Multiple Degree of Freedom Optical Pattern Recognition

D. Casasent
Carnegie Mellon University
Pittsburgh, PA 15213

ABSTRACT

Three general optical approaches to multiple degree of freedom object pattern recognition (where no stable object rest position exists) are advanced. These techniques include: feature extraction, correlation, and artificial intelligence. The details of the various processors are advanced together with initial results.

1. INTRODUCTION

This paper addresses object pattern recognition for *multiple degree of freedom* (M-DOF) image cases. This is defined as the recognition and identification of an object with no stable rest position. We emphasize *optical pattern recognition* (OPR) techniques and research for this problem, with recent results obtained at the *Center for Excellence in Optical Data Processing* at Carnegie Mellon University. Three different optical processing techniques are addressed and highlighted. These include: feature extraction (Section 2), correlation (Section 3) and optical artificial intelligence (Section 4).

2. OPTICAL FEATURE EXTRACTION FOR M-DOF PATTERN RECOGNITION

The general feature extraction approach to pattern recognition [1] is diagramed in Figure 1. In this section, we emphasize different feature spaces that can be optically realized. The feature extraction and classification techniques are established [2]. All feature spaces we consider are in-plane distortion-invariant. We achieve 3-D M-DOF distortion-invariance by training sets and use of *linear discriminant functions* (LDFs).

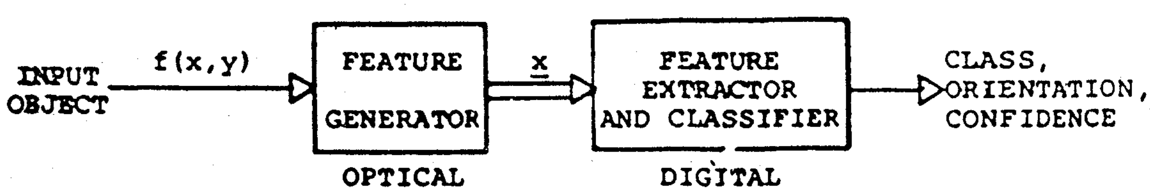

Figure 1: *Hybrid Optical/Digital Feature Extraction Processor Block Diagram*

2.1 CHORD DISTRIBUTION FEATURE SPACE

This feature space consists of the distributions h(r) of the length (r) and the distributions h(θ) of the angles (θ) of all chords associated with an input object. We allow gray-level objects, internal object points, and all chords associated with these input object points (if the internal object points are reliable) in our synthesis algorithm. We achieve generation of this feature space [3-4] with the system block diagram in Figure 2. This feature space provides in-plane distortion-invariance. We achieve out-of-plane distortion-invariance by the use of training set data and LDFs. The case studies for which this feature space has been tested included a set of ship data and a set of aircraft imagery [3,4]. The LDFs used were Fisher vectors and dominant Karhuhen-Loeve eigenvectors. Most attractive results ($\simeq$ 95% correct classification) were obtained.

Figure 2: *Optical Chord Transform Feature Space Generation Block Diagram*

2.2 SPACE-VARIANT FEATURE SPACE

An attractive feature space that is in-plane distortion-invariant can be obtained from the Fourier transform of coordinate transformed in-plane data [5]. The resultant system has a different impulse response at each spatial point in the system. The coordinate transform is chosen to make the features invariant to different geometrical distortions. A polar transform results in rotation invariance. If the logarithm of the axes is taken, the features are scale invariant and a Mellin transform results. If we log the radial axis in polar space, scale and rotation invariance are both achieved. To obtain shift-invariance, the object must be centered (by moments, etc.) or the coordinate transform operations can be performed on the magnitude Fourier transform of the input data. Figure 3a shows an optical system to achieve this. The *coordinate transformation* (CT) is performed by a *computer generated hologram* (CGH) at P_2. The output feature space at P_3 can be operated on in parallel by optical LDFs implemented on another CGH. In this case, the class of the input object is determined by the location of a peak in P_4 on a particular detector, or by the binary-encoded output value from a set of N detectors. Figure 3b shows the block diagram of this space-variant processor [6].

As a demonstration of the use of this architecture for M-DOF object identification, we consider a set of 9 different aircraft. These objects have no stable rest position and thus represent an attractive application for an M-DOF processor. Since the feature space at P_3 is in-plane (scale, rotation and translation) invariant, we use a training set to provide out-of-plane invariance (in pitch and roll of the aircraft). A relational graph was devised to identify the class of the aircraft. At the first level of the graph, a decision is made on the sub-class of the object (e.g. commercial, fighter, etc.). A *synthetic discriminant function* (SDF) LDF was used at this node for this decision. At subsequent nodes, the name class (F104, DC10, etc.) of the aircraft is determined. This represents a multi-class graph (with greater than one decision, i.e. one of three choices, made per node). A second binary graph (with one of only two decisions made per node) was then devised using Fisher LDFs. In both graphs, different features (the

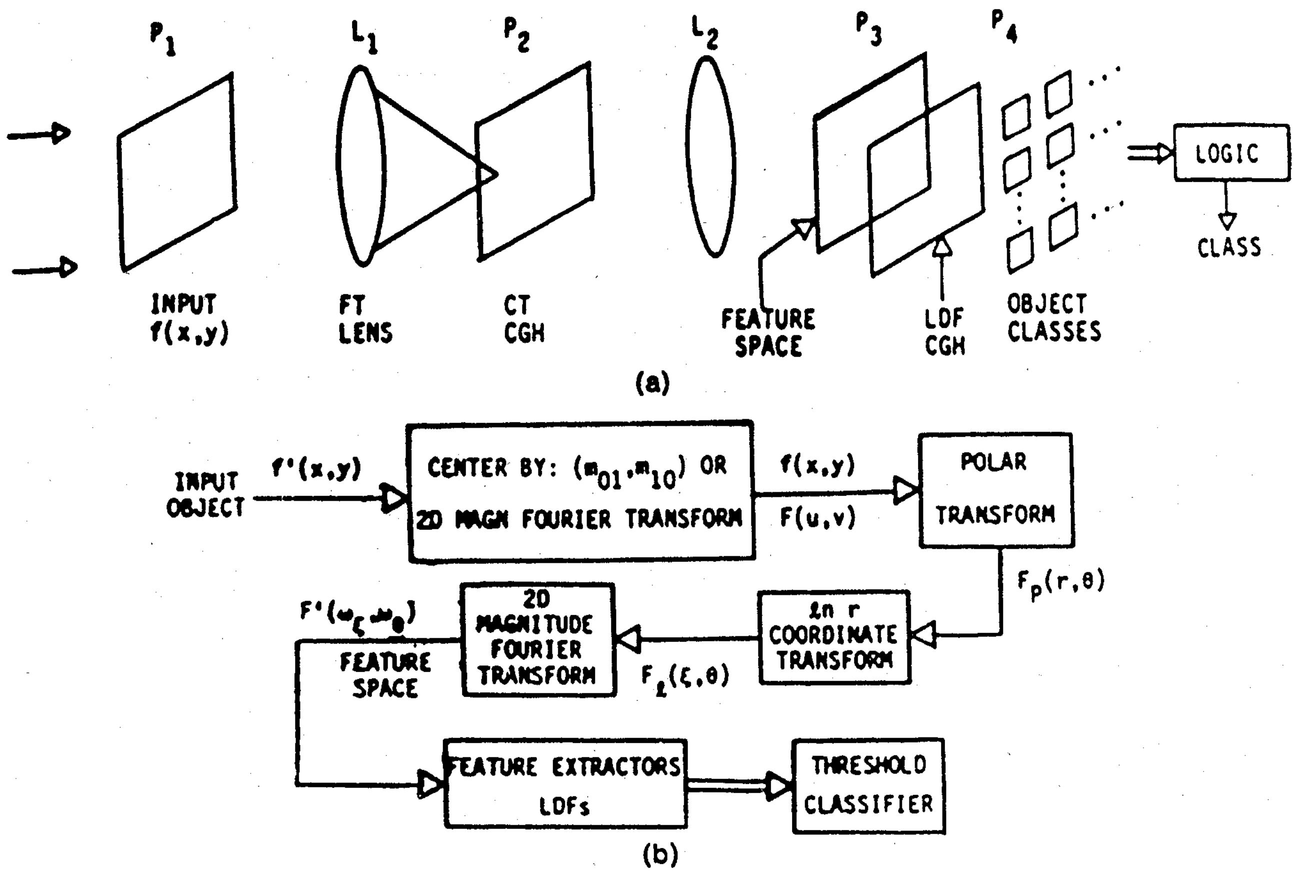

Figure 3: *Optical Space-Variant Feature Space. (a) Optical System; (b) Block Diagram*

optimal ones) are used at different nodes. The training set consists of 5 images per object class (aircraft name class) at 0° and $\pm 20°$ rotations in pitch and roll (recall that the feature space is invariant to yaw, as well as scale differences). The graphs were then tested versus 0°, $\pm 10°$, $\pm 20°$ and $\pm 30°$ in pitch and roll. (These are distortions for which the feature space is not automatically invariant. In other tests on in-plane distortions, all results were positive and thus are not included in these M-DOF tests.) The full test set thus consisted of 13 images for 9 different aircraft (117 images). The results obtained were approximately 99% and 95% correct recognition for the two graphs. This demonstrates the M-DOF performance of this feature space processor.

