N93-24552

A STUDY OF VIDEO FRAME RATE ON THE PERCEPTION OF MOVING IMAGERY DETAIL

Richard F. Haines
RECOM Technologies
Ames Research Center - NASA
Moffett Field, California 94035

Sherry L. Chuang
Spacecraft Data Systems Research Branch
Ames Research Center - NASA
Moffett Field, California 94035

Abstract

The rate at which each frame of color moving video imagery is displayed was varied in small steps to determine what is the minimal acceptable frame rate for life scientists viewing white rats within a small enclosure. Two, twenty five second-long scenes (slow and fast animal motions) were evaluated by nine NASA principal investigators and animal care technicians. The mean minimum acceptable frame rate across these subjects was 3.9 fps both for the slow and fast moving animal scenes. The highest single trial frame rate averaged across all subjects for the slow and the fast scene was 6.2 and 4.8, respectively. Further research is called for in which frame rate, image size, and color/gray scale depth are covaried during the same observation period.

Introduction

The perception of moving detail(s) on a computer monitor or TV screen is a complex function of many optical, visual, and cognitive variables; disagreement remains concerning the impact of specific variables. For example Farrell and Booth (1984) reported that decreasing video bandwidth produces relatively little reduction in subjectively determined image acuity for moving objects while Connor and Berrang's (1974) data suggest a linear relationship between increased bandwidth and increased judged image quality. Some investigators feel that this linear relationship results from an improvement in perceptibility due to increasing speed of image motion across the screen. However, given the same amount of bandwidth reduction and speed of image motion, the impairment of image quality is greater for images having many vertically oriented edges of high contrast than for images with only a few such edges. So both the contrast and orientation of the objects are important.

Initially we assumed that those who work with small animals prefer to see smoothly moving images rather than disjointed, choppy motion since smooth motion supports improved image recognition and more correct interpretation of behavioral functions and interactions.

A number of other earlier studies have been performed on the effect of varying frame rate on image usefulness. Ranadive (1979) reported that video bandwidth was directly proportional to the product of *resolution* (height x width; pixels per frame), *frame rate* (fps), and *gray scale* (bits/pixel). When the viewer varied one of these three parameters at a time (while watching his own motions controlling a robot in order to perform a simple task), it was found that he could carry out the assigned task relatively well even though these image parameters were degraded significantly. Performance was defined as the quotient Tt/Td where Tt is the time to accomplish the task using full video (i.e., no degradation) and Td is the time required to accomplish the task using degraded video. He found that when only one of the three parameters was systematically reduced performance remained at acceptable levels until a point was reached where the task could no longer be accomplished at all. He also found that frame rate and gray scale could be degraded by larger amounts than resolution before the critical performance limit was reached. Since the total bits associated with the frame rate parameters in Ranadive's study was only 42 percent of the total bits associated with the other two parameters this suggests that frame rate is a very attractive candidate for reducing video bandwidth under these viewing conditions.

Deghuee (1980) had an operator adjust resolution, frame rate, and gray scale during manual robotic control operations under total bit rate constraints. Dynamically changing these three parameters in real time influenced performance *although lower bit rates did not result in reduced performance*. Since only two bit rates were studied (10 kbps and 20 kbps) it is possible that these total bit rate conditions were not sufficiently small enough and/or sufficiently different from one another to produce significant decrements in performance. Deghuee also reported that the operators did not adjust the three parameters to achieve an image with some "optimal" quality but, rather, set each parameter to achieve some predetermined combination of settings of the three available parameters. Because his operators were sufficiently familiar with the appearance of changes in each of the three parameters separately they were (probably) able to adequately anticipate the appearance of a predetermined combination of them. Deghuee also found that the type of manipulation task undertaken yielded the most significant differences in performance which is what we found when comparing different levels of video compression (Haines and Chuang, 1992).

None of the studies cited above varied frame rate systematically while viewers evaluated the health and behavior of small animals as will be done in future Space Station Freedom experiments. This paper describes a study of the relationship between video frame rate and perceived quality and acceptability to life scientists of moving imagery of white rats. It is another in a continuing series of studies related to remote monitoring between earth orbit and the ground where transmission bandwidth is limited and must be used optimally.

As Haskell and Steele (1981) state, "Only when perception is properly understood will we have accurate objective measures. However, the day when we can, with confidence, objectively evaluate a new impairment without recourse to subjective testing seems very remote." The interested reader should consult (Gonzalez and Wintz, 1987; Watson, 1987; Watson et al., 1983; Wood et al., 1971) for further information on this issue.

Method

Experimental Design and Variables. The experimental design used may be characterized as a 2 x 3 x 2 x 9 parametric design having the following factors:

2 levels of direction of change of frame rate (increasing; decreasing)
3 levels of frame rate change resolution (5, 2, 1 fps)
2 scenