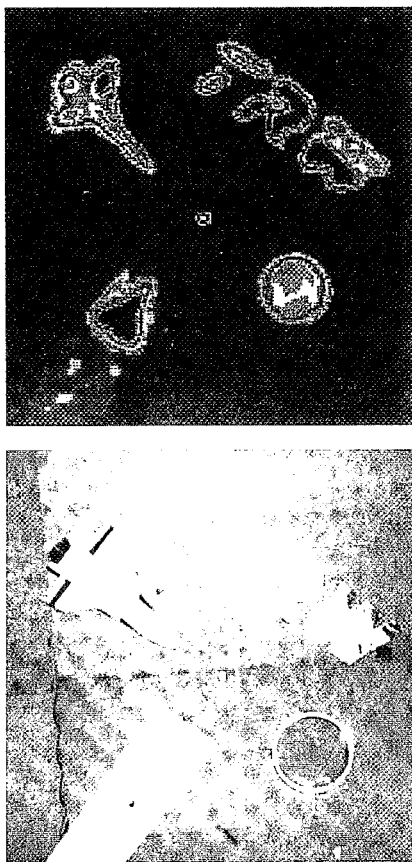
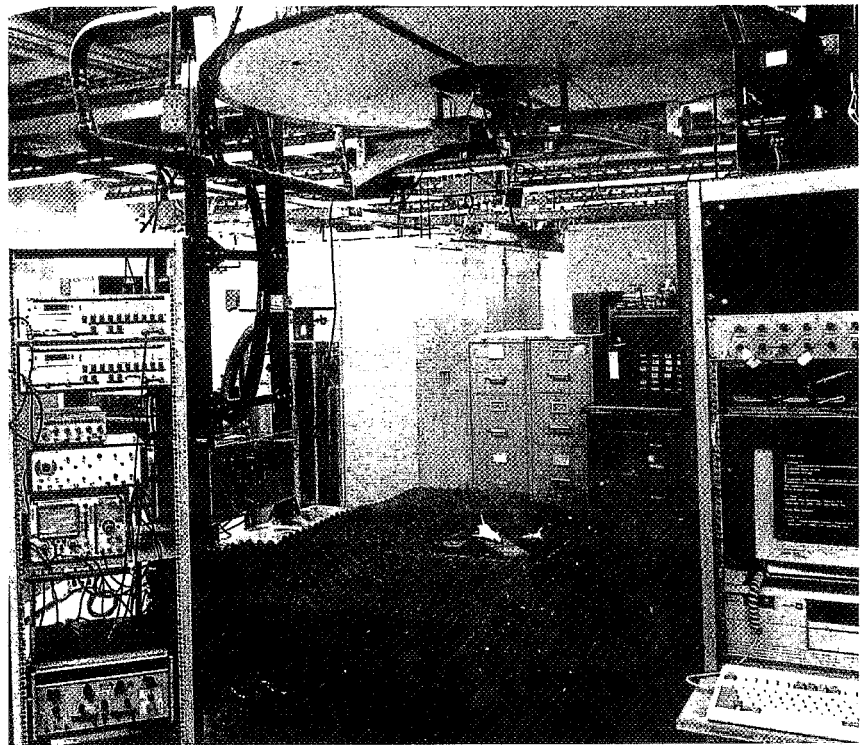


Figure 12. The photo on the right shows the high-resolution interferometer testbed facility. A model interferometer image is shown top left; the visible scene is shown bottom left.



the antennas. This technique has been successfully employed in radio astronomy for high-resolution mapping of extra-terrestrial radio sources, and the resolution now exceeds that achieved by optical telescopes. The application of this technique to Earth observation is now of increasing interest. At TRW, we are investigating appropriate sampling and reconstruction methods.

Our interferometer testbed operates in several frequency bands and contains pairs of millimeter-wave radiometers, a positioning rail for baseline variation, elements of a simulated scene, data acquisition and display electronics, and a cryogenic model sky, the temperature of which can be controlled to simulate illumination conditions. The testbed and sample image data are shown in Figure 12.

Systems Applications

There are a multitude of applications that would benefit from a passive millimeter-wave imaging (PMMWI) system. PMMWI systems can be configured in various ways, depending on the application. A separation into one-dimensional, two-dimensional, and sparse-array designs distinguishes between three general system classes based not only on hardware complexity, but also on the missions to be achieved by each configuration. The first two designs improve the ability to acquire faster frames while keeping good radiometric sensitivity with longer integration times. In other words, they can produce higher-sensitivity images at faster frame rates. The sparse-array design improves the spatial resolution of the imaging, much as is done with high-resolution millimeter-wave radio astronomy.

One-dimensional arrays are used for both fast and slower frame-imaging systems. When used on board a flying platform, a downlooking one-dimensional array is fixed to the aircraft in the cross-track position; the second dimension of the image is obtained by the aircraft motion along-track. The image

obtained is similar to the two-dimensional array image. Its line-scan rate is variable, depending on proper matching of aircraft speed and altitude with sensor aperture. One-dimensional array systems are also used on the ground and other fixed platforms requiring slower frame rates: the pushbroom array uses either mechanically scanned optics or is itself mechanically scanned.

Two-dimensional arrays are usually used when a PMMWI system requires imaging at frame rates similar to visual video cameras, i.e., between 10 and 30 Hz. It takes 5 minutes to obtain one frame of a 100-by-100-pixels image with a dwell time of 30 msec per pixel with a single receiver channel scanning the full 10,000 pixels. With a one-dimensional array of 100 pixels scanning vertically in a pushbroom fashion it takes 3 sec to obtain the frame; with a two-dimensional array staring at the scene it takes 30 msec to obtain the same frame. This latter choice has the distinct advantage of providing real-time imaging similar to visual and IR video cameras.

Finally, an array farm, a distribution of either one- or two-dimensional multiple arrays with a baseline between each, forms a sparse array that can be used for high-resolution imaging. The technique is similar to radio astronomy and is employed in instances where a very large, solidly filled aperture cannot be implemented to support the required spatial resolution. The following paragraphs describe sample applications that show the utility of PMMWI systems.

PMMWI for the Landing Mission. The ability to take off, land, roll, and taxi in fog and low cloud ceilings has long been a high priority for both military and commercial aviation. Such capabilities hold high tactical military value as well as significant commercial gain for the airline industry. Attempts to achieve this mission have been made in the past, but none holds as much promise as millimeter-wave imaging, because it can be an autonomous method with the unique advantage of giving the pilot an image of the forward-looking scene that he otherwise would not have in adverse weather. Equipped with a millimeter-wave sensor, accidents caused by fog and low-visibility conditions, either in the air or on the ground, could be avoided.

Currently, commercial jet aviation can land in low-visibility conditions (Cat III weather) only with planes equipped with an auto-pilot landing system and on runways equipped with two Instrument Landing Systems (ILSs), also called Category 2-type runways. In Cat III weather, the autopilot, using the double ILS electronic guidance, controls the hydraulic systems of the aircraft and brings it down on the runway automatically without the pilot being 'in the loop,' because he cannot see the forward-looking scene. Not only are these landings uncomfortable to pilots and limited to Category 2-equipped airports (and there are only thirty-five in the U. S.), but they are also not economical for the airline industry because of costs associated with tighter instrument tolerances, higher levels of equipment maintenance, and pilot training, as well as the limited availability of equipped aircraft/facilities.

The proposed concept for a pilot-in-the-loop, adverse-weather system for take-off and landing is a millimeter-wave sensor operating at any of the propagation windows of 35, 94, 140, or 220 GHz. Most of the currently proposed systems lie in the 35- or 94-GHz frequency windows because the millimeter-wave electronics hardware at these frequencies is both more mature and less expensive than at 140 or 220 GHz, and fog penetrability is greater.

In 1989, the Federal Aviation Administration, together with the Air Force, issued a program research and development announcement, called Synthetic Vision, to solicit bids for millimeter-wave sensors capable of carrying out the mission. TRW's PMMWI camera concept was one of the four winners selected for the first study phase.

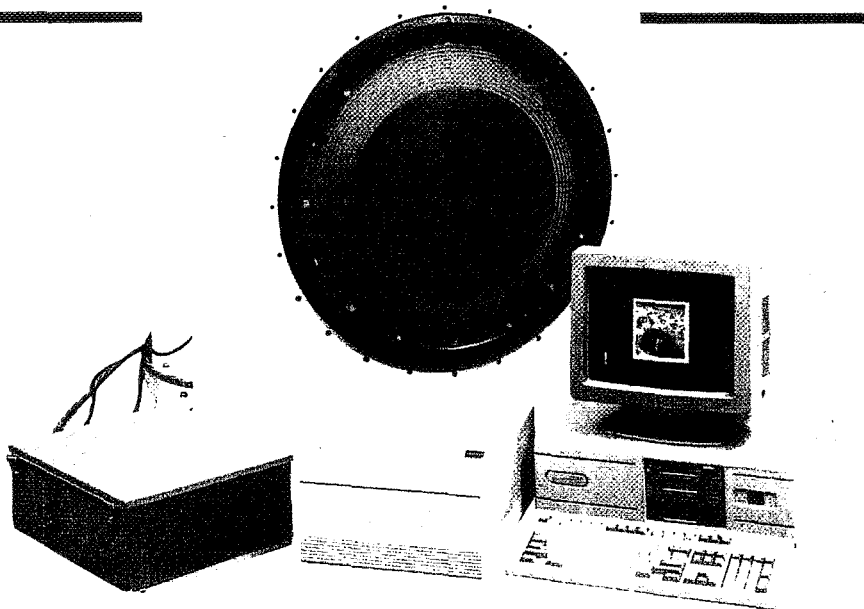
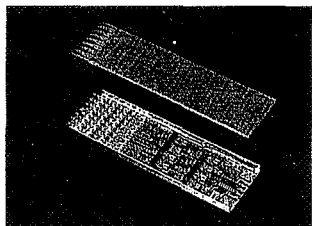
The civilian take-off and landing mission can be met with different types of millimeter-wave sensors. For the airline industry, both an autonomous and a beacon-aided system have been suggested. Some active systems use stored maps and a terrain-reconnaissance/terrain-mapping radar similar to those used in seeker missiles. Millimeter-wave beacons can be used on the ground similar to landing lights at night. While both of these schemes are feasible when the landings occur on specific major airfields, generalizing the concepts to all airfields is almost impossible because of the high cost involved. General aviation, which is most of the non-airline part of the civilian sector, would not benefit from these systems. For example, air carriers of overnight delivery packages use many non-major airfields and such systems would be too expensive for them.

The TRW PMMWI system, however, has the unique capability of giving the pilot a literal, visual-like image of the forward-looking scene. It is autonomous in that it needs no ground assistance or other knowledge-based system; it can, if needed, operate with the assistance of ground-based beacons, an on-board flight-guidance system, or in conjunction with other imaging sensors such as IR or visual cameras. Thus, the TRW concept is a general one suitable for multiple users and missions. The TRW PMMWI video camera is designed to respond to all the requirements of the take-off and landing mission: operate in fog, low visibility, and adverse weather conditions; provide the pilot with a good resolution image of the forward-looking scene; provide adequate field-of-view for runway acquisition, landing, roll-out, and taxi; and provide real-time quality display of the acquired images.

The millimeter-wave radiometric image is displayed to the pilot on a heads-up display that allows him to see through and recover the visual scene whenever fog subsides and visibility conditions improve during the landing. This gradual transition from millimeter-wave to visible image is only possible with radiometric sensors like passive millimeter wave and IR because of their visual-like image; active radars cannot provide this capability for the look angles required during landing and take-off. The TRW concept is a two-dimensional staring focal-plane array, operating at the 94-GHz propagation window frequency, using a lens with a resolution < 6 milliradians, a field-of-view as large as 30° horizontal by 20° vertical, and an adjustable frame rate of 10 to 30 Hz. With an aperture resolution of 6 milliradians, the number of focal-plane-array receivers required to yield the full field-of-view is 80 by 56, or just under 5,000 receiver pixels. To prove the concept's feasibility, TRW and Millitech Corporation implemented an 8-by-8-pixels bread-board demonstration camera that performs most of the features of a large-array camera. Figure 13 shows the camera with its 24-in. transmission optic lens and its PC-based data-acquisition system.

Other applications. The TRW PMMWI sensor is the ideal sensor for many military missions. A major feature of the landing sensor is its covertness: it produces almost no emanations, which makes it highly desirable for military applications. We envision multiple applications for the PMMWI sensor and are working with the services to determine their specific requirements.

Figure 13. The photo on the right shows the TRW 8-x-8 pixel breadboard demonstration camera. The photo below shows the 8-pixel focal-plane array module.



Similar to visual and IR video cameras, the PMMWI video camera is a great asset for the surveillance mission. It can perform many of the missions that visual and IR cameras cannot perform during fog and poor-visibility conditions. While the price is usually decreased resolution, in many of the applications of interest the resolution is good enough for the detection of targets of interest. Some examples of these applications include ground surveillance of traffic in airports, at borders, at harbors and water channels, and on-board ships and armored vehicles. The camera can also be used for remote sensing, for Earth monitoring, and for ground or sea surveillance. In these applications, aperture synthesis may be needed, depending on the resolution required.

Millimeter-wave radiometric images discriminate between various vegetation canopies, sand, concrete, asphalt, metals, ice, snow, and water. An air- or spaceborne sensor can also discriminate between different states of some materials: old and new ice, for example, coniferous trees with needle-like leaves and trees with flat leaves, dry and wet snow, and calm and agitated seas. The ability of millimeter-wave radiometry to discriminate between different fluids is useful in locating oil spills at sea, and in determining relative thickness and volume.

Sector Involvement

It is important to note that the technology and implementation of passive millimeter-wave imaging is not limited to the Applied Technology Division; there is a broad-based involvement by all the groups in the Space & Defense Sector. For example, many segments of the Space & Technology Group will be working on PMMWI sensors, sparse-array technology, space payloads and missions, and analyses and systems engineering efforts. Insertion of MMIC technologies in the sensor's hardware design and VHSIC technology for real-time image processing and display are tasks for the Electronic Systems Group. The Avionics & Surveillance Group is currently directing the aircraft landing mission and is chartered to implement airborne surveillance applications as well.

In all of these efforts, TRW's work—in the investigation of millimeter-wave phenomenology, the development of imaging systems, and the demonstration of systems—is enabling a whole new generation of low-cost, compact, imaging applications.

Recommended Reading

P. Bhartia and I. J. Bahl, *Millimeter Wave Engineering and Application*, John Wiley and Sons, Inc., New York, NY, 1984.

J. C. Fischer, Ed., *Passive Microwave Observing from Environmental Satellites*, NOAA Technical Report NESDIS 35, November 1987.

M. I. Skolnick, Ed., "Passive Detection," by D. D. King, *Radar Handbook*, McGraw Hill, New York, NY, 1970.

A. R. Thompson, J. R. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, John Wiley and Sons, Inc., New York, NY, 1986

F. T. Ulaby, R. K. Moore, and A. K. Fung, *Microwave Remote Sensing*, Addison-Wesley Publishing Co., Reading, MA, 1981.

The Authors

Stephen K. Young is manager of the Electromagnetic Technology Department in S&TG's Applied Technology Division where he is responsible for developing active and passive millimeter-wave technologies and systems applications of strategic interest to TRW. These activities range from oil-spill surveillance to aircraft identification. Young has a BS from St. Joseph's University, Pennsylvania, and an MS and a PhD from the University of Maryland, all in physics.

Roger A. Davidheiser is a senior staff scientist with responsibilities in passive millimeter-wave technology and interferometry, as well as high-temperature superconducting electronics. He has a BS in physics from the California Institute of Technology and a PhD in electrical engineering from the University of Southern California.

Bruce Hauss, who is the area leader for computational physics in the Electromagnetic Technology Department, is the leader of the TRW passive millimeter-wave modeling effort. He has a BS in physics and a PhD in nuclear engineering, both from the University of California at Los Angeles.

Paul S. C. Lee is a project engineer in the Electromagnetic Technology Department and is responsible for designing PMMW systems for ground-based, airborne, and spaceborne applications. He received a BS, an MS, and a PhD, all in physics, from the University of California at Los Angeles.

Michael Mussetto is the assistant manager of the Electromagnetic Technology Department and is responsible for the development, test, and demonstration of passive millimeter-wave field- and flight-imaging systems. He is also the test director for flight operations. Mussetto received a BS from the Massachusetts Institute of Technology and a PhD from Yale University, both in engineering and applied physics.

Merit M. Shoucri is manager of the Air Force program for autonomous landing using a passive millimeter-wave sensor. He was formerly manager of the AF/FAA Synthetic Vision Sensor Development program. Shoucri received a BA in electrical engineering from Alexandria University and a PhD, also in electrical engineering, from the University of Southern California.

Larry Yujiri is a staff engineer/scientist responsible for the design and test of millimeter-wave imaging systems used in phenomenology investigation and applications demonstrations, and a team leader on image-acquisition field tests. He received a BS in physics from Harvey Mudd College and an MS and a PhD in nuclear engineering/plasma physics from the University of Wisconsin-Madison.



**(Back row) Merit M. Shoucri, Larry Yujiri, Bruce Hauss, and Roger A. Davidheiser.
(Front row) Michael Mussetto, Stephan K. Young, and Paul S. C. Lee**

This article, and the work it describes, is the result of the combined efforts of the whole ATD passive millimeter-wave team. The authors would therefore like to express special thanks to R. Ackerman, H. Agravante, D. Arnush, H. Boehmer, C. Carlson, J. Cragin, T. Christensen, S. Covington, D. Demers, D. Dixon, A. Eskovitz, M. Figueroa, L. Harnett, K. Herrin, O. Horne, R. Horton, C. Jackson, S. Korn, M. Iskra, C. Leavitt, K. Luebke, R. Marohn, R. Mendoza, P. Moffa, D. O'Keefe, J. Pearlman, G. Robertson, T. Romesser, G. Rosiak, H. Rosner, T. Samec, S. Salem, B. Stark, W. Steele, S. Talmadge, S. Texter, G. Uyeno, R. Wrape, M. White, C. Woo, K. Yokoyama, and J. Yaung. We would also like to thank the following people for their contribution: B. Allen, D. An, J. Berenz, C. C. Chen, D. Chow, S. Dow, M. Huang, and W. Lam of ESG; M. Hoover, L. Peterson, D. Spencer of ASG; and R. Huguenin and his technical staff at Millitech Corporation.

Application of Aircraft Navigation Sensors to Enhanced Vision Systems

Barbara T. Sweet
NASA Ames Research Center

249517
S5-04
205 735
P-17

ABSTRACT

In this presentation, the applicability of various aircraft navigation sensors to enhanced vision system design is discussed. First, the accuracy requirements of the FAA for precision landing systems are presented, followed by the current navigation systems and their characteristics. These systems include Instrument Landing System (ILS), Microwave Landing System (MLS), Inertial Navigation, Altimetry, and Global Positioning System (GPS). Finally, the use of navigation system data to improve enhanced vision systems is discussed. These applications include radar image rectification, motion compensation, and image registration.

Application of Aircraft Navigation Sensors to Enhanced Vision Systems

Barbara Sweet

Flight Human Factors Branch

80



Ames Research Center

Outline

- **Current Accuracy Requirements**
- **Current Precision Landing Systems**
- **Inertial Navigation**
- **Altimetry**
- **GPS**
- **Image Processing Applications**

FAA Requirements for Navigational System Accuracy:

Non-Precision Approach:

Limited to 250 ft above surface

100 m 2 drms lateral position accuracy

Precision Approach:

Category I:

Vertical: +/- 1.4 m 2 sigma

Lateral: +/- 17.1 m 2 sigma

Decision Height 200 ft/61 m

Category II:

Vertical: +/- 1.7 m 2 sigma

Lateral: +/- 5.2 m 2 sigma

Decision Height 100 ft/30 m

Category III:

Vertical: +/- .6 m 2 sigma

Lateral: +/- 4.1 m 2 sigma

Decision Height 50 ft/15 m

drms = distance root mean square



Ames Research Center

Current Precision Approach Systems

Instrument Landing System

Come in three categories (I, II, III)

Straight-in Approach to Airport

Requires Glideslope & Localizer Transmitter for each runway threshold with an ILS approach

Microwave Landing System

Come in three categories (I, II, III)

Supports both Straight-in & Curved Approaches

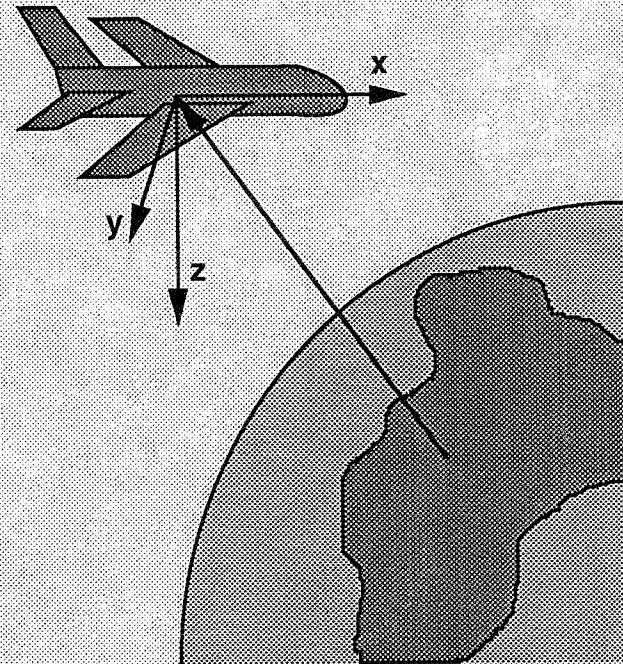
Requires Glideslope & Localizer Transmitter for each runway threshold with an MLS approach

Inertial Navigation

Method:

Inertial Measurement Unit (IMU) measures accelerations and angular rates with respect to three orthogonal axes.

Coordinate transformations/integrations to determine position, attitude with respect to the earth.



Types:

Platform & Strapdown

Limitation:

Lateral positioning only. Vertical position not feasible.

Inertial Navigation

Accuracies output from the IMU:

Acceleration: 6 g to 14 bit plus sign resolution = .00037 g

Angular rates: 256 deg/sec to 14 bit plus sign resolution = .015 deg/sec

Groundspeed: 6 knots

Position: drift rate 1 nm/hr

Pitch, Roll attitude: .25 deg

True Heading: 10 arc min

Track: function of groundspeed, nominally 3 degrees at approach

all accuracies 1 sigma

Altitude Measurement

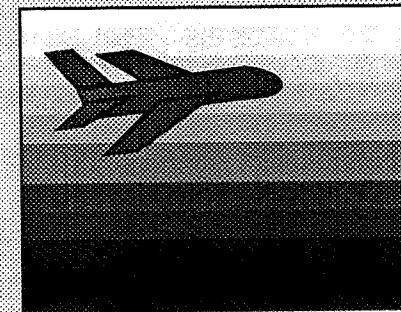
Barometric Altimeter

Indicates altitude based on standard atmosphere

Dependent on accurate altimeter setting

Error at surface: 50 ft (1 sigma)

Error at 40,000 ft: 200 ft (1 sigma)



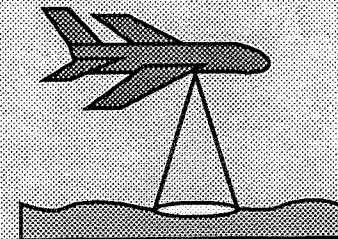
Radio Altimeter

Calibrated to read zero when wheels touch at nominal landing attitude (3 degrees)

Gives elevation above terrain (directly below aircraft)

Operate from 0 to 3000 ft above ground

Accuracy: 2 ft below 40 ft, 2.5 % of height above 40 ft (1 sigma)



Global Positioning System (GPS)

- Satellite based system, no ground-based aids required
- Give range/range-rate to each satellite received
- Positioning in two levels of accuracy:

Precision Code (P-code)

17.8 m 2d rms lateral

27.7 m 2 sigma vertical

**Available only to
military users**

Course Acquisition Code (C/A-code)

100 m 2d rms lateral

156 m 2 sigma vertical

**Available to
civil community**

- Limitations due to masking of satellites can be encountered
- Typical update rates of 1 hz

Differential GPS

Ground-based receiver at surveyed location calculates ranges to all satellites in view

Range corrections broadcast to users

Demonstrated Accuracies:

P-code:

.91 m rms horizontal

2.7 m rms vertical

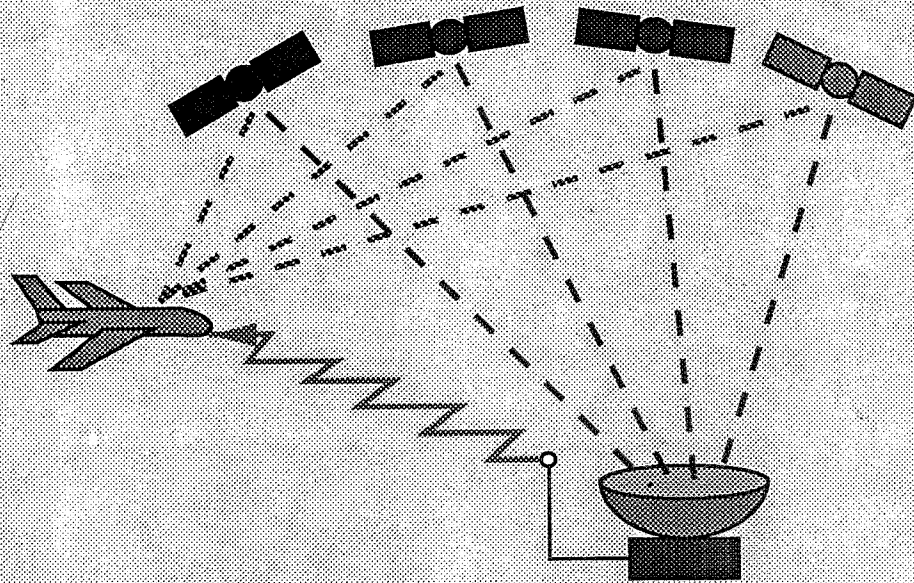
C/A-code:

7.6 m rms horizontal

8.5 m rms vertical

Pseudolite at differential station can improve vertical position

Carrier wave tracking shows promise for improving performance

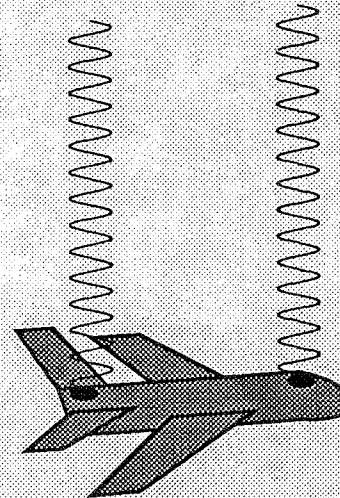


Other GPS Applications

Carrier Wave Attitude Determination

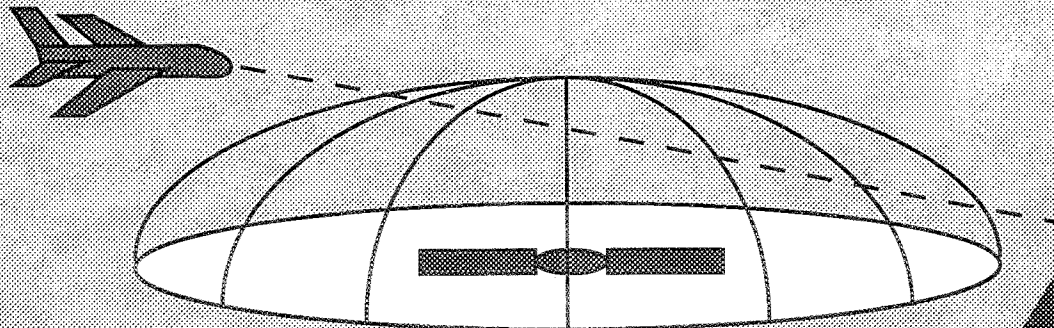
Multiple GPS antennae on aircraft allows measurement of phase differential

Accurate to .05 deg 1-sigma



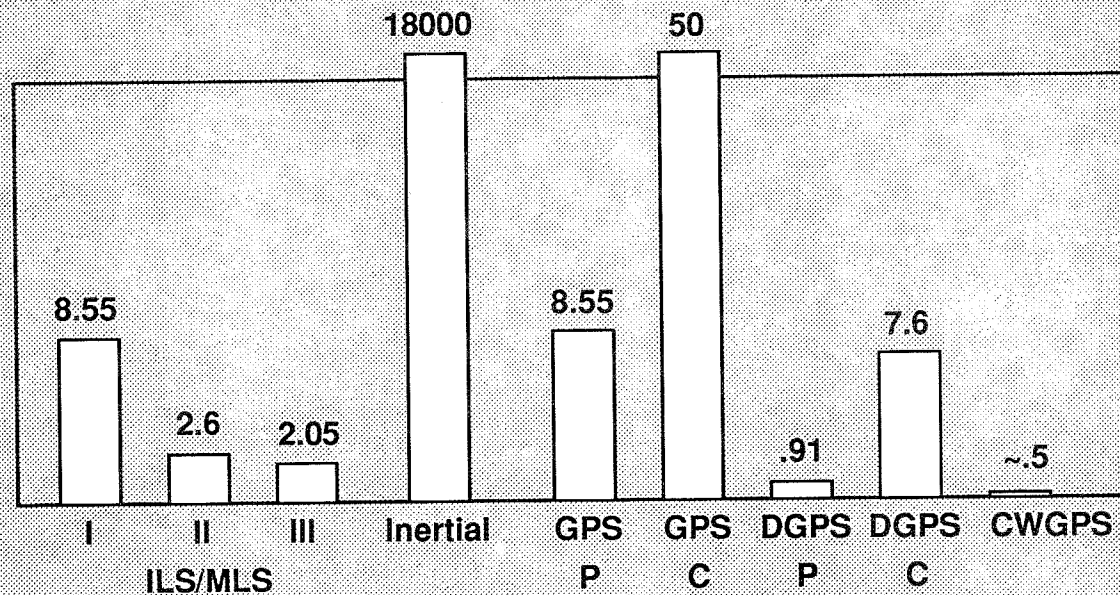
Carrier Wave/Pseudolite Navaid:

Demonstrated Accuracies in range of pseudolite of 5 cm

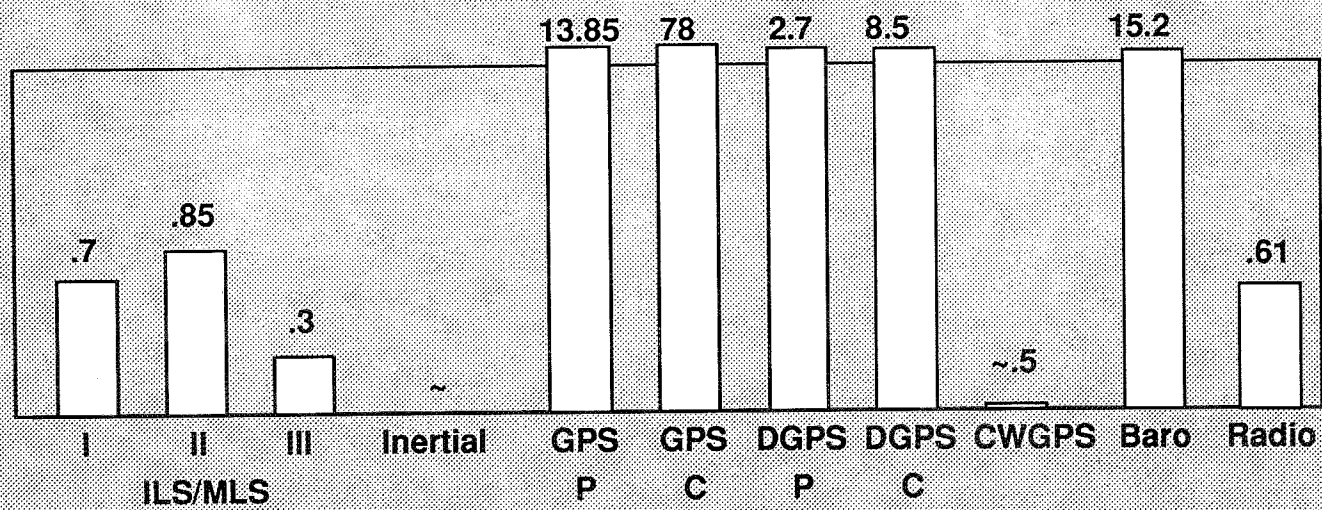


Summary of Navigational Accuracies

Lateral



Vertical



Applications to Enhanced Vision

- Radar Image Rectification
- Motion Compensation
- Image Registration

Radar Image Rectification

Issues

- Accurate altitude is key to producing rectified radar image
- Altitude is difficult to measure accurately
- Potential Energy vs Kinetic Energy

$mgh = \text{potential}$
 $mv^2 = \text{kinetic}$  50 ft potential energy is equivalent to 24 knots!

- Possible issue for certifying under current criteria

Motion Compensation

Issues

- Aircraft motion can cause blurring/distortion of radar image for slower scanning rates
- Improvement of image from motion compensation will be limited by accuracy of path measurement

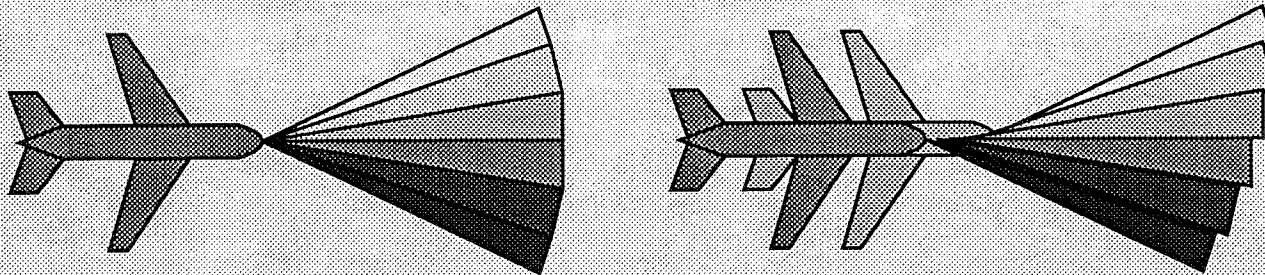


Image Registration

Issues

- Aircraft state information can affect registration times & registration accuracy
- In order to fuse images, registration is necessary
- Database image dependent upon position/attitude of aircraft
- Accuracy of position/attitude will affect feasibility of database fusion

Conclusions

- Accuracies of aircraft state measurements need to be accounted for in enhanced vision designs
- Techniques to extract state information from the image should be investigated

OMH

II. SENSOR MODELING

The IRGen Infrared Data Base Modeler

Uri Bernstein
Technology Service Corporation

ABSTRACT

IRGen is a modeling system which creates three-dimensional IR data bases for real-time simulation of thermal IR sensors. Starting from a visual data base, IRGen computes the temperature and radiance of every data base surface with a user-specified thermal environment. The predicted gray shade of each surface is then computed from the user-specified sensor characteristics. IRGen is based on first-principles models of heat transport and heat flux sources, and it accurately simulates the variations of IR imagery with time of day and with changing environmental conditions.

The starting point for creating an IRGen data base is a visual faceted data base, in which every facet has been labeled with a material code. This code is an index into a material data base which contains surface and bulk thermal properties for the material. IRGen uses the material properties to compute the surface temperature at the specified time of day. IRGen also supports image generator features such as texturing and smooth shading, which greatly enhance image realism.

Imaging IR Sensors

Imaging IR sensors (also called FLIR's), generate high-resolution video-rate images. The images displayed by an IR sensor are radiance maps of the scene viewed by the sensor. In the thermal (mid-IR and long-IR) bands, the radiance from a surface contains both emitted and reflected radiance. The emitted term depends on the surface temperature, and thus most IR images show a scene.

Since an imaging IR sensor displays the radiance from the scene, the appearance of a scene varies significantly with time of day, and with environmental conditions. Contrast reversals are frequently observed over the diurnal temperature cycle.

Atmospheric attenuation is a significant factor in the thermal IR bands. Attenuation varies dramatically with local meteorological factors such as humidity, fog, and rain.

Imaging IR Sensors

- **High-resolution image, updated at video rates.
The image is a radiance map of the scene.**
- **Total scene radiance includes both emitted and reflected radiance**
- **The appearance of the scene can vary significantly with time of day**
- **Atmospheric attenuation is significant and can vary dramatically with local meteorological conditions**

IR Simulation Methodologies

Several modeling methodologies have been used to generate data bases or images for IR sensor simulation. The simplest technique complements the intensity of the visible scene so that surface which is bright in the visible scene appears dark in the corresponding IR scene. In some cases, this technique has been elaborated by using a color table for visual-to-IR conversion. This technique is obviously limited (an asphalt road and a lake could be rendered with the same IR gray shade), and cannot handle diurnal variations.

At the other end of the IR simulation spectrum are models which have very elaborate models of heat transfer, and which may include time-dependent shadows, specialized natural feature models, and angle-dependent surface emissivity and reflectivity. This complexity may be necessary when an accurate signature is required for a particular object or natural feature. However, these models are very complex to set up, and require a long time to generate a single image.

IR Simulation Methodologies

- **Simplest - transform a visible image**
 - **complement visible color**
 - **color table**
- **Most complex**
 - **detailed thermal model with CFD for air flow, vegetative evaporation, etc.**
 - **time-dependent shadows**
 - **angle-dependent emissivity**

IRGen Principles

- **IR simulation intended for real-time simulation and training applications. Compatible with standard modeling and simulation software.**
- **Three-dimensional faceted data bases, including moving targets, structures, and terrain**
- **First-principles models of heat transfer, radiation, and atmospheric propagation**
- **Easy to use; user can control all the parameters of the materials, environment, and sensor**

IRGen Principles (continued)

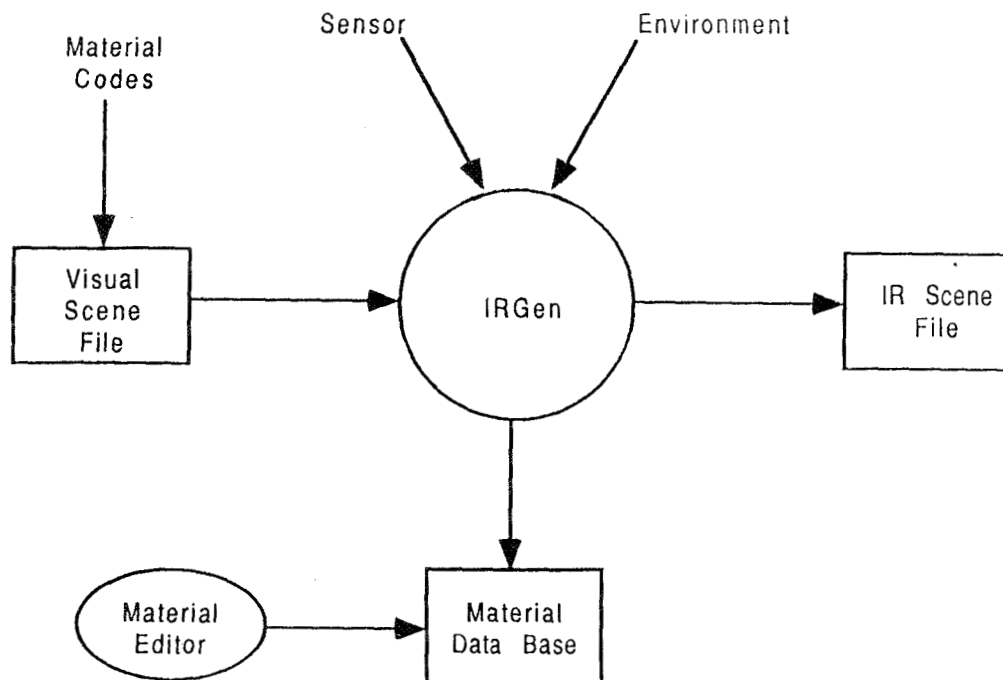
- **Built-in support for real-time graphics features**
 - **texturing**
 - **smooth shading**
- **Flexible and extendable**
 - **New materials can be created by the user**
 - **User-defined thermal models can be called by IRGen**
- **Reasonable setup and run time**

IRGen Data Diagram

This diagram shows the inputs and outputs of the IRGen program. The main input is a visual data base whose surfaces have been given material codes. Other inputs include the environment, atmospheric and sensor parameters.

The main outputs of IRGen is an IR data base whose geometry is identical to the visual data base geometry, but which has IR gray shades instead of the visual color. Other outputs include auxiliary graphics information such as texture maps and atmospheric attenuation information. The surface radiance and temperature values are accessible within the data base and are also recorded in a separate data file.

IRGen DATA DIAGRAM



IRGen INPUTS

- **3-D visual data base from a data base creation program**
- **Each surface facet labeled with material code**
- **Environment**
 - **Thermal**
 - **Atmospheric**
- **Sensor characteristics**

IRGen OUTPUTS

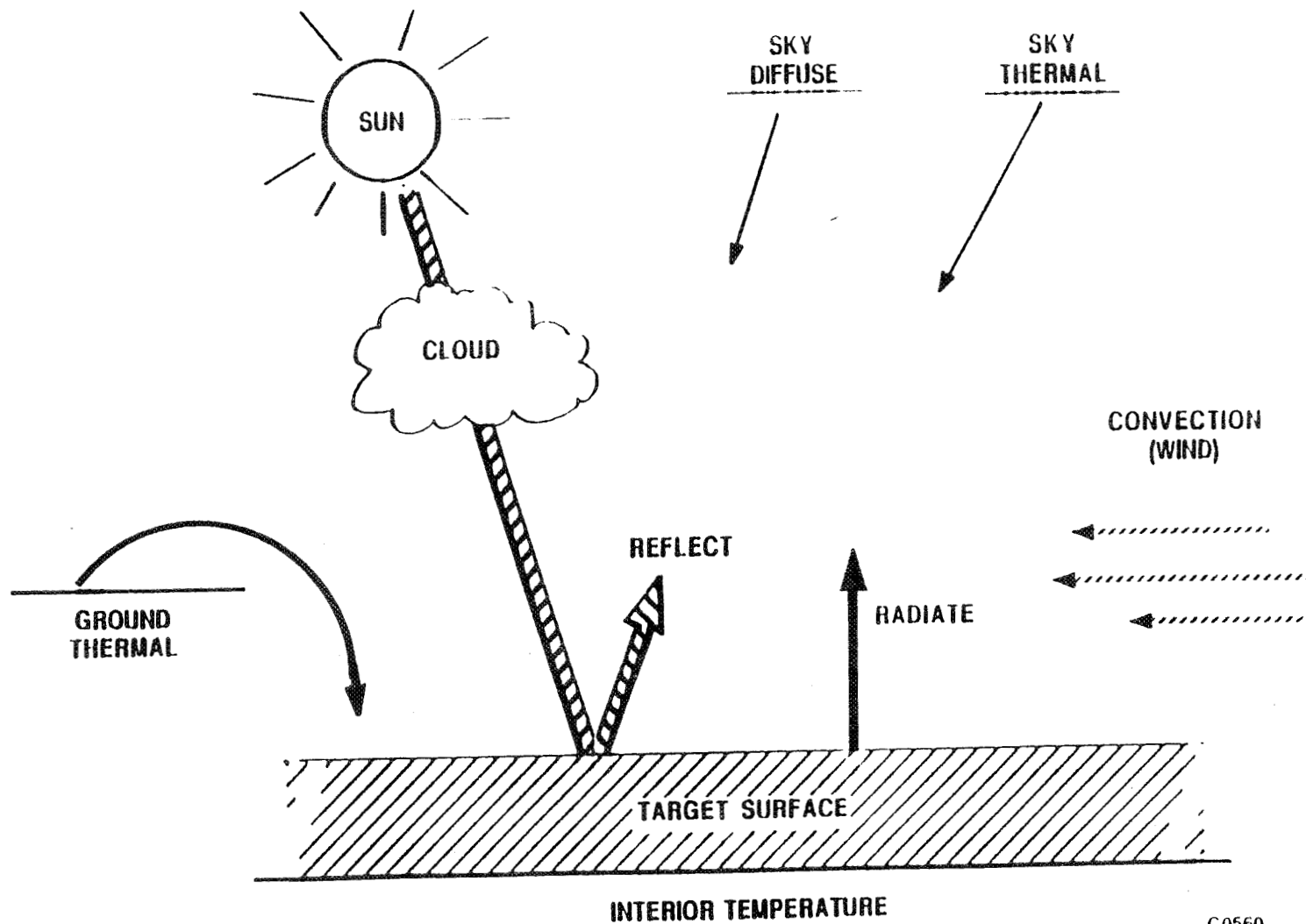
- **3-D IR data base with surface colors replaced by IR gray shades; radiance values stored as surface data**
- **Other real-time graphics data (attenuation, texture, etc.)**
- **Radiance data file**

IRGen Thermal Model

This diagram shows the main sources of heat flux for the IRGen thermal model. Heat flow normal to the surface is simulated by integrating the one-dimensional heat transport equation, using a finite-difference method. External sources of heat flux include direct and diffuse solar radiation, sky and ground thermal radiation, and convection. Internal sources of heat flux include interior convection and conduction.

The surface radiance include both surface thermal emission, and reflected sky and ground radiance.

TSC THERMAL MODEL



C0560

IRGen MODULES

- **Material Data Base**
 - **Surface and bulk thermal properties of scenario materials**
- **Thermal Model**
 - **Computation of surface temperature and radiance for every surface in the scenario**
 - **Integration of heat transport equation**
 - **External and internal heat flux sources**

IRGen MODULES (continued)

- **Environment Model**
 - **Solar, convective, sky, and ground heat sources**
- **Sensor Model**
 - **Sensor response function**
- **Atmospheric Model**
 - **LOWTRAN 7 (U.S. DoD standard model) integrated with graphics generator fog/haze function**

IRGen Operating Environment

IRGen currently generates data bases for both Silicon Graphics and Star Graphicon image generators. The latest version will run on any Silicon Graphics workstation.

Since IRGen requires a geometric data base, it must be used in conjunction with a geometric modeling program. The preferred modeling programs are MultiGen® and ModelGen™ (from Software Systems, San Jose, CA) which support the full set of image generator features such as level-of-detail, texture, and smooth shading. These modeling programs allow the user to enter a material code for each surface in a special data field that is reserved for IRGen.

An alternative version of IRGen runs with the AutoCAD® modeling program (Autodesk, Sausalito, CA).

IRGen OPERATING ENVIRONMENT

- **Hardware Platforms**
 - **Silicon Graphics (SGI) Workstation**
 - **Star Graphicon 2000**
- **Modeling Interface**
 - **MultiGen (SGI) modeling system - standard modeling system for real-time visual simulation. Supports level-of-detail, hierarchical data bases, texture.**
 - **AutoCAD (PC)**

IRGen Options

IRGen has several options for special applications. The Defense Mapping Agency (DMA) data option allows the use of the material codes provided by DMA digital feature analysis data (DFAD). With this option, the user does not have to enter any material codes. Note that DMA digital terrain elevation data (DTED) can be polygonized by MultiGen DTED option, and passed through IRGen into the IR data base.

The texture option allows the creation of IR textured data bases with thermally accurate texture maps. Textures are particularly important for realistic low-altitude flight simulation over terrain and water surfaces. The textures can come from three sources: (1) the visual data base, (2) a scanned IR image, or (3) statistical texture creation program.

The special effects option creates translucent and smooth-shaded surfaces.

IRGen OPTIONS

- **U.S. Defense Mapping Agency (DMA) Data Interface**
- **Digital Terrain Elevation Data (DTED) - can be polygonized by the MultiGen DTED Option**
- **Digital Feature Analysis Data (DFAD) - automatically convert DFAD material codes and feature IDs to IRGen material codes.**

IRGen OPTIONS (continued)

- **Texture**
 - **Generate thermally accurate textured surfaces. Important for terrain, sea, and cloud backgrounds**
 - **Scanned images or synthetic textures**
- **Special Effects**
 - **Translucent surfaces (exhaust plumes, obscurants)**
 - **Vertex shading (temperature gradients)**

IRGen Material Data Base Parameters

Properties of IRGen materials are stored in the material data base, which is accessed by the material code. The user can modify material properties or add new materials. Material parameters 1 and 2 serve to identify the material. Parameters 3 through 17 are used for the temperature and radiance computations. ("Number of nodes" refers to the finite-difference method.) Parameters 18 through 20 are used to implement intersurface thermal coupling when computing smooth shading, and parameter 21 identifies the texture map for textured surfaces.

IRGen MATERIAL DATA BASE PARAMETERS

Identification:

1. material code
2. label

Thermal model parameters:

3. 3-5 micron emissivity
4. 8-12 micron emissivity
5. solar reflectivity
6. integration time increment
7. integration settling time
8. interior temperature
9. interior conductive/convective flag
10. interior thermal coupling
11. two-sided surface flag
12. shadow surface
13. number of nodes
14. node heat capacity array
15. node conductive transport array
16. node radiative transport array
17. node conductive coefficient array
18. node radiative coefficient array

Intersurface thermal coupling:

19. read/write flag for vertex thermal coupling
20. vertex coupling file number
21. vertex coupling flags

Textured materials:

22. name of thermal texture file

The Radar Image Generation (RIG) Model

**Anthony J. Stenger
Technology Service Corporation**

205737

P 11

249919

ABSTRACT

RIG is a modeling system which creates synthetic aperture radar (SAR) and inverse SAR images from 3-D faceted data bases. RIG is based on a physical optics model and includes the effects of multiple reflections. Both conducting and dielectric surfaces can be modeled; each surface is labeled with a material code which is an index into a data base of electromagnetic properties. The inputs to the program include the radar processing parameters, the target orientation, the sensor velocity, and (for inverse SAR) the target angle rates.

The current version of RIG can be run on any workstation, however, it is not a real-time model. We are considering several approaches to enable the program to generate real-time radar imagery.

In addition to its image generation function, RIG can also generate radar cross-section (RCS) plots as well as range and doppler radar return profiles.

RADAR IMAGERY GENERATOR (RIG)

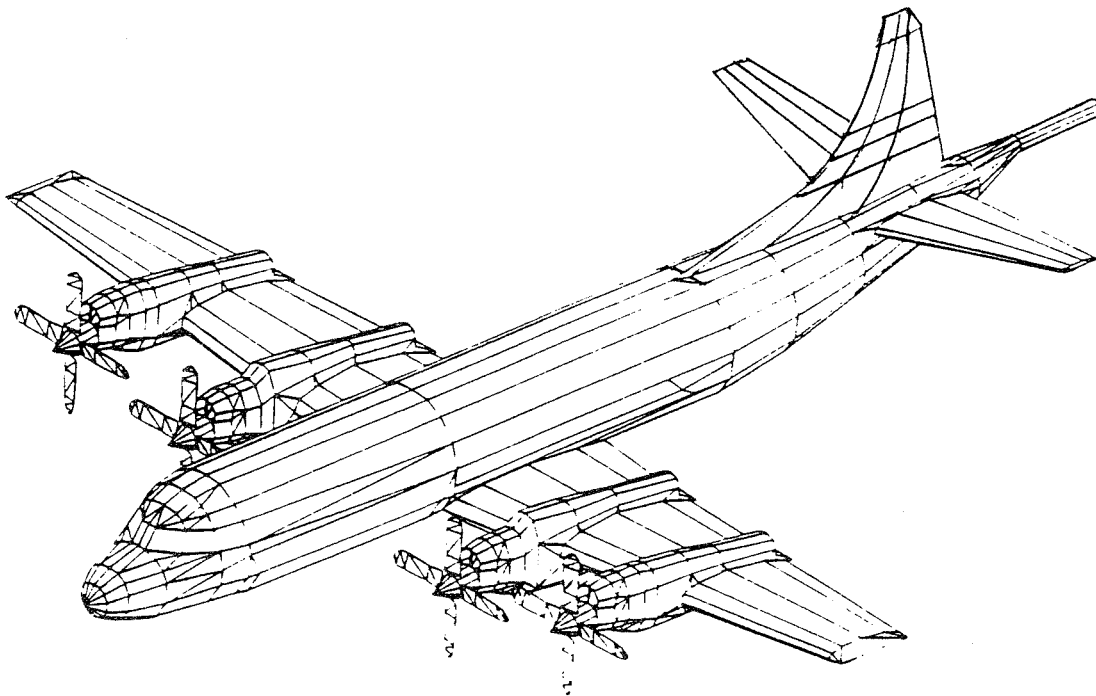
The Radar Imagery Generator (RIG) simulates the image from a synthetic aperture radar (SAR) or an inverse SAR (ISAR). The target model for RIG is a 3-D geometric data base. RIG uses a physical optics model to calculate the radar return from conductive and dielectric surfaces. RIG uses a ray tracing method to calculate the coherent path to each surface. Multiple bounces from non-contiguous objects as well as dihedral and monostatic returns are modeled.

The user can define the radar parameters, e.g. wavelength, polarization, range resolution and doppler bandwidth. The target is defined by its orientation and speed, or in more detail, by its complete motion cycle in roll, pitch and yaw.

	RIG RADAR IMAGERY GENERATOR
--	---

- **INTEGRATED TOOL FOR CREATING SYNTHETIC APERTURE RADAR (SAR) AND INVERSE SAR (ISAR) IMAGERY**
- **PHYSICAL OPTICS MODELING OF CONDUCTIVE AND COATED MATERIALS**
- **MONOSTATIC AND DIHEDRAL BOUNCE MODELING**
- **3-D FACETED DATA BASES OF AIRBORNE, LAND, AND SEA BASED TARGETS**
- **CONTROL OF RADAR PLATFORM AND TARGET POSITIONS**
- **USER-DEFINED DIFFUSE GROUND TOPOGRAPHY**

P-3C "ORION" TARGET MODEL



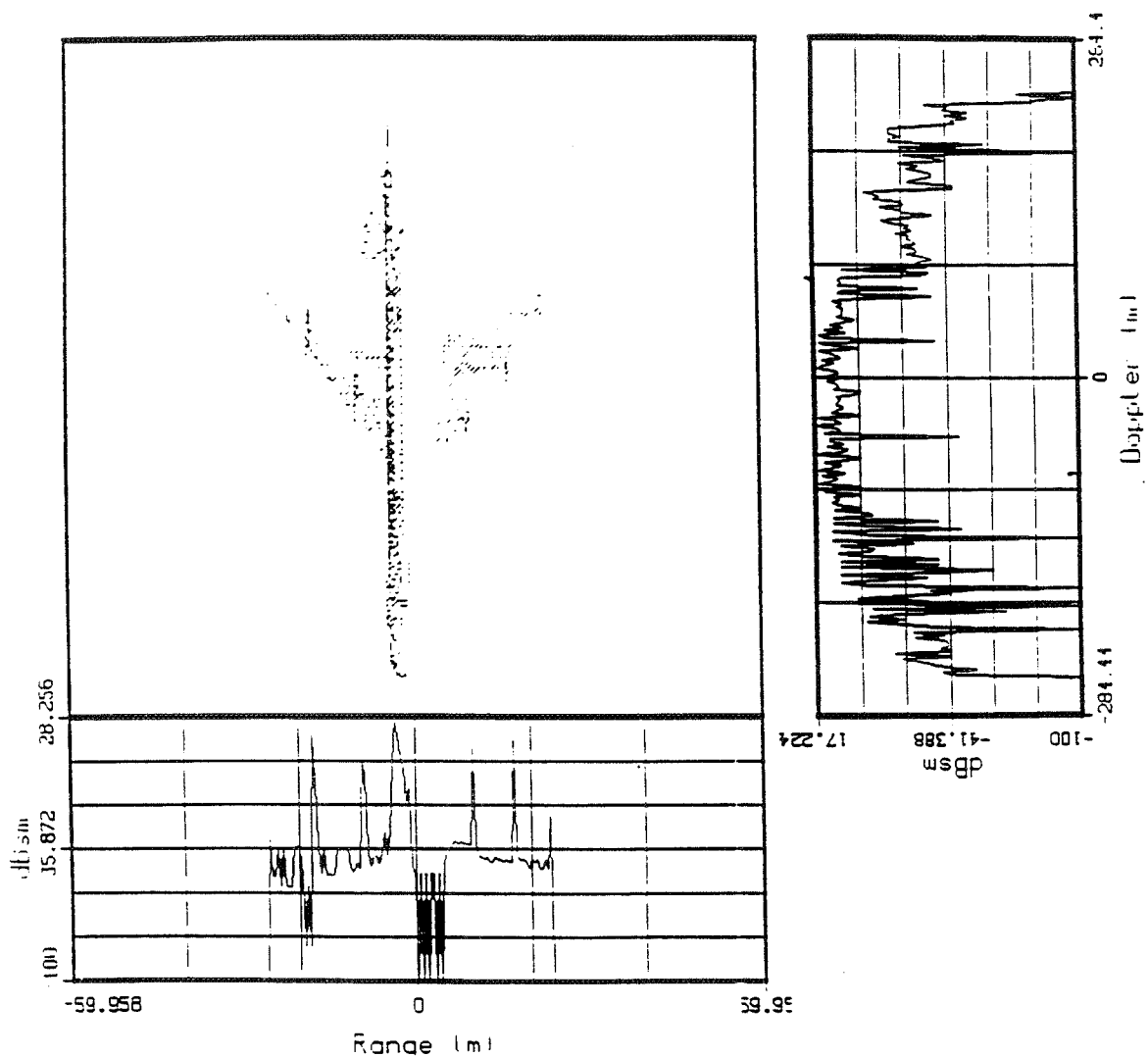
- 1344 POLYGONS
- NO LANDING GEAR
- ALL FLAPS, SLATS IN CLEAN POSITIONS
- PROPELLERS ORTHOGONALLY ORIENTED

SIMULATED IMAGERY

The returns from several surfaces that appear in a given range/doppler cell are coherently integrated to generate the SAR or ISAR image. The RCS profile as a function of range (doppler) is generated by summing in the doppler (range) dimension.

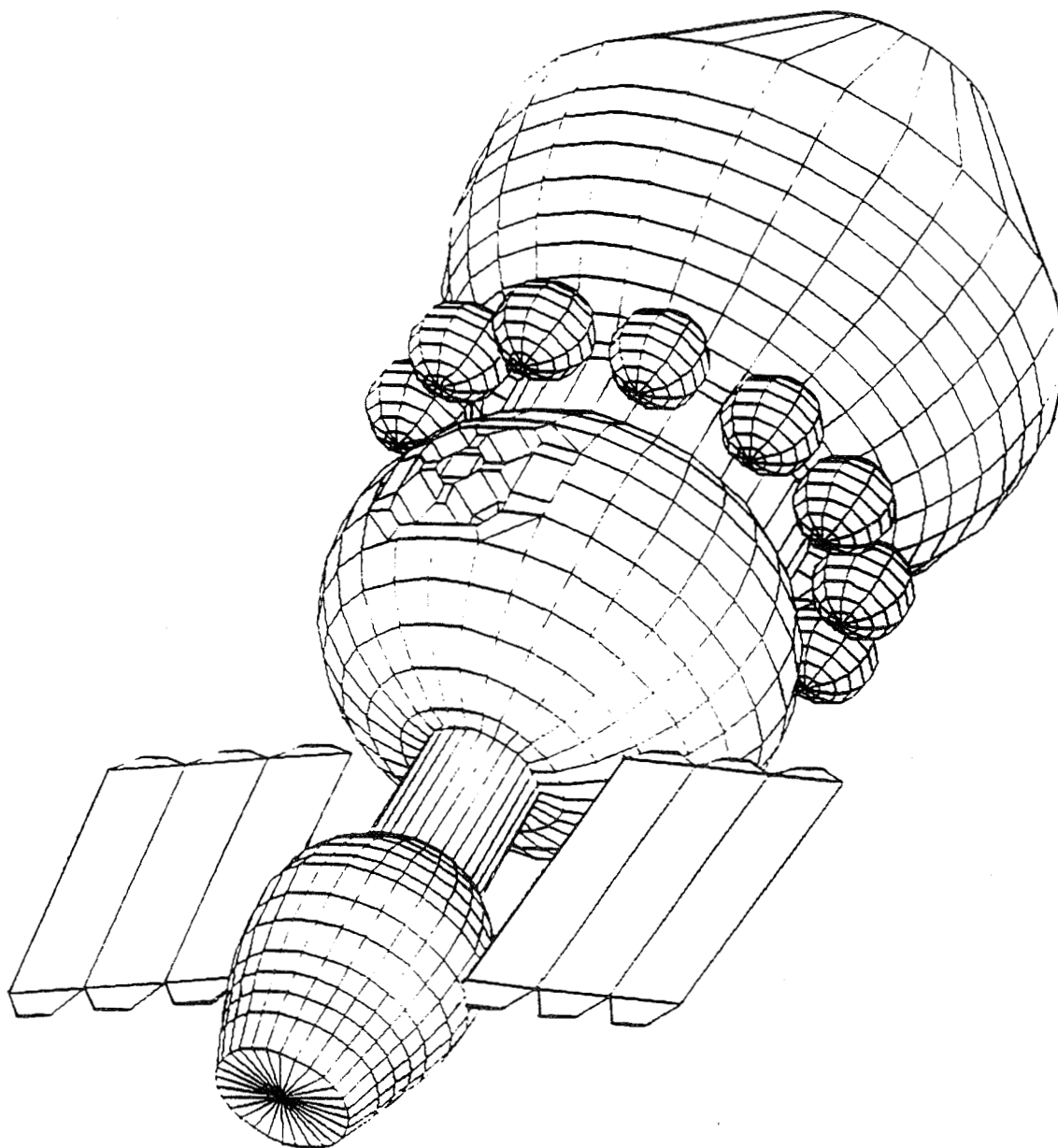
The final step of RIG is to convolve the radar response function (that models the antenna, range and doppler response characteristics) with the ideal RCS image. The images and profiles provided in the Figure are the ideal RCS and do not show the results of the convolution.

IMAGERY WITH RCS SIMULATION



- RCS PROFILES ARE COHERENTLY SUMMED WITHIN EACH RANGE BIN AND EACH DOPPLER FILTER.

RECONNAISSANCE SATELLITE

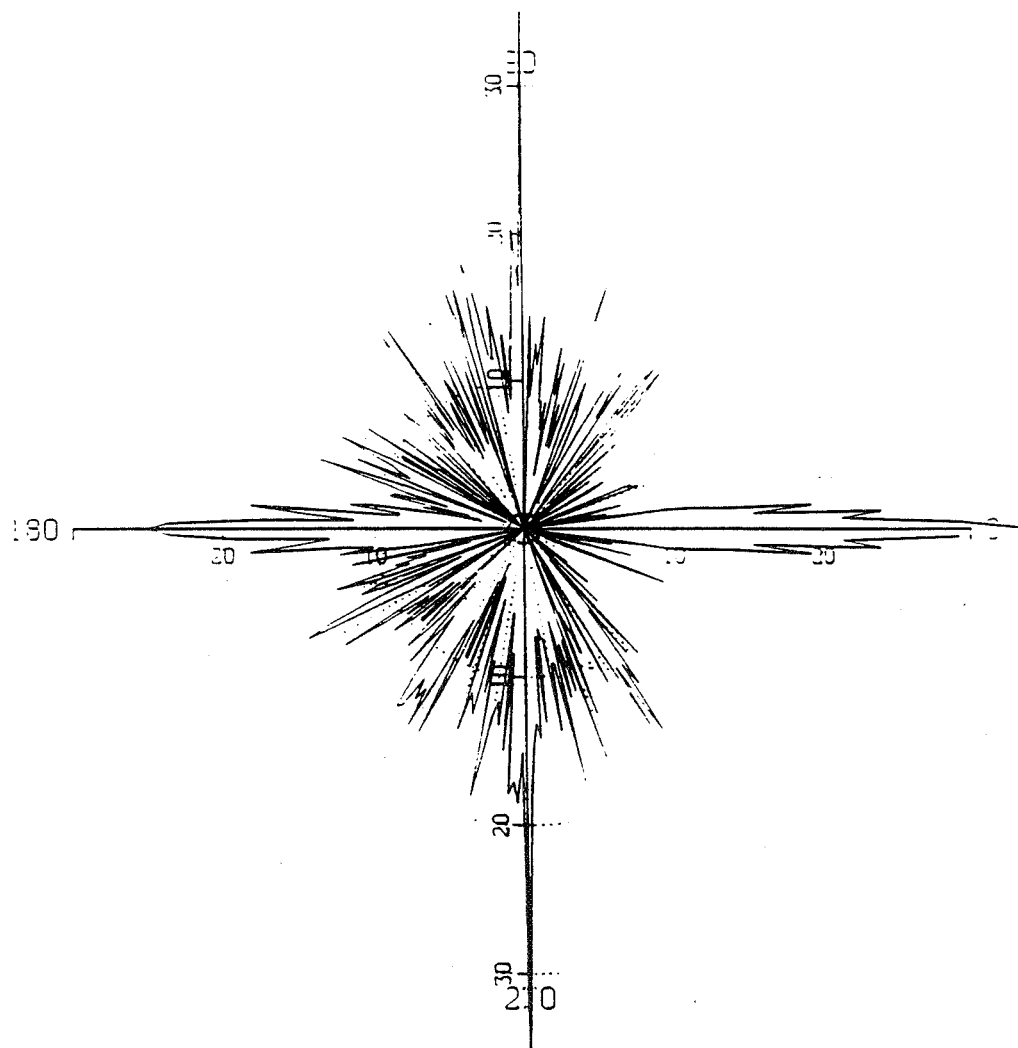


TOTAL RCS

RIG also generates the total RCS of the target by coherently summing over all range and doppler cells. The RCS of a satellite is given in the Figure as a function of aspect. Angle is defined in a plane perpendicular to the solar panels, with 0° looking toward the panels. The RCS without convolution with the radar response is provided.

TOTAL RCS OF SPACECRAFT

RCS(dBsm) vs Aspect Angle (deg)



MULTISPECTRAL SIMULATION

RIG is the radar equivalent of IRGen that is described in a companion paper. Together both programs can generate multispectral imagery from the same geometric data base. The combined system would simulate the visible, infrared and radar image of the same scene for the same viewing and atmospheric conditions.

MULTISPECTRAL SIMULATION

GEOMETRIC
3D MODELER

Create Imagery for Sensor
Design and Evaluation

DATA BASE

SOLID RENDERING
SOFTWARE

RIG

IRGen

VISUAL IMAGE

Real Beam, SAR
or ISAR

IR
Imagery

Sun Position
Other Light Sources
Color Rendering

Propagation Effects
Rain
Terrain
Obstacle

3-5 um
8-12 um
Solar Heating
Internal Sources (engines)
Aerodynamic Heating
Plumes

Advanced Radiometric & Interferometric Millimeter-Wave Scene Simulations

B. I. Hauss, P. J. Moffa, W. G. Steele, H. Agravante, R. Davidheiser, T. Samec
and S. K. Young
TRW Space and Electronics Group

1.0 Introduction:

Smart munitions and weapons utilize various imaging sensors (including passive IR, active and passive millimeter-wave, and visible wavebands) to detect/identify targets at short standoff ranges and in varied terrain backgrounds. In order to design and evaluate these sensors under a variety of conditions, a high-fidelity scene simulation capability is necessary. Such a capability for passive millimeter-wave scene simulation exists at TRW. TRW's Advanced Radiometric Millimeter-Wave Scene Simulation (ARMSS) code is a rigorous, benchmarked, end-to-end passive millimeter-wave scene simulation code for interpreting millimeter-wave data, establishing scene signatures and evaluating sensor performance.

In passive millimeter-wave imaging, resolution is limited due to wavelength and aperture size. Where high resolution is required, the utility of passive millimeter-wave imaging is confined to short ranges. Recent developments in interferometry have made possible high resolution applications on military platforms. Interferometry or synthetic aperture radiometry allows the creation of a high resolution image with a sparsely filled aperture. Borrowing from research work in radio astronomy, we have developed and tested at TRW scene reconstruction algorithms that allow the recovery of the scene from a relatively small number of spatial frequency components.

In this paper, the TRW modeling capability is described and numerical results are presented.

2.0 The ARMSS Code:

The radiometric signature of a man-made, highly reflecting target depends sensitively on the target geometry and the background (sky and/or terrain) brightness temperatures which happen to lie along the specular reflection path. It is thus critical to describe these elements accurately. To model the interaction between the target, the sky/terrain background and the radiometer, TRW has developed ARMSS, a rigorous, benchmarked, end-to-end passive millimeter-wave scene simulation code. Many of the physics models employed are "first principles"-models, requiring only measurable physical conditions to accurately predict millimeter-wave scene signatures. In addition, our models offer a true

3-D scene simulation capability, allowing the complex interactions between the various elements of the scene to be correctly described. This is required at millimeter-wave frequencies both because the downwelling atmospheric radiation varies dramatically with zenith angle and because the emissivity/ reflectivity of most terrain materials has a significant dependence on incidence angle. This is especially true near grazing incidence, where scattering and emission are further complicated on rough surfaces by multiple scattering and shadowing effects.

The four major components of the ARMSS code are shown in Figure 2.1. The first and primary component of this end-to-end code is a rigorous description of the passive mm wave phenomenology. This encompasses state-of-the-art physics models describing: emission from the scene constituents, scattering of the downwelling sky radiation by the scene, propagation/attenuation of the electromagnetic energy from the scene to the sensor, and upwelling atmospheric radiation between the scene and the sensor. More specifically, the phenomenology model includes sub-models for atmospheric propagation effects and meteorology, surface/terrain physics describing the mix of emission and scattering from scene constituents, ray-tracing algorithms for efficient but accurate solution of the radiative transfer equation, and the use of combinatorial geometry for constructing complex three-dimensional scenes, Figure 2.2. Each aspect of the phenomenology model has been individually benchmarked against both measured data and other models in the literature. In addition, the phenomenology model as a whole has been benchmarked against the field-imaging data which we have collected.

The second component of the end-to-end simulation code, the sensor model, takes output from the phenomenology model (i.e., the very high resolution, radiometric image in front of the sensor) and constructs the actual image as seen by the sensor, based on diffraction optics and including such effects as lens aberrations, finite detector size, and noise. This allows us to assess sensor performance and perform design tradeoffs. Again, all aspects of the sensor model have been benchmarked.

Next, to evaluate the ability of real-time image enhancement and restoration techniques to improve image quality, thereby allowing tradeoffs to be made with the sensor design requirements, an image processing capability has been included in the end-to-end code. This takes as input raw data from the sensor and applies noise filtering, upsampling, temperature bandpass filtering, global and hybrid histogram equalization, and edge-operator sharpening techniques to enhance the resulting image and thereby allow some relaxation of the sensor design requirements.

The display model, the final component of the end-to-end code, captures the enhanced images, frame-by-frame, on video tape for replay at the frame-rate for which the images were produced. This allows us to perform those sensor design tradeoffs which involve frame-rate, where higher frame rates normally result in a poorer signal-to-noise ratio.

Because of their importance to the accurate generation of passive millimeter-wave scenes, a more detailed description of the models describing atmospheric propagation and the

calculation of the sky radiometric temperature profile, terrain emissivity/ scattering, and the construction of the background-target scene geometry will be given in sub-sections 2.1-2.3 below.

2.1 Atmospheric Propagation and Sky Radiometric Temperature Calculations

The sky radiometric temperature profile (a function of zenith angle) is calculated within the ARMSS code based on computations of the downwelling atmospheric radiation. These calculations begin with a determination of the specific attenuation rates in the atmosphere. To this end, the propagation effects model developed by the Institute for Telecommunication Sciences (Reference 1) has been implemented in the code. The model calculates the specific attenuation rates as a function of measurable meteorological parameters (pressure, thermometric temperature, relative humidity, hydrosol concentration and rain rate) and has a range of validity from 0 to 1000 GHz. The model includes pressure broadened resonance lines for water and oxygen, continuum absorption due to non-resonant oxygen, pressure induced nitrogen absorption, Rayleigh absorption for haze, fog and clouds, and a parameterized power-law rain attenuation model to simulate Mie scattering and absorption by a distribution of droplet sizes corresponding to a measured rain rate. The model accurately compares with published and measured data for clear-air, fog, and rain attenuation, Figure 2.1.1.

To provide meteorological properties as a function of altitude for diverse geographic and seasonal changes in atmospheric conditions, the ARMSS code makes use of any of ten synthetic atmospheric databases compiled by the Air Force Geophysics Laboratory. This allows the code to accommodate a diverse range of climatological and weather conditions, ranging from subtropical to arctic and in various seasons. In addition, plane-stratified (i.e., layer) models for clouds, fog, haze and rain are included in the code to allow study of their effects, both individually and collectively.

The sky radiometric temperature profile is calculated by a detailed evaluation of the radiative transfer equation for the downwelling atmospheric radiation, taken from 30 km above sea-level. The highly efficient ray tracing solution permits some 60,000 rays to be processed in only 7 minutes on a Silicon Graphics Personal Iris. Benchmarks with the literature and field measurements, using the 1976 U.S. Standard Atmospheric data base to provide meteorological properties, have been performed, Figure 2.1.2.

The models described above are also used in computing both the upwelling atmospheric radiation and the attenuation of the scattered and/or emitted radiation between elements of the scene and the sensor. A benchmark of these calculations, including the contributions due to terrain emission and scattering is discussed in the following sections.

2.2 Terrain Emissivity/Scattering Calculation

Terrain emissivities/reflectivities are calculated within the ARMSS code based on the dielectric properties of the terrain layer(s) and their surface/subsurface geometry. For a

single smooth (i.e., specular) layer, emissivity/reflectivity is determined from a straightforward calculation of the Fresnel reflection coefficient, which depends only on the angle of incidence and the complex dielectric constant of the terrain material.

The emissivity/reflectivity for multiple smooth dielectric layers is obtained from a calculation of either the coherent or incoherent multiple layer effective reflectivity, depending on whether phase coherence is maintained within the layers (i.e., whether volume scattering within the layers is significant). The coherent reflectivity is calculated by rigorously solving for the electromagnetic fields in each dielectric layer and then employing a matrix technique to combine their individual effects, always requiring phase accountability, to give the effective field reflection coefficient at the terrain surface. Squaring the magnitude of this quantity then gives the coherent power reflection coefficient. For the calculation of the incoherent reflectivity, reflections from each layer are treated as an incoherent process, avoiding phase effects by basing all calculations on the power (i.e., Fresnel) reflection coefficient for each layer. This calculation is carried to infinite order in the number of reflections at the layer boundaries. For the three-layer problem, this results in a closed-form expression for the effective surface power reflection coefficient. Finally, assuming that the thermometric temperature is the same for all the terrain layers, the emissivity for either the coherent or incoherent process is the difference between unity and the calculated reflectivity.

For the rough surface emissivity, we employ either the semi-empirical model of Choudhury and Wang (Reference 2), with roughness parameters chosen to give the best fit to measured data, or Wagner-Lynch (Reference 3) scattering theory for an anisotropic, random rough surface characterized by Gaussian statistics. This latter approach is based on a geometrical-optics theory of emission and scattering. A complete ray treatment is provided in the sense that single-scatter and bistatic shadowing effects are included in a consistent manner for a general two-dimensional rough surface. To conserve energy to a relatively high degree of approximation for all observation angles, a double-scatter approximation is usually required. However, the single-scatter approximation employed in the code provides predicted radiometric temperatures within a few Kelvin of the true temperatures over most observation angles, Figure 2.2.1.

A data-base of models describing the dielectric properties of naturally occurring and man-made terrain materials (water-fresh and sea, ice-fresh and sea, snow, various types of soils, asphalt, concrete, etc.) has been developed for use in calculating terrain emissivities. For the majority of materials, these models are given as a function of frequency, physical temperature, density, and water content. The bulk dielectric mixing models for some materials are setup using a specified material makeup (e.g., the various soil categories use specified bulk densities and percentages of sand, silt, and clay) as a user convenience. This convention is easily modified to allow any appropriate combination of parameters as determined by measurement of the local properties. These models have been successfully compared to published data, Figure 2.2.2.

2.3 Three-Dimensional Background-Target Scene Generation

Atmospheric propagation and terrain surface interaction models are joined through the use of a true 3-D ray tracing solution of the radiative transfer equation. This model determines ray paths through the atmosphere and ray intercepts with scene objects. The model first employs a backward tracing of the ray paths, from the sensor, through multiple reflections off scene objects and upward through the atmosphere. A forward integration of the radiative transfer equation along the calculated ray path then gives the radiometric temperature at a single point in the infinite resolution image at the pupil plane in front of the sensor. Figure 2.3.1 shows four snapshot simulations of an aircraft landing on a concrete runway surrounded by dirt. The weather conditions are heavy fog with wet ground surfaces. A plane is parked on an adjacent taxi-way, with its reflected image on the nearby terrain surface. The important point to note is that this is a complex scene viewed at near grazing incidence on both specular and rough terrain surfaces which is realistically modeled.

The fidelity of the combined models for atmospheric propagation, terrain emission and scattering, and the numerical solution of the radiative transfer equation has been extensively benchmarked by comparisons with field measurements, Figures 2.3.2. These results indicate that the models are not only qualitatively correct, but also quantitatively accurate.

To achieve an efficient and highly accurate 3-D scene description, the ARMSS code employs combinatorial geometry (also known as constructive solid geometry) to model both elements of the terrain and high-value targets in the scene. The mathematical description of each object in the scene is achieved through the orderly combination of any of eight basic solid geometric primitives; rectangular parallelepiped, box, sphere, right circular cylinder, right elliptical cylinder, truncated right angle cone, ellipsoid of revolution, and right angle wedge. A scene object's location and shape is described by selecting the appropriate geometric primitives and specifying their location, dimensions, and how to combine them (given in terms of the unions, intersections, and exclusions, of their individual volumes), Figure 2.3.3. As can be seen from the constructed models for the BMP-1 troop transport, the T-72 tank and the SS-24 missile and mobile launcher (Figure 2.3.4), this approach affords an accurate representation of scene objects, with true surface curvatures which would be extremely difficult to achieve from a faceted geometry model. The requirement to accurately predict the millimeter-wave scene obviously dictates the need for this accurate treatment of the scene geometry.

In addition to determining the path length from the ray's current position to its next intersection with a scene surface, the geometry package also identifies the code surface element intersected, the angle of the incident ray to the surface, and the normal to the surface at the point of intersection. This information is necessary in modelling the contributions to the radiometric temperature from the terrain surface. In particular, the identification of the code surface element intersected provides the terrain/surface physics models with the particular surface and subsurface properties (specified as input for each surface element) at the point of intersection. These properties include the number of

dielectric layers for the surface element, specification of either coherent or incoherent scattering/emission (for code surface elements having multiple dielectric layers), layer material type, layer water content, layer density, surface thermometric temperature, and parameters specifying the surface rms roughness slope.

2.4 Real-Time Passive Millimeter Wave Scene Simulation:

As part of a joint program with NASA LaRC, TRW has been developing a real-time, passive millimeter wave scene simulation capability. The general approach taken to achieve real-time operation has been to identify the necessary passive millimeter wave phenomenology models from TRW's ARMSS code and implement these in an approximate fashion into NASA's visible flight simulator. The primary requirement on this process was that it maintains reasonable scene fidelity without sacrificing real-time performance. The approximations made are summarized in Table 2.4.1 and described briefly below.

First, the Constructive Solid Geometry (CSG) description of the terrain scene was replaced with a polygonal tessellation. This allowed us to replace the high ray sampling of the CSG scene with a much reduced (by a factor of 1000 or more) ray tracing only to the verticies of the polygonal scene elements. Polygon shading between the verticies is performed by simple shading models implemented in the Silicon Graphics firmware. This introduces a small interpolation error in the scene radiances between polygon verticies; however, the magnitude of this interpolation error is easily controlled by reducing the size of the scene polygons. A second problem introduced by the polygonal scene element approach is the difficulty in simulating multiple reflections and shadowing effects, although a method has been devised for implementing these as well.

The second group of approximations which were required to achieve real-time passive millimeter wave scenes were the use of lookup tables. The real-time code employs lookup tables for the sky temperature profile, the emissivity/reflectivity of specular-surface scene elements versus

incidence angle, and the apparent temperature of rough-surface terrain elements as a function of the angle of observation and assuming a horizontal mean ground-plane. These tables are computed at the beginning of the simulation based on the input atmospheric and terrain conditions. This use of lookup tables eliminates the need for repetitive calculations of the downwelling atmospheric radiation and the emitted and scattered radiation from the scene elements for each ray. There is a small price incurred in terms of interpolation error, but as will be illustrated in the following talk from NASA LaRC, these errors are negligible.

A significant improvement in performance, which allowed real-time operation, resulted from the approximation for the upwelling atmospheric radiation from a scene element to the sensor. Since the sensor is continuously moving and viewing different elements of the terrain, this calculation could not be handled using a lookup table. The approximation employed makes use of the fact that the temperature lapse rate in the troposphere is small, only 6.5K/km. This means that over a plane stratified layer of perhaps a few tenths of kilometers in height, the thermometric temperature is essentially constant. Considering that most of the landing simulations will involve sensors within 0.2km of the ground, the integral of the path radiance from the scene element to the sensor,

$$\int_0^L \alpha(s') T(s') \exp\left[-\int_{s'}^L \alpha(s'') ds''\right],$$

can be reduced to a simple algebraic form

$$T_m \left\{ 1 - \exp[-\tau(0, L)] \right\},$$

where T_m is the effective or mean thermometric temperature along the path and

$$\tau(0, L) \equiv \int_0^L \alpha(s'') ds'' = \sec\theta \tau(0, Z)$$

is the cumulative optical thickness. A lookup table of $\tau(0, z)$ is

computed at the beginning of the simulation, and used to further speedup the calculation. As can be seen from Figure 2.4.1, the difference between a brute-force numerical integration of the path radiance and the above constant temperature approximation is negligible; however, the approximate solution is easily two-orders of magnitude faster.

The final approximation employed in the real-time model is the restriction to a single specular reflection from an element of the scene. The model assumes that any reflection off a scene-element which results in the ray going back towards the terrain will be reflected from the terrain as if from a perfectly conducting horizontal ground plane. This approximation was implemented as a temporary measure until there was sufficient resources to implement a multiple reflection model. A method for implementing multiple reflections and shadowing in real-time using the polygonal model described earlier has been devised, but not yet implemented. The current approach does not correctly treat the interaction between elements of the 3-D scene.

We have benchmarked the real-time passive millimeter wave scene simulation against TRW's ARMSS code, and have found it to be accurate to within a few Kelvin throughout the entire scene. The details of this comparison and a live demonstration of the real-time passive millimeter wave flight simulator will be presented in the following talk by NASA LaRC. The principal planned upgrade to the real-time simulator is the implementation of models for multiple reflection and shadowing, allowing the correct treatment of the interaction of the 3D scene elements.

3.0 Interferometric Modeling:

Interferometry is a technique for trying to achieve the resolution of a large aperture by only sparsely covering the equivalent area with much smaller apertures. The Van Cittert-Zernike Theorem (see for example Reference 4) relates the correlations (called visibilities, V) as measured by each antenna pair of the interferometer with the scene intensity (brightness, I). The visibilities are functions of the two spatial frequencies u and v . These are the x and y components respectively of the antenna spacing (baseline) divided by the wavelength. The Theorem states that V and I are a Fourier pair and thus a simple inversion can be utilized to recover the scene intensity. (Figure 3.1) The sparse array of antennas produces, however, only a fraction of the Fourier coefficients. The modeling techniques described in this section addresses the issue of image reconstruction based on an incomplete Fourier transform. To increase the number of Fourier coefficients measured, or the coverage, one can increase either the number of antennas or the bandwidth. In the latter case, the received bandwidth must be subdivided or channelized to provide discrete Fourier coefficients. The design of an interferometric system relies on striking a balance between hardware and processing.

Besides the problem of trying to determine the scene content by only measuring a fraction of the Fourier coefficients, there is a calibration concern. Errors in each antenna measurement can be attributed to uncertainties in its location relative to the other antennas, atmospheric effects on the signal propagation and errors introduced by hardware imperfections. These errors must be removed through processing.

The Astronomical Image Processing System (AIPS) was acquired from the National Radio Astronomy Laboratory. It contains state-of-the-art algorithms developed by the radio astronomy community for image formation, image processing and self-calibration. (See Reference 5.)

There is a penalty paid for trying to recreate the resolution of a large aperture by only sparsely filling the area with antennas. Large, deterministic but confusing, sidelobes appear in the interferometric image. The radio astronomers have descriptively termed this unprocessed image a "dirty" image. The large sidelobes arise since many of the Fourier coefficients necessary to fully determine the image have not been measured. In the inverse Fourier transform performed to create the image, these unmeasured terms are set to zero. The dirty beam is defined to be the dirty image of a point source at the image center. It is equivalent to the point spread function in optics. It is determined by setting all of the measured correlations to one and then Fourier transforming. It is the response of the interferometer to a point source and is fully deterministic.

The dirty image can be thought of as the convolution of the dirty beam with all the sources in the scene. Clearly, the large sidelobes associated with each of the stronger sources will tend to cover the image and mask the weaker sources. The deconvolution of this dirty beam from the dirty image will lead to a "cleaner" representation of the sources in the scene. This is the goal of the nonlinear deconvolution techniques developed by the radio astronomers. (See, for example, Reference 5.) The two principal ones are CLEAN and MEM (maximum entropy method).

3.1 CLEAN and MEM

CLEAN is a straightforward iterative method for removing the sidelobes from the dirty image and uncovering the true sources. In its simplest form, the pixel with the largest amplitude is located; a dirty beam scaled to a fraction of the peak amplitude (that fraction is termed the gain) and located at the peak is subtracted from the dirty image; a tally of the location and strength of the peak is kept; and the process is repeated until the remaining image (called the residual image) is either flat enough or small enough. At that point, all of the point values stored from the found peaks are combined, convolved with an appropriate "clean" beam, and added to the residual image; The result is the "clean" image. As the stronger sources are located and their associated dirty beams are subtracted, the weaker sources emerge from the sea of sidelobes and image fidelity is dramatically improved.

A more sophisticated version of CLEAN, the Clark algorithm, has been implemented in AIPS. The CLEANING iteration has been split into major and minor cycles, in order to speed up execution. Usually, thousands of iterations are necessary.

The second approach for image cleaning is MEM. It is mathematically more complicated than CLEAN. Unlike CLEAN, which has an underlying assumption that the scene is made up of discrete isolated sources, MEM is a much more general nonlinear deconvolution technique. The premise on which it is based states that there are an infinite number of choices for the values of the unmeasured Fourier coefficients and that setting them to zero, as is done in the dirty image formation, is not the optimum choice. MEM is a prescription for choosing the unknown Fourier coefficients.

With the MEM algorithm, an entropy-like function of the image pixel intensities is constructed. This can be related to the information content of the scene. MEM then chooses the values of the unmeasured Fourier terms by maximizing the "entropy", with the constraint that the measured Fourier coefficients match the Fourier transform of the MEMed image to within the noise. This multi-dimensional, constrained maximization has been implemented in AIPS in an iterative scheme that converges rapidly, usually in ten's of iterations.

The radio astronomers have taken advantage of the fact that the main errors arising in interferometric data collection are associated with each antenna. Since correlations are formed pair-wise, there are many more correlations than errors. An iterative technique,

known as self-calibration, has been developed to remove these errors from the data. This algorithm is included in the AIPS package.

3.2 Modeling Results

In Figure 3.2.1, we show an airport scene generated by the phenomenology module of the ARMSS code. For each specific interferometric configuration, a "mask" depicting the corresponding u-v plane coverage is produced. (See Figure 3.2.2) Using this mask, the appropriate Fourier components that the interferometer will measure are filtered out and stored in a file suitable for input into an image processing code such as AIPS. This scene generation procedure is summarized in Figure 3.2.3. The unprocessed and the processed images (using the CLEAN and the MEM algorithms respectively) of the scene are shown in Figure 3.2.4. Finally, to illustrate self-calibration, random phase noise is injected into the received signals in order to corrupt the interferometric image. The self-calibration algorithms allow for the recovery of the original image as shown in Figure 3.2.5..

4.0 Conclusion:

An end-to-end passive millimeter wave system modeling capability has been developed at TRW and state-of-the-art interferometric image processing codes have been acquired. These codes have been applied extensively to the design of radiometric and interferometric imaging systems for diverse commercial and military applications (Reference 6).

References:

1. H. Liebe, "An updated model for millimeter wave propagation in moist air," *Radio Sci.* 20 (1985) 1069
2. F.T. Ulaby, R.K. Moore and A.K. Fung, *Microwave Remote Sensing*, Volume 3, Addison-Wesley Publishing Co., Reading, MA, 1981
3. P.J. Lynch and R.J. Wagner, "Rough-surface scattering: shadowing, multiple scatter, and energy conservation," *Journ.Math.Phys.* 11 (1970) 3032
4. A.R. Thompson, J.R. Moran, and G.W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, John Wiley and Sons, Inc., New York, NY 1986
5. R.A. Perley, F.R. Schwab, and A.N. Bridle, editors, *Synthesis Imaging*, distributed by the National Radio Astronomy Observatory, 1986
6. S.K. Young, R.A. Davidheiser, B. Hauss, P. Lee, M. Mussetto, M.M. Shoucri, and L. Yujiri, "Passive millimeter-wave imaging", *TRW Quest Mag.*, vol.13, no.2, pp 3-20, Winter 1990/1991

APPROXIMATIONS TO PHENOMENOLOGY FOR REAL-TIME OPERATION

Polygonal tessellation of scene elements, with tracing only to polygon vertices	Much fewer rays to trace (by factor of 1000 or more) permits near-real-time operation	Less accurate description of scene geometry Difficult to simulate reflection Introduces interpolation error in computed temperatures between polygon vertices
Lookup table for sky temperature vs. zenith angle, computed at start of simulation	Saves repetitive integration of rays from top of troposphere for downwelling atmospheric radiation	Negligible interpolation error in sky temperature at arbitrary zenith angle Limited to azimuthally symmetric sky conditions (i.e., no patchy clouds)
Lookup table for emissivity of specular surfaces as a function of incidence angle, computed at start of simulation	Saves repetitive calculation of dielectric properties and single or multiple layer emissivities for terrain surface elements	Negligible interpolation error in computed terrain emissivity at arbitrary incidence angle

Table 2.4.1

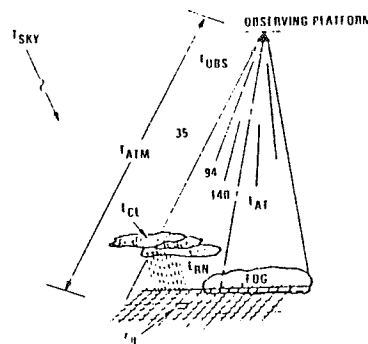
APPROXIMATIONS TO PHENOMENOLOGY FOR REAL-TIME OPERATION

Lookup table for rough surface apparent temperature as a function of observation angle (with normal to mean ground plane pointed towards zenith), computed at start of simulation	Saves repetitive calculation of multidimensional integrals for emitted and scattered radiation from anisotropic random rough surfaces	Restricted to ground planes which are close to horizontal Doesn't allow for shadowing by other scene elements Negligible interpolation error in computed apparent temperature at arbitrary angle
Approximate method for treating upwelling atmospheric radiation	Much shorter computation time for evaluation of upwelling atmospheric radiation (by a factor of at least 100) permits real-time operation	Negligible integration error introduced when sensor platform height is within a few kilometers of ground
Single specular reflection model, which assumes second reflection off terrain is from a perfectly conducting, horizontal ground plane	Some computational savings in not having to follow multiply reflected rays	Will not correctly treat the interaction between elements of 3-D terrain and obstacles

Table 2.4.1 (Cont.)