2.3 MOMENT FEATURE SPACE

A moment-based system (block diagramed in Figure 4) has also achieved M-DOF recognition [7,8]. In this system, the moments are optically generated. The first level classifier uses the ratio μ_{20}/μ_{02} to estimate the aspect of the object and a hierarchical tree to estimate the object class. The results from these first-level estimators are used to access 21 moments for each object class. These are then used in an iterative second-level estimator to confirm the object class, its distortion parameters and the confidence of the estimates. This M-DOF processor has been successfully tested on data bases of pipe parts [7] and ship data [8].

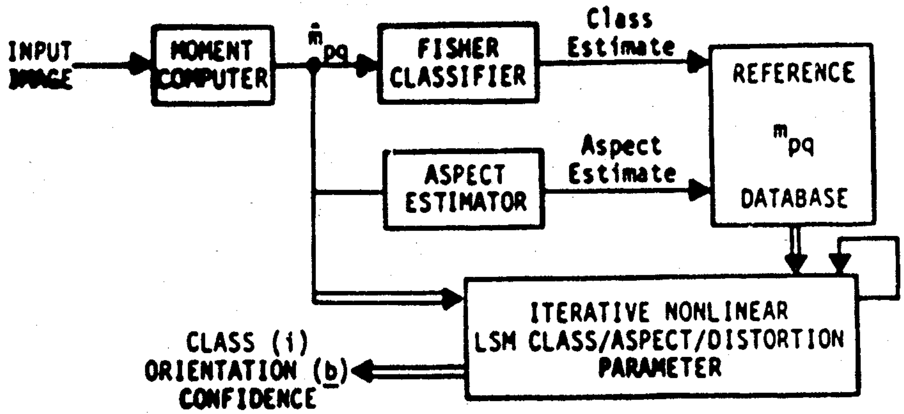

FIGURE 4: *M-DOF Moment-Based Pattern Recognition System Block Diagram*

3. M-DOF OPTICAL CORRELATORS

Optical correlators represent one of the most powerful optical systems. They provide shift-invariant recognition of multiple objects in parallel in the presence of high-clutter. With space-multiplexed filters, one can correlate an input scene versus several filter functions in parallel and either produce multiple output correlation planes or superimposed multiple correlation plane outputs. Frequency-multiplexed filters also enable multiple output correlation planes. One can employ frequency-multiplexed filters at each spatial-multiplexed filter location. With *holographic optical elements* (HOEs) lenses on each filter, various summations of multiple output correlation planes are possible. These architectures are limited in practice by the number of 2-D correlation planes one can read out in parallel and by the lack of distortion-invariance in correlation *matched spatial filters* (MSFs). We now discuss distortion-invariant MSFs, a hierarchical correlator and a symbolic correlator for M-DOF processing.

3.1 DISTORTION-INVARIANT FILTERS

We have devised various techniques to synthesize distortion-invariant correlation filters from training set data [9,10]. These are referred to as SDFs. We can specify the peak value of the correlation output in most of these filters. The three types of filters are: projection filters (these specify only the correlation peak value), output correlation filters (these specify the shape of the correlation function), and *peak to sidelobe ratio* (PSR) filters (these maximize PSR, but cannot control the correlation peak value). These filters have been synthesized to recognize an object independent of its aspect view. Initial tests have been most successful for ATR, ship and aircraft targets.

3.2 HIERARCHICAL CORRELATORS

These distortion-invariant filters allow one filter to recognize an object independent of distortions. They thus significantly reduce the number of filters necessary and hence correlation planes to be analyzed. The use of K multiple filters with binary encoding of the outputs enables K filters to recognize 2^K object classes. Control of the filter peak outputs to L levels allows F filters to handle L^F object classes. Thus, these filters allow large object class problems with a small number of filters and with the other advantages of a correlator. In extensive tests, we find that as the size of the problem to be solved increases, the filter's performance degrades. A proposed solution to this is a hierarchical correlator [11]. In the first level of this system, PSR filters are used to locate *regions of interest* (ROIs) or candidate objects in the scene. Correlation filters are then used in the second level to test each location and the shape of the correlation peak there. In the final level of the hierarchical system, projection filters are used to confirm the object class and to identify it and to determine its orientation.

3.3 SYMBOLIC CORRELATORS

One can view correlation outputs from multiple filters as a symbolic description of the input object. The use of multiple multiplexed filters with symbolic post-processing offers significant potential for M-DOF multiple object pattern recognition in parallel.

4. OPTICAL ARTIFICIAL INTELLIGENCE

Various optical AI processors have recently emerged and have been advanced. Initial remarks on each are now advanced. More extensive tests on all are necessary to more fully assess each. The relational graph processor in Figure 3 is one approach. The automatic organization of data into sub-classes as employed in this processor is a useful technique for any knowledge base or inference system. A model-based description of objects is another approach that is most attractive because of its efficient storage and its ability to easily generate different object aspect views. A reference function generator using this concept is quite general purpose and useful for filter synthesis and generation of the filters for correlators and for the memory matrices in associative processors. Successful initial tests on aircraft data has been most attractive using these approaches. We expect future work to concentrate on optical AI techniques, hopefully with attention to system realization and to more extensive testing.