PHENOMENOLOGY MODEL

- Atmospheric Propagation Model
- Atmospheric Weather Model
- Surface/Terrain Physics Model
- Ray Tracing Algorithm



SENSOR MODEL

- Sensor Optics Model
- Detector Model
- Mechanical/Electrical Effects Model

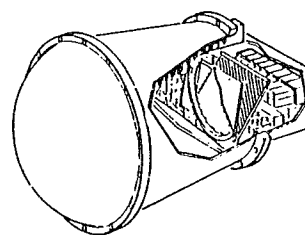
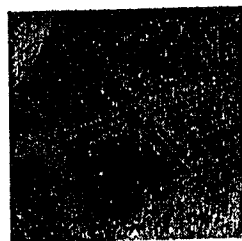


IMAGE PROCESSING MODEL

- Image Enhancement Techniques
- Image Restoration Techniques
- Definition of Real-Time Algorithms and Hardware



DISPLAY MODEL

- Frame-by-Frame Animation

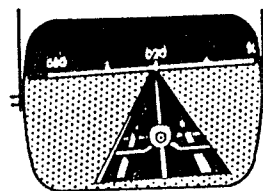


Figure 2.1 Principal Components of TRW's End-to-End Advanced Radiometric Millimeter Wave Scene Simulation Code (ARMSS)

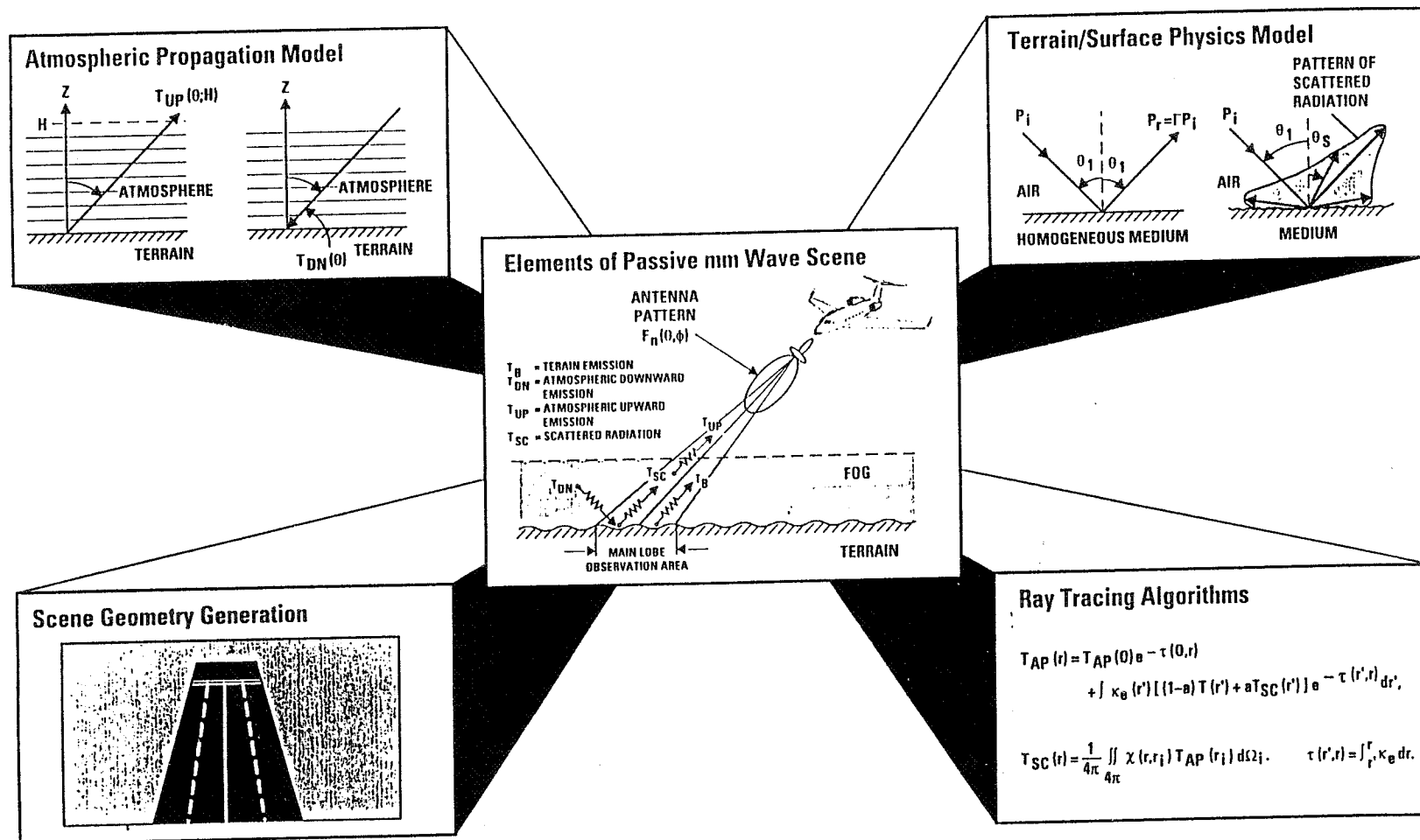


Figure 2.2 Basic Elements of the Passive Millimeter Wave Phenomenology Models in ARMSS

- MOST RECENT VERSION OF POINT ATTENUATION RATE MODEL FROM ITS (H. UEBE)
- CALCULATES RADIO PATH PARAMETERS (ATTENUATION AND PROPAGATION PATH DELAY EFFECTS) FROM METEOROLOGICAL DATA (P - T - RH - W - RR)
- MODEL'S RANGE OF VALIDITY IS 0 - 1000 GHz
- MODEL INCLUDES:
 - PRESSURE BROADENED RESONANCE LINES FOR H_2O (22-997 GHz) AND O_2 (49-834 GHz)
 - CONTINUUM ABSORPTION DUE TO H_2O LINES ABOVE 1 THz AND EMPIRICAL CORRECTIONS REQUIRED BY V-W LINE SHAPES AWAY FROM RESONANCE
 - CONTINUUM DUE TO NON-RESONANT O_2 AND PRESSURE INDUCED N ABSORPTION
 - HYDROSOL ATTENUATION MODEL (RAYLEIGH ABSORPTION FOR HAZE, FOG, AND CLOUDS)
 - PARAMETERIZED, POWER-LAW RAIN ATTENUATION MODEL TO SIMULATE Mie SCATTERING

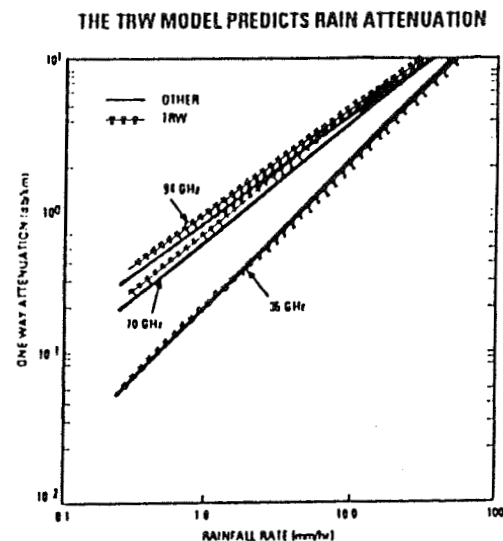
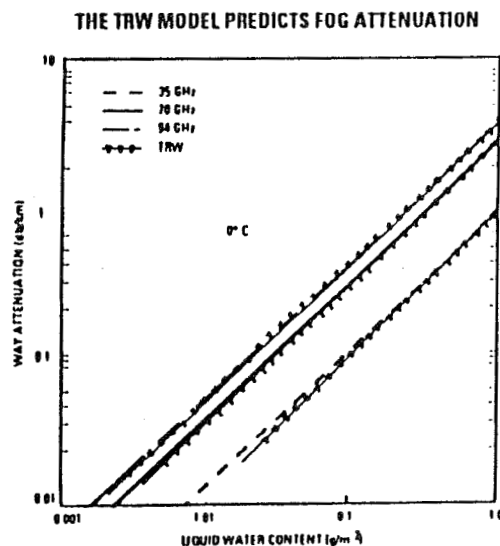
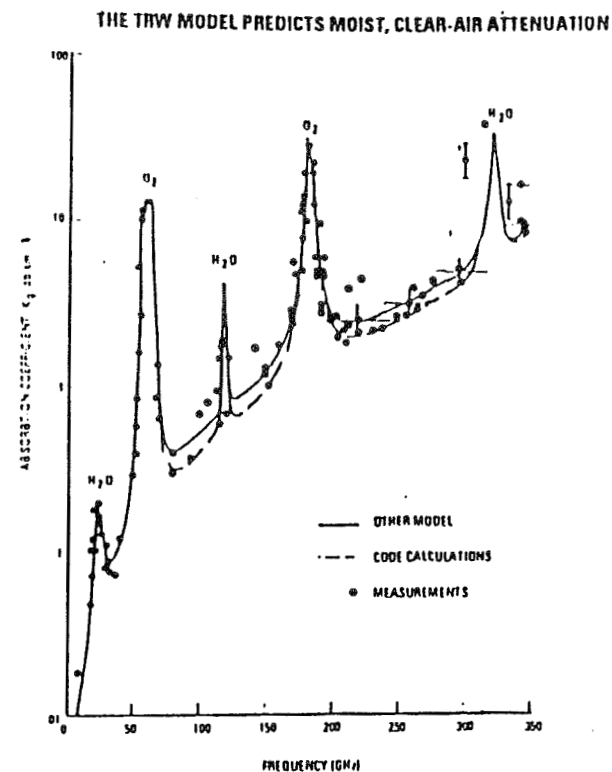


Figure 2.1.1 Comparisons of ARMSS-Predicted Atmospheric Attenuation With Published Data

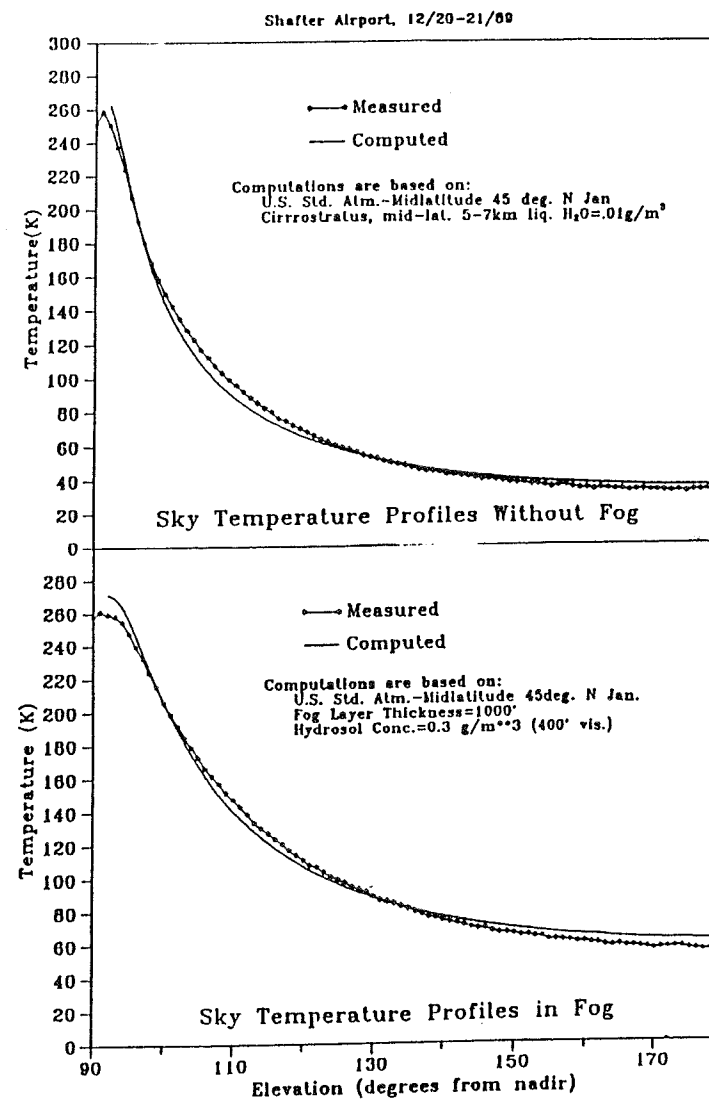
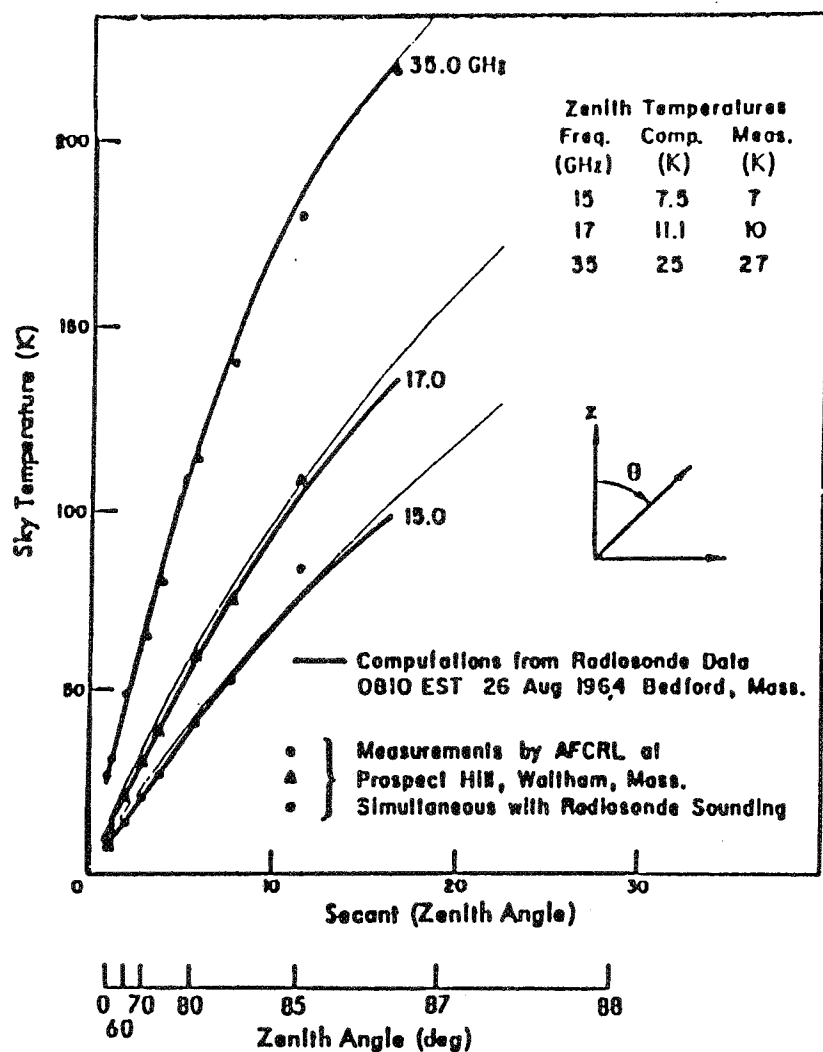


Figure 2.1.2 ARMSS-Predicted Sky Temperature Profiles Under Clear Air and Fog Conditions

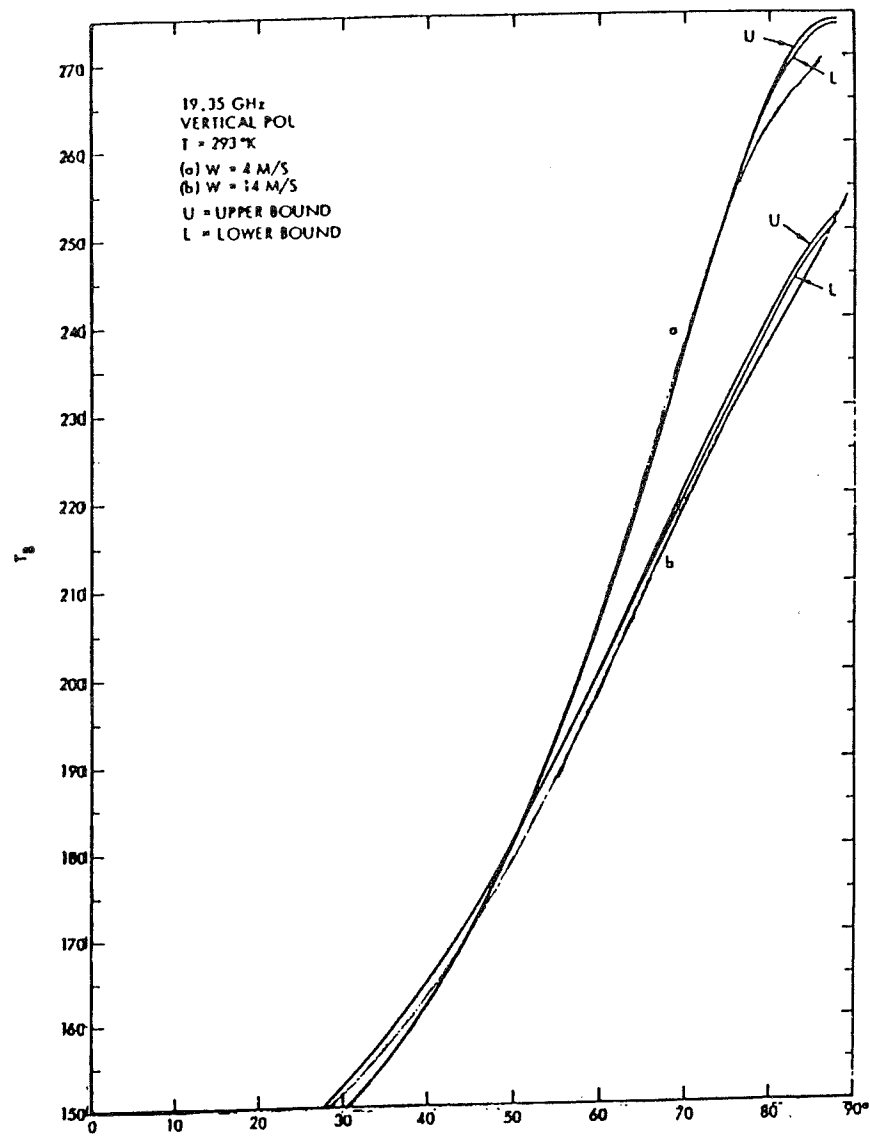
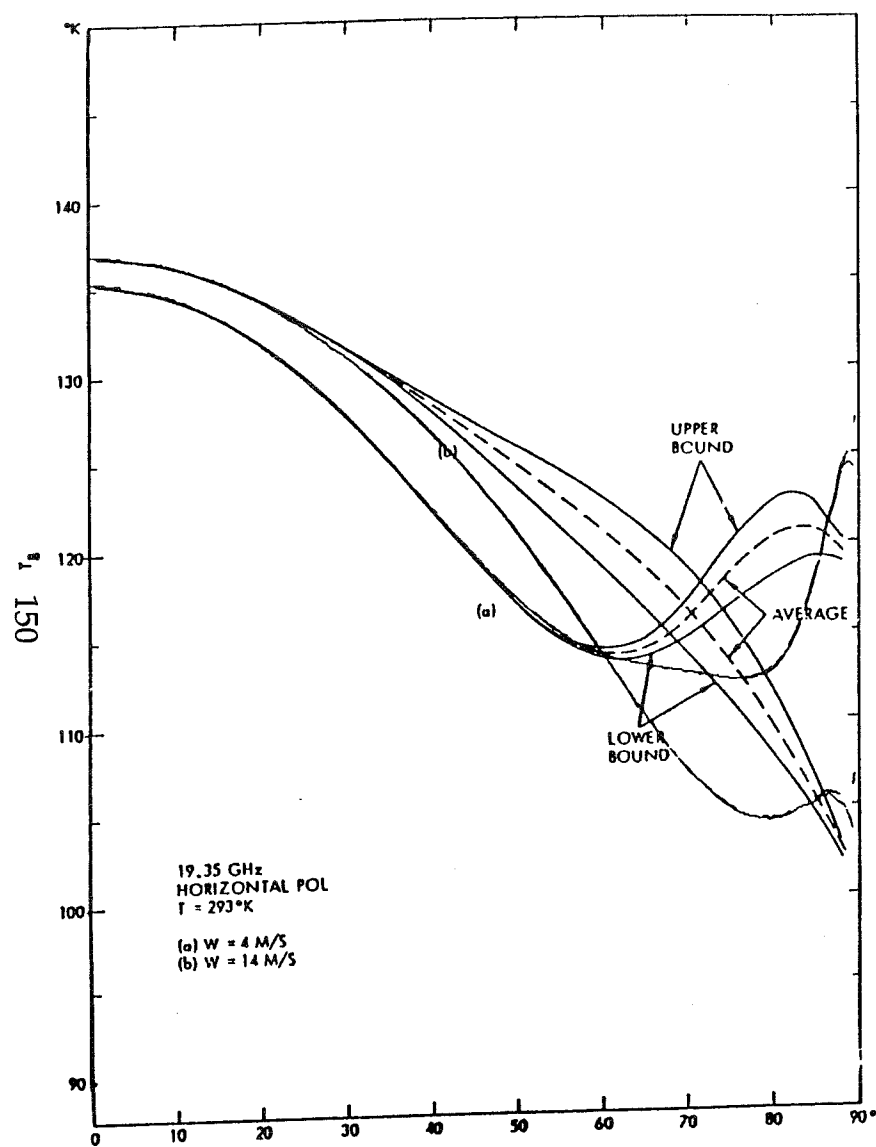


Figure 2.2.1 Comparison of ARMSS-Scatter and a Double-Scatter Surface Model for (a) Horizontal Polarization and (b) Vertical Polarization

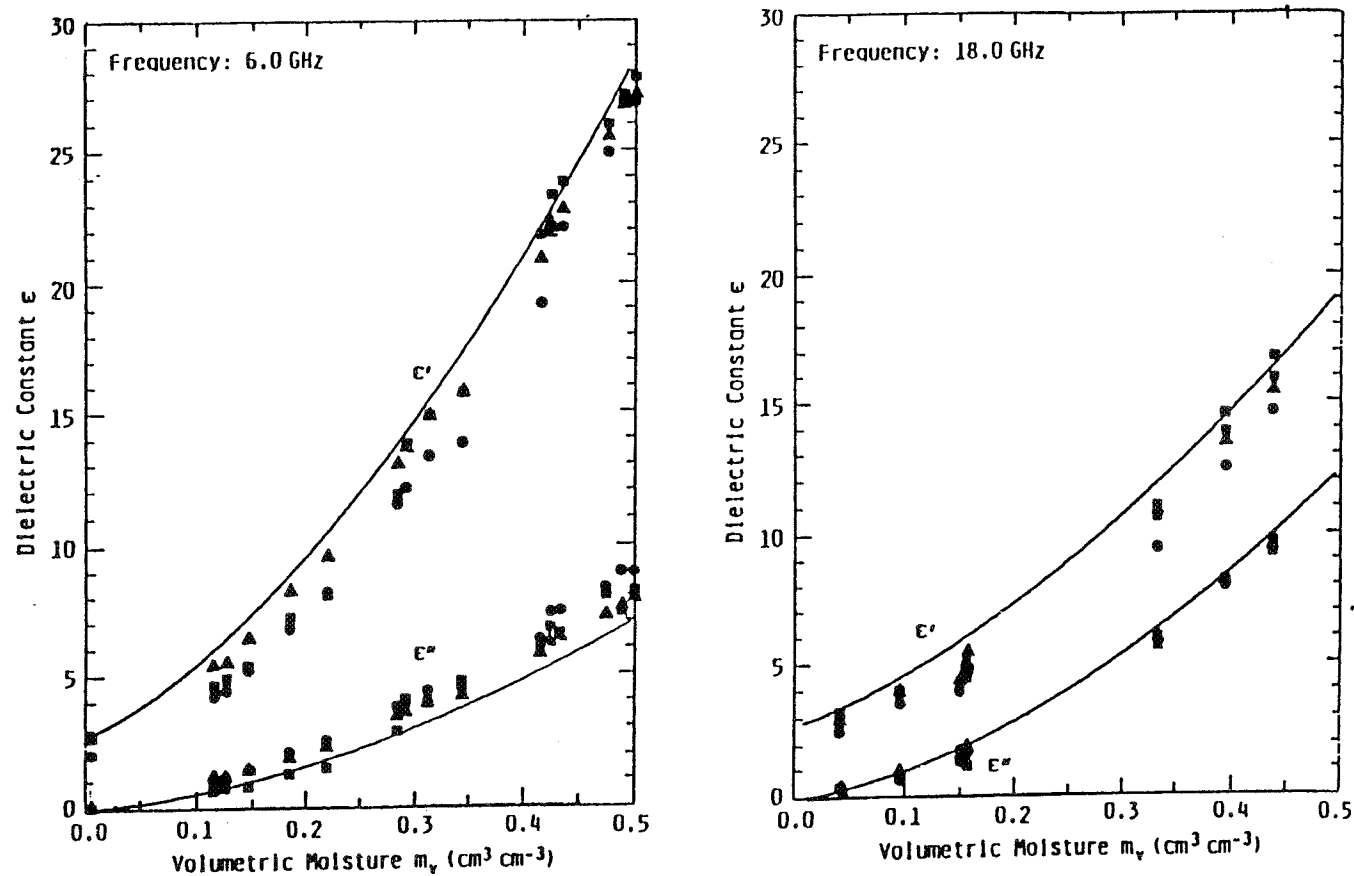


Figure 2.2.2 ARMSS-Predicted and Measured Dielectric Properties for a Silty Clay at (a) 6 GHz and (b) 18 GHz

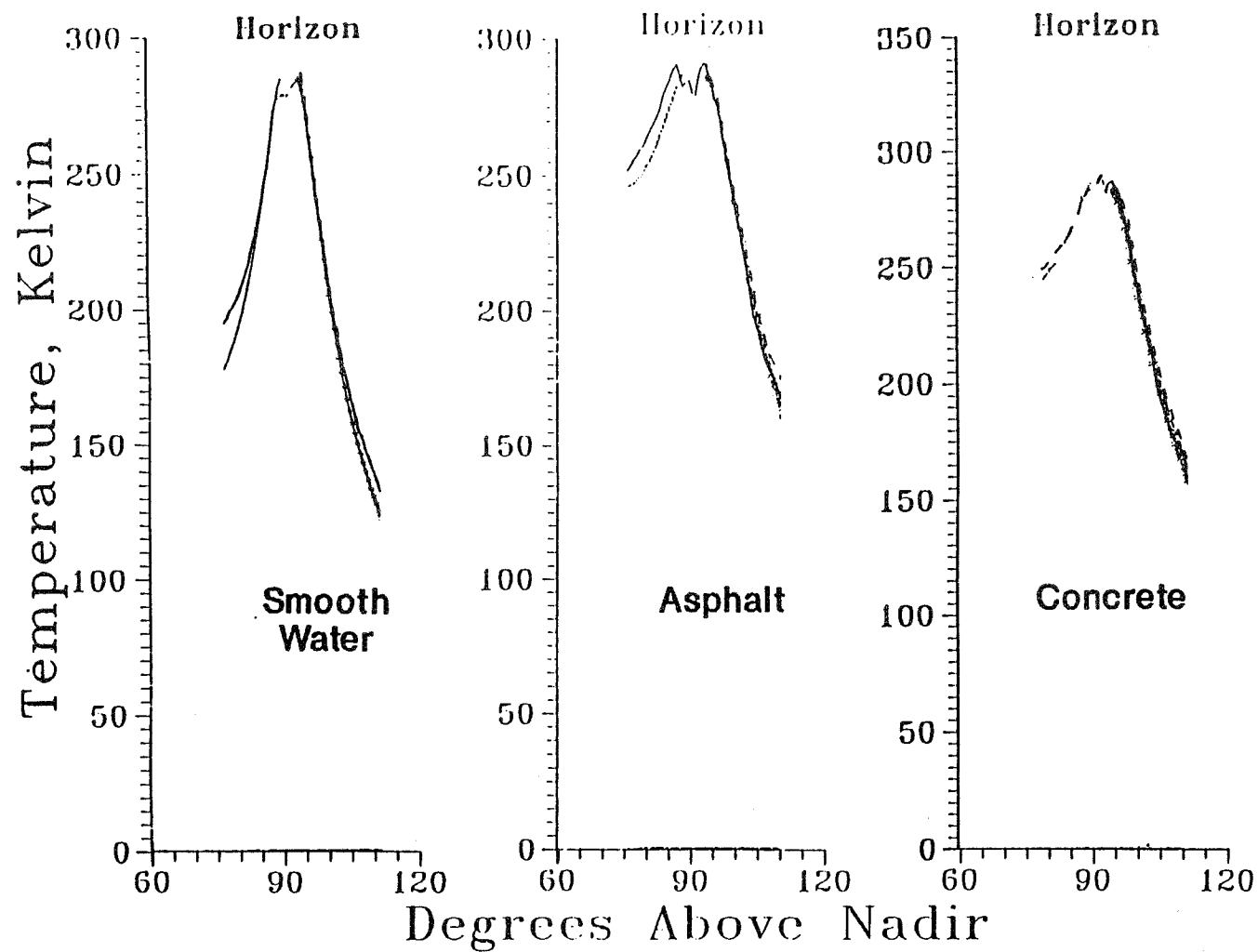


Figure 2.3.2 Measured and ARMSS-Predicted Radiometric Scans of Various Terrain Media

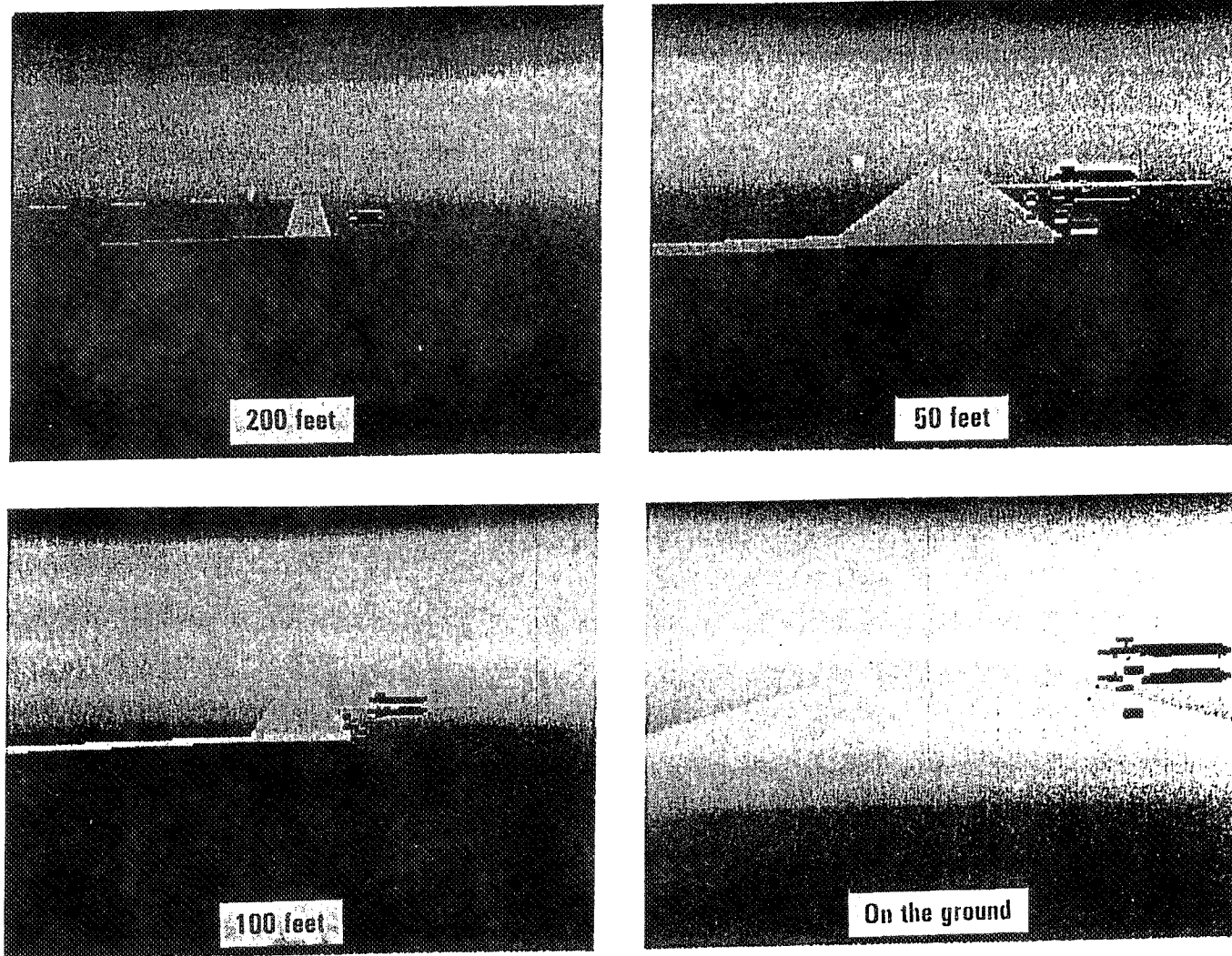
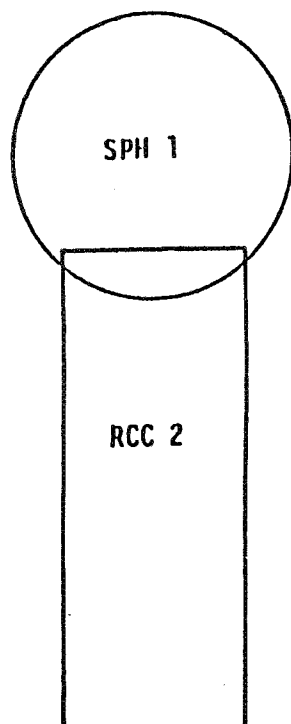
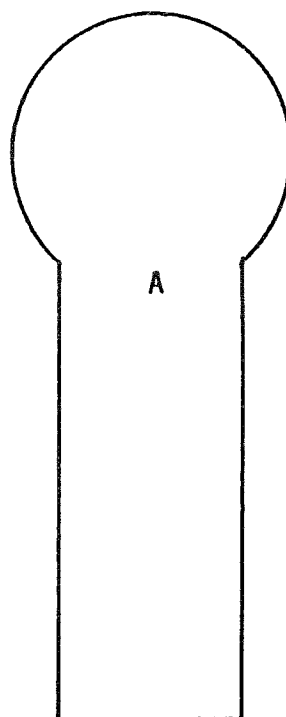


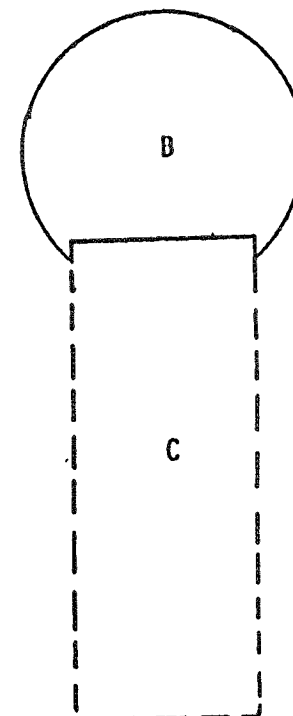
Figure 2.3.1 ARMSS-Generated High Resolution Millimeter Wave Runway Scenes Under Heavy Fog and Wet Surface Conditions



(a)



(b)



(c)

Figure 2.3.3 Three-Dimensional Objects Formed from (b) Union and (c) Exclusion of a Sphere and Cylinder

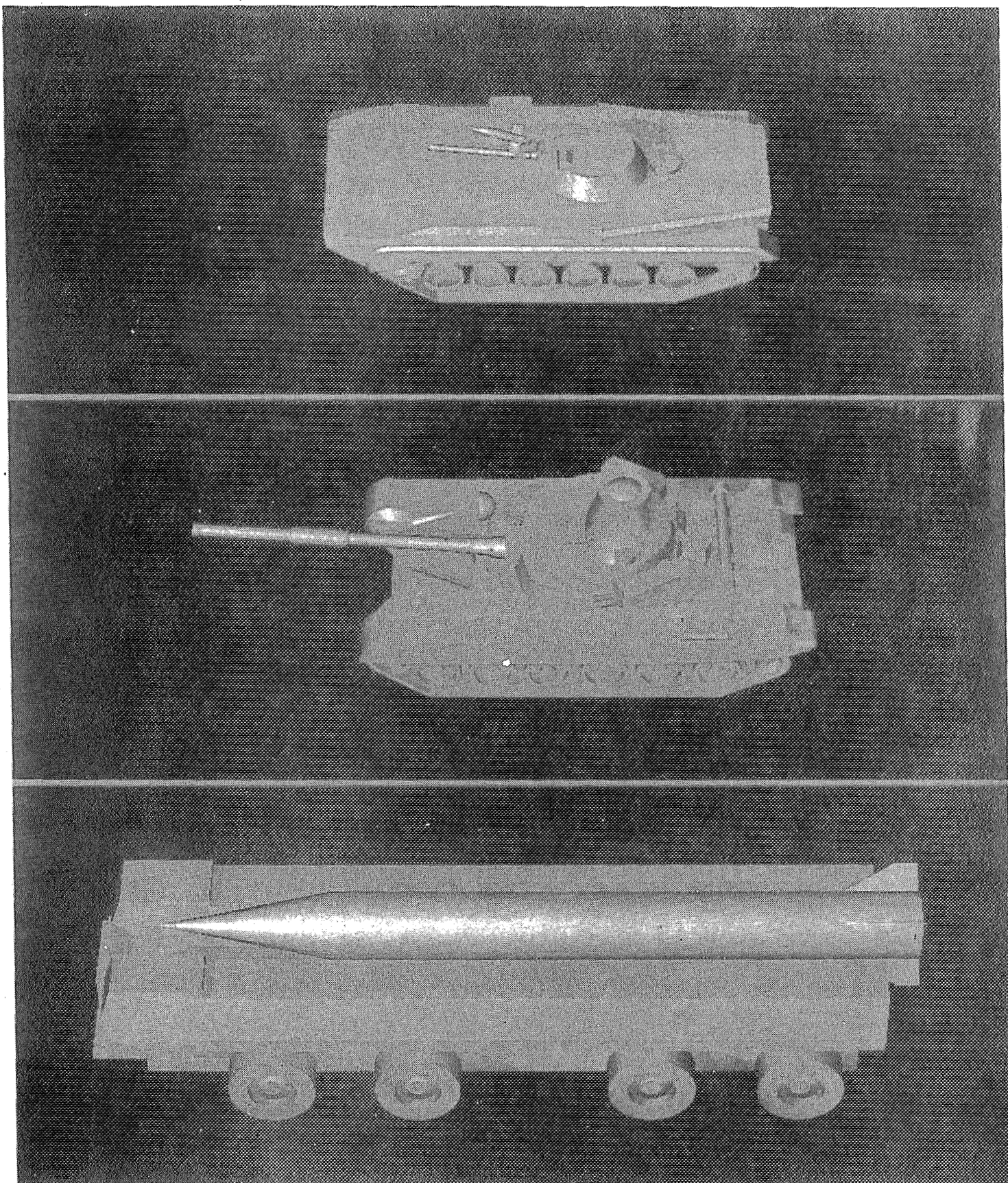


Figure 2.3.4 Combinatorial Geometry Models of BMP-1 Troop Transport, T-72 Tank, and SS-24 Missile and Mobile Launcher

Methods for Including Upwelling Atmospheric Radiation

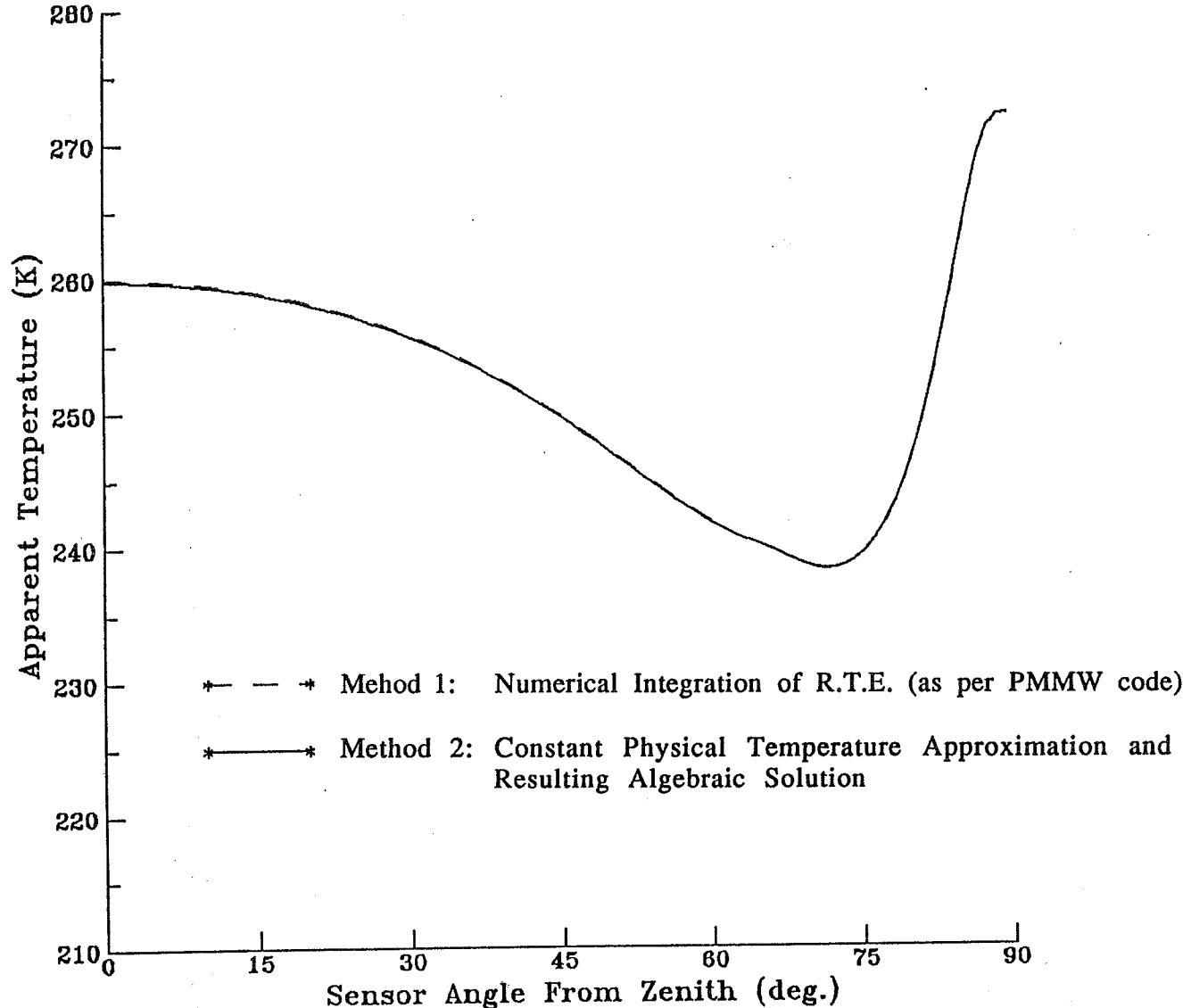
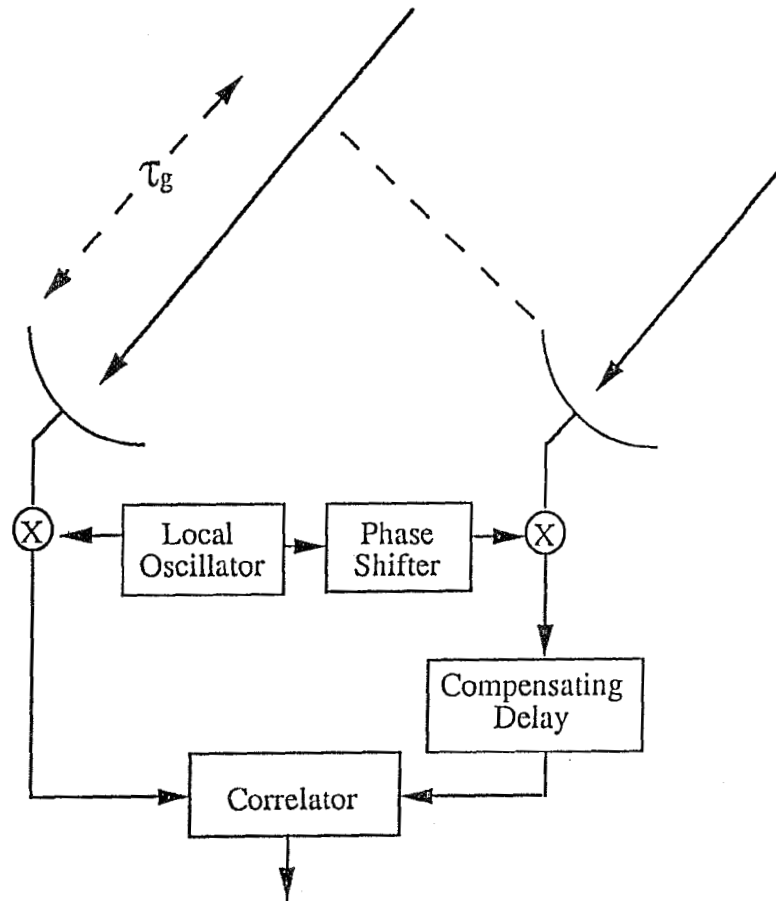


Figure 2.4.1

Figure 3.1

Two-element Interferometer Is The Building Block For The Interferometric Imager



- Correlator output "samples" a scene spatial frequency component
- Correlations between different apertures or at different frequencies produce additional "samples"
- Image is generated by Fourier Transform of "samples"
- Enhancement techniques applied to compensate for incomplete, or phase corrupted "samples"

$$V(u,v) = \iint A(l,m) I(l,m) e^{-2\pi i(ul+vm)} dl dm$$

Figure 3.2.1

LOW DENSITY AIRPORT SCENE

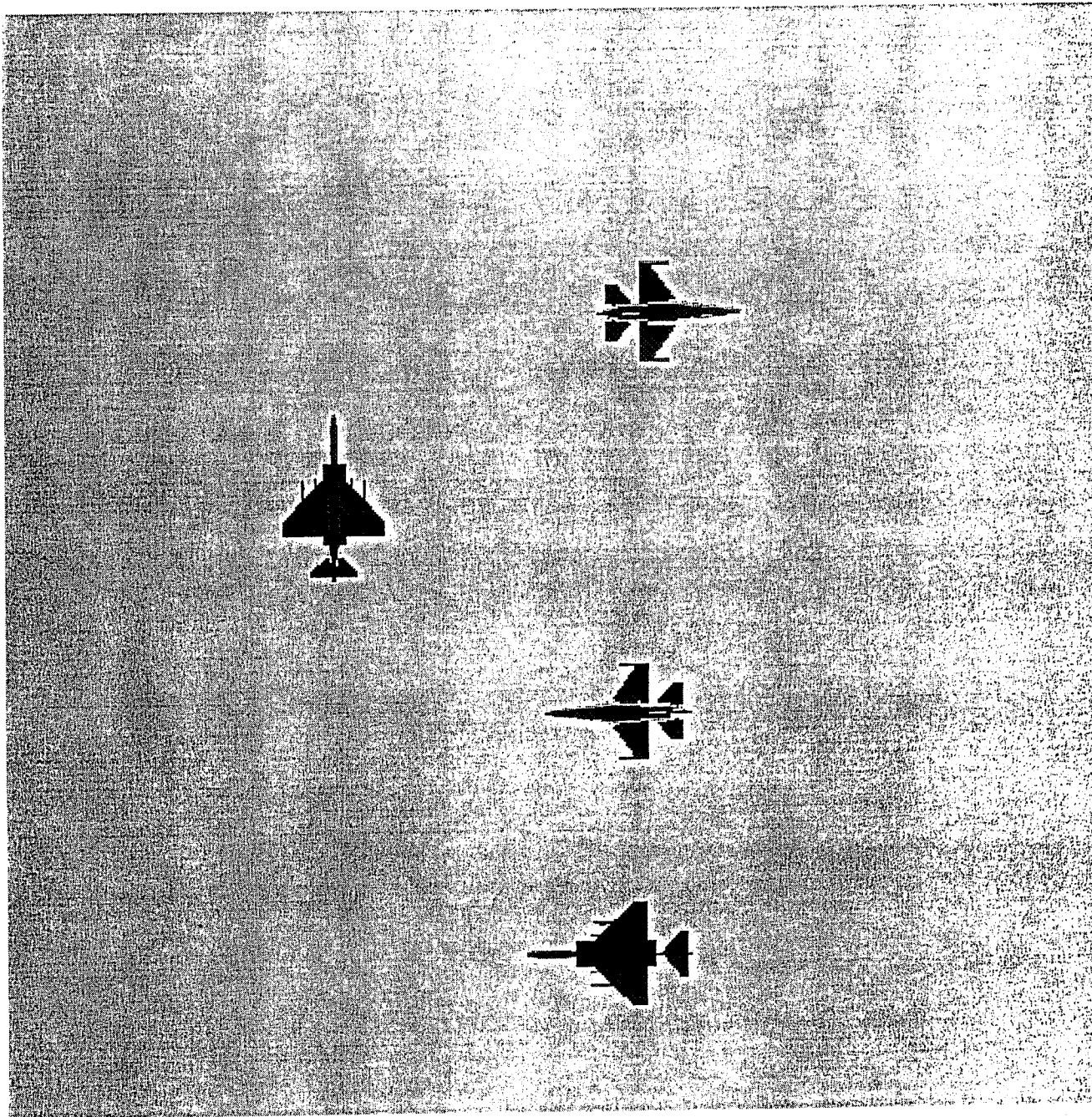


Figure 3.2.2

INTERFEROMETRIC SAMPLING

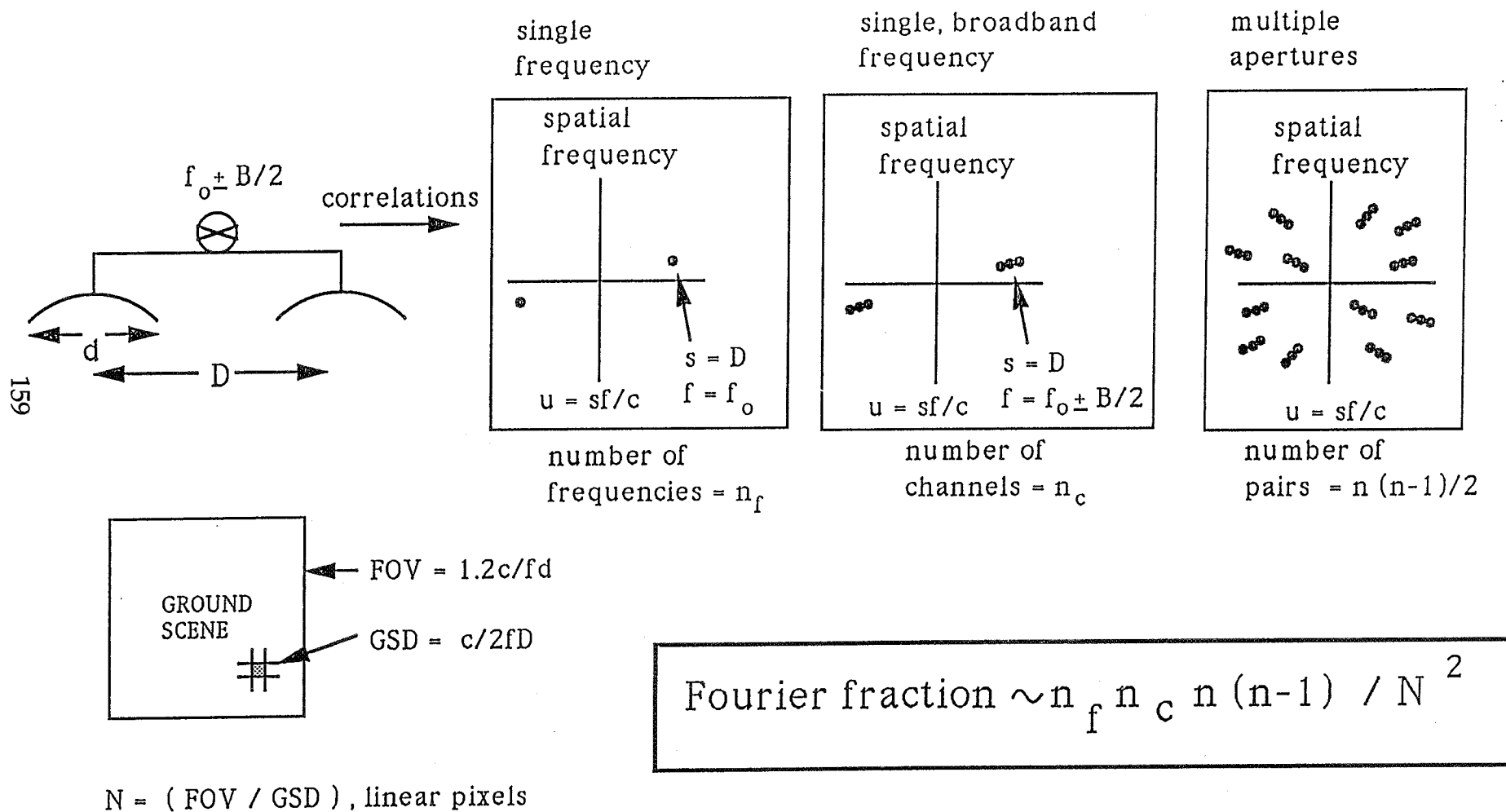


Figure 3.2.3

Procedure

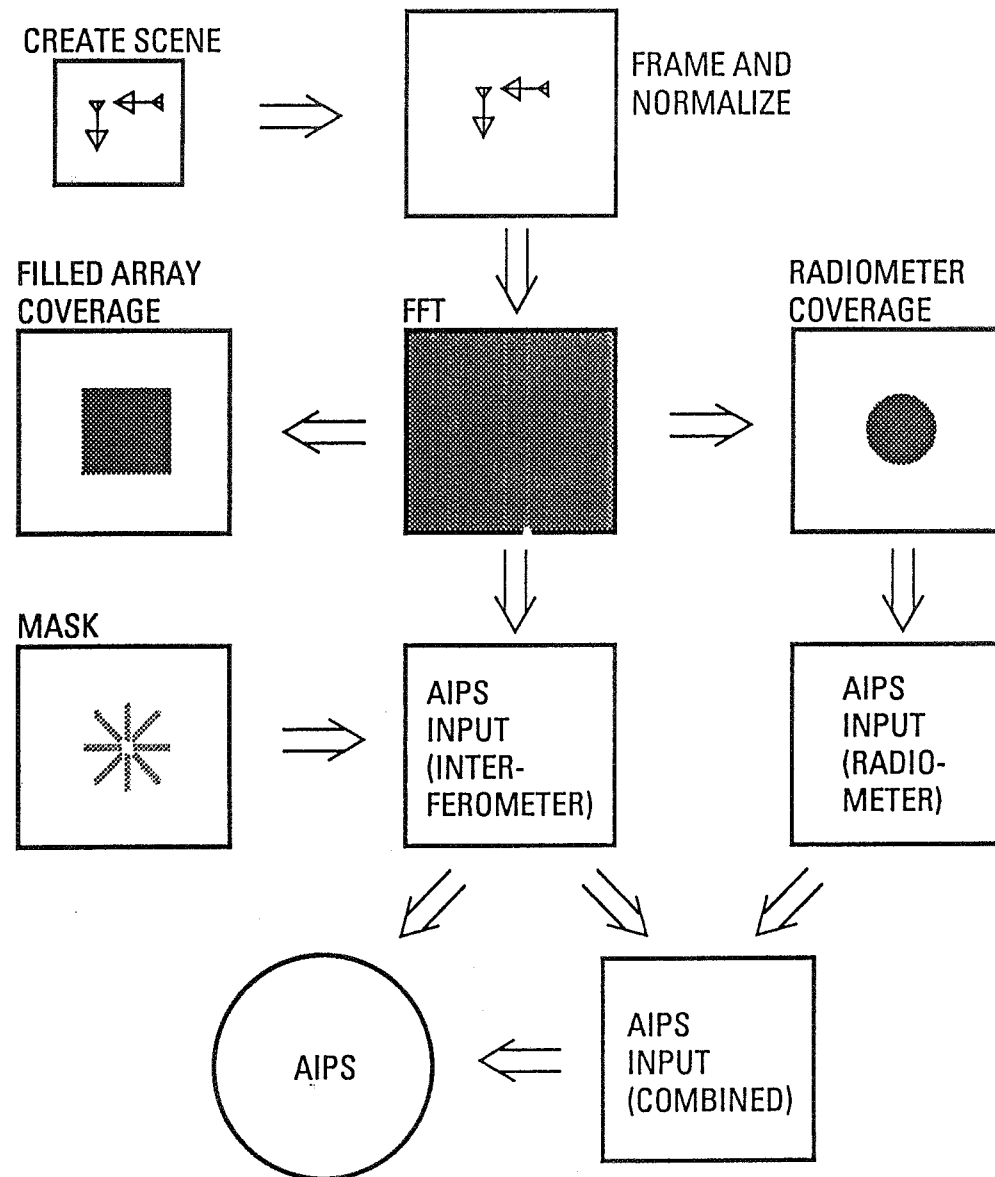
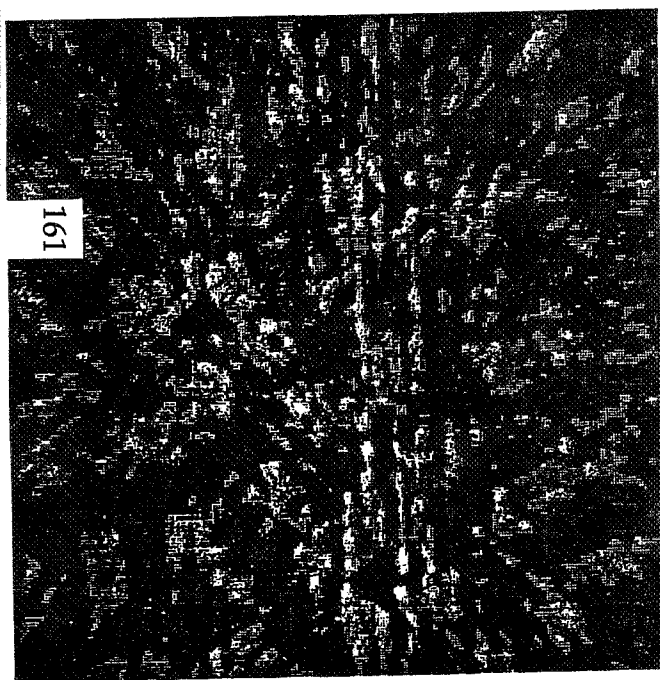
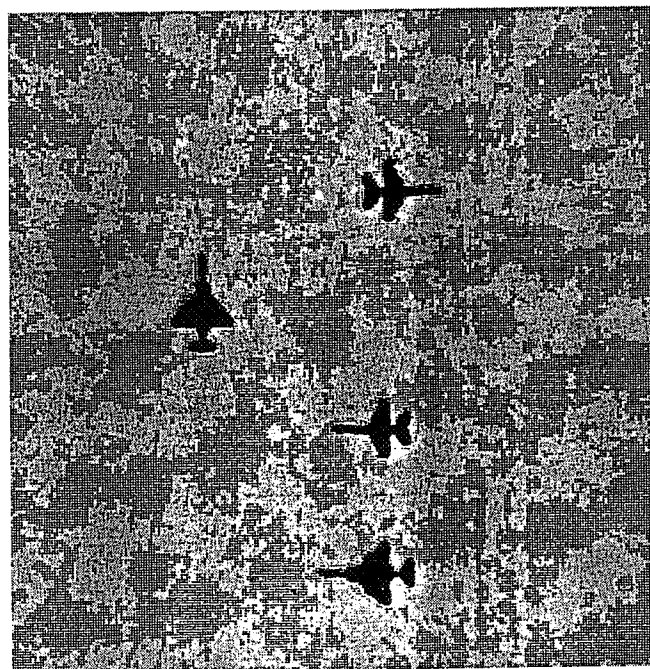


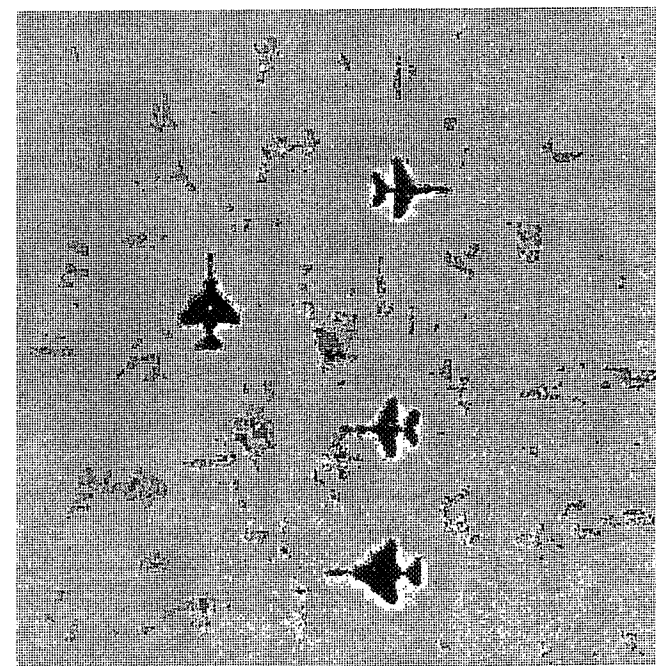
Figure 3.2.4 Unprocessed and Processed Images



IMAGE



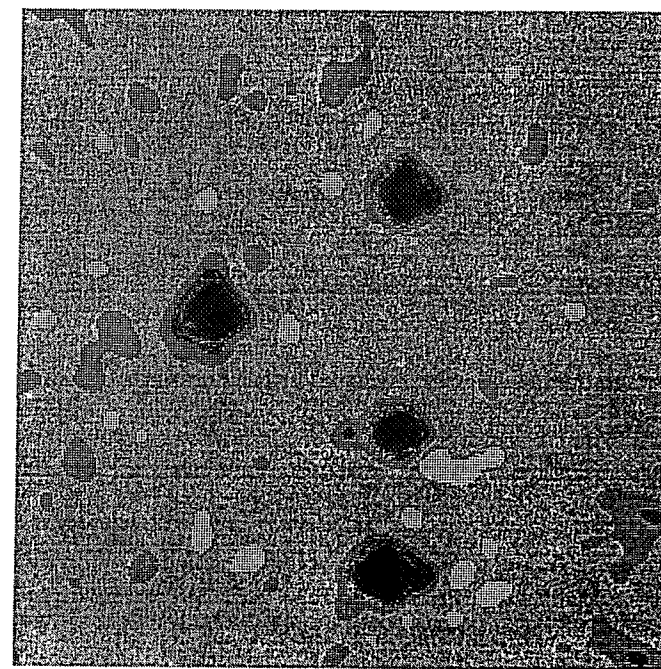
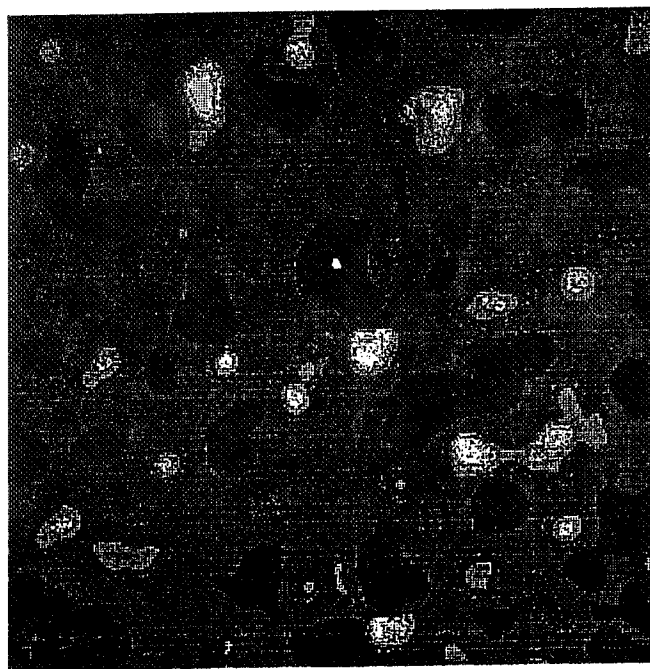
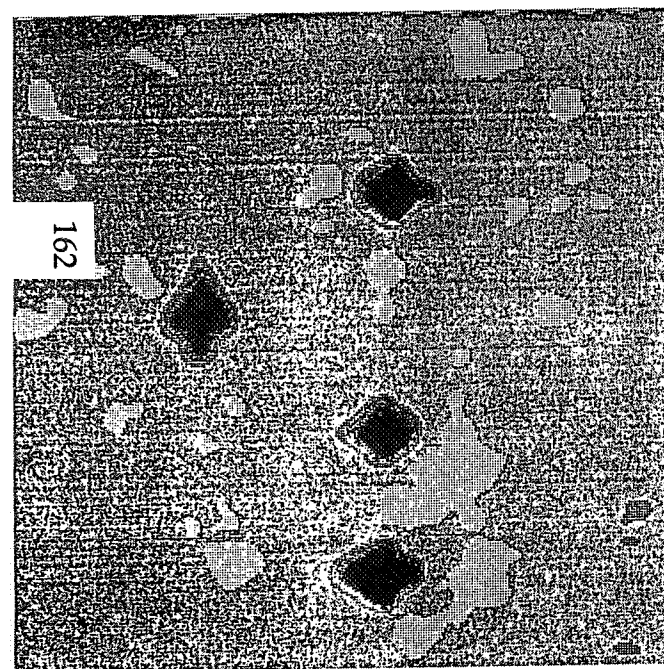
CLEAN



MEM

Figure 3.2.5

SELF-CALIBRATION CLEANED IMAGES



UNCORRUPTED

CORRUPTED

CALIBRATED

Simulation of a Passive Millimeter Wave Sensor 205739

William W. Kahlbaum
Lockheed Engineering and Science Corporation

 P- 18
249922**ABSTRACT**

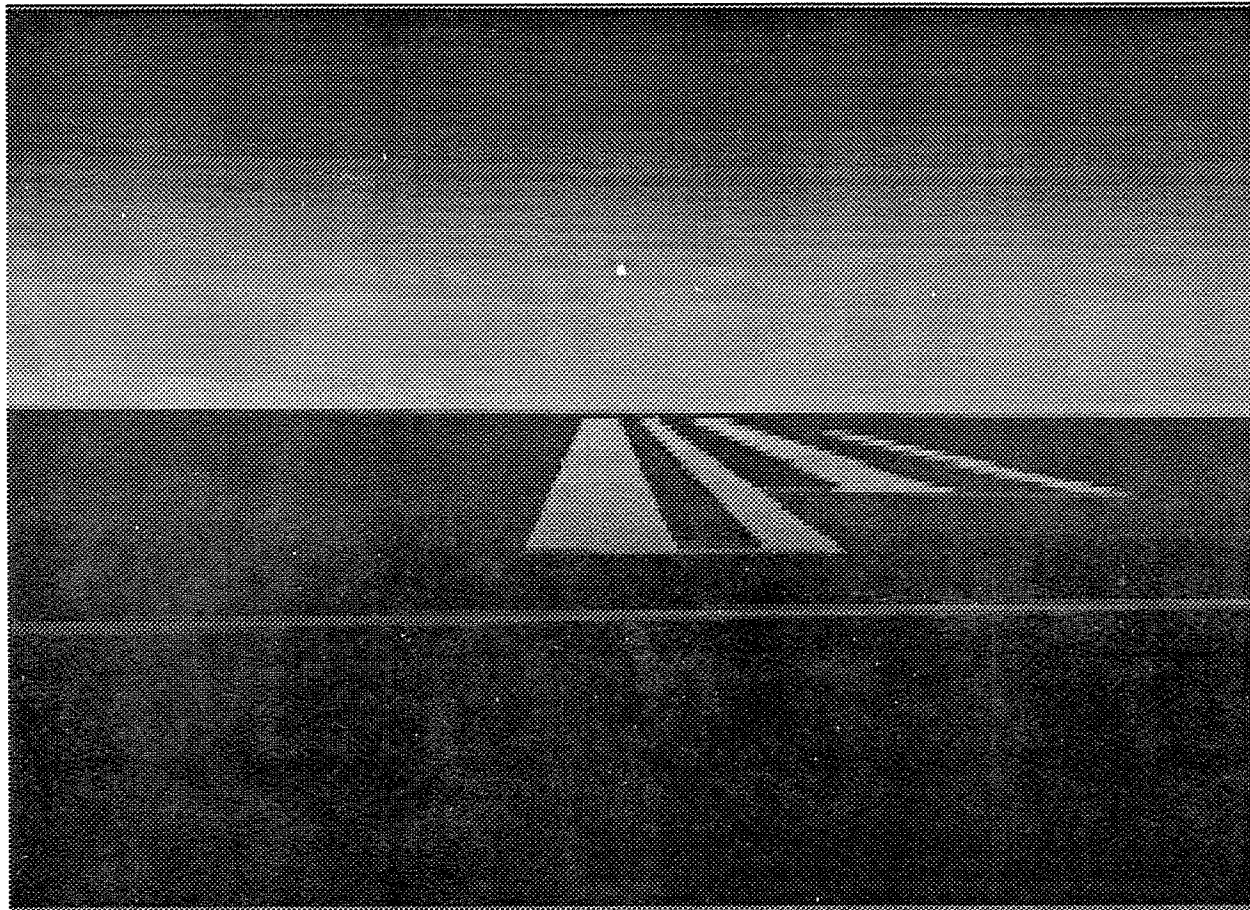
The visual display expected to be generated by a Passive Millimeter Wave (PMMW) camera and sensor system has been simulated on a Silicon Graphics IRIS workstation at the NASA Langley Research Center (LaRC). The low resolution of the sensor has been simulated by graphically manipulating the scene as it is being drawn by the IRIS in real time. Camera field of view, sensor resolution, and sensor update rate are the controllable parameters. Physical effects such as lens model, radome effects, and noise have not been included at this time. An approximate dynamic model of the atmospheric phenomenology has been included which generates the gray-scale intensity values in real time for the simulated image. The gray-scale values are proportional to temperature. A snapshot capability which captures individual image frames during real-time operation has been included. These images were used to validate the approximate phenomenology model against a more rigorous physical model.

Introduction

- **Graphics techniques used to create Passive Millimeter Wave (PMMW) display**
- **Interface with the atmospheric phenomenology model**
- **Solutions to problems which were encountered**

Simulated Passive Millimeter Wave Display

- Low Resolution Corresponds to Anticipated Sensor Element Density
- Gray Scale Corresponds to Temperature Value



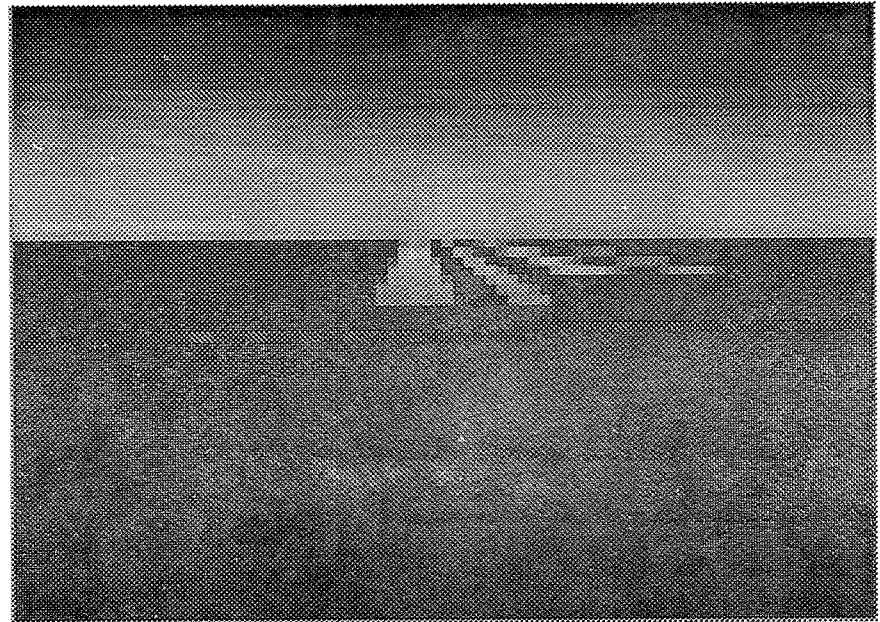
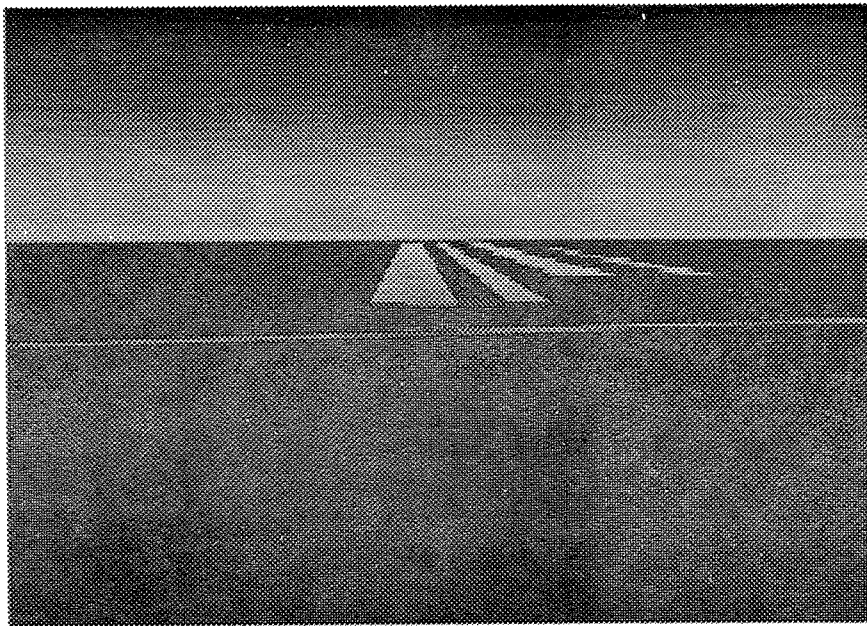
Controllable Parameters

- **Sensor field of view**
- **Sensor resolution**
- **Sensor update rate**

Variation of Resolution with Constant Field of View

- 30 X 24 degree Field of View
- 0.1 degrees/pixel

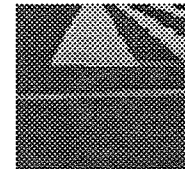
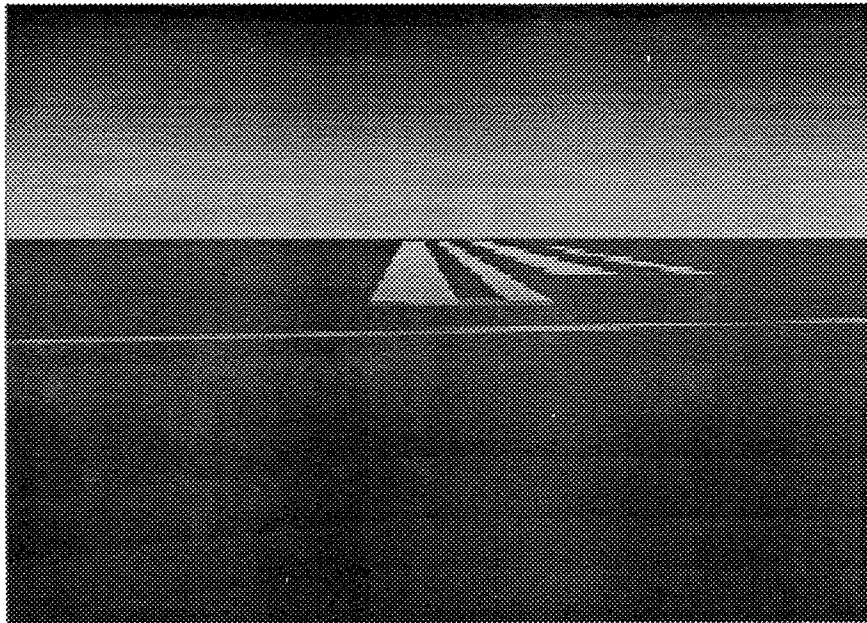
- 30 X 24 degree Field of View
- 0.34 degrees/pixel



Variation of Field of View with Constant Resolution

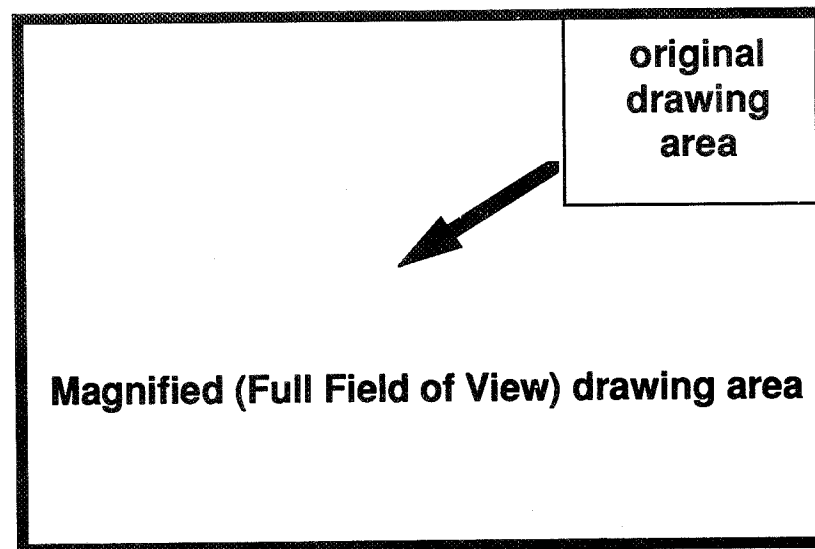
- 30 X 24 degree Field of View
- 0.1 degrees/pixel

- 5 X 5 degree Field of View
- 0.1 degrees/pixel



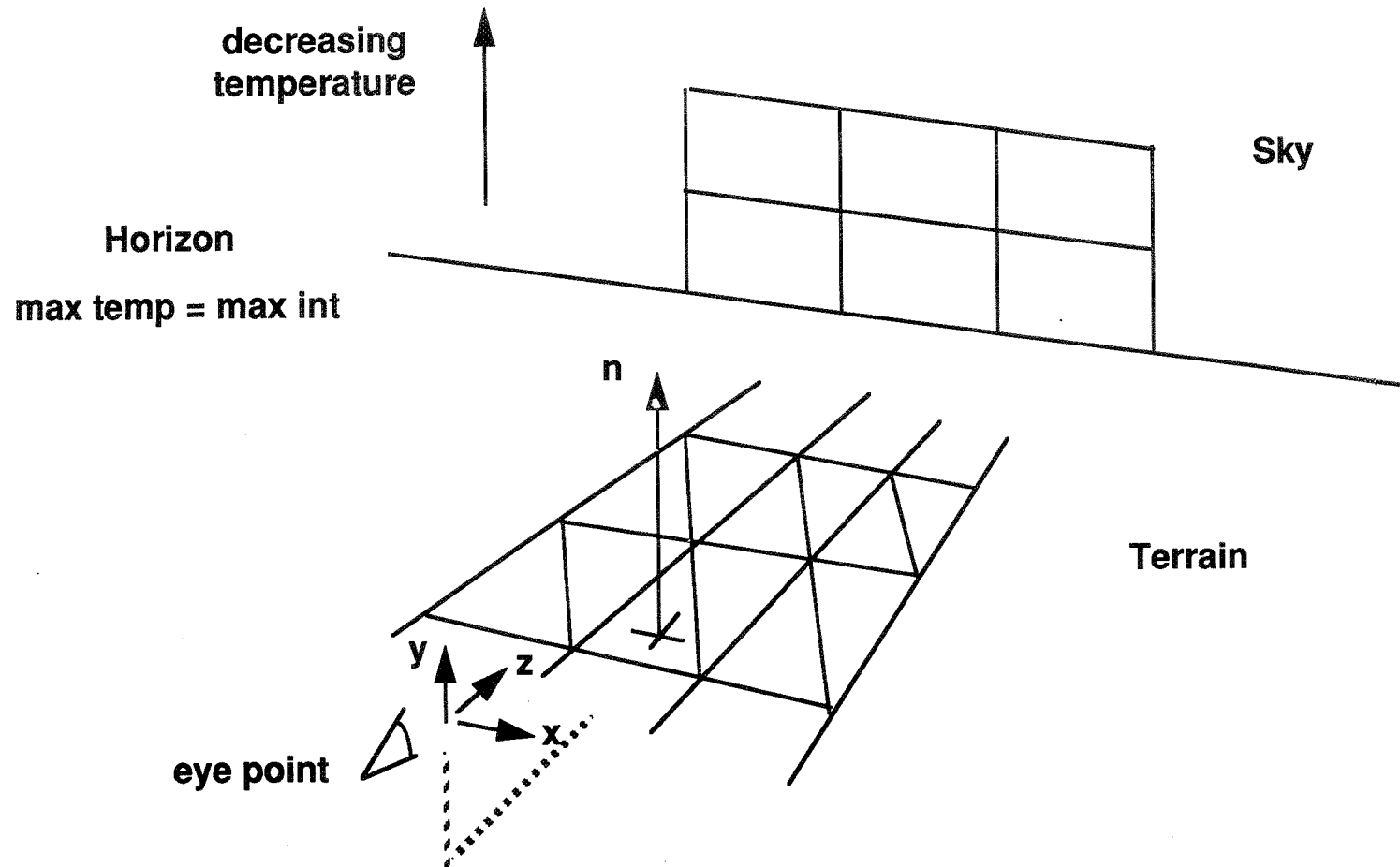
Resolution

- Image drawn in original drawing area at normal screen resolution
- Image copied to full field area with zoom
- Full field area may or may not cover entire screen, depending on the field of view
- Update rate controlled by changing the animation update rate of the IRIS

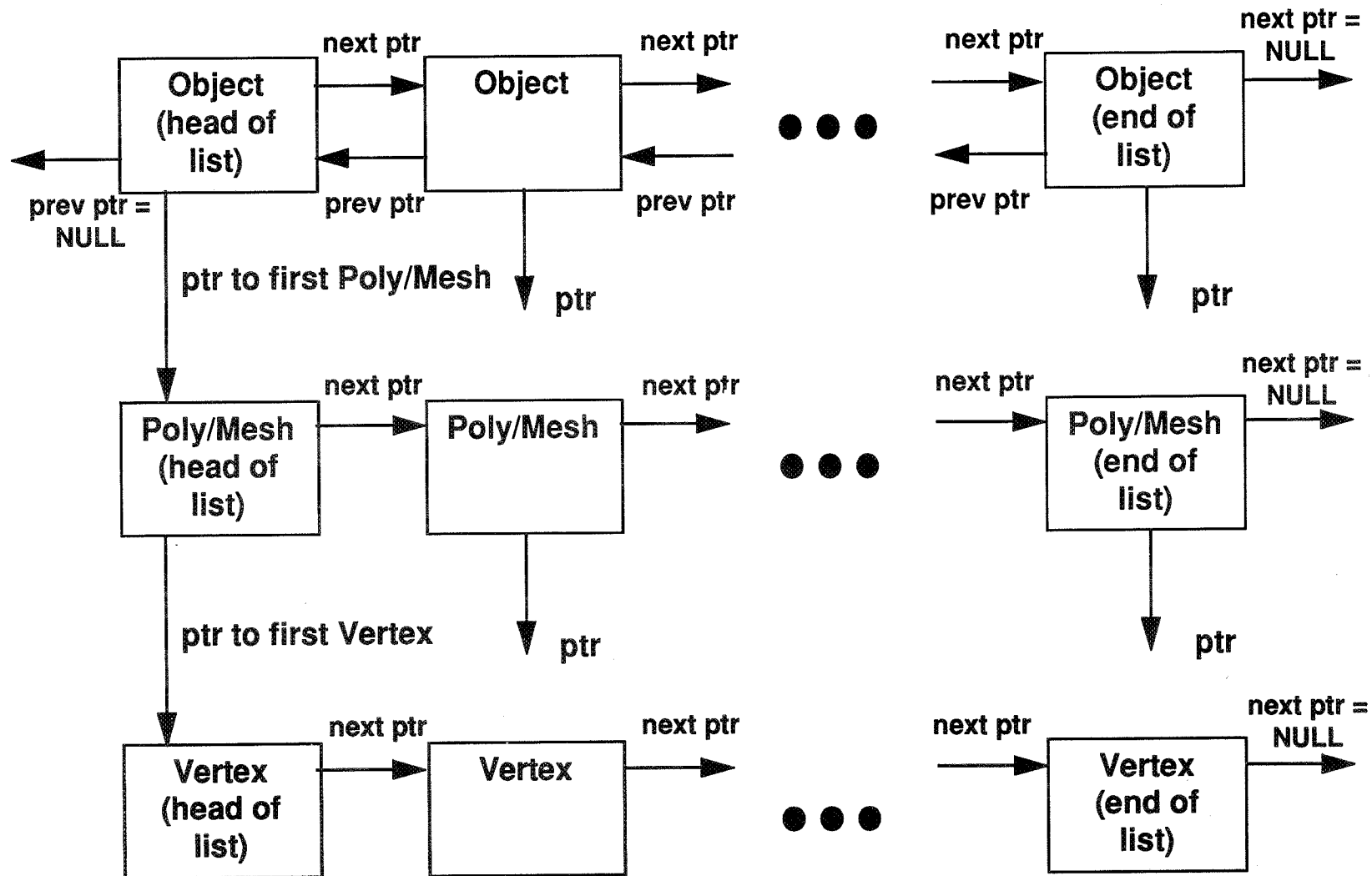


Polygonal Data Base

Sky and Terrain

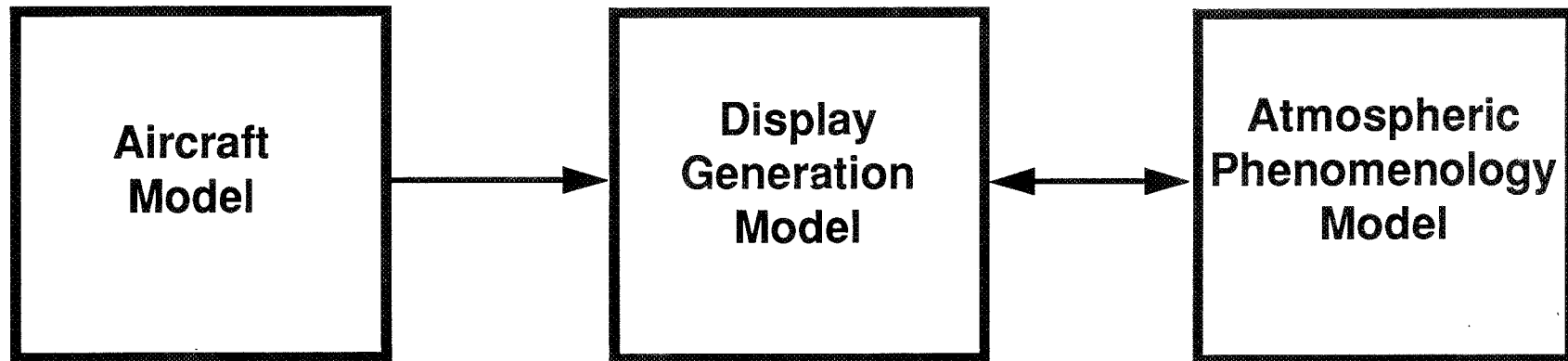


Data Base Tree Structure



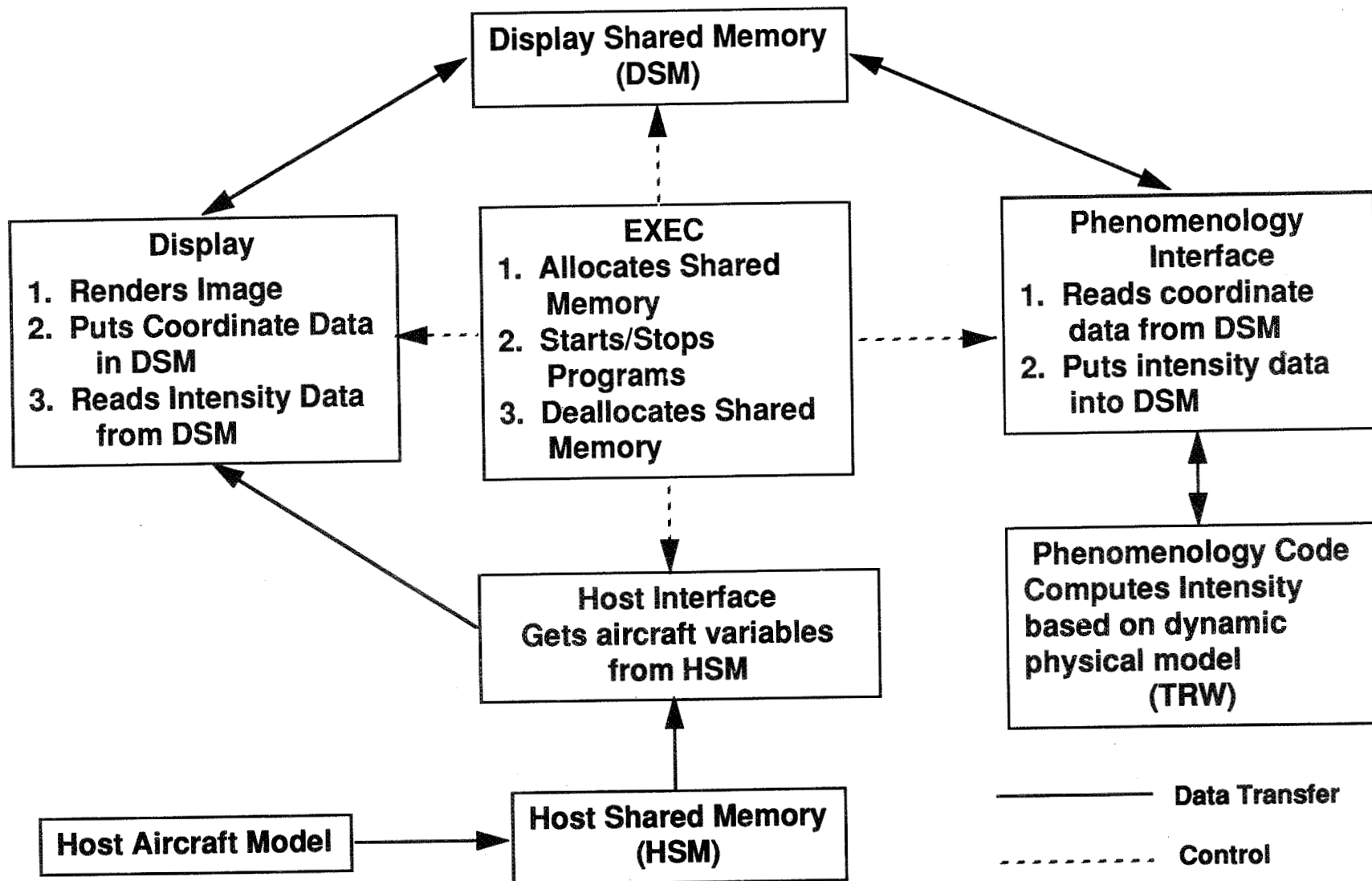
Simulation Models

172



Real-Time Program Flow Diagram

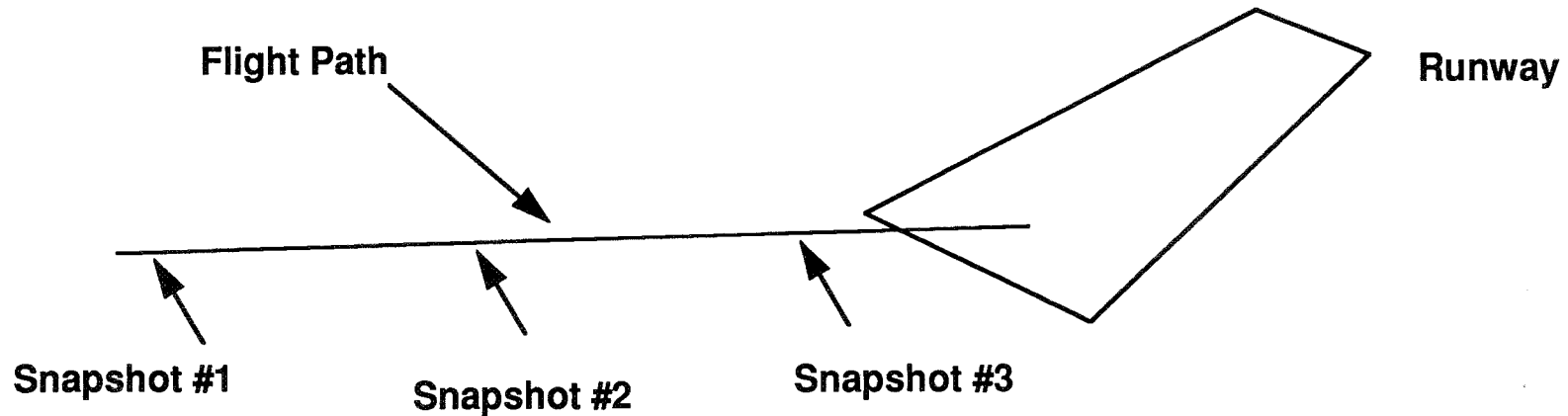
173



Task Summary

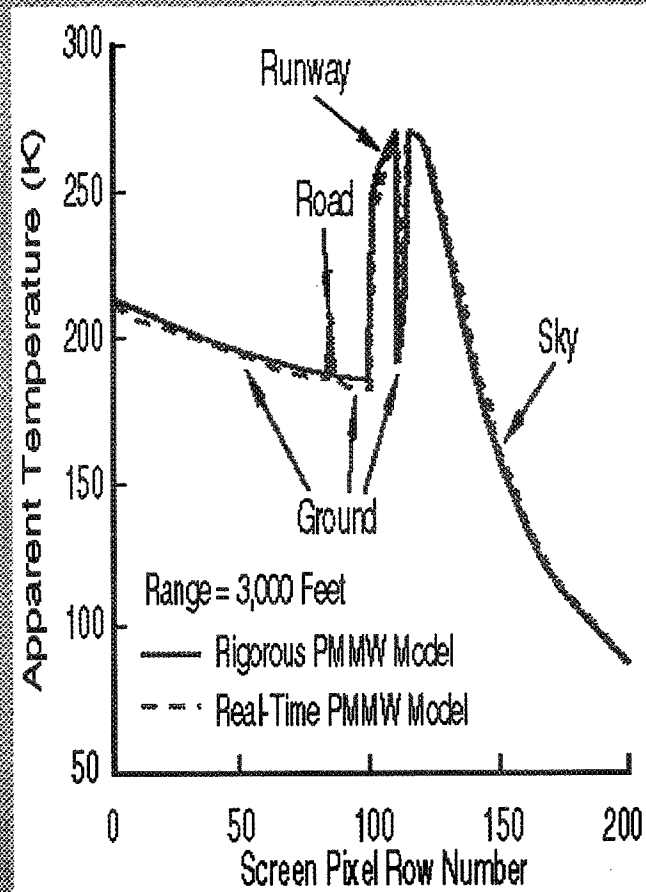
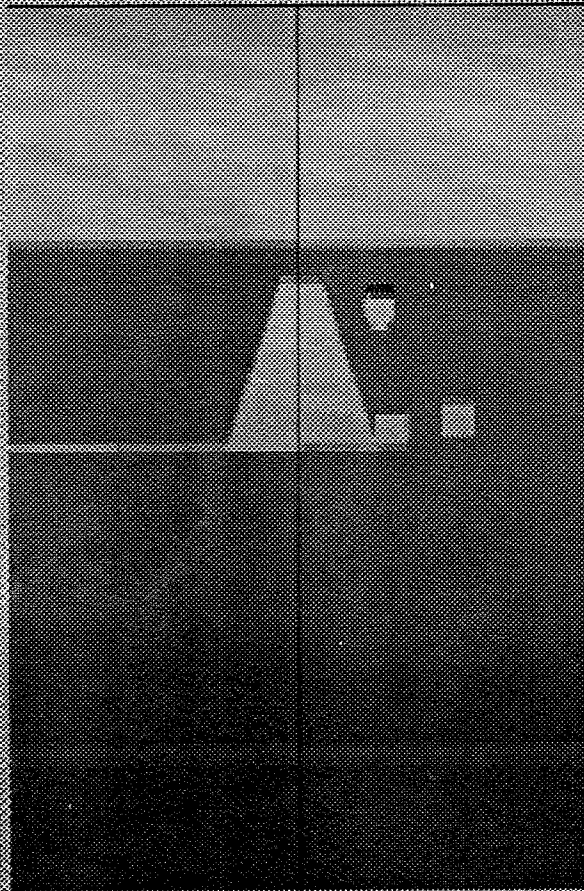
- **Four tasks running on IRIS 480 Reality Engine**
- **Each task (EXEC, Display, Host Interface, Phenomenology) is running on a separate CPU**
- **Update rate is 30 Hz for all field of view and resolution combination**

Snapshots

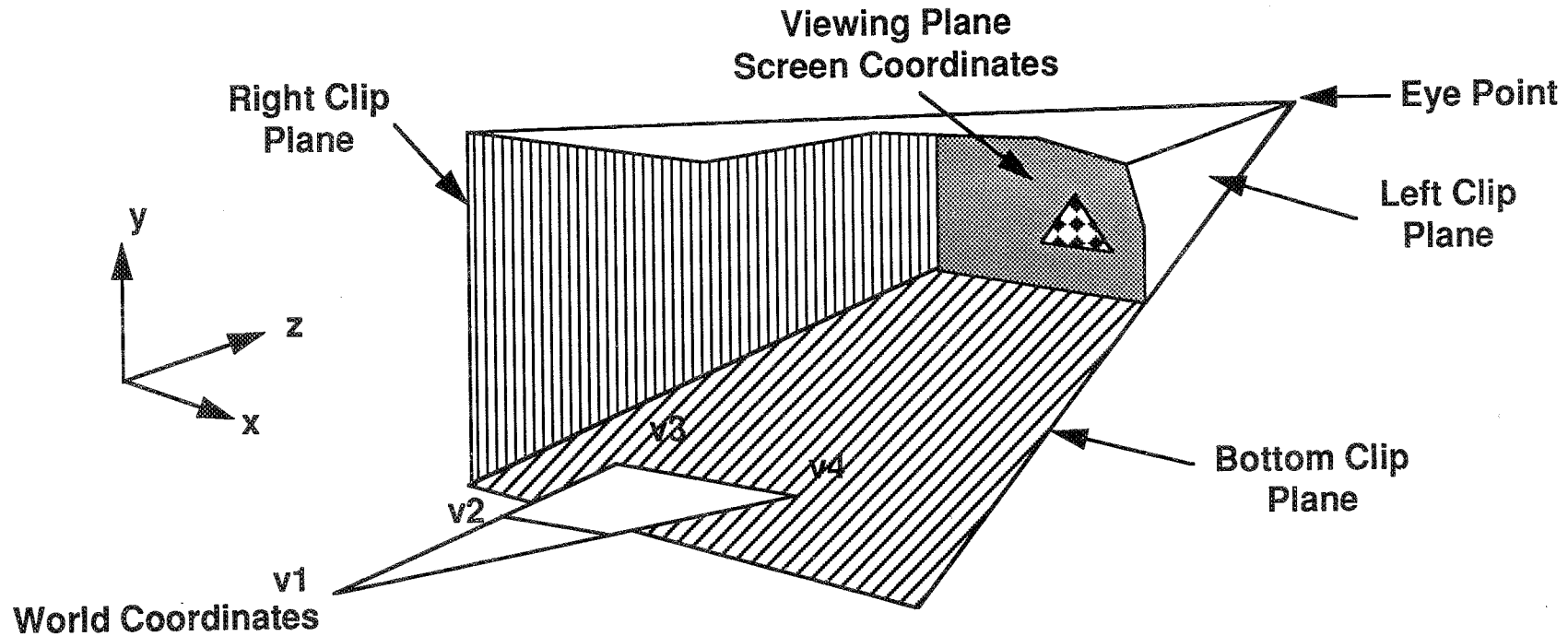


- Captured snapshot frames during real-time flight
- View snapshots using offline program
- Analyze Intensity Profiles and generate graphs
- Intensity Profiles were used for validation of model
- Captured sequence of images used by Dr. Kasturi at Penn State

Validation of Real-Time PMMW Model



Clipping Planes



- Intensities are computed by the phenomenology at vertices in world coordinates
- Gouraud shading performs a bilinear interpolation between vertices but it is done in Viewing Plane coordinates
- This is a good approximation as long as polygon is inside the field of view

Solutions to the Shading Problem

- **Clipping polygons and interpolating intensities in 3 dimensional world coordinates prior to rendering image with Gouroud shading**
- **Increasing the number of polygons in the data base**

Recent Embellishments and Future Plans

- **Revised Display task to use MultiGen data bases**
 - Ease of modifying data base
 - More control over sorting and culling of data base
 - Ease of control of moving objects
 - Future addition of Levels of Detail
 - **Investigate partitioning of Display task to relieve CPU bottleneck**
 - **Addition of Shadows**
-

Summary

- **Described the graphical techniques used in the PMMW Sensor Simulation**
 - **Described the interface with the dynamic phenomenology model**
 - **Discussed problems and solutions**
-

Engineering Workstation: Sensor Modeling

M. Pavel and B. Sweet
NASA Ames Research Center

ABSTRACT

The purpose of the engineering workstation is to provide an environment for rapid prototyping and evaluation of fusion and image processing algorithms. Ideally, the algorithms are designed to optimize the extraction of information that is useful to a pilot for all phases of flight operations. Successful design of effective fusion algorithms depends on the ability to characterize both the information available from the sensors and the information useful to a pilot.

The workstation is comprised of subsystems for simulation of sensor-generated images, image processing, image enhancement, and fusion algorithms. As such, the workstation can be used to implement and evaluate both short-term solutions and long-term solutions. The short-term solutions are being developed to enhance a pilot's situational awareness by providing information in addition to his direct vision. The long term solutions are aimed at the development of complete synthetic vision systems.

One of the important functions of the engineering workstation is to simulate the images that would be generated by the sensors. The simulation system is designed to use the graphics modeling and rendering capabilities of various workstations manufactured by Silicon Graphics Inc. The workstation simulates various aspects of the sensor-generated images arising from phenomenology of the sensors.

In addition, the workstation can be used to simulate a variety of impairments due to mechanical limitations of the sensor placement and due to the motion of the airplane.

Although the simulation is currently not performed in real-time, sequences of individual frames can be processed, stored, and recorded in a video format. In that way it is possible to examine the appearance of different dynamic sensor-generated and fused images.

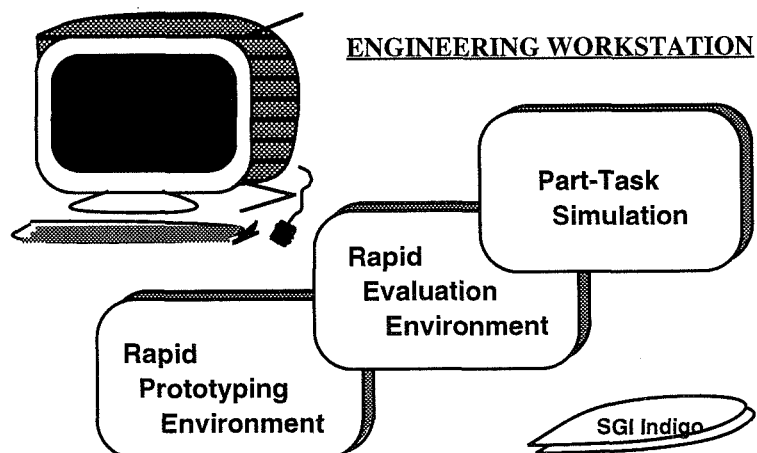
510-54
205740
p. 10
249923

GOALS

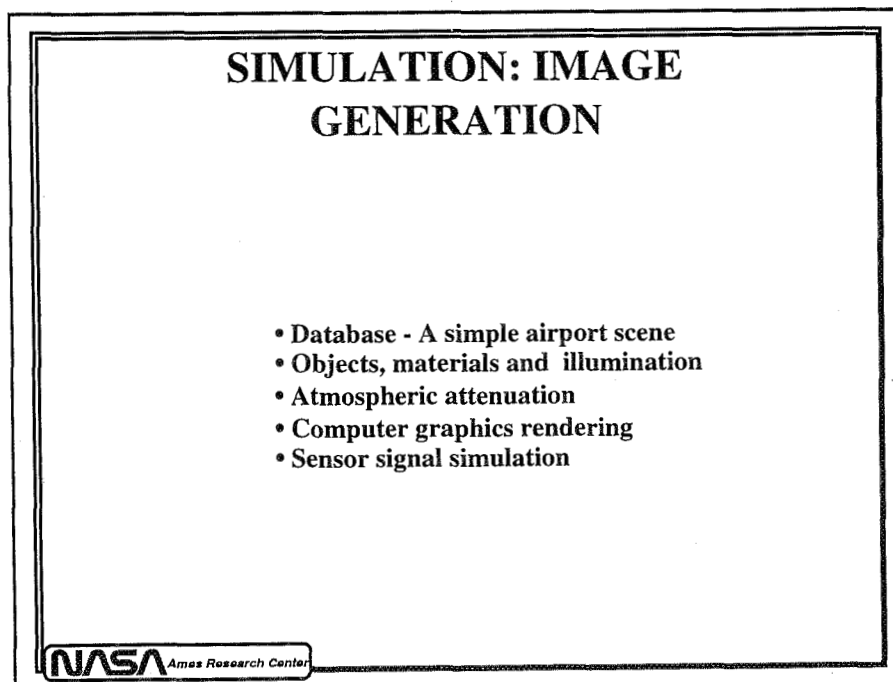
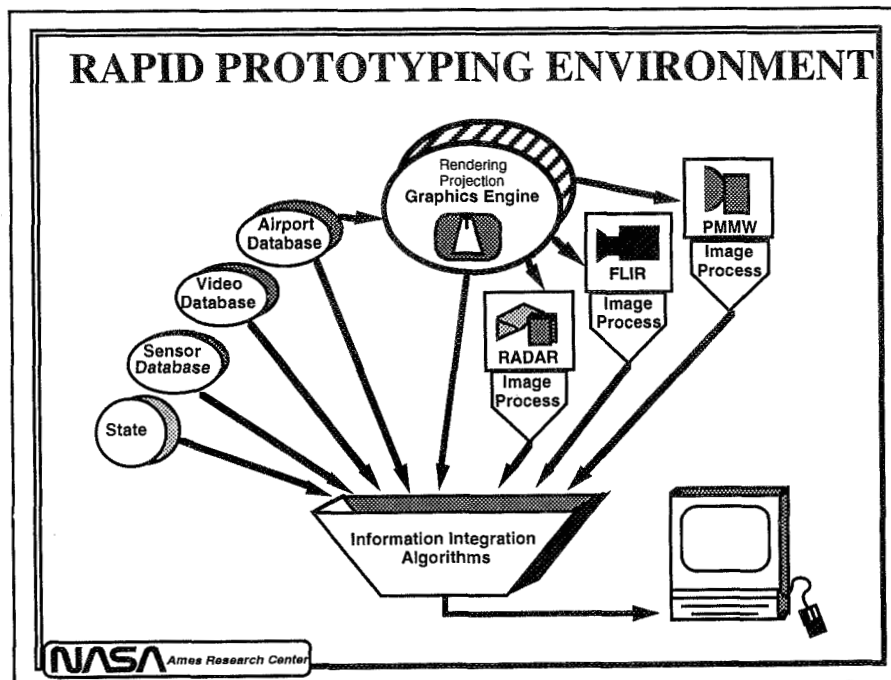
- Tools for rapid development and evaluation of augmented vision systems
- Development of short-term solutions
- Simulation of sensor signals
- Signal and image processing
- Simulation of algorithms
- Error analysis
- Easy-to-use interface

NASA Ames Research Center

SYSTEM DEVELOPMENT ENVIRONMENT



NASA Ames Research Center



VISUAL IMAGE

- Simple airport scene
- Polygonal representation
- Simple lighting model
- Color image rendering



SENSOR SIMULATION PHILOSOPHY

- Goal: Reduce simulation complexity
- Simulate critical characteristics
- Restricted viewing conditions
- Restricted environmental conditions
- Material specification -> Signal

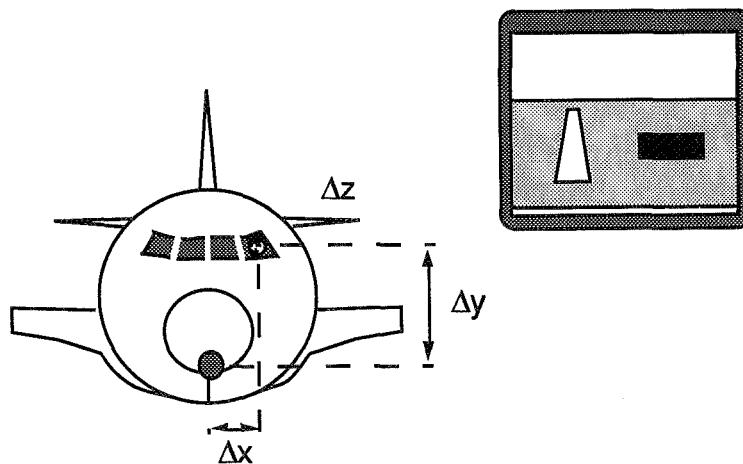


SENSOR CHARACTERIZATION

- Relationship between a visual and a sensor image
- Spatial response characteristics
- Temporal response characteristics
- Sensitivity and signal-to-noise ratio
- Stability: drift, changes in gain
- Atmospheric effects and attenuation
- Inhomogeneity of sensor image

NASA Ames Research Center

GEOMETRIC DISTORTIONS



NASA Ames Research Center

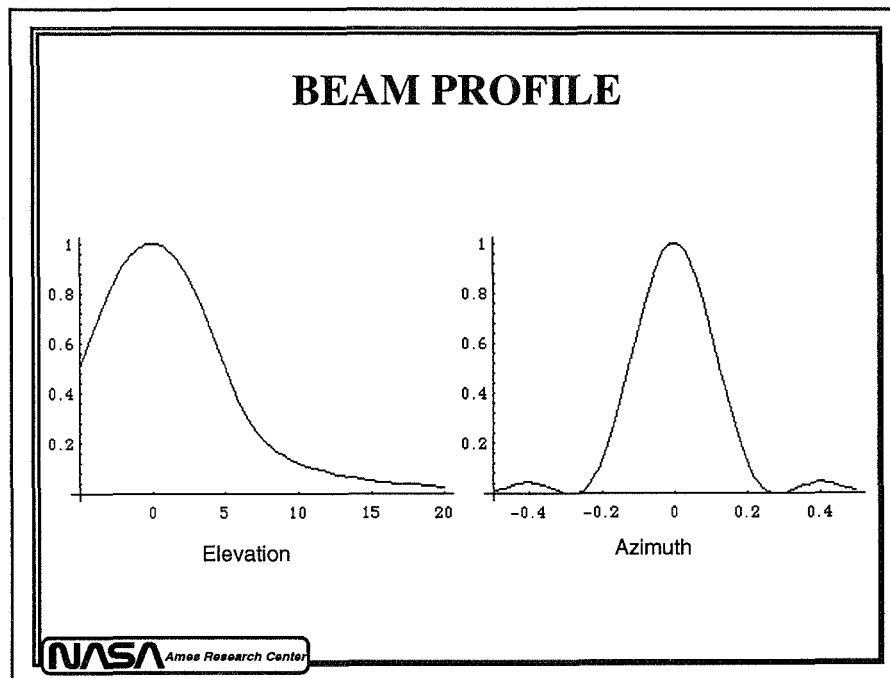


IMAGE PROCESSING

- HIPS Image Processing System
- Image Processing
- Special Algorithms
- Fusion

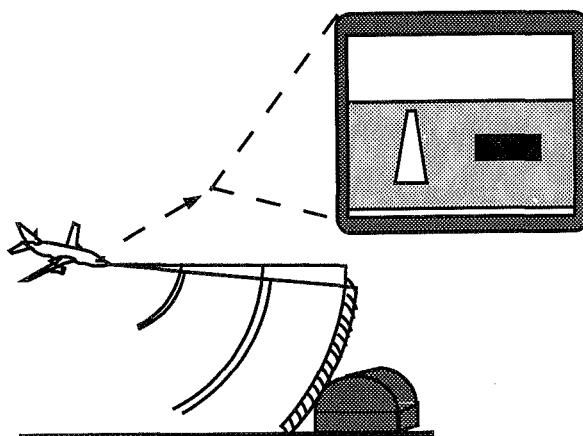
NASA Ames Research Center

RADAR SIMULATION

- Assignment of material or radar cross section (RCS)
- Computer generated image - Rendering
- Beam profile calculations
- Compute Range using Hardware Z-buffer
- Scattering variability
- Gain control

NASA Ames Research Center

RADAR SIMULATION



NASA Ames Research Center

Passive Millimeter Wave (PMMW) SENSOR CHARACTERISTICS

- The following are examples of particular implementations of selected sensor models
- 16 x 16 Focal plane array
- Operating Frequency: 94 GHz
- Spatial Resolution: 6 Milliradians (1/3 degree)
- Minimum Resolvable Temperature: 1 Deg K
- Update rate: 10 Hz
- Noise Figure



Passive Millimeter Wave (PMMW) SENSOR SIMULATION

Assumptions

- Uniform hot sky
- Runway, grass -> reflectivity specification
- Spatial modulation transfer function (MTF)
- Gaussian noise

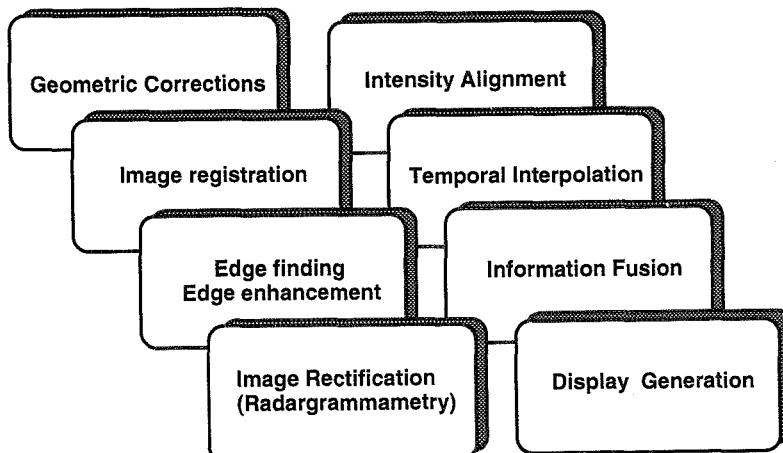


HIPS IMAGE PROCESSING SOFTWARE

- Modular, UNIX-based system
- Modifyable source code
- Self-documenting image files
- Built-in functions:
 - Filtering, edge-detection
 - Image transformations
 - Image statistics
 - Image compression

NASA Ames Research Center

EXAMPLES OF ALGORITHMS



NASA Ames Research Center

USER INTERFACE

- **Generate sequences of frames**
- **Menu-based interactions**
 - Stop, examine a frame
 - Generate fog
 - Render PMMW image
 - Render radar image
 - Modify parameters
- **Save images in HIPS Format**



ONT

III. SENSOR FUSION

IMAGE FUSION

M. Pavel¹

Western Aerospace Laboratories
and
NASA Ames Research Center

1 Motivation

In order for a pilot to fly an airplane, she or he must combine information from a large number of different sources. Useful information for this purpose may be available as readouts from avionics instruments, symbology on a HUD, or from the image of an airport scene seen through a window. The workload of the pilot is frequently increased as the number of sources of information and the complexity of the data increases. Because humans do not necessarily combine information optimally, effective automatic combination of the data may lower the load and thereby free the pilot to be ready if necessary to make critical decisions. The combined data are frequently more useful because the combination may reduce variability, or use complementary information from the different sources.

It is interesting to note that fusion of information is a common process in both natural and machine vision. Consider these examples of fusion:

1. Combining images obtained from different locations, e.g., binocular stereopsis.
2. Combining images obtained from different sources — flight instruments and an image of a scene.
3. Combining information from one source over time, i.e., temporal filtering.
4. Combining information from one source over space, i.e., spatial filtering.

¹This work was supported in part by a grant NCC 2-486 from NASA to the Western Aerospace Laboratories

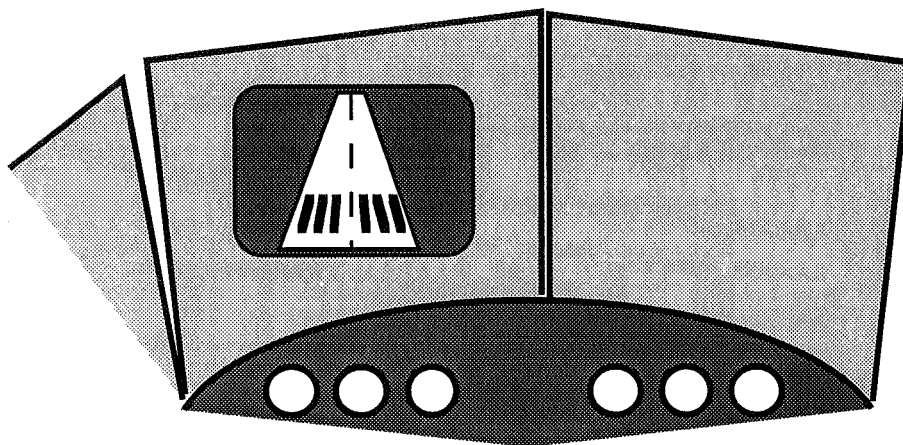


Figure 1: Schematic representation of the HUD arrangement.

These considerations are among those motivating the development of systems that augment the traditional display system. One approach, schematically depicted in Figure 1, illustrates one possible implementation of the AVID system.

2 System Overview

Figure 2 illustrates the basic components of a system designed to improve the ability of a pilot to fly through low-visibility conditions such as fog.

The underlying principle is based on the fact that atmospheric attenuation is greatly reduced for millimeter waves (MMW) relative to the radiation in the visible spectrum. In the proposed system the information (images) from sensors operating in the MMW regime are combined with other information such as a global positioning system (GPS) and a stored database. The fusion process is necessary because the spatial and temporal resolution of the MMW sensors is greatly limited.

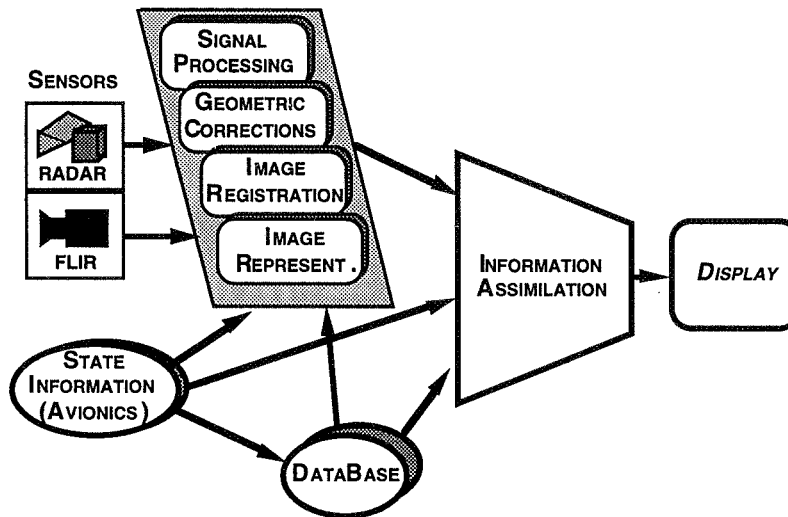


Figure 2: Diagram of ESAS.

2.1 Role of Visual Sciences

A successful design of a system such as the one illustrated in Figure 2 requires a combination of expertise ranging from radar engineering to human factors and psychology.

Life sciences are critical for the development and design of such a system in at least three ways. First, knowledge of the visual system must be used to optimize the design of displays used by the pilot in all phases of flight operations. Second, understanding the human visual information processing can guide the development of solutions to many system design problems. For example, biological fusion may be used in the process of reverse engineering to guide the design of fusion algorithms. Finally, psychology of measurement, combined with the models of the visual system, can be used to develop methodology for evaluation of the complete system.

It is also important to note that the solution of the particular problems associated with AVID gives rise to questions whose answers will enhance our basic understanding of the human visual system. For example, displaying information *on* a HUD without impairing significantly the information viewed *through* the HUD requires a good understanding of perception of transparent images. Although recent results[2] provide useful information for the

designer, additional basic research is required to develop a model of transparency perception.

2.2 Fusion Issues

The first prerequisite for a successful design and evaluation of fusion algorithms is a definition of a goal specified in terms of desired images and an objective function. The ultimate desired image is one that contains all necessary information for flight control. To achieve (or to approximate) this goal requires a convenient representation of data, optimal fusion algorithms, and a effective display of the resulting images. System evaluation can be performed by comparing the obtain image to the desired one with respect to the objective function.

Unfortunately, our knowledge to date is not sufficiently complete to specify a unique desired image and an objective function. Rather, we define a gray-level image $s(x,t)$ to be an image that would be obtained under uniform illumination with unlimited visibility. Using simulator test results, one can easily demonstrate that this image is sufficient, but not necessary, for a pilot to land an airplane.

3 Sources of Information

There are many sources of information that could be used to support the functions of the enhanced situational awareness. For the purpose of this project, we consider the following sources of information:

- High resolution sensors of visible spectrum (Video)
- High resolution sensors of infrared spectrum (IR)
- Low resolution millimeter wave sensors (Radar, PMMW)
- Terrain database
- Inertial navigation system (INS)
- Global positioning system (GPS)

3.1 Sensor Characterization

Effective fusion of information from different sources requires the comprehensive characterization of the sources. The following is a list of sensor characteristics that are important in the design of image processing and fusion algorithms.

3.1.1 Signal Characteristics

These characteristics describe the properties of the signals generated by the sensor:

- Spatial and temporal transfer functions
- Sensitivity
- Relationship between visual and sensor images
- Noise, drift, changes in gain
- Atmospheric attenuation
- Temporal sampling / dynamics
- Inhomogeneity of sensor image

3.1.2 Geometric Properties

Knowledge of the imaging geometry of the sensor is critical in order to generate conformal images from different sources. In addition to the imaging geometry of each sensor, its location and orientation is also critical. These effects are illustrated in Figure 3. Geometric corrections to compensate for the variety of geometric distortions can be implemented, for most sensors, by simple transformations. One notable exception is an active radar which requires special considerations.

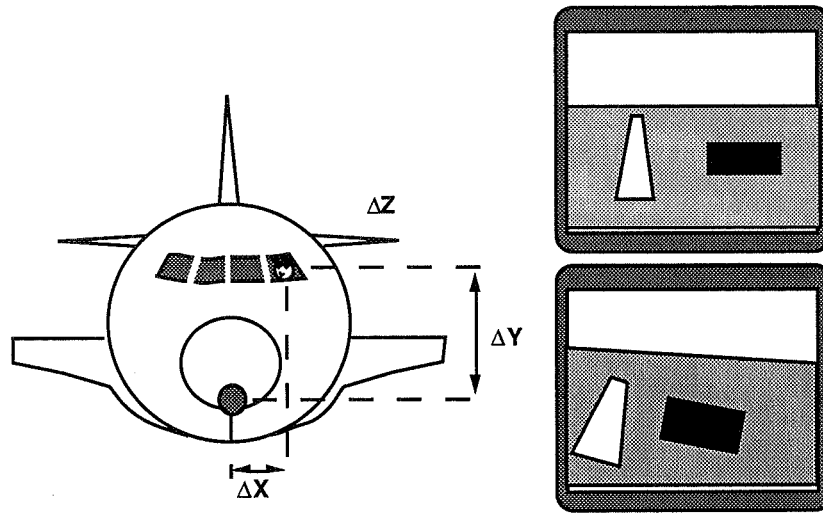


Figure 3: Diagram of geometric distortions due to sensor viewpoint placement

3.1.3 Imaging Radar Distortions

Radar is an active device that illuminates a scene, detects reflections, estimates delays associated with the reflections, and thereby estimates the distances of the reflecting objects. Since a radar measures ranges (b-scope representation), a geometric transformation is necessary to convert the range image to a perspective projection of the scene (c-scope image). As shown in Figure 4, this transformation is, unfortunately, underconstrained because measured distances do not specify position uniquely.

A typical solution, used to regularize this problem, is to assume that all reflections are from objects located on the surface of flat earth. Of course the flat-earth assumption results in errors whenever the actual reflections are generated by objects at some vertical distance from the earth surface (Figure 4).

Recently we have been able to demonstrate a theoretical approach to reduce the problem by eliminating the flat earth assumption. The computational method is based on integrating information from multiple frames of b-scope images. We are currently examining the practical implications of these theoretical efforts.

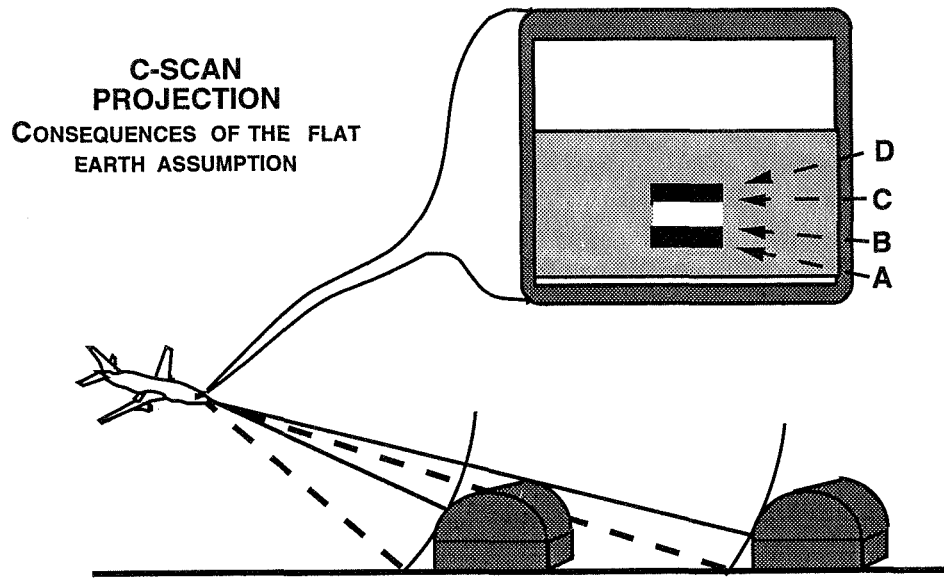


Figure 4: An illustration of the effects of flat-earth assumption in the rectification of returns from two elevated structures.

3.2 Simplified Sensor Model

Under the assumption that it is possible to correct all geometric distortions in images obtained from a sensor, the output of the sensor can be approximated by

$$m(\vec{x}) = h * \{a[r(\vec{x})] b(\vec{x}) s(\vec{x}) + n_m(\vec{x})\} \quad (1)$$

where m is the sensor image

$\vec{x}$ image coordinates

h spatial impulse response

a atmospheric attenuation

r range (distance) from sensor to an object

b sensor-to-visual factor

s objective image

n_m noise

3.3 Database

The database (DB) consists of the best available information (model) of the landing terrain. The database includes the airport, the runway, and some surrounding stationary objects. The models of the objects are represented in terms of polygons. The geometric model of the terrain includes color information and it is rendered by the geometry engine of a graphics workstation, such as the Silicon Graphics Inc. (SGI) machine.

When the rendered scene is converted to a gray-level representation of the landing scenario, the resulting image can be approximated by:

$$d(\vec{x}) = [1 - c(\vec{x})] s(\vec{x}) + c(\vec{x}) g(\vec{x}) + n_d(\vec{x}) \quad (2)$$

where

d computer generated image obtained from the DB

c obstacle indicator function

s objective image

g obstacle image

n_d] noise, quantification of DB inaccuracy.

In this simple model, the difference between a real image of the scene and the DB rendering is expressed by the noise term in equation (2).

4 Image Processing

Prior to fusion, information from each sensor is processed by algorithms specialized for that sensor. These algorithms are designed for:

1. Noise reduction: Linear and non-linear filtering
2. Image enhancement: Histogram equalization, edge enhancement.
3. Uncertainty (Noise) Estimation: Estimation of variability and consistency within and across sources.
4. Prediction: Recursive estimation of expected and observed image.

5 Image Fusion

There are many ways to combine information from different sources. The optimal technique to be selected depends on prior knowledge of the signal characteristics, the objective, and the required robustness. The following is a list of examples of candidate techniques:

1. Additive, linear combination
2. Selection (1/0)
3. Additive, nonlinear combination
4. Bayesian update of information

I will first discuss briefly the first two techniques which have been considered by several investigators [1, 3].

5.1 Linear Additive Combination

Linear additive rule is a pixel by pixel combination of two sources that can be expressed by

$$\langle s(\vec{x}) \rangle = \alpha d(\vec{x}) + \beta m(\vec{x}).$$

There are several reasons why a linear additive combination is particularly important. First, additive combination is an optimal rule when the individual sources can be characterized by normal distributions. Second, additive combination is easily implemented in real-time hardware. Finally, additive combination occurs naturally when an image is displayed on a HUD.

5.2 Disadvantages of Additive Fusion

There are several shortcomings of the simple linear additive approach:

Obstacle Detection: Whenever information is present in one, but not in the other image, the fused signal-to-noise ratio is lower than that in the original image with the signal.

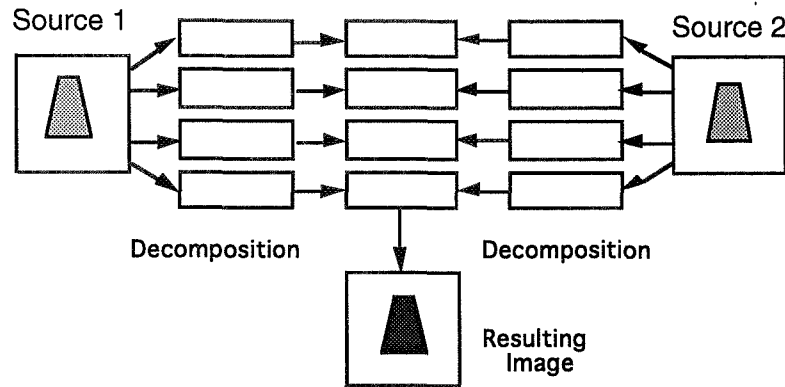


Figure 5: A diagram of fusion by components.

Polarity Changes: The relationship between the polarity of two images may vary for different locations and may depend on environmental conditions.

Spatial Frequency: Signal-to-noise ratio may vary for different spatial frequency bands and different spatial locations.

Because of these shortcomings of the linear additive rule, we consider more complex, nonlinear rules.

5.3 Fusion by Components

One approach that can be used to remedy the disadvantages of the linear additive rule is to decompose each image into components and then perform the combination by combination rules specific to the components. This general approach is shown in Figure 5.

Depending on the specific application, there are numerous ways of decomposing images into components. Multiresolution representation of images is one way of decomposing images into its components.

5.4 Multiresolution Representation

A typical multiresolution representation can be thought of as a decomposition of an image into a set of spatial frequency bands as illustrated in Figure 6.

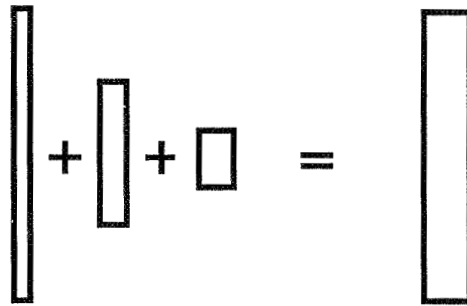


Figure 6: Illustration of a pyramidal representation.

The size of the blocks in the diagram in Figure 6 indicates that the lower spatial resolution bands require fewer samples.

One way to construct such representation consists of recursive applications of the following steps:

1. low-pass filter,
2. subsample,
3. interpolate,
4. compute difference between two adjacent levels, until the representation reduces to a single sample.

In this particular multiresolution representation, each resolution level is insensitive to local orientation of features. There are other schemas for the decomposition such that the information at each resolution level is further decomposed to several subimages, one for each of a set of directions [1, 4].

Given the multiresolution representation, there are many alternative ways to fuse the images.

5.5 Sample Selection

One way to fuse two images consists of examining each pixel in both images at each level, and selecting the pixel with a particular property. For example, one can select the pixel with the greater gray level value [1]. Alternatively, it is possible to compute contrast at each level and select the pixel with

greater contrast value [3]. Although these methods have been shown to be successful they do not eliminate all the problems listed in Section 5.2. We are, therefore, considering a more general, statistical approach to fusion.

5.6 Optimal Fusion Approach

The goal of the optimal fusion approach is to use the best models of the sources together with the desired image and determine the combination that minimizes the difference between the fused and the desired images. Although there are questions concerning the particular metric to be used for the measurement of the difference, our initial development is based on maximizing aposteriory probability.

This approach requires either prior knowledge or on-line estimation of the variability of the sensor images. Limited spatial resolution and the physical phenomena underlying some sensors, e.g., MMW radar, results in spatial correlation that can be utilized in fusion.

Our current approach consists of the following steps:

1. Compute multiresolution pyramid for each image.
2. Predict image from the database.
3. Predict image from prior frames.
4. Estimate the variances at each pixel $\vec{x}$ at each level l .
5. Estimate correlation with the expected image from the database.
6. Combine pixels using optimal weights for each pixel and each level.

To the extent that the underlying assumptions are valid, this approach determines statistically optimal fused images. In addition, this statistically-based approach can be used directly to identify specific features of interest, for example, unexpected obstacles or runway incursions.

References

- [1] P. Burt. A gradient pyramid basis for pattern-selective image fusion. In *Proceedings of the Society for Information Display*, pages 467–470. SPIE, 1992.

- [2] J. B. Mulligan. Nonlinear combination rules and the perception of motion transparency. *Vision Research*, 33:2021–2030, 1993.
- [3] A. Toet. Hierarchical image fusion. *Machine Vision and Applications*, 3:1–11, 1990.
- [4] A. B. Watson. Efficiency of a human image code. *Journal of the Optical Society of America A*, 4:2401–2417, 1987.

Enhanced Image Capture Through Fusion

Peter J. Burt
Keith Hanna
Raymond J. Kolczynski
David Sarnoff Research Center

249928

S/2-35

205742

P-18

ABSTRACT

Image fusion may be used to combine images from different sensors, such as IR and visible cameras, to obtain a single composite with extended information content. Fusion may also be used to combine multiple images from a given sensor to form a composite image in which information of interest is enhanced.

We present a general method for performing image fusion and show that this method is effective for diverse fusion applications. We suggest that fusion may provide a powerful tool for enhanced image capture with broad utility in image processing and computer vision.

The Fusion Task

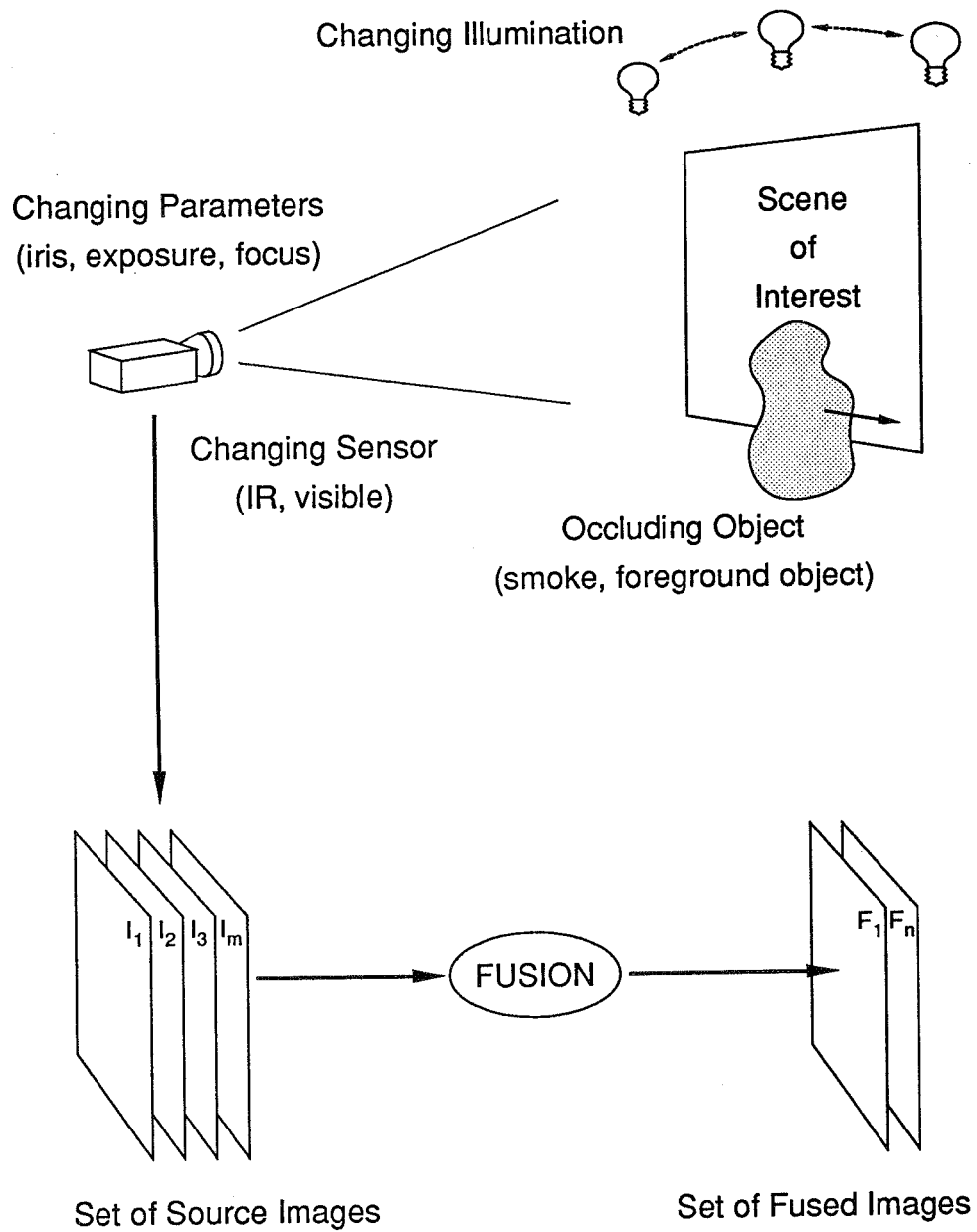
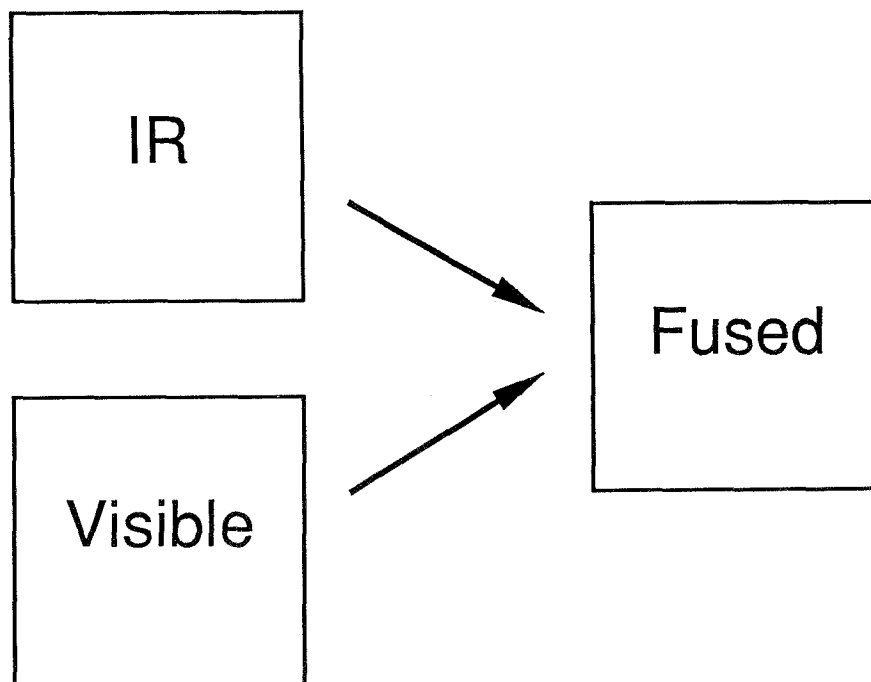


Image Fusion: Objectives

Combine two or more source images to obtain a single composite with extended information content.

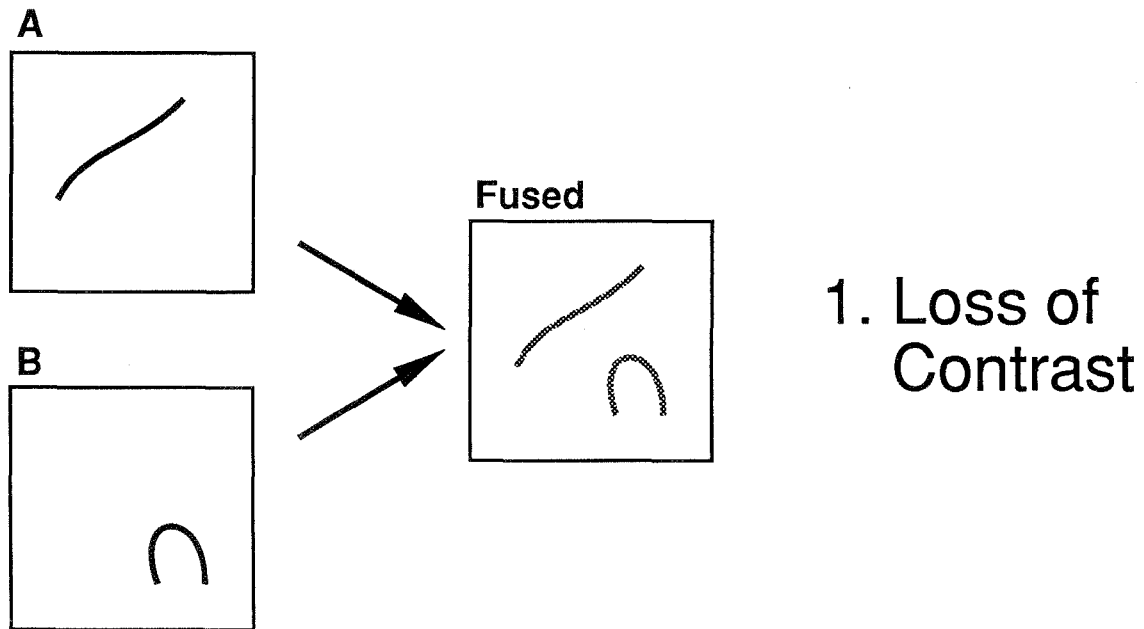


Requirements

- retain all useful information from the source images
- not introduce fusion artifacts into the combined image
- look "natural"

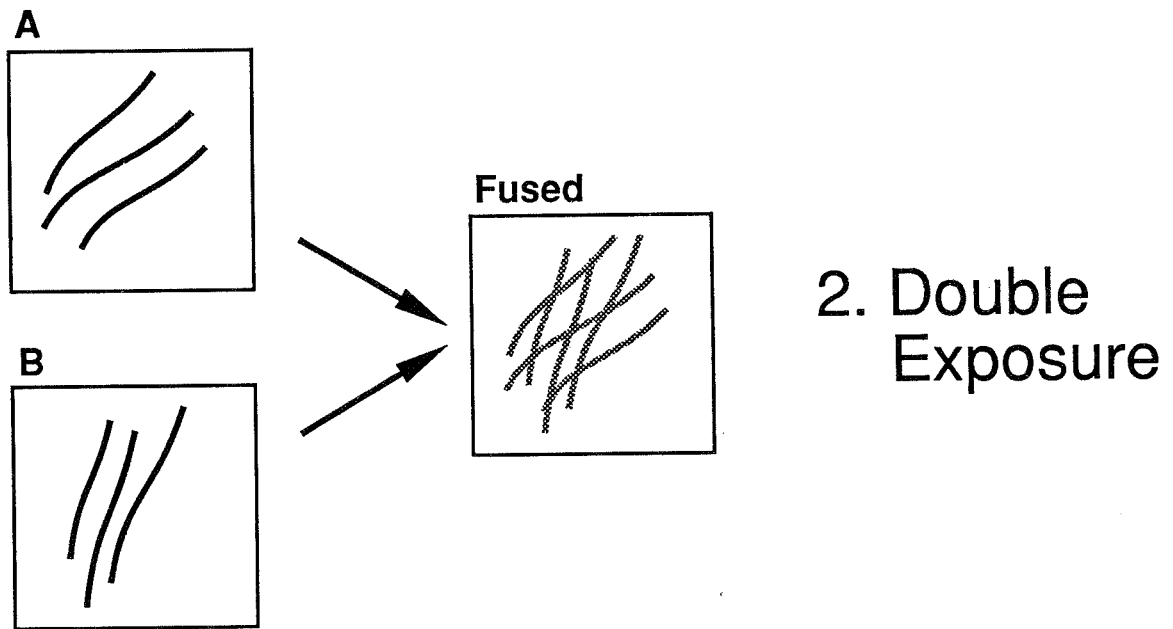
Technical Challenge

Pixel averaging, results in ...

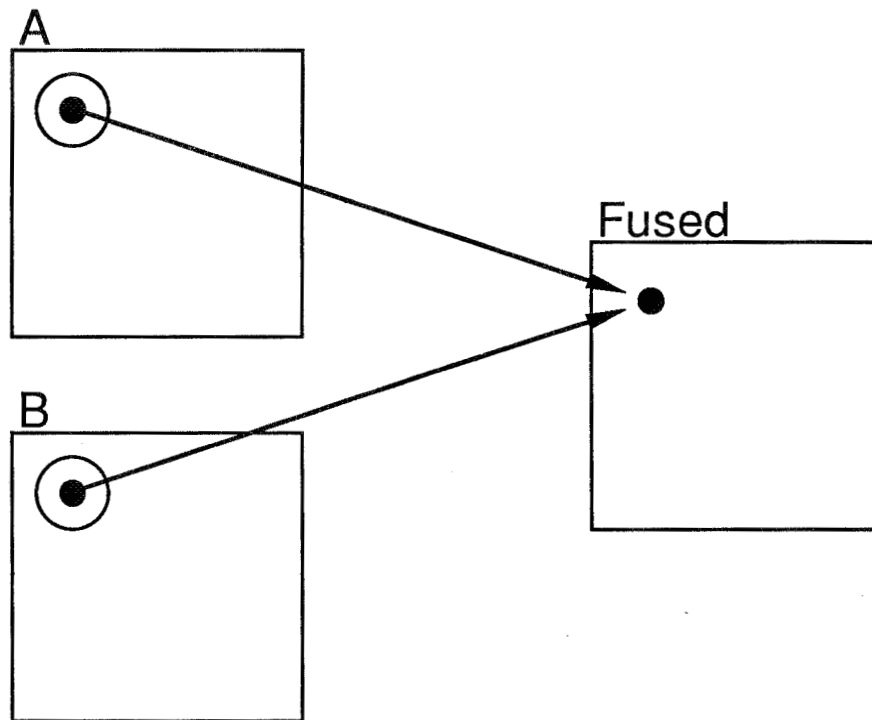


Technical Challenge

Pixel averaging, results in ...

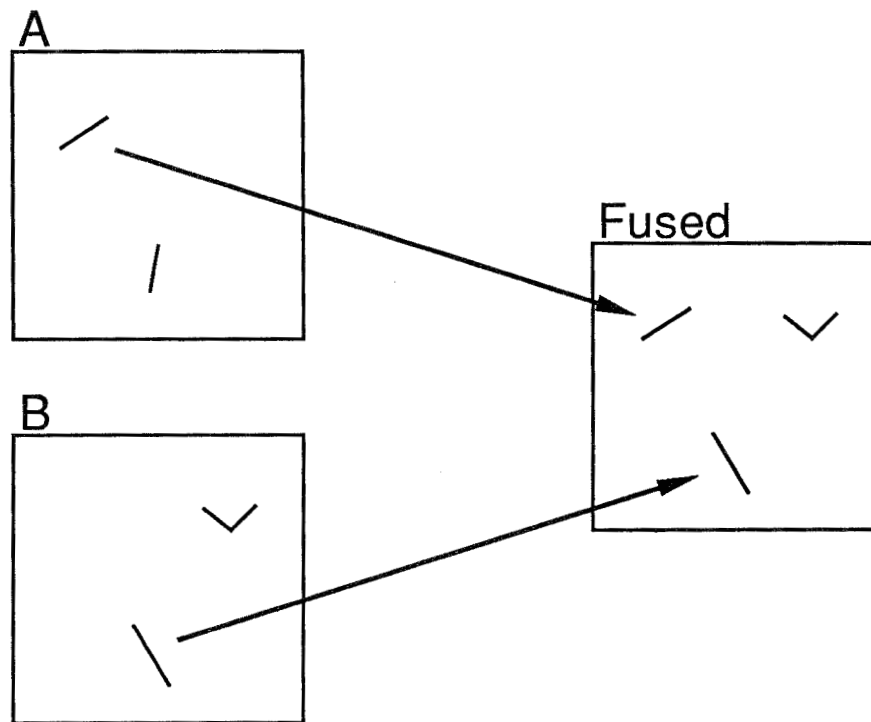


Pixel-Based Approach



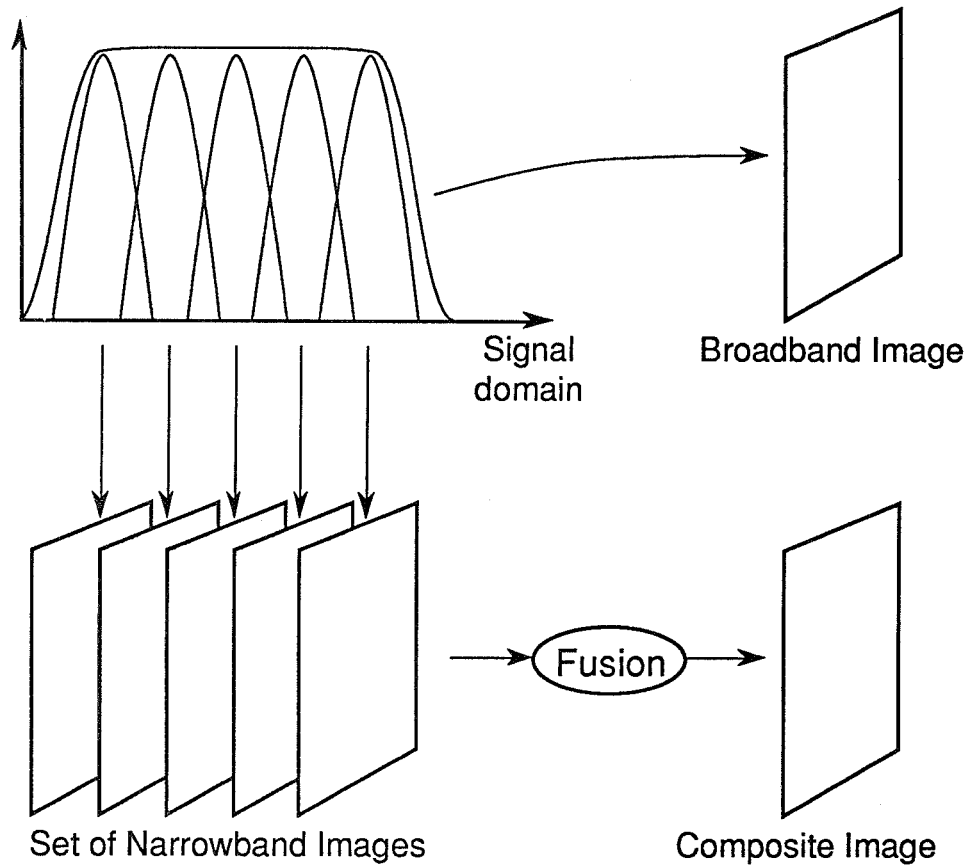
- each output pixel is computed separately
- based on the corresponding source image pixels
- or neighborhoods of corresponding pixels

Pattern-Selective Approach



- copy a pattern at a time
- select most salient patterns only

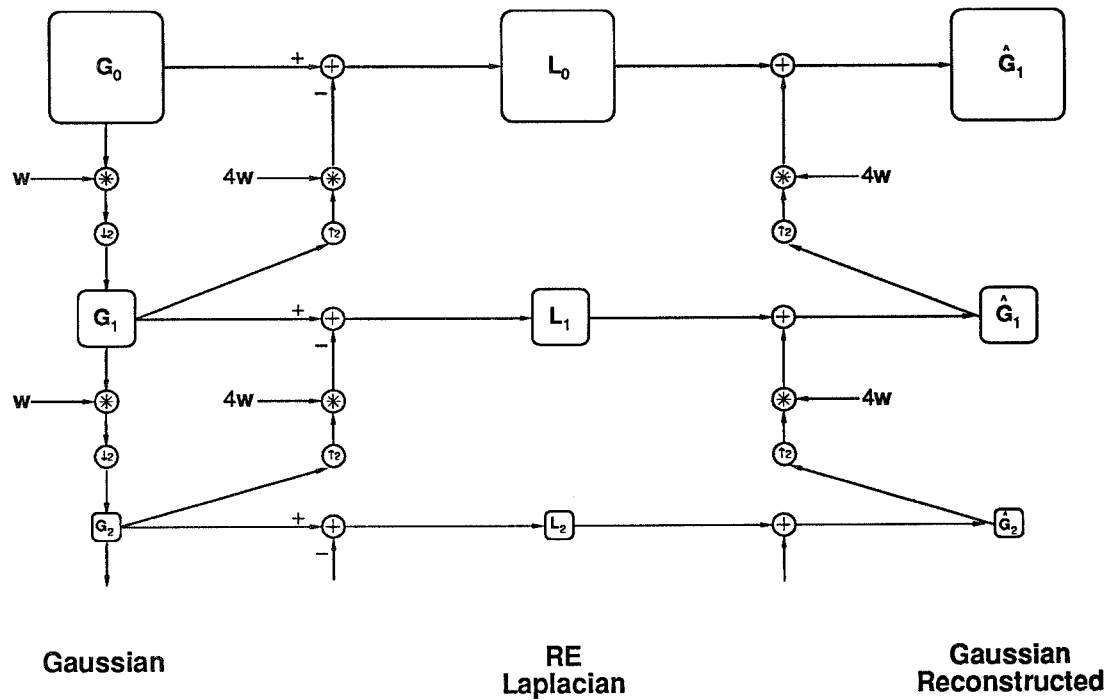
Composite Imaging



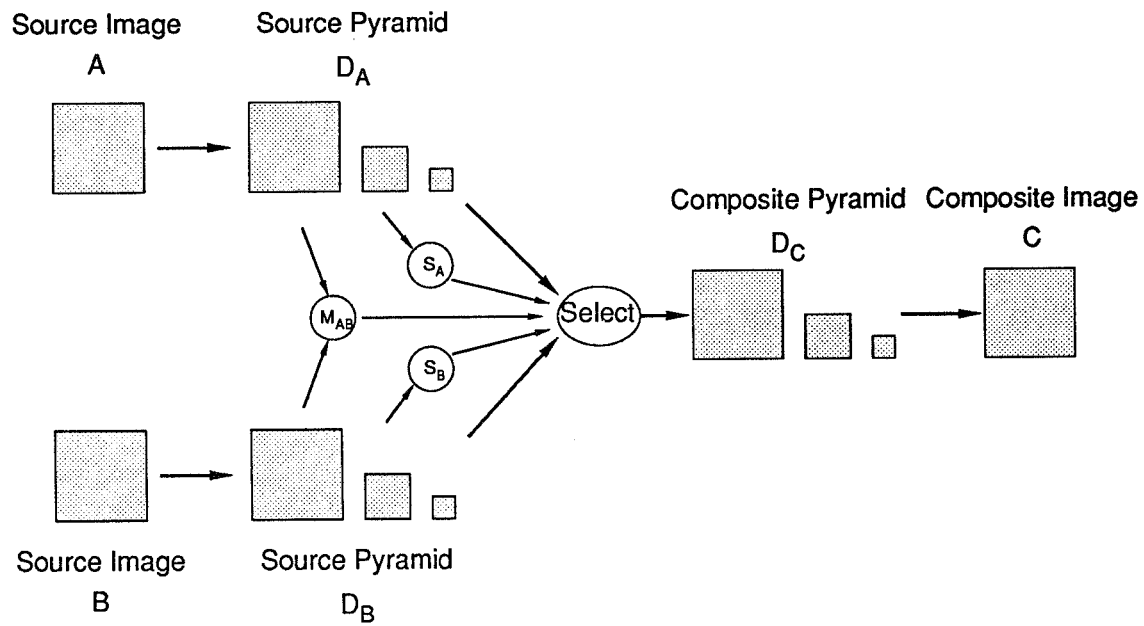
Pyramid-Based Fusion: Some History...

1983	Burt	model of human binocular vision
1984	Adelson	multi-focus
1990	Toet	IR and visual images
1991	Pavel,..	noise model
1992	Tinkler	TI method

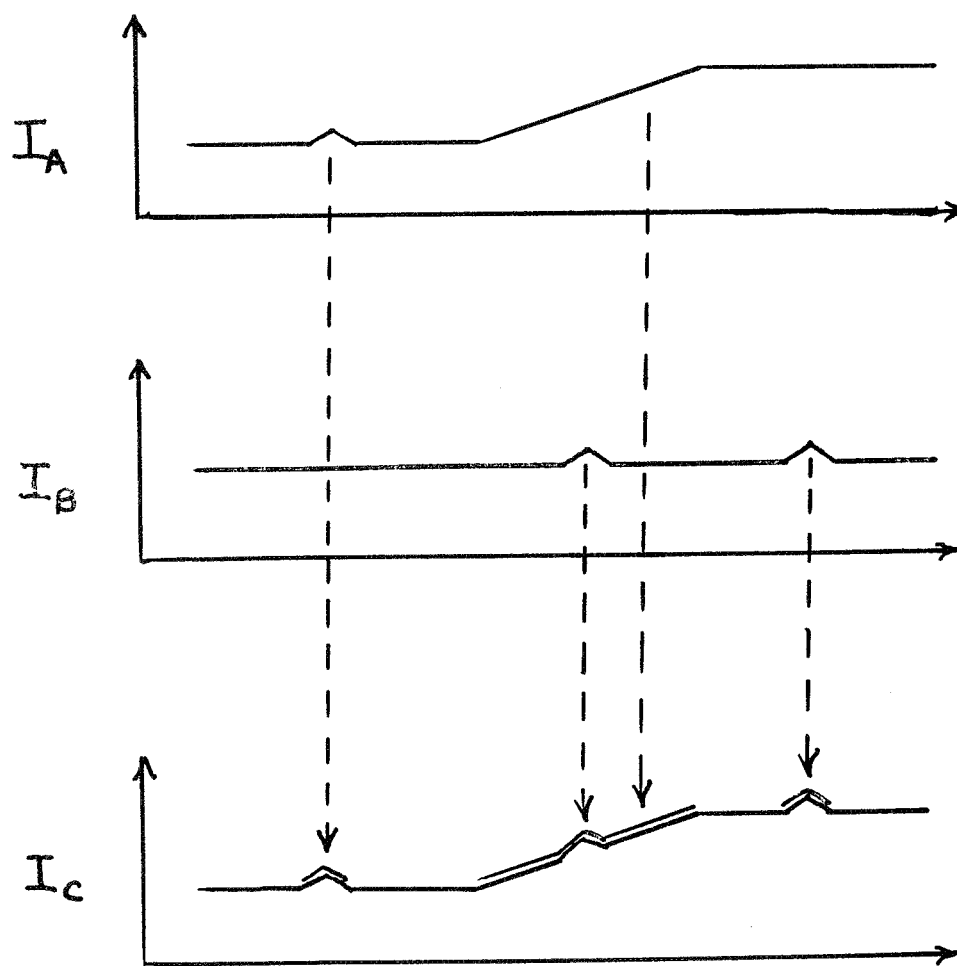
Laplacian Pyramid Transform



Gradient Pyramid Framework for Image Fusion



Composition Based on Scaled Relative Gradients



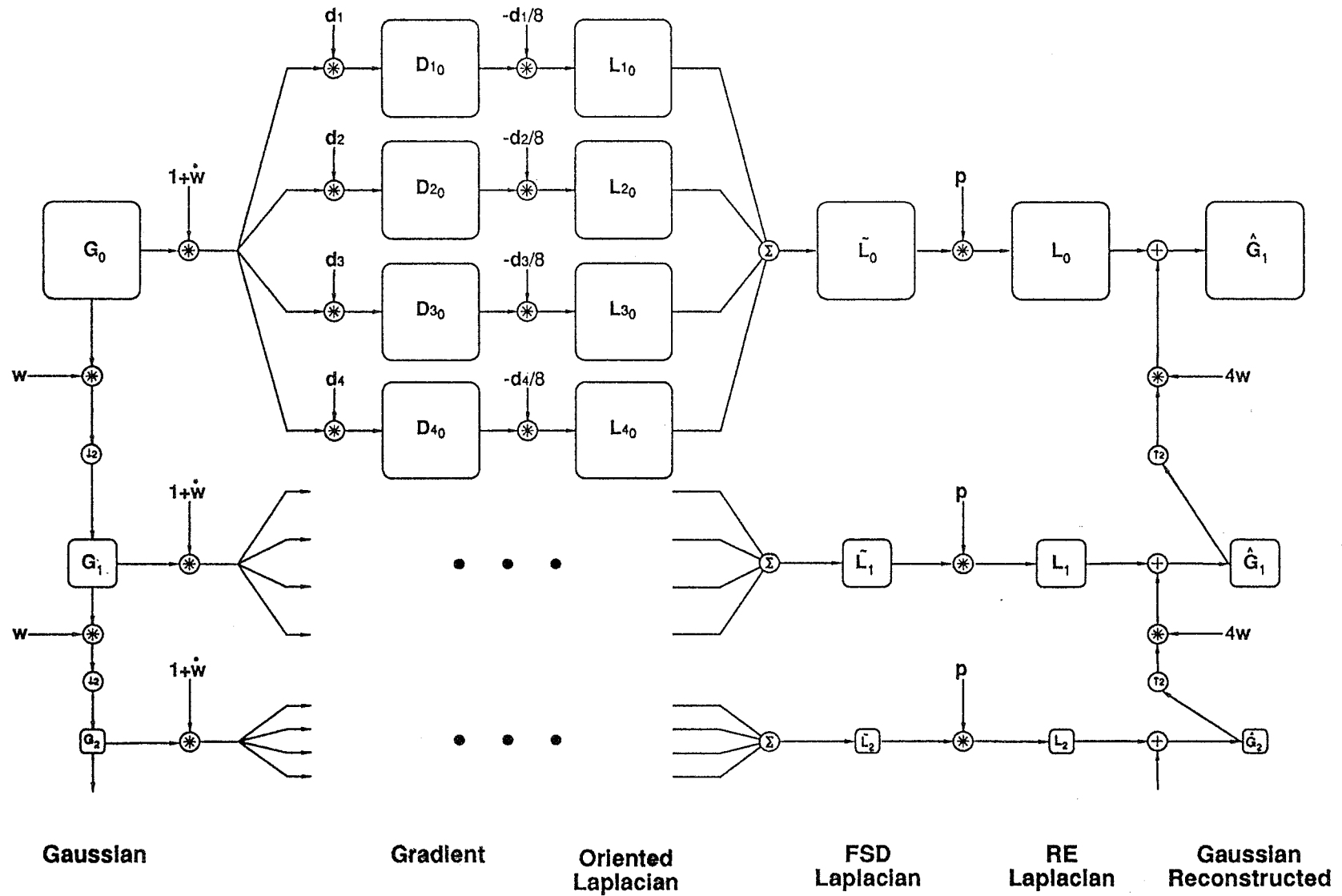
Scaled Gradient Operators

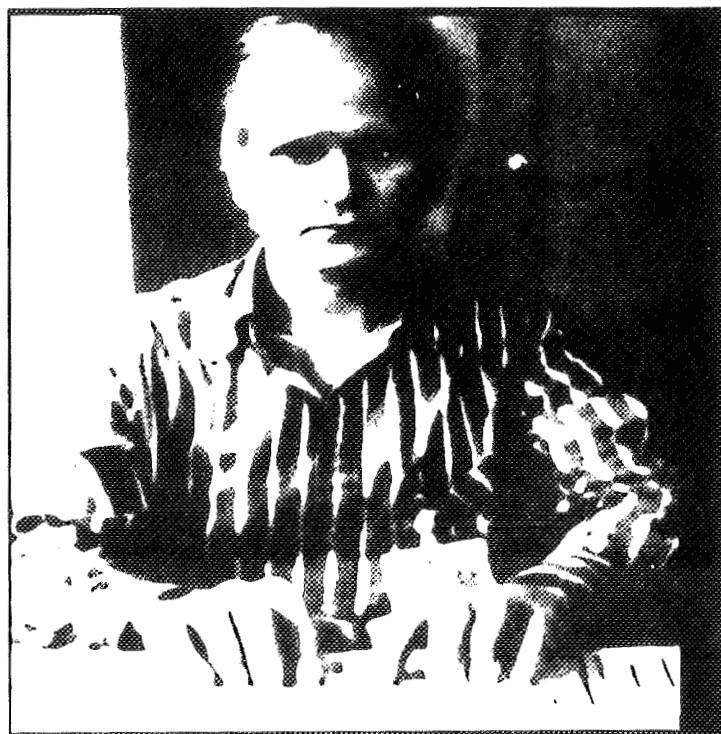
$G(x, \sigma)$

$\frac{d}{dx} G$

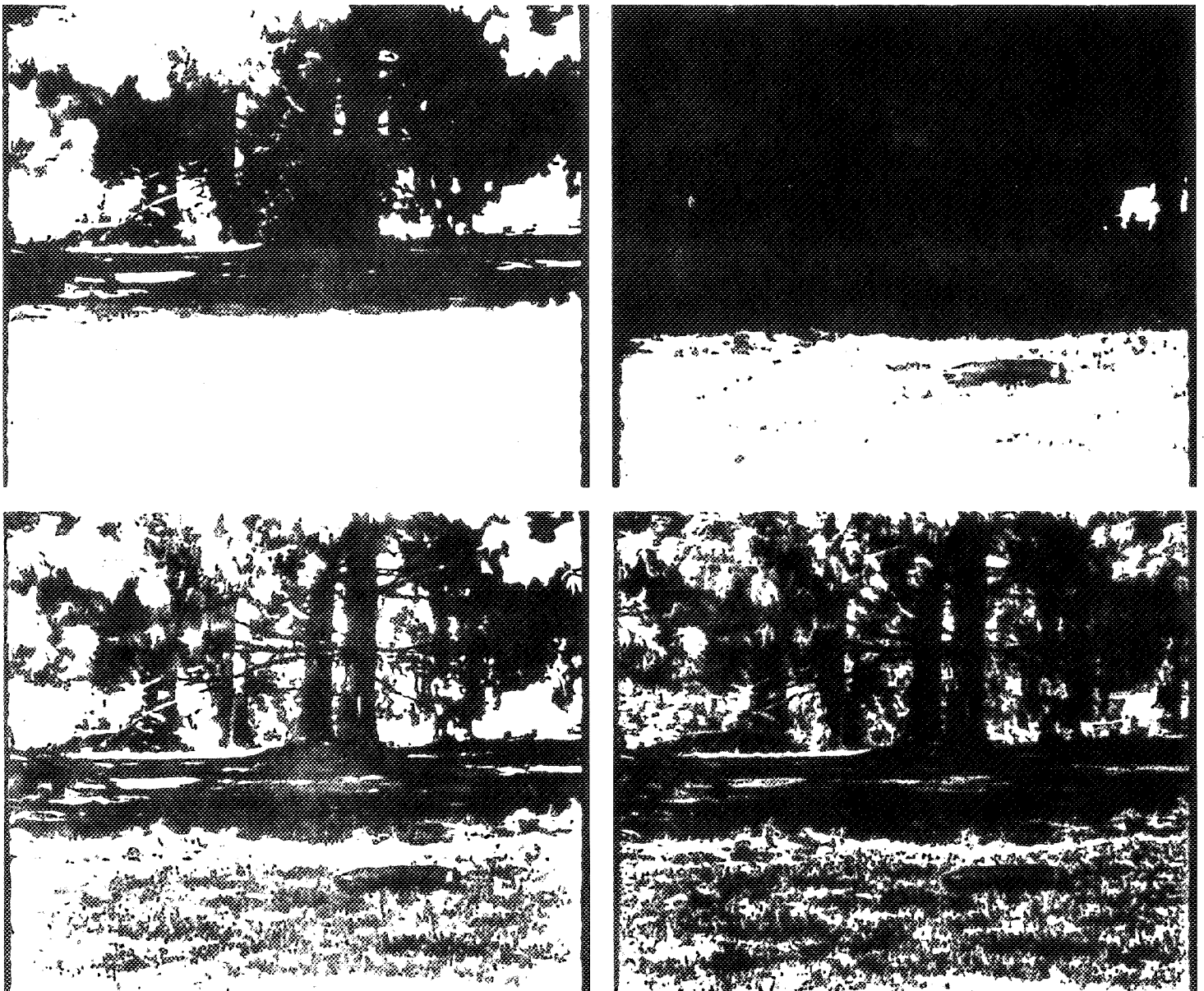


Gradient Pyramid Transform

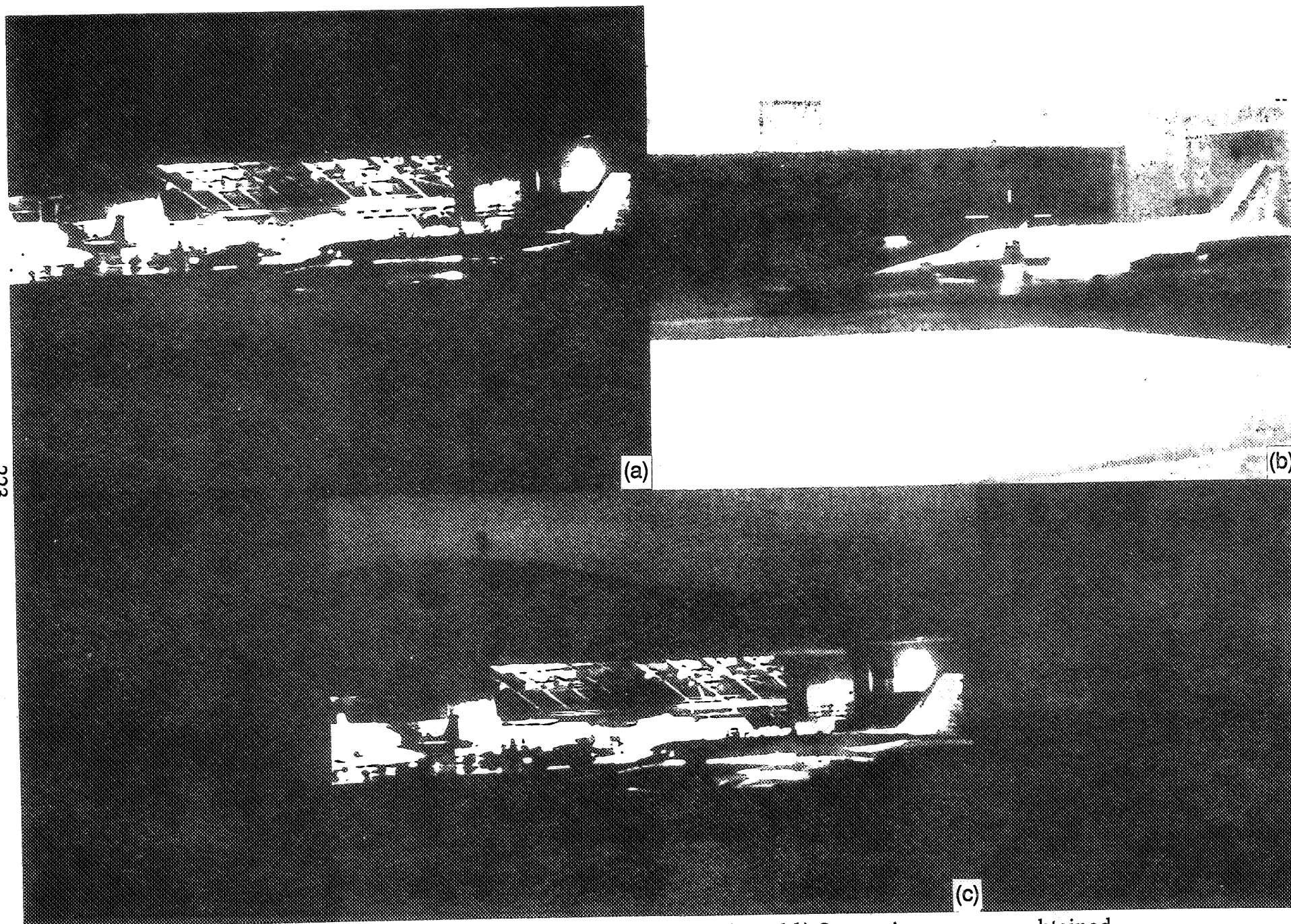




Multi-focus example of gradient pyramid fusion. (*a* and *b*) Source images obtained with a camera lens set to focus at different distances. (*c*) The fused image has an extended depth of field.



Multi-exposure example of gradient pyramid fusion. (*a* and *b*) Source images obtained with different camera exposure settings to observe patterns in shaded regions (*a*) and bright, sun-lit regions (*b*). (*c*) The fused image includes detail from both regions. (*d*) Pyramid samples values are normalized and quantized to just 4 bits to demonstrate that a broad dynamic range scene can be represented by a narrow dynamic range signal without loss of critical detail.



A multi-sensor example of gradient pyramid fusion. (a and b) Source images were obtained from a visible light camera (a) and an infrared camera (b). (c) The fused image includes details

Summary

- **Enhance image capture by combining observations**
- **Combine to preserve contrast (max gradient)**
- **Gradient pyramid framework (multiscale)**
- **Deliberately limit each observation (narrow band)**

Radar E-O Image Fusion

William F. O'Neil
Westinghouse Electric Corporation

205943
1-8
249930

Abstract

The fusion of radar and electro-optic (E-O) sensor images presents unique challenges. The two sensors measure different properties of the real three-dimensional (3-D) world. Forming the sensor outputs into a common format does not mask these differences. In this paper, the conditions under which fusion of the two sensor signals is possible are explored. The program currently planned to investigate this problem is briefly discussed.

Introduction

Westinghouse has been developing novel adverse weather landing aids for commercial and military aircraft. We have concluded that it will be necessary to use a multiple sensor suite to provide both an active radar imaging sensor, and a passive imaging E-O sensor. The radar imager provides excellent penetration of adverse weather, but has limited angular resolution. The E-O sensor provides very good angular resolution but is severely affected by adverse weather such as fog, rain or snow. The fundamental property that distinguishes the two sensor classes is operating wavelength. This is both the driver on adverse weather penetration, and the driver on angular resolution. When the wavelength is greater than the size of atmospheric aerosols and raindrops, the penetration is good. When the wavelength is small compared to the receiving aperture, the resolution is good.

For the current paper, an equally important distinction is the difference between active sensing and passive sensing. An active sensor provides its own illumination of the scene to be imaged, while a passive sensor depends on either some external illuminator, or on self-emitted radiation of the objects being imaged. An active sensor has an advantage in that the properties of the illuminating waveform can be exploited for coherent detection of reflected energy. This dependence on reflected (i.e. back scattered) energy determines how the active sensor images a real 3-D scene. Specifically, electromagnetic properties that are determined by the surface to some depth are important in determining the reflection characteristics. In addition, macroscopic scale features are important since energy can experience multiple reflections before being returned to the receiver.

For the E-O sensor the considerations are very different. Few surfaces are optically smooth. Thus the behavior of such surfaces in reflected light is significantly different than the behavior of the self-emitted energy. Multiple reflections of emitted or reflected energy play a minimal role in determining signal. The properties that determine reflection or absorption are not well correlated with the bulk properties that determine reflection and absorption at radar wavelengths.

The third distinction between radar and E-O sensors is that different evolutionary paths have resulted in radar providing very precise range and range rate measurements with only limited emphasis on received signal strength which is the only property usually quantified with an E-O sensor. For the application at hand, the radar image is returned as a range versus azimuth angle using an antenna that is mechanically scanned, and which has a shaped beam pattern designed to minimize the variation of signal with elevation angle, under the assumption of flight nearly parallel to the ground where returns originate. The use of range versus angle as opposed to signal return level versus angle presents some challenges. Height of a return source above terrain is lost. Converting from an azimuth/range/intensity image to an azimuth/elevation/intensity image requires an assumption about the height of the return sources. Figure 1 shows an E-O sensor image of a runway at the Salisbury MD airport. Figure 2 shows the same runway as viewed using an X-Band (10 GHz) radar operated in the Monopulse Ground Map (MGM) mode. Figure 2 was derived from Figure 3 (azimuth/range/intensity) by assuming that all reflecting elements are in a ground plane which has a known orientation with respect to the flight path. As shown in Figure 4, each range cell in the radar return is assigned an elevation angle on the basis of the aircraft height above the ground plane. While there are a number of important error sources which must be accounted for in this process, for the purpose of this paper, it is sufficient to assume that those difficulties will be overcome, and that a proper image in angle/angle/intensity format will be achieved.

Fusion Technique

Westinghouse has approached the task of Radar E-O image fusion as an evolution of previously developed technology. The MGM mode for the radar, coupled with a transformation from azimuth/range to azimuth/elevation produces an image which has a compatible format with standard E-O images and displays. Westinghouse has also been participating with the David Sarnoff Research Center in a program that uses pyramid decomposition of visible and IR E-O images to construct fused images. That program has advanced to the point where real time operation at television rates and resolutions will be possible in the very near future. Combining these two developments provides a path to the desired Radar E-O fusion. The paper by Dr. Hannah at this workshop describes the pyramid fusion technique for visible and IR images. The interested reader will find additional information in references 1-3.

Figure 5 shows the general arrangement of a postulated Radar E-O image fusion system. The Radar is operated in the MGM mode and creates angle/range/intensity images at a low frame rate. These are converted to angle/angle/intensity images using a combination of on-board inertial and altitude sensors. The images are used to generate a 30 Hz image stream by motion compensation plus image extrapolation. This step may occur either before or after pyramid decomposition, depending on engineering details. The Radar images are decomposed using pyramid decomposition. The E-O images are similarly decomposed, so that features from both images can be identified, matched, and registered. Feature blending/selection is used to produce the composite image in transform space. This image is then inverse transformed, using the merged pyramids to construct the angle/angle/intensity image. Standard processes, such as gain and level adjustment, are then used to correct that image prior to display.

The pyramid decomposition has the effect of generating intermediate images which contain a limited range of spatial frequencies. Thus, the decomposition of a high resolution E-O image will result in transformed images that have resolution compatible with the Radar resolution. By suitable choice of scan angles, sampling rates, and optical design, the reduced resolution image will match the resolution of the Radar such that direct comparison and fusion of features will be possible. Figure 6 shows the decomposition and feature match processes. The fusion process, represented by a single block, is a variant of the previously published work.

Each cycle of the pyramid decomposition produces a bandpass image (the Laplacian) that contains one octave of spatial frequency data, plus a residue image that contains all spatial frequencies from zero to the lower limit of the bandpass image. The two image sources can, by suitable choice of sampling grids, provide bandpass images that share a common range of spatial frequencies. It is also a property of the pyramid decomposition process that the spatial coordinates of each feature are preserved in the transform process. Thus, each feature will be represented by both spatial coordinates and spatial frequency content. Relatively simple operations such as rectification and thresholding permit the determination that the feature is present. If such a test is satisfied in both images, then the features can be fused into a single feature that can be displayed. In addition, a feature present in one image, but not the other, can be used in the composite image. This will provide an image containing the information from both sources.

Fusion Issues

The above discussion of Radar E-O fusion has glossed over several potential difficulties. The most obvious is commensurability. Are the features in a Radar image sufficiently similar in size, shape, location, or intensity to be clearly identifiable as the same feature by some analytic rule? Is the only answer to this question anecdotal, or is there a formal method for resolving this issue?

One approach to the commensurability is shown in Figure 7. Both scenes are derived from the same 3-D real world. Each of the sensors has performed a transform into one or more spaces depending on where we choose to view the image. If we can add a transform to one or both images which produces intermediate images which are demonstrably the same for equal real world inputs, then, in that transform space, they are commensurable and can be merged. As inspection of Figure 7 shows, it is a generalization of Figure 5 which is the particular transform path we are exploring.