5. SUMMARY AND CONCLUSION

Figure 5 shows one version of three levels of the hierarchical vision processing system, the scene and object elements involved in each and the type of processing employed at each level. As seen and as was briefly described above, there is a significant role for optics in each level of vision.

ACKNOWLEDGMENTS

The support of future work in this area by a NASA JPL contract is gratefully acknowledged.

REFERENCES

1. D. Casasent, "Hybrid Optical/Digital Image Pattern Recognition: A Review", Proc. SPIE, Vol. 528, pp. 64-82, January 1985.
2. R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, Wiley & Sons, New York, 1973.
3. D. Casasent and W.T. Chang, "Generalized Chord Transformation for

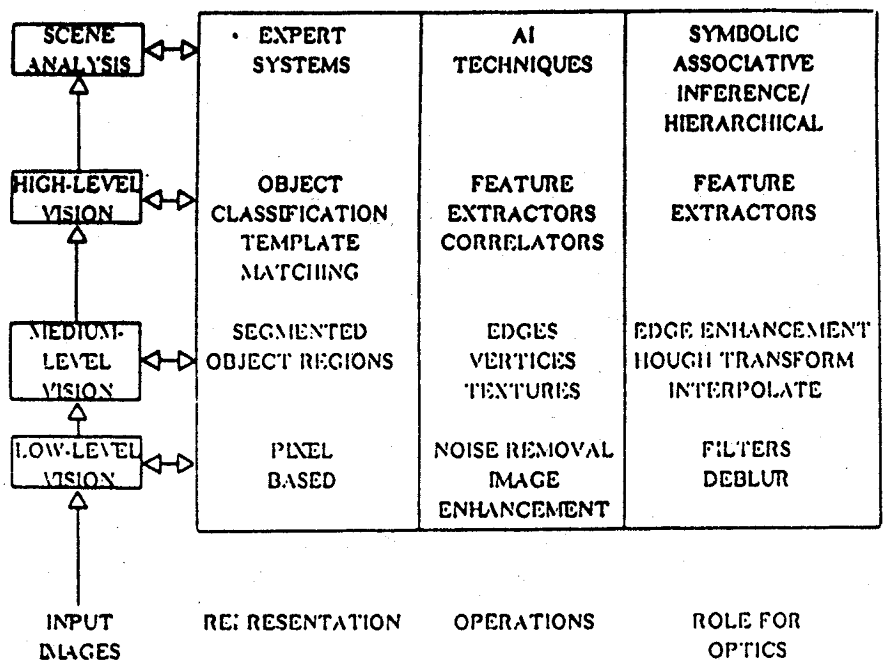

Figure 5: *Hierarchical Levels of Vision, Scene, and Object Recognition and the Role for Optics in Each*

Distortion-Invariant Optical Pattern Recognition", Applied Optics, Vol. 22, pp. 2087-2094, July 1983.

4. D. Casasent and W.T. Chang, "Parameter Estimation and In-Plane Distortion Invariant Chord Processing", Proc. SPIE, Vol. 579, pp. 2-10, September 1985.
5. D. Casasent and D. Psaltis, "Deformation-Invariant, Space-Variant Optical Pattern Recognition", Chapter in Progress in Optics, E. Wolf Ed., Vol. XVI, Holland Pub. Co., New York, 1978, pp. 291-356.
6. D. Casasent and A.J. Lee, "An Optical Relational-Graph Rule-Based Processor for Structural-Attribute Knowledge Bases", Applied Optics, Vol. 15, , pp. 3065-3070, 15 September 1986.
7. D. Casasent and R.L. Cheatham, "Hierarchical Pattern Recognition Using Parallel Feature Extraction", Proc. ASME, *Computers in Engineering 1984*, Vol. 1, pp. 1-6, August 1984.
8. R.L. Cheatham and D. Casasent, "Hierarchical Fisher and Moment-Based Pattern Recognition", Proc. SPIE, Vol. 504, pp. 19-26, August 1984.
9. D. Casasent, "Unified Synthetic Discriminant Function Computational Formulation", Applied Optics, Vol. 23, pp. 1620-1627, May 1984.
10. D. Casasent and W.T. Chang, "Correlation Synthetic Discriminant Functions", Applied Optics, Vol. 25, pp. 2343-2350, 15 July 1986.
11. D. Casasent, "Optical AI Symbolic Correlators: Architecture and Filter Considerations", Proc. SPIE, Vol. 625, pp. 220-225, January 1986.