Another issue might be called "fusability". If we identify a feature from both sensors, and can conclude that it is the same feature, we are still left with the need to transform the features in such a way as to provide commensurability in intensity space. We have not envisioned an alternative since the objective is to provide an intensity/angle/angle image for a pilot. The fusability issue is also linked with the issue of deciding which sensor contributes how much to the final image. The visible IR fusion effort used a binary decision rule, but we anticipate that a blending rule will prove advantageous in the present case. Some departure from current radar practice may be needed to assess the image quality of the radar signal, and assign the transformed image an equivalent intensity for a blending rule.

Still another issue of concern is the subject of clutter. Spatial clutter is a potential problem for both sensors, while temporal clutter is observed in Radar images. Such clutter complicates the processing task, since it represents additional features which must be analyzed. Applying image extrapolation to achieve compatibility with the 30 Hz video, may aggravate clutter as a distraction to the pilot. The low sample rate which is provided by the radar is effectively aliased into higher temporal frequencies by any extrapolation algorithm.

Current Plans

Westinghouse is engaged in an analytic and experimental program to investigate these issues. The analytic program includes development of basic theoretical models for the sensor phenomenology, as well as investigations using simultaneous data from multiple sensors. To address these issues requires that a significant data base be available. Westinghouse has an instrumented aircraft that provides both radar and E-O sensors with digital data collection. Initial efforts will include collecting data from the Westinghouse MODARS weather radar together with visible and IR E-O data. This will be processed in our image processing laboratory to evaluate algorithms and assess fundamental problems which must be solved. From these results, we plan to formulate a program where the fusion process can be implemented as a real time airborne process.

Conclusions

The fusion of Radar and E-O sensor data will provide the ability to select an optimum mix of resolution and penetration for each weather condition that will be encountered. To be effective, the fundamentals of fusion across different image domains must be established so that a fully automated fusion system can be implemented. The spatially coherent pyramid decomposition technique appears to offer significant benefits in this fusion effort. There are fundamental unanswered questions which must be addressed. In addition, the experimental data base required to assess alternative theories has not been obtained. Westinghouse has initiated a program that will address the theoretical and experimental issues of Radar E-O fusion.

References:

1. Burt, Peter J. and Edward H. Adelson, "Merging Images Through Pattern Decomposition", SPIE National Summer Conference, San Diego, CA Aug. 18-23, 1985
2. Burt, Peter J. and Edward H. Adelson, "The Laplacian Pyramid as a Compact Image Code", IEEE Transaction on Communications, Vol COM-31, No. 4, April 1983
3. Pavel, Misha and James Larimer, Editors, Session 25 of the Society for Information Display Meeting of May 20, 1992 contains three papers on image fusion.

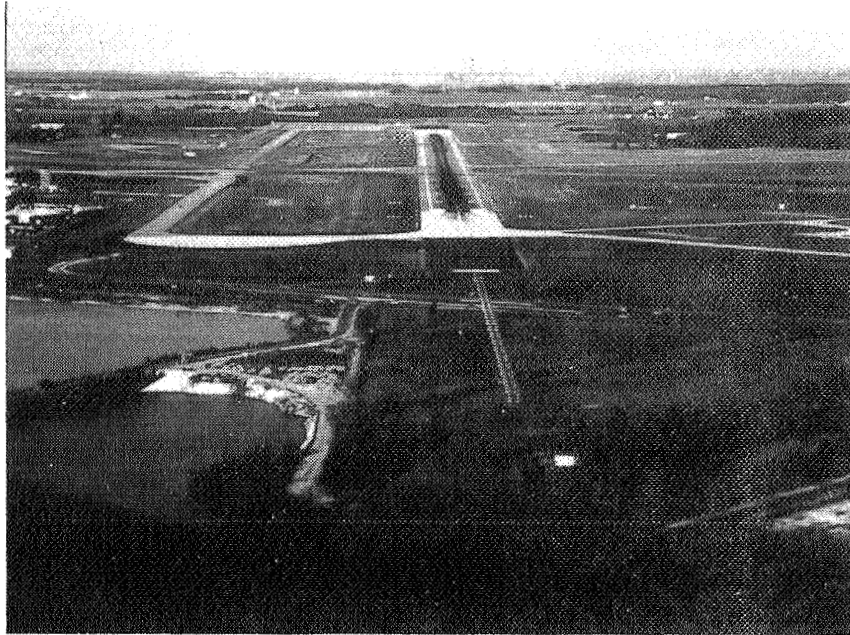


Figure 1 Visible Sensor Image of Runway and Environs

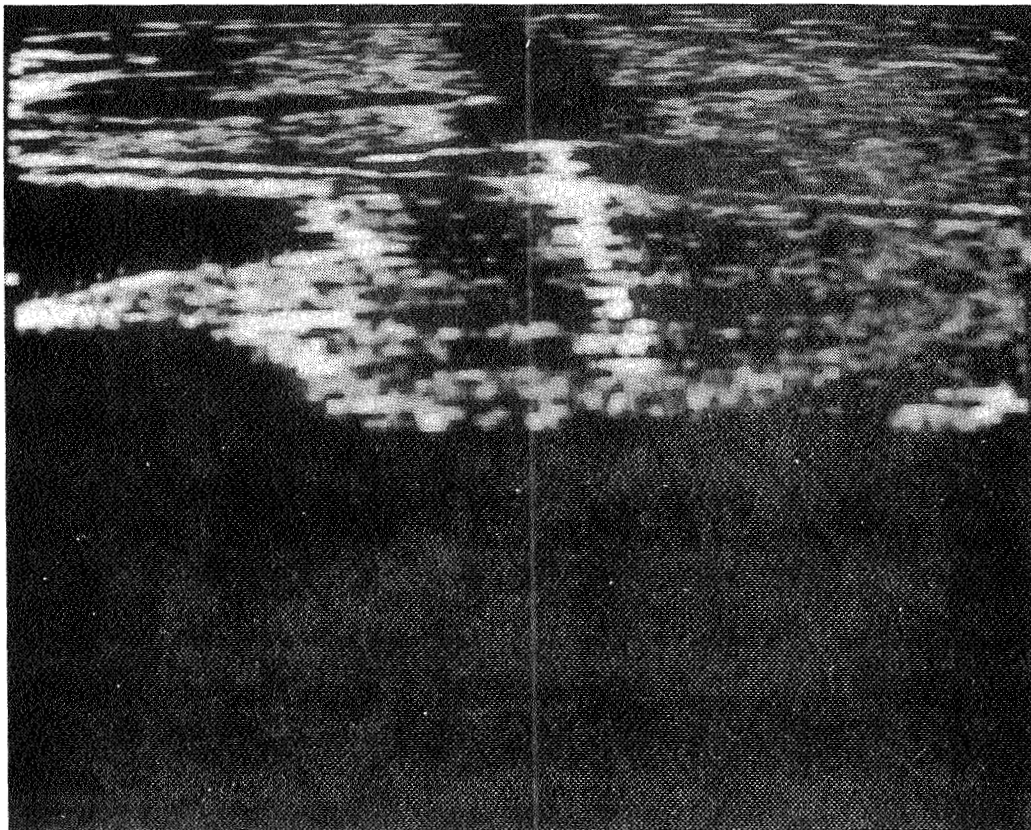


Figure 2 Angle/Angle Radar Image of Runway and Environs

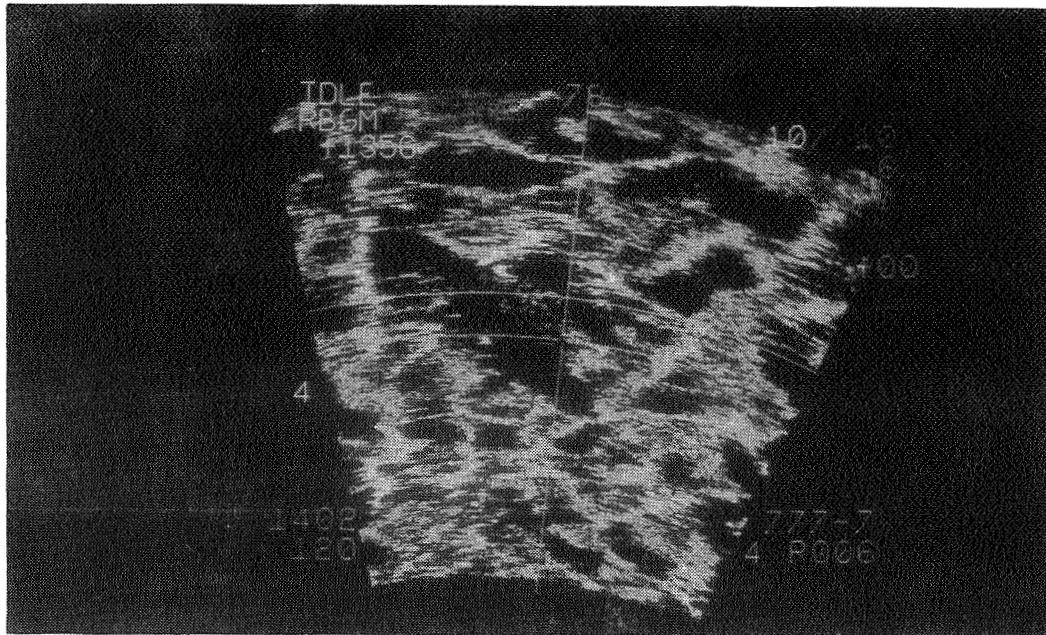


Figure 3 Azimuth/Range Radar Image of Runway and Environs

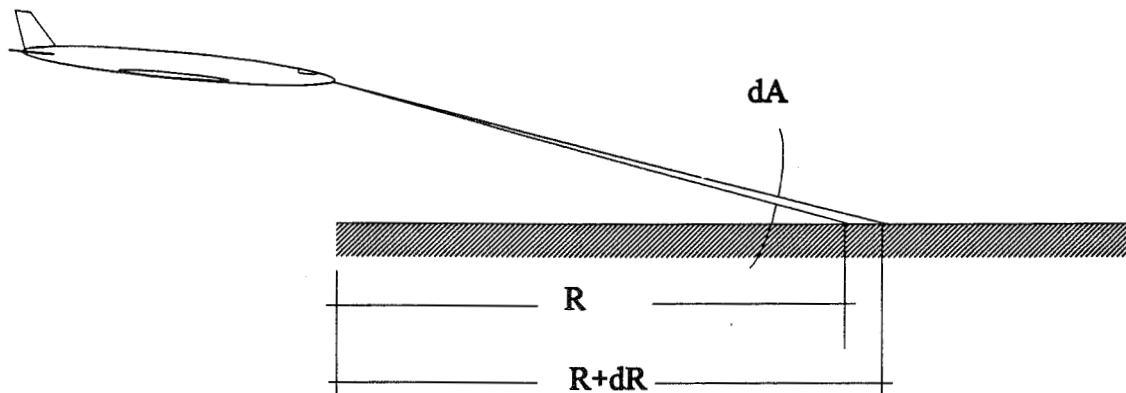


Figure 4 Conversion from Radar Range to Elevation

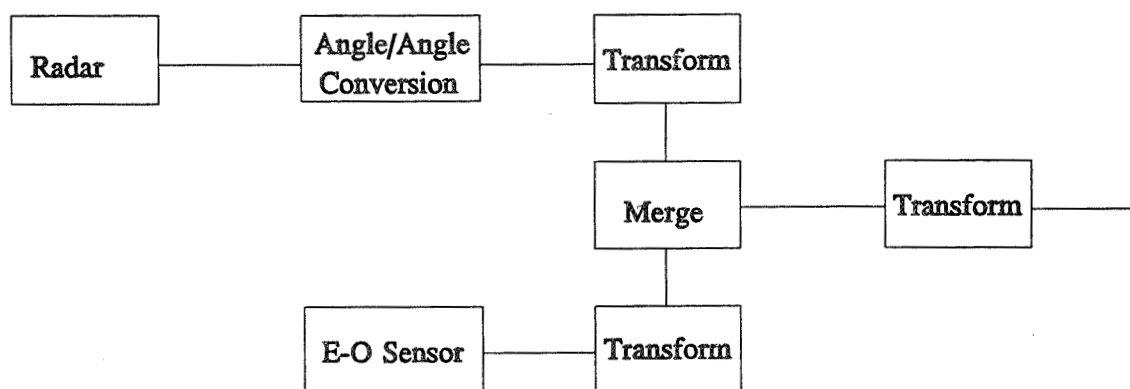


Figure 5 Radar/E-O Fusion Using Pyramid Transform

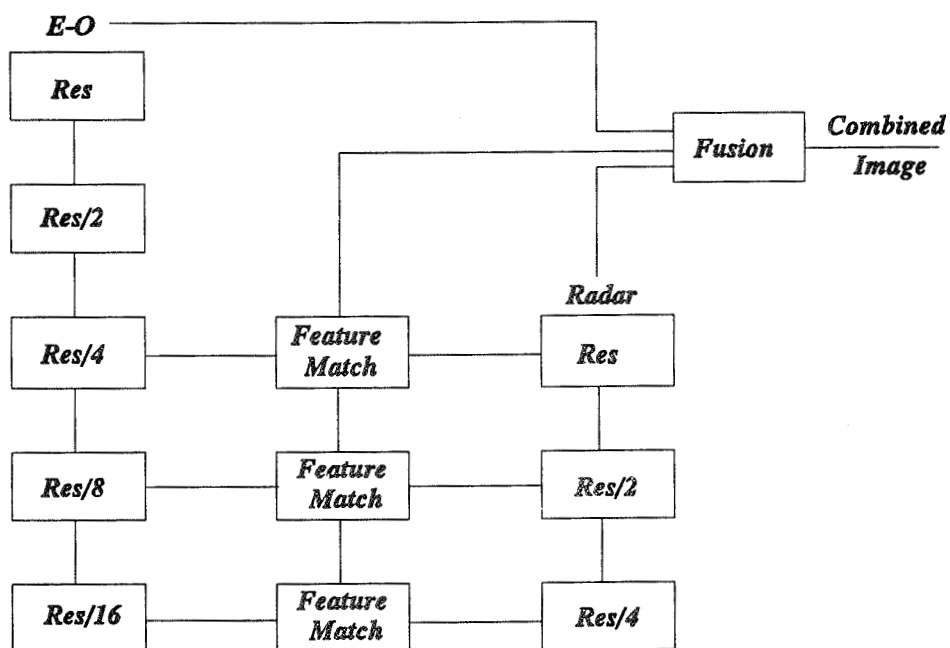


Figure 6 Reduced E-O Resolution Matches Radar Resolution

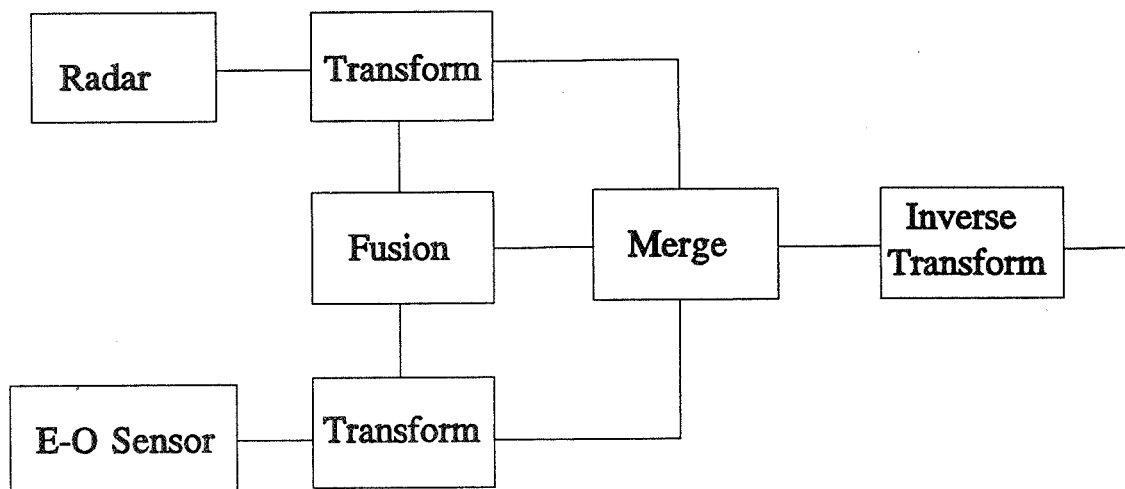


Figure 7 Generalized Radar E-O Fusion Using Transforms

omit

IV. IMAGE PROCESSING: COMPUTER VISION

~~PRECEDING PAGE BLANK NOT FILMED~~

Expansion-Based Passive Ranging

Yair Barniv
NASA Ames Research Center

514-04

205744

P- 52

249931

SUMMARY

This paper describes a new technique of passive ranging which is based on utilizing the image-plane expansion experienced by every object as its distance from the sensor decreases. This technique belongs in the feature/object-based family.

The motion and shape of a small window, assumed to be fully contained inside the boundaries of some object, is approximated by an affine transformation. The parameters of the transformation matrix are derived by initially comparing successive images, and progressively increasing the image time separation so as to achieve much larger triangulation baseline than currently possible. Depth is directly derived from the expansion part of the transformation.

To a first approximation, image-plane expansion is independent of image-plane location with respect to the focus of expansion (FOE) and of platform maneuvers. Thus, an expansion-based method has the potential of providing a reliable range in the difficult image area around the FOE. In areas far from the FOE the shift parameters of the affine transformation can provide more accurate depth information than the expansion alone, and can thus be used similarly to the way they have been used in conjunction with the Inertial Navigation Unit (INU) and Kalman filtering. However, the performance of a shift-based algorithm, when the shifts are derived from the affine transformation, would be much improved compared to current algorithms because the shifts—as well as the other parameters—can be obtained between widely separated images.

Thus, the main advantage of this new approach is that, allowing the tracked window to expand and rotate, in addition to moving laterally, enables one to correlate images over a very long time span which, in turn, translates into a large spatial baseline—resulting in a proportionately higher depth accuracy.

ACRONYMS USED IN TEXT

FOE	- Focus of Expansion
FOV	- Field of View
INU	- Inertial Navigation Unit
LOS	- Line of Sight
OF	- Optical Flow
PSF	- Point-Spread-Function
SNR	- Signal-to-Noise Ratio
PSR	- Peak-to-Sidelobes Ratio
TBD	- Track Before Detect
TTC	- Time To Collision
3-D	- 3-Dimensional
AFTR	- Affine Transformation
CW, CCW	- Clockwise, counterclockwise

1 INTRODUCTION

Passive ranging is an area of considerable interest for applications such as obstacle avoidance for rotorcraft nap-of-the-earth navigation and spacecraft landing. Two main passive-ranging methods can potentially be employed for this purpose; one based on motion and the resulting image-plane optical flow, and the other based on stationary stereo. Both methods can be thought of as special cases of a more general triangulation method known as “bearing-only” or “direction-of-arrival” (*e.g.*, [1, 2, 3, 4]). In this paper we chose to concentrate on monocular OF-based ranging.

The motion of an imaging sensor causes each imaged point of the scene to correspondingly describe a time trajectory on the image plane. The trajectories of all imaged points constitute what’s called the “optical flow” (OF). A forward-looking imaging sensor, such as a TV camera or a Forward-Looking Infra-Red (FLIR), is typically used as the source of optical flow data. The various methods of extracting depth information from the OF can be classified as belonging into three main classes as we did in [5]. The method described in this paper can be considered object-based or feature-based depending on the definition of features. If the features are chosen based on some local image property, such as texture or edge, then we are dealing with a feature-based method. If, the feature is chosen through some pre-segmentation to be wholly contained inside a physical object, then our new technique can be considered to be object-based; that is how we chose to regard it in this paper.

Like all other passive-ranging methods, we assume that the scene and its illumination sources are temporally constant (see [6]). We also assume that all points belonging to the same object share the same range. In [5] we differentiated between detect-then-track and track-before-detect (TBD) algorithms (akin to filtering and smoothing respectively) and pointed out the advantages of the TBDs in terms of SNR-performance and robustness (see [7, 8, 9]). We will return to this subject after presenting our new algorithm, and show that it is a TBD one.

The OF at any given point in the image plane consists of three kinds of motion: lateral translation, expansion (or divergence), and rotation (curl). When considering a window of some finite size, one can approximately describe its time evolution by an affine transformation which, in the most general case, has six parameters: four belonging to the 2×2 multiplying matrix, and two belonging to the vector of lateral translation. Most depth-estimation methods, such as described in [10, 11], make use of the lateral motion alone. Two basic limitations are implicit in these methods. First, they cannot perform in the image plane close to the FOE, and second, they can only use a relatively short triangulation baseline because far-apart images would not correlate due to the misadjustment in the other components of the affine transformation (besides the shifts) which are not accounted for. “Triangulation baseline” is the term we use for the distance the platform travels between the frames to be correlated. As we have shown in our earlier work [12], the depth-error is inversely proportional to the triangulation baseline (see (18)

ahead).

In this work I will discuss methods of extracting depth information from the divergence of the OF as it is approximately obtained from the affine transformation matrix. I use the term “divergence” (or “local divergence”) to refer to the mathematical definition of the derivative-vector operator denoted by ∇ which, in this case, scalar-multiplies the velocity vector at a point. Divergence is thus defined for an infinitesimal area and time. We use “expansion”, or “global divergence”, as a short-hand for the “rate of area expansion” to denote the *average* divergence over the area of some finite-size window or of an object. We will soon see why the divergence of the velocity vector, $\nabla \cdot \mathbf{v}(p)$ at some point p actually measures the rate of area expansion (which explains the above proliferation of terms).

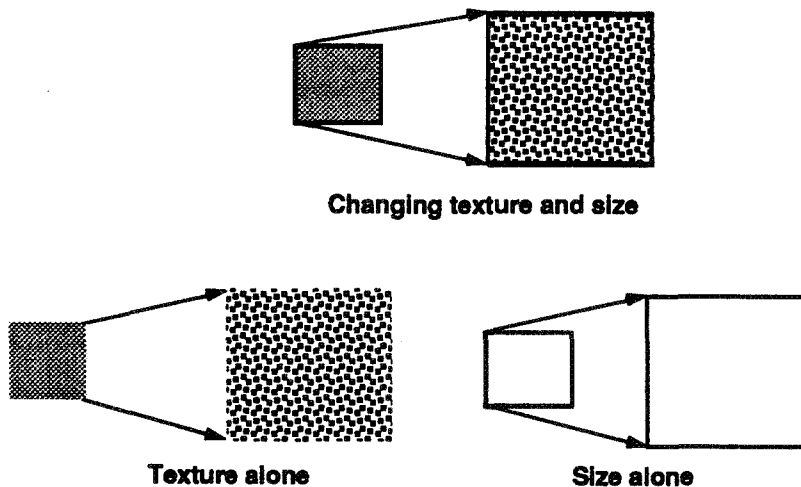


Figure 1: Texture and size cues.

The idea of using divergence as a source of depth information is not new. The works of Longuet-Higgins and Prazdny [13], Prazdny [14, 15], Koenderink [16], Koenderink and van Doorn [17, 18], and Nelson and Aloimonos [19] elaborate quite extensively on this subject. Recently, an interesting extension to these works was reported by Ringach and Baram in [20]; although it is field-based, it explicitly assumes that the scene is composed of objects (defined by their borders) and derives the global divergence for all objects without the need to actually delineate or identify them. The local- and global-divergence methods are intended for different kinds of objects as exemplified in figure 1. The local-divergence method is intended for textured objects with no well-defined edges, whereas the global-divergence method is intended for objects with little or no texture but having well-defined edges. In this paper we rely upon the objects being textured, so our algorithm roughly derives the equivalent of local divergence.

If we examine a window centered on the FOE, its translational motion is zero by definition, but it still expands as the depth decreases, and this expansion is left as the only source of depth information. Thus, there are two new aspects to our work; one is the direct derivation of depth from expansion, and the other is enabling the use of a long triangulation baseline for

even using just the conventional translation-based methods (as well as for the expansion-based ones). The later feature is the one that transforms this algorithm from a track-then-detect to a TBD, because the accuracy (or SNR) of the final result is based on the largest available baseline, as opposed to (Kalman) filtering of results that were individually obtained based on a single-interframe baseline.

This is why one can consider this work to represent an extension of the existing translation-based algorithms such as the one developed by Sridhar, Phatak, and Cheng in [10, 11] and Sridhar, Suorsa and Hussien in [21] which derive the image-plane translations of "points of interest" (small windows) through spatial cross-correlation between consecutive images and subsequent Kalman filtering of their image-plane trajectories.

In order to round off the picture, we also need to refer to another closely-related area of research represented by the work of Merhav and Bresler (see [22, 23, 24, 25]). The first three papers primarily address image-plane motion estimation, which is, of course, equivalent to depth. Also, they rely upon the assumption (that we do not need to make) that the image statistics in the X and Y directions are separable. The fourth paper suggests a stochastic-gradient approach to image-plane motion estimation which can be thought of as a precursor of the work reported here.

As a last comment, it is noteworthy that utilizing divergence (or expansion) for depth derivation has been largely motivated by advances in the understanding of visual processing in humans and primates. For example, experiments with humans suggest the existence of divergence (looming) detectors in the human visual system [26, 27, 28] as well as *vorticity* detectors [28, 29, 30].

The organization of this report is as follows. Section 2 contains the theory relating Divergence, Expansion, and Depth. Section 3 presents the idea of using the affine transformation to relate objects in different frames. Section 4 presents simulation results. Section 5 presents the practical algorithm that iterates over increased frame separation. Section 6 discusses the error analysis.

2 OPTICAL FLOW, DIVERGENCE AND EXPANSION

The basic equations for the divergence in the image plane are derived in this section. This derivation is based on prior work described in [13] to [20].

It is convenient to think for a moment of imaging the outside scene onto a spherical surface because such projections are identical irrespective of the camera-axis direction. In fact, with

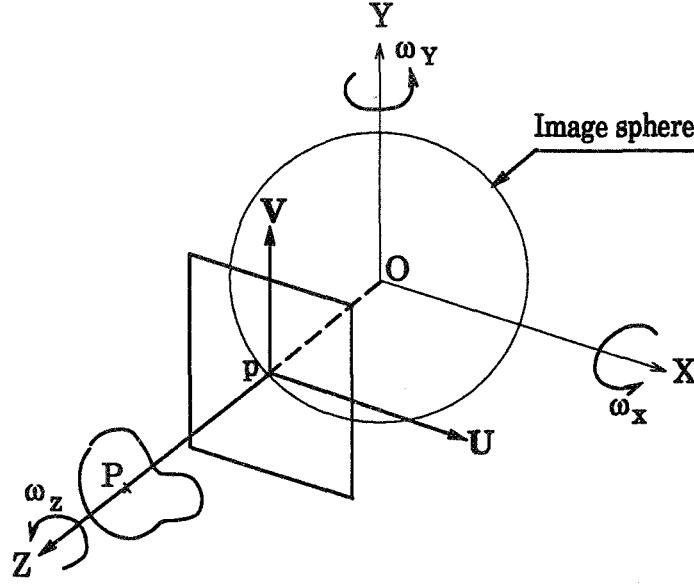


Figure 2: The geometry of projection onto the image plane.

such geometry, the camera axis is *defined* to coincide with the line-of-sight (LOS) from the center of the sphere to any imaged point as seen in figure 2. Another motivation for regarding the image plane as a sphere is that this geometry is similar to that of imaging the world by a lens onto a spherical retina, *e.g.*, in the human eye. Let us define the coordinate system of the spherical camera to have its origin at the sphere's center and its Z axis to pass through the imaged point P of some object. The sphere is defined to be of unity radius. Consider the projection of P onto the sphere at point p . At that point define the origin of an (U, V) plane tangent to the sphere which is called *local projective image plane* (image plane, for short); this image plane approximates the sphere at the point of tangency. Let us assume that P is found on a smooth surface described by some function $z = f(x, y)$ so that its gradient $\nabla z = (z_X, z_Y)$ exists. The distance of any point on that surface from the sphere's center can then be approximated in the neighborhood of P by

$$z \approx z_0 + \nabla z \cdot (x, y), \quad (1)$$

where z_0 is the distance between points O and P. The relative motion of the camera with respect to the scene is defined by its translational velocity $\mathbf{V} \triangleq (V_X, V_Y, V_Z)$ and rotational velocity $\boldsymbol{\omega} \triangleq (\omega_X, \omega_Y, \omega_Z)$. It is convenient to normalize $\mathbf{V}$ by z_0 and define $(v_X, v_Y, v_Z) \triangleq (V_X, V_Y, V_Z)/z_0$.

The motion of the camera causes the stationary point P and its surrounding to describe a retinal velocity field (or optical flow) around p on the image plane. We denote image-plane projections by (u, v) , to correspond with (U, V) , and their temporal derivatives by (u_t, v_t) . Thus, the image-plane velocity vector at p is defined as $\mathbf{v}(p) \triangleq (u_t, v_t)|_p$. The spatial partial derivatives of (u_t, v_t) are denoted by $u_{tx}, u_{ty}, v_{tx}, v_{ty}$. From [13] we know that the following equations hold at p ,

$$u_t = -v_X - \omega_Y, \quad v_t = -v_Y + \omega_X,$$

$$\begin{aligned}
u_{tx} &= v_Z + v_X z_X, & v_{ty} &= v_Z + v_Y z_Y, \\
u_{ty} &= \omega_Z + v_X z_Y, & v_{tx} &= -\omega_Z + v_Y z_X
\end{aligned} \tag{2}$$

Using the above equations, the divergence at p (denoted $\text{div}(p)$) can be expressed as

$$\text{div}(p) = \nabla \cdot \mathbf{v}(p) \triangleq u_{tx} + v_{ty} = 2v_Z + \nabla z \cdot (v_X, v_Y) \tag{3}$$

To interpret the above equation, suppose that the camera only moves in the Z direction. In that case $v_X = v_Y = 0$ and $\nabla \cdot \mathbf{v}(p) = 2v_Z = 2V_Z/z_0$, that is, $\text{div}(p)$ is twice the reciprocal of the time-to-collision (TTC) of P with the camera's center. Because of this interpretation, $\text{div}(p)$ is termed "immediacy" in [16] and other papers, that is, it measures the immediacy of an imminent collision. In the opposite case, when $(v_X, v_Y) \neq (0, 0)$ and $v_Z = 0$, there can still be a relative depth change between the camera and the patch because it is generally slanted. In other words, $\text{div}(p)$ will still have the same interpretation as before, except that the imminent collision is going to be with some point on the plane tangent to the patch at P and not with the point P itself. Thus both terms of the immediacy have a valid physical interpretation. Notice that the rotational velocities do not appear in $\text{div}(p)$. This is a very important (and well-known) observation because it says that *the TTC information is wholly contained in the imagery; no additional information is needed* (such as from the INU).

Nelson and Aloimonos describe in [19] a straight-forward mechanism for evaluating the divergence from a sequence of images. In practice, this algorithm can only provide a rough estimate of the local divergence.

The global divergence is defined (see [20]) as the *average* divergence over the area of each object, and denoted by $\chi(R)$ for an object whose projection onto the image plane is R (assuming, for the moment, that its boundary ∂R is well defined). Thus,

$$\chi(R) \triangleq \frac{1}{A(R)} \int_R \text{div}(p) \, ds = \frac{1}{A(R)} \int_{\partial R} \nabla \cdot \mathbf{v}(p) \, ds = \frac{1}{A(R)} \int_{\partial R} \mathbf{v}(p) \cdot \mathbf{n} \, dl, \tag{4}$$

where $A(R)$ is the object area, ds is the elemental area, dl is the elemental length along ∂R , $\mathbf{n}$ is a unit vector normal to ∂R , and the equality is based on the divergence theorem. In words, the average divergence equals the line integral of the normal component of the velocity vector at the edge along the edge of the object. This line integral can easily be shown (see [20]) to have an intuitive interpretation, that is,

$$\chi(R) = \frac{1}{A(R)} \frac{dA(R)}{dt}, \tag{5}$$

i.e., the global divergence equals the temporal rate of change of the normalized object area.

To find the relationship between global divergence, expansion, and time-to-collision, consider the similar-triangles equation relating the image-plane projection at $(u, 0)$ of some point

similar to P but located at $(x = l, y = 0, z = z_0)$ in figure 2,

$$u = \frac{l}{z_0} \quad (6)$$

Taking the derivative of u with respect to z_0 , we find that

$$du = -\frac{u}{z_0} dz_0 = \frac{u}{z_0} V_Z dt = uv_Z dt, \quad (7)$$

Thus,

$$\frac{1}{u} \frac{du}{dt} = v_Z \quad (8)$$

If we repeat this derivation for an *area* change, dA , rather than for a length change, du , we would find, using $dA/A = 2du/u$, that

$$\frac{1}{A} \frac{dA}{dt} = 2v_Z \quad (9)$$

Comparing (9) to (5), it is seen that χ has the interpretation of twice the TTC. Thus, the normalized (by the area) temporal rate of change of the projected area A of some object, that is, its *rate of area expansion*, equals twice the TTC.

3 ESTIMATING THE RATE OF EXPANSION

In this section we introduce the affine transformation, and develop the algorithm necessary to estimate the object's rate of expansion.

3.1 The affine transformation

The affine transformation (AFTR) can be used to relate object's projections at different frames (or times); its most general form is defined by six parameter. However, we intuitively judged that four parameters should suffice because they directly convey the physically-interpretable changes one would expect to occur. We thus define our specific AFTR by

$$\begin{bmatrix} \hat{u} \\ \hat{v} \end{bmatrix} = s \begin{bmatrix} \mathcal{O} & -\mathcal{S} \\ \mathcal{S} & \mathcal{O} \end{bmatrix} \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \begin{bmatrix} a + u_0 \\ b + v_0 \end{bmatrix}, \quad (10)$$

where s is a scaling (or expansion) factor, $\mathcal{O} \triangleq \cos(\theta)$ and $\mathcal{S} \triangleq \sin(\theta)$, and θ is the angle by which the object in I_1 is CW rotated with respect to its original orientation in I_0 . Thus, this AFTR maps points (u, v) from one frame (I_0) onto the corresponding points $(\hat{u}, \hat{v})$ in another frame (I_1). In figure 3 we notice that, first, the object expanded about 50%, second, it rotated about 25° CCW, and third, it moved up and right. This is indeed the order of mappings conveyed

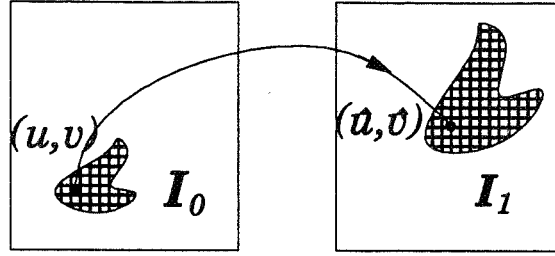


Figure 3: Mapping of a point through the affine transformation.

by the above definition although the order of scaling and rotation is immaterial. Notice that scaling and rotation is performed around the arbitrarily-defined center point of the object located at (u_0, v_0) , and shifting is performed later—back to the original center point plus an incremental shift by the vector (a, b) .

3.2 Vehicle's maneuvers and image-plane motion

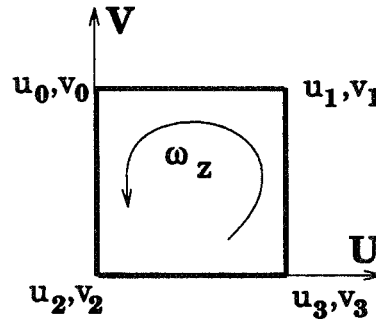


Figure 4: Window.

In this subsection we calculate the transformation that an object's projection undergoes as a result of platform maneuvers so as to relate it with the AFTR as defined above. To do that, we start from the well-known equations for the temporal derivatives of the image-plan projections (u, v) . Repeated as in [21], and adapted to our earlier notation, we have

$$\begin{aligned} u_t &= -fv_X + uv_Z + \omega_X \frac{uv}{f} - f\omega_Y \left(1 + \frac{u^2}{f^2}\right) + v\omega_Z \\ v_t &= -fv_Y + vv_Z - \omega_Y \frac{uv}{f} + f\omega_X \left(1 + \frac{v^2}{f^2}\right) - u\omega_Z, \end{aligned} \quad (11)$$

where f is the focal length. Now consider the shifts experienced by the corners of the window shown in figure 4. The differences between their shifts can be used to yield rotation and expan-

sion. The rotation of the upper side of the square (where $v_1 = v_0$) during some interframe time can be approximated by

$$\frac{\Delta v_1 - \Delta v_0}{u_1 - u_0} = -\omega_Z - \frac{v_0 \omega_Y}{f} \quad (12)$$

When the point (u_0, v_0) coincides with the FOE, this reduces to $-\omega_Z$. The rotation of the left side of the square (where $u_0 = u_2$) is similarly found as

$$\frac{\Delta u_2 - \Delta u_0}{v_0 - v_2} = -\omega_Z - \frac{u_0 \omega_X}{f}, \quad (13)$$

which also reduces to $-\omega_Z$ at the FOE. We have used the rotations of vertical and horizontal lines to show that, first, they rotate slightly differently, that is, in principle, the square distorts, and, second, this rotation approximately equals the platform roll. Comparing the two terms on the right of (12) for equal platform roll and yaw, the yaw term is smaller by a factor of f/v_0 . At a distance of, say, 50 pixels from the FOE, and with $f=622$ (using our camera as an example), this factor is 12.4. Since the expansion-based algorithm suggested here is intended to mainly enhance depth derivation around the FOE, we conclude that image-plane rotation is reasonably approximated by platform roll.

Next, let us analyze the expansion factor. For the upper side of the square it is approximated by

$$\frac{\Delta u_1 - \Delta u_0}{u_1 - u_0} = v_Z + \frac{v_0 \omega_X}{f} - \frac{(u_0 + u_1) \omega_Y}{f} \quad (14)$$

and for the left side of the square by

$$\frac{\Delta v_0 - \Delta v_2}{v_0 - v_2} = v_Z - \frac{u_0 \omega_Y}{f} + \frac{(v_0 + v_2) \omega_X}{f}, \quad (15)$$

At the FOE, both expressions approach v_Z as the square size goes to zero. Again, horizontal and vertical lines expand slightly differently, but, to a good approximation, this expansion equals v_Z (the TTC). The superfluous terms are an order of magnitude smaller than v_Z for areas up to 50 pixels from the FOE and small angular speeds.

Our conclusion is that, over the expected range of flight scenarios, the affine transformation represents a good approximation to the actual mapping that is taking place between different frames. If this approximation is not adequate, one can always use the full 6 degrees of freedom available in the general affine transformation.

3.3 What happens when scaling and rotation are ignored

In this subsection we elaborate on the importance of using the AFTR even for an algorithm which calculates depth based on the shifts alone. Ignoring the AFTR amounts to taking it to be

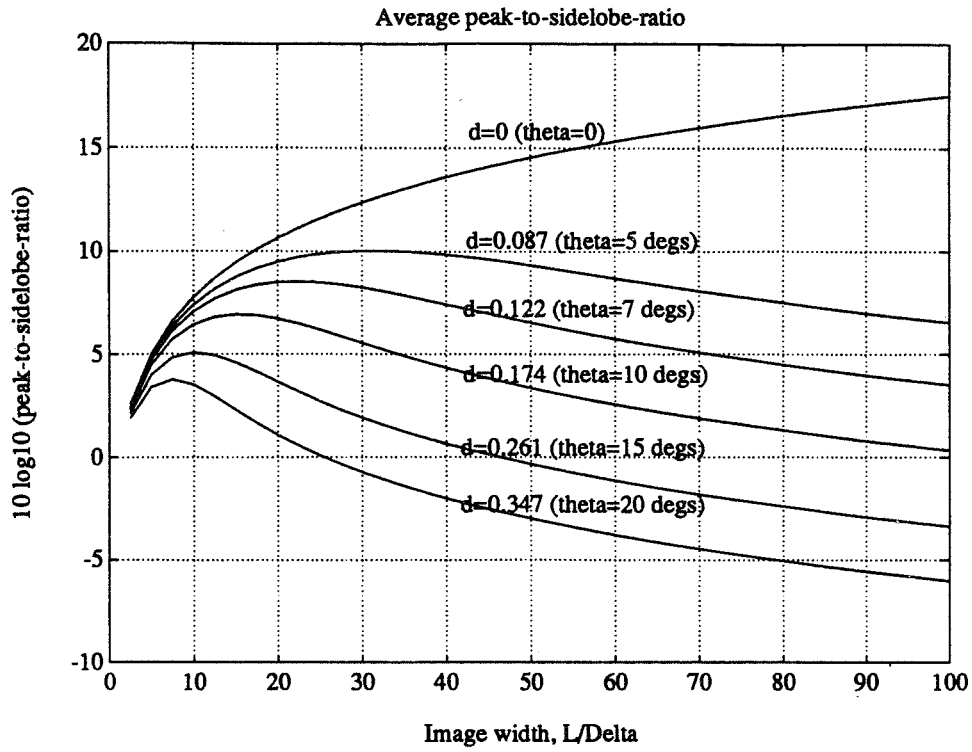


Figure 5: Average peak-to-sidelobes ratio as a function of image size for different distortions.

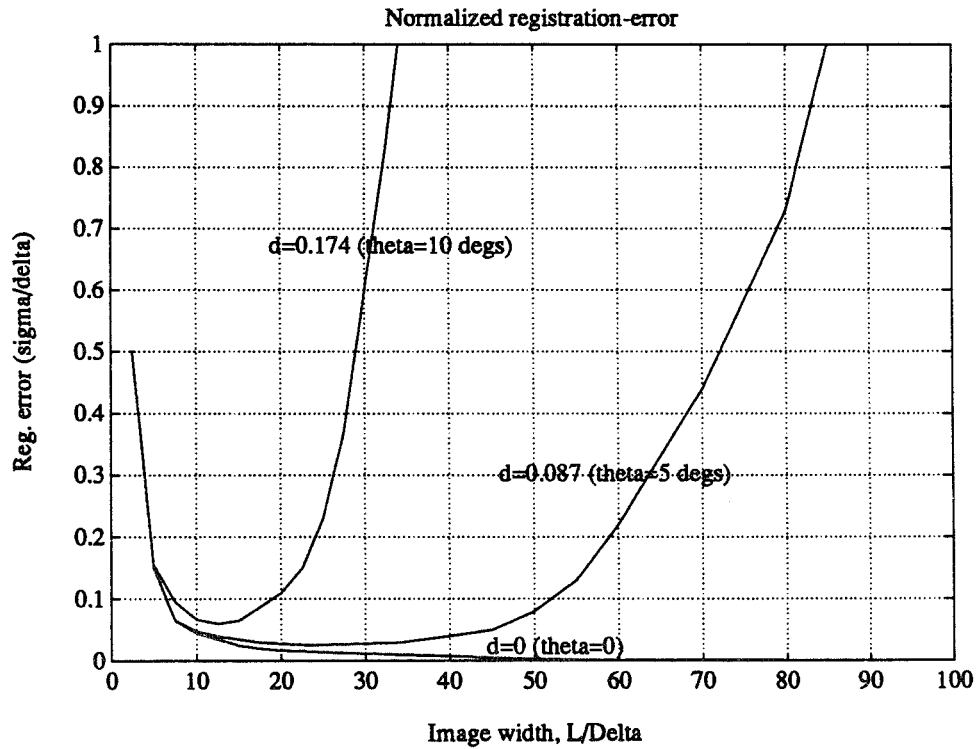


Figure 6: Registration-error standard deviation as a function of image size for different distortions.

a unity matrix. This question has been investigated quite extensively by Mostafavi and Smith in [31, 32]. For completeness, we summarize their results here.

For images having a circularly symmetric Gaussian correlation function,

$$R(\tau_u, \tau_v) = \exp \left\{ -\frac{1}{2\Delta^2} [\tau_u^2 + \tau_v^2] \right\}, \quad (16)$$

where τ_u, τ_v are the spatial shifts, and Δ the “correlation width”, the effects of non-compensated rotation (by θ) and/or scaling (by s) are determined by the combined geometric-distortion parameter d ,

$$d \triangleq \sqrt{|1 - 2s \cos \theta + s^2|} \approx \sqrt{(1 - s)^2 + \theta^2} \text{ for small } \theta \text{ and } s \approx 1 \quad (17)$$

Figure 5 shows the effect of d on the peak-to-sidelobes ratio (PSR). Peak stands for the maximum value of the cross-correlation function between two frames, and “sidelobes” stands for the standard deviation of the cross-correlation function far from the peak. The reference image is taken as a square of size $L \times L$. The other (sensed) image is much larger than the reference image. In the figure, L appears normalized by the correlation width because what counts is the effective number of “independent” image objects. Six graphs are shown for different d values. The graph for $d = 0.087$, for example, can be used for rotation alone (of 5°), or for scaling alone ($s = 1.087$), or for any of their combinations such that (17) yields $d = 0.087$. Figure 6 similarly shows the behavior of the registration error.

Let us use the following example to demonstrate the effect of uncompensated rotation or scaling errors. Take speed $V_X = 25$ m/s, depth $z_0 = 120$ m, image-plane location 10 pixels from the FOE, a rolling maneuver of $\omega_Z = 20^\circ/\text{s}$, $L = 21$, $\Delta = 1.5$ pixels, and frame rate of 2 fr/s. This low frame rate is used to achieve a large triangulation baseline as will be explained later. Only two consecutive frames are used in this example.

In a single interframe time the platform rotates 10° and there is an expansion by a factor of $s = 120/(120 - 25 \cdot 0.5) = 1.1163$, so that $d = 0.21$. The PSR will incur a loss of ≈ 3 (6 dB in PSR power)—as read from figure 5. This is why, without using the AFTR, one needs to use a higher frame rate, say, 10 fr/s. The registration error, as extrapolated from figure 6, will increase from 0.025 to 0.070 pixels. In [12] we have found the depth error:

$$\sigma_z = \frac{\sqrt{2} z_0^2 \sigma_u}{b u}, \quad (18)$$

where b is the triangulation baseline. Thus, the depth error incurred by a geometrically-compensated algorithm ($b = 12.5$ m) is 4.1 m while that incurred by a non-compensated algorithm ($b = 2.5$ m) is 57 m (out of 120 m!).

This example shows that, even in the conventional shift-based algorithm, neglecting to compensate for the AFTR in the process of cross-correlation is costly in two ways. First, it

either degrades the PSR which may hinder locking onto the correct peak (false alarm) or impose a short b , and second, even when correct peak detection is achieved, the depth error would increase around tenfold.

3.4 Converging on the correct affine transformation

In this subsection we derive the equations and algorithm necessary to obtain the correct affine transformation. The basic idea is to use Newton's equation (see [33]) iteratively to converge from the initially-assumed transformation matrix into the correct one by minimizing an appropriate error measure, or cost-function.

We thus start by defining the cost-function, J , as the integral over the window area, A , of the squared difference of image gray levels, that is,

$$\epsilon \triangleq I_1(\hat{u}, \hat{v}) - I_0(u, v); \quad J \triangleq \frac{1}{2A} \iint_A \epsilon^2 dA \quad (19)$$

If all points (u, v) inside the window (defined in I_0) are correctly mapped into $(\hat{u}, \hat{v})$ of I_1 , then the above cost should equal zero. In practice, however, we can only expect to *minimize* this cost albeit not to drive it to zero. Our plan is to find the gradient and second derivatives of J so that we can use Newton's method to *solve* for the minimum assuming that the cost-function is quadratic in the four parameters to be estimated. Since this assumption only holds approximately, it is necessary, in practice, to iterate a few times until the solution converges. The iterative update equation for the estimated parameter vector $\hat{X}(k)$ becomes

$$\hat{X}(k+1) = \hat{X}(k) - \{\nabla^2 J[\hat{X}(k)]\}^{-1} \nabla J[\hat{X}(k)], \quad (20)$$

where

$$X(k) \triangleq \begin{bmatrix} a \\ b \\ s \\ \theta \end{bmatrix} \quad (21)$$

The four components of the cost-function gradient are calculated next. Starting with the first shift-parameter, a ,

$$\frac{\partial J}{\partial a} = \frac{1}{A} \iint_A \epsilon \frac{\partial \epsilon}{\partial a} dA = \frac{1}{A} \iint_A \epsilon \frac{\partial I_1}{\partial a} dA, \quad (22)$$

because only the $I_1(\hat{u}, \hat{v})$ part of ϵ depends on a through $\hat{u}, \hat{v}$. Developing that relationship,

$$\frac{\partial I_1}{\partial a} = \frac{\partial I_1}{\partial \hat{u}} \frac{\partial \hat{u}}{\partial a} + \frac{\partial I_1}{\partial \hat{v}} \frac{\partial \hat{v}}{\partial a} \quad (23)$$

Similar equations are obtained for the other three parameters by substituting them in place of a in (23). The above four equations require the partials of $\hat{u}, \hat{v}$ with respect to all four parameters. These are obtained by differentiating the two scalar equations obtained from (10), that is,

$$\begin{aligned}\hat{u} &= s[\mathcal{C}\theta(u - u_0) - \mathcal{S}\theta(v - v_0)] + u_0 + a \\ \hat{v} &= s[\mathcal{S}\theta(u - u_0) + \mathcal{C}\theta(v - v_0)] + v_0 + b,\end{aligned}\quad (24)$$

so that

$$\begin{aligned}\frac{\partial \hat{u}}{\partial a} &= 1; & \frac{\partial \hat{v}}{\partial a} &= 0 \\ \frac{\partial \hat{u}}{\partial b} &= 0; & \frac{\partial \hat{v}}{\partial b} &= 1 \\ \frac{\partial \hat{u}}{\partial s} &= \mathcal{C}\theta(u - u_0) - \mathcal{S}\theta(v - v_0); & \frac{\partial \hat{v}}{\partial s} &= \mathcal{S}\theta(u - u_0) + \mathcal{C}\theta(v - v_0) \\ \frac{\partial \hat{u}}{\partial \theta} &= -s[\mathcal{S}\theta(u - u_0) + \mathcal{C}\theta(v - v_0)]; & \frac{\partial \hat{v}}{\partial \theta} &= s[\mathcal{C}\theta(u - u_0) - \mathcal{S}\theta(v - v_0)]\end{aligned}\quad (25)$$

We now need the ten second derivatives of the symmetrical matrix $\nabla^2 J[\hat{X}(k)]$. In order to simplify notation, we will drop the “dA” from the integrals, the subscript 1 from I , and the hats from u, v ; these will now be understood whenever not specified. Let us start with one of the mixed second derivatives, say that of a and θ . We thus have

$$\frac{\partial^2 J}{\partial a \partial \theta} = \frac{\partial}{\partial a} \left(\frac{\partial J}{\partial \theta} \right) = \frac{1}{A} \iint_A \frac{\partial \epsilon}{\partial a} \frac{\partial \epsilon}{\partial \theta} + \epsilon \frac{\partial}{\partial a} \left[\frac{\partial I}{\partial u} \frac{\partial u}{\partial \theta} + \frac{\partial I}{\partial v} \frac{\partial v}{\partial \theta} \right] \quad (26)$$

After some more algebra, we get

$$\frac{\partial^2 J}{\partial a \partial \theta} = \frac{1}{A} \iint_A U \frac{\partial u}{\partial a} \frac{\partial u}{\partial \theta} + V \frac{\partial v}{\partial a} \frac{\partial v}{\partial \theta} + W \left[\frac{\partial u}{\partial a} \frac{\partial v}{\partial \theta} + \frac{\partial u}{\partial \theta} \frac{\partial v}{\partial a} \right] + \epsilon \left[\frac{\partial I}{\partial u} \frac{\partial^2 u}{\partial a \partial \theta} + \frac{\partial I}{\partial v} \frac{\partial^2 v}{\partial a \partial \theta} \right], \quad (27)$$

where

$$U \triangleq \left(\frac{\partial I}{\partial u} \right)^2 + \epsilon \frac{\partial^2 I}{\partial u^2}; \quad V \triangleq \left(\frac{\partial I}{\partial v} \right)^2 + \epsilon \frac{\partial^2 I}{\partial v^2}; \quad W \triangleq \epsilon \frac{\partial^2 I}{\partial u \partial v} + \frac{\partial I}{\partial u} \frac{\partial I}{\partial v} \quad (28)$$

The other mixed second derivatives of J are similar and can be obtained by substituting the other parameters in place of a and θ in (28). The second (non-mixed) derivatives can, of course, be obtained by substituting the same parameter twice. For example, the second derivative of J with respect to a is

$$\frac{\partial^2 J}{\partial a^2} = \frac{1}{A} \iint_A U \left(\frac{\partial u}{\partial a} \right)^2 + V \left(\frac{\partial v}{\partial a} \right)^2 + 2W \frac{\partial u}{\partial a} \frac{\partial v}{\partial a} + \epsilon \left[\frac{\partial I}{\partial u} \frac{\partial^2 u}{\partial a^2} + \frac{\partial I}{\partial v} \frac{\partial^2 v}{\partial a^2} \right] \quad (29)$$

Notice that the above equations require two kinds of building blocks; these are the first and second (also mixed) spatial derivatives of the I_1 image as well as the first and second (also mixed)

derivatives of $\hat{u}$ and $\hat{v}$ with respect to the four transformation parameters. The image spatial derivatives are calculated by convolving it with a simple Sobel-operator-type 3×3 window. The first derivatives of $\hat{u}$ and $\hat{v}$ where already calculated for the gradient as in (25). Differentiating (25) once again yields 10 second derivatives for $\hat{u}$ and 10 for $\hat{v}$. Out of these 20, all turn out to be zero except the following four:

$$\begin{aligned} \frac{\partial^2 u}{\partial s \partial \theta} &= -\mathcal{S}\theta(u - u_0) - \mathcal{O}\theta(v - v_0) = -\frac{\partial v}{\partial s}; & \frac{\partial^2 v}{\partial s \partial \theta} &= \mathcal{O}\theta(u - u_0) - \mathcal{S}\theta(v - v_0) = \frac{\partial u}{\partial s} \\ \frac{\partial^2 u}{\partial \theta^2} &= s[-\mathcal{O}\theta(u - u_0) + \mathcal{S}\theta(v - v_0)] = -\frac{\partial v}{\partial \theta}; & \frac{\partial^2 v}{\partial \theta^2} &= -s[\mathcal{S}\theta(u - u_0) + \mathcal{O}\theta(v - v_0)] = \frac{\partial u}{\partial \theta} \end{aligned} \quad (30)$$

At this point all the components necessary for a single iteration on the Newton's solution have been derived.

4 SIMULATIONS OF THE COST-FUNCTION AND ITS DERIVATIVES

We now want to examine the behavior of the cost-function and its derivatives as a function of the four parameters in open loop, that is, without trying to correct the errors yet. For the following experimental results we used simulated imagery where the scene is composed of a wall normal to the initial flight trajectory. This wall is painted with a random Gaussian colored noise having spatial correlation width of 2 pixels in each of the two spatial dimensions. In this section we describe the main features of our Flight/Vision simulation and the open-loop error measurements.

4.1 Flight/Vision simulation

We have developed a simple simulation that enables us to generate a sequence of images (imagery) as obtained from an optical sensor that travels and maneuvers as prescribed. This simulation is described here.

The scenery is composed of a flat wall oriented normal to the initial LOS. The gray levels of the wall are derived by passing a white Gaussian noise through a two-dimensional Gaussian-shaped low-pass filter of some desired width. The wall is densely sampled by "wall-pixels" which, when imaged onto the camera's focal plane, are much finer than the "chip-pixels" of the camera. Typically 25 wall-pixels fall inside a single chip-pixel at the beginning of the run; this is chosen so that the wall can approximately be considered continuous. The correlation width of the low-pass filter above is chosen in terms of equivalent chip-pixels. In all simulations described later we chose correlation width of 2 chip-pixels because that is a typical width for the lens'

point-spread-function. The number of wall-pixels impinging on each chip-pixel is proportional to the depth squared because the camera uses a fixed angular field of view. Thus, to maintain a constant wall brightness on the image plane, we have to factor the wall brightness (or gray-level) by the depths-squared inverse. This compensation is nothing more than simulating the dependence of light radiation (power-per-area) on the inverse of the range squared.

The camera is initially located across from (and pointing to) the wall center at a distance of z_0 m. It is generally flying towards the wall center and can perform any desired maneuvers on its way. Each ray from the center of a wall-pixel to the camera's focal point (in world coordinates) gets transformed into the camera's coordinates through the 3×3 rotation matrices corresponding to yaw, pitch, and roll (*e.g.*, see [34]). The camera coordinates of the ray are used in the projection equations to yield the image coordinates of the ray's piercing point, that is,

$$u = f \frac{x}{z}; \quad v = f \frac{y}{z}, \quad (31)$$

We now assume a point-spread-function (PSF), having the shape of a chip-pixel and centered on the (non-integer) (u, v) point, to impinge upon the grid of chip pixels. This is where interpolation becomes necessary.

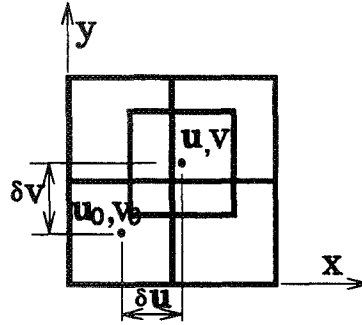


Figure 7: The interpolation method.

The method of interpolation is explained with the help of figure 7. The (u, v) point falls at a distance of $(\delta u, \delta v)$ from some integer point (u_0, v_0) . We thus assign the PSF areas intersected by each of the 4 chip-pixels to these pixels. The corresponding areas are thus assigned as follows:

$$\begin{aligned} (1 - \delta u)(1 - \delta v) & \quad \text{to pixel } (u_0, v_0); \\ (1 - \delta u)\delta v & \quad \text{to pixel } (u_0, v_0 + 1); \\ \delta u(1 - \delta v) & \quad \text{to pixel } (u_0 + 1, v_0); \\ \delta u\delta v & \quad \text{to pixel } (u_0 + 1, v_0 + 1), \end{aligned} \quad (32)$$

where $(\delta u, \delta v)$ are derived as

$$\delta u = u - \text{int}(u); \quad \delta v = v - \text{int}(v), \quad (33)$$

and $\text{int}(\cdot)$ is a function that rounds off its argument to the nearest lower integer.

As we have said above, there are, around 25 such partial contributions into every chip pixel—each contribution resulting from the center of a wall-pixel being projected onto some different (u, v) point. We have found that choosing the ratio between the sides of a chip-pixel and a wall-pixel to equal 5, using a chip-pixel-size PSF, and interpolating as prescribed above, results in a realistically-appearing textural behavior of the wall as it gets closer to the camera during the simulated flight. Examples showing the time evolution of the imaged wall will be shown in the sequel.

4.2 Simulation of the error equations

The error equations are, in principle, simulated as prescribed by equations (19) to (30). However, since we are dealing with a spatially-discretized images, it is necessary to implement these equations in a discrete form as well. There are no conceptual problems associated with replacing integrals by summations. However, all we know about the real physical image values comes from the pixels' gray-level data. It is important to understand that the gray-level of a pixel represents the value of a double integral over its area (average), where the spatially-continuous radiation emanating from the scene serves as the integrand. Another way to put it is that each pixel collects all the photons impinging anywhere within its boundaries during its integration time (interframe time).

Differentiating between a pixel's gray-level and the actual value of the scene at any (continuous) location on the image plan is important in estimating the scene values $I_1(\hat{u}, \hat{v})$ as required in (19) because $(\hat{u}, \hat{v})$ are generally non-integers. There is no such problem in estimating $I_0(u, v)$ because, by definition, we start from the pixel's center (integer) and thus take its gray-level as the best estimate of the scene value at this pixel's center. For the estimation of $I_1(\hat{u}, \hat{v})$, we use an interpolation method that looks identical to the one used for the imagery generation, although it is conceptually completely different.

Referring once more to figure 7, here is the problem. Say we have an estimate for the value of the scene at the center point of some initial pixel, that is, we have $I_0(u_0, v_0)$. This point has been mapped into location $(\hat{u}, \hat{v})$ in image I_1 , and we want to estimate $I_1(\hat{u}, \hat{v})$. The relevant information available from image I_1 is its pixel values for the four pixels shown in the figure because these are directly affected by the original scenery patch (of pixel size). We can think of the value of each such pixel as a random variable crosscorrelated with $I_0(u_0, v_0)$ in proportion with the intersected areas as defined by (32). This led us to use the rather ad hoc interpolation method:

$$\begin{aligned} I_1(\hat{u}, \hat{v}) \triangleq & (1 - \delta u)(1 - \delta v)I_1(\hat{u}_0, \hat{v}_0) + \delta v(1 - \delta u)I_1(\hat{u}_0, \hat{v}_0 + 1) \\ & + \delta u(1 - \delta v)I_1(\hat{u}_0 + 1, \hat{v}_0) + \delta u\delta v I_1(\hat{u}_0 + 1, \hat{v}_0 + 1) \end{aligned} \quad (34)$$

This method has the advantage that it yields the expected results when $(\hat{u}, \hat{v})$ take on integer values, and it provides a continuous estimate inside the convex hull defined by the values of the four nearest pixels. The same interpolation method is used for estimating the image values as well as their first and second derivatives.

4.3 Open-loop error measurements

In the first set of open-loop error simulations we investigated the error sensitivity to the scaling factor s in isolation as a function of window size. The flight trajectory used for this set is non-maneuvering and constant-velocity towards the center of the wall starting from a depth of 150 m at a speed of 1 m/fr. The set of 3 images (number 0, 12, and 24) are shown in figure 8 to demonstrate the effect of expansion as the depth decreases from 150 to 138 to 126 m. Figure 9 shows the case of a 11×11 window size which is centered on the FOE. The first and fifth frames are used for I_0 and I_1 respectively so that the baseline is $b = 4$ m. The figure shows four curves. Three curves belong to the cost-function and its first and second derivatives as derived in the previous section. The fourth curve shows the correction for s as calculated by the Newton's algorithm of (20), that is, the third component of $\{\nabla^2 J[\hat{X}(k)]\}^{-1} \nabla J[\hat{X}(k)]$. The four graphs in each figure are scaled as necessary for convenient presentation. Figure 10 and figure 11 only differ from figure 9 by the window size as indicated in their titles. Figures 12 and 13 represent contraction—as opposed to expansion—and they serve to verify symmetry in comparison with figures 10 and 11 respectively.

The following observations are noteworthy.

1. The absolute values of all four variables increase monotonically with the window size. The reason is that, since the free variable is an expansion factor, it causes each pixel of the window to shift in linear proportion to its distance from the center of the window. Thus, the larger the window, the larger are the shift errors experienced by its pixels.
2. The values of the cost-function and its first and second derivatives roughly agree; this is not obvious because each derivative is obtained directly from the corresponding image derivatives. Low-pass-filtering of the image derivatives and the fact that we deal with discrete pixel values and have to resort to interpolation, can account for the numerical disparities.
3. The actual value of s , to be denoted s_a , is shown by the vertical bars in all figures. It is noticed that, in all 5 cases it falls closer to the minima of the cost-functions than to the zero crossings of the first derivatives. We do not have a satisfactory explanation for this behavior except to assume that these are noise-like inaccuracies resulting from the quantization and interpolation operations; they clearly diminish as the window becomes larger. It warrants commenting here that it is the zero crossing of the derivative which

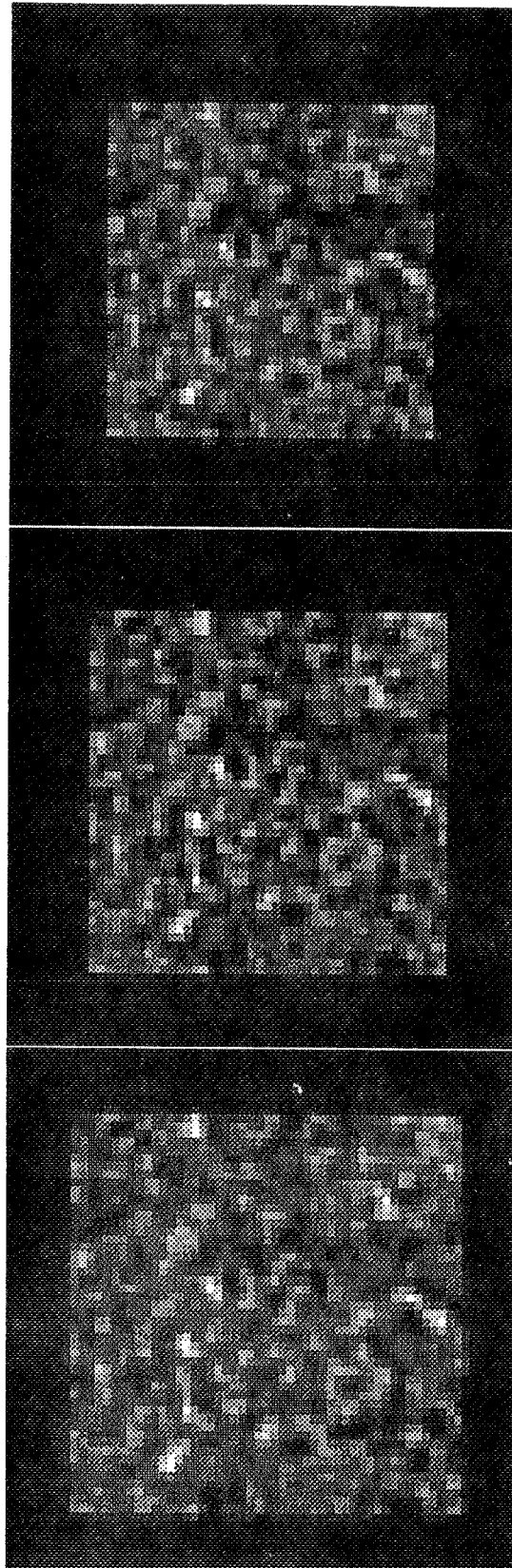


Figure 8: Frames 0, 12, and 24 of simulated textured wall seen while flying forward.

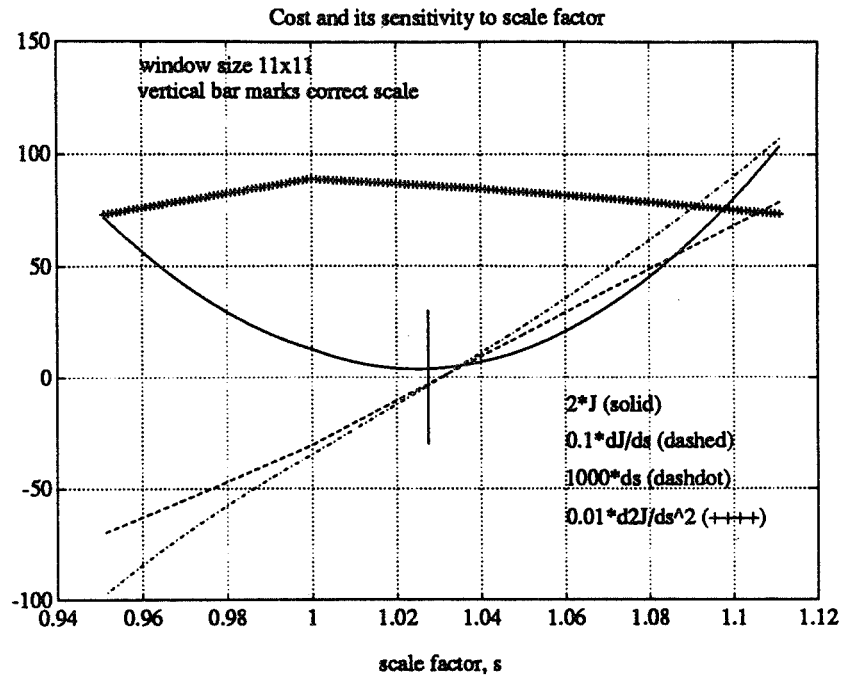


Figure 9: Sensitivity of the cost-function and its derivatives to the scale factor (11×11 window).

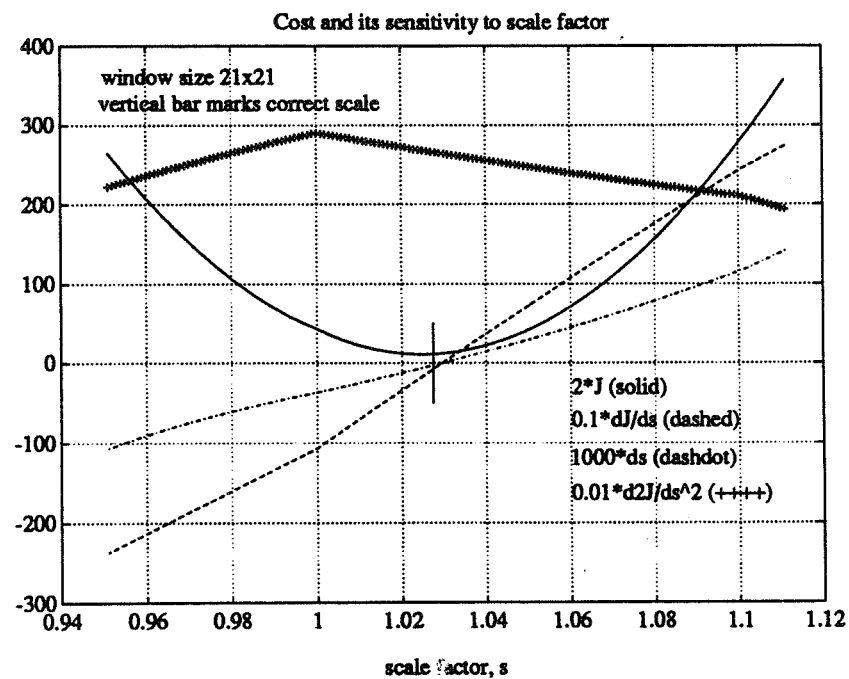


Figure 10: Sensitivity of the cost-function and its derivatives to the scale factor (21×21 window).

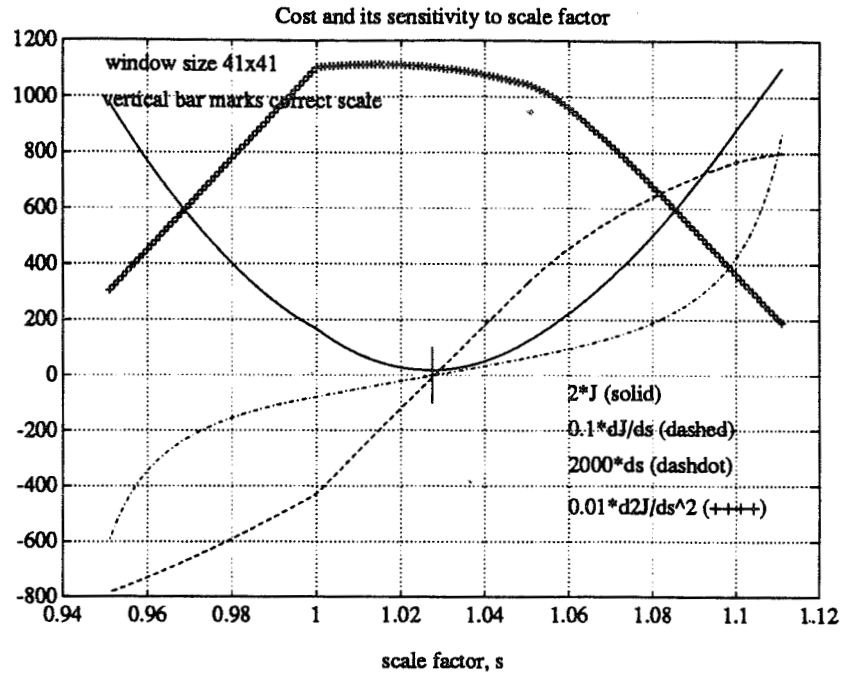


Figure 11: Sensitivity of the cost-function and its derivatives to the scale factor (41×41 window).

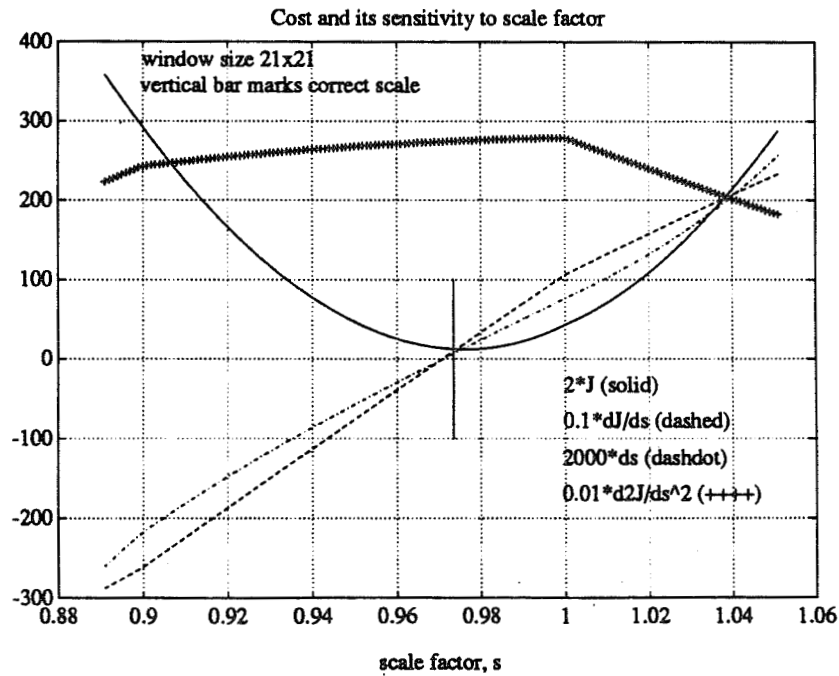


Figure 12: Sensitivity of the cost-function and its derivatives to the scale factor (21×21 window).

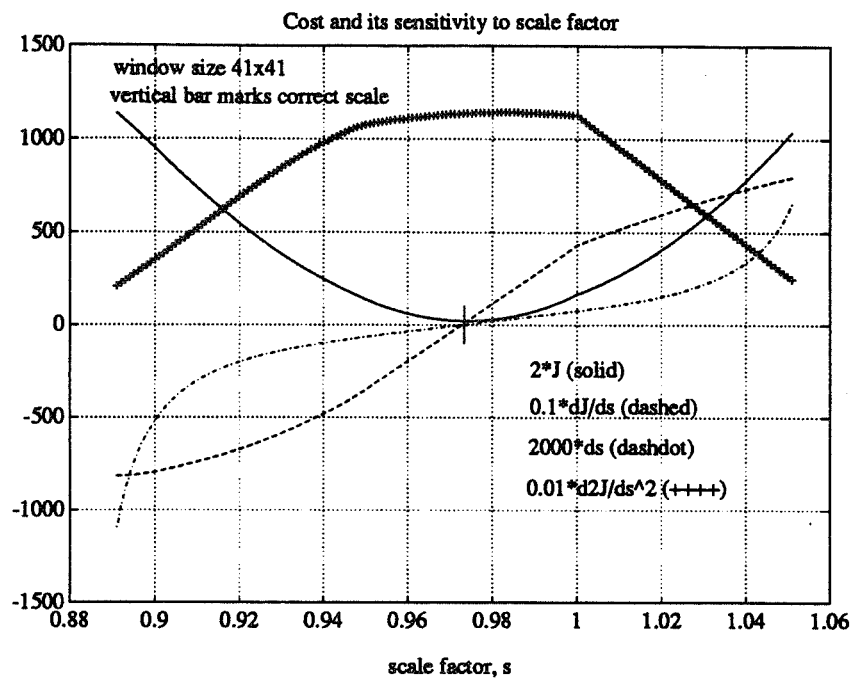


Figure 13: Sensitivity of the cost-function and its derivatives to the scale factor (41×41 window).

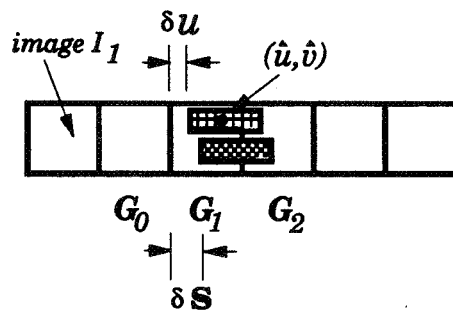


Figure 14: Interpolation example.

matters and not the minimum of the cost-function because that is where the closed-loop system would converge to.

4. The second derivative shows a sharp slope change at $s = 1$; the first derivative and the cost-function itself show corresponding behavior. The reason for that is explained by analyzing our interpolation method as shown in figure 14 for a simple one-dimensional case. The black dot represents the center point, $(\hat{u}, \hat{v})$, of one of the I_0 pixels that got shifted—as a result of expansion by some factor $s > 1$ —to its new location in image I_1 . The rectangle centered on the dot represents the original I_0 pixel. This new location is shifted by δu with respect to where it would fall if s equaled unity. Let us take the gray-level of this particular I_0 pixel as unity with all its neighbors being zeroes. This pixel will cause the gray-levels of image I_1 to become

$$G_0 = 0; \quad G_1 = (1 - \delta u); \quad G_2 = \delta u \quad (35)$$

In order to generate the error curves, we sweep the value of s over some range around $s = 1$. The lower rectangle in the figure represents the location of the corresponding swept pixel for some $s > 1$ (denoted by s_s) which is different from the actual s_a . This swept pixel is shown shifted by δs . Interpolating for the current value of $s = s_s$, we have

$$I_1(\hat{u}, \hat{v}) = G_1(1 - \delta s) + G_2\delta s = (1 - \delta u)(1 - \delta s) + \delta u\delta s = 1 - \delta s - \delta u + 2\delta u\delta s \quad (36)$$

When s sweeps through values less than unity, i.e., $s_s < 1$, we have

$$I_1(\hat{u}, \hat{v}) = G_1(1 - |\delta s|) + G_0|\delta s| = G_1(1 - |\delta s|) = (1 - \delta u)(1 - |\delta s|), \quad (37)$$

which is always less than the corresponding result for a positive δs .

We thus conclude that, for an expansion, when the actual s is larger than 1, sweeping s_s over values of $s_s > 1$ always results in $I_1(\hat{u}, \hat{v})$ larger than those resulting from symmetrical (around $s = 1$) values of $s_s < 1$. This effect becomes more pronounced as the window size increases because the window pixels are, on the average, farther from its center and they experience larger δu shifts. When the actual s is smaller than 1, we see the opposite behavior as exemplified by figures 12 and 13. In these, the actual s is $s_a = 146/150 = 0.9733$. It is important to realize that, since the closed-loop algorithm performs around s_a and not around $s = 1$, it is not affected by the above phenomenon.

5. The curves of ds give the calculated correction for the case where the error occurs (through sweeping) in s alone. In such a case, the correction part of equation (20) simplifies to

$$ds = \frac{dJ/ds}{d^2J/ds^2} \quad (38)$$

It can be seen from the figures that ds approximately agrees with this equation. Also, the discontinuities in the first and second derivatives at $s = 1$ cancel each other in (38) so that the ds graphs do not show any discontinuity.

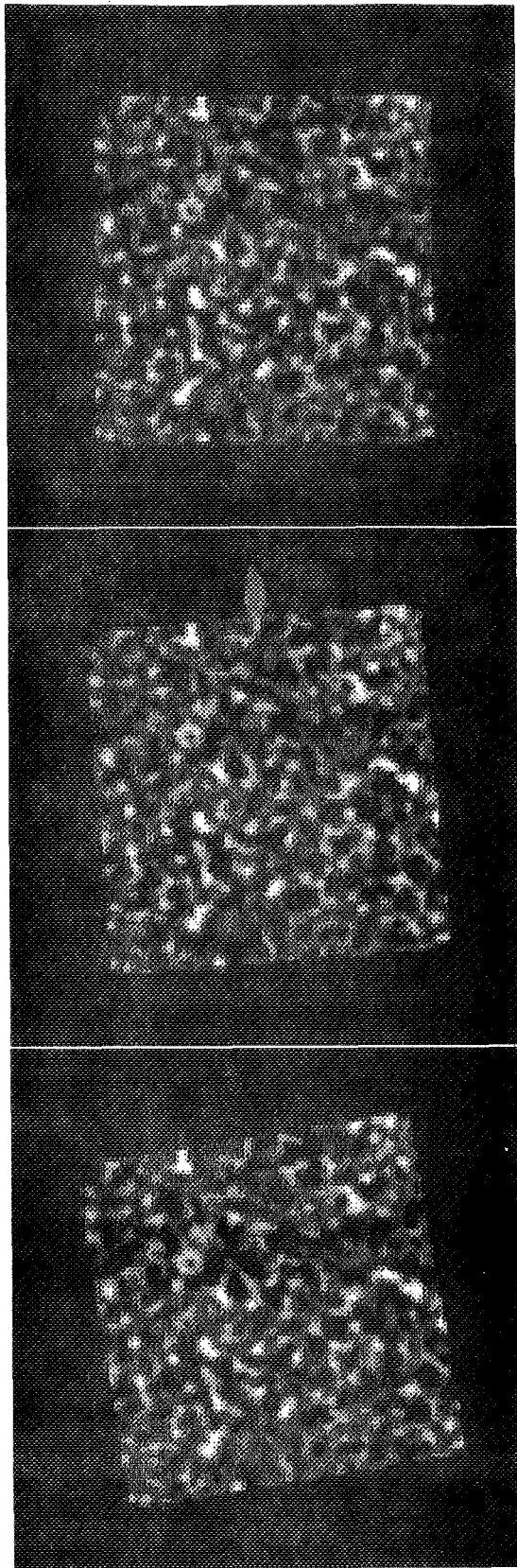


Figure 15: Frames 0, 4, and 8 of simulated textured wall as seen while rolling with no lateral motion.

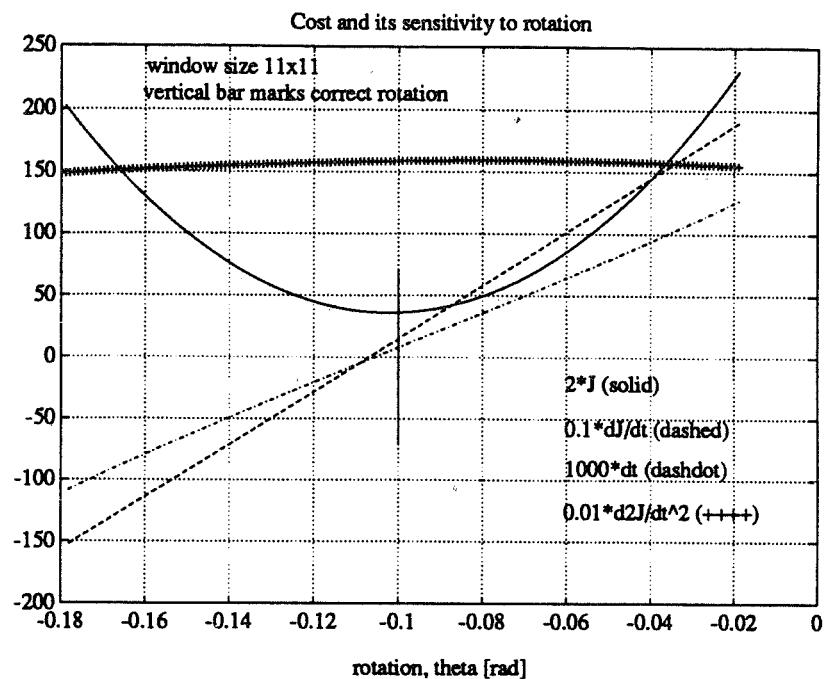


Figure 16: Sensitivity of the cost-function and its derivatives to rotation (11×11 window).

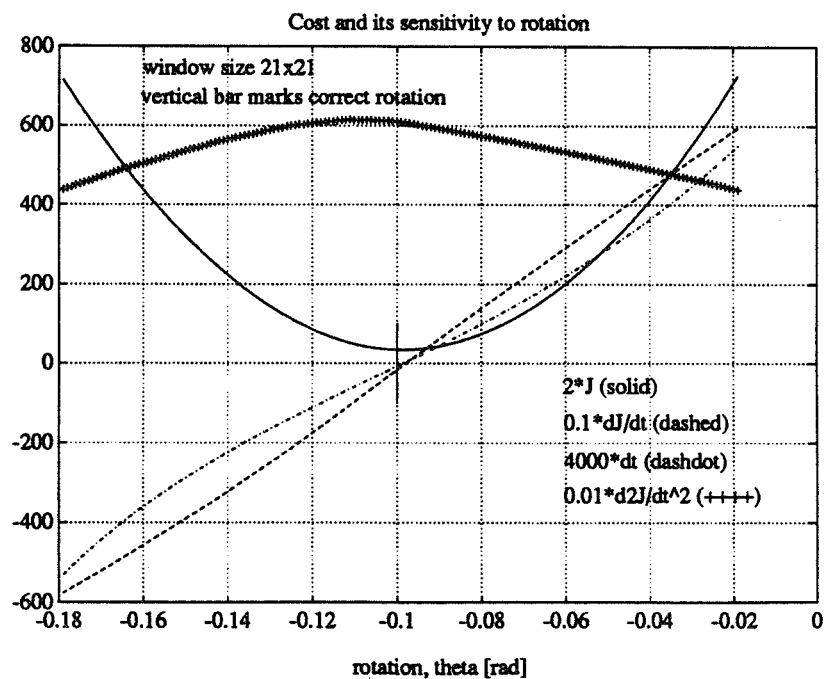


Figure 17: Sensitivity of the cost-function and its derivatives to rotation (21×21 window).

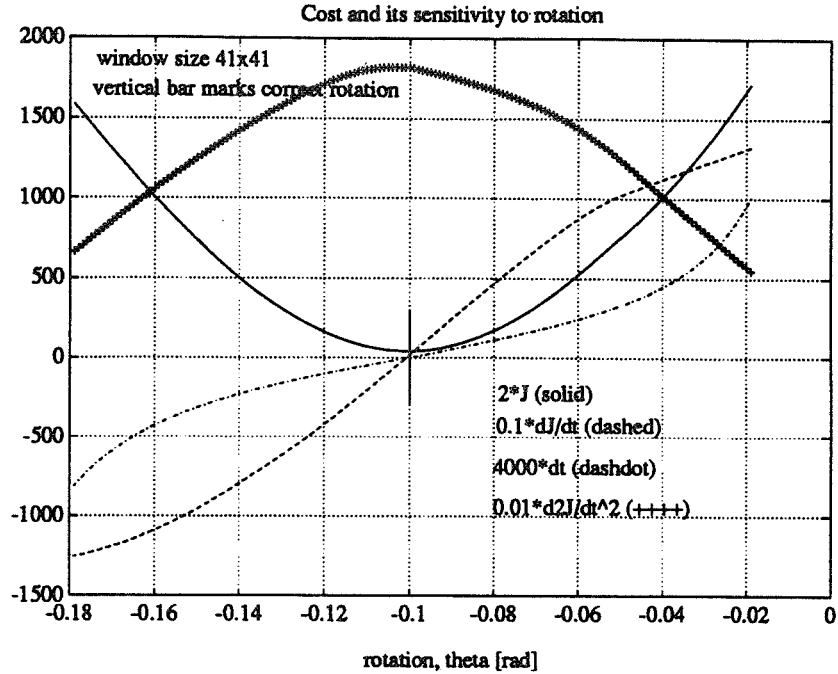


Figure 18: Sensitivity of the cost-function and its derivatives to rotation (41×41 window).

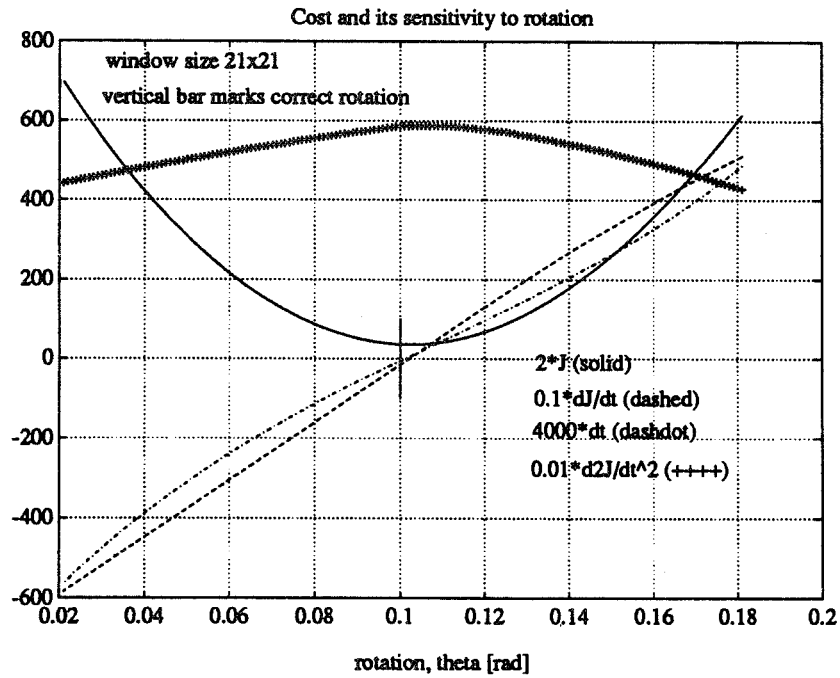


Figure 19: Sensitivity of the cost-function and its derivatives to rotation, 21×21 window, positive rotation.

In the next set of error measurements we investigated the error sensitivity to rotation angle θ in isolation as a function of window size. For this set the camera does not travel laterally; it only rolls at -0.02 rad/fr while pointing towards the center of the wall from a constant depth of 150 m. The set of 3 images (number 0, 4, and 8) are shown in figure 15 to demonstrate the effect of rotation. Figure 16 shows the case of a 11×11 window size which is centered on the FOE. Figures 17 and 18 correspond to windows of size 21 and 41 respectively. The first and sixth frames are used for I_0 and I_1 respectively so that the total roll used in generating the first 3 figures is of -0.1 rad. Figure 19 shows a roll in the opposite direction for a symmetry check. The same four curves as before are shown in all figures.

The following observations can be made.

1. The absolute values of all four variables increase monotonically with the window size. The reason here is the same that applied to the scaling-only cases. The larger the window the larger the shifts experienced by pixels which are farther from the window center.
2. The values of the cost-function and its first and second derivatives roughly agree as for the s curves.
3. The actual value of θ is shown by the vertical bars in the figures. It is noticed that the bars fall close to the minima of the cost-functions and also to the zero crossings of the first derivatives. The larger the window, the more accurate these results are.
4. There are no marked discontinuities as found in the s curves because the reason that caused it there does not apply here.
5. The curves of $d\theta$ give the calculated correction for the case where the error occurs (through sweeping) in θ alone. In such a case equation (20) simplifies to

$$d\theta = \frac{dJ/d\theta}{d^2J/d\theta^2} \quad (39)$$

In the figures $d\theta$ approximately agrees with this equation.

In the next set of error measurements we investigated the error sensitivity to image-plane shifts, a , in isolation as a function of window size. For this set the camera is stationary except that it is panning at 0.0005 rad/f while pointing towards the center of the wall from a constant depth of 150 m. Images numbers 0 and 4 are used for I_0 and I_1 respectively. The panned images are not shown because they look quite indistinguishable—being shifted only by about a pixel. Figure 20 shows the case of a 21×21 -size window (top) and 41×41 -size window (bottom) when both are centered on the FOE. The following observations can be made.

1. As opposed to the previous cases, where s or θ served to generate the errors, here there is very little sensitivity to the window size because the shifts are equal for all pixels within the window of any size.

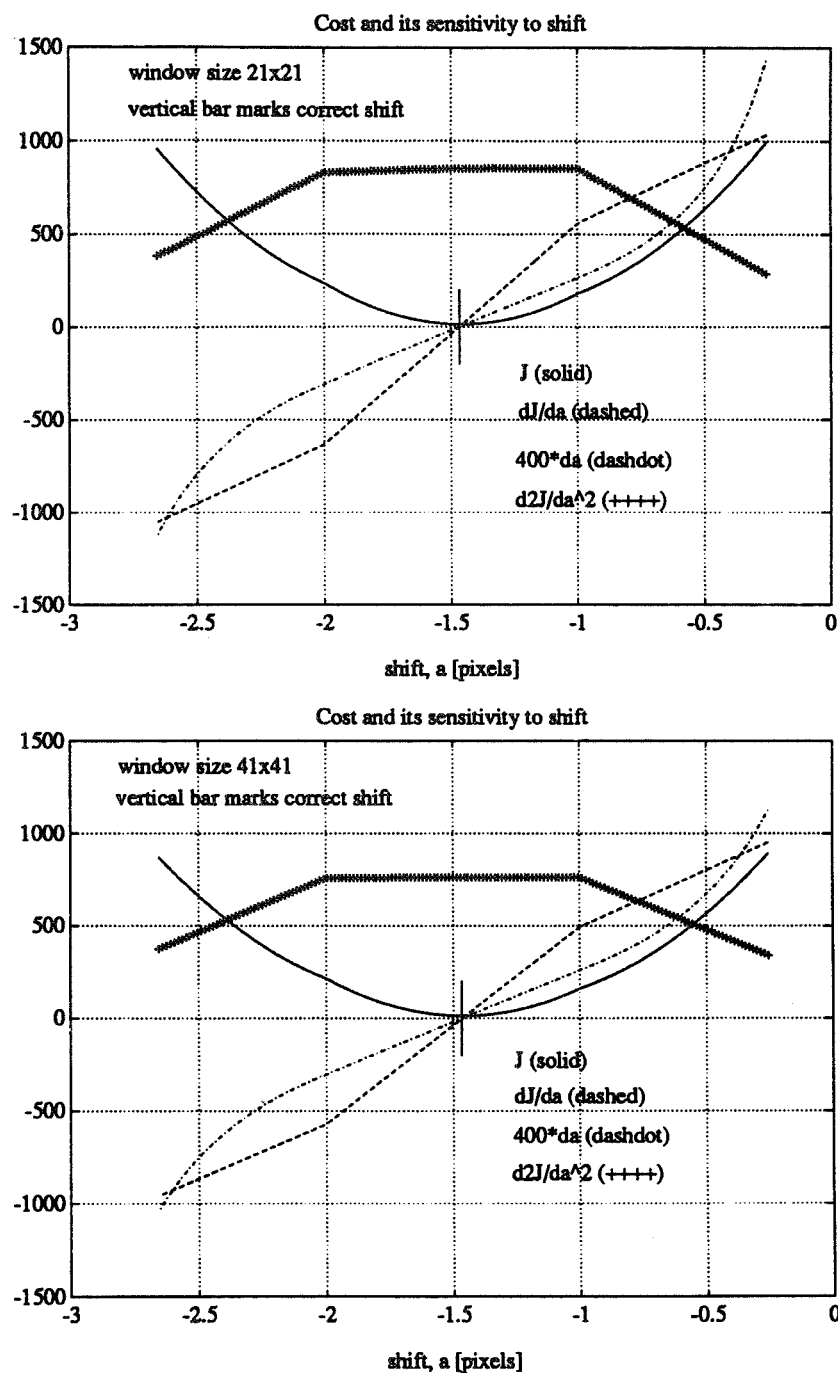


Figure 20: Sensitivity of the cost-function and its derivatives to shift; $L = 21$ (top), $L = 41$ (bot).

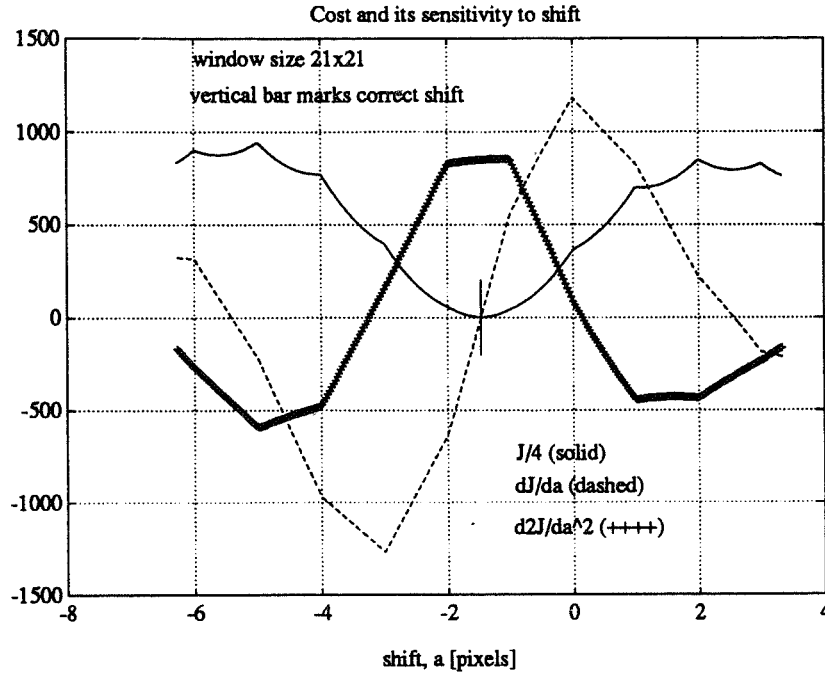


Figure 21: Sensitivity of the cost-function and its derivatives to shift over a wide range (21×21 window).

2. The actual value of a is marked by the vertical bars in all figures. These bars fall close to the minima of the cost-functions and also to the zero crossings of the first derivatives. As before, the larger the window, the more accurate these results are.
3. The second-derivative discontinuities at the integer pixel shifts can be explained by arguments similar to those used in the case of the s curves.
4. The curves of da give the calculated correction for the case where the error occurs (through sweeping) in a (or b) alone. In such a case, the correction part of equation (20) simplifies to

$$da = \frac{dJ/da}{d^2J/da^2} \quad (40)$$

In the figures, da approximately agrees with this equation.

5. Figure 21 shows the behavior of the cost-function curve for large shifts—where it becomes highly non-linear. The Newton's solution loses much of its value at such large errors. However, convergence is still possible inside the error region defined by the nearest zero-crossing of the first derivative on either side of the zero-error point (± 4 pixels here). Inside this region the correction still shows the right sign.

4.4 Closed-loop performance

In this subsection we summarize the results of closed-loop runs. These runs are divided into four groups. The first three groups parallel the open-loop cases of forward-flying, rolling, and panning (yaw). In the fourth run there are maneuvers in all variables so we could test the most general case. Within each group there are two kinds of parameters. One parameter is the window size, and the other is the location of the window with respect to the FOE.

In each run the errors are corrected using the Newton's method for six iterations. Theoretically, Newton's method should "converge" in one shot for any ideal parabolic cost-function. We allow for discrepancies from the ideal by (1) iterating on the solution more than once, (2) factoring the corrections by an experimental factor of 0.75 to prevent overshoots, and then, (3) bounding δs by ± 0.03 , $\delta\theta$ by ± 0.03 rad, and δa , δb by ± 0.75 pixels.

Each of the graphical results for all runs include five curves to show the convergence of the cost-function, J , and the four parameters: s , θ , a , and b . In addition, there are four bars (arbitrarily located between iteration number 4 and 5) whose ordinates show the ground-truth values of the four parameters for ready visual comparison. The bars are marked by the parameter symbols.

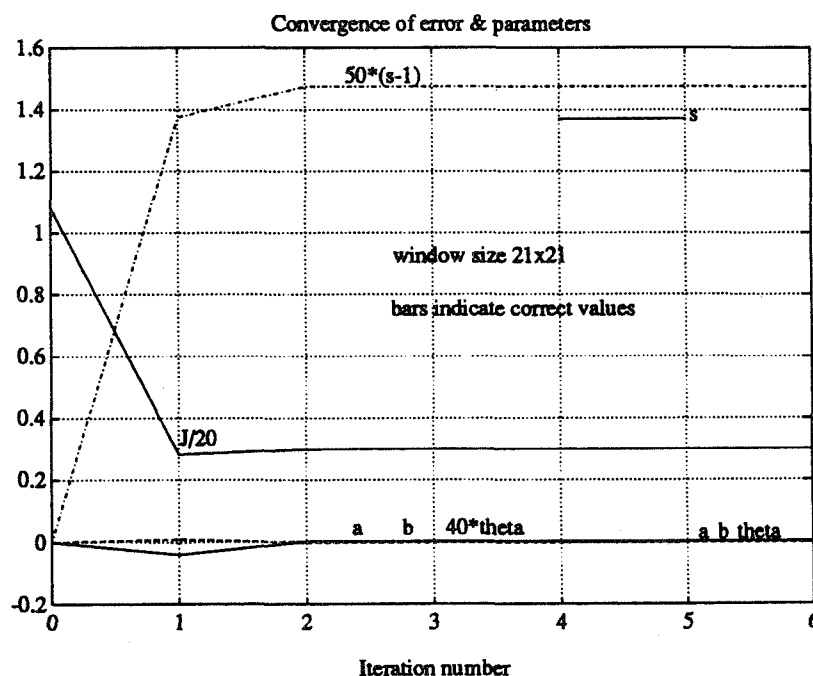


Figure 22: Convergence for forward flying and no maneuvers at the FOE (21×21 window).

Let us start with the results for forward-flying with no maneuvers. The initial depth is 150 m and the velocity is 1 m/fr towards the center of the wall. The transformation parameters are calculated at the time of frame number 4 by comparing it to frame number 0 (skipping the

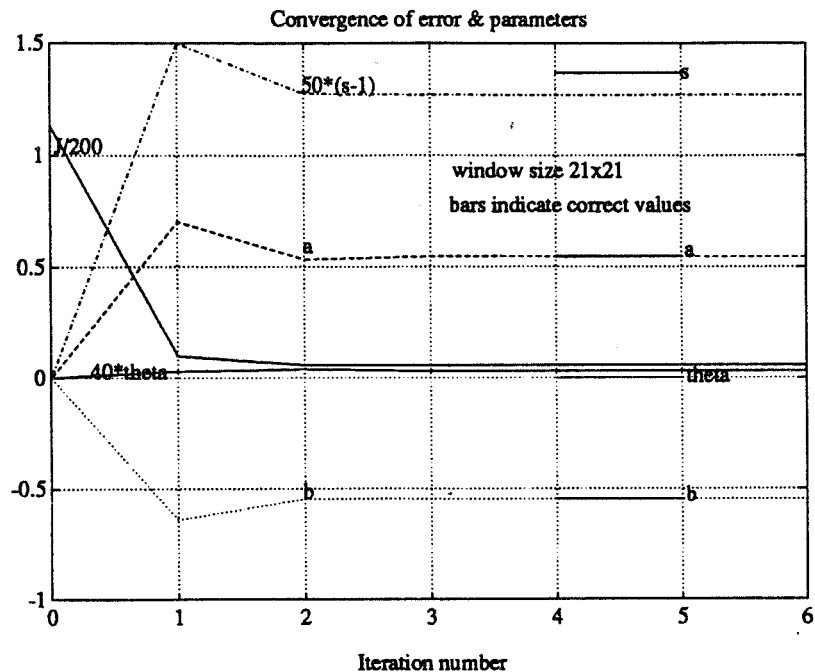


Figure 23: Convergence for forward flying and no maneuvers at (20,20) from the FOE (21×21 window).

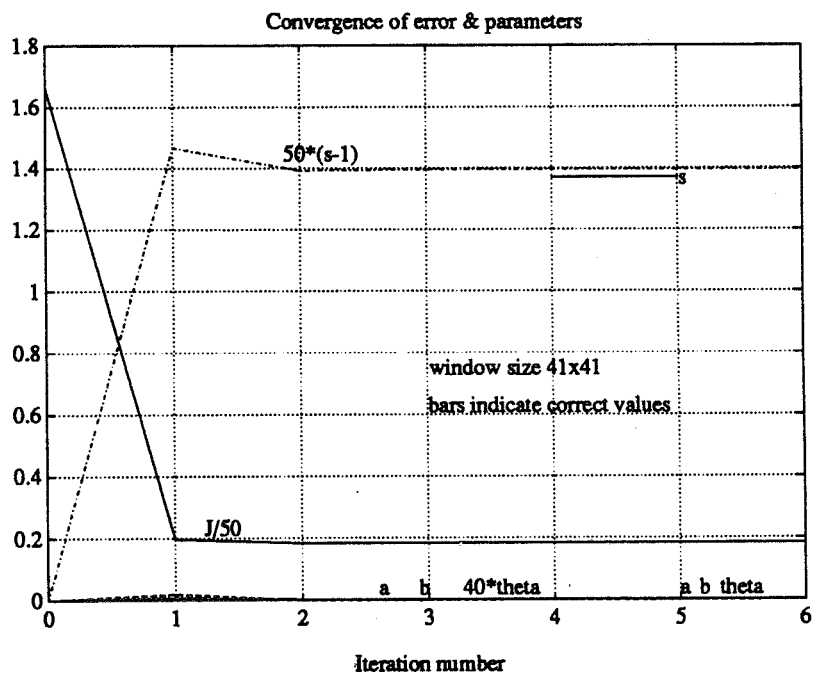


Figure 24: Convergence for forward flying and no maneuvers at the FOE (41×41 window).

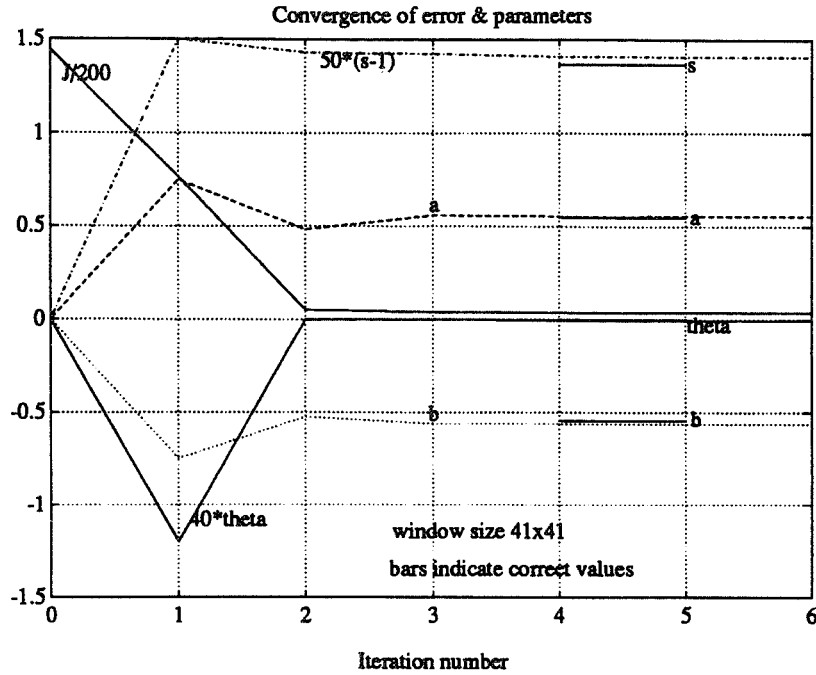


Figure 25: Convergence for forward flying and no maneuvers at (20,20) from the FOE (41×41 window).

intermediate frames). These runs are intended to demonstrate expansion alone for a window centered on the FOE, and expansion-plus-shift for a window centered on the point (20,20) with respect to the FOE. The following observations can be made:

1. The cost-function and all parameters practically converge in two iterations. When no parameter correction hits its bounds, convergence is achieved in a single iteration.
2. The accuracies—especially for s —improve noticeably as the window size doubles (4 times the window area), but they are still very good for the 21×21 -size window. For example, from figure 22, the correct expansion (indicated by the s bar) is $150/146=1.0274$, which corresponds to 146 frames-to-collision, whereas the converged value is $s = 1.0296$ which corresponds to 135 frames-to-collision.
3. The converged shifts for the (20,20) point practically show no error. This is especially impressive because these shifts are small—only (0.548,0.548) pixels.

Next, we present the results for roll-only flying without any forward or lateral motion. The depth is constant at 150 m. The transformation parameters are calculated at the time of frame number 2 by comparing it with frame number 0. the roll-angle difference is 0.04 rad between these two frames. In these runs we demonstrate rotation alone for a window centered on the FOE, and rotation-plus-shift for a window centered on the point (20,20) with respect to the FOE. The following observations can be made:

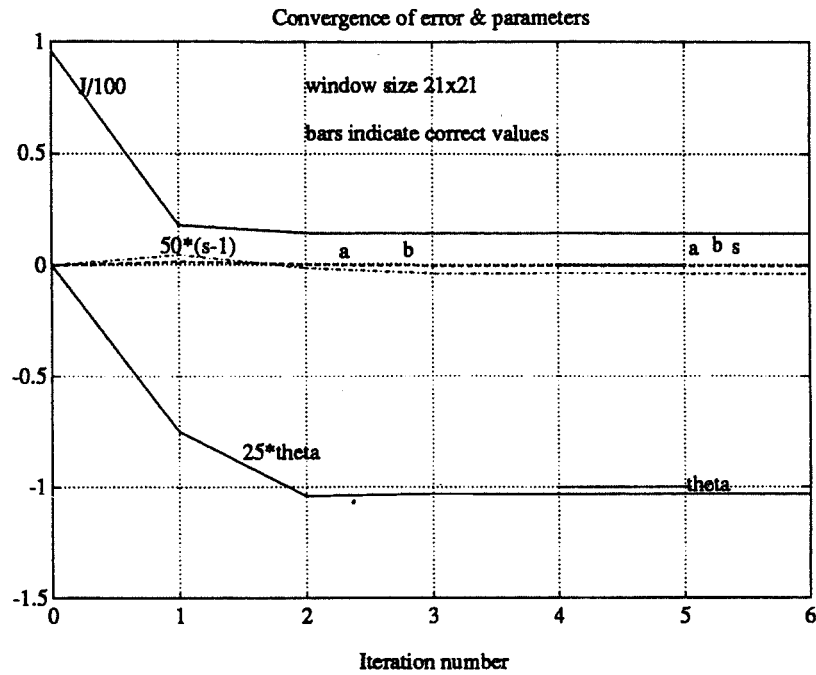


Figure 26: Convergence for roll-only maneuver at the FOE (21×21 window).

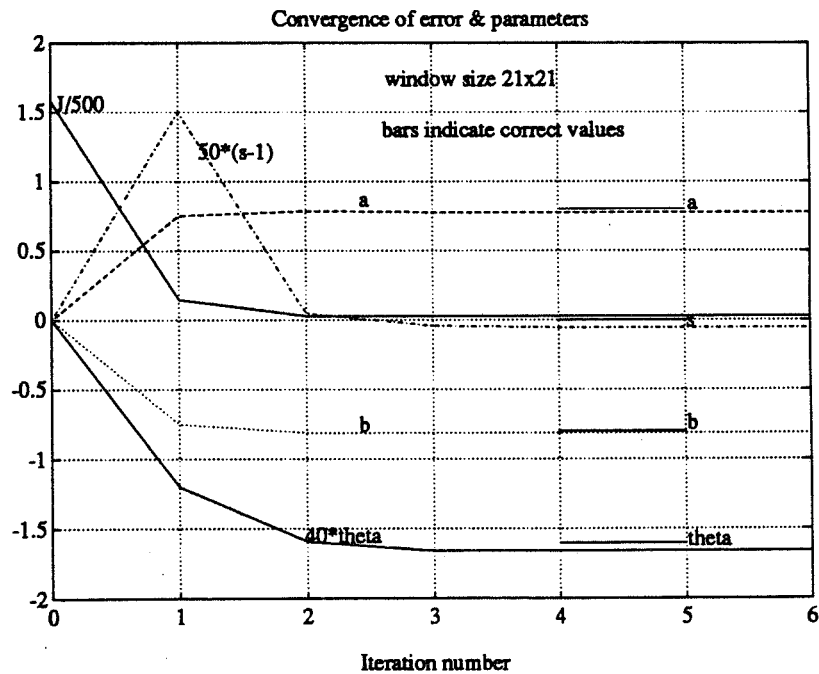


Figure 27: Convergence for roll-only maneuver at (20,20) from the FOE (21×21 window).

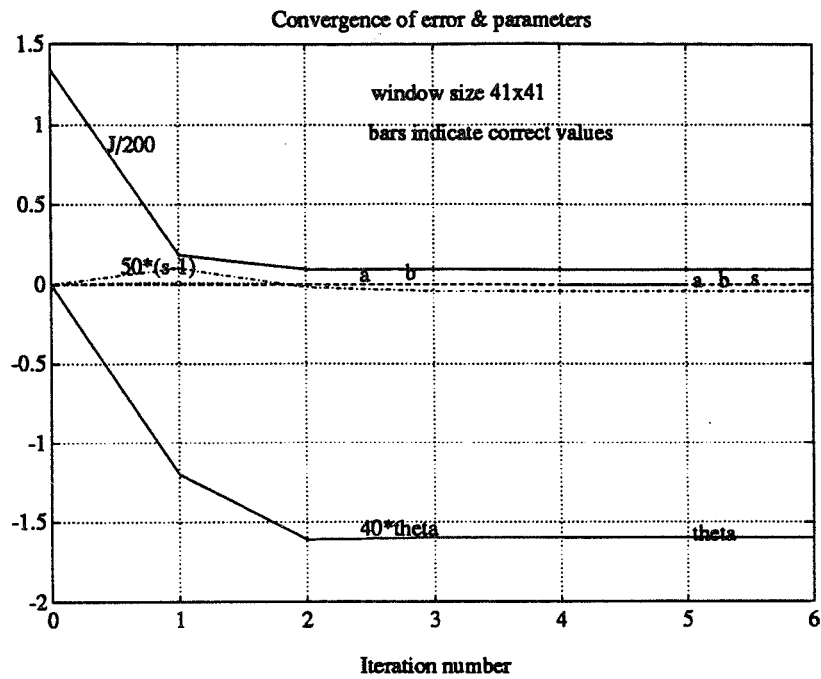


Figure 28: Convergence for roll-only maneuver at the FOE (41×41 window).

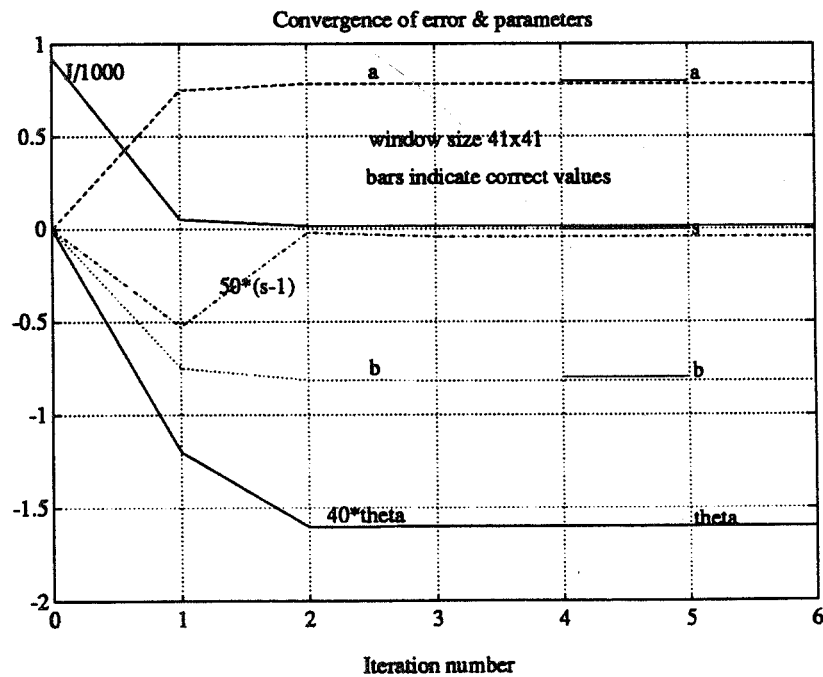


Figure 29: Convergence for roll-only maneuver at (20,20) from the FOE (41×41 window).

1. As before, the system practically converge within two iterations.
2. Although the cost-function—especially in figure 26—does not converge as close to zero as in all other case, the parameters still converge accurately to their respective values.
3. The accuracies improve noticeably as the window size doubles. From figure 24 and figure 25, θ virtually has zero error, while its error increases to 3.6% for the 21×21 window.
4. The expansion shows a transient for the (20,20) point, but it settles to zero after 2 iterations.
5. The converged shifts at the (20,20) point are remarkably close to the correct ones of (0.8,0.8) pixels.

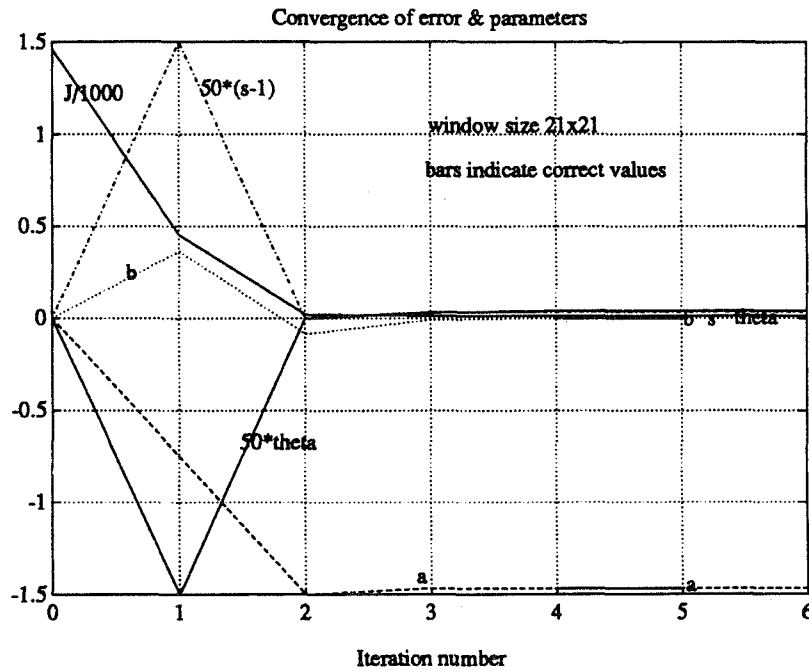


Figure 30: Convergence for yaw-only maneuver at the FOE (21×21 window).

Next, we present the results for yaw-only flying with no forward or lateral motion. The depth is constant at 150 m. The transformation parameters are calculated at the time of frame number 4 by comparing it with frame number 0; the yaw-angle difference is 0.002 rad. We translate this yaw angle by using the fact that, in our Flight/Vision simulation, the camera's FOV is taken as 10 degrees, and it corresponds with an image of size 128×128 . This means that the expected shift is $\delta u = 1.467$ pixels. Thus, in these runs, we demonstrate δu -shift alone for a window centered on the FOE or on the point (26,26) with respect to the FOE. The following observations can be made:

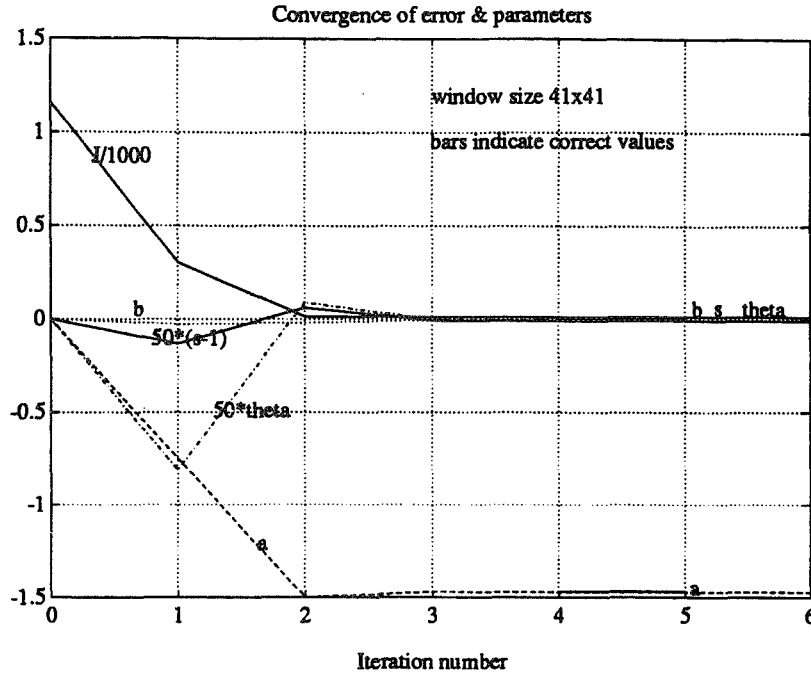


Figure 31: Convergence for yaw-only maneuver at (26,26) from the FOE (41×41 window).

1. Irrespective of the window size, or the location of the image-point with respect to the FOE, all converged parameters are close to being error free.
2. The expansion and rotation show transients which decay to zero after two iterations.

Lastly, we present the results for a general maneuver where the velocity is 1 m/s (starting from 150 m depth), pitch and yaw rates are 0.0005 rad/s each, and the roll-rate is 0.02 rad/s. The transformation parameters are calculated at the time of frame number 2 in figures 32, 33, and 35, and at frame number 4 in figure 34 by comparison with frame number 0. The following observations can be made:

1. The system converges within two iterations.
2. Generally, the accuracies improve with the window size.
3. The accuracy of s is around 6% for the FOE point—irrespective of the window size (21 to 61)—and it drops to 16% for the (20,20) point.

Summarizing the simulation results, we can conclude that the basic idea and algorithm are solid and perform very well. Although these simulations were done in apparently noise-free situation, they do get affected by the noise inherent in the pixel quantization.

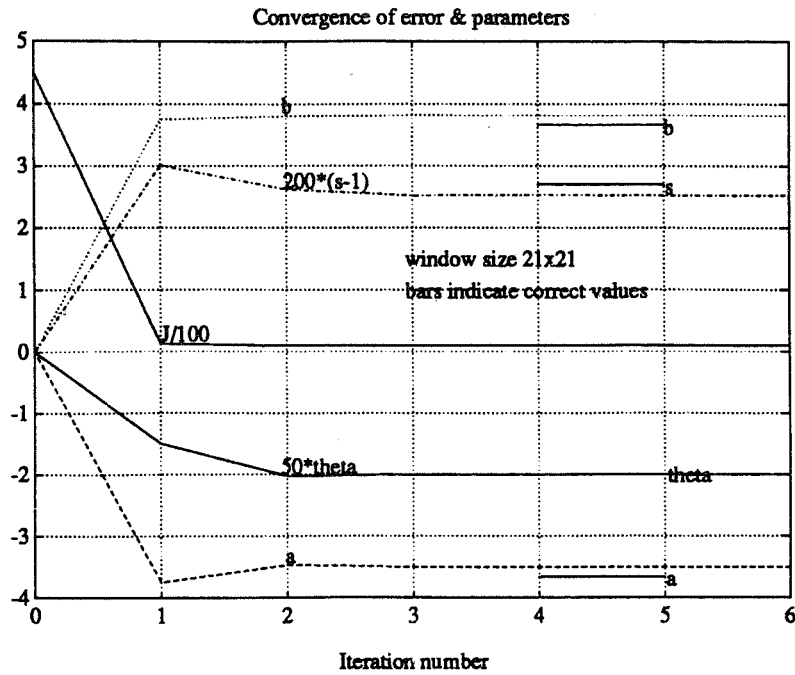


Figure 32: Convergence for general maneuvers at the FOE (21×21 window, 2-frames difference).

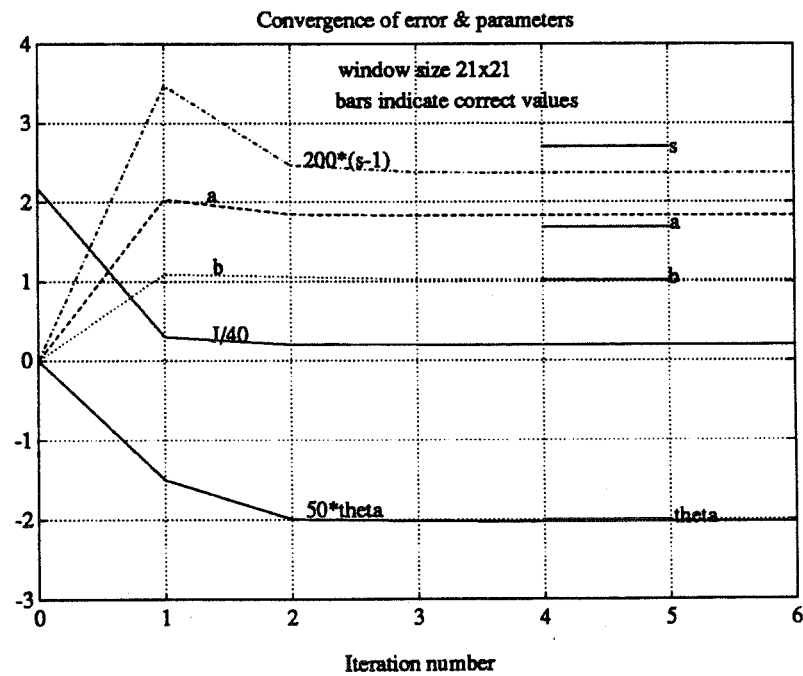


Figure 33: Convergence for general maneuvers at (20,20) from the FOE (21×21 window, 2-frames difference).

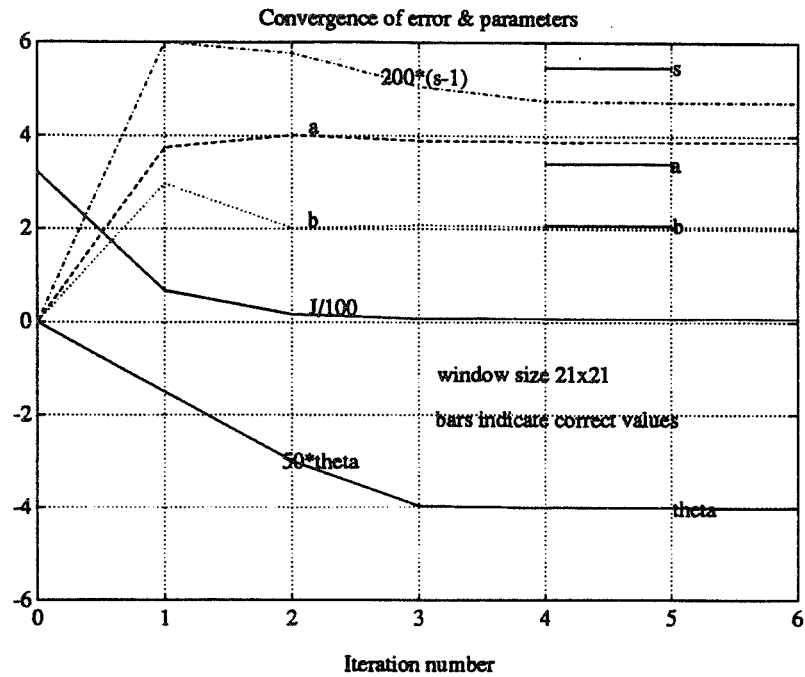


Figure 34: Convergence for general maneuvers at (20,20) from the FOE (21×21 window, 4-frames difference).

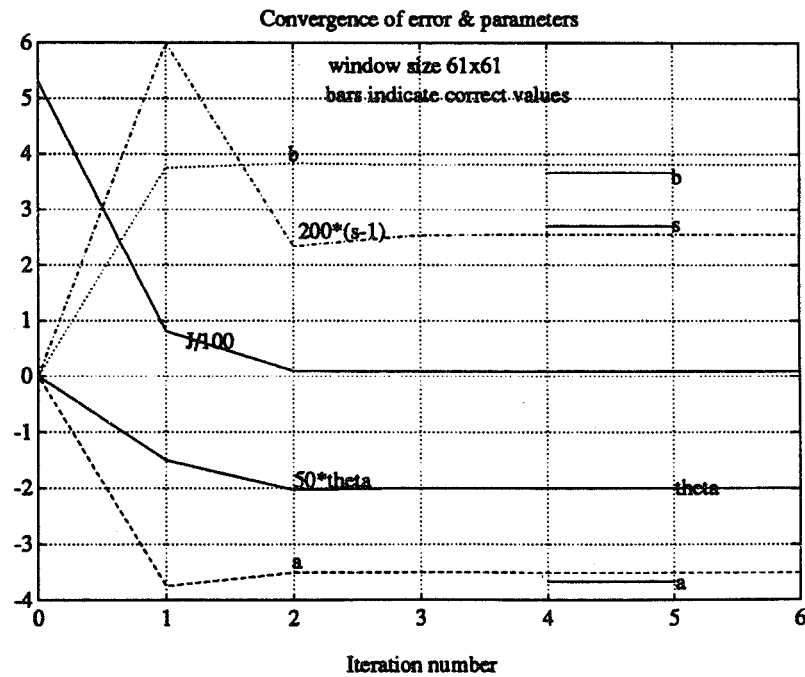


Figure 35: Convergence for general maneuvers at the FOE (61×61 window, 2-frames difference).

5 INCREASING THE TRIANGULATION BASE-LINE

In this section we use the above algorithm as the core on which a farther layer is to be built with the intention of increasing the accuracy and robustness of the practical algorithm. The implicit assumption here is that the flight trajectory is basically non-maneuvering, or, in other words, it is the maneuvers which will determine the maximum usable triangulation baseline.

5.1 The capture zone

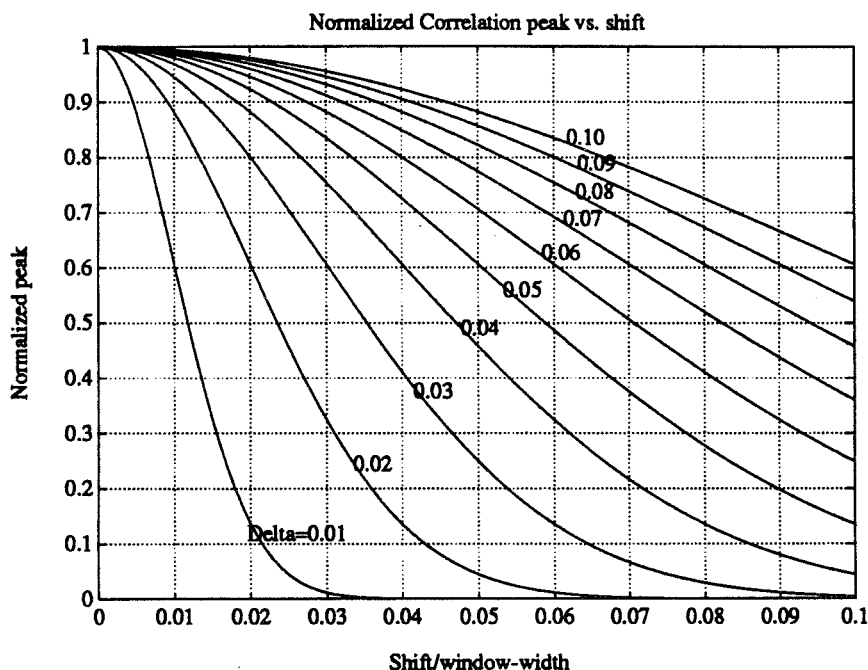


Figure 36: Average normalized correlation peak vs. shift, Delta in image-width fraction.

We have touched on the question of convergence in regard to figure 21. In that figure the “capture zone” is of ± 4 pixels—meaning that, as long as the error is within this zone, it always has the correct sign to drive it towards the stable solution. Thus, convergence is assured inside this zone, although its width is not usually known—especially when more than a single parameter is involved. It is possible, however, to estimate some lower bounds on the capture zone for each one of the four parameters. Estimating the width of the capture zone is based on the bandwidth or correlation width of the images. For that, we used $\Delta = 1.5$ pixels in conjunction with figures 5, 6, 36, and 37. What it means is that image-plane locations 1.5 pixels apart have gray-levels correlated with a correlation coefficient of $\exp\{-0.50\} = 0.606$ (see (16)).

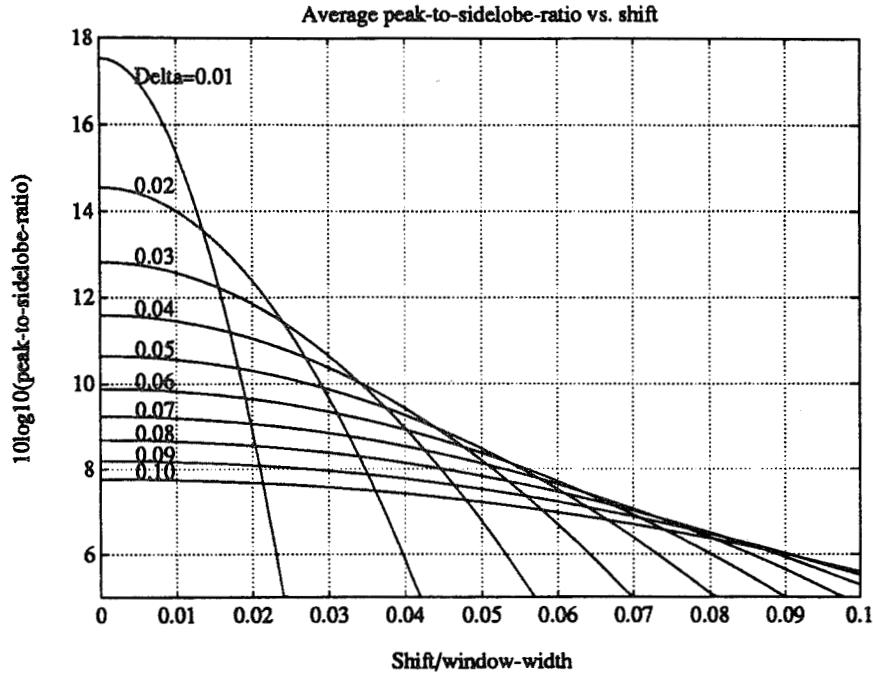


Figure 37: Average peak-to-sidelobe-ratio vs. shift, Delta in image-width fraction.

To estimate the capture zone, we arbitrarily assume that a PSR=7.5 is acceptable to provide a high enough probability of detecting the correct correlation peak and a low enough probability of false alarm (locking onto a wrong peak). This figure is equivalent to 15 dB in power ratios. Let us assume that the window size is 21×21 ; then Δ of 1.5 pixels is ≈ 0.07 of the window-size. From the corresponding graph in figure 37 we read that a PSR=7.5 is achieved for shifts less than 0.063 of the window size, i.e., ± 1.32 pixels. Repeating this exercise for $\Delta = 2$ would result in a smaller capture zone of only 0.97 pixels.

A word about figures 36 and 37 is now in place. Figure 36 shows that the correlation peak drops slowly with the shift when Δ is large—as expected from (16). However figure 37 shows that the higher the Δ , the higher the PSR's initial value is, and the sharper its drop. This result is attributed to the fact that, when Δ is large, the effective number of independent image areas (objects) decreases. That has no effect on the mean correlation peak but it increases the sidelobes variance. The sidelobes variance of the cross-correlation, $C(\tau_u, \tau_v)$, is given by equation (A19) of [31],

$$\begin{aligned} \text{var}\{C(\tau_u, \tau_v)\} = & L^{-2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(u, v) R(u, v) R(\hat{u}, \hat{v}) du dv + \\ & L^{-2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(u, v) R(u - \tau_u, v - \tau_v) R(u + \tau_u, v + \tau_v) du dv, \end{aligned} \quad (41)$$

where

$$g(u, v) = \begin{cases} (1 - |u/L|)(1 - |v/L|), & (u, v) \in [-L, L] \times [-L, L] \\ 0 & \text{otherwise,} \end{cases} \quad (42)$$

is a triangular window that weighs the integrand, L^2 is the window area used for normalization, and $(\hat{u}, \hat{v})$ are the transformed (u, v) of (10). Far from the crosscorrelation peak, only the first term of (41) prevails, and that is what we used for constructing the above figures.

To estimate the capture zone for the expansion factor, s , and the rotation, θ , we refer back to figure 5. For the same case $L/\Delta = 14$, and we see from the figure that this is achieved with $d = 0.148$ which is equivalent to 8.5° of rotation or $s = 1.148$ of expansion. Overall, we have shown that the capture zone is quite wide, and there is some optimal window size that can be chosen for any given correlation width. Images from real scenes are highly non-stationary in the sense that Δ might be small for one part of the image and large for another. However it can never be smaller than the PSF which is why we used $\Delta = 1.5$ as a PSF-width estimate.

5.2 The iterative algorithm

In the iterative algorithm we start with frames which are close enough in time to ensure that the errors in the four parameters fall inside the worst-case capture zone. Let us say that we initially use frame-0 and frame-1, so the frame separation is one. The Newton's equations are iterated upon until the error converges. The converged parameters are then used to predict the initial values for a larger frame separation, say, between frame-0 and frame-4 (notice that the first frame of the pair is fixed here). The same is now repeated for this new frame separation. Thus, there are two nested iteration loops; the inner one iterates on the Newton's equations until convergence is achieved for some fixed frame separation; the outer loop iterates through increased frame separation. The algorithm can be summarized by the following pseudo code.

```

frame_separation = 1;
frame_0 = 0;
frame_1 = frame_0 + frame_separation;

while(frame < last_frame) {
    while(error has not converged) {
        solve Newton's eqs. to update
        a, b, s, theta;
    }
    if(final error is low) {
        increase frame separation;
        frame_1 = frame_0 + frame_separation;
        save last parameter values;
        predict initial parameter values for new frame separation;
    }
    else { declare previous-iteration results as final; }
}

```

}

When running in batch for some fixed number of frames, the outer iteration loop must stop when the frame separation cannot be increased any more. Another condition to stop is that the converged results of the last frame-separation are not satisfactory, as judged by some criteria.

The prediction of initial parameter values for the next (larger) frame separation is calculated from the converged parameters of the previous frame separation using the projection equations (31). Let us project an object of length l onto the image plane so that its projection is defined as unity. After decreasing the depth from z_0 to z_1 , the projection changes to s_1 . For a frame separation of t_1 , that can be written as

$$1 = \frac{fl}{z_0}; \quad s_1 = \frac{fl}{z_1}; \quad z_1 = z_0 - V_Z t_1, \quad (43)$$

from which

$$s_1 = \frac{z_0}{z_0 - V_Z t_1}; \quad z_0 = \frac{s_1 V_Z t_1}{s_1 - 1} \quad (44)$$

Rewriting the last equation for some s_2 , t_2 instead of for s_1 , t_1 , and solving for s_2 , we get

$$s_2 = \frac{s_1 t_1}{t_2 - s_1(t_2 - t_1)} \quad (45)$$

This is how the current expansion estimate (for the current frame separation) is used to predict the expansion estimate for a larger frame separation, t_2 . The other three parameters are predicted based on linear extrapolation, so that

$$a_2 = a_1 t_2 / t_1; \quad b_2 = b_1 t_2 / t_1; \quad \theta_2 = \theta_1 t_2 / t_1 \quad (46)$$

After the algorithm stops, (44) is used to calculate the current best estimate of the *initial* depth z_0 based on the last pair of s_i , t_i which corresponds to the largest triangulation baseline that yielded convergence.

5.3 Performance of the iterative algorithm

First we ran the iterative algorithm on our simulated imagery, and then on some real imagery.

Let us start with a typical run on the simulated imagery. It is a non-maneuvering, forward-flying case with velocity of 2 m/s. The first frame pair is made up of frame-0 and frame-2. The window of size 21×21 is initially centered on pixel (74,74) which is 10 pixels away from the FOE (which is at (64,64)) in u and v . There are 40 frames in the set. The following screen output reports progress in the estimation of the initial depth of 150 m. Each table-like block of numbers reports the convergence of the inner loop for the current frame separation. The inner-loop iteration number is k and the error is denoted by err .

Opened forward_flying frame 0

Opened forward_flying frame 2

k,a,b,s,theta err = 0	0.000000	0.000000	1.000000	0.000000	267.672333
k,a,b,s,theta err = 1	0.440236	0.406381	1.030000	-0.005372	141.133240
k,a,b,s,theta err = 2	0.236247	0.248784	1.021997	0.000912	79.217903
k,a,b,s,theta err = 3	0.293832	0.275524	1.020823	-0.000823	83.259949
k,a,b,s,theta err = 4	0.278676	0.270192	1.020453	-0.000340	82.172791
k,a,b,s,theta err = 5	0.283285	0.271448	1.020556	-0.000496	82.425667
k,a,b,s,theta err = 6	0.281893	0.271135	1.020525	-0.000446	82.349678
k,a,b,s,theta err = 7	0.282312	0.271216	1.020533	-0.000462	82.372147
k,a,b,s,theta err = 8	0.282187	0.271194	1.020531	-0.000457	82.365257
k,a,b,s,theta err = 9	0.282224	0.271201	1.020531	-0.000458	82.367386
k,a,b,s,theta err = 10	0.282213	0.271199	1.020531	-0.000458	82.366638

Current estimate of initial depth = 198.825650

Opened forward_flying frame 5

k,a,b,s,theta err = 0	0.705533	0.677997	1.052959	-0.001145	74.528183
k,a,b,s,theta err = 1	0.708271	0.690493	1.069563	-0.000779	55.280674
k,a,b,s,theta err = 2	0.712015	0.685594	1.065833	-0.000990	55.973583
k,a,b,s,theta err = 3	0.710209	0.685661	1.066386	-0.000896	55.718666
k,a,b,s,theta err = 4	0.710757	0.685674	1.066294	-0.000920	55.761593
k,a,b,s,theta err = 5	0.710619	0.685669	1.066310	-0.000915	55.753265
k,a,b,s,theta err = 6	0.710652	0.685670	1.066307	-0.000916	55.754974
k,a,b,s,theta err = 7	0.710644	0.685671	1.066308	-0.000916	55.754597

Current estimate of initial depth = 160.812401

Opened forward_flying frame 10

k,a,b,s,theta err = 0	1.421288	1.371342	1.142033	-0.001831	73.417000
k,a,b,s,theta err = 1	1.565485	1.555250	1.153044	0.000626	28.408524
k,a,b,s,theta err = 2	1.531177	1.529034	1.151938	-0.000007	27.785374
k,a,b,s,theta err = 3	1.540096	1.533858	1.152508	0.000302	27.518284
k,a,b,s,theta err = 4	1.537325	1.532594	1.152262	0.000157	27.572006
k,a,b,s,theta err = 5	1.538222	1.532950	1.152353	0.000221	27.549091
k,a,b,s,theta err = 6	1.537919	1.532843	1.152320	0.000194	27.556232
k,a,b,s,theta err = 7	1.538024	1.532876	1.152332	0.000205	27.553633
k,a,b,s,theta err = 8	1.537986	1.532865	1.152328	0.000200	27.554605

k,a,b,s,theta err = 9 1.538000 1.532869 1.152329 0.000202 27.554232

Current estimate of initial depth = 151.294483

Opened forward_flying frame 16

k,a,b,s,theta err = 0 2.460800 2.452590 1.268244 0.000323 122.735245
k,a,b,s,theta err = 1 2.784964 2.769740 1.269186 -0.001672 24.976557
k,a,b,s,theta err = 2 2.683191 2.691992 1.270512 0.000779 17.046618
k,a,b,s,theta err = 3 2.703243 2.702418 1.268434 0.000409 16.436787
k,a,b,s,theta err = 4 2.697401 2.699811 1.268962 0.000566 16.535572
k,a,b,s,theta err = 5 2.698910 2.700440 1.268833 0.000553 16.499908
k,a,b,s,theta err = 6 2.698529 2.700282 1.268864 0.000553 16.508102
k,a,b,s,theta err = 7 2.698624 2.700322 1.268857 0.000553 16.506060
k,a,b,s,theta err = 8 2.698600 2.700312 1.268859 0.000553 16.506615
k,a,b,s,theta err = 9 2.698606 2.700315 1.268858 0.000553 16.506516

Current estimate of initial depth = 151.021904

Opened forward_flying frame 22

k,a,b,s,theta err = 0 3.710583 3.712933 1.411131 0.000760 268.700623
k,a,b,s,theta err = 1 4.293034 4.237146 1.417225 -0.000956 30.175304
k,a,b,s,theta err = 2 4.125841 4.143232 1.415531 -0.001404 8.244106
k,a,b,s,theta err = 3 4.139481 4.153663 1.415352 -0.000450 8.060862
k,a,b,s,theta err = 4 4.137228 4.152442 1.415281 -0.000596 8.049483
k,a,b,s,theta err = 5 4.137534 4.152607 1.415305 -0.000568 8.049872
k,a,b,s,theta err = 6 4.137496 4.152581 1.415299 -0.000574 8.049752

Current estimate of initial depth = 149.947839

Opened forward_flying frame 28

k,a,b,s,theta err = 0 5.265904 5.285103 1.596075 -0.000731 522.403687
k,a,b,s,theta err = 1 6.015904 6.035103 1.601842 -0.001724 25.510233
k,a,b,s,theta err = 2 5.956887 5.961528 1.600377 0.000404 18.381941
k,a,b,s,theta err = 3 5.961518 5.965533 1.599840 0.000178 18.452717
k,a,b,s,theta err = 4 5.960965 5.965257 1.599872 0.000228 18.443174
k,a,b,s,theta err = 5 5.961043 5.965267 1.599865 0.000224 18.443617
k,a,b,s,theta err = 6 5.961033 5.965264 1.599866 0.000224 18.443457

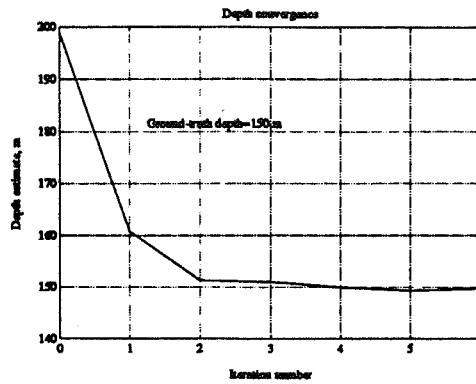


Figure 38: Depth convergence with iterations (increased triangulation baseline).

Current estimate of initial depth = 149.354233

Opened forward_flying frame 34

```
k,a,b,s,theta err = 0 7.238398 7.243535 1.835851 0.000272 938.172974
k,a,b,s,theta err = 1 7.988398 7.993535 1.834242 -0.009328 124.417595
k,a,b,s,theta err = 2 8.284004 8.308021 1.832178 -0.000755 30.430639
k,a,b,s,theta err = 3 8.285615 8.306216 1.832047 -0.000582 30.408218
k,a,b,s,theta err = 4 8.285925 8.305839 1.831991 -0.000556 30.394806
k,a,b,s,theta err = 5 8.285982 8.305765 1.831979 -0.000549 30.392294
k,a,b,s,theta err = 6 8.285994 8.305748 1.831976 -0.000548 30.391562
k,a,b,s,theta err = 7 8.285996 8.305745 1.831976 -0.000548 30.391680
```

Current estimate of initial depth = 149.733168

Final estimate of initial depth = 149.733168

There are a few interesting observations to make:

1. Frame-0 is always used as the basis for comparison —initially with frame-2, then with frames 5, 10, 16, 22, 28, and 34. The depth estimate improves with the frame separation as shown in figure 38
2. Notice that the first line of each block represents the initial conditions for a , b , s , and θ . In the first block, these are 0.0, 0.0, 1.0, 0.0 because we do not know any better. The last line of each block represents the converged values which are used to predict the initial conditions for the next block.
3. The error in each block starts from some value and usually drops and stabilizes. If the initial guess falls far from the minimum but inside the capture zone, then the error starts

from a large value and drops sharply. If the initial guess happened to be good, then the errors are already “converged”; this is exemplified by the second block belonging to the frame pair (0,5).

4. The final result was obtained from the image pair (0,34)—which does not necessarily represent the maximum frame separation possible. We have thus effectively used a triangulation baseline of 68 m which constitutes a substantial fraction of the initial depth of 150 m. This is the reason why we regard this algorithm as a track-before-detect one. In this example, the accuracy of the final result is 0.178 percent.

We have run the algorithm on various other simulated cases—at and around the FOE. Generally, the depth accuracies are better than 2%, and they improve as we get closer to the FOE.

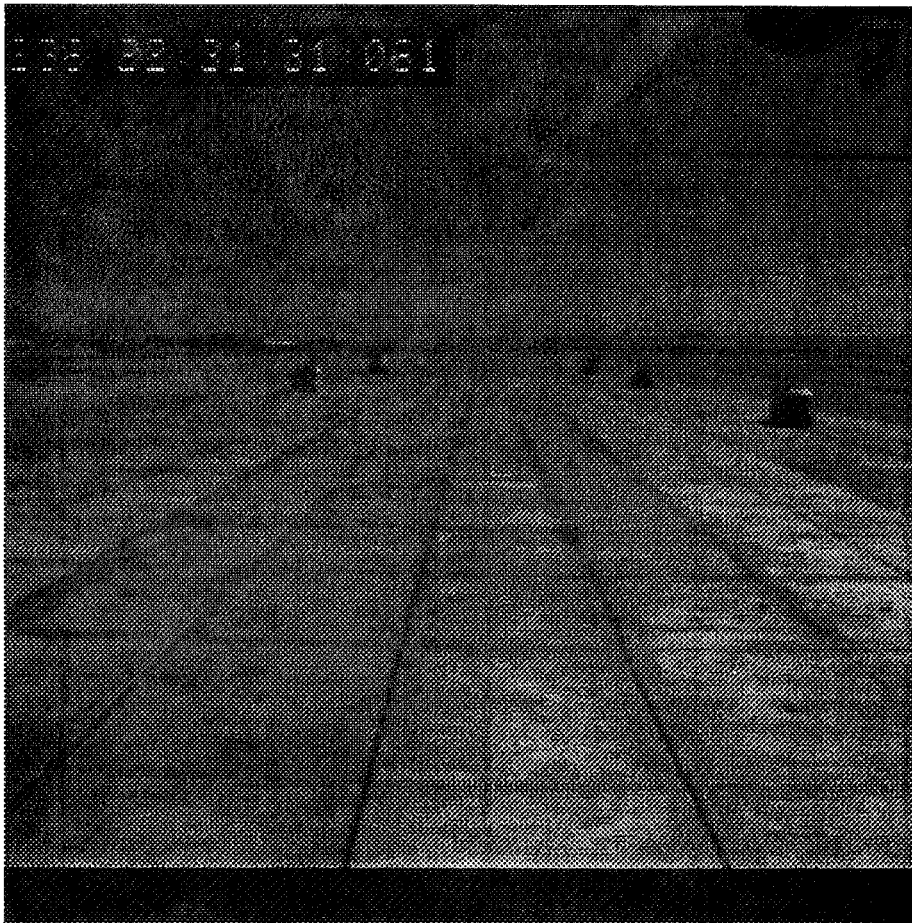


Figure 39: The first “newline” image.

We now present real-data cases from our imagery set “newline”; the first image of this sequence is shown in figure 39. The scene is that of a runway with a few surveyed trucks. The images are of size 512×512 , the speed is 30.17 ft/s, and the frame rate 30 per second. There are only minor maneuvers in this flight. The convergence curve is shown in figure 40 for the

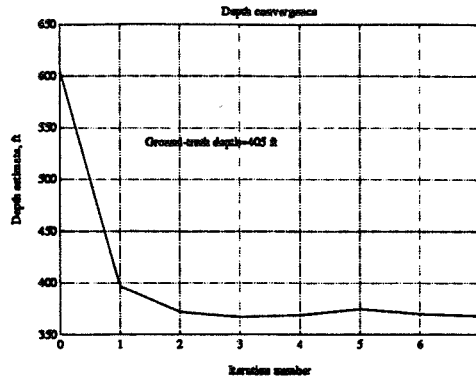


Figure 40: Depth convergence with iterations for “newline” leftmost truck.

leftmost truck which is at depth of 405 ft. The frame pairs used are: frame-0 with 2, 5, 10, 16, 22, 28, 34, and 40. Each iteration uses the next-larger frame separation. The converged depth resulting from the algorithm is 368, so that the accuracy here is of 9%. For the farther truck on

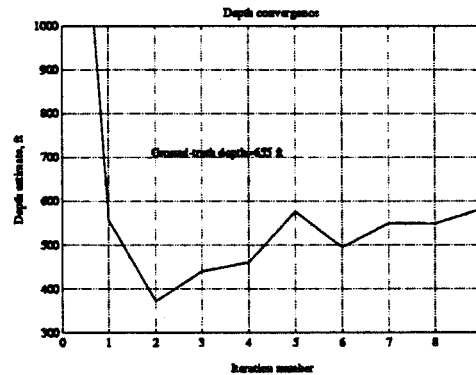


Figure 41: Depth convergence with iterations for “newline” leftmost truck.

the left, the algorithm ran ten iterations (last frame pair was (0,52)) and converged on a depth of 583 ft, where the ground-truth depth is 655 ft—accuracy of 11%. The convergence curve is shown in figure 41. The objects in these two examples show very little texture, and they are also small and far ($TTC \approx 10$ s) which may explain why the algorithm does not perform that well. Still the results can be considered satisfactory.

6 ERROR ANALYSIS

In this section we analyze the depth error as achieved by combining the depth results from lateral translation and those from expansion. We have already discussed the accuracy of the depth derived from lateral translation which is given by (18) where σ_u is given by figure 6.

The accuracy of the depth derived from expansion is determined by that of the expansion factor. When all the (four) parameters have converged, and thus compensated for, the case becomes that of nominally zero distortion and shifts. Therefore we have to examine the sensitivity of the correlation peak value to residual errors in the expansion factor alone. This accuracy is determined by the additive noise at the peak (denoted by $C_N(0,0)$). Notice that, so far, we have neglected this noise because it is practically much smaller than the *sidelobe* noise which results from the randomness of the image itself. The additive noise at the peak is given by equation (19) of [31] which is similar to (41) but with $\tau_u = \tau_v = 0$ and one of the correlation functions replaced by that of the noise, $R_N(u,v)$, that is,

$$\text{var}\{C_N(0,0)\} = L^{-2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(u,v) R(u,v) R_N(u,v) du dv \quad (47)$$

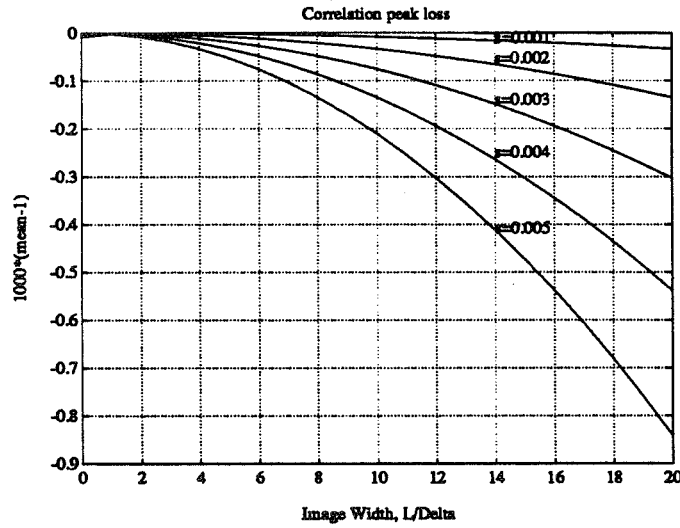


Figure 42: Loss in correlation peak value due to residual errors in scaling factor.

For simplicity we use equal $R(\tau_u, \tau_u)$ and $R_N(\tau_u, \tau_u)$ as given by (16). The question is now: what is the change in the expansion factor which causes a change in the correlation peak equal to the standard deviation, $\sqrt{\text{var}\{C_N(0,0)\}}$. The correlation peak, as given by (5) of [31], is plotted in figure 42. For the same example used earlier, where $L/\Delta = 14$, and assuming an image signal-to-noise ratio of a 100, it is found from (47) that $\sqrt{\text{var}\{C_N(0,0)\}} = 0.000177$. In the figure, the point having $L/\Delta = 14$ and an ordinate of -0.177 falls between the graphs of $s = 0.003$ and $s = 0.004$. Interpolating between these, results in $s = 0.00325$.

The relationship between the s error and the depth error is derived from (44), where we had

$$z_0 = \frac{sV_Z\Delta t}{s-1}, \quad (48)$$

so that

$$\frac{dz_0}{z_0} = -\frac{ds}{s(s-1)} \approx -\frac{ds}{s-1} \quad (49)$$

We can thus express the expansion-based depth standard deviation as

$$\sigma_{zs} = \frac{\sigma_s z_0}{s - 1} \quad (50)$$

For $s = 1.0274$, as was used to create figure 10 ($z_0 = 150$ m), and with $\sigma_s = 0.00325$, (50) yields $\sigma_{zs} = 17.8$ m which is close to the simulation results.

The depth information contained in the expansion factor, s , and that contained in the shifts (a, b) , is likely to be correlated because it is the same additive noise that causes inaccuracies in both measurements. Developing the necessary covariance matrix that relates their errors is not an easy task, and we thus forego that job here. However, we can still write down the combining algorithm for the initial-depth unbiased estimate, $\hat{z}_0$, as (see [33])

$$\hat{z}_0 = kz_s + (1 - k)z_t, \quad (51)$$

where z_s is the expansion-based depth measurement and z_t the translation- (or shifts-) based one. k is determined by the variances, σ_{zs}^2 of z_s and σ_{zt}^2 of z_t , and by their correlation coefficient ρ , as

$$k = \frac{\sigma_{zt}^2 - \rho\sigma_{zt}\sigma_{zs}}{\sigma_{zs}^2 + \sigma_{zt}^2 - 2\rho\sigma_{zt}\sigma_{zs}}, \quad (52)$$

and the minimum error—using this k —is then

$$E\{e^2\} \triangleq E\{(z_0 - \hat{z}_0)^2\} = \frac{\sigma_{zt}^2\sigma_{zs}^2(1 - \rho^2)}{\sigma_{zs}^2 + \sigma_{zt}^2 - 2\rho\sigma_{zt}\sigma_{zs}} \quad (53)$$

We know that, close to the FOE, $\sigma_{zs} \ll \sigma_{zt}$ so that, irrespective of ρ , $k \Rightarrow 1$, and vice versa. This means that, even if we use some guessed ρ of, say, 0.5 at this point, we will still be combining the measurements in a consistent way; that is the accurate measurement will contribute more than the inaccurate one—although, without knowing ρ , the proportions will not be optimal.

7 CONCLUDING REMARKS

In this paper we developed a new expansion-based passive-ranging algorithm that can complement the existing shift-based algorithm in the image areas near the FOE. We presented simulation and real-data results and compared them with the analysis results.

In the future we intend to develop this algorithm in two directions. One is to make it process an image sequence in real time and produce range maps. The other is to use it to segment an image into objects.

REFERENCES

- [1] L.H. Wegner. On the accuracy analysis of airborne techniques for passively locating electromagnetic emitters. Report R-722-PR AD 729 767, NTIS ASTIA D.C., Rand Corp, 1971.
- [2] J.L. Poirot and G.V. McWilliams. Application of linear statistical models to radar location techniques. *IEEE Trans. on Aerospace and Electronic Systems*, 10(6):830-834, November 1974.
- [3] D.J. Torrieri. Statistical theory of passive location systems. *IEEE Trans. on Aerospace and Electronic Systems*, 20(2):183-198, March 1984.
- [4] M. Gavish and E. Fogel. Effect of bias on bearing-only target location. *IEEE Trans. on Aerospace and Electronic Systems*, 26(1):22-25, January 1990.
- [5] Y. Barniv. Neural networks application to divergence-based passive ranging. Technical Memorandum 103981, NASA, Ames Research Center, Moffett Field, CA, August 1992.
- [6] B.K.P. Horn and B.G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(3):185-203, August 1981.
- [7] Y. Barniv. Dynamic programming solution for detecting dim moving targets. *IEEE Trans. on Aerospace and Electronic Systems*, 21(3):144-156, March 1985.
- [8] Y. Barniv and O. Kella. Dynamic programming solution for detecting dim moving targets part II: Analysis. *IEEE Trans. on Aerospace and Electronic Systems*, 23(6):776-788, November 1987.
- [9] Barniv, Y. (Ch. 4) (Yaakov Bar-Shalom, Editor). *Multitarget-Multisensor Tracking: Advanced Applications*. Artech House, 1990.
- [10] B. Sridhar and A.V. Phatak. Simulation and analysis of image-based navigation system for rotorcraft low-altitude flight. In *Proceedings of the AHS Meeting on Automation Application for Rotorcraft*, Atlanta, GA, April 1988.
- [11] B. Sridhar, V.H.L. Cheng, and A.V. Phatak. Kalman filter based range estimation for autonomous navigation using imaging sensors. In *Proceedings of the 11th Symposium on Automatic Control in Aerospace*, Tsukuba, Japan, July 1989.
- [12] Y. Barniv. Error analysis of combined optical-flow and stereo passive ranging. *IEEE Trans. on Aerospace and Electronic Systems*, to be published, October 1992.
- [13] H. C. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proc. R. Soc., London B*, 208:385-397, 1980.

- [14] K. Prazdny. Determining the instantaneous direction of motion from optical flow generated by a curvilinear moving observer. *Computer Vision, Graphics, and Image Processing*, 17:238–248, 1981.
- [15] K. Prazdny. Egomotion and relative depth map from optical flow. *Biological Cybernetics*, 36:87–102, 1980.
- [16] J. Koenderink. Optic flow. *Vision Research*, 26(1):161–180, 1986.
- [17] J.J. Koenderink and A.J. van Doorn. Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, 22(9):773–791, 1975.
- [18] J.J. Koenderink and A.J. van Doorn. Local structure of movement parallax of the plane. *Journal of Optical Society of America*, 66(7):717–723, July 1976.
- [19] R. C. Nelson and J. Aloimonos. Obstacle avoidance using flow field divergence. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(10):1102–1106, 1989.
- [20] D. L. Ringach and Y. Baram. A diffusion mechanism for obstacle detection from size-change information. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 5(5):6–7, 1992.
- [21] B. Sridhar, R.E. Suorsa, and B. Hussien. Passive range estimation for rotocraft low altitude flight. *Journal of Machine Vision and Applications*, 1993. Technical Memorandum 103899, NASA Ames Research Center. October 1991.
- [22] S. Merhav and Y. Bresler. On-line vehicle motion estimation from visual terrain information, part i: Recursive image registration. *IEEE Trans. on Aerospace and Electronic Systems*, pages 588–605, September 1986.
- [23] Y. Bresler and S.J. Merhav. On-line vehicle motion estimation for visual terrain information. part ii: Ground velocity and position estimation. *IEEE Trans. on Aerospace and Electronic Systems*, AES-22(5):588–603, September 1986.
- [24] Y. Bresler and S.J. Merhav. Recursive image registration with application to motion estimation. *IEEE Trans. on ASSP*, pages 70–86, January 1987.
- [25] S. Merhav. Air-to-surface range estimation by a stochastic gradient approach. Technical Memorandum MEMO, NASA, Ames Research Center, Moffett Field, CA, October 1992.
- [26] D. Reagan and K.I. Beverley. Looming detectors in the human visual pathway. *Vision Research*, 18:415–421, 1978.
- [27] D. Reagan and K.I. Beverley. Visual responses to changing size and sideways motion for different directions of motion in depth: Linearization of visual responses. *Journal of Optical Society of America*, 70(11):1289–1297, November 1980.

- [28] M. Hershenson. Visual system responds to rotational and size-change components of complex proximal motion patterns. *Perception and Psychophysics*, 42(1):60–64, 1987.
- [29] P. Cavanagh and O.E. Favreau. Motion aftereffect: A global mechanism for the perception of rotation. *Perception and Psychophysics*, 9:175–182, 1980.
- [30] D. Reagan and K.I. Beverley. Visual responses to vorticity and the neural analysis of optic flow. *Journal of Optical Society of America*, 2(2), February 1985.
- [31] Mostafavi H. and F.W. Smith. Image correlation with geometric distortion, part i: Acquisition performance. *IEEE Trans. on Aerospace and Electronic Systems*, 14(3):487–493, May 1978.
- [32] Mostafavi H. and F.W. Smith. Image correlation with geometric distortion, ii: Effect on local accuracy. *IEEE Trans. on Aerospace and Electronic Systems*, 14(3):494–500, May 1978.
- [33] A. Gelb. *Applied Optimal Estimation*. The MIT Press, 1974.
- [34] P.N. Smith. Flight data acquisition for validation of passive ranging algorithms for obstacle avoidance. In *Proceedings of the American Helicopter Society Forum*, Washington, D.C., May 1990.

Multiple-Camera/Motion Stereoscopy for Range Estimation in Helicopter Flight

Phillip N. Smith, Banavar Sridhar, and Raymond E. Suorsa
NASA Ames Research Center

24982
515-54
205745
P. 5

Abstract

Pilot aiding to improve safety and reduce pilot workload to detect obstacles and plan obstacle-free flight paths during low-altitude helicopter flight is desirable. Computer vision techniques provide an attractive method of obstacle detection and range estimation for objects within a large field of view ahead of the helicopter. Previous research has met considerable success by using an image sequence from a single moving camera in solving this problem. The major limitations of single camera approaches are that no range information can be obtained near the instantaneous direction of motion or in the absence of motion. These limitations can be overcome through the use of multiple cameras. This paper presents a hybrid motion/stereo algorithm which allows range refinement through recursive range estimation while avoiding loss of range information in the direction of travel. A feature-based approach is used to track objects between image frames. An extended Kalman filter combines knowledge of the camera motion and measurements of a feature's image location to recursively estimate the feature's range and to predict its location in future images. Performance of the algorithm will be illustrated using an image sequence, motion information, and independent range measurements from a low-altitude helicopter flight experiment.

1 Introduction

To increase safety and improve mission effectiveness during low-altitude helicopter flight, NASA Ames Research Center in conjunction with the U.S. Army has been developing automation tools to assist pilots in detecting obstacles and planning obstacle-free flight paths. The most challenging mode of low-altitude flight is Nap-of-the-Earth (NOE) flight, characterized by lateral maneuvers below tree-top level in order to conceal the helicopter behind available terrain or man-made objects. An on-line sensor to gather obstacle information is required for pilot-aiding during NOE flight because existing *a priori* terrain data such as digital maps (1) suffer from inaccuracies larger than the vehicle's altitude, (2) have insufficient resolution to show obstacles such as trees and buildings, and (3) cannot easily account for changes in the terrain such as the growth of new trees or the construction of new buildings. Vision sensors are desirable for obtaining the online obstacle information due to their passive nature and relatively large field of view.

The classification of obstacles is unnecessary for accomplishing the obstacle avoidance task because it is sufficient to avoid *all* obstacles regardless of identity. It is

therefore required only that the vision system provide position information for each object in the field of view. In practice the vision system attempts to compute a range map depicting the distance to the terrain for each point in the field of view.

A common approach to this problem makes use of an image sequence collected from a single moving camera and in some cases the camera's motion information. Small regions of interest (called features) are identified in an image, the feature's location is tracked in successive images, and a recursive filter is used to estimate range and/or camera motion [1, 2, 3]. The authors have previously developed an algorithm of this class and evaluated its performance with helicopter flight data as described in [4, 5, 6]. A major limitation of this approach is that range information cannot be obtained along the instantaneous direction of motion and, in practice, reliable range information cannot be obtained even for objects lying near the direction of motion. This limitation can be overcome through the use of multiple cameras mounted so their baseline is roughly normal to the motion direction [7, 8]. A hybrid motion/stereo algorithm is presented in this paper which allows range refinement through recursive range estimation while avoiding loss of range information in the direction of travel.

The extended Kalman filter provides a convenient structure for the implementation of motion/stereo range estimation. The Kalman filter allows for range refinement through recursive estimation. Furthermore, the range prediction generated during the time update serves to constrain the search area required to locate the feature in future images.

A low-altitude helicopter flight experiment has been conducted to obtain realistic data for evaluating the motion/stereo algorithm. The flight experiment provides video imagery from two monochrome video cameras, helicopter motion data, and camera calibration information. True range measurements have been obtained using a laser tracker to allow evaluation of the algorithm's performance.

The purpose of this paper is to describe a Kalman filter based motion/stereo ranging algorithm and to present preliminary results obtained using data from a helicopter flight experiment. Section 2 will discuss the Kalman filter implementation of the motion/stereo ranging algorithm. Section 3 will describe the helicopter flight experiment and calibration of the camera system. In Section 4, preliminary results obtained using the experimental data and the motion/stereo algorithm will be presented. Finally, Section 5 will complete the paper with a brief discussion and concluding remarks.

2 Kalman Filter

The proposed algorithm uses a feature-based method in which the image is treated as a collection of tokens or features and information (such as range) is computed only for the individual features rather than for every point in the image. Currently, features are defined to be 11×11 square pixel image patches which exhibit a sufficiently high intensity variance. A feature's location in another image is determined by correlation of the feature's intensity surface with the intensity surface of the other image. The correlation surface is then interpolated in the region near its peak, and the location of the resulting peak is taken to be the feature's location to subpixel accuracy. Features can be born with each new image, and old features die when they fail to be tracked between images. Further discussion of feature detection and tracking can be found in Ref. [4, 9].

In our implementation, a Kalman filter is associated with each feature for determining the location of the object which gives rise to the feature. The motion/stereo Kalman filter is an extension of the monocular range estimation Kalman filter derived in an earlier work [10]. Both filters rely on the assumptions that all objects of interest are stationary in an Earth-fixed frame, and that measurements of the camera's linear and angular velocities are available (from an inertial navigation system, for example). The resulting state equation is an expression of the Coriolis equation:

$$\dot{X} = -[\omega_s]X - V_s \quad (1)$$

where

$$[\omega_s] = \begin{bmatrix} 0 & -\omega_{zs} & \omega_{ys} \\ \omega_{xs} & 0 & -\omega_{xs} \\ -\omega_{ys} & \omega_{xs} & 0 \end{bmatrix}$$

$X = [x_s, y_s, z_s]^T$ is the object position relative to the camera, $\omega_s = [\omega_{xs}, \omega_{ys}, \omega_{zs}]^T$ is the camera's angular velocity, and V_s is the linear velocity. The measurement equation accounting for perspective projection of the object onto the image plane is given below

$$Z = h(X) = [fy_s/x_s, fz_s/x_s]^T$$

where $Z = [u, v]^T$ is the location of the object on the image plane and f is the camera's focal length. Here the camera axes have been defined with the x_s axis passing through the focal point and perpendicular to the sensor array, and y_s and z_s in the direction of the rows and columns of the sensor array, respectively. The extended Kalman filter is formed by linearizing $h(X)$ about the current state estimate yielding

$$\begin{aligned} Z &= HX \\ H &= \partial h(X)/\partial X \\ &= f \begin{bmatrix} 1/x_s & 0 & -y_s/x_s^2 \\ 0 & 1/x_s & -z_s/x_s^2 \end{bmatrix} \end{aligned}$$

To extend the Kalman filter characterization for two cameras we need additional measurement equations relating $Z' = [u', v']^T$, the image location of the same object in the second camera. Let X' be the object position relative to the second camera. The relationship between the cameras is of the form

$$X' = RX + T$$

where R is a 3×3 matrix and T is a vector representing the relative rotation and translation, respectively, between the two cameras' coordinate systems and centers of reference. Then the measurement Z' can be written as follows

$$Z' = h(X') = [fy'_s/x'_s, fz'_s/x'_s]^T$$

As above, we can derive a linearized measurement equation of the following form

$$\begin{aligned} Z' &= H'X \\ H' &= \partial h(X')/\partial X \end{aligned}$$

The Kalman filter can be computed for the system using the state equation (1) and the composite linearized measurement

$$Z_c = \begin{bmatrix} H \\ H' \end{bmatrix} X$$

Thus, the Kalman filter measurement update may be performed based on the obstacle location in any imaging sensor provided the location and orientation of the sensor are known relative to the reference sensor system. The stereo system has four measurements and the same state equations as the monocular system. Based upon the given state and measurement equations, the full discrete-time extended Kalman filter equations can be derived in the standard manner. This method can be extended in the same way to any number of cameras.

The range estimation process begins when a feature is identified in the image from one camera. A stereo match is determined by searching an area in the image from the second camera which is constrained by *a priori* values of the minimum and maximum range of interest. The resulting stereo range estimate is used to initialize the Kalman filter. The initial value of the Kalman filter's state covariance matrix may also be estimated or chosen *a priori*. The range estimate is then propagated forward in time by the Kalman filter, and the predicted state vector and state covariance matrix give rise to a search area to be traversed in locating the feature in the next image [9]. The Kalman filter uses the matched feature locations to perform its measurement update. As the Kalman filter converges, the value of the state covariance matrix decreases leading to smaller search areas and reduced computational effort. Given images from the two cameras over time, a variety of tracking schemes are possible. The currently implemented approach is to match each feature (1) from the left camera at the current time to the left camera at the next time and (2) from the left camera at the current time to the right camera at the next time. The above procedure is repeated for each feature until such time as the feature fails to be matched.

3 Flight Experiment

The helicopter flight experiment conducted to provide raw data and independent truth measurements for development and validation of passive ranging algorithms is illustrated in Figure 1. The resulting data set includes video imagery from two monochrome video cameras, helicopter motion data from an onboard inertial navigation system (INS), true range measurements obtained with a

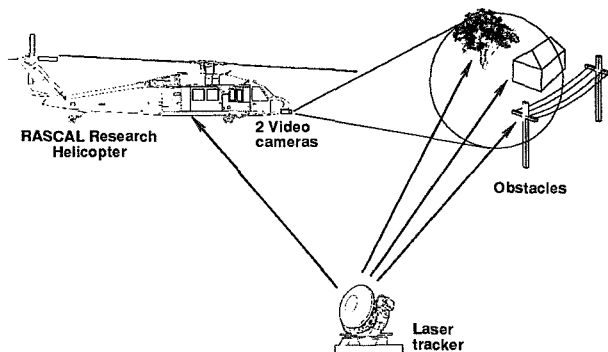


Figure 1: Flight Experiment Overview

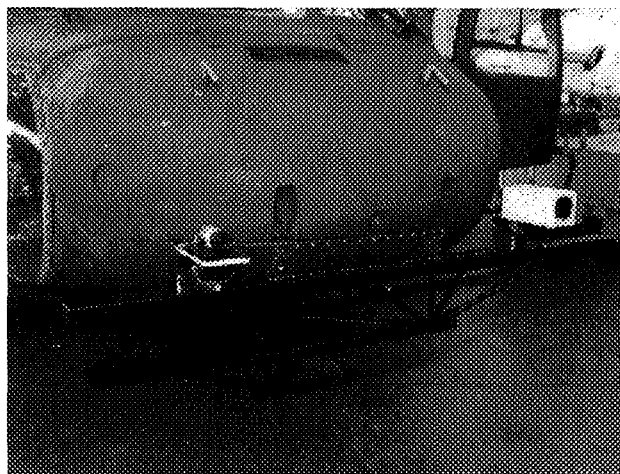


Figure 2: Camera Installation

laser tracker, and experimentally determined camera calibration parameters which characterize the geometry and imaging properties of the camera system.

The test apparatus consists of two Cohu 6410 monochrome interlaced video cameras mounted 1 meter apart on a horizontal bar attached to the nose of a UH-60 Blackhawk helicopter as shown in Figure 2. The cameras have a focal length of 6 mm, a field of view of 58×45 degrees and they are electronically shuttered with a $1/1000$ sec exposure time to reduce image smear due to camera motion. The video imagery from each camera is time-tagged using a Datum 9550 video time inserter unit and recorded using a Sony VO-9600 U-matic SP video recorder onboard the helicopter. The images are acquired at the rate of 30 frames/sec per camera. The helicopter's motion state is measured by a Litton LN93 inertial navigation system (INS) and also recorded onboard the helicopter. A laser tracker measures the helicopter's position during flight and also measures the location of the (stationary) obstacles of interest. Synchronization of the various data sources is accomplished by recording a master time index along with each element of the data set.

Post-flight processing consists of digitizing the recorded video data into 512×512 pixel images with 256 levels of gray. In addition, INS-derived motion data and laser-tracker-derived position data are processed together using a forward-backward filtering technique [11] to en-

sure kinematic consistency and to identify and correct for any sensor bias or scale factor errors. The resulting uncertainty in the motion data is approximately ± 2 ft in position, ± 0.01 deg in orientation, ± 0.25 ft/sec in velocity, and ± 0.3 deg/sec in angular velocity. Filtered motion data is desirable for development of the ranging algorithm, but in an operational system the motion state would be acquired directly from the INS.

The camera calibration parameters which characterize the camera system consist of two sets: the external parameters which include the geometrical description of the camera system, and the internal parameters which describe the imaging properties of the cameras. The external parameters allow the motion state measurements to be transformed from the helicopter body axes (as defined by the INS) to the sensor axis system (as defined by the cameras) for input to the Kalman filter. Similarly, using the external parameters, range estimates can be transformed back from sensor axes to body axes where they are more useful to the pilots or to an obstacle-avoidance guidance system. In addition, the external parameters provide the cameras' relative orientation, which is required for the stereo component of the ranging algorithm. The internal parameters define the mapping from points in the sensor axis system to pixel row and column coordinates in a digitized image. Internal parameters include the focal length, the pixel location where the x_s axis passes through the image plane, the effective dimensions of the pixels including any stretching effects caused by the recording and digitization process, and any distortion effects. There are a total of six external parameters and 5 internal parameters (assuming no distortion) for each camera. We have not yet found it necessary to model distortion terms with the ranging algorithms we have tested.

A separate experiment has been performed to determine the calibration parameters. Camera calibration has not received much attention in the literature but plays a central role in the performance of operational vision system. Some treatment of calibration techniques can be found in [12, 13, 14]. The approach taken here has been to (1) place a grid of target points within the cameras' field of view, (2) measure the locations of target points relative to the helicopter body axes, (3) determine the pixel locations of the target points in a digitized image taken with the camera, and (4) estimate the camera calibration parameters relating the two sets of measurements by solving a nonlinear cost minimization problem.

The calibration procedure uses a grid of horizontal and vertical lines, the 99 intersections of which serve as the calibration targets. A surveyor's transit is used to determine the target locations in the helicopter body axis system with an accuracy of approximately ± 3 mm. Five target points are measured directly, from which the remaining target locations can be interpolated. The entire grid assembly is stationed at four different distances in front of the cameras ranging between eight and 22 feet.

From a digitized image, the target pixel locations are found with subpixel accuracy by computing the intersections of curves fit to each of the grid lines. First, the intensity distribution perpendicular to one of the grid lines at some station is examined. The intensity peak, which is determined by locally fitting the intensity distribution with a parabola, defines one point on the grid line. The

process is repeated for several stations along each grid line, and the resulting points are fit with a curve (a line or a higher-order polynomial depending on the significance of image distortion). The curves' intersections are determined mathematically to give the target locations to subpixel accuracy.

In the final step, the parameters are estimated by minimizing a cost function which is a sum of squared errors terms. Two general approaches were taken: estimating the parameters for each camera separately and estimating the parameters for both cameras simultaneously. In the first case the cost function is the sum of errors in distance between the measured target pixel locations and the estimated pixel locations based on the measured body-axis locations and postulated parameter values. In a variation of this cost function, penalty terms were included for violation of Tsai's radial alignment constraint [12]. This calibration procedure resulted in RMS errors of approximately 0.4 pixel. However, using the resulting calibration parameters with the measured target pixel locations to estimate the corresponding body-axis locations using stereo leads to large errors. By estimating the calibration parameters for both cameras simultaneously the stereo ranging errors can be reduced through augmentation of the cost function. Several variations of the cost function were implemented, but little difference was observed in the result so long as terms were included for errors in the location of target points in the image plane and in the body axes. Weighting an error of 0.5 pixel in the image plane equivalently with a 0.25 inch error in the body axes leads to an RMS error of approximately 0.5 pixel and 0.5 inch, respectively.

4 Results

The image sequence used in generating the results given in this section was taken with the helicopter following a nominally straight flight path at a velocity of about 25 knots (42 ft/sec) 20 feet above a runway. Six trucks were positioned along the runway to serve as obstacles, initially ranging between 500 and 1100 feet from the helicopter. Figure 3 shows the first and last images in a sequence of 180 frames taken with the left camera. It is noted that in spite of the nominally straight line flight path, the FOE (depicted by crosshairs in Figure 3) travels 30 pixels in both the horizontal and vertical directions throughout the image sequence.

The image sequence is processed with the motion/stereo algorithm of Section 2 giving the range estimates to approximately 300 features in each image. To evaluate the algorithms performance, the average of the range estimates for all features belonging to each truck is computed. These preliminary results for the five closest trucks are given in Table 1 along with the true range at frame numbers 1, 60, 120, and 180. For reference, the corresponding results obtained with the earlier monocular ranging algorithm are also shown in Table 1. The preliminary results show that the initial range estimates are significantly better using the stereo method as expected since the trucks are both far away and close to the FOE. Over time, the additional measurements lead to improved range estimates and the results of both

Table 1: Preliminary Range Results

Truck	Frame	Range, ft		
		Truth	Monocular	Motion/Stereo
A	1	488	171	489
	60	399	405	431
	120	316	335	350
	180	235	227	247
B	1	614	270	785
	60	525	568	587
	120	443	462	463
	180	363	364	341
C	1	741	267	739
	60	650	519	498
	120	568	606	565
	180	487	514	486
D	1	860	138	N/A
	60	770	618	594
	120	688	653	799
	180	609	534	671
E	1	991	122	955
	60	899	995	813
	120	817	594	698
	180	736	863	722

methods converge toward the true range. Note that the motion/stereo case sometimes produces less accurate results, potentially due to the following characteristics of the currently-implemented algorithm. Range estimates are not always available using the stereo-motion method. In fact there are only half as many features resulting from the motion/stereo method as from the monocular method, indicating fewer (though hopefully stronger) feature matches. Sometimes even apparently strong features may fail to match in both cameras which on further examination is attributed to small-scale differences between the images from the two cameras due to image noise and the differences in the cameras themselves. A modification of the tracking scheme to match only between images taken with the same camera or between images taken at the same time may lead to better matching. Even if matching cannot be improved, the motion/stereo results could be enhanced by allowing range estimates to be propagated based on monocular motion only rather than killing the feature in the event that a stereo match cannot be made. In this way, the motion/stereo algorithm gracefully degrades to the monocular algorithm when stereo matches cannot be obtained, but stereo information is utilized when it is available.

5 Concluding Remarks

A hybrid motion/stereo range estimation algorithm has been described which combines the strengths of stereo methods (i.e., ranging without motion and ranging to objects near the FOE) and monocular methods (i.e., recursive range refinement). This motion/stereo algorithm has been implemented as a Kalman filter. A helicopter flight experiment was conducted to collect data for validation of the algorithm. Preliminary results indicate that initial motion/stereo range estimates are an improvement over initial monocular estimates and that both methods give range results which generally approach the true range over time. It was noted that some improvement in the robustness of the motion/stereo algorithm could be obtained by



Figure 3: First and Last Images of Helicopter Sequence

allowing it to degrade to the monocular algorithm for a given feature when a stereo match cannot be established. In the future we plan to continue refinement of the motion/stereo algorithm and to test it with flight sequences having curvilinear motion and images of natural terrain.

References

- [1] Chellappa, R., and Shekar, C., "Passive Ranging Using a Moving Camera," *Journal of Robotic Systems*, vol. 9, 1992.
- [2] Bhanu, B., Roberts, B., and Ming, J., "Inertial Navigation Sensor Integrated Motion Analysis," *Proceedings of the DARPA Image Understanding Workshop*, Palo Alto, CA, May 1989.
- [3] Kumar, R., and Hansen, A.R., "Model Extension and Refinement Using Pose Recovery Techniques," *Journal of Robotic Systems*, vol. 9, 1992.
- [4] Sridhar, B., Suorsa, R., and Hussien, B., "Passive Range Estimation for Rotorcraft Low-altitude Flight," NASA TM, October 1990; to appear in *Journal of Machine Vision and Applications*, March 1993.
- [5] Smith, P.N., "Flight Data Acquisition for Validation of Passive Ranging Algorithms for Obstacle Avoidance," *Journal of the American Helicopter Society*, vol. 37, No. 4, pp. 32-37, October 1992.
- [6] Smith, P.N., Sridhar, B., and Hussien, B., "Vision-Based Range Estimation Using Helicopter Flight Data," *Proceedings of The 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp 202-208, Champaign, IL, June 1992.
- [7] Sridhar, B. and Suorsa, R., "Integration of Motion and Stereo Sensors in Passive Ranging Systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 27, No. 4, pp. 741-746, July 1991.
- [8] Kriegman, D.J., Triendl, E., and Binford, T.O., "Stereo Vision and Navigation in Buildings for Mobile Robots," *IEEE Transactions on Robotics and Automation*, vol. 5, pp. 792-803, 1989.
- [9] Suorsa, R.E., and Sridhar, B., "A Parallel Implementation of a Multisensor Feature-Based Range-Estimation Method," NASA TM 103999, January 1993.
- [10] Sridhar, and Phatak, A., "Simulation and Analysis of Image-Based Navigation system for Rotorcraft Low-Altitude Flight," *American Helicopter Society Specialists' Meeting on Automation Applications of Rotorcraft*, Atlanta, GA, April 1988; *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, no. 2, pp. 96-101, March/April 1992.
- [11] Bach, R.E., Jr., "State Estimation Applications in Aircraft Flight-Data Analysis (A User's Manual for SMACK)," NASA RP 1252, March 1991.
- [12] Tsai, R.Y., "A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 4, Aug. 1987, pp. 323-344.
- [13] Slama, C.C. (editor), *Manual of Photogrammetry, Fourth Edition*, American Society of Photogrammetry and Remote Sensing, Falls Church, Virginia, 1980.
- [14] Beyer, H.A., "Accurate Calibration of CCD-Cameras," *Proceedings of The 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp 202-208, Champaign, IL, June 1992.

**Validation of Vision-Based Range Estimation Algorithms
Using Helicopter Flight Data**

Phillip N. Smith
NASA Ames Research Center

249938
516-74
205746
p. 19

Objective

Demonstrate effectiveness of an optic flow method for passive range estimation using a Kalman-filter implementation with helicopter flight data

Overview

- Ranging Algorithm
- Flight Experiment
- Analysis Methodology
- Results
- Concluding Remarks

Range Estimation Algorithm

- Monocular range-from-motion
- Camera motion state assumed known
- Feature-based approach for optic flow
- Correlation method for feature matching
- Kalman-filter implementation for range estimation

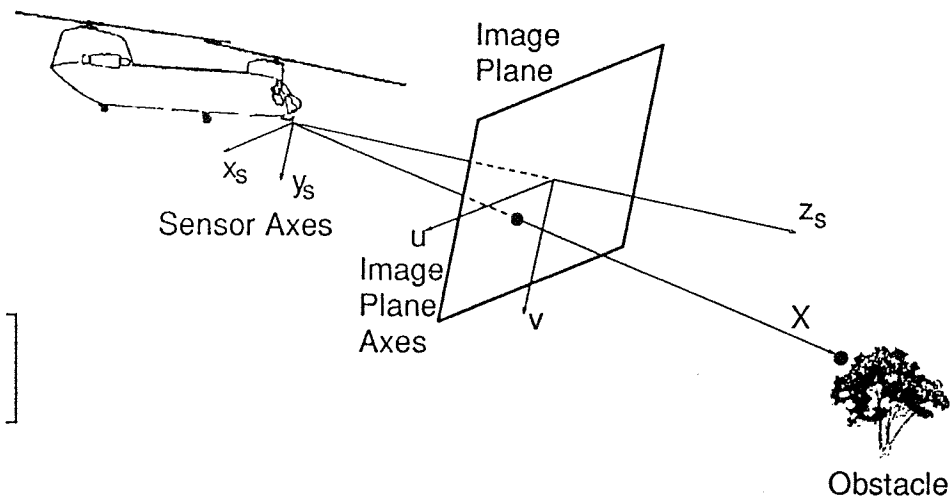
$$\dot{X} = -V_s - \omega_s \times X$$

$$Z = HX$$

$$X = \{x_s, y_s, z_s\}^T$$

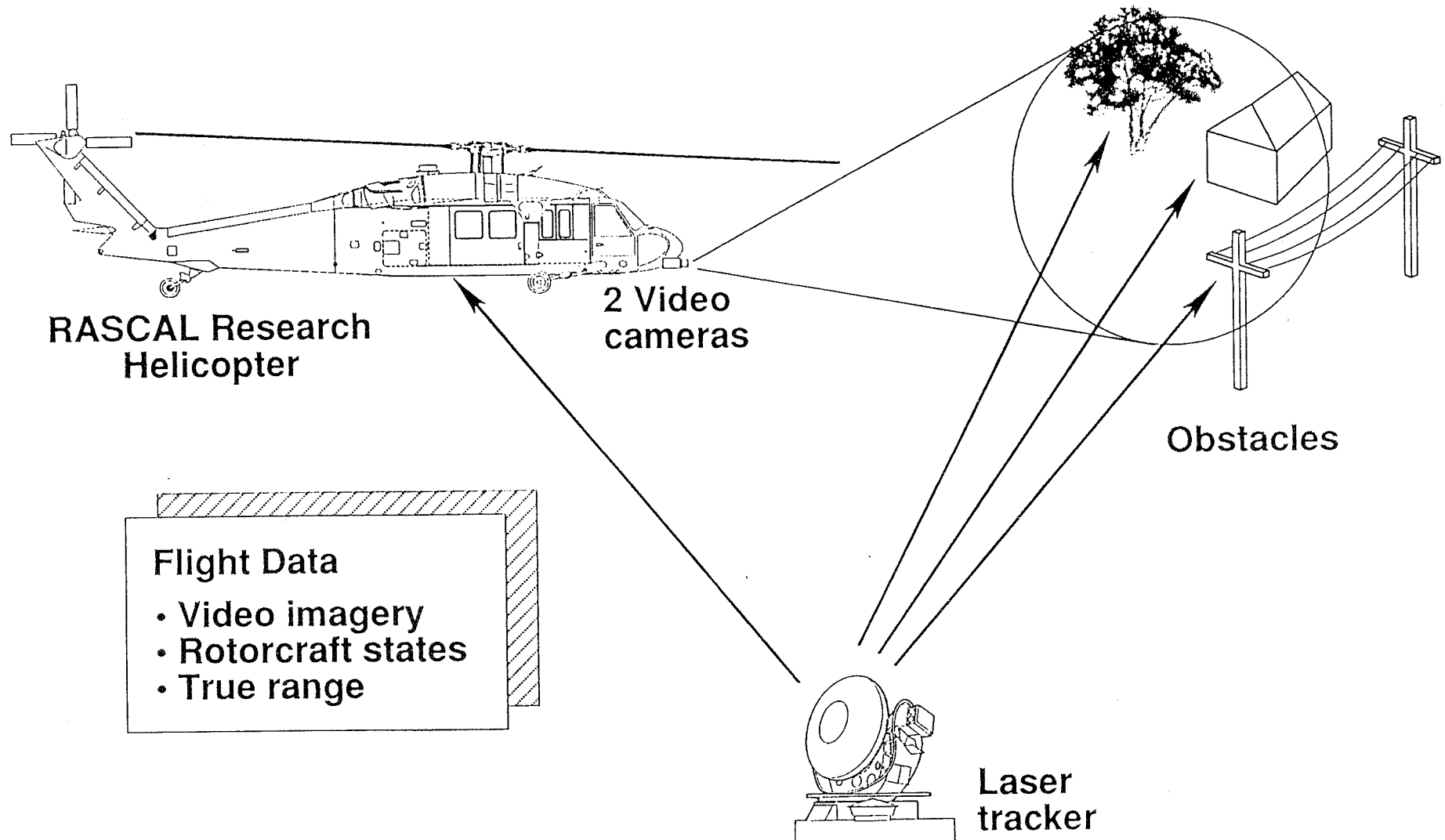
$$Z = \{u, v\}^T$$

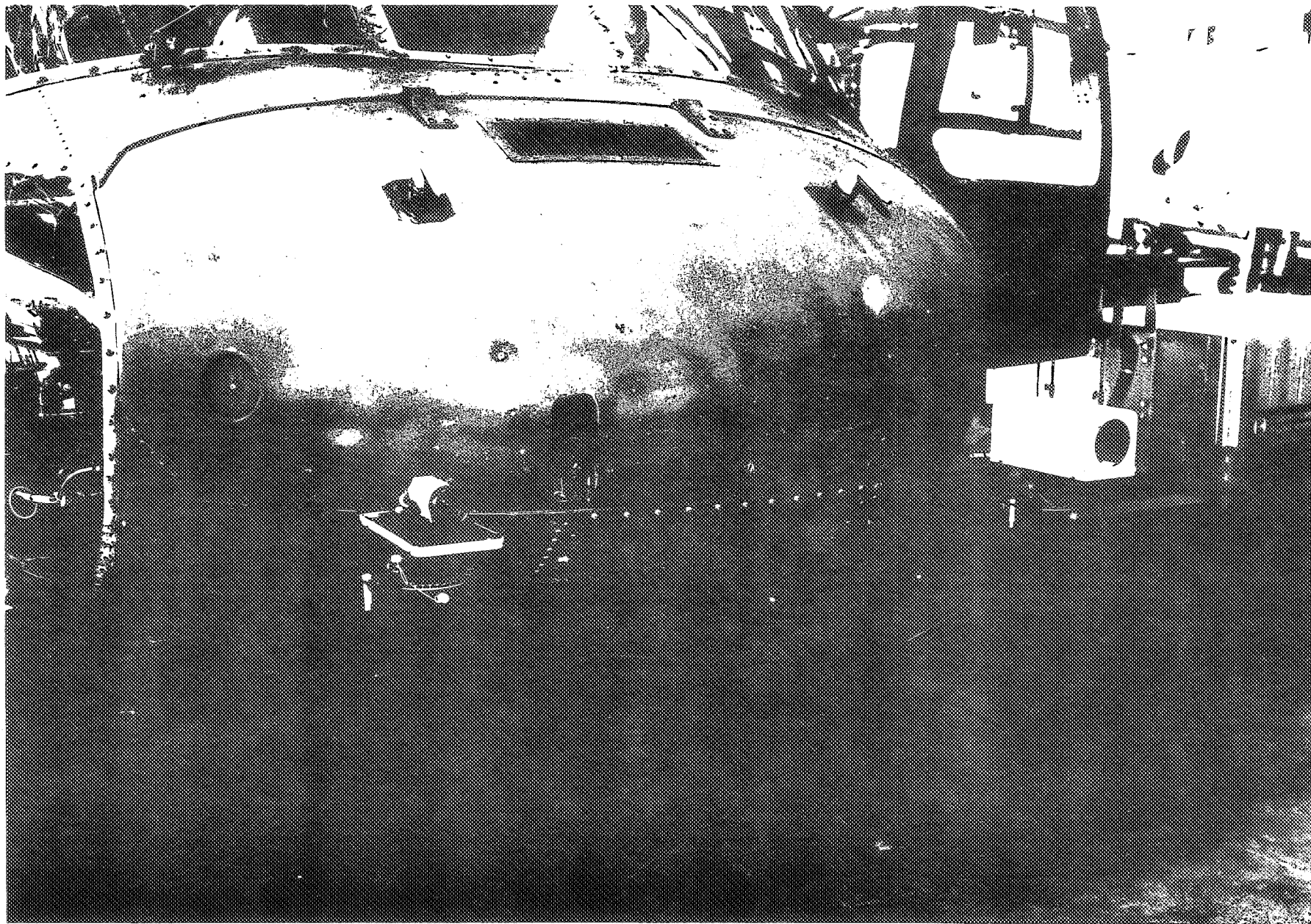
$$H = f \begin{bmatrix} 1/z_s & 0 & -x_s/z_s^2 \\ 0 & 1/z_s & -y_s/z_s^2 \end{bmatrix}$$



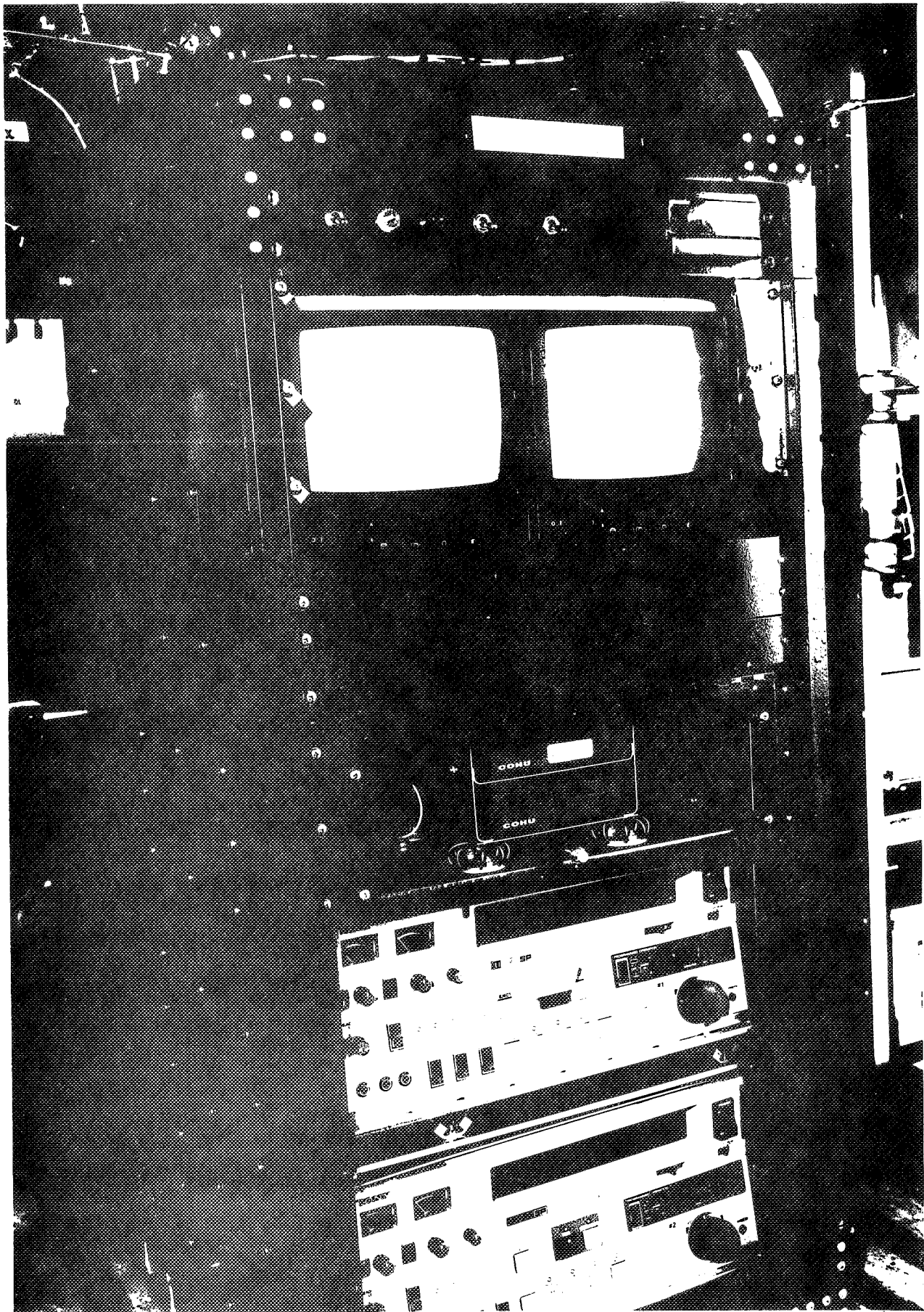
FLIGHT EXPERIMENT DESCRIPTION

295





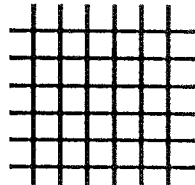
Camera Installation



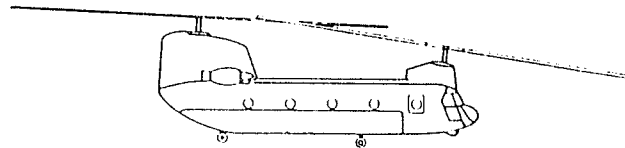
On-Board Operator Station
297



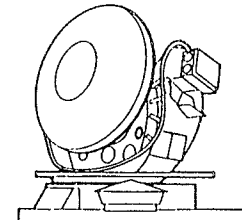
Image data



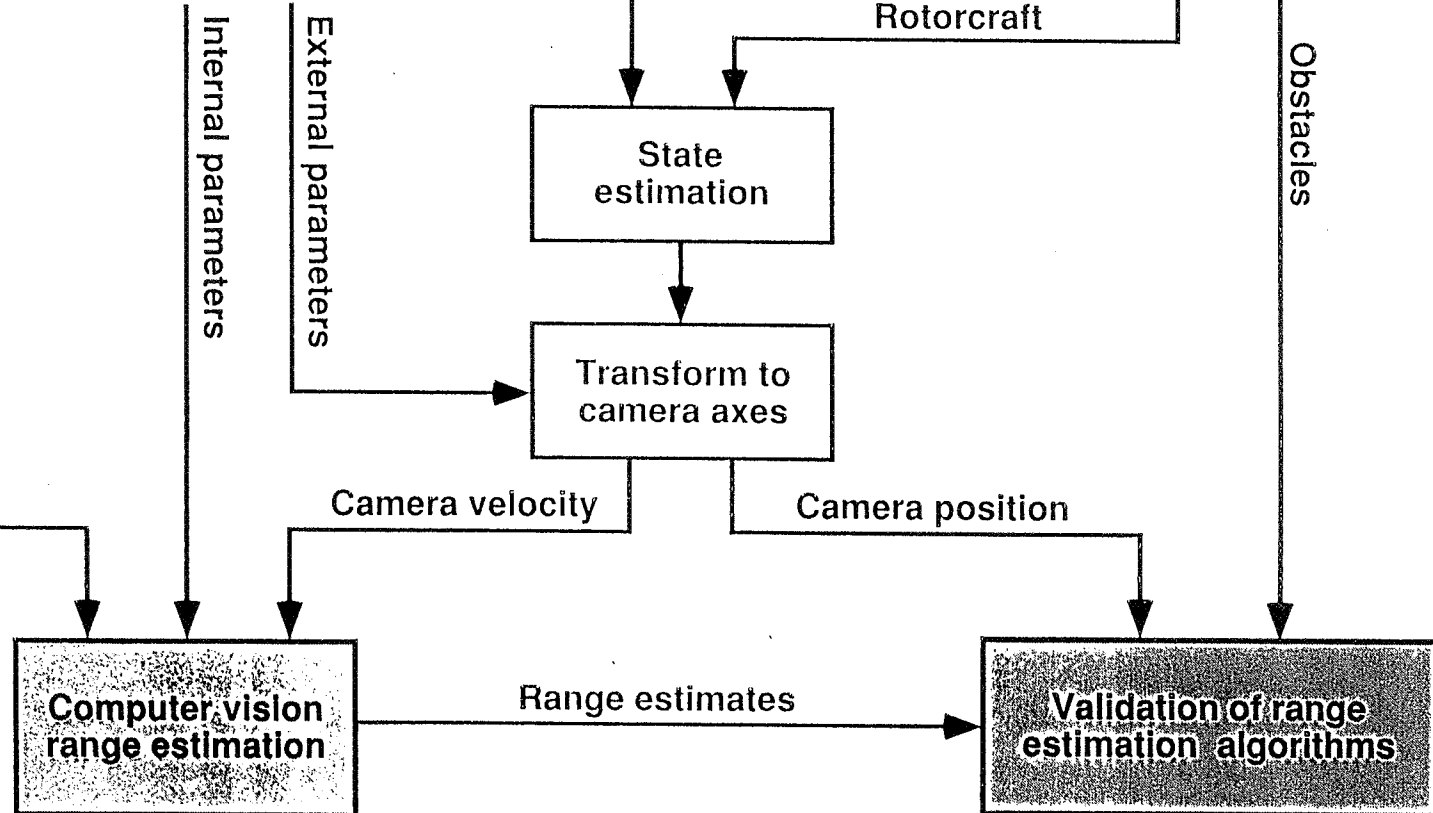
Camera calibration



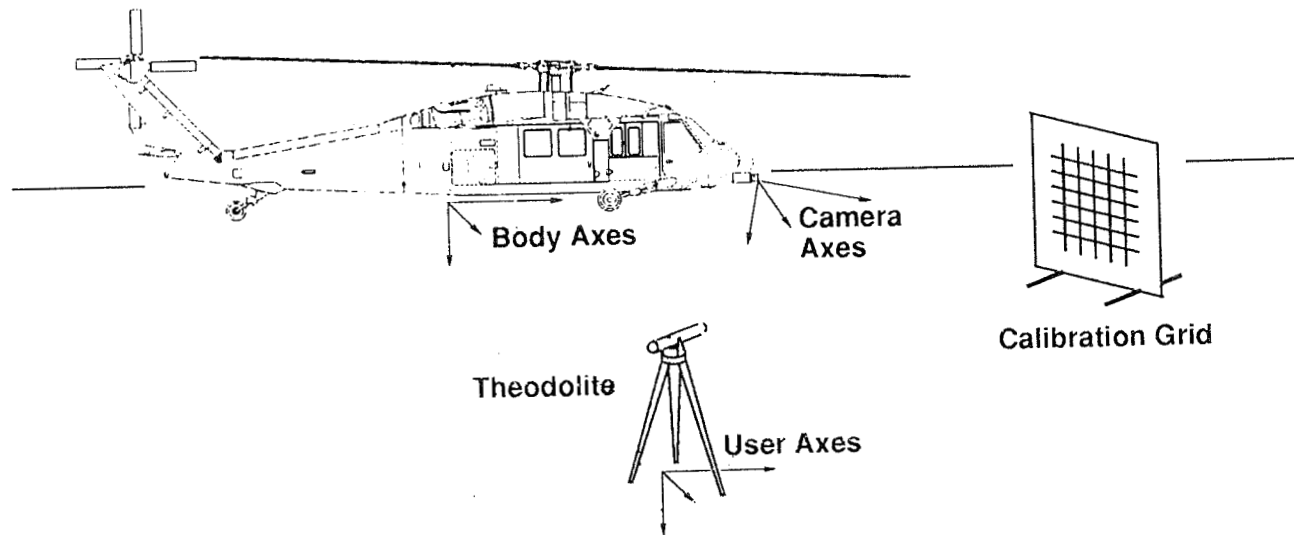
Rotorcraft data



Position data



Calibration

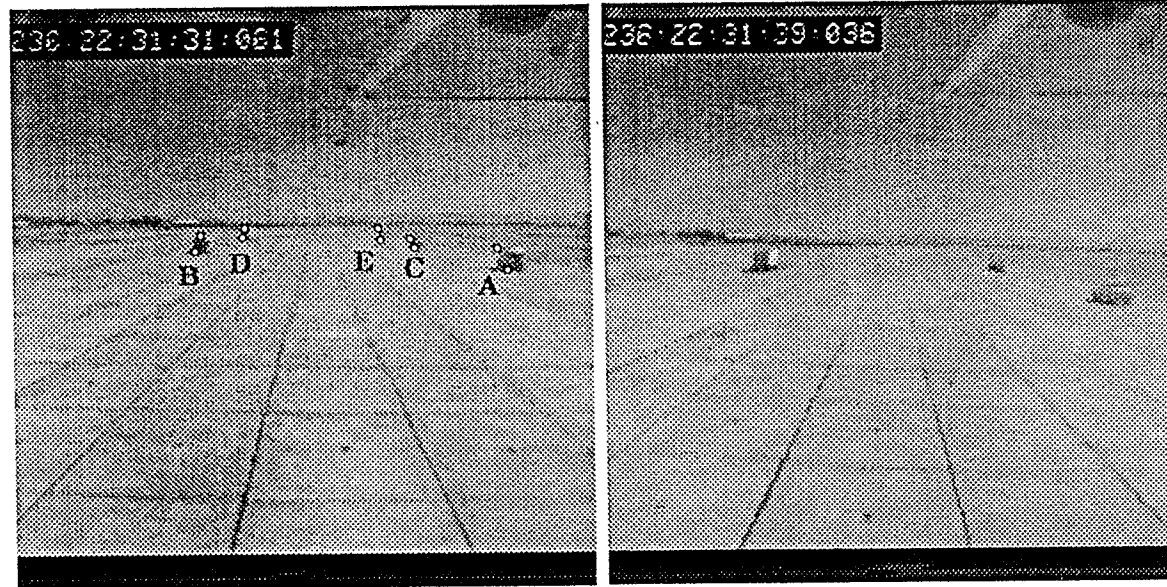


- Calibration parameters define the transformation of object points onto the image plane
- Calibration parameters
 - Position/orientation (6 unknowns)
 - Imaging model (4 unknowns)

Analysis Methodology

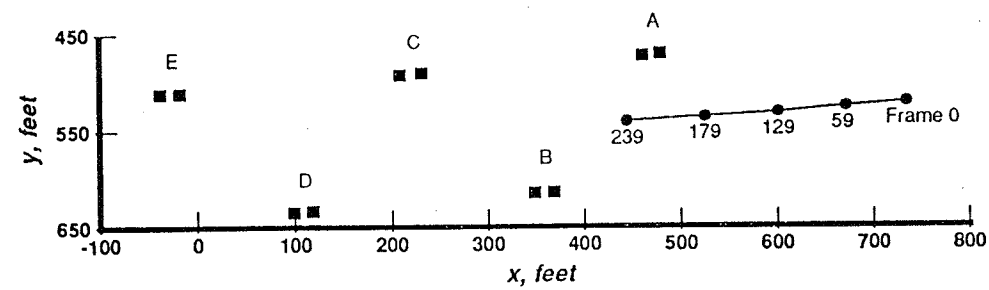
- Validate algorithm with different flight sequences
- Points of Comparison
 - Range accuracy
 - Convergence time
 - Range at convergence

Line Sequence

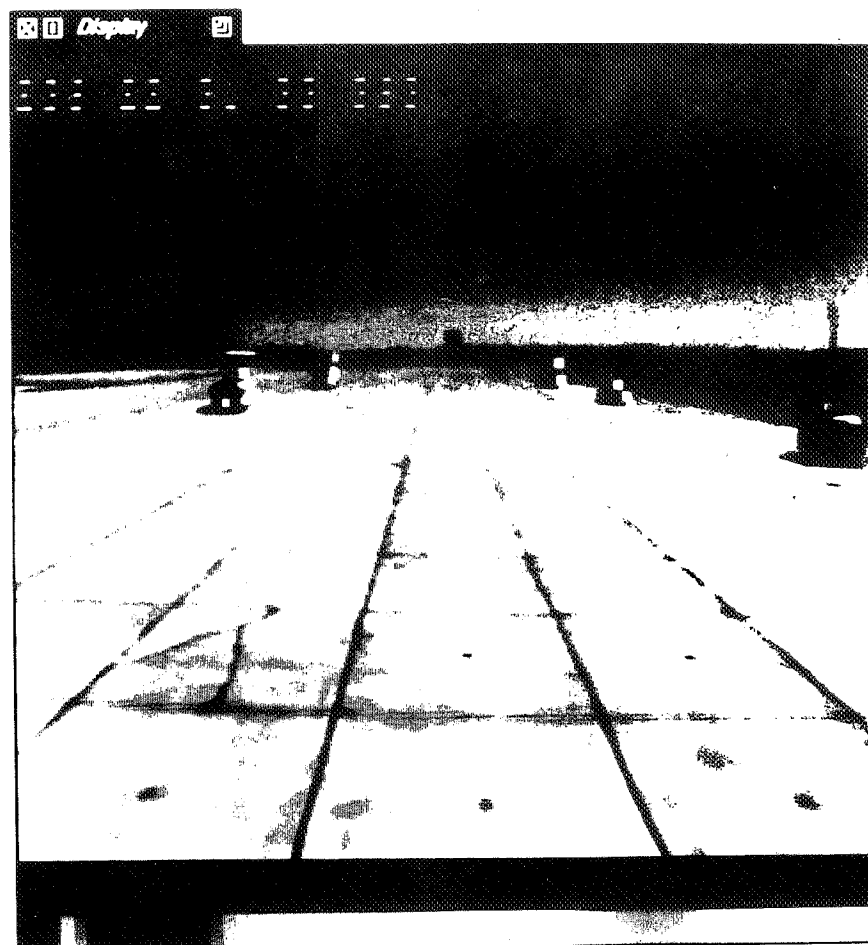
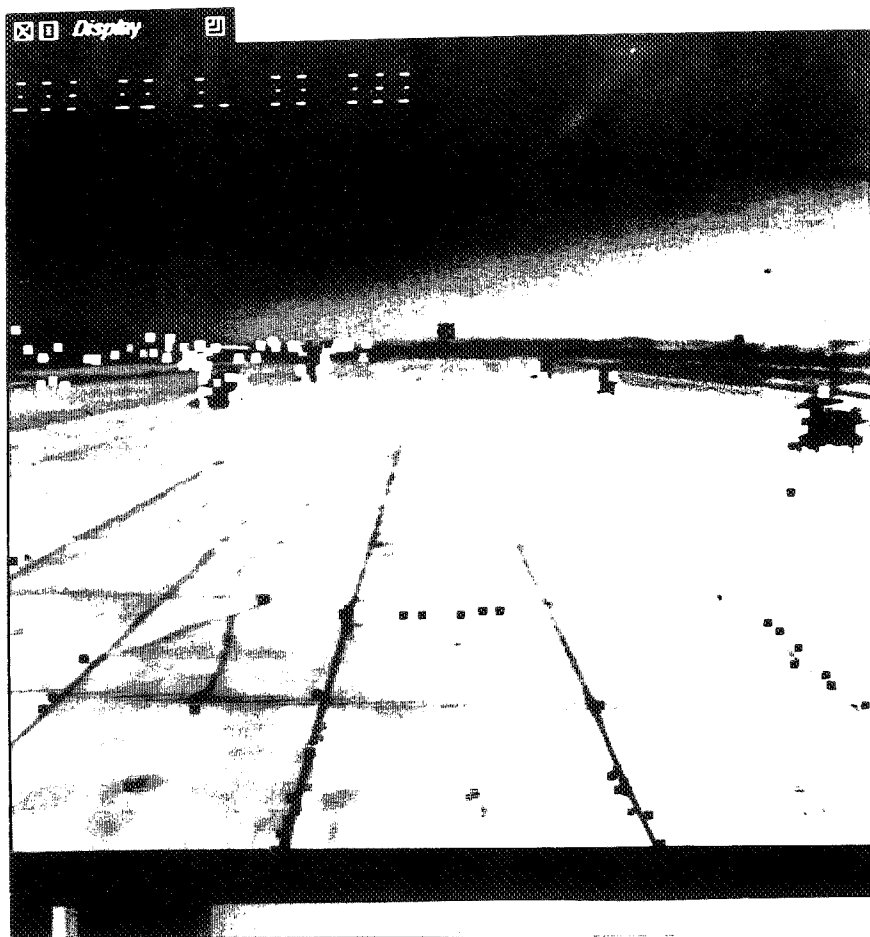


Frame 0

Frame 239

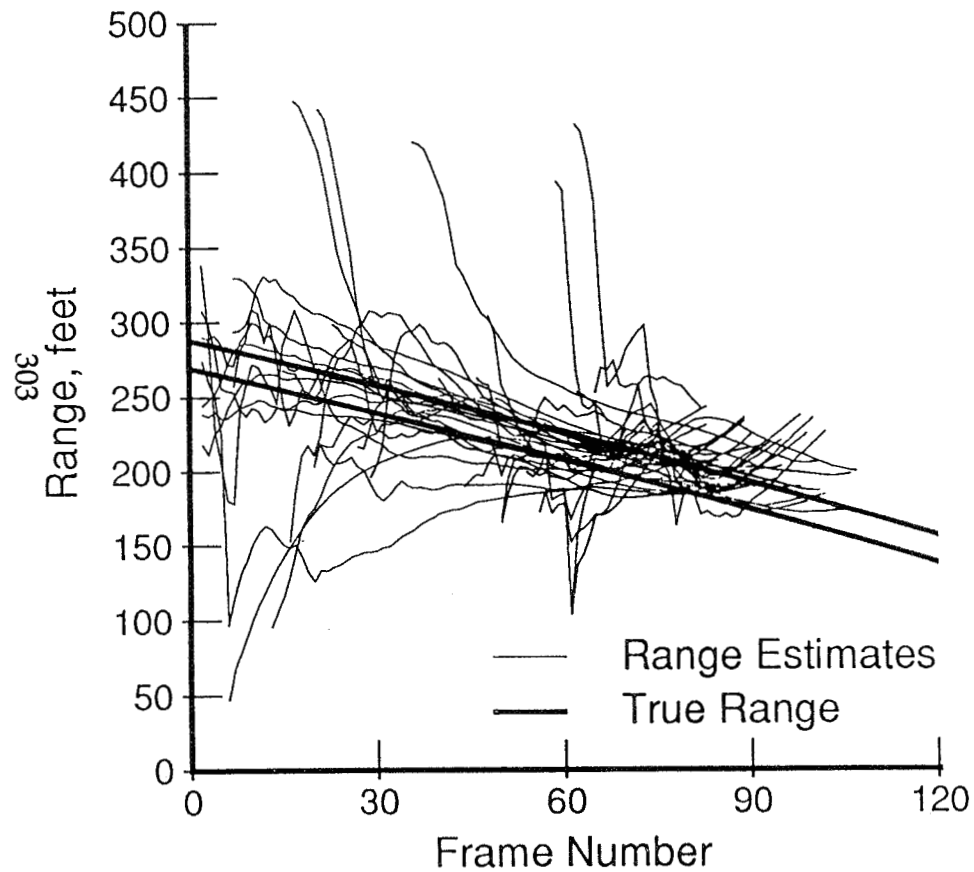


Camera Path and Measured Truck Locations

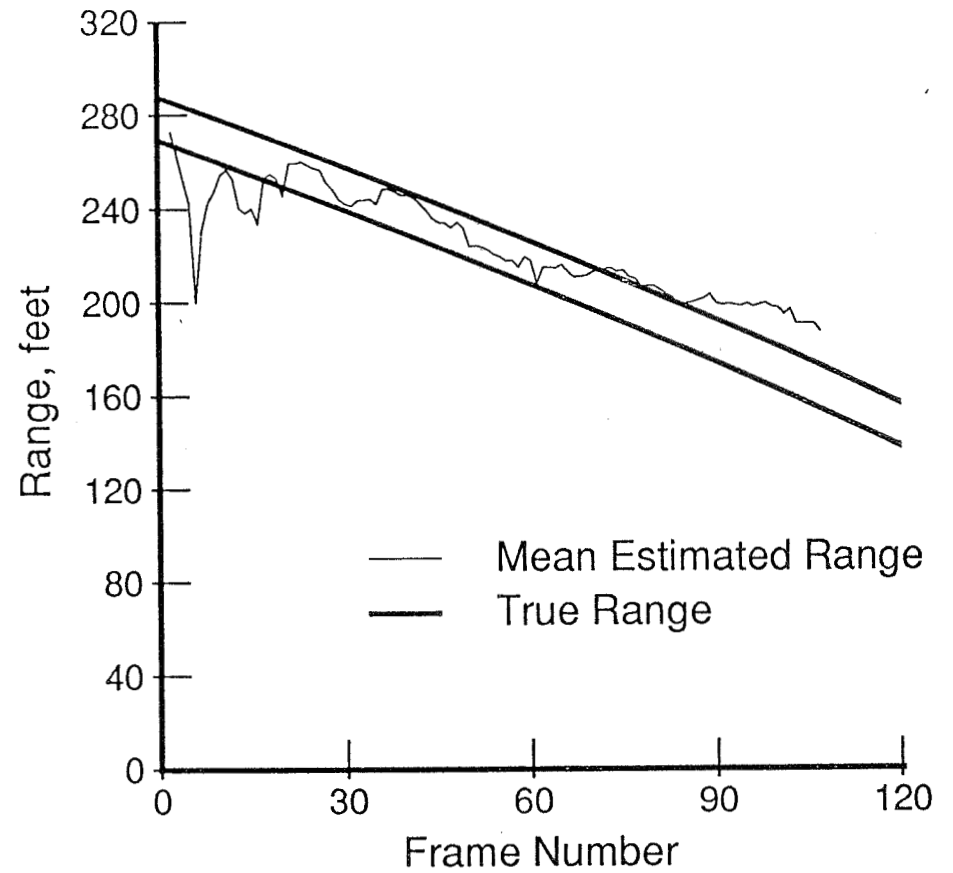


Helicopter Sequence Images with Features

Range Estimates for Truck A

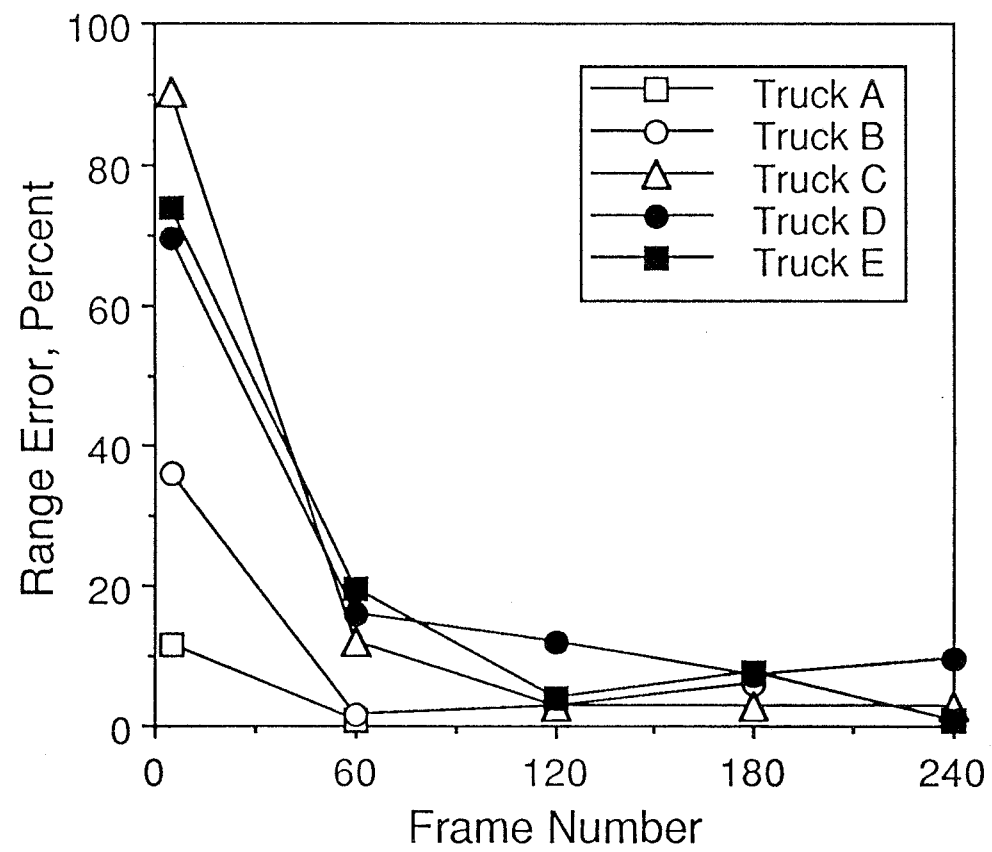


Individual Range Estimates

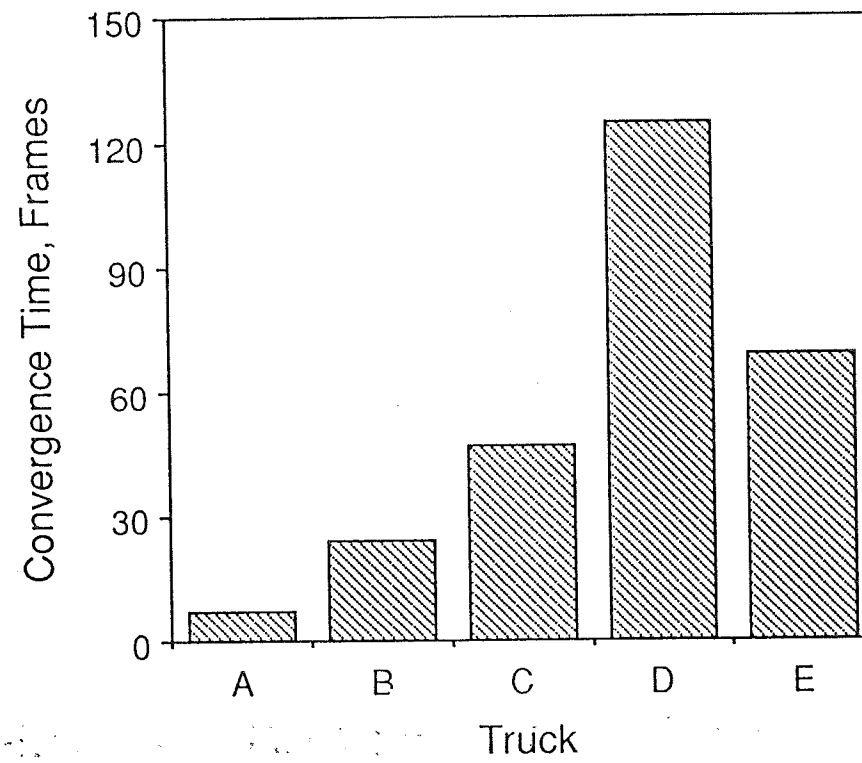


Mean Range Estimate

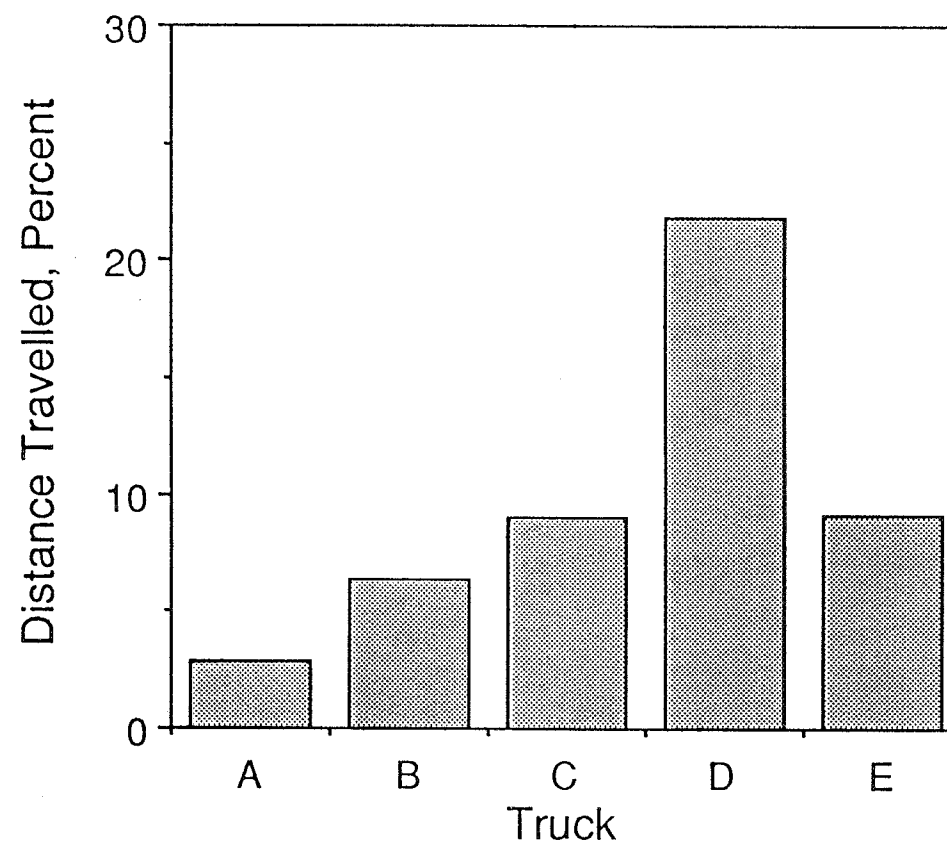
Range Accuracy



Time for Convergence



Distance Travelled for Convergence



Stereo Line Sequence

307



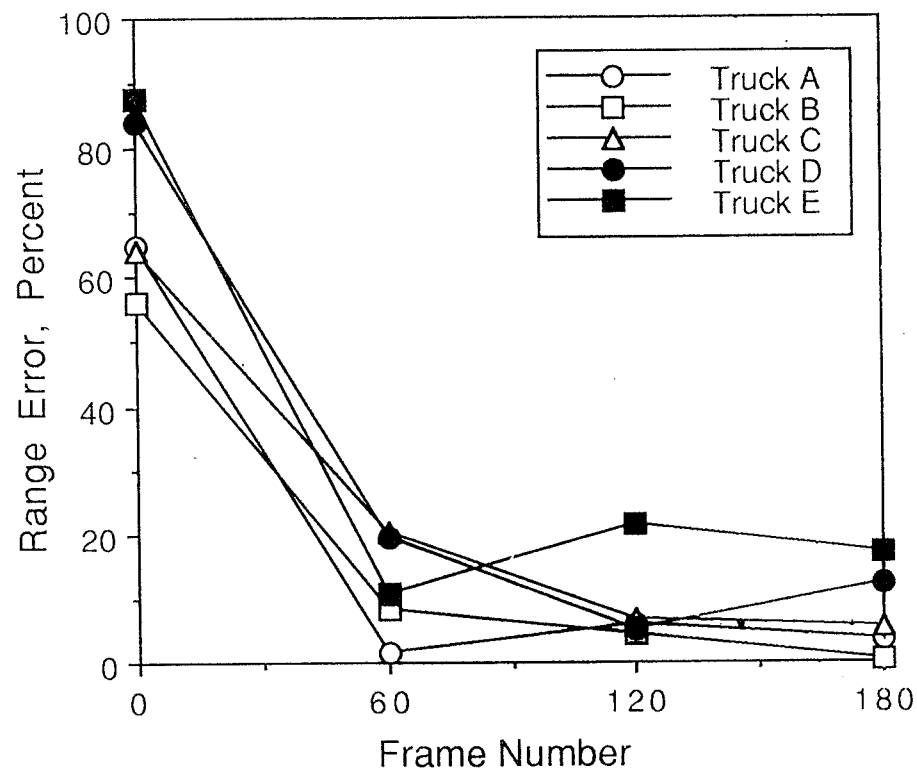
Frame 0



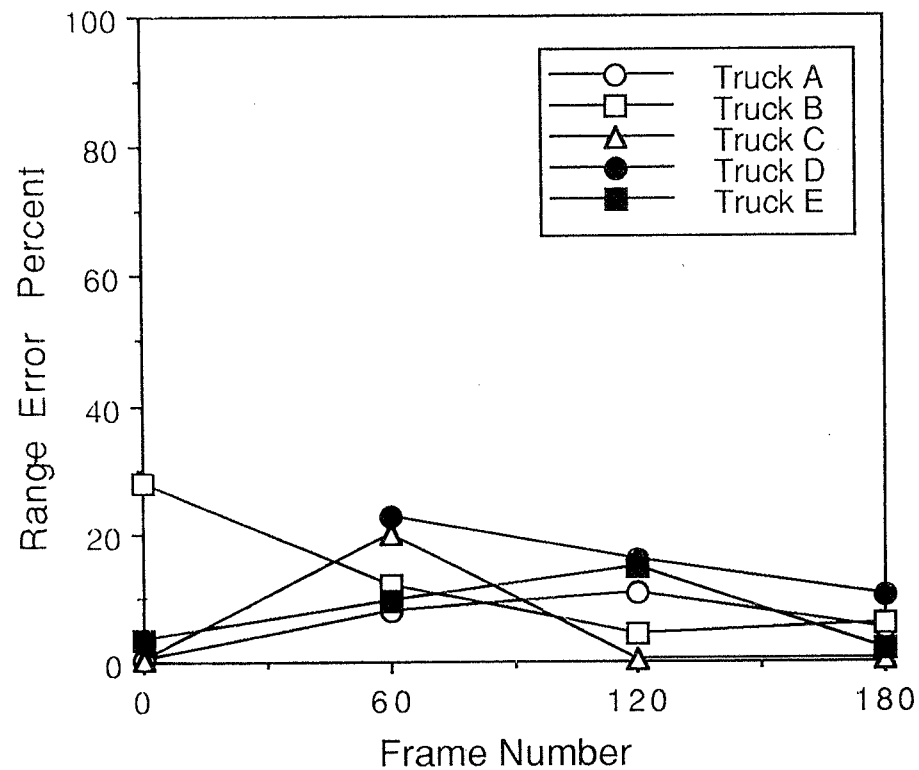
Frame 179

Range Accuracy

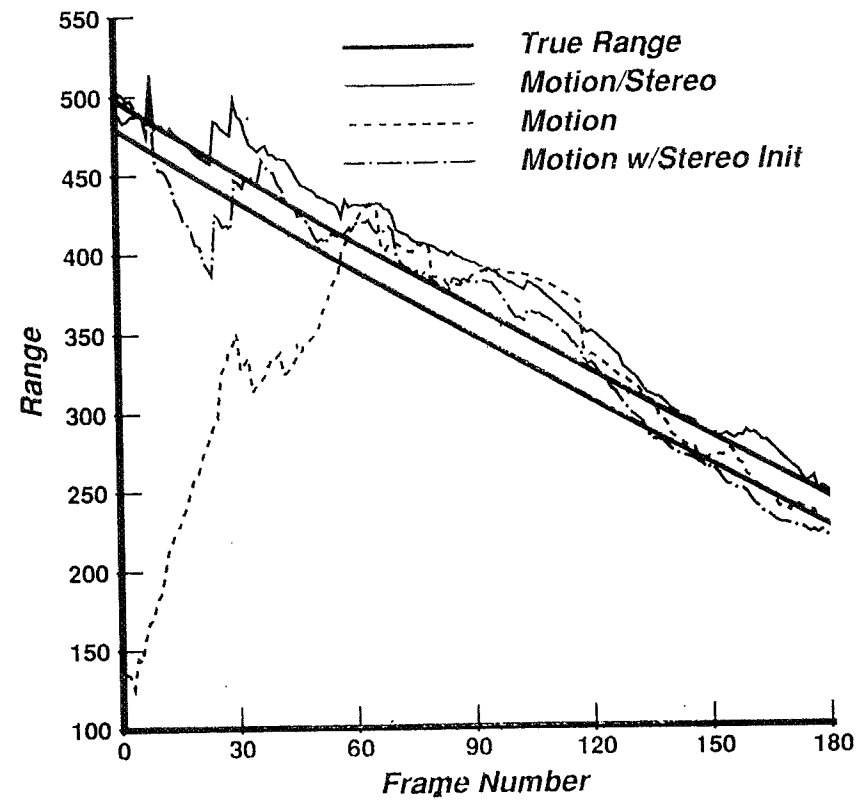
Motion Only

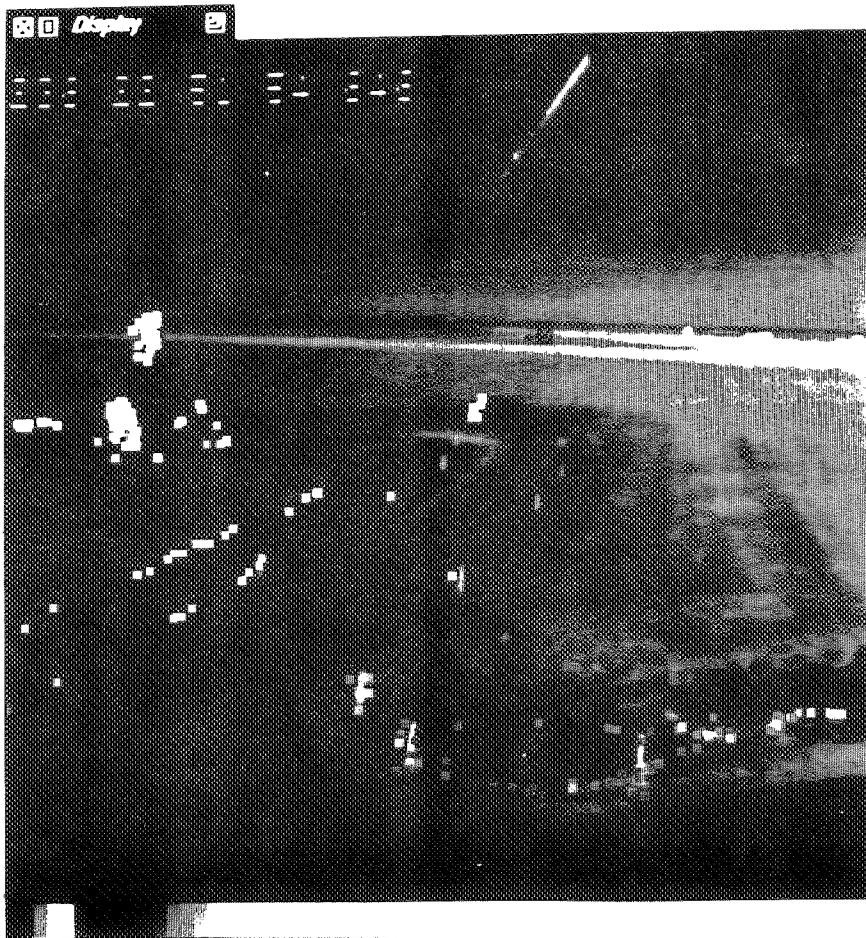


Motion/Stereo



Comparison of Ranging Methods for Truck A





Feature mapping on open-field images.

Concluding Remarks

- Successful validation of vision-based ranging algorithms using helicopter data
 - General sensor motion
 - Realistic sensor vibration
- Algorithm demonstrates robust performance with range accuracy of about 10% for objects whose range is up to 10 times the distance travelled
- Research issues
 - Combination of motion and multisensor ranging methods
 - Frame rate selection
 - Calibration
 - Generalized motion
- Future Work
 - Further processing of multicamera sequences
 - Infrared image sequences

A Model-Based Approach for Detection of Objects in Low Resolution Passive-Millimeter Wave Images*

Yuan-Liang Tang, Sadashiva Devadiga, and Rangachar Kasturi**

The Pennsylvania State University

and

Randall L. Harris, Sr.
NASA Langley Research Center

Abstract

We describe a model-based vision system to assist the pilots in landing maneuvers under restricted visibility conditions. The system has been designed to analyze image sequences obtained from a Passive Millimeter Wave (PMMW) imaging system mounted on the aircraft to delineate runways/taxiways, buildings, and other objects on or near runways. PMMW sensors have good response in a foggy atmosphere; but their spatial resolution is very low. However, additional data such as airport model and approximate position and orientation of aircraft are available. We exploit these data to guide our model-based system to locate objects in the low resolution image and generate warning signals to alert the pilots. We also derive analytical expressions for the accuracy of the camera position estimate obtained by detecting the position of known objects in the image.

I. Introduction

Federal regulations specify the minimum visibility conditions under which airlines may take off and land. These minima are a function of the types of airplane and airport equipment. Therefore, there is a great deal of interest in imaging sensors which can *see through* fog and produce a *real world* display which, when combined with symbolic or pictorial guidance information, could provide the basis for a landing system with lower visual minimum capability than those presently being used [1].

Since the energy attenuation in the visible spectrum due to fog is very large [2] (Fig.1), sensors are being designed to operate at lower frequencies (e.g. 94 GHz) where the attenuation is lower providing the ability to see through fog. NASA Langley Research Center, in cooperation with industry, is performing research on an on-board imaging system using a passive sensor operating at this frequency. Images from such sensors are of very low spatial resolution (Fig. 2). However, additional supporting information in the form of knowledge about the airport and the position, orientation and velocity of aircraft is generally available. Thus a model-based image analysis approach is feasible to segment the image and to detect and track objects on the ground. Information

* This research is supported by a grant from NASA Langley Research Center, Hampton, Virginia (NAG-1-1371).

** Address all correspondence to Professor Rangachar Kasturi (kasturi@cmpe.psu.edu).

extracted from such an analysis is useful to generate warning signals to the pilot of any potential hazards. This paper describes such a model-based technique, which makes use of a priori information about the geometric model of the airport and camera position and attitude data provided by the Global Positioning System (GPS) and other instruments.

The geometric model of the airport contains positions of the runways/taxiways and buildings, the navigation instruments provide the position of the aircraft, and on-board instruments provide the orientation of the aircraft (yaw, pitch and roll). We use this information to define regions of interest in the image where important features such as runways/taxiways, the horizon, etc. are likely to be present. Edges corresponding to these features of interest are detected within these regions. After delineating regions representing runway/taxiways, we look for objects inside and outside these regions.

The data from radio navigation instruments are known only upto a certain accuracy depending upon the type of radio navigation instruments. For example, GPS data is updated once every second and it is likely that a few such updates are missed making camera position data to be a few hundred feet off. On-board instrument data is generally useful to obtain more accurate camera position data than the GPS-based data. An alternative approach is to use the information about the location of detected objects in the images with known world coordinates (e.g. intersection of runways/taxiways, corners of buildings, etc.) to obtain an improved estimate of the camera position. This requires an analytical study of the relationships among the camera parameters, the resolution of the images, and the distances between the aircraft and objects.

In Section II we present a block diagram of the complete system. In Section III we describe the analytical model that establishes the relationship between the position, orientation and other physical parameters of the camera and the attributes of the captured images. This model is useful to calculate the accuracy of camera position estimation using image based features. In Section IV we present the method for defining the regions of interest in the image using the camera parameters and airport model. Section V includes image processing steps that are used to find regions corresponding to major features in the image and to detect objects in these regions. Experimental results are presented in Section VI. We conclude the paper with a summary and a brief description of future work.

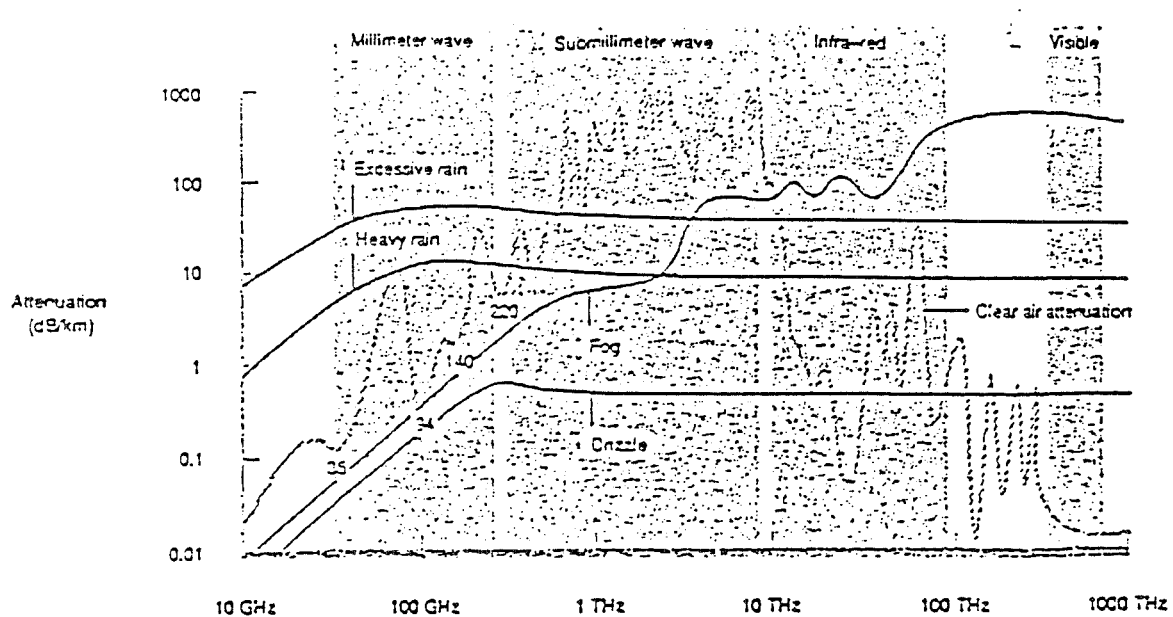


Fig. 1. Atmospheric effects on electromagnetic radiation [2].

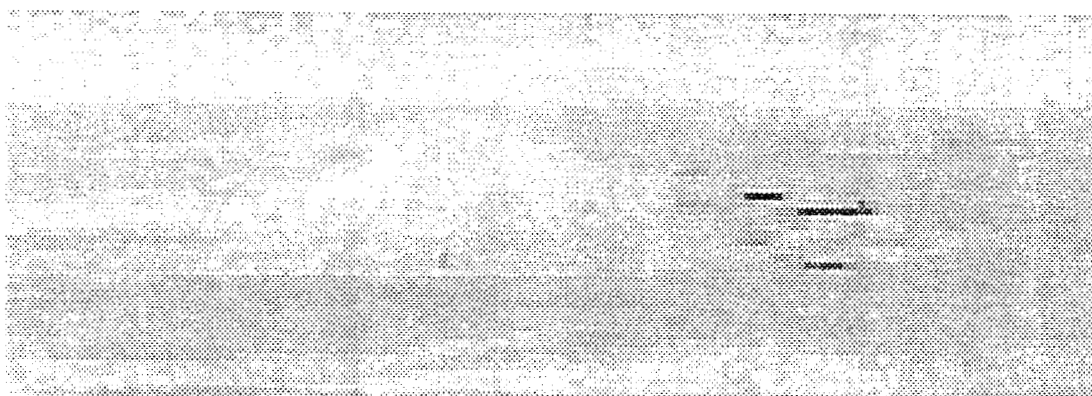


Fig. 2. The Passive Millimeter Wave image.

II. System Description

In this section, we describe the functions of various modules of the system shown in Fig.3 and the interactions between them. The input model of the airport contains positions of the runways/taxiways, and buildings. The *model transformation* module will take this model and the camera state information (position and orientation) as inputs to define the regions of interest in the image plane.

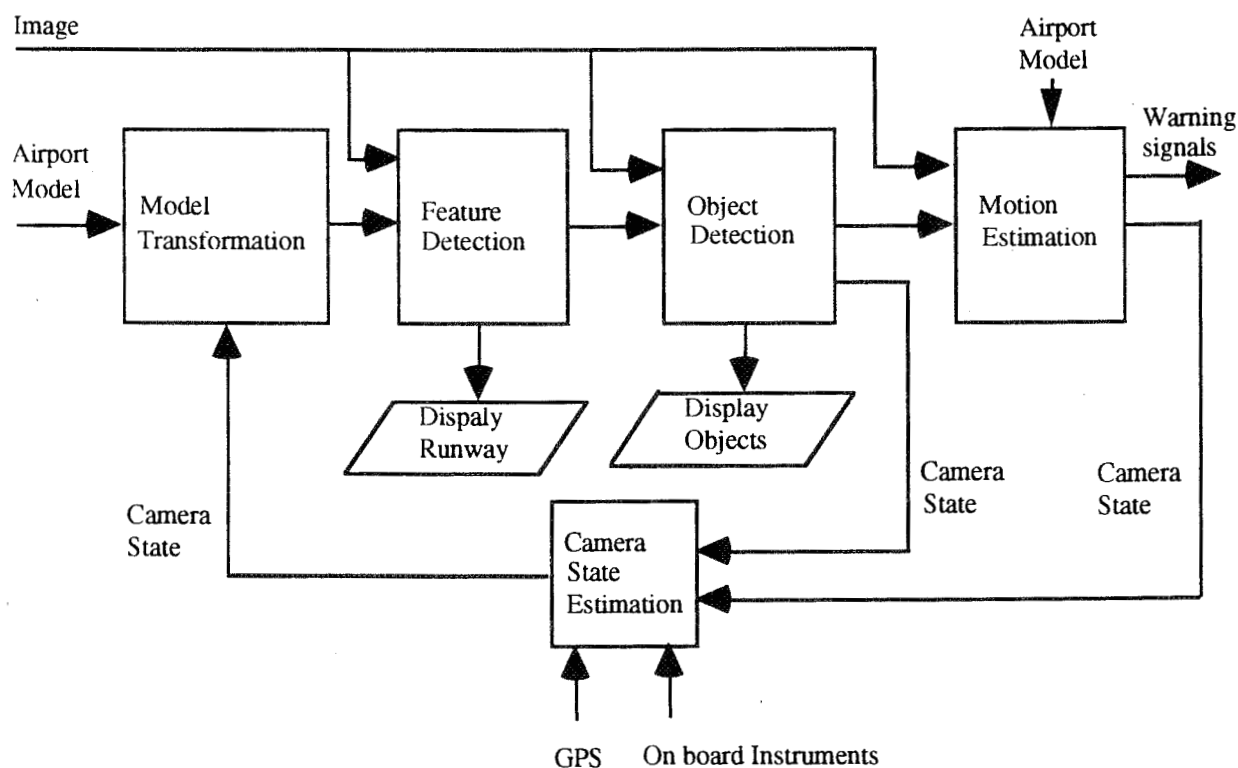


Fig. 3. System block diagram.

The image processing algorithms in the *feature detection* module operates within these regions of interest to detect the edges of the runway, horizon, etc. in the image. An edge is fitted to the edge pixels if enough edge pixels are found within the region of interest. The module outputs parameters which define major regions in the input image.

The *object detection* module detects objects in the image using different thresholds for each region. For example, since detection of objects on the runway is extremely important, a lower threshold is used to flag every object even if the contrast is low whereas a higher threshold is used to detect objects which are outside the runway such as buildings, etc. Locations of detected objects with known world coordinates is useful to estimate camera state parameters.

The *motion estimation module* uses dynamic scene analysis methods to estimate camera state parameters as well as to detect velocities of objects on the ground. The outputs from this module will be useful to detect potential collisions and generate warning signals as appropriate.

The *camera state estimation* module integrates information obtained about the position and velocity of the aircraft from various sensors and modules and outputs necessary data to the *model transformation* module.

III. Accuracy of Camera State Estimation from Image-based Features

As we need to use the camera state estimated from locating features of known objects in the image during the period when the GPS is not updated, it is necessary to know the accuracy of such estimated positions and the factors that decide the accuracy. Hence, an analytical model that establishes the relationship between the camera parameters and the attributes of captured images is necessary for guiding the image analysis system. Sensor positional parameters include range (distance from the aircraft to the runway threshold), cross range (distance from the aircraft to the runway center line), altitude, and pitch, roll and yaw angles. Sensor imaging attributes include the number of pixels in the image and the optical angular view measured in degrees. We derive the inter-relationships among these parameters. Using these relationships we calculate the accuracy of the estimate of camera position based on a minimum resolvable movement of features by one pixel in the image. We obtain these accuracies for three different types of cameras (PMMW, FLIR, HDTV) at six ranges.

A. Analysis

Throughout the analysis, for convenience, we assume that the sensor is located at the center of gravity of the airplane. Hence we can use the terms sensor position and aircraft position interchangeably. We also neglect the effect of curvature of the earth. The system of reference axis that forms the basis of system of notations used to describe the position of the sensor is shown in Fig. 4. The figure shows an airplane with three mutually perpendicular axes — pitch, roll and yaw — passing through the center of gravity of the airplane. The three angular displacements are termed pitch, roll and yaw as shown in Fig. 4. The image plane is assumed to be perpendicular to the rolling axis with its vertical and horizontal axes coinciding with the yawing and the pitching axis of the airplane, respectively.

Fig. 5 shows an imaging situation during landing where the aircraft is at (X_c, Y_c, Z_c) , with pitching angle θ , zero yaw and zero roll angle. Let $\alpha = 90 - \theta$. The field of view of the camera is determined by two viewing angles: $\Delta\alpha$ defined in the same plane as θ and $\Delta\beta$ at right angles to $\Delta\alpha$ ($\Delta\alpha$ determines the vertical extent of the image and $\Delta\beta$ its horizontal extent). Even though the image obtained by the sensor is always a rectangle, the ground area captured by the sensor is a trapezoid $ABCD$ whose side length and area depends on $\Delta\alpha$, $\Delta\beta$ and various other sensor

parameters like position, orientation etc. Note that a pixel in the image plane corresponds to a patch on the ground plane. We refer to this as a pixel-patch (see Fig. 6).

Consider a point feature which has been detected at some pixel (p, q) . Let the actual world coordinates of this feature be $(P, Q, 0)$. Since a pixel represents a patch on the ground, the camera could change in its position by certain amount while still retaining the image of the feature at the same pixel (p, q) . Hence a camera pose estimation by passive triangulation will always give the same camera pose for nearby camera positions unless the change in camera position is large enough for the feature to be observed in the neighboring pixel. We define this minimum change in camera displacement as the sensitivity of the camera. Note that this is a measure of accuracy of camera position estimate and is a function of the camera, image size in number of pixels, angular resolution, and the pixel location (p, q) in the image plane.

Let N_x and N_y represent the number of pixels in the vertical and horizontal directions, respectively. The pixels are numbered $-N_x/2, \dots, 0, \dots, N_x/2-1$ in the vertical direction and $-N_y/2, \dots, 0, \dots, N_y/2-1$ in the horizontal direction. The rolling axis of the plane is assumed to pass through the bottom right corner of the patch on the ground plane which corresponds to the center pixel in the image plane. Other pixels are referenced in a similar manner. The coordinates of the reference corner of the ground area covered by a pixel (p, q) can be estimated by the following relations.

$$\begin{aligned} X &= X_c + Z_c \tan(\alpha + p \frac{\Delta\alpha}{N_x}) \\ Y &= Y_c + \frac{Z_c}{\cos(\alpha + p \frac{\Delta\alpha}{N_x})} \tan(q \frac{\Delta\beta}{N_y}) \end{aligned} \quad (1)$$

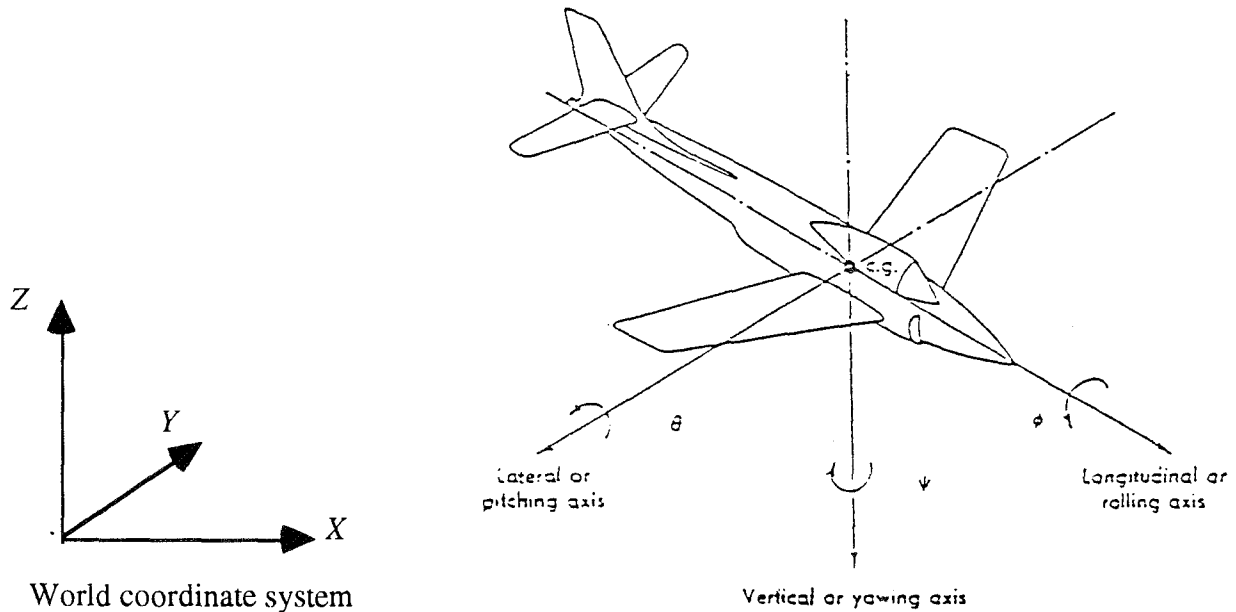


Fig. 4. Airplane-body axis (Reproduced from "Airplane Aerodynamics" by Dommasch and Danieol Otto [ed. 1967]).

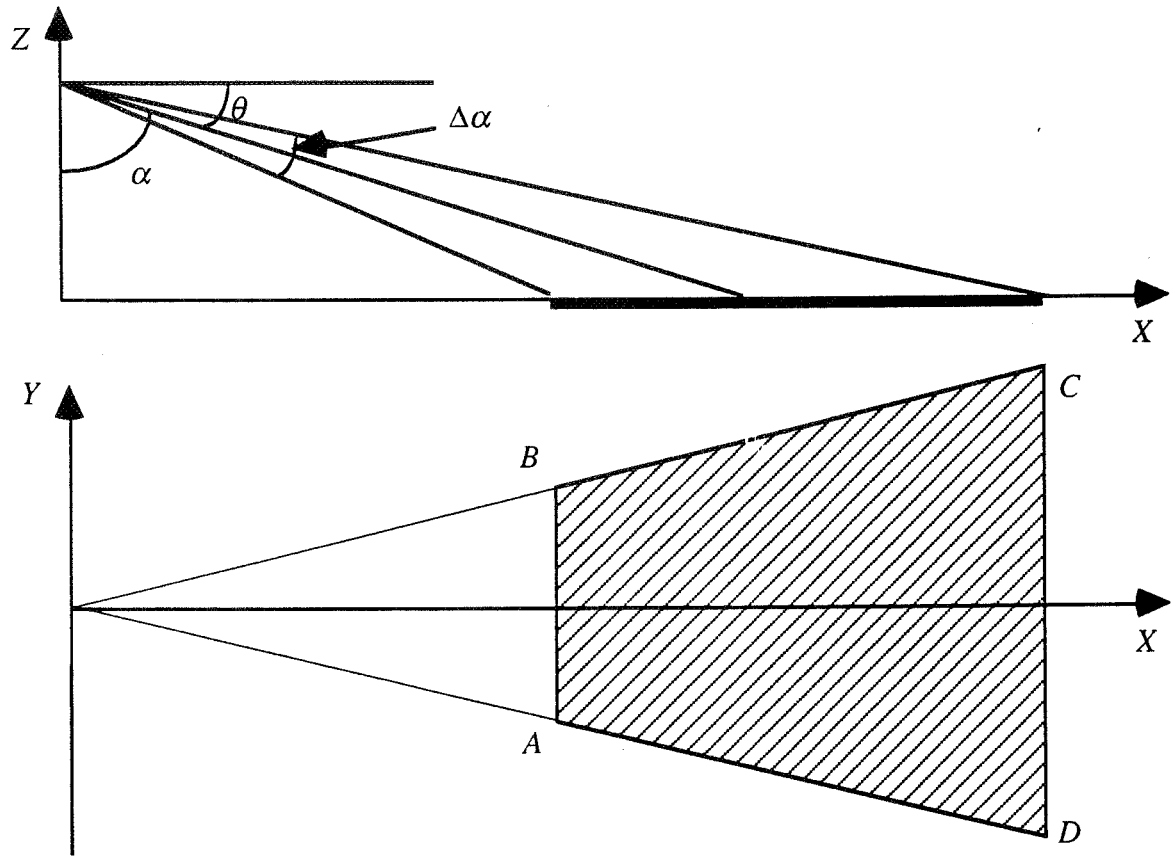


Fig. 5. Image obtained by the sensor is projected towards the ground.
Hatched portion is the ground area covered by the sensor.

For a non zero rolling angle ϕ , the ground coordinates (X', Y') which corresponds to a pixel (p, q) in the image plane are obtained by replacing (p, q) in the above equation by (p', q') , where

$$\begin{aligned} p' &= p \cos \phi - q \sin \phi, \\ q' &= p \sin \phi + q \cos \phi. \end{aligned} \quad (2)$$

Since a pixel-patch is referenced by its bottom right corner of the pixel, the other three corners become the reference of its three neighboring pixels-patch as shown in Fig. 7. Thus, the four corners of this pixel-patch, (X_i', Y_i') , $i=1,2,3,4$, are obtained by using Eq. (1), where (p, q) are replaced by (p_i', q_i') , where

$$\begin{aligned} p_i &= p_i \cos \phi - q_i \sin \phi, \\ q_i &= p_i \sin \phi + q_i \cos \phi. \end{aligned} \quad (3)$$

and $(p_1, q_1) = (p, q)$, $(p_2, q_2) = (p+1, q)$, $(p_3, q_3) = (p+1, q+1)$, and $(p_4, q_4) = (p, q+1)$.

Eq. (1) explicitly gives the relationship between the camera parameters (X_c, Y_c, Z_c) , θ , ϕ , and a ground point corresponding to a pixel (p, q) . We are now interested in computing the sensitivity of the imagery sensor. This is defined as the minimum change in a camera parameter that

would move a fixed ground point to the next pixel in the image plane. We obtain this by taking the partial derivative of X_I' and Y_I' with respect to the corresponding parameter. For example,

$$D_{X_c}^X = \frac{\partial X_1''}{\partial X_c}, \text{ and } D_{X_c}^Y = \frac{\partial Y_1'}{\partial X_c}. \quad (4)$$

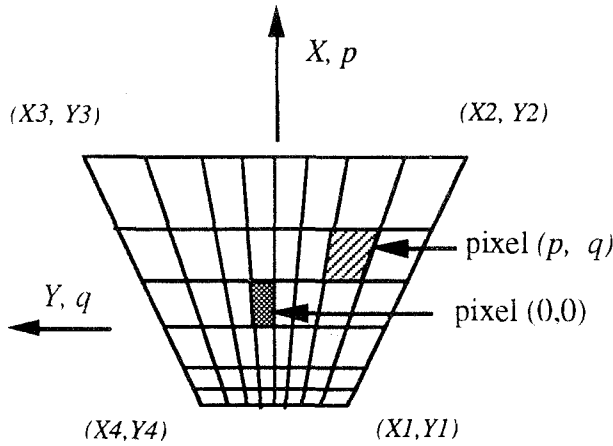


Fig. 6. Ground area covered by the sensors. Each small trapezoid corresponds to a pixel in the actual image.

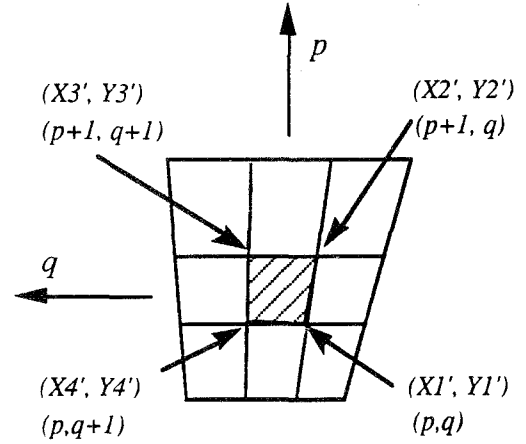


Fig. 7. A pixel (p, q) projected towards the ground.

This derivation is an approximation to the amount of change in X_1' for unit change in X_c . Thus we estimate that the amount of change in X_c in order to change X_1' to X_2' , or Y_1' to Y_4' (which define the corners of adjacent pixels) as

$$S_{X_c} = \frac{(X_2' - X_1')}{D_{X_c}^X}, \text{ and } S_{X_c}^Y = \frac{(Y_4' - Y_1')}{D_{X_c}^Y}. \quad (5)$$

Note that $S_{X_c}^Y = \infty$, as expected. Sensitivity with reference to other parameter is defined in a similar manner. These are summarized in Table I.

Sensor sensitivity is a function of various sensor parameters and sensor attitudes. Since the sensor plane is inclined to the ground plane, the sensitivity varies in the vertical and horizontal direction along the sensor plane and hence is a function of pixel number (p, q). Equivalently, the accuracy of estimation of sensor position using ground truth data is a function of pixel position as well as other parameters. For a given range, the estimation using features that are observed at the top half of the sensor are less accurate because of the large ground area represented by these pixels. Also for a given p , the accuracy decreases as we move towards the border of the sensor in the horizontal direction. In summary, the accuracy of estimation is a function of sensor characteristic and the ratio of the sensor view angle to the number of pixels in the image.

SPP		Sensor Sensitivity at (p, q)	Sensor Sensitivity at (0, 0) with $\phi = 0$
X_C	$S_{X_c}^X$	$2 Z_C \sin(\cos\phi, \Delta\alpha/N_X) / \{\cos[2\alpha + \Delta\alpha/N_X \{(2p+1)\cos\phi - 2q\sin\phi\}] + 1\}$	$2 Z_C \sin(\Delta\alpha/N_X) / \{\cos[2\alpha + \Delta\alpha/N_X] + 1\}$
	$S_{X_c}^Y$	∞	∞
Y_C	$S_{Y_c}^X$	∞	∞
	$S_{Y_c}^Y$	$\{Z_C \tan(q_4' \Delta\beta / N_Y) / \cos(\alpha + p_4' \Delta\alpha / N_X)\} - \{Z_C \tan(q_1' \Delta\beta / N_Y) / \cos(\alpha + p_1' \Delta\alpha / N_X)\}$	$2 Z_C \sin(\Delta\beta/N_Y) / \{\cos\alpha \cdot \cos(\Delta\beta/N_Y) + 1\}$
Z_C	$S_{Z_c}^X$	$S_{X_c}^X / \tan(\alpha + p_1' \Delta\alpha / N_X)$	$2 Z_C \sin(\Delta\alpha/N_X) / \sin(2\alpha + \Delta\alpha/N_X)$
	$S_{Z_c}^Y$	$S_{Y_c}^Y \cos(\alpha + p_1' \Delta\alpha / N_X) / \tan(q_1' \Delta\beta / N_Y)$	∞
θ	S_{θ}^X	$S_{X_c}^X \cos^2(\alpha + p_1' \Delta\alpha / N_X) / Z_C$	$\sin(\Delta\alpha/N_X) / \{\cos\alpha / \cos(\alpha + \Delta\alpha/N_X)\}$
	S_{θ}^Y	$S_{Y_c}^Y \cos^2(\alpha + p_1' \Delta\alpha / N_X) / Z_C \tan(q_1' \Delta\beta / N_Y) \sin(\alpha + p_1' \Delta\alpha / N_X)$	∞
ϕ	S_{ϕ}^X	$S_{X_c}^X \cos^2(\alpha + p_1' \Delta\alpha / N_X) / (Z_C \Delta\alpha / N_X) (-p \sin\phi - q \cos\phi)$	∞
	S_{ϕ}^Y	$S_{Y_c}^Y / Z_C [A \delta B / \delta\phi + B \delta A / \delta\phi]$	∞
$A = 1/\cos(\alpha + p_1' \Delta\alpha / N_X)$; $B = \tan(q_1' \Delta\beta / N_Y)$; $\delta B / \delta\phi = (p \cos\phi - q \sin\phi) (\Delta\beta / N_Y) \cos^2(q_1' \Delta\beta / N_Y)$ $\delta A / \delta\phi = \tan(\alpha + p_1' \Delta\alpha / N_X) (-p \sin\phi - q \cos\phi) (\Delta\alpha / N_X) \cos(\alpha + p_1' \Delta\alpha / N_X)$; $\alpha = 90 + \theta$; $(p_1, q_1) = (p, q)$; $(p_4, q_4) = (p, q+1)$; $p_1' = p_1 \cos\phi - q_1 \sin\phi$; $p_4' = p_4 \cos\phi - q_4 \sin\phi$; $q_1' = p_1 \sin\phi + q_1 \cos\phi$; $q_4' = p_4 \sin\phi + q_4 \cos\phi$;			
SPP: Sensor Positional Parameters		Sensor Characteristics	
(X_C, Y_C, Z_C)	Sensor position	Vertical	Horizontal
θ	Pitch angle	$\Delta\alpha$	$\Delta\beta$
ϕ	Roll angle	N_X	N_Y
Sensitivity: Minimum change in the sensor positional parameters ($X_C, Y_C, Z_C, \theta, \phi$) that will make the object to appear in the next pixel either in the vertical (X; hence called as sensitivity in x direction) or in the horizontal (Y; hence called as sensitivity in y direction) direction of the sensor plane. S_i^j : Sensitivity in the direction 'j' due to the sensor positional parameter 'i' computed at pixel (p, q) in the image plane.			

Table I. Sensor positional sensitivity equations.

B. Quantitative Results and Discussions

The sensitivity analysis described in the previous section was applied to three different sensors at six different positions (Table II). Sensitivities $S_{X_c}^X$, $S_{Y_c}^Y$ and $S_{Z_c}^X$, at the aim point (i.e., $p=0$, $q=0$) for various sensor positions are plotted in Figs. 8, 9 and 10 respectively. Note that $S_{Z_c}^Y$ is larger than $S_{Z_c}^X$ at (0, 0) and hence a feature would move to the next horizontal pixel before it moves to the next vertical pixel. Thus only $S_{Z_c}^X$ is important.

As expected, the sensitivity is the best for the sensor with the highest pixel resolution. Sensitivity also improves as the sensor is moved closer to the ground. It becomes poor for the features that are located at the far end of the vertical axis (top of the sensor), i.e., for the objects that are located at the far end of the runway. Thus, as expected, the position and velocity of the aircraft can be computed to a better accuracy by knowing the position of stationary objects on the ground that are closer to the aircraft.

The results indicate that the accuracy of camera state estimation would be no better than the GPS data unless a high resolution sensor is employed. Note that these results do not consider potential improvements that can be obtained by motion stereo techniques using a large number of image frames. We are presently investigating the possibility of improving the accuracy of the computed sensor positional parameters by extending our analysis using this method.

Sensor Characteristic

<i>Sensor type</i>	<i>Pixel (H x V)</i>	<i>Field of View (H x V) deg.</i>
HDTV	1920 x 1035	30 x 24
FLIR	512 x 512	28 x 21
MMW	80 x 64	27 x 22

Sensor Positions

<i>Location</i>	<i>Range in ft.</i>	<i>Altitude in ft.</i>
Threshold	0.0	50.0
CAT II - DH	908.1	100.0
CAT I -DH	2816.2	200.0
Middle Marker	4500.0	288.2
1000' Altitude	18081.1	1000.0
Outer Marker	29040.0	1574.3

In all the above six cases
 Pitch angle -3.0 degree
 Roll angle 0.0 degree
 Cross Range 0.0 ft.

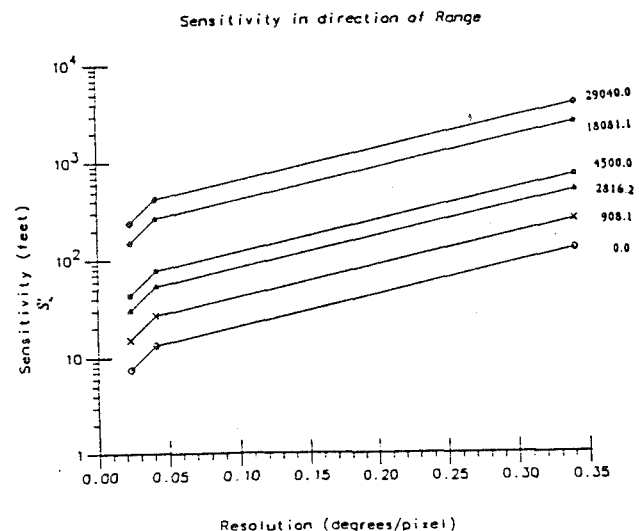


Table II.

Fig. 8.

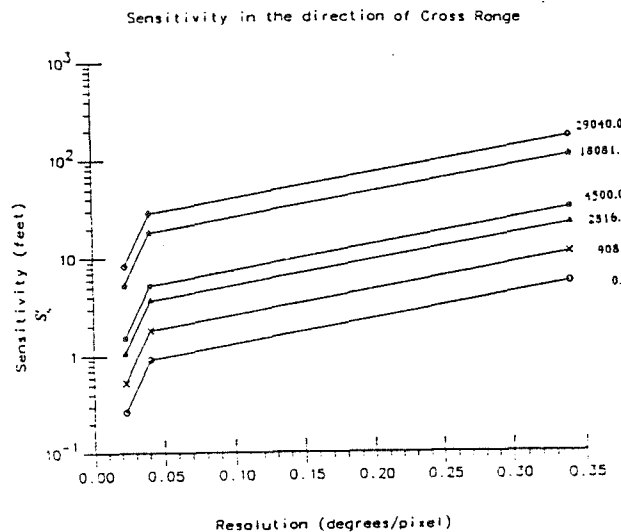


Fig. 9.

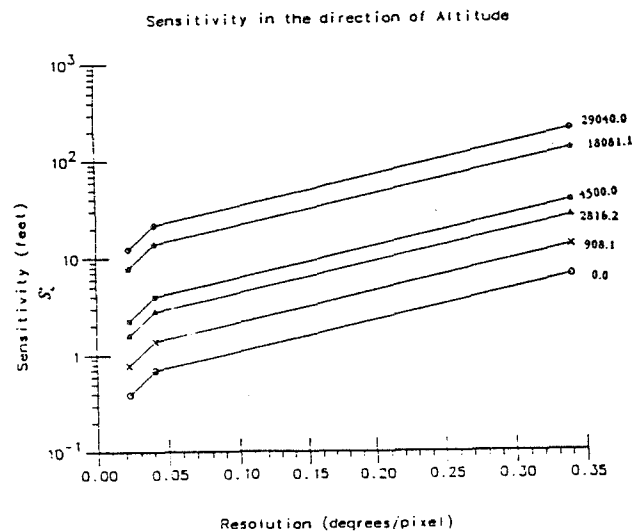


Fig. 10.

IV. Model Transformation

As noted earlier, the PMMW images are low contrast-low resolution images. Simple edge detection techniques on these images generate many noisy edge pixels in addition to those belonging to the true edges such as runways, sky etc. This problem is alleviated by defining regions where the true edges are expected to occur using knowledge about the aircraft position and a model of the airport. The main functions of the *model transformation* module is to define a region of interest on the ground plane for each feature in the model and to perform 3D to 2D transformation. It also defines a region in the image plane where the horizon line should occur.

A. Defining Regions of Interest for Runway Edges

The error in the expected location of a feature and its actual position in the image depends on several factors, most notably the accuracy of the camera position parameters used by the model transformation module. Furthermore, it is evident from our earlier analysis (Fig.6) that the ground area covered by a pixel is a function of the position of the pixel in the image. Thus it is not reasonable to define the search space for each feature as a fixed number of pixels centered around the expected location in the image plane. Hence we define the region of interest in the 3D space and then apply transformation to get the corresponding region of interest in the image. The extent of the search space in the 3D space is determined by the estimated error in camera positional parameters (which are based on GPS and on-board instrument data).

The geometric model of the airport contains a sequence of 3D coordinates of the vertices of the runway/taxiways, which forms a polygon with n vertices:

$$runway = \{P_i\}, i=1, 2, \dots, n,$$

where $P_i = (X_i, Y_i, Z_i)^T$ is one of the vertices of the polygon. Note that $Z_i = 0$. $P_i P_{i+1}$ specifies an edge of the polygon. The region of interest is defined as a rectangle on the ground which encloses the edge. Therefore, each edge $P_i P_{i+1}$ of the polygon is associated with the region of interest

defined by four points $b_j = (X_j, Y_j, Z_j)^T, j=1, \dots, 4$, and $Z_j = 0$.

The width of the region of interest is defined as a function of the width of the runway/taxiway, w , accuracy of the GPS data, g ($g \leq 1$), and the accuracy of the on-board instruments, d ($d \leq 1$). Note that g and d are determined by the specification and characteristics of these instruments. This relationship is given by

$$\text{width}(w, g, d) = \frac{0.2w}{gd}. \quad (6)$$

Note that the minimum width is $0.2w$ when $g=d=1$, which corresponds to $\pm 10\%$ potential displacement of runway edge feature. To limit the search area from being a large fraction of the runway width we limit the search width to $0.4w$ even if $gd < 0.5$.

After defining the region of interest for each edge, 3D to 2D coordinate transformation is performed using the following homogeneous equation [3]:

$$\begin{bmatrix} \lambda \rho \\ \lambda q \\ \lambda r \\ \lambda \end{bmatrix} = [P][R][T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \quad (7)$$

where

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{f} & 0 & 0 & 0 \end{bmatrix} \quad (8)$$

$$R = \begin{bmatrix} -\cos(\psi)\cos(\theta) & -\sin(\psi)\cos(\theta) & -\sin(\theta) & 0 \\ \cos(\psi)\sin(\theta)\sin(\phi) - \sin(\psi)\cos(\phi) & \sin(\psi)\sin(\theta)\sin(\phi) + \cos(\psi)\sin(\phi) & -\cos(\theta)\sin(\phi) & 0 \\ \cos(\psi)\sin(\theta)\cos(\phi) + \sin(\psi)\sin(\phi) & \sin(\psi)\sin(\theta)\cos(\phi) - \cos(\psi)\cos(\phi) & -\cos(\theta)\cos(\phi) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (9)$$

$$\text{and } T = \begin{bmatrix} 1 & 0 & 0 & -X_c \\ 0 & 1 & 0 & -Y_c \\ 0 & 0 & 1 & -Z_c \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (10)$$

are the perspective projection, rotation and translation transformation matrices, respectively, and f is the focal length. After perspective projection, we need to consider the following special cases:

- A. the region of interest degenerates to a line in the image plane because the region is too far from the camera,

B. the region of interest in the image plane becomes very large because the edge is very close to the camera.

For case A, a minimum width in the image plane is assigned in order to provide some search space for the feature detector. For case B, a maximum width in image space is defined to further restrict the region. In our experiment, for the aforementioned extreme cases, the minimum and maximum width of a region of interest are set to be 10 and 20 pixels, respectively.

B. Defining Search Space for Horizon Line

When the vertical angular field of view is larger than 2θ , then a horizon line appears in the image (Fig. 11). The horizon is an important clue in estimating the camera orientation since it gives the roll angle information directly. Search space in the image plane is defined to locate this line.

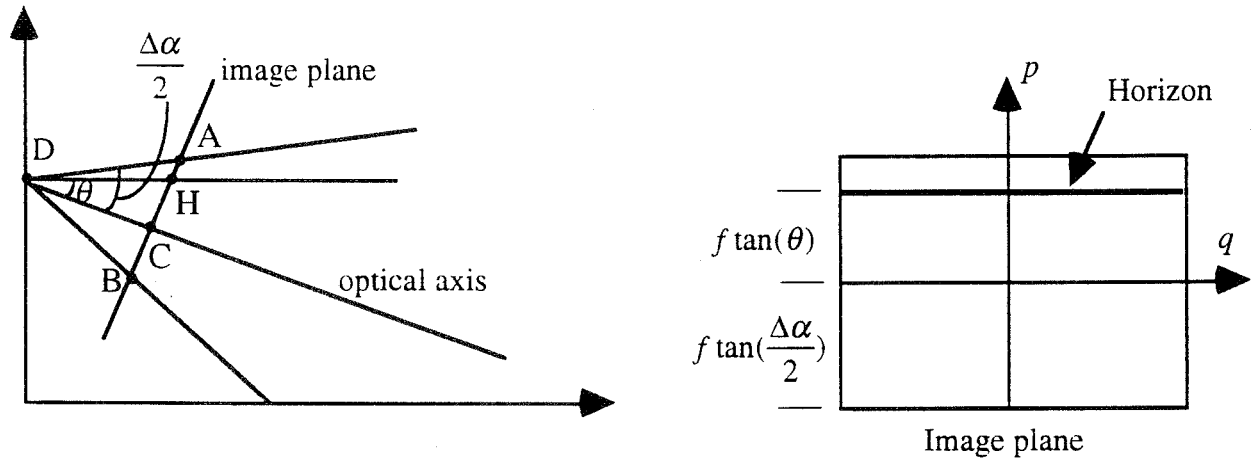


Fig. 11. Horizon line in the image.

Without loss of generality, consider the situation when the aircraft is heading towards the X axis of the world coordinate system. Assume the camera is located at point D (see Fig. 11) with pitch angle θ , and zero yaw and roll angles. Points A and B are on the top and bottom edge of the image, respectively. The horizon will then appear horizontally in the image plane as shown. The distance between this line and the center line of the image is given by $\overline{HC} = f \tan(\theta)$. Since in the above analysis roll angle has been assumed to be zero, the horizon appears parallel to the horizontal axis of the image plane. For any non zero roll angle, a simple roll transformation on this line will give the horizon in the image. The associated region of interest is defined to be 10 pixels centered around the expected horizontal line in the image.

It is possible for the projection of the region of interest onto the image plane to be partially outside the image boundary. In such cases, we need to clip these regions so that the search space always remains within the confines of the image. This is done using the “polygon clip and fill” algorithm [4]. The regions of interest for both the runway and the horizon of the image sequence used in these experiment are shown in Fig.12.

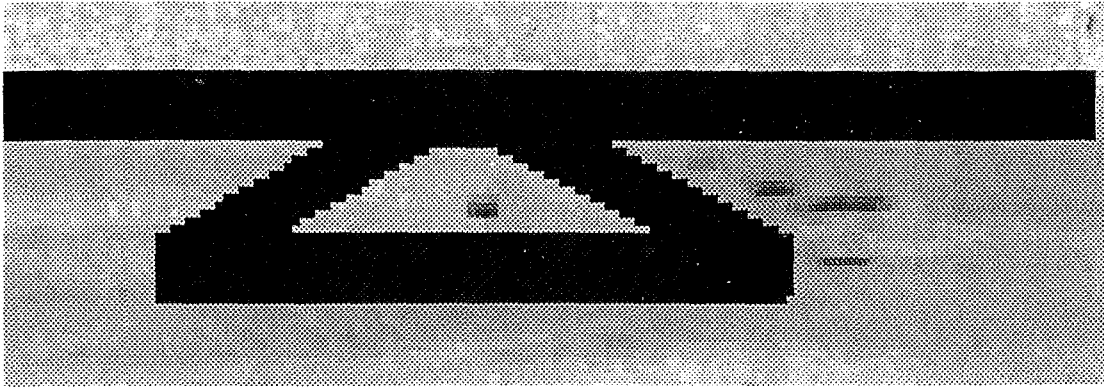


Fig. 12. Regions of interest.

V. Runway Localization and Object Detection

A. Runway Localization

In this part of the system we search for the expected features within the region of interest, defined by the previous module. This will significantly reduce the search time and also avoid the spurious response which is likely in such a low resolution input image. An accurate localization of the feature is necessary for estimation of motion parameters and camera pose.

A Sobel edge detector is applied to the sensor image. We then select one of the four scanning directions (-45° , 0° , 45° , 90°) which is approximately orthogonal to the direction of the expected edge. Along each scan line we locate pixels with greatest edge strength. As the runway edge is supposed to be a straight line we fit a best line to these pixels. We also associate a measure of confidence for these detected edges based on the number of edge pixels detected along the line.

B. Object Detection

In this section, the region inside and outside the runway/taxiways are separately checked for the existence of any stationary or moving objects. The image has three homogeneous regions, namely the sky, the runway/taxiways and the region outside the runway/taxiways. Any objects on or outside the runway/taxiways are expected to have some deviation in graylevel from their respective homogeneous background. Hence, we use histogram-based thresholding for object detection. The thresholds which determine this deviation are set to be different for different regions.

We generate a mask image which represents three homogeneous regions. Using this mask image, we generate the histogram and compute its standard deviation for each region separately (except for the sky region). The threshold value is determined as a function of the mean and the standard deviation, and any area which has graylevel lower than the threshold is considered as object regions. An object is assumed to have a reasonable size. This size restriction on the object can be used to ignore spurious responses resulting from the thresholding. Each object is then labeled based on 4-connectivity.

VI. Experimental Results

We have tested our algorithm on a test image provided by the TRW. This image was obtained using a single pixel camera located at a fixed point in space (a camera with an array of pixels is under development). The camera was mechanically scanned to obtain a 50X150 pixel image. This is the image shown in Fig. 2. We were also provided with the model of the runway giving the 3D world coordinates of the runway corners, locations of the buildings etc. Using these data and the single image, we created a sequence of 30 frames to simulate the images from a moving camera. Frames 1 (original), 5, 10, and 15 from this sequence is shown in Fig.13(a). Edge enhanced images corresponding to these frames are shown in Fig.13(b). The regions of interest defined on these frames are shown in Fig.13(c). Delineated features superimposed on the images are shown in Fig.13(d). Although all the edges are detected accurately in this example, it is likely that one or more edges of a polygon are not detected. To handle such situations we associate a degree of importance for each edge. For example, runway edges which are closer to the camera must be detected in the image whereas those corresponding to the far end of the runway are usually very short and may or may not be detected. And overall confidence measure is associated with each detected region.

Objects detected on the runway in Frame 1 and those outside the runway are shown in Fig.14. Warning signals are generated for each object on or near runway. Algorithms to track these in successive frames and estimate camera state using motion stereo are under development.

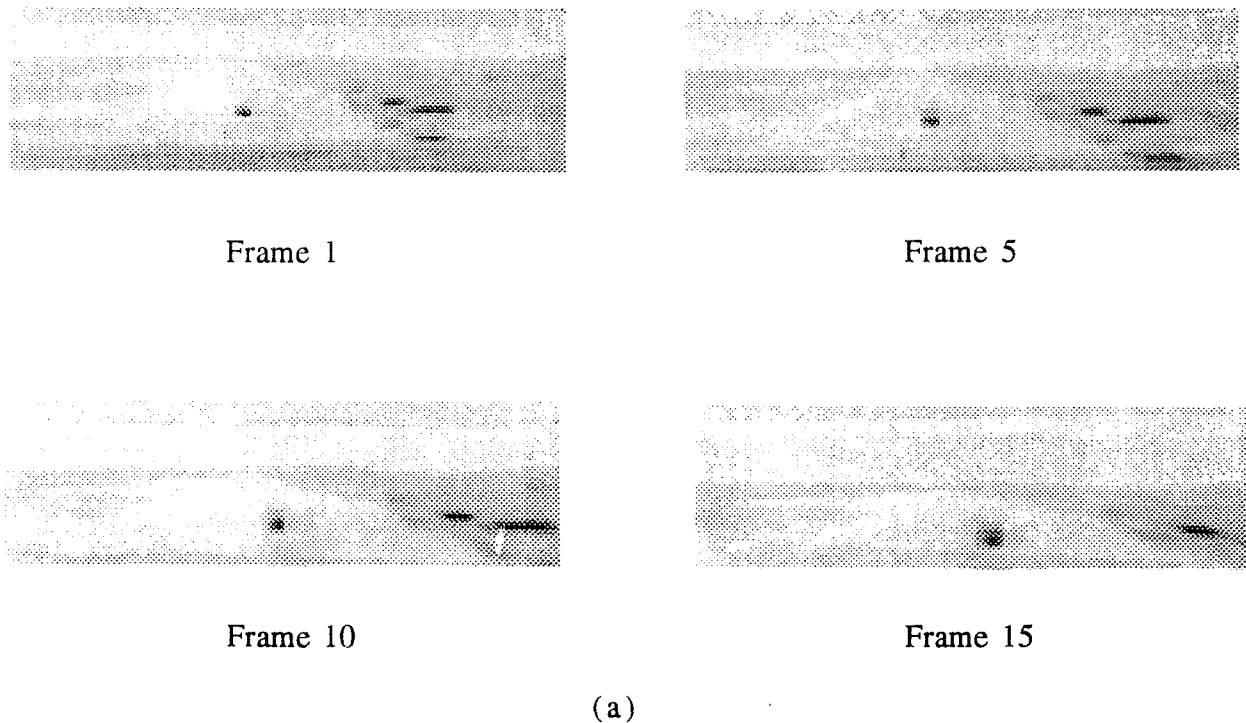
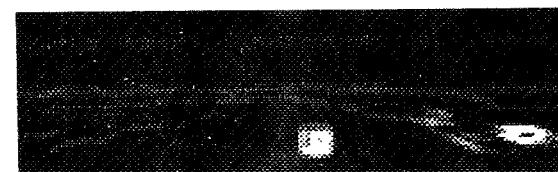
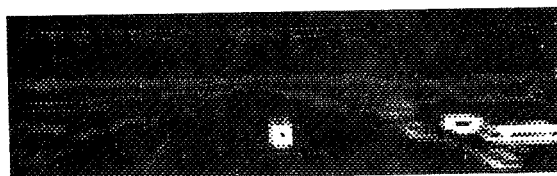
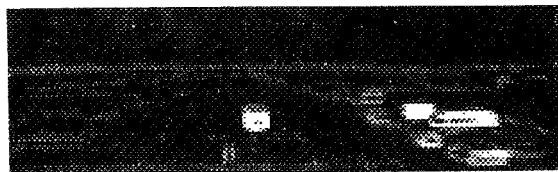
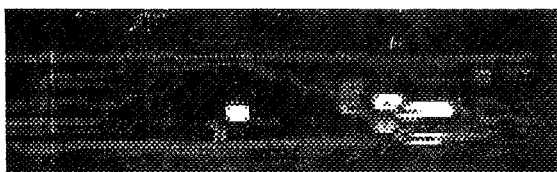
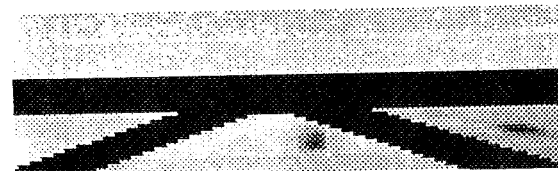
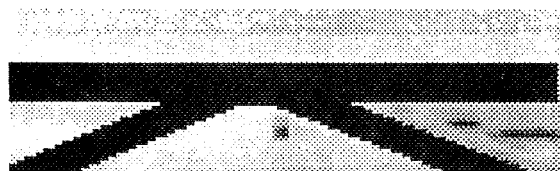
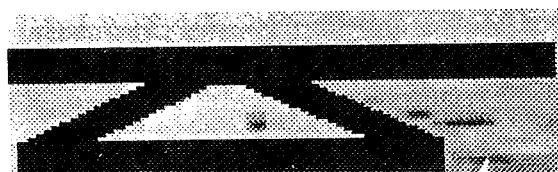
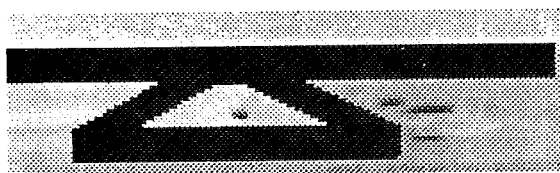


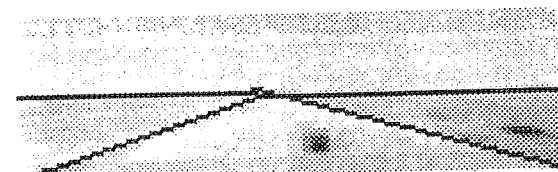
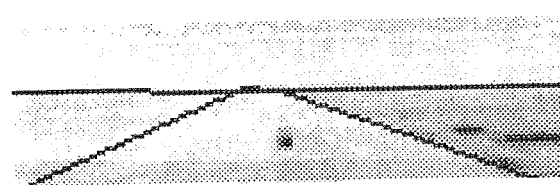
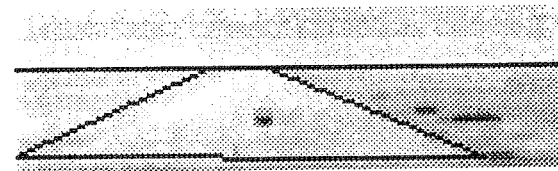
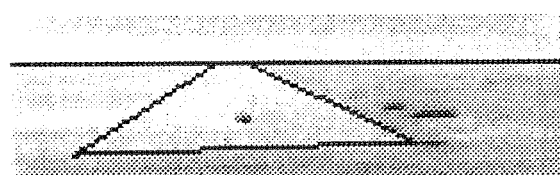
Fig. 13. The input images (a), edge images (b), regions of interest (c), and detected features superimposed on the original images (d).



(b)



(c)



(d)

Fig. 13. (continued)

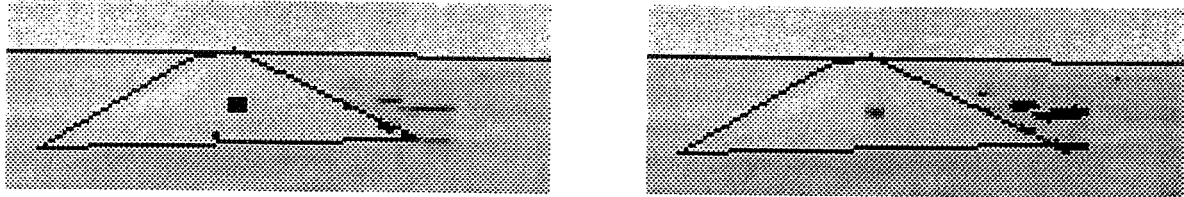


Fig. 14. Detected objects inside (left) and outside (right) the runway.

VII. Future Work and Conclusions

In this paper, we have described a vision-based system to assist pilots during landing under restricted visibility conditions. The images obtained by a passive sensor is processed to detect major regions such as runways and objects inside and outside these regions. The image resolution is very poor; however, additional information in the form of airport geometric model, and camera position parameters are available to guide the segmentation algorithms. Objects are detected in each of these regions using thresholds computed separately for each region. Our results show that the model-based feature detection approach is quite accurate and the homogeneity assumption on regions for object detection is reasonable. The success of this model-based approach clearly depends upon the accuracy of the camera position parameters used to define search regions in the image. One of the methods for updating camera position information is triangulation using known objects. We have derived the accuracy of such an update as a function of camera characteristics and image parameters.

At this stage, our system is able to detect the runway/taxiways and the objects inside and outside the runway/taxiways in each frame and to report their positions in the image. Since we have a moving camera, moving object situation, even the stationary objects appear to be moving in the image. Work is in progress to estimate the egomotion of the camera, to distinguish moving objects from stationary ones and to estimate the velocities of the moving objects. There is also potential to obtain more accurate camera state estimation using motion stereo from image sequences compared to using GPS data alone.

References

- [1]. Hatfield, J.J. and R.V. Parrish, "Advanced Cockpit Technology for Future Civil Transport aircraft," *Proc. 11th Annual IEEE/AESS Dayton Chapter Symposium*, Dayton Ohio, Nov. 1990.
- [2]. Young, S.K., R.A. Davidheiser, B. Hauss, P.S.C. Lee, M. Mussetto, M.M. Shoucri, and L. Yujiri, "Passive Millimeter-Wave Imaging," *TRW Space & Defense — Quest*, Vol. 13, No.2, pp. 3–20, Winter 1990/1991.
- [3]. Smith, P.N., "NASA Image Data Base User's Guide," NASA Ames Research Center, Moffett Field, CA., Version 1.0, 1990.
- [4]. Foley, J.D., A. van Dam, S.K. Feiner, and J.F. Hughes, "Computer Graphics—Principles and Practice," Addison-Wesley Publishing Co., 2nd. ed., 1990.

Image Processing for Flight Crew Enhanced Situational Awareness

Barry Roberts
Honeywell Systems and Research Center

ABSTRACT

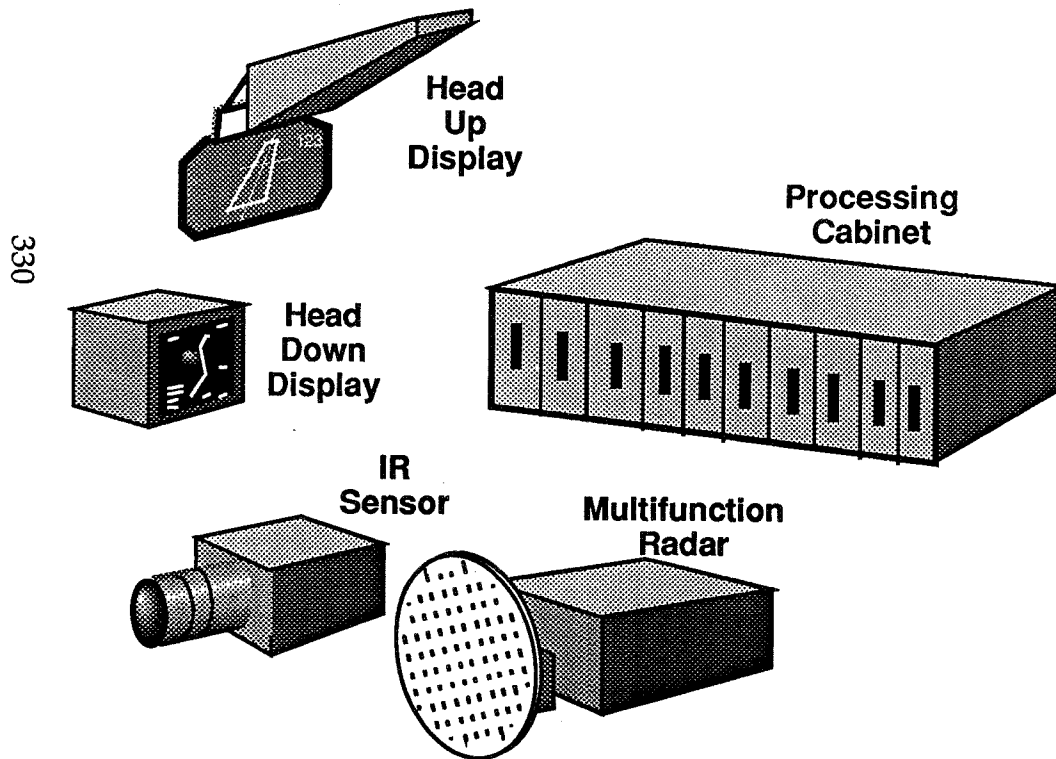
This presentation describes the image processing work that is being performed for the Enhanced Situational Awareness System (ESAS) application. Specifically, the presented work supports the Enhanced Vision System (EVS) component of ESAS.

518-54/
205 748
p. 33
24837

Enhanced Situation Awareness System (ESAS)

ESAS Functions

- ⇒ SVS/EVS
- ⇒ Ground Taxi & Takeoff
- ⇒ Weather RADAR
- ⇒ Windshear/Microburst Detection
- ⇒ CAT Detection
- ⇒ CFIT Avoidance
- ⇒ Wake Vortex Detection
- ⇒ Dry Hail Detection



The processing of imagery and its display to the flight crew will enhance aircraft operation in three areas.

- ⇒ Airplane pitch and roll stabilization
- ⇒ Adverse weather landing guidance
- ⇒ Runway incursion/obstacle avoidance

Currently, the information in the processed imagery is conveyed to the flight crew through a display of raster imagery and/or extracted features on a Head Up Display (HUD).

Image Processing for the Enhanced Vision Function of ESAS

✈ Information obtained from imaging sensors is displayed to the flight crew to enhance aircraft operation

- ⇒ Airplane pitch and roll stabilization
- ⇒ Adverse weather landing guidance
- ⇒ Runway incursion/obstacle avoidance

✈ Information presentation methods

- ⇒ Enhanced sensor images for HUD raster presentation
- ⇒ Sensor derived information for symbolic/synthetic HUD stroke presentation

Sensor Image Enhancement is required to compensate for sensor artifacts, reduce image noise level, etc.

Image Metrics are computed to aid in decisions regarding the application of sensor fusion, assessment of image quality and image information content, etc.

Sensor Feature Extraction produces image features corresponding to runway and taxiway edges/boundaries, the location of runway lights, etc. The features that are produced are subsequently displayed in symbolic form on a HUD or are used in registration of images on the HUD.

Areas of Image Processing Research for ESAS

✈ **Sensor Image Enhancement**

✈ **Image Metrics**

✈ **Sensor Feature Extraction**

The end result of sensor image enhancement is improved imagery for HUD display. Also, the resulting imagery is used by the feature extraction and image metric algorithms.

Sensor Image Enhancement

✈ **Beam sharpening for MMW Radar data**

⇒ Compensate for transfer function of wide beam antenna

✈ **Contrast enhancement and noise cleaning for MMW Radar data**

⇒ Range adaptive contrast enhancement

⇒ Noise filtering and edge preserving smoothing

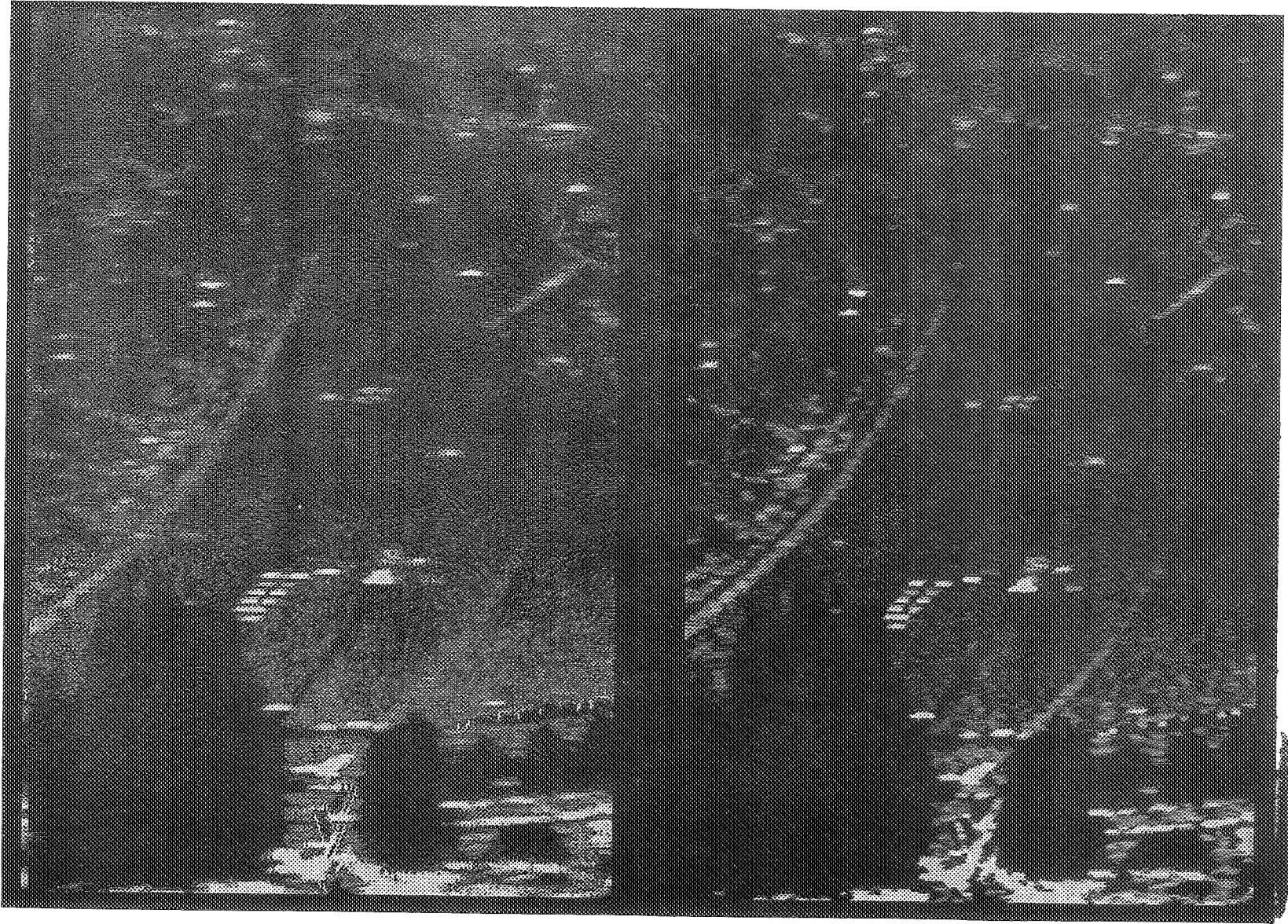
Right image: the raw imagery prior to beam sharpening

Left image: the beam sharpened image.

The Right image provides a factor of 2 improvement in image clarity/resolution which yields an improvement in the ability to distinguish adjacent objects in the radar image.

(The photo copying process doesn't do justice to the imagery.)

Beam Sharpening Results

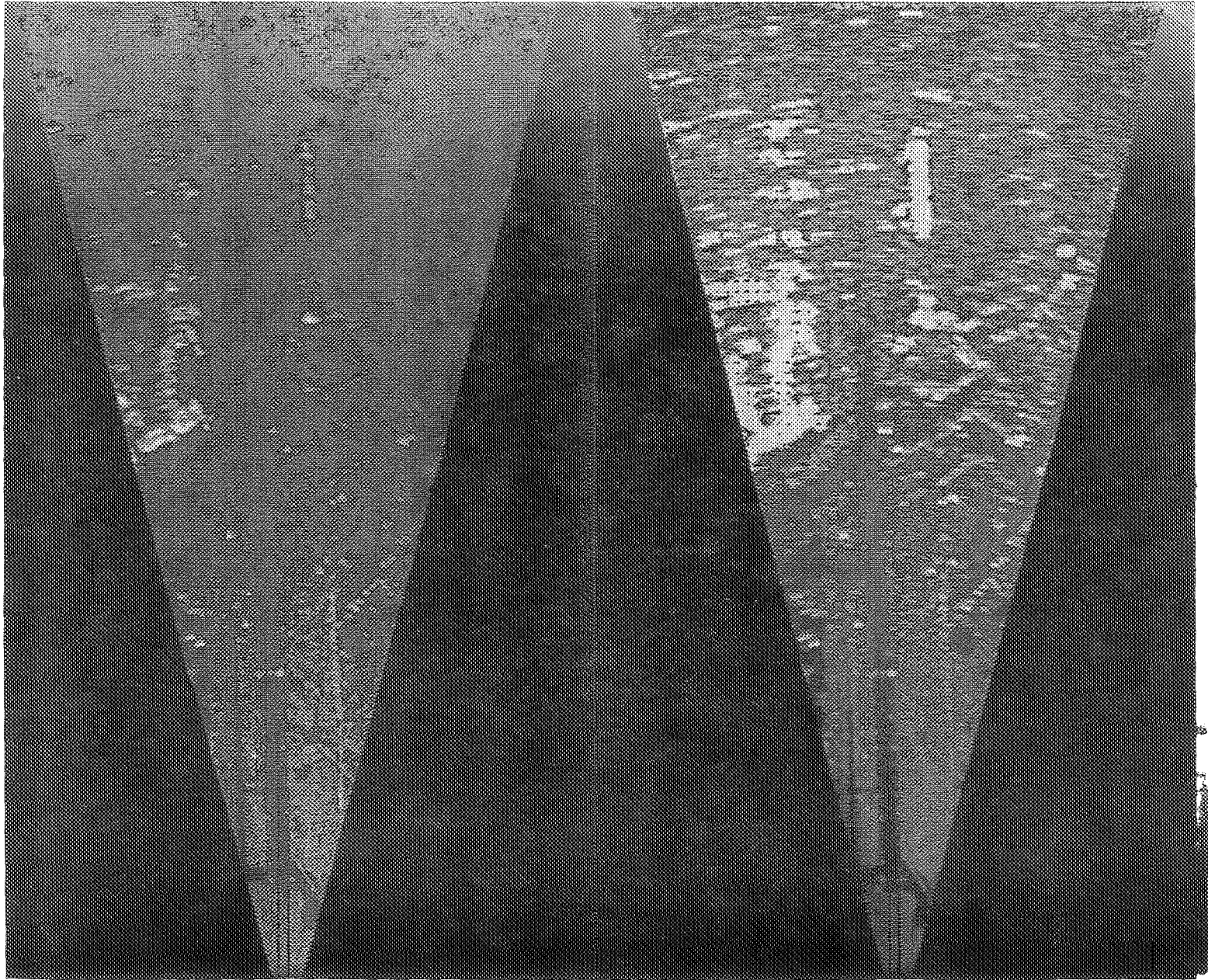


Left image: original radar image in B scan display (obtained from 35 GHz imaging radar; the site is Pt. Magu NAS).

Right image: range-adaptive, contrast enhanced image.

This contrast enhancement process provides range-adaptive gain control, which is determined from an empirically verified sensor model, to yield improved detail in the image at the far ranges.

Range Adaptive Contrast Enhancement



Multiple image metrics are being considered for use in characterizing/evaluating image quality/content. One application of these metrics is to control the application of a sensor fusion process.

Edge energy metric = local average of edge magnitude (the edge image is produced by a Sobel operator applied to the original image).

Contrast metric = convolution of two windows; one inside of the other.

$$Metric = \left[\frac{[(\text{Mean pixel value within large window}) - (\text{Mean pixel value within small window})]^2}{(\text{Standard deviation of pixel values within large window})} \right]^2$$

Sensor Image Evaluation

✈ Sensor Image Metrics

- ⇒ Edge energy
- ⇒ Spatial frequency content
- ⇒ Local variance
- ⇒ Contrast metrics
- ⇒ Texture metrics (e.g., co-occurrence matrices)

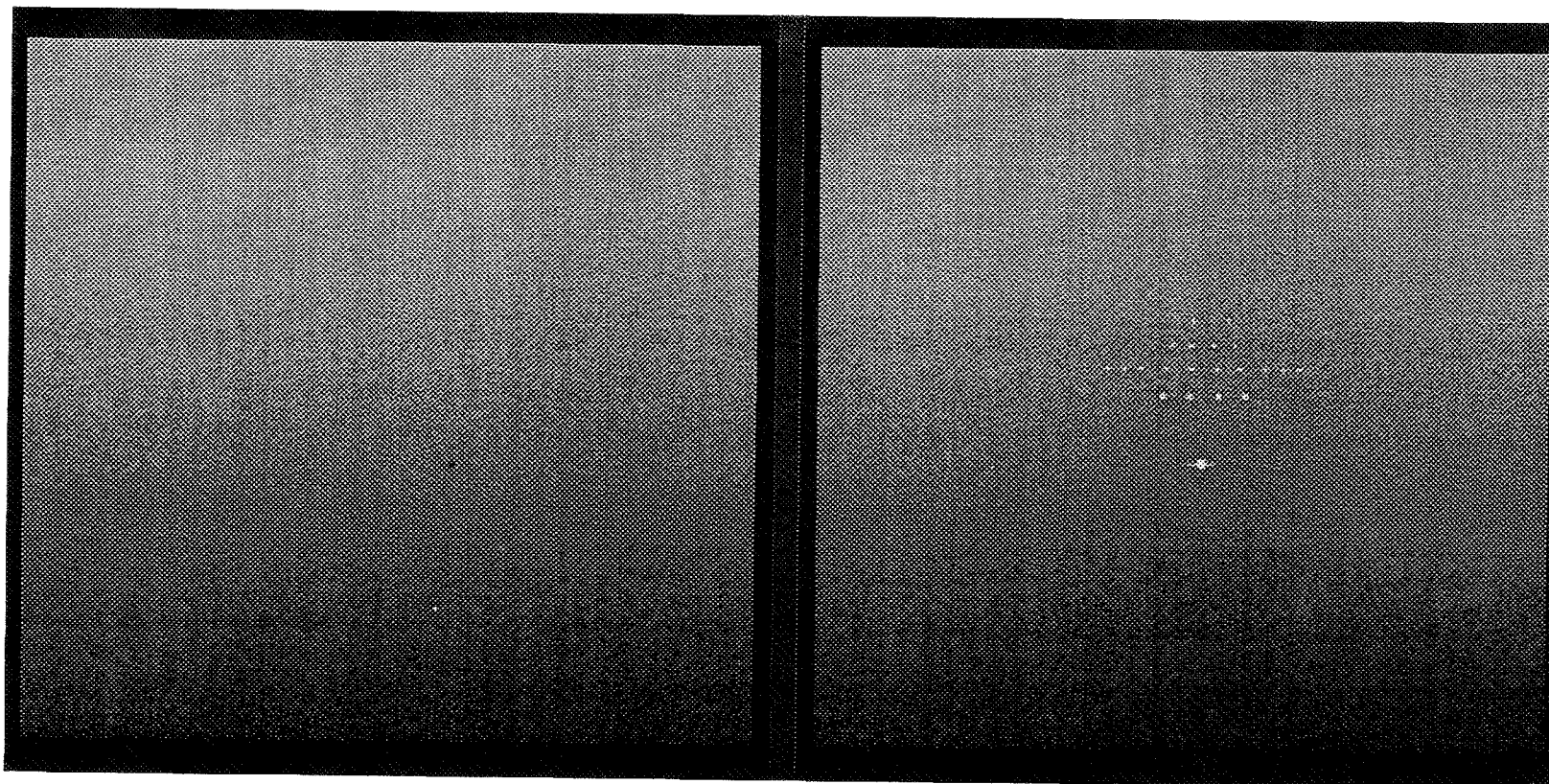
A collection of test images were synthesized for testing of the various forms of image metrics.

Set 1:

Left image: a test image consisting of pure fog.

Right image: an image of landing lights in fog.

Image Metric Test Images



Additional synthesized test images.

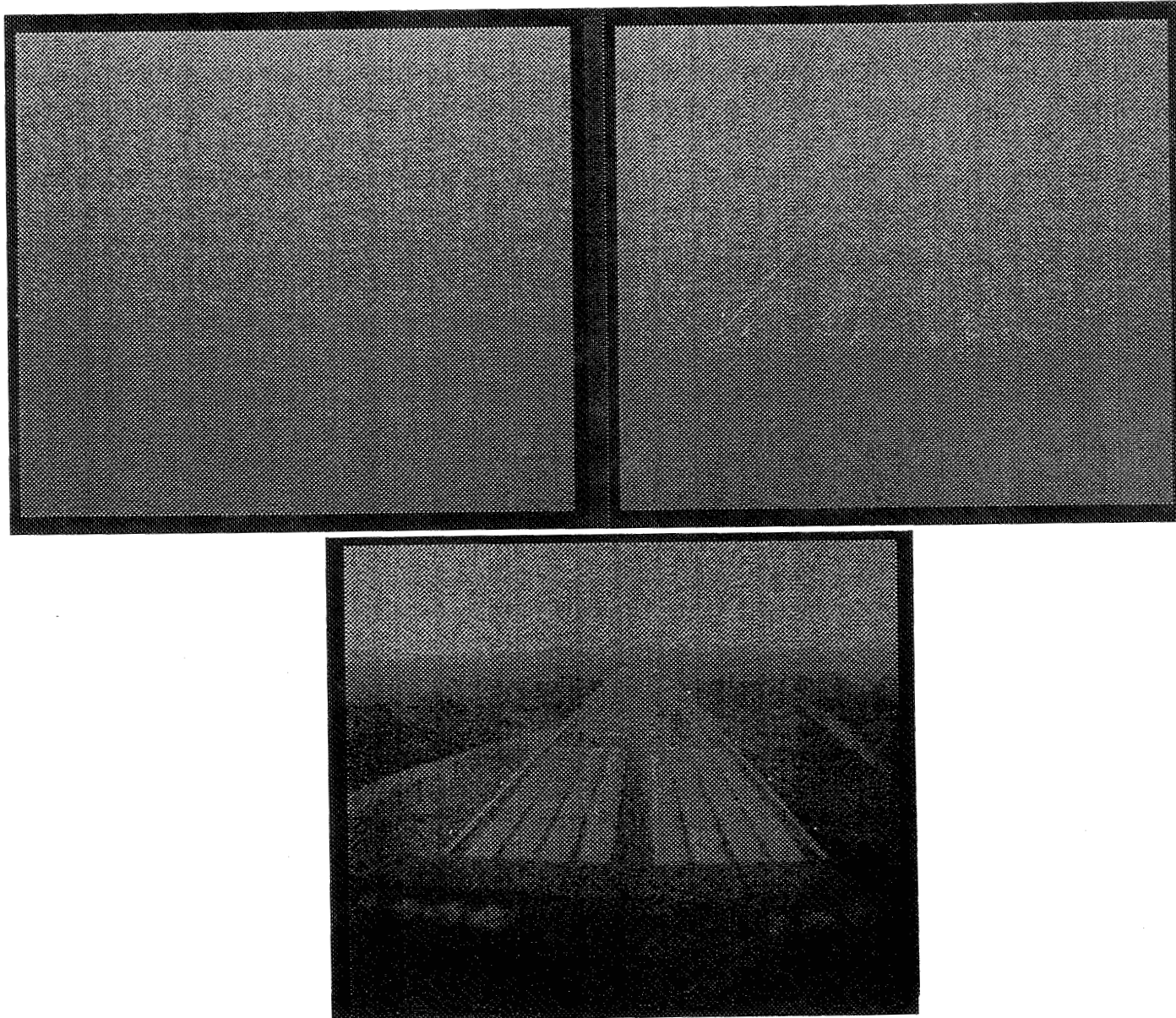
Set 2:

Top Left image: a test image consisting of pure fog.

Top Right image: an image with runway lights and markings showing through heavy fog.

Bottom image: an image of the same runway as above but with light fog.

Image Metric Test Images



Results of the edge energy metric as applied to the first set of test images.

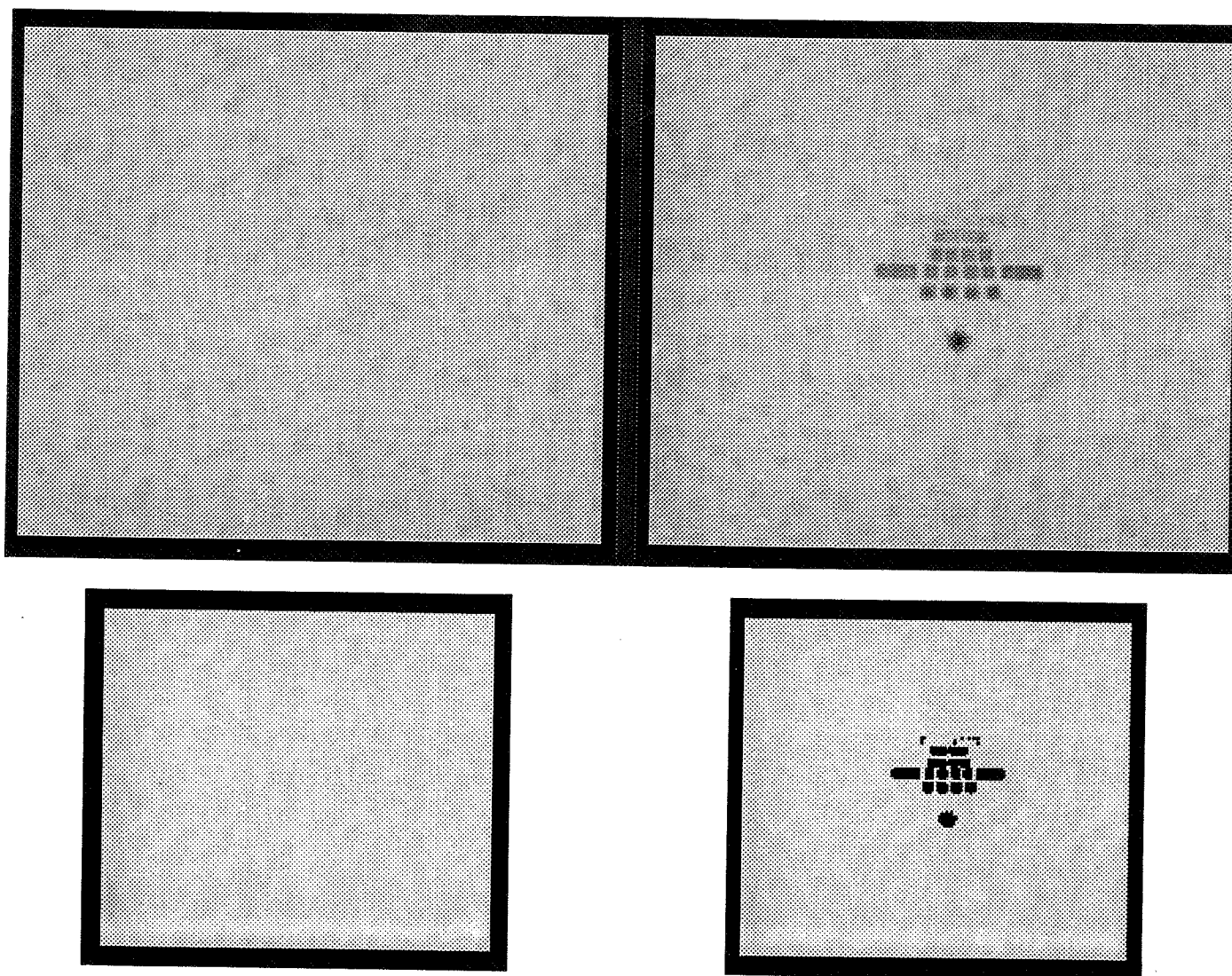
Top Left image: metric values of the pure fog image.

Top Right image: metric values for an image of landing lights in fog.

Bottom Left image: thresholded metric values for the top left image. (Note: the image is zero valued)

Bottom Right image: thresholded metric values for the top right image.

Edge Energy Image Metric Results



Edge metric applied to the second set of test images.

Top Left image: metric values of the pure fog image.

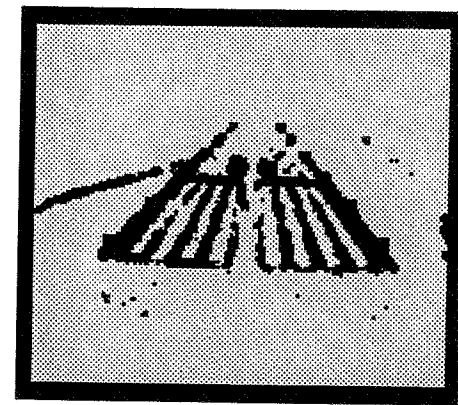
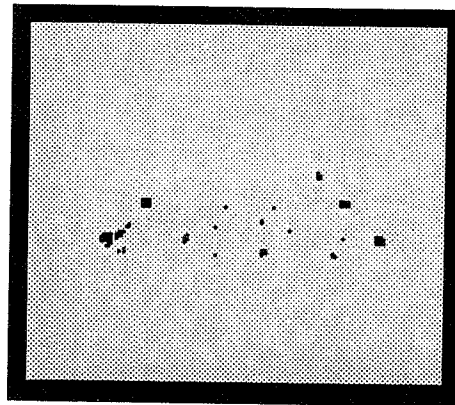
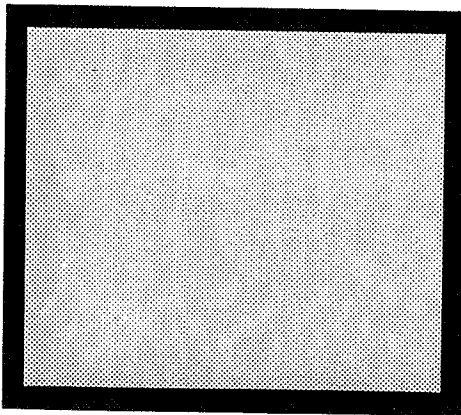
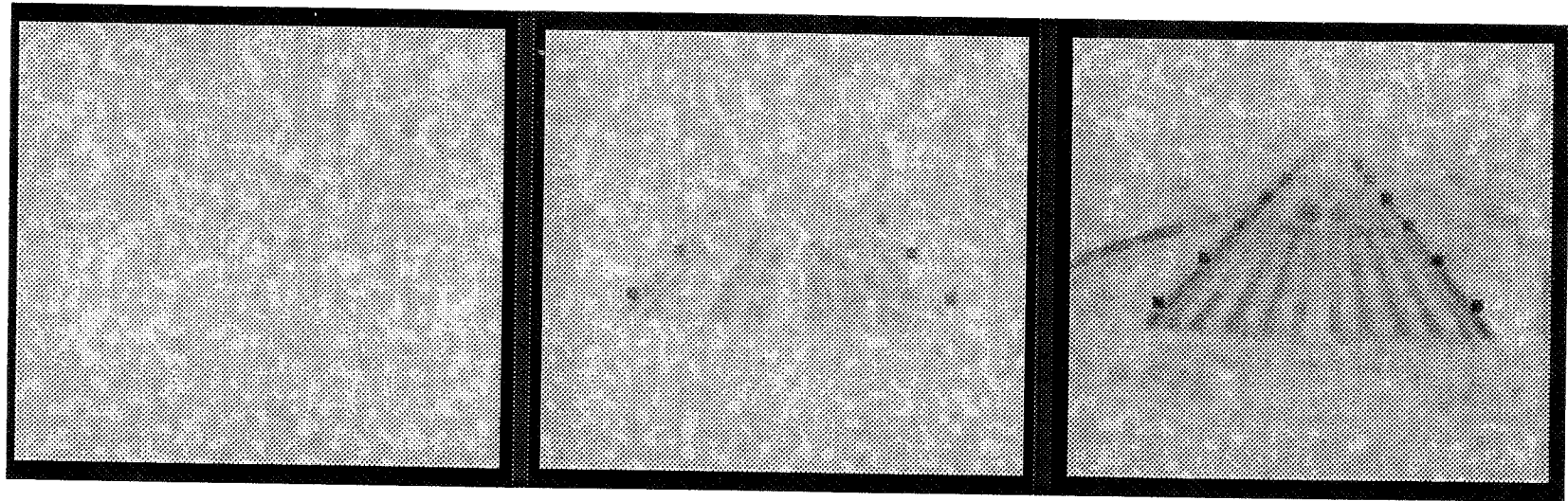
Top Middle image: metric values of an image with runway lights and markings showing through heavy fog.

Top Right image: metric values of an image of the same runway as above but with light fog.

Bottom images: thresholded versions of the Top row images.

(Note: the bottom left image is zero valued)

Edge Energy Image Metric Results



Contrast-based metric applied to the test images.

Top Left image: metric values for an image with runway lights and markings showing through heavy fog.

Top Right image: metric values for an image of the same runway as above but with light fog.

Bottom Left image: metric values for an image of pure fog.
(Note: this image is zero valued)

Bottom Right image: metric values for an image of landing lights in fog.

Contrast-Based Image Metric Results

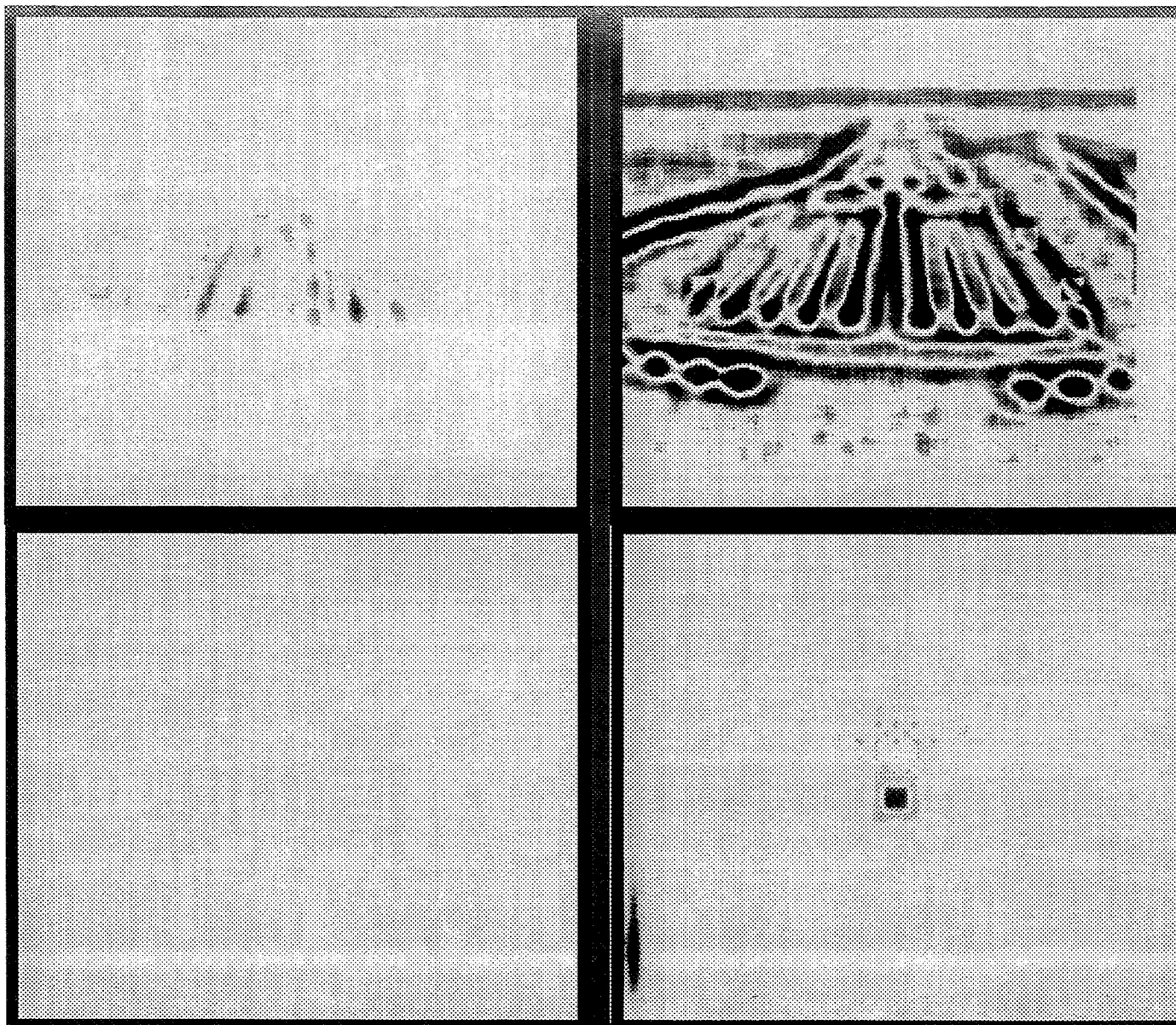


Image features are extracted for runways and taxiways for subsequent display on a HUD, etc.

Such features lead to improved situational awareness and can potentially lead to automatic performance of key functions:

- ✈ Runway/Taxiway Detection
- ✈ Runway Augmentation
- ✈ Runway Incursion/Obstacle Detection

Sensor Feature Extraction

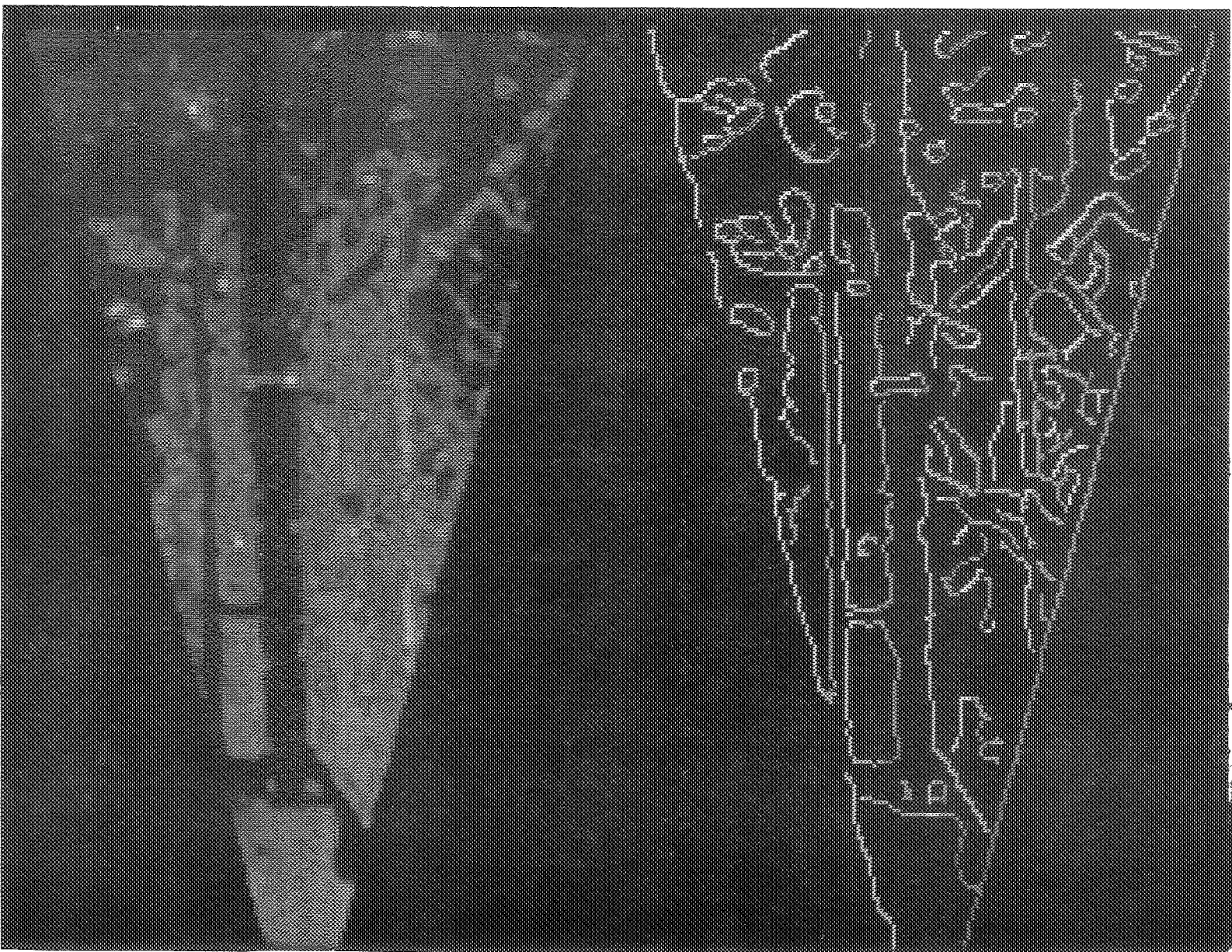
- ✈ **Runway/Taxiway Detection**
- ✈ **Runway Augmentation**
- ✈ **Runway Incursion/Obstacle Detection**

Left image: a subimage of the previous radar image of Pt. Magu NAS as produced by a 35 GHz radar.

Right image: edges extracted using a multi-threshold, edge linking algorithm.

Subsequent processing of the edge image will lead to runway boundaries being extracted and displayed on a cockpit HUD.

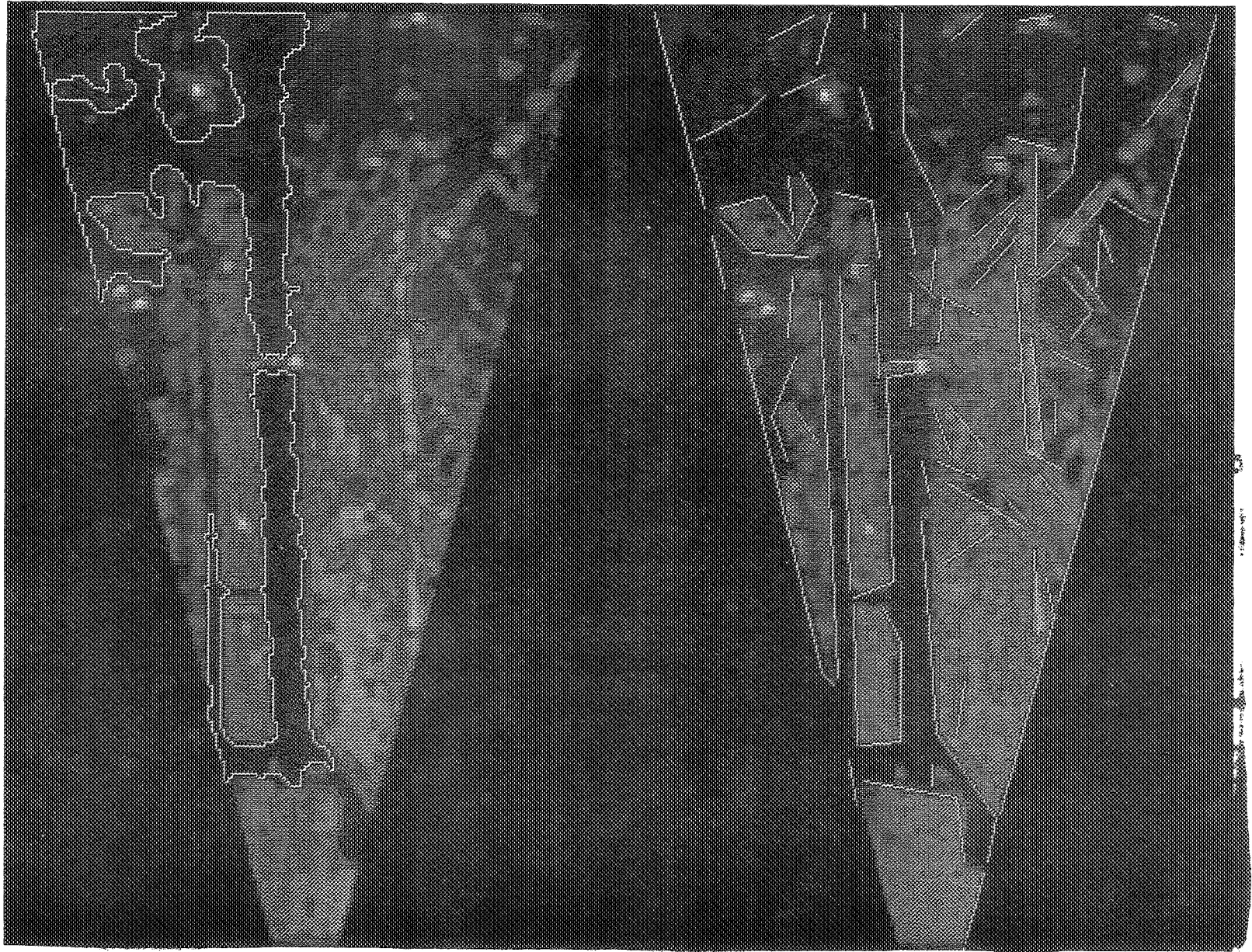
Runway/Taxiway Detection Results



Left image: a segmentation of the radar image as produced by a region growing algorithm. The white lines outline a nice definition of the runway and runway-like regions

Right image: the white lines represent line segments that have been fit to the edge contours shown on the previous viewgraph.

Runway/Taxiway Detection Results



Note that all of the processed imagery shown, constitute our initial explorations in these various areas of image processing. Much research remains to be done in these areas.

Conclusions

- ✈ **Image processing will provide important contributions to flight crew enhanced situational awareness.**
- ✈ **Ongoing efforts concentrate on techniques that deliver maximum performance and allow cost effective, real-time, implementation.**

omit

V. IMAGE PROCESSING: HUMAN VISION

DCT Quantization Matrices Visually Optimized for Individual Images

Andrew B. Watson
NASA Ames Research Center

ABSTRACT

This presentation describes how a vision model incorporating contrast sensitivity, contrast masking, and light adaptation is used to design visually optimal quantization matrices for Discrete Cosine Transform image compression. The Discrete Cosine Transform (DCT) underlies several image compression standards (JPEG, MPEG, H.261). The DCT is applied to 8x8 pixel blocks, and the resulting coefficients are quantized by division and rounding. The 8x8 "quantization matrix" of divisors determines the visual quality of the reconstructed image; the design of this matrix is left to the user.

Since each DCT coefficient corresponds to a particular spatial frequency in a particular image region, each quantization error consists of a local increment or decrement in a particular frequency. After adjustments for contrast sensitivity, local light adaptation, and local contrast masking, this coefficient error can be converted to a just-noticeable-difference (jnd). The jnds for different frequencies and image blocks can be pooled to yield a global perceptual error metric. With this metric, we can compute for each image the quantization matrix that minimizes bitrate for a given perceptual error, or perceptual error for a given bitrate.

Implementation of this system demonstrates its advantages over existing techniques. A unique feature of this scheme is that the quantization matrix is optimized for each individual image. This is compatible with the JPEG standard, which requires transmission of the quantization matrix.



DCT Quantization

- Uniform quantization by division and rounding

$$u_{ijk} = \text{Round}\left[c_{ijk}/q_{ij}\right]$$

- Quantization Matrix: 8 x 8 matrix of divisors q_{ij}
- Quantization error

$$e_{ijk} = c_{ijk} - q_{ij} u_{ijk}$$



Perceptual Approach to QM Design

- Maximum error is $q_{ij}/2$
- Measure detection thresholds t_{ij}
- Develop comprehensive formula

$$t_{ij} = ap[i, j, L, px, py]$$

- Set maximum error to detection threshold
- Equivalently, set QM to twice threshold

$$q_{ij} = 2 t_{ij}$$



Shortcomings

- Image dependence
- Luminance Masking
- Contrast Masking
- Error Pooling
- Quality Metric
- Image-Dependent Perceptual Approach



Luminance Masking

- Thresholds increase with block luminance
- Define Luminance masked thresholds t_{ijk}
- Use comprehensive formula

$$t_{ijk} = \text{ap}[c_{00k}, i, j]$$

- Or use power-law approximation

$$t_{ijk} = t_{ij}(c_{00k}/\bar{c}_{00})^{a_T}$$



Contrast Masking

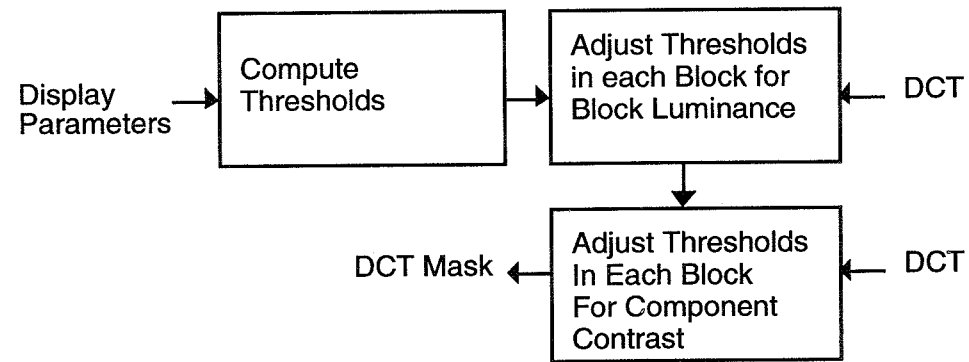
- Thresholds increase as contrast increases
- Masking greatest within block and coeff
- Define Contrast-Masked threshold

$$m_{ijk} = \text{Max} \left[t_{ijk}, |c_{ijk}|^{w_{ij}} t_{ijk}^{1-w_{ij}} \right]$$

- w_{ij} (0 - 1) defines strength of masking
- w_{ij} may differ for different frequencies i,j



Computing the DCT Mask





Perceptual Error

- Define elementary perceptual error as jnd
- Quantization error divided by masked threshold

$$d_{ijk} = e_{ijk} / m_{ijk}$$



Spatial Error Pooling

- Minkowski metric to pool between blocks

$$p_{ij} = \left(\sum_k |d_{ijk}|^{\beta_s} \right)^{1/\beta_s}$$

- Result is “Perceptual Error Matrix”
- Describes the jnds pooled over all blocks at each frequency
- β_s defines nature of spatial pooling



Frequency Error Pooling

- Pool over frequencies to get total perceptual error P
- Minkowski metric

$$P = \left(\sum_{ij} p_{ij}^{\beta_f} \right)^{1/\beta_f}$$

- When $\beta_f = \infty$, is max-of pooling

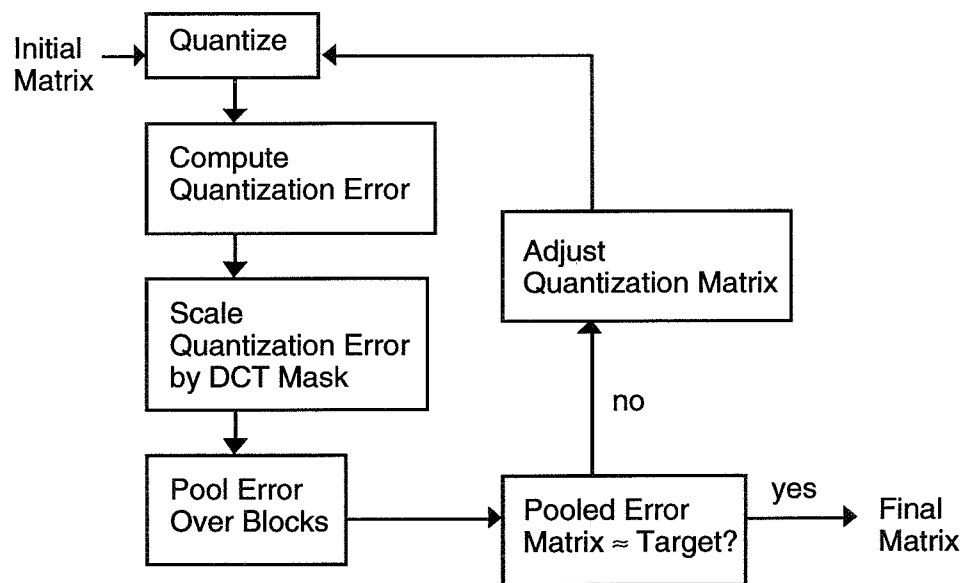


Optimizing the Quantization Matrix

- Goal: Minimize total perceptual error for given bitrate
- When $\beta_f = \infty$, optimum is when
$$p_{ij} = \psi \quad \forall i, j$$
- Intermediate goal: find q_{ij} for which
$$p_{ij} = \psi \quad \forall i, j$$
- Note that $p_{ij} = f_{ij}(q_{ij})$



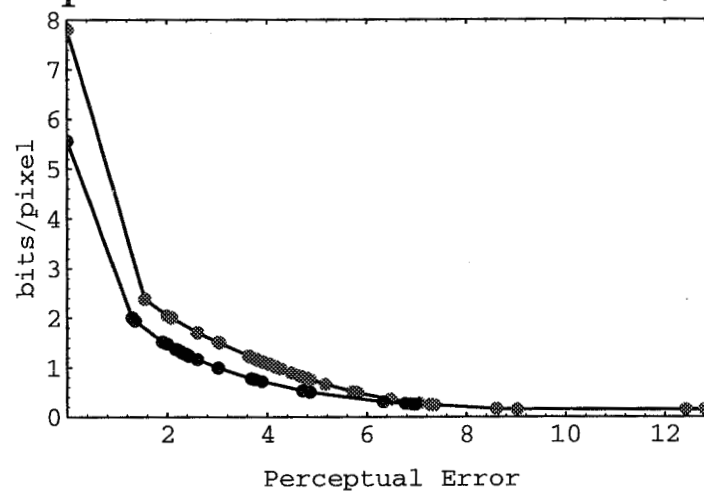
Inner Optimization Block Diagram





Optimizing for Given Bit Rate

- Samples from function $\text{bitrate}(\psi)$



-
- Iteratively estimate ψ yielding desired bitrate



Summary

- Perceptual error metric based on DCT
- Incorporates luminance masking, contrast masking, and error pooling
- Offers plausible “quality factor”
- Allows simple optimization of QM
- Compatible with JPEG standard
- Can incorporate color & alternate visual models
- Consider the alternatives



Summary (cont.)

- Use in adaptive DCT schemes
 - MPEG
 - thresholding
- Use in wavelet schemes
- "Free"

Extracting Heading and Temporal Range from Optic Flow: Human Performance Issues

Mary K. Kaiser
John A. Perrone
Leland Stone
NASA Ames Research Center

Martin S. Banks
James A. Crowell
University of California, Berkeley

520-54
205750
p. 18
249940

ABSTRACT

Pilots are able to extract information about their vehicle motion and environmental structure from dynamic transformations in the out-the-window scene. In this presentation, we focus on the information in the optic flow which specifies vehicle heading and distance to objects in the environment (scaled to a temporal metric). In particular, we are concerned with modeling how the human operators extract the necessary information, and what factors impact their ability to utilize the critical information. In general, the psychophysical data suggest that the human visual system is fairly robust to degradations in the visual display (e.g., reduced contrast and resolution, restricted field of view). However, extraneous motion flow (i.e., introduced by sensor rotation) greatly compromises human performance. The implications of these models and data for enhanced/synthetic vision systems are discussed.

INTRODUCTION

The out-the-cockpit scene provides a variety of visual cues to aid the pilot with vehicular control. As Walter Johnson discussed in his talk, some of these can be considered as static (e.g., horizon ratios), whereas others are dynamic or time-varying (e.g., change in the splay angle of the runway). Our research examines the control relevant information carried in the optic flow. Optic flow is the visual streaming of visible points, edges, and objects that results when one moves through a stationary, structured environment. During transport flight, relevant optic flow occurs primarily below the horizon line -- it is defined by textures and objects on the ground plane.

Optic flow is represented as a field of vectors, with the length of each vector representing the speed at which an element moves relative to the vantage point of the sensor (e.g., the human eye). For linear motion with a fixed-orientation sensor, the focus of expansion of the vector field defines the heading. If the sensor rotates as it translates (e.g., if it fixates on a point in the environment), this adds a common motion component to all the vectors which needs to be factored out before heading can be recovered. Once heading is extracted, the angle objects form relative to the heading (and the rate of change of this

angle) define their temporal range. Thus, heading extraction is a critical component to range extraction as well. In this presentation, we describe a model of heading extraction by human observers which is both physiologically plausible and consistent with psychophysical data. We then discuss the psychophysical findings from our laboratories concerning what factors do and do not degrade heading and temporal range extraction.

HEADING EXTRACTION

Many algorithms have been proposed for solving the self-motion estimation problem (for reviews, Warren, Morris, & Kalish, 1988; Warren & Hannon, 1990). Some of these use the image motion from a small number of points to solve a set of nonlinear equations (e.g. Longuet-Higgins & Prazdny, 1980; Ballard & Kimball, 1983). Such techniques tend to be sensitive to noise in the image motion measurements and must rely on iterative methods to arrive at a solution. Others make use of differential invariants of the flow field and are based on spatial derivatives (e.g. Koenderink & van Doorn, 1975). In addition to being sensitive to noise, these methods require locally continuous flow fields and a smoothness constraint for environmental surfaces. One of the more popular approaches to the self-motion problem makes use of the fact that image motion resulting from rotation is independent of the depth of points in the scene, while that resulting from translation is not (Longuet-Higgins & Prazdny, 1980). Therefore, the difference between flow-field vectors at adjacent points at different depths yields information related to the translation only. Rieger and Lawton (1985) developed a model which uses this principle, but which is able to use flow-field vectors from nearby points on the image plane rather than points that were exactly adjacent or overlapping. This "local differential motion model" is currently the most popular candidate for the algorithm underlying human self-motion perception (see Warren & Hannon, 1990; Hildreth, 1992). However, psychophysical studies at Ames Research Center by Perrone and Stone (Perrone & Stone, 1991; Stone & Perrone, 1991, 1993) have shown that heading can still be estimated correctly in situations that lack the local differential image motion necessary for the Reiger-Lawton model to work properly.

To explain their psychophysical findings, Perrone and Stone (Perrone, 1992; Perrone & Stone, 1992a, 1992b) have recently proposed an altogether different "physiologically-based" approach to solving the self-motion problem (Figure 1). The rationale for using a physiologically-based system is two-fold. First, it is more likely to allow extrapolation to a wider range of human performance and secondly, such "reverse engineering" will hopefully eventually lead to the design of artificial vision systems that are as robust and as fast as the human brain. One of the model's strengths is that it is based on known physiological properties of motion sensitive neurons in the Middle Temporal (MT) area of the primate visual cortex known to be involved in motion processing (Zeki, 1980; Maunsell & Van Essen, 1983; Albright, 1984; Newsome, Wurtz, Dursteler & Mikami, 1985; Newsome, Britten, & J. A. Movshon, 1989; Salzman, Britten, & Newsome, 1990) and proposes a theoretical framework for how neurons in the Medial Superior Temporal (MST) area might use the output from MT cells to extract heading. In the model, MT-like units carry out the local analysis of the 2-D image motion using direction and speed tuned "sensors" (Figure 2). The outputs from specific sets of MT-sensors are then summed to produce the output for a

specialized MST-like "detector" which is "tuned" to a particular pattern of self-motion produced image motion and responds much like actual MST neurons (Saito, Yukie, Tanaka, Hikosaka, Fukada, & Iwai, 1986; Tanaka, Hikosaka, Saito, Yukie, Fukada, & Iwai, 1986; Duffy & Wurtz, 1991). These MST-like detectors sum MT-like sensor outputs over a large portion of the visual field and act as templates searching for specific patterns of global retinal image motion (Figure 3). The most active detector, within a map of possible combined translation-rotations, identifies what self-motion is most consistent with the image flow and, hence, solves the self-motion problem.

Comparison of human psychophysical data with simulations of the Perrone-Stone model (Figure 4) demonstrates that the model is consistent with known properties of visual heading perception and, in particular, that the model can provide a quantitative estimate of the break down of human performance at higher rotation rates seen by both Perrone and Stone (Perrone & Stone, 1991; Stone & Perrone, 1991) and Banks and colleagues (Royden et al., 1992). This approach is therefore very promising, although further psychophysical validation and refinement will be necessary before it can be used as an engineering design tool. In particular, the model does not attempt to include non-visual signals that are likely to contribute to human perception (Royden et al., 1992). However, the output-map structure of the Perrone-Stone model lends itself well to the incorporation of such additional non-visual information.

The Perrone-Stone model predicts, and psychophysical evidence demonstrates, that heading extraction is impaired when rotation (without non-visual information about rotation) is added to the visual display. Banks and his colleagues have also examined whether two aspects of display quality, resolution and contrast, affects people's ability to determine their heading from optic flow. Displays were presented both foveally and peripherally (40° nasal). Three levels of crab-angle (i.e., heading relative to the center of the display) were used: 0°, 20°, and 70°. In a reduced contrast study, Weber contrast was varied between 1 and 40 (0.85 is the contrast threshold for central vision, 3.10 is contrast threshold for 40° nasal). As shown in Figure 5, heading threshold varied as a function of crab angle; headings were harder to discriminate during higher crab angles. But heading extraction was fairly robust to contrast level, at least for supra-threshold contrast levels. For centrally viewed displays, performance did not improve with the Weber contrast levels increasing beyond five. In a visual acuity (resolution) study (Figure 6), there was a similar effect for crab angle, and some effect for resolution. Still, performance with the 0° crab angle, centrally viewed display was fairly accurate (threshold < 2°) even with 20/100 resolution.

TEMPORAL RANGE ESTIMATE

Given that people can extract heading from the optic flow, it is possible, in principle, to then determine the temporal range to any object in the environment (Kaiser & Mowafy, in press). For objects lying on the flight vector (Figure 7), the time to contact (TTC) is specified by the angular extent of the object, θ , divided by the rate of change of the angle, $\delta\theta/\delta t$. That is:

$$TTC = \theta / \delta\theta/\delta t \quad (1)$$

For objects lying off the heading vector, an analogous derivation is possible, using the angle between the object and the tract vector, ϕ , and its rate of change, $\delta\phi/\delta t$. The ratio of these terms specifies time to passage (TTP), which is the time until the object intersects the eye-plane perpendicular to the heading vector (Figure 8):

$$TTP = \phi / \delta\phi/\delta t \quad (2)$$

Most empirical work on people's sensitivity to this optical information has focused on the TTC situation, and the use of these cues for coordinating motor activity such as hitting and catching approaching objects (see Tresilian, 1991 for a review). However, the TTP case is more germane for most flight control regimes; the pilot needs to estimate the time to various way-points for navigation, control, and execution of maneuvers (e.g., flare). Kaiser and her colleagues (Kaiser & Mowafy, in press) have recently examined people's sensitivity to TTP information. In the experimental paradigm, observers viewed a translation through a volume of point lights, and either judged which of two targets would pass their eye plane first (relative judgment task) or indicated when a target which had left the field of view would pass their eye plane (absolute judgment task). In both relative and absolute judgment tasks, people were able to perform reliably. Judgments of relative TTP were precise to around 600 msec and were comparable for narrow (19°) and wide (46°) fields of view (Figure 9). Absolute TTP judgments were reliable even in the absence of feedback (Figure 10), indicating that people's temporal estimates are "pre-calibrated."

One manner in which pilots might use this TTP information for flight control is illustrated in Figure 11. For any assigned altitude, the distance along a particular gaze angle is constant in eye-heights (i.e., the ground plane along the 45° gaze angle is one eye-height distant, the ground plane along the 26.5° gaze angle is two eye-heights, etc.). Pilots may seek to maintain a constant temporal distance (i.e., lead time) to objects along a given gaze angle. This will result in appropriate flight control for some regimes (e.g., rotorcraft landing, where speed is reduced proportional to distance-to-go), but will cause an inappropriate bias when speed should be held constant during altitude change. Also, pilots may misjudge their taxi speeds if they perform ground operations in a variety of vehicles with very discrepant eye-heights (Figure 12).

IMPLICATIONS FOR ENHANCED/SYNTHETIC VISION SYSTEMS

Optic flow provides a critical source of visual information for vehicular control. If proposed sensor displays for enhanced/synthetic vision systems do not adequately preserve optic flow information, pilot performance may be impaired. Also, the noise from some sensor systems can mask or distort flow patterns. Empirical findings and performance models suggest that such extraneous pseudo-motion signals might seriously compromise human optical flow processing. In such cases where natural motion cues are degraded or distorted, pilots may require other visual cue augmentations (e.g., flare cues) to compensate.

REFERENCES

- T. D. Albright, "Direction and orientation selectivity of neurons in visual area MT of the macaque," *J. Neurophysiol.* 52, 1106-1130 (1984).
- D. H. Ballard and O. A. Kimball, "Rigid body motion from depth and optical flow," *Comp. Vis. Graph. Image Process.* 22, 95-115 (1983).
- C. J. Duffy and R. H. Wurtz, "Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli," *J. Neurophys.* 65, 1329-1345 (1991).
- E. Hildreth, "Recovering heading for visually-guided navigation," *Vis Res.* 32, 1177-1192 (1992).
- M. K. Kaiser and L. Mowafy, "Optical specification of time-to-passage: Observers' sensitivity to global tau," *J. Exp. Psychol. Hum. Percept. Perform.* (in press).
- J. J. Koenderink and A. J. van Doorn, "Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer," *Optica Acta* 22, 773-791 (1975).
- H. C. Longuet-Higgins and K. Prazdny, "The interpretation of moving retinal images," *Proc. R. Soc. London Ser B* B208, 385-387 (1980).
- J. H. R. Maunsell and D. C. Van Essen, "Functional properties of neurons in the middle temporal visual area of the Macaque monkey. I. Selectivity for stimulus direction, speed, orientation," *J. Neurophysiol.* 49, 1127-1147 (1983).
- W. T. Newsome, K. H. Britten, & J. A. Movshon, "Neuronal correlates for a perceptual decision," *Nature*, 341: 52-54 (1989).
- W. T. Newsome, R. H. Wurtz, M. R. Dursteler, & A. Mikami, "Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey," *J. Neurosci.* 5: 825-840 (1985).
- J. A. Perrone, "Model for the computation of self-motion in biological systems," *J. Opt. Soc. of Am. A*, 9:177 (1992).
- J. A. Perrone and L. S. Stone, "Using the properties of MT neurons for self-motion estimation in the presence of eye movements," *Perception (Suppl)* 21: 64-65 (1992a).
- J. A. Perrone and L. S. Stone, "A possible role for the speed tuning properties of MT cells in the recovery of depth from motion," *Invest. Ophthalmol. & Vis. Sci. (Suppl)* 33:1141 (1992b).
- J. A. Perrone and L. S. Stone, "The perception of egomotion: Global versus local mechanisms," *Invest. Ophthalmol. & Vis. Sci. (Suppl)* 32: 957 (1991).

- J. H. Rieger and D. T. Lawton, "Processing differential image motion," *J. Opt. Soc. Am. A* 2, 354-360 (1985).
- C. S. Royden, I. V. Laudeman, J. A. Crowell, & M. S. Banks, "The influence of eye movements on heading judgments," *Invest. Ophthalmol. & Vis. Sci. (Suppl)* 33: 1051 (1992).
- H. Saito, M. Yukie, K. Tanaka, K. Hikosaka, Y. Fukada, and E. Iwai, "Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey," *J. Neurosci.* 6, 145-157 (1986).
- C. D. Salzman, K. H. Britten, & W. T. Newsome, "Cortical microstimulation influences perceptual judgments of motion direction," *Nature* 346, 174-177 (1990).
- L. S. Stone and J. A. Perrone, "Human heading perception during combined translational and rotational self-motion," *Soc. for Neurosci. Abstr.* 17: 847 (1991).
- L. S. Stone and J. A. Perrone, "Human heading perception cannot be explained using a local differential algorithm," *Invest. Ophthalmol. & Vis. Sci. (Suppl)* 34: 1229 (1993).
- K. Tanaka, K. Hikosaka, H. Saito, M. Yukie, Y. Fukada, and E. Iwai, "Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey," *J. Neurosci.* 6, 134-144 (1986).
- J. R. Tresilian, "Empirical and theoretical issues in the perception of time to contact," *J. Exp. Psychol. Hum. Percept. Perform.* 17, 865-876 (1991).
- W. H. Warren and D. J. Hannon, "Eye movements and optical flow," *J. Opt. Soc. Am. A* 7, 160-168 (1990).
- W. H. Warren, M. W. Morris, and M. Kalish, "Perception of translational heading from optical flow," *J. Exp. Psychol. Hum. Percept. Perform.* 14, 646-660 (1988).
- S. M. Zeki, "The response properties of cells in the middle temporal area (area MT) of owl monkey visual cortex," *Proc. R. Soc. London Ser B* 207, 239-248 (1980).

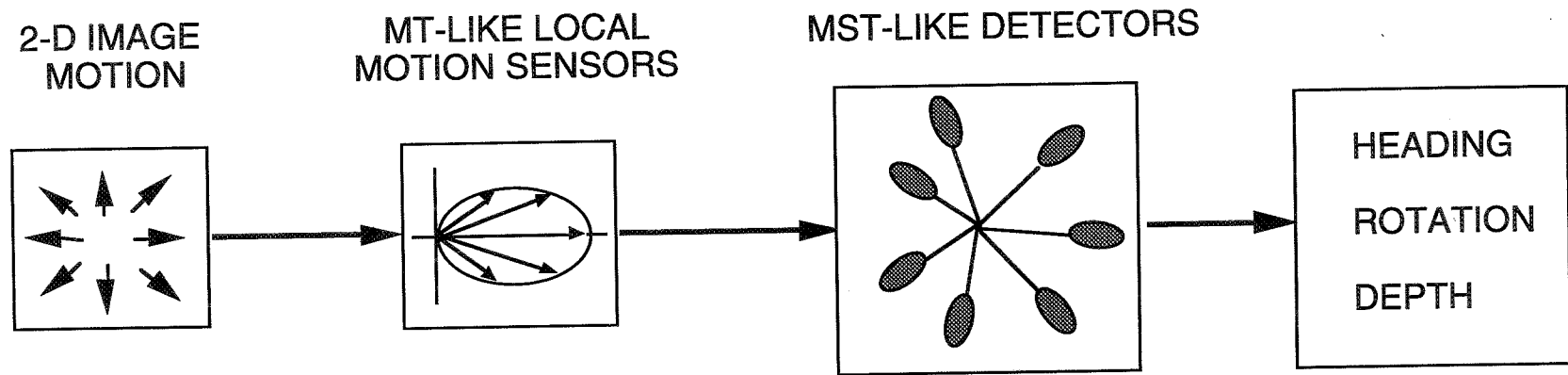
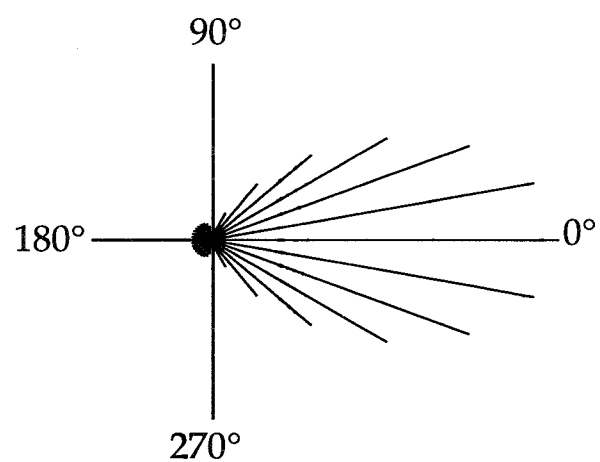


Figure 1. Overall structure of template model.

(a)



(b)

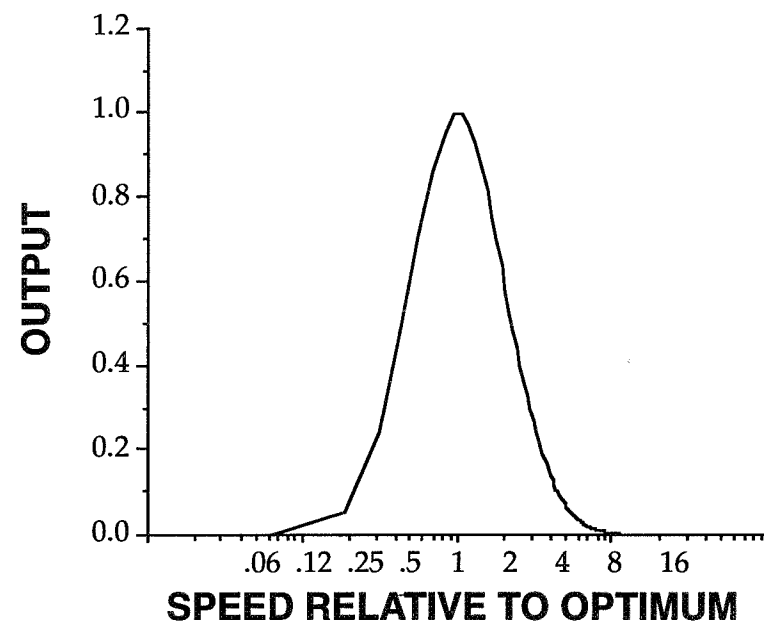


Figure 2. Idealized MT neuron responses. **a)** Direction tuning curve in polar plot form. **b)** Speed tuning curve.

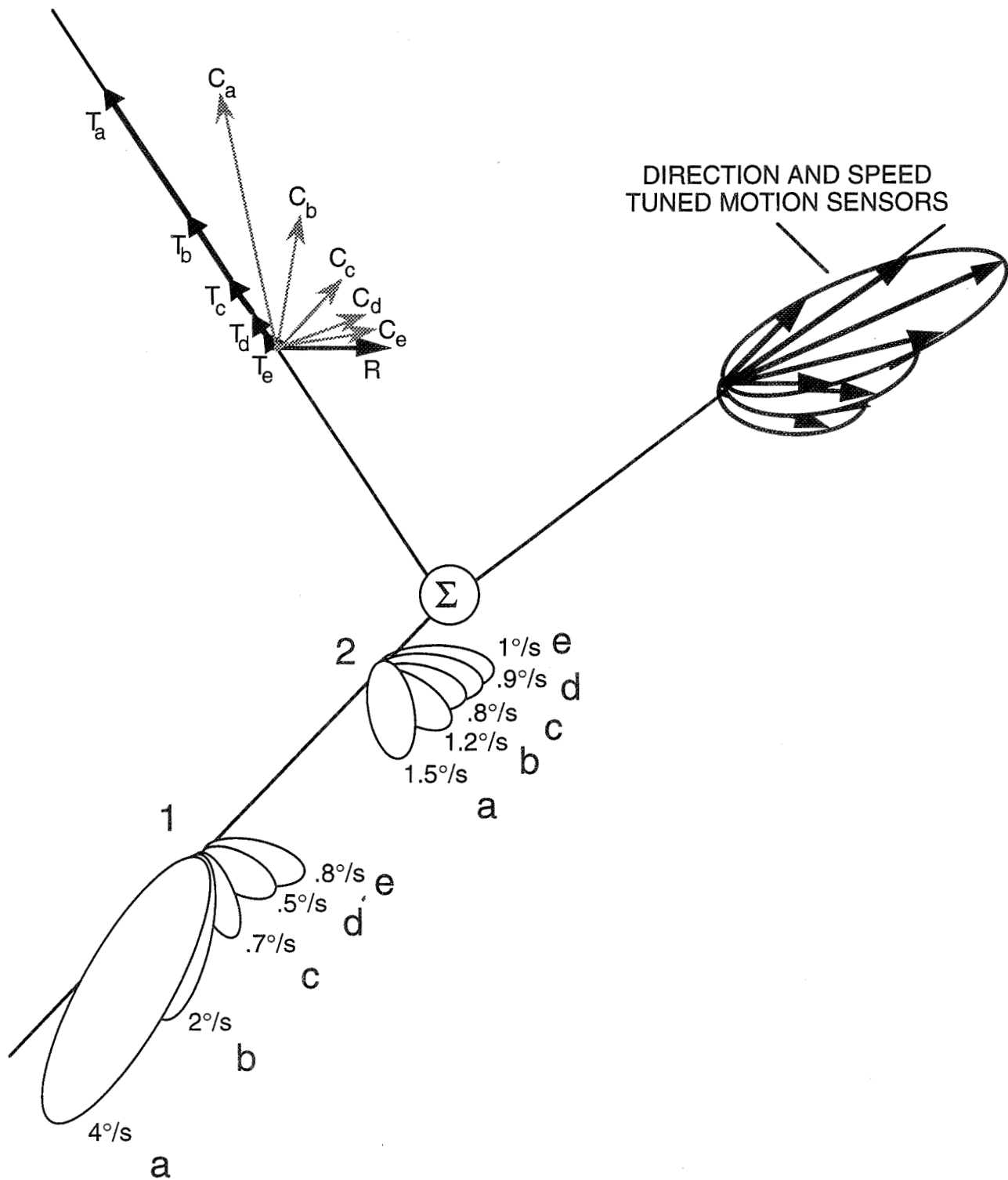


Figure 3. MST-like detector which acts as a template for a specific heading-rotation combination. The activity of groups of MT-like sensors at various locations in the visual field is summed, with the speed and direction-tuning of each sensor set to respond to the image motion, $C = T$ (translation) + R (rotation), associated with a specific depth plane (a through e).

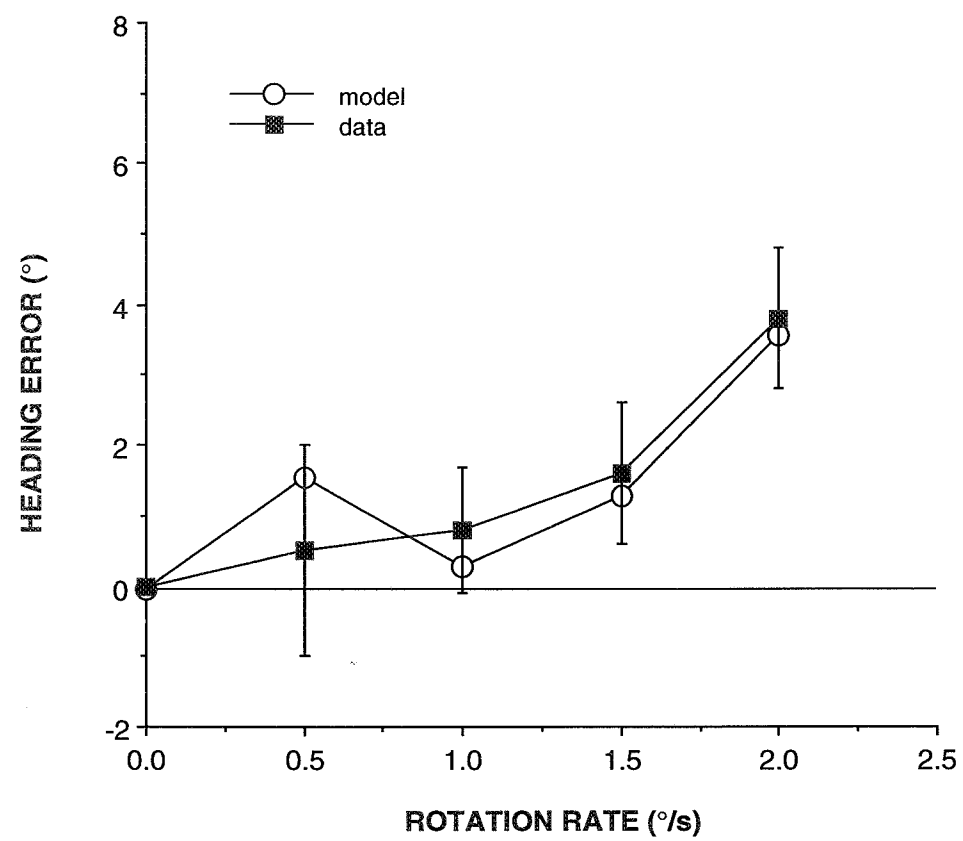


Figure 4. Comparison of heading error vs rotation rate for human observers and for the Perrone-Stone model.

Reduced Contrast Sensitivity

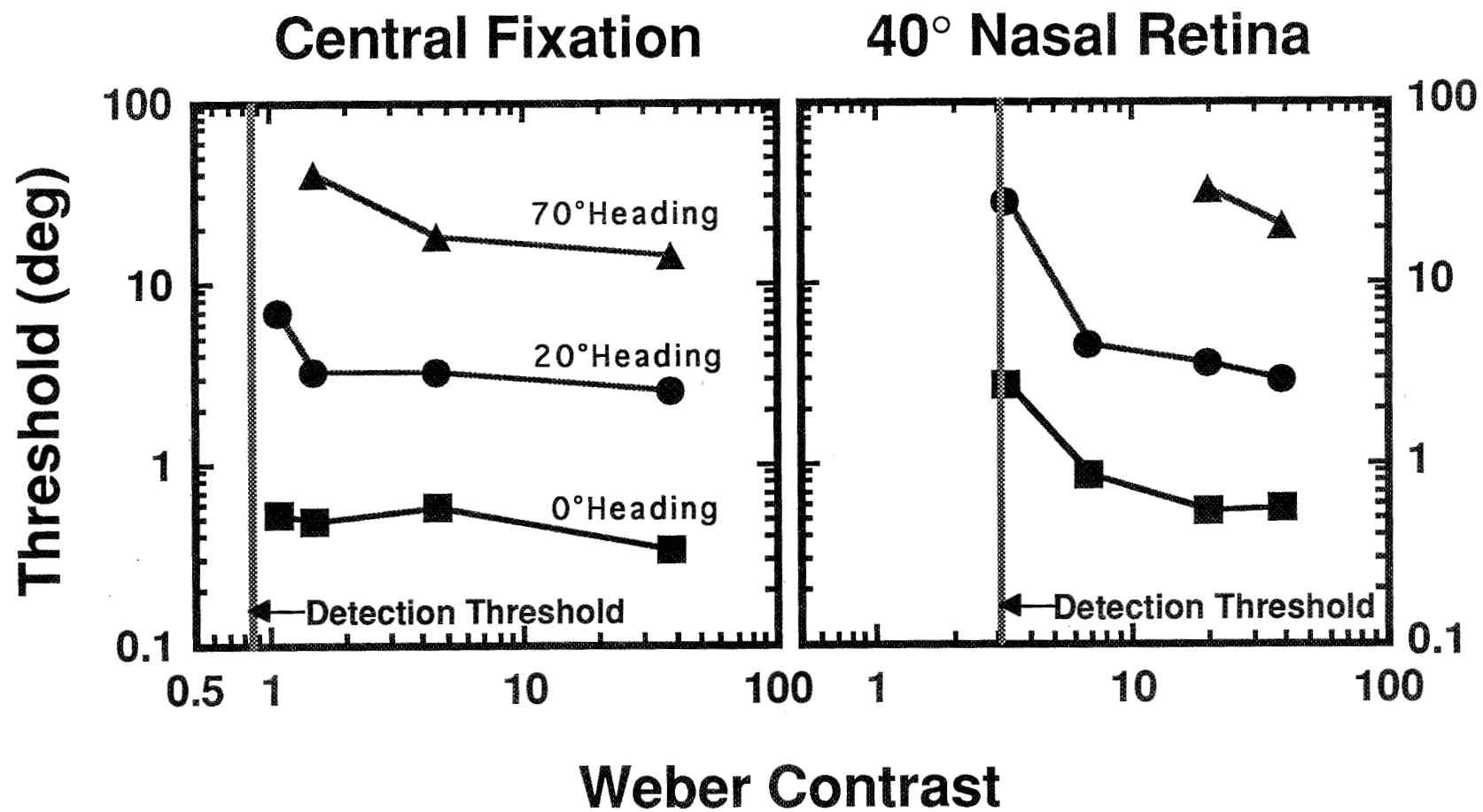


Figure 5. Heading threshold as a function of Weber contrast, eccentricity, and crab angle.

Reduced Visual Acuity

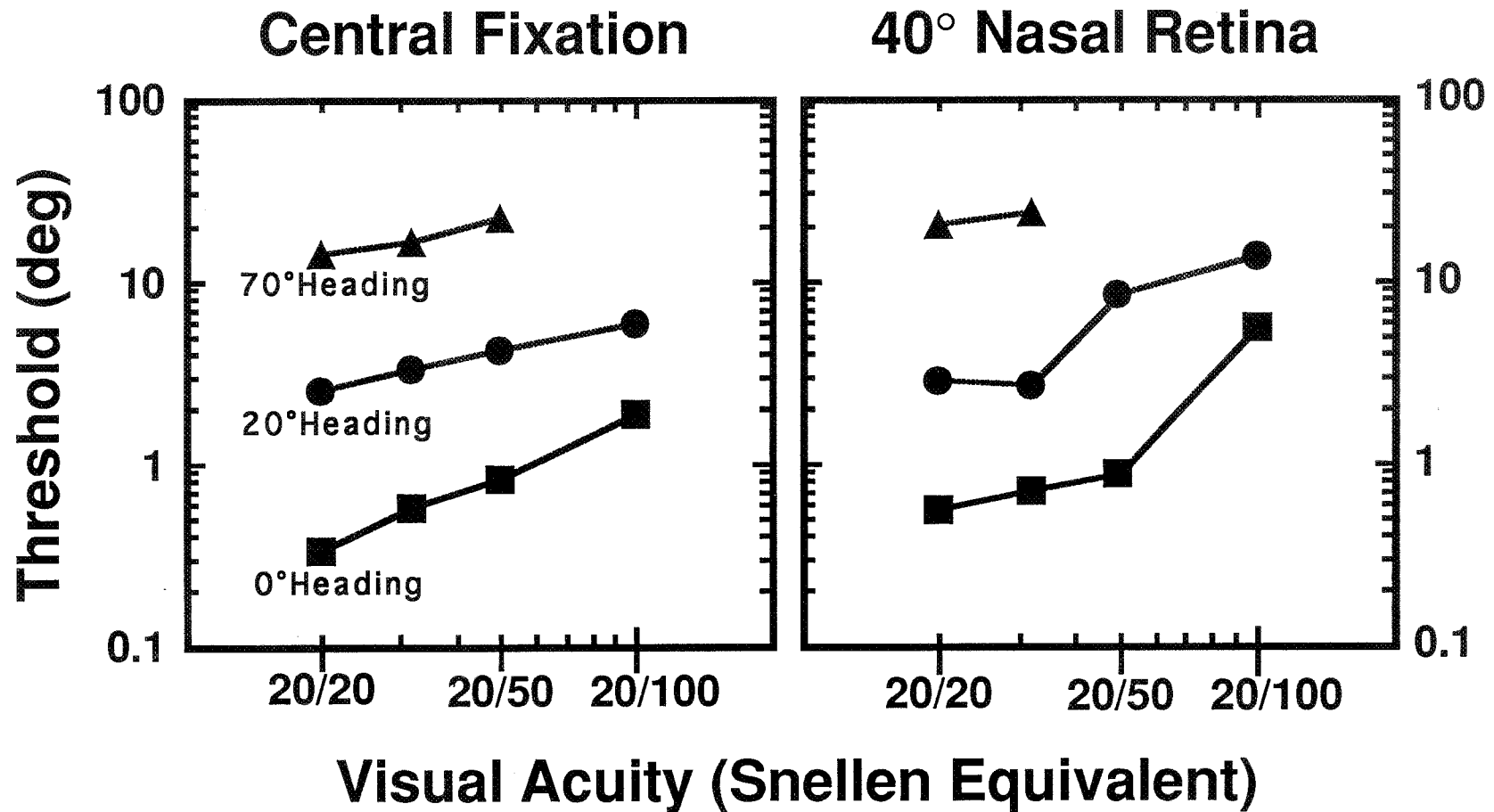


Figure 6. Heading threshold as a function of visual acuity, eccentricity, and crab angle.

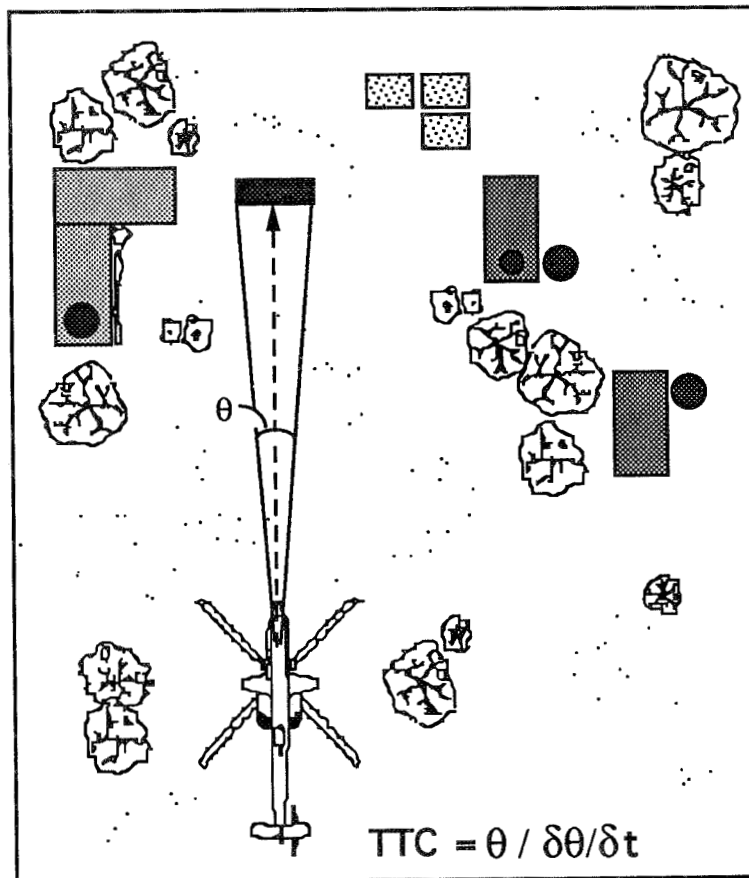


Figure 7. Geometry of the Time-to-Contact (TTC) situation. θ is the visual angle an object subtends.

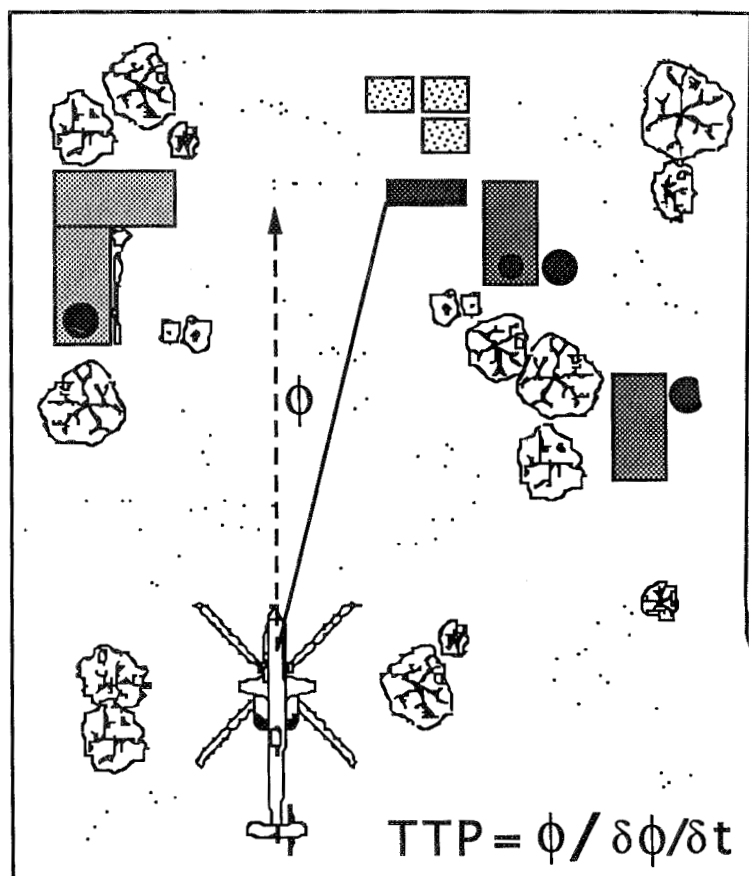


Figure 8. Geometry of the Time-to-Passage (TTP) situation. ϕ is the visual angle between an object and the heading vector.

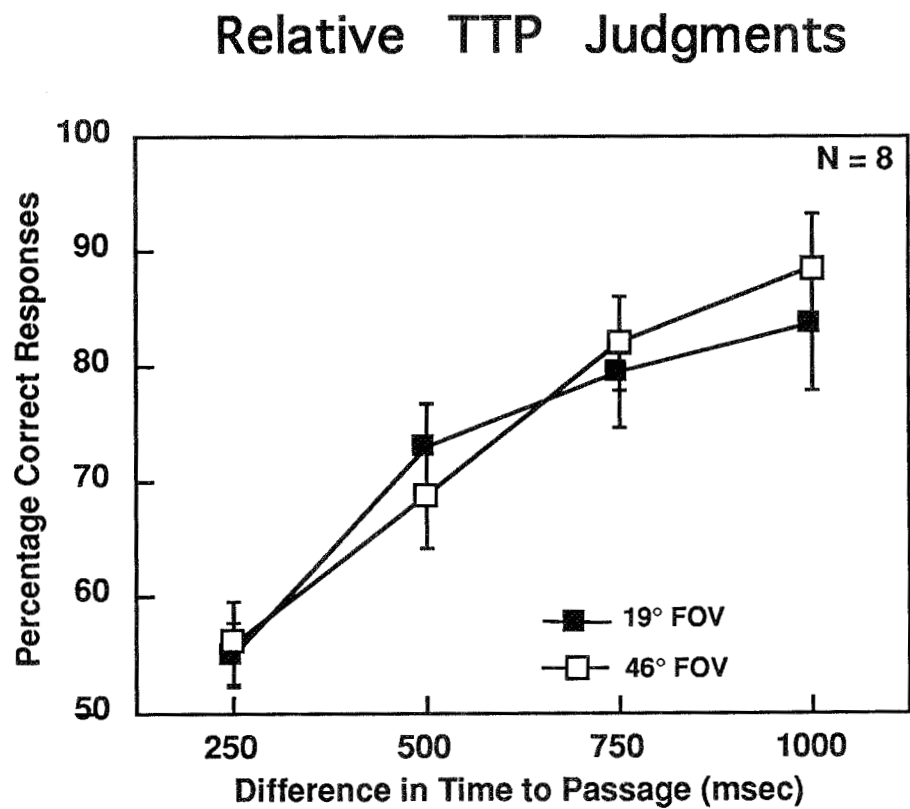


Figure 9. Relative Time-to-Passage (TTP) judgments for narrow (19°) and wide (46°) fields of view (FOV).

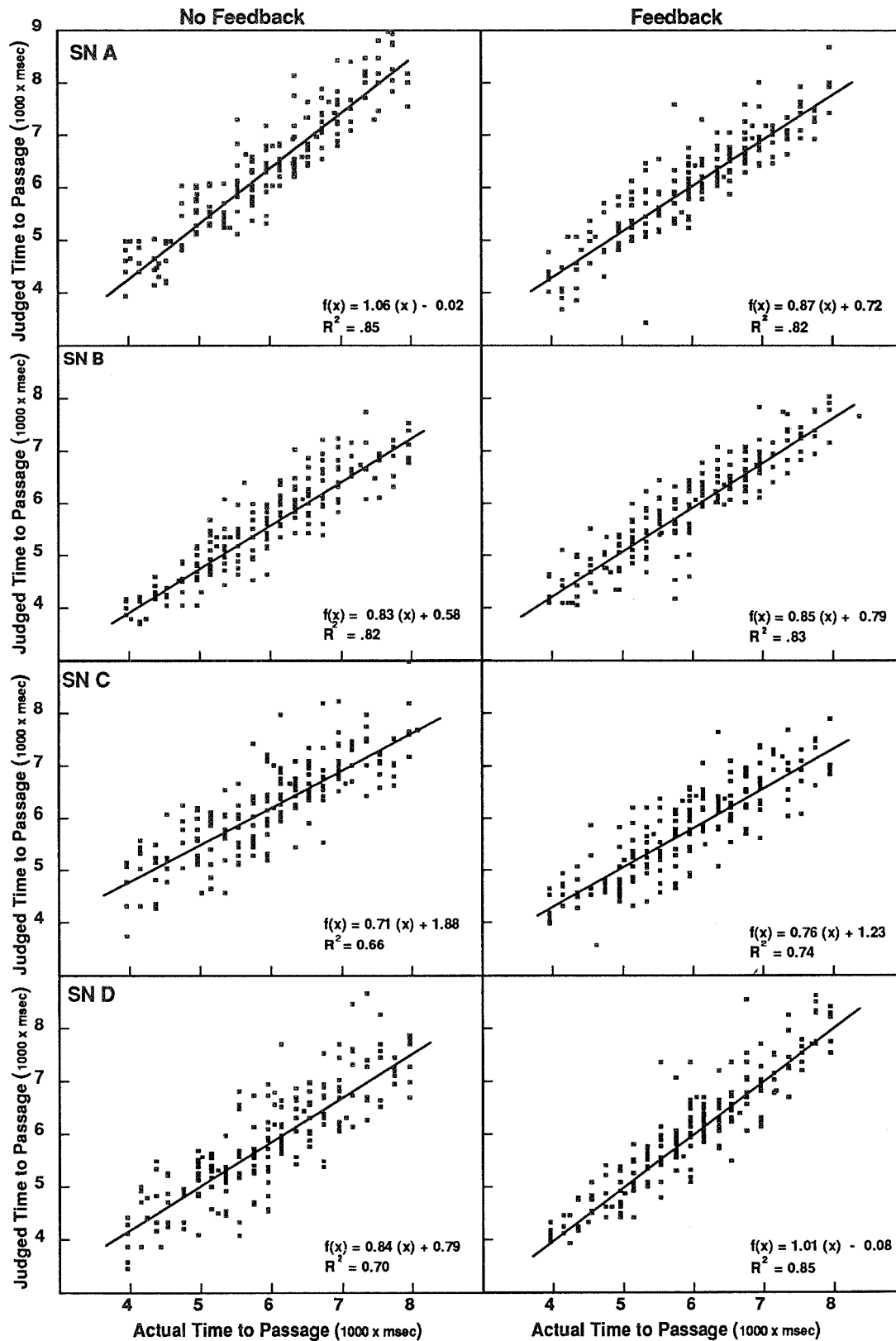


Figure 10. Relative Time-to-Passage (TTP) judgments in the presence and absence of feedback.

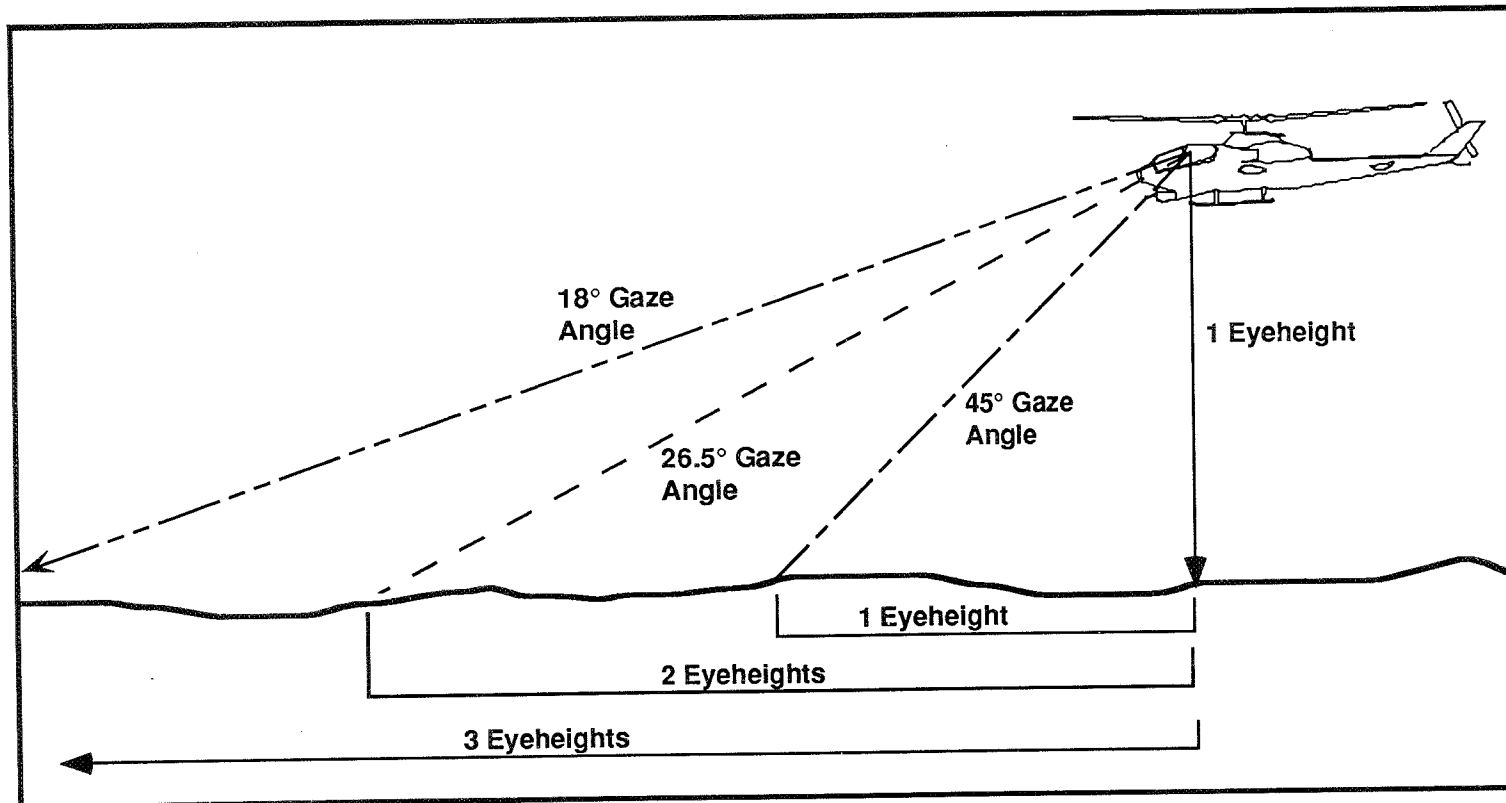
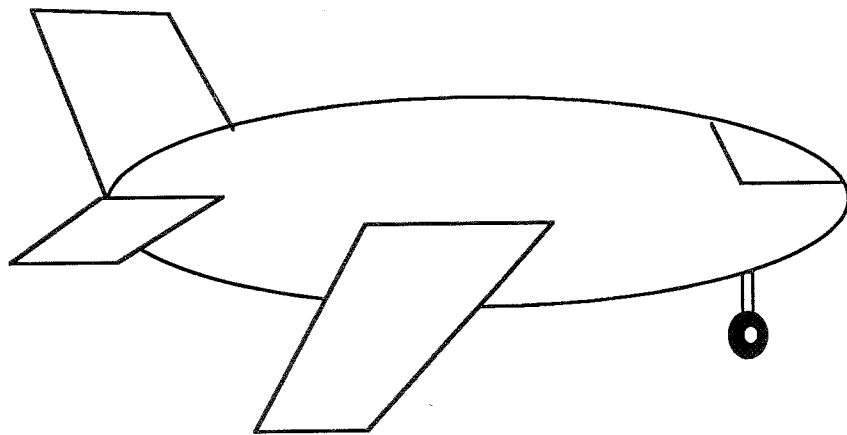


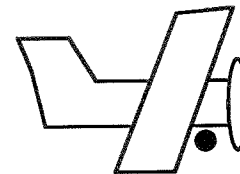
Figure 11. Eyeheight geometry. Distance along a given gaze angle is constant in eyeheight units.

Jumbo Jet
Eyeheight = 6 m



1 Eyeheight/sec
= 12 knots

Cessna 152
Eyeheight = 2 m



1 Eyeheight/sec
= 4 knots

Figure 12. Speed corresponding to 1 eyeheight/second for two sample eyeheights.

Optical Information in Landing Scenes

Walter W. Johnson
NASA Ames Research Center

ABSTRACT

During landing, the visual scene contains optical information about speed, altitude, glideslope, and track that is useful for the maintenance of spatial orientation and awareness. This information, embedded in the structure and transformations of the optical patterns, may be globally, regionally, or locally available. Global changes occur everywhere in the visual field during landing and include such information as flow rate acceleration due to changing speed and/or altitude. Regional changes occur within a more restricted area and include such information as horizon line motion due to aircraft pitching and rolling. Locally available changes are the most restricted and include such information as changes in runway form ratios due to changing glideslopes. Thus, within partially or fully synthetic displays, or within sensor-driven displays, preservation of flow rate and horizon motion information requires a minimum of knowledge about the details of the airport layout, while runway outlines do require much more knowledge of the layout. All may be important, however, and these, as well as other sources of optical information, can provide a pilot with his most natural framework for maintaining orientation.

521-54
205951
p. 10
249941

Optical Information Analysis

Properties of Optical Information

- *Optical Patterns* - Structure and transformations of the optical geometry
- *Optic Regions* - Where the relationship can be viewed (Elevation & Azimuth)
- *Information Content* - Flight path properties (e.g. speed, closure rate, sink rate) that covary with changes in optical patterns
- *Ecological Constraints* - Restrictions under which the optical information analysis holds (e.g. flat level earth)

Applications of Information Analysis

- **Airport/Vertiport Design - Layout of landing surfaces, surface markings, and approach lighting**
- **Display Design - Determination of important format and content considerations**

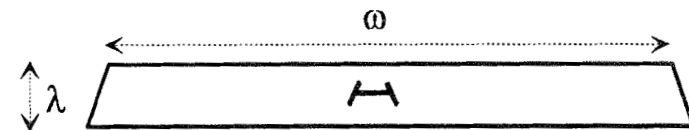
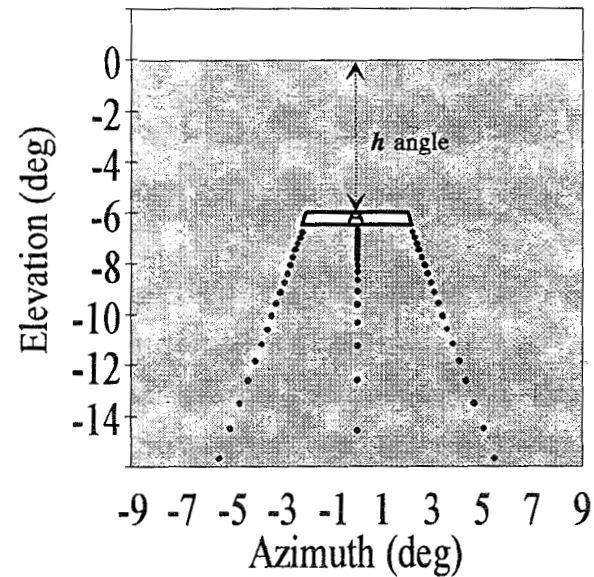
Information for Glideslope

$\alpha = \text{Glideslope}$

$$\alpha = \tan^{-1} \frac{z}{x}$$

$$\alpha = h$$

$$\alpha \approx \frac{W}{L} \bullet \frac{\lambda}{\omega}$$



Form Ratio = λ/ω

Constraints

h angle : correct horizon

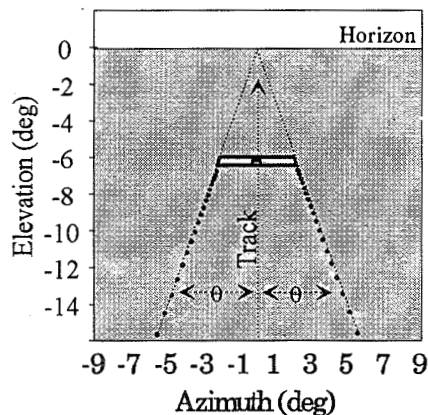
Form Ratio : Experience with pad dimensions

Optical Splay Rate

$$\text{Splay Rate} = \dot{\theta} = \frac{\dot{h}}{h} \cdot \sin\theta \cdot \cos\theta$$

θ = Angle between track vector and location on the ground, h = height above ground

Splay rate is globally modified by, and is useful for controlling, sink rate scaled in altitude units. It specifies rate of closure, or time to contact, with the ground.



All locations along paths parallel to the track vector have the same splay angle and splay rate.

Optical Flow Rate

$$\left(\frac{\dot{s}}{h}\right) \cdot \sin^2 \beta \cdot \sqrt{\left((\sin^2 \alpha \cdot \csc^2 \beta + \cos^2 \alpha) \cdot \cos^2 \gamma + \cot^2 \beta \cdot \sin^2 \gamma - \cos \alpha \cdot \cot \beta \cdot \sin \gamma \cdot \cos \gamma\right)}$$

a = azimuth, b = elevation, g = glideangle, $\dot{s}$ = path speed, h = height above ground

Optical flow rate, as defined above, is the angular speed of optical elements associated with points on a level ground plane. Flow rate is globally modified as a function of path speed scaled in altitude units, and is therefore useful for controlling this parameter.

Optical Edge Rate

$$\frac{\dot{x}}{\text{size}}$$

$\dot{x}$ = ground speed, size = size/spacing of salient ground objects

403

Optical edge rate is the rate (frequency) with which optical discontinuities pass across an optical region or location. Edge rate is a function of groundspeed scaled in terms of the size or spacing of salient ground objects, and therefore useful for controlling this parameter.

Relative Optical Expansion Rate (Tau)

$$\frac{\dot{\theta}}{\theta} \approx \frac{\dot{s}}{r}$$

θ = Optical (angular) size of object being approached, $\dot{s}$ = path speed, r = range to the object

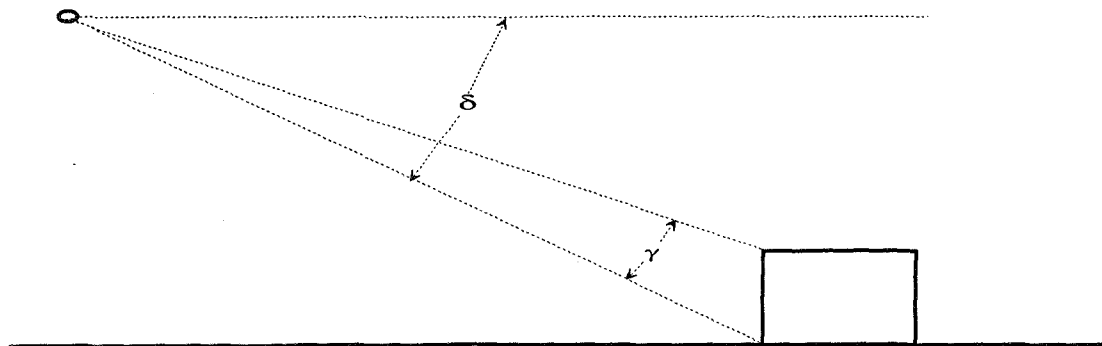
404

The relative optical expansion rate is a function of path speed and range to the object being approached, and is the relative (%/s) rate at which the angular size is changing. The inverse of this parameter is the projected time to arrival (tau). Therefore this is useful for controlling these quantities.

Information for Altitude

Horizon ratio $R = \gamma/\delta$

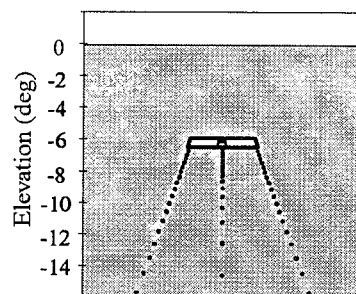
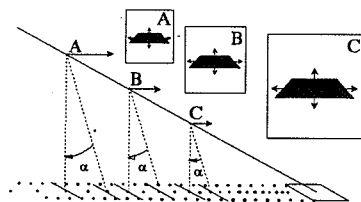
Horizon ratio is the ratio of the optical height of an object to the optical separation of the object base from the horizon. The horizon ratio is a function of the observer altitude and the object height, and approximates height above ground scaled in object height units.



Constraints
Correct Horizon

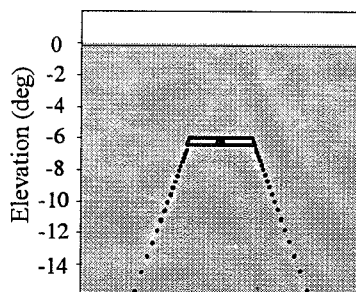
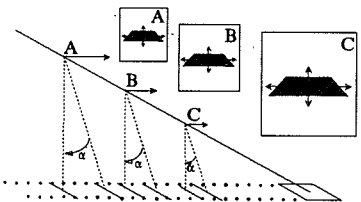
VMS Approach Lighting Study

Exponentially plus regular spaced lights



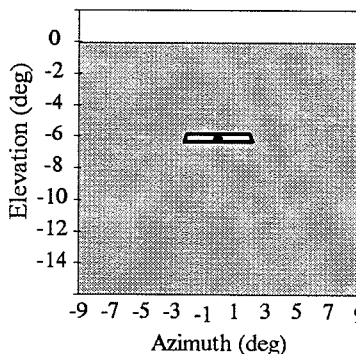
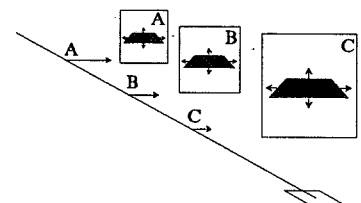
Near & Far Optic Flow Rate - Path speed/Altitude
Optic Edge Rate - Groundspeed, GroundSpeed/Range
Optical Expansion Rate - Path Speed/Range
Near & Far Optic Splay - Sink Rate/Altitude
h angle - Glideslope
Form Ratio - Glideslope

Regularly spaced lights



Near & Far Optic Flow Rate - Path speed/Altitude
Optic Edge Rate - Groundspeed
Optical Expansion Rate - Path Speed/Range
Near & Far Optic Splay - Sink Rate/Altitude
h angle - Glideslope
Form Ratio - Glideslope

No Lights



Far Optic Flow Rate - Path speed/Altitude
Optical Expansion Rate - Path Speed/Range
Far Optic Splay - Sink Rate/Altitude
h angle - Glideslope
Form Ratio - Glideslope

6MIT

VI. IMAGE EVALUATION AND METRICS

Sensor Fusion Display Evaluation Using Information Integration Models in Enhanced/Synthetic Vision Applications

David C. Foyle
NASA Ames Research Center

ABSTRACT

Based on existing integration models in the psychological literature, an evaluation framework is developed to assess sensor fusion displays as might be implemented in an enhanced/synthetic vision system. The proposed evaluation framework for evaluating the operator's ability to use such systems is a normative approach: The pilot's performance with the sensor fusion image is compared to models' predictions based on the pilot's performance when viewing the original component sensor images prior to fusion. This allows for the determination as to when a sensor fusion system leads to: 1) poorer performance than one of the original sensor displays (clearly an undesirable system in which the fused sensor system causes some distortion or interference); 2) better performance than with either single sensor system alone, but at a sub-optimal (compared to model predictions) level; 3) optimal performance (compared to model predictions); or, 4) super-optimal performance, which may occur if the operator were able to use some highly diagnostic "emergent features" in the sensor fusion display, which were unavailable in the original sensor displays.

INTRODUCTION

Many different types of imaging sensors exist, each sensitive to a different region of the electromagnetic spectrum. Passive sensors, which collect energy emitted or reflected from a source, include television (visible light), night-vision devices (intensified visible and near-infrared light), passive millimeter wave sensors, and thermal imaging (infrared) sensors. Active sensors, in which objects are irradiated and the energy reflected from those objects is collected, include the various bands of radar (radio waves), such as x-band and millimeter wave.

These imaging sensors were developed because of their ability to increase the probability of identification or detection of objects under difficult environmental conditions. Because each sensor is sensitive to different portions of the spectrum, the resultant images contain different information when used under the same conditions. In order to present this information to an operator, image processing algorithms are being developed in many laboratories to "fuse" the information into a single coherent image containing information from more than one sensor (Toet, 1990; Pavel, Larimer & Ahumada, 1992). These displays are referred to as sensor fusion displays.

Sensor fusion displays are being considered in enhanced or synthetic vision systems for civil transport use. These displays would allow pilots to detect runway features and incursions during landing, and would aid in detecting obstacles and traffic in taxi (Foyle, Ahumada, Larimer & Sweet, 1992). Such sensor systems would allow continued operation in low-visibility weather conditions (i.e., the sensors would "see" through the fog).

Much of the role of enhanced and synthetic vision systems with sensor fusion can be characterized as a detection task for the pilot. These systems must allow the pilot to detect runway incursions by ground vehicles and by other aircraft, and to detect obstacles in taxi to the gate. Additionally, in order to complete an approach at an airport, the pilot must verify (detect) any of ten different visual references (see Table 1).

VISUAL REFERENCES TO COMPLETE APPROACH (FROM ACJ-OPS 1-3.20001 AND SIMILAR TO FAR 91.175)
THE APPROACH LIGHT SYSTEM
THE THRESHOLD
THE THRESHOLD MARKING
THE THRESHOLD LIGHTS
THRESHOLD IDENTIFICATION LIGHT
THE VISUAL GLIDE SLOPE INDICATOR
THE TOUCHDOWN ZONE OR TOUCHDOWN ZONE MARKINGS
THE TOUCHDOWN ZONE LIGHTS
THE RUNWAY LIGHTS

Table 1. Visual references required to be seen by the pilot at decision height to complete an approach under current FAA rules.

The work described in this paper was conducted to guide the development of such sensor fusion displays. An engineer developing such a system constantly reviews the resulting display and underlying algorithms on a subjective basis. More formal testing is also necessary. Suppose, for example, that two sensor sources individually allow the pilot to achieve 0.70 probability of runway incursion detection under some particular environmental conditions. What, then, is the expected probability of runway incursion detection when the two sensors are combined according to some image processing technique? If observed runway incursion detection improves with a sensor fusion system to 0.80, is that a large improvement, or should one actually expect more? The ability to answer these types of questions can lead to a better human-machine system in two ways: Proposed sensor integration hardware and software can be evaluated both relatively, by determining which sensors and algorithm combinations are better than others, and absolutely, by comparing system (pilot/display) performance to theoretical expectations.

INFORMATION INTEGRATION MODELS

Previous work has been conducted on the topic of how operators integrate the information from multicomponent auditory signals, from the visual and auditory senses, and from multiple observations over time (Green, 1958; Craig, Colquhoun & Corcoran, 1976; Green & Swets, 1966/1974). These models all predict operator integration performance as a function of the operator's performance with the individual stimuli

comprising the integration task. Two classes of models have been developed: Decision combination models and observation integration models (Swets, 1984). The decision combination models assume that in the integration task the operator makes an individual decision about each aspect of the combined display and then combines those decisions to yield one final decision. At the time of the final decision, only the previous decisions are available, and not the information that led to the individual decisions. The observation integration models, on the contrary, assume that the operator does have access to that information. The internal representations of the individual observations (e.g., likelihood ratios) are then combined, yielding only one decision.

The simplest version of a decision combination model is the probability summation, or statistical summation, model. It is derived from the independence theorem of probability theory and was first proposed by Pirenne as a perceptual model (Pirenne, 1943; Swets, 1984). In its simplest form, the two information sources are assumed to be independent and uncorrelated. It states that performance with a complex stimulus is predictable from the performance with the individual stimuli according to the following equation:

$$p_{12} = p_1 + p_2 - p_1 p_2 \quad (1)$$

where p_1 and p_2 represent detection probabilities for the two stimuli presented in isolation, and p_{12} is the detection probability when both stimuli are available.

The most cited version of the observation integration model is derived from the theory of signal detectability and was originally proposed by Green (1958). As in Pirenne's (1943) model, in its most simple form, the information from the two sources is also assumed to be independent and uncorrelated. The model is stated in terms of the sensitivity measure, d' :

$$d'_{12} = \sqrt{(d'_1)^2 + (d'_2)^2} \quad (2)$$

where d'_1 and d'_2 , and d'_{12} , respectively, represent performance with the two stimuli presented in isolation, and when both stimuli are available.

Swets has noted that the statistical summation model fits simple detection data fairly well when the observed detection probabilities are corrected for chance success (Swets, 1984). Similarly, in the experiments in which it has been applied, the observation integration model well represents the data.

The two integration models presented here have been incorporated into the development of a framework which can be used to evaluate combined human-machine performance for sensor fusion displays.

PROPOSED EVALUATION FRAMEWORK

A sensor fusion display typically refers to the combined image display resulting from the application of an image processing technique on two or more individual sensor

images. The proposed framework for evaluating the operator's ability to use such systems is a normative approach: The operator's performance with the sensor fusion display can be compared to performance on the individual sensor displays comprising that display and to various optimal models of integration.

Typically, as the environmental conditions change in which the individual sensor operates, so does the information content of that image. The information content of the image can be "scaled" by the operator's ability to perform a target identification or discrimination task (e.g., detecting a runway incursion). One would expect task performance with a sensor fusion display formed from two low information content (hence poor performance) images, to still be relatively poor. Similarly, two high information content (high performance) sensor images should yield good performance

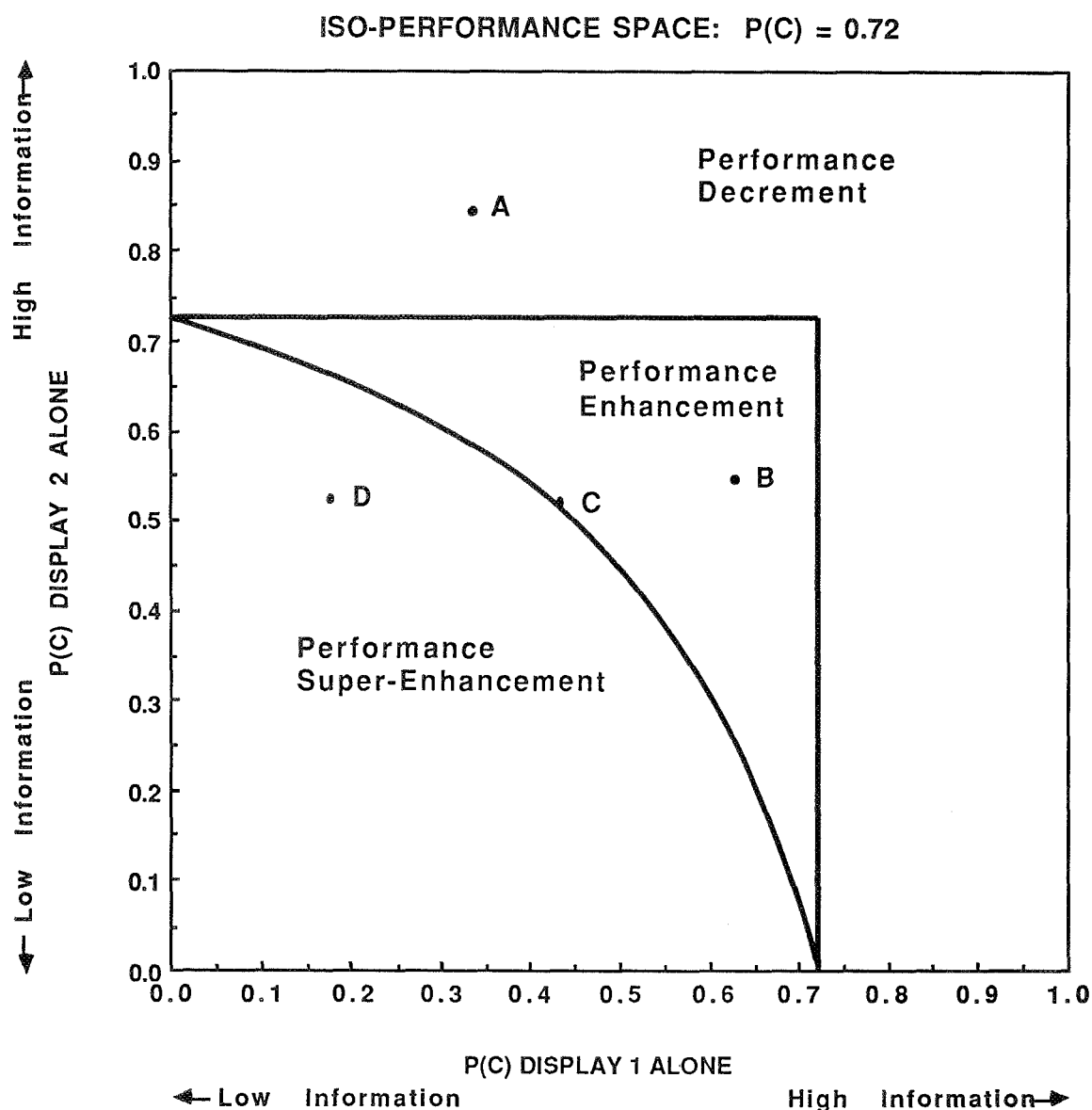


Fig. 1. A proposed evaluation framework for sensor fusion displays. All data points represent $P(C) = 0.72$ for the dual display or sensor fusion display task.

when combined into a sensor fusion display. Assuming that there was some independent information in the two individual sensor images, one would also expect performance with the sensor fusion display to be better than with either of the two individual sensors alone. This results in a 3-dimensional performance space: Performance with the sensor fusion image is a function of the performance levels associated with the two individual sensor images.

Figure 1 shows part of this performance space associated with a sensor fusion display. The abscissa and the ordinate result from the stimulus-performance scaling for Sensor Display 1 and Sensor Display 2, respectively, when viewed by an operator in isolation. The figure shows the iso-performance horizontal "slice" through the

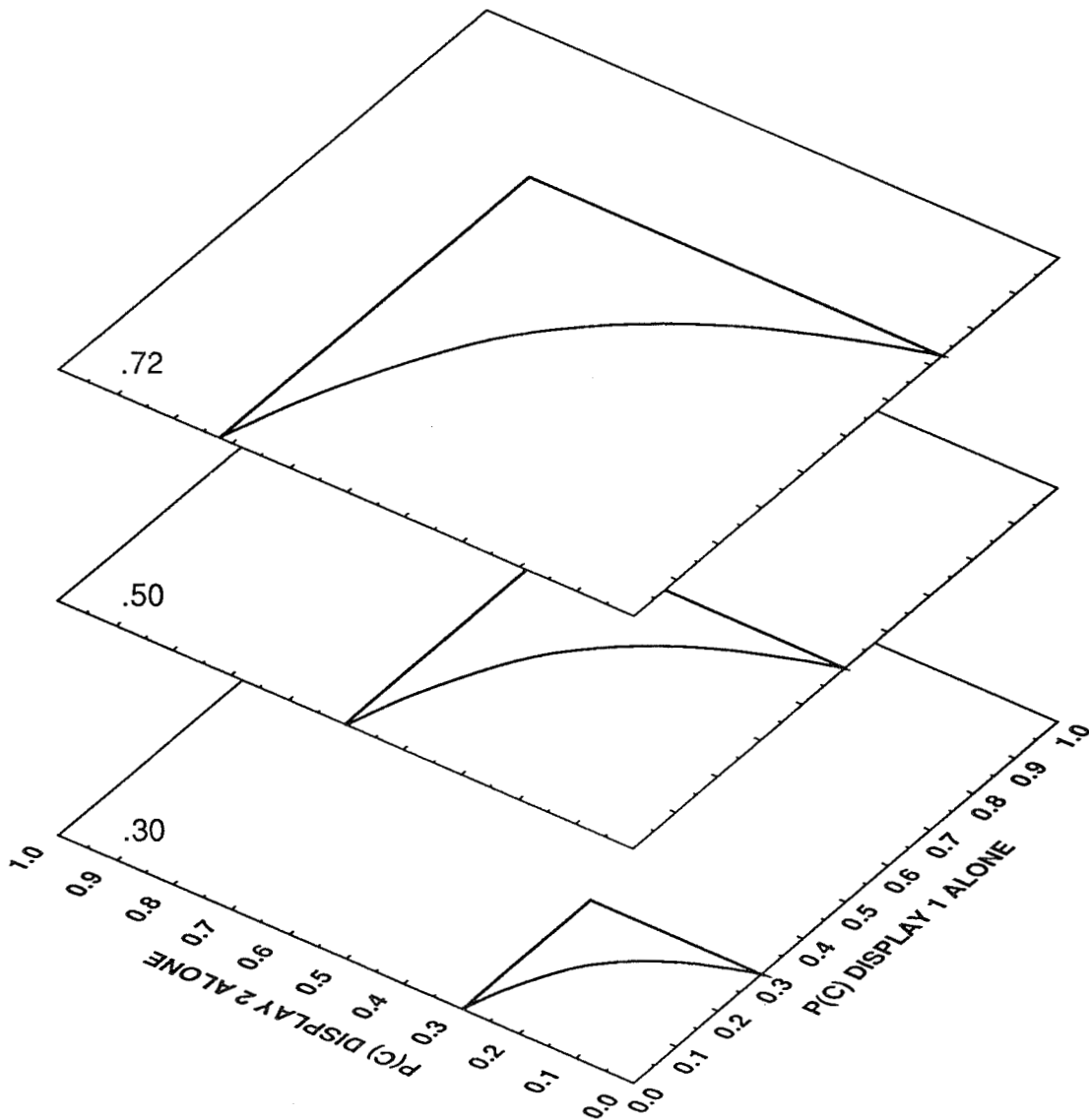


Fig. 2. Three example horizontal slices through the 3-dimensional performance space. The value on each overlay represents the performance level, in $P(C)$, for the sensor fusion display task.

3-dimensional space in which all performance data points represent 0.72 (corrected for chance) detection probability using a sensor fusion display. As noted above, the actual performance space is 3-dimensional and is represented in Figure 2 by similar-appearing "slices" for three example performance levels.

Because the sensor fusion display data are plotted as iso-performance slices, data points near the origin represent better performance than away from the origin. For the same level of performance, a data point near the origin represents a condition in which very little information was available in the two displays, whereas a data point away from the origin refers to a condition in which relatively more information was available in the separate displays. Thus, for a given resultant sensor fusion performance level (i.e., "horizontal slice") data points near the origin represent better sensor fusion displays.

In these figures and all remaining references, $P(C)$ refers to the proportion of correct responses with a correction for chance applied. A correction for chance is necessary when measuring performance in $P(C)$ units because the integration models require that a performance level of zero be associated with the operator receiving no information from the display. No such correction is necessary when measuring performance in d' units since $d' = 0$ refers to chance performance.

As can be seen from the two figures, the sensor fusion performance space can be divided into three separate areas, Performance Decrement, Performance Enhancement, and Performance Super-Enhancement, each with unique interpretations if data points lie in those areas. The two right-angle lines dividing the Performance Decrement and Performance Enhancement areas are determined by the horizontal and vertical lines crossing the axes at the level of performance [$P(C) = 0.72$ in Figure 1] for the sensor fusion display. The smooth curves separating the Performance Enhancement and Performance Super-Enhancement areas are the predictions of the statistical summation model (see eq. 1) where $p_{12} = 0.72$ in Figure 1 and 0.30, 0.50, and 0.72 in Figure 2. Because these two models predict optimal performance (that is, they both assume ideal observers with no memory limitations, etc., with independent and uncorrelated information in the separate displays) their predictions can be used as an upper bound against which to measure integration performance. The interpretation of the data points falling into the three areas is best illustrated by example.

Performance decrement

Suppose under a given environmental condition, an operator achieved runway incursion detection performance of $P(C) = 0.33$ when viewing Sensor 1 in isolation and $P(C) = 0.84$ when viewing Sensor 2 in isolation. When these two sources are both available (separately on two monitors, or fused on a single monitor according to a sensor fusion algorithm) to the operator and performance is $P(C) = 0.72$, the resultant data point would be the one labeled "A" in Figure 1. Obviously, in this situation, the sensor fusion display has not improved the pilot's overall runway incursion detection performance. In fact, performance in the sensor fusion display case has now decreased to only $P(C) = 0.72$, whereas previously the operator was able to use Sensor 1 in isolation and reach $P(C) = 0.84$ performance. Such a performance decrement could be the result of the deletion of necessary information by the sensor fusion algorithm, or could represent a cognitive limitation on the part of the pilot.

Performance enhancement

Data point "B" in Figure 1 would result if $P(C) = 0.72$ performance obtained using the sensor fusion display, when Sensors 1 and 2 yielded $P(C) = 0.63$ and $P(C) = 0.55$ in isolation. In this case, performance has improved, since the pilot is now doing better with the sensor fusion display (0.72) than with either of the two sources alone (0.55, 0.63). However, the two models of information integration predict a larger improvement in this case. Thus, for data points falling in this region, there is performance improvement, but one would expect more. Data point "C", lying on the statistical summation model curve, represents optimal integration performance, in which sensor fusion display performance of 0.72 is expected if performance on Sensor 1 were 0.42 and Sensor 2 performance was 0.52.

Pilot detection performance occurring in this region would occur when some of the information in the two sources is redundant (correlated and not independent), or when the sensor fusion algorithm integrates the information suboptimally. The statistical summation model (as well as the observation integration model) can be viewed as an upper limit of integration: It assumes that the information in the two sources is independent and non-redundant, and does not assume any decrease in performance due to the limits of cognitive processes (i.e., memory, workload, or suboptimal strategies).

Performance super-enhancement

Data point "D" would result when the individual runway incursion detection performance for the two sensors alone was $P(C) = 0.17$ and $P(C) = 0.52$ and sensor fusion display performance was $P(C) = 0.72$. Data points falling in this region between the model prediction and the origin represent improved performance that is better than is predictable from the model. That is, when the sensor fusion display is viewed, some new, previously unusable, information emerges which results in much better performance.

The random-dot stereogram display can be thought of as an example of a sensor fusion display that has these properties (Julesz, 1971). In these displays, random dots are offset differentially yielding a perception of an object in the third dimension. In such a stereogram there is no information whatsoever in the individual halves of the stereogram, but only in differences between the two displays. The object is observable only by stereoscopically fusing the two halves of the stereogram or analytically determining the differences. In fact, if one conducted an experiment in which subjects had to state the "floating" shape, one would obtain chance performance when viewing only one stereogram half and perfect performance when both stereogram pairs are viewed. This represents Performance Super-Enhancement because based on chance performance with the stereogram halves, one would conclude that they contain no information. This would lead one to predict chance performance when both halves are available, which obviously is not the case. Conditions in which Performance Super-Enhancement occurs could be capitalized upon to produce useful sensor fusion techniques. The proposed evaluation framework provides for the ability to recognize and quantify such conditions.

Evaluation framework implementation

In order to evaluate human performance with a sensor fusion system using the proposed evaluation framework, the following steps must be taken:

Performance scaling of Sensor 1. Determine the psychometric function relating task performance (e.g., runway incursion detection, runway lights detection) to the environmental conditions of interest. For example, infrared imagery is degraded by increasing atmospheric moisture. The information content of each sensor image varies with the environmental conditions, and in a sense, this scaling estimates the amount of information available to the operator with Sensor 1 alone under those conditions.

Performance scaling of Sensor 2. Similar to Sensor 1.

Performance with sensor fusion display. For various combinations of environmental or sensor conditions previously evaluated in isolation, determine task performance using the proposed fusion algorithm and associated display.

Performance with operator integration. As in the sensor fusion evaluation phase, determine task performance with both sensors but with either two displays or a split screen. This step acts as a control condition, and essentially allows the operator to integrate the information from the two sensors. A sensor fusion algorithm should yield better task performance than when the operator uses two displays or a split-screen display.

SENSOR FUSION EVALUATION: FOYLE (1992)

To illustrate how the evaluation framework would be used the results from an experiment are briefly presented. In an experiment reported in Foyle (1992), subjects had to integrate the information in two sensor displays to detect a target. As an experimental convenience, combinations of separate sensor sources yielding an iso-performance level [$P(C) = 0.72$] of integration performance were determined (with both sensor sources available on multiple screens, analogous to performance with a sensor fusion display). These combinations were then plotted on the evaluation framework graph.

Figure 3 shows combinations of the individual sensor sources, in $P(C)$ units as scaled by $P(C)$ psychometric functions, yielding $P(C) = 0.72$ dual-display (sensor fusion) performance. The two curves represent predictions of the two optimal integration models (statistical summation and observation integration) as described by the equations shown in the figures. For illustration purposes, note the right-most (also lower-most) data point for subject 4. That data point shows that $P(C) = 0.72$ detection performance obtained when viewing two sensor displays simultaneously: A Sensor 1 image display which yielded $P(C) = 0.60$ probability of detection alone, and a Sensor 2 image display which yielded $P(C) = 0.36$ probability of detection alone.

Analyzing the results of this experiment using this method, Foyle (1992) concluded that ten of the eighteen data points in Figure 3 lie in the triangular "performance enhancement" region when plotted onto the evaluation framework graph. For those conditions, the subjects were able to integrate the images from the two displays and performed better than when only one of those displays was available. The conditions that led to integration occurred when Sensor Display 1 yielded moderate detection performance (approximately $P(C) = 0.50$ in Figure 3). When a low-quality image (yielding about $P(C) = 0.30$) was presented as Sensor Display 1, the images in Sensor Display 2 were required to be of very high-quality in order to yield $P(C) = 0.72$ with both displays. In fact,

they were of such high quality that when presented in isolation, they would have yielded performance of $P(C) = 0.80$ or 0.90 . The subjects would have done better in those conditions if they had simply ignored the low-quality images on Sensor Display 1 and based their responses only on the images on Sensor Display 2. (Graphically, that would have forced the data points onto the horizontal straight line in Figure 3.)

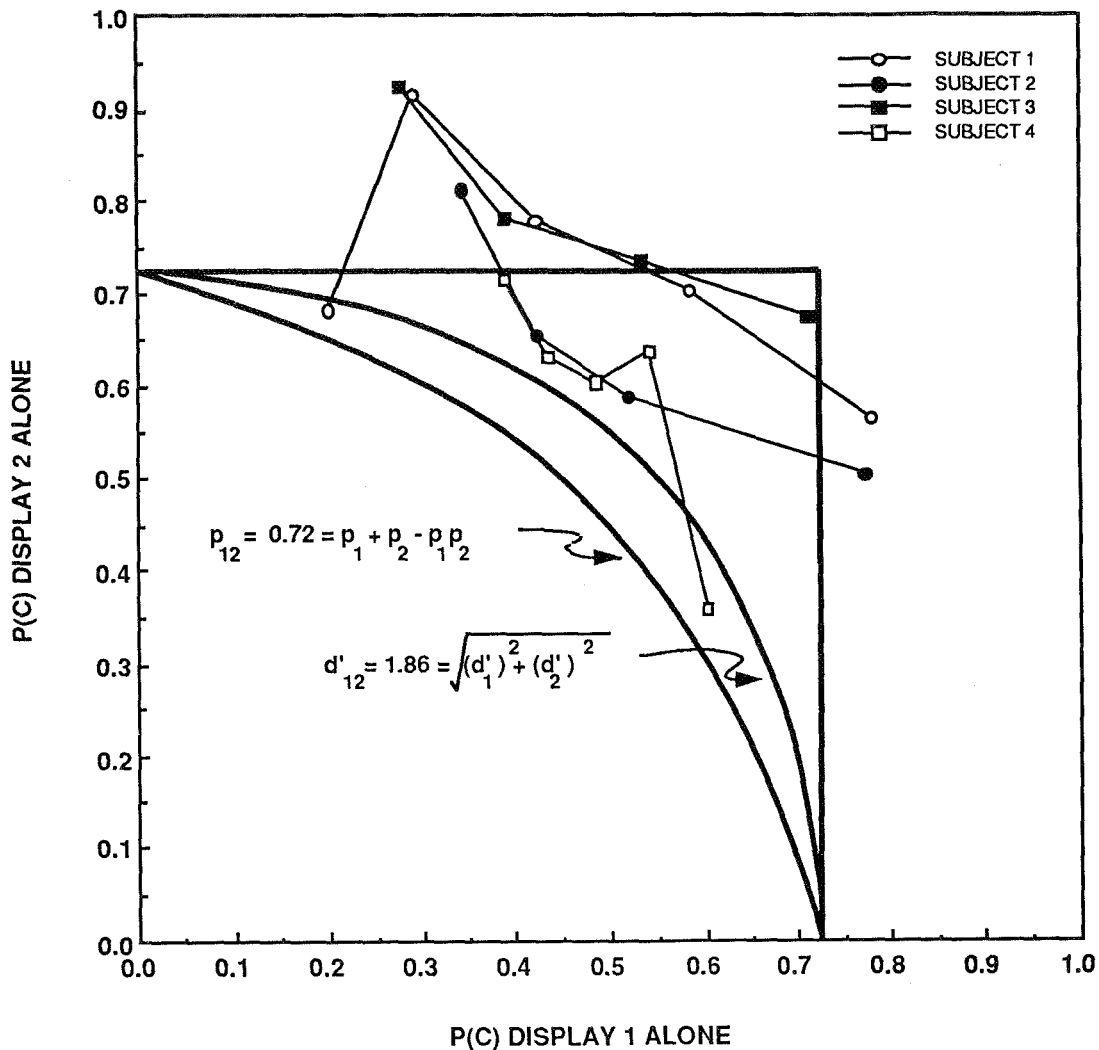


Fig. 3. Experimental data from Foyle (1992), in corrected-for-chance $P(C)$, overlaid on the proposed evaluation framework. All data points in this "horizontal slice" through the 3-dimensional space represent $P(C) = 0.72$ detection probability performance.

These data were explained by a model in which subjects always give equal weight to the information in the two displays despite the image quality level. The effect may be similar to that noted by Tversky and Kahneman (1974) in which subjects weighted obviously irrelevant information equally with relevant information. The conditions under which subjects are able to integrate display information, and those that do not facilitate, and actually decrease performance clearly warrant more investigation. As stated earlier, the statistical summation and observation integration models can be viewed as an upper bound to normal (not Performance Super-Enhancement) information integration. In this

particular experiment, the model predictions were not only an upper bound on performance in general, but in fact were appropriate predictions since the information in the dual-display condition was independent and uncorrelated. The models' failure to predict the data establishes the existence of the subjects' cognitive limitations in this particular task.

CONCLUSIONS AND SUMMARY

For a sensor fusion display in an enhanced or synthetic vision system, much of what the pilot must do with the system is to detect traffic and detect certain visual references in order to complete an approach and land. The evaluation framework described in this paper allows system engineers and researchers to evaluate pilot-in-the-loop performance with the sensor fusion algorithms and display against a theoretical optimal benchmark. By using such a benchmark, the system engineer can ensure that the important features available in the sensor imagery prior to fusion are preserved.

In summary, the evaluation framework developed herein has been demonstrated to be a useful tool to evaluate pilot's ability to extract information from a sensor fusion display or to integrate information from two displays. The techniques discussed allow the evaluation of sensor fusion displays by comparing sensor fusion display performance to the predictions of existing optimal integration models and to multiple display presentations. This evaluation allows the human factors engineer to recognize in an absolute sense, as well as relative, whether the proposed sensor fusion display does what it was designed to do: integrate the sensor information and present it well.

REFERENCES

- Craig, A., Colquhoun, W.P. and Corcoran, D.W.J. (1976). Combining evidence presented simultaneously to the eye and the ear: A comparison of some predictive models. *Perception & Psychophysics*, 19, 473-484.
- Foyle, D.C. (1992). Proposed evaluation framework for assessing operator performance with multisensor displays. SPIE Volume 1666: *Human Vision, Visual Processing, and Digital Display III*, 514-525.
- Foyle, D.C., Ahumada, A.J., Larimer, J. and Sweet, B.T. (1992). Enhanced/synthetic vision systems: Human factors research and implications for future systems (Paper 921968). *Enhanced Situation Awareness Technology for Retrofit and Advanced Cockpit Design*, SAE SP-933: Conference Proceedings of the SAE Aerotech '92 Meeting (Human Behavioral Technology/Aerospace Technologies Activity).
- Green, D.M. (1958). Detection of multiple component signals in noise. *Journal of the Acoustical Society of America*, 30, 904-911.
- Green, D.M. and Swets, J.A. (1974). *Signal Detection Theory and Psychophysics*, Krieger, New York, 1974. (Originally published 1966 by Wiley, New York.)
- Julesz, B. (1971). *Foundations of Cyclopean Perception*, University of Chicago Press, Chicago.

- Pavel, M., Larimer, J., and Ahumada, A. (1992). Sensor fusion for synthetic vision. *SID International Symposium Digest of Technical Papers*, 23, 475-478.
- Pirenne, M.H. (1943). Binocular and unocular thresholds in vision. *Nature*, 152, 698-699.
- Swets, J.A. (1984). Mathematical models of attention. In *Varieties of Attention*, R. Parasuraman and D.R. Davies (Eds.), 183 - 242, Academic Press, New York.
- Toet, A. (1990). Hierarchical image fusion. *Machine Vision and Applications*, 3, 1-11.
- Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty. *Science*, 185, 1124-1130.

Quality Metrics for Sensor Images

Al Ahumada
NASA Ames Research Center

523-54

205753

24944 p. 2

Methods are needed for evaluating the quality of Augmented Visual Displays (AVID). Computational quality metrics will help summarize, interpolate, and extrapolate the results of human performance tests with displays. The FLM Vision group at NASA Ames has been developing computational models of visual processing and using them to develop computational metrics for similar problems, for example,

- 1) Display modeling systems use metrics for comparing proposed displays (Martin, Ahumada, and Larimer, 1992; Lubin, 1993).
- 2) Halftoning optimizing methods use metrics to evaluate the difference between the halftone and the original (Mulligan and Ahumada, 1992).
- 3) Image compression methods minimize the predicted visibility of compression artifacts (Peterson, Ahumada, and Watson, 1993; Watson, 1993).

The visual discrimination models take as input two arbitrary images A and B, and compute an estimate of the probability that a human observer will report that A is different from B. If A is an image that one desires to display and B is the actual displayed image, such an estimate can be regarded as an image quality metric reflecting how well B approximates A (Watson, 1983; Nielsen, Watson, and Ahumada, 1985).

There are additional complexities associated with the problem of evaluating the quality of radar and IR enhanced displays for AVID tasks.

One important problem is the question of whether intruding obstacles are detectable in such displays. Although the discrimination model can handle detection situations by making B the original image A plus the intrusion, this detection model makes the inappropriate assumption that the observer knows where the intrusion will be. Effects of signal uncertainty as studied by Pelli (1985), for example, need to be added to our models.

A pilot needs to make his decisions rapidly. Our models need to predict not just the probability of a correct decision, but the probability of a correct decision by the time the decision needs to be made. That is, the models need to predict latency as well as accuracy. Luce and Green have generated models for auditory detection latencies. Similar models are needed for visual detection.

Most image quality models are designed for static imagery. Watson has been developing a general spatial-temporal vision model to optimize video compression techniques. These models need to be adapted and calibrated for AVID applications.

Radar images especially are characterized by high levels of noise. Although detection and discrimination models have been developed for noisy images (Legge, Kersten, and Burgess, 1987; Barrett, 1992), their features have not been integrated into our current models.

Models have been developed within our group to predict a pilot's 3D heading estimate from a video display (Perrone, 1992; Heeger and Jepson, 1992). These models can be developed into quality measures relating to the pilot's ability to gather dynamic orientation information from such displays.

References

- H. H. Barrett (1992) Evaluation of image quality through linear discriminant models. In J. Morreale (Ed.), *Society for Information Display Digest of Technical Papers*, 871-873, Playa del Rey, CA: Society for Information Display.
- D. J. Heeger and A. D. Jepson (1992) Subspace methods for recovering rigid motion I: Algorithm and Implementation. *Intern. J. Computer Vision*, 7, 95-117.
- G. E. Legge, D. Kersten, and A. E. Burgess (1987) Contrast discrimination in noise. *J. Opt. Soc. America A*, 4, 381-404.
- J. Lubin (1993) The use of psychophysical data and models in the analysis of display system performance. In A. B. Watson (Ed.), *Visual Factors in Electronic Image Communications*, Cambridge, MA: MIT Press.
- R. A. Martin, A. J. Ahumada, Jr., and J. O. Larimer (1992) Color matrix display simulation based upon luminance and chromatic contrast sensitivity of early vision. In B. E. Rogowitz (Ed.), *Human Vision, Visual Processing, and Digital Display III*, Proc. 1666, 336-342, Bellingham, WA: SPIE.
- J. B. Mulligan and A. J. Ahumada, Jr. (1992) Principled halftoning based on models of human vision. In B. E. Rogowitz (Ed.), *Human Vision, Visual Processing, and Digital Display III*, Proc. 1666, 109-121, Bellingham, WA: SPIE.
- K. R. K. Nielsen, A. B. Watson, and A. J. Ahumada, Jr. (1985) Application of a computable model of human spatial vision to phase discrimination. *J. Opt. Soc. America A*, 2, 1600-1606.
- D. G. Pelli (1985) Uncertainty explains many aspects of visual contrast detection and discrimination. *J. Opt. Soc. America A*, 2, 1508-1532.
- J. A. Perrone (1992) Model for the computation of self-motion in biological systems. *J. Opt. Soc. America A*, 9, 177-194.
- H. Peterson, A. Ahumada, and A. Watson (1993) An improved detection model for DCT coefficient quantization. In B. Rogowitz and J. Allebach (Eds.), *Human Vision, Visual Processing, and Digital Display IV*, Proc. 1913, Bellingham, WA: SPIE.
- A. B. Watson (1983) Detection and recognition of simple spatial forms. In O. J. Braddick and A. C. Sleigh (Eds.), *Physical and Biological Processing of Images*, Berlin: Springer-Verlag.
- A. B. Watson (1993) DCT quantization matrices visually optimized for individual images. In B. Rogowitz and J. Allebach (Eds.), *Human Vision, Visual Processing, and Digital Display IV*, Proc. 1913, Bellingham, WA: SPIE.

Evaluation of Image Quality

M. Pavel
NASA Ames Research Center
and
Western Aerospace Laboratories

524-54

205754

p-9
249945

ABSTRACT

This presentation outlines a general approach to the evaluation of display system quality for aviation applications. This approach is based on the assumption that it is possible to develop a model of the display which captures most of the significant properties of the display. The display characteristics should include spatial and temporal resolution, intensity quantizing effects, spatial sampling, delays, etc. The model must be sufficiently well specified to permit generation of stimuli that simulate the output of the display system.

The first step in the evaluation of display quality is an analysis of the tasks to be performed using the display. Thus, for example, if a display is used by a pilot during a final approach, the aesthetic aspects of the display may be less relevant than its dynamic characteristics. The opposite task requirements may apply to imaging systems used for displaying navigation charts. Thus, display quality is defined with regard to one or more tasks.

Given a set of relevant tasks, there are many ways to approach display evaluation. The range of evaluation approaches includes visual inspection, rapid evaluation, part-task simulation, and full mission simulation.

The work described today is focused on two complementary approaches to rapid evaluation. The first approach is based on a model of the human visual system. A model of the human visual system is used to predict the performance of the selected tasks. The model-based evaluation approach permits very rapid and inexpensive evaluation of various design decisions.

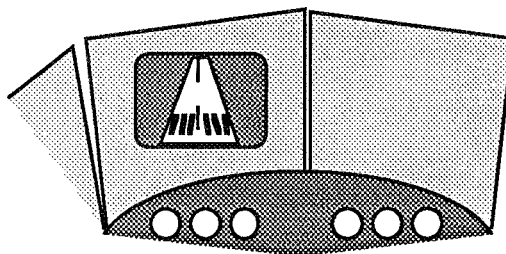
The second rapid evaluation approach employs specifically designed critical tests that embody many important characteristics of actual tasks. These are used in situations where a validated model is not available. These rapid evaluation tests are being implemented in a workstation environment.

EVALUATION

- Task Analysis
- Model-Based Evaluation
- Visual Inspection
- Rapid Laboratory Evaluation ←
- Part-Task Simulation
- Full Mission Simulation
- Flight Tests

NASA Ames Research Center

BASIC CONCEPT



NASA Ames Research Center

IMAGE QUALITY

Insertion Impairment

IMAGE → OBSERVER

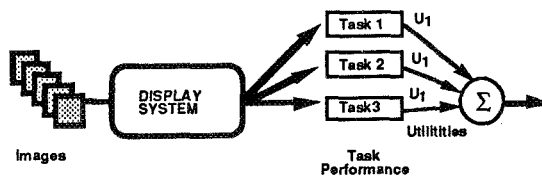


Examples: Insertion loss (attenuation), noise, delays, geometric distortions, etc.

NASA Ames Research Center

TASK-BASED IMAGE QUALITY METRIC

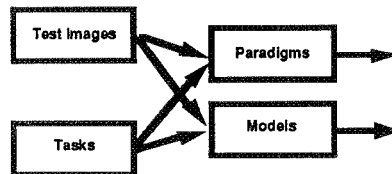
- Representative sample of images
- Performance measures
- Task utility



NASA Ames Research Center

RAPID EVALUATION WORKBENCH

- Images: Test patterns
- Tasks: Detection, alignment
- Models: Prediction of performance
- Tests: Empirical Paradigms



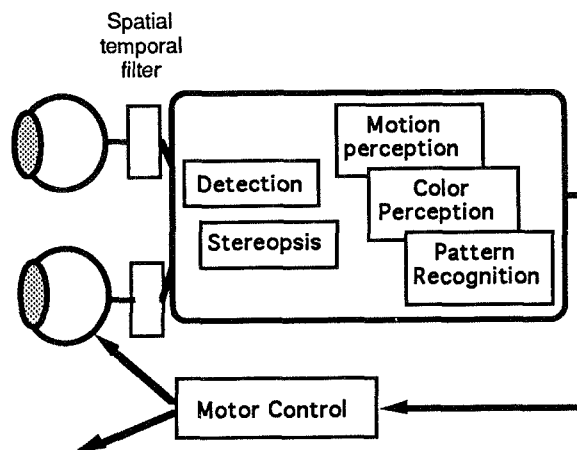
NASA Ames Research Center

TASK ANALYSIS

- Runway acquisition (at distance 10,000 ft)
- Runway identification (at distance 6,000 ft)
- Runway location
- Runway orientation
- Aimpoint estimation
- Traffic detection
- Hazard (e.g., runway incursion) detection

NASA Ames Research Center

MODEL BASED EVALUATION HUMAN VISUAL SYSTEM



NASA Ames Research Center

EMPIRICAL TASKS

- Bar detection in noise
- Edge orientation
- Visual search
- Vernier alignment
- Optic flow perception (self-motion)
- Motion perception

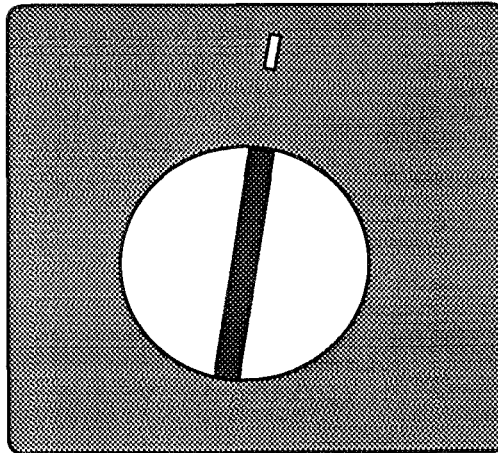
NASA Ames Research Center

DISPLAY CHARACTERISTICS

- Field of view, perspective, symbology
- Temporal Resolution, update rate, delay
- Quantization (spatial & gray-level)
- Spatial resolution, stroke, raster
- Reliability, noise, masking
- Contrast, brightness, color
- Geometric distortions
- Display stabilization
- Registration

NASA Ames Research Center

Rapid Evaluation Example



Task: Alignment of the bar and with the probe.

NASA Ames Research Center

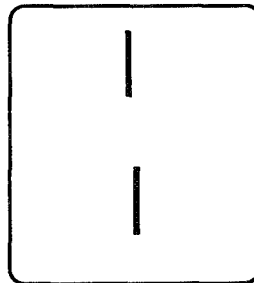
EXAMPLE: SEARCH



Task: To find a target — the lighter bar

NASA Ames Research Center

EXAMPLE: VERNIER ALIGNMENT



Task: To judge the relative position of the two vertical lines.

NASA Ames Research Center

SIMULATOR/FLIGHT PERFORMANCE

- Situational awareness
- Landing performance
- Landing dispersions
- Breaking performance
- Glideslope alignment
- Workload
- Training (Learning curves, retention)

NASA Ames Research Center

ARTIFACTS

- Geometric Illusions
 - Size
 - Distance
- Color illusions
 - Brightness
 - Color
- Motion illusions
 - Direction of moving objects
 - Direction of selfmotion

NASA Ames Research Center

SYMBOLOLOGY

- **TYPE OF INFORMATION**
 - Pitch bars
 - Glide slope
 - Velocity vector
 - Energy management
 - Wind conditions
 - Predicted path
- **SYMBOL DESIGN AND SELECTION**
- **SYMBOLOLOGY CLUTTER**

NASA Ames Research Center

APPENDIX: AVID WORKSHOP ATTENDEES

OMIT TO
END

Government

Dr. Randall L. Harris, Sr.
Mail Stop 152E
NASA Langley Research Center
Hampton, VA 23665-5225
(804) 864-6641
FAX: (804) 864-8294
r.l.harris@larc.nasa.gov

Dr. William M. Kahlbaum
Mail Stop 489
NASA Langley Research Center
Hampton, VA 23665-5225
(804) 864-4044
wmk@virisi.larc.nasa.gov

Ames Research Center
(Moffett Field, CA 94035-1000
415-604-XXXX)

Dr. Albert Ahumada
MS 262-2
4-6257
FAX: 4-3323

Dr. Yair Barniv
MS 210-9
4-5451

Mr. Vernol Battiste
MS 262-3
4-3666

Dr. Victor Cheng
MS 210-9
4-5424
FAX: 4-0174

Dr. Dallas Denery
MS 210-9
4-5427
FAX: 4-0174

Dr. R. Key Dismukes
MS 262-1
4-0150

Dr. David Foyle
MS 262-3
4-3053
FAX: 4-3323
foyle@eos.arc.nasa.gov

Ms. Sandra Hart
MS 262-3
4-6072

Dr. Bassam Hussien
MS 210-9
4-5439

Dr. Walter Johnson
MS 262-3
4-3667
FAX: 4-3323
wwj@eos.arc.nasa.gov

Dr. Mary Kaiser
MS 262-3
4-4448
FAX: 4-3323
moose@eos.arc.nasa.gov

Dr. J. Victor Lebacqz
MS 262-4
4-5792

Dr. Robert McCann
MS 262-3
4-0052
rsm@eos.arc.nasa.gov

Dr. Jeffrey Mulligan
MS 262-2
4-3745

Dr. Cynthia Null
MS 262-2
4-1260

Dr. Misha Pavel
MS 262-2
4-1447
mis@vision.arc.nasa.gov

Mr. Phillip Smith
MS 210-9
4-5469

Mr. C. Thomas Snyder
MS 200-3
4-5066

Dr. Banavar Sridhar
MS 210-9
4-5450
FAX: 4-0174
bsridhar@windchime.arc.nasa.gov

Mr. Raymond Suorsa
MS 210-9
4-6334
grendel@windchime.arc.nasa.gov

Dr. Irving Statler
MS 262-1
4-5729

Ms. Barbara Sweet
MS 262-4
4-0006
FAX: 4-3729
sweet@vision.arc.nasa.gov

Dr. Andrew B. Watson
MS 262-2
4-5419
FAX: 4-3323
beau@vision.arc.nasa.gov

Mr. Richard Zelenka
MS 210-9
4-5433
FAX: 4-0174

Industry

Mr. Yair Alon
Program Manager
Lear Astronics Corporation
3400 Airport Blvd
Santa Monica, CA 90406
(310) 915-6772
FAX: (310) 915-8182

Mr. Michael Bazakos
Honeywell SRC
3660 Technology Drive
Minneapolis, MN 55418
(612) 951-7533
FAX: (612) 951-7438

Mr. Earl T. Benser
Honeywell SRC
3660 Technology Drive
Minneapolis, MN 55418
(612) 951-7710
FAX: (612) 951-7438
benser@src.honeywell.com

Dr. Uri Bernstein
Technology Services Corp.
2950 31st Street
Santa Monica, CA 90405
(310) 450-9755
FAX: (310) 452-3175

Mr. Sidney A. Connor
Maryland Adv. Development Laboratory
6411 Ivy Lane, Suite 110
Greenbelt, MD 20770
(301) 345-8664
FAX: (301) 345-7305

Dr. Keith Hanna
David Sarnoff Research Center
CN 5300
Princeton, NJ 08562
(609) 734-2159
FAX: (609) 734-2662
kjh@vision.sarnoff.com

Dr. Bruce I. Hauss
TRW Applied Technology Division
1 Space Park
Building R1, Room 2128
Redondo Beach, CA 90278
(310) 812-0261
FAX: (310) 812-1277

Dr. J. Richard Kerr
Vice President, Engineering
FLIR Systems Inc.
16505 S.W. 72nd Ave.
Portland, OR 97224
(503) 684-3731
FAX: (503) 684-7046

Dr. Anthony Lambregts
Boeing Commercial Airplane Group
P.O. Box 3707
Seattle, WA 98124
(206) 965-5521

Mr. William F. O'Neil
Westinghouse ESG
P.O. Box 1693
Mail Station G8
Baltimore, MD 21203
(410) 765-0039
FAX: (410) 765-0013

Mr. Jeffrey Radke
Honeywell SRC
10701 Lyndale Ave South
Bloomington, MN 55442
(612) 887-4098
FAX: (612) 887-4517

Dr. Barry Roberts
Honeywell SRC
3660 Technology Drive
Minneapolis MN 55418
(612) 951-7535
FAX: (612) 951-7438

Mr. Anthony Stenger
Technology Service Corp.
2950 31st Street
Santa Monica, CA 90402
(310) 450-9755
FAX: (310) 452-3175

Mr. David G. Strong
Westinghouse Electric Corp.
P.O. Box 746
Mail Station 200
Baltimore, MD 21108
(410) 993-7114
FAX: (410) 993-7685

Military

Mr. Robert Smith
WL/FIGS Bldg. 146
2210 Eighth St. Ste. 11
Wright Patterson AFB, OH 45433
(513) 255-3546
FAX: (513) 255-5297
smithr@batman.flight.wpafb.af.mil

Academia

Dr. Rangachar Kasturi
Dept. of Computer Engineering
The Pennsylvania State University
University Park, PA 16802
(814) 863-4254
FAX: (814) 865-0192
kasturi@cmpe.psu.edu

Dr. P. K. Menon
Santa Clara University
Santa Clara, CA
(415) 493-5567

Private

Mr. Richard Bray
880 Helena Drive
Sunnyvale, CA 94087
(408) 245-1156
FAX: (415) 604-4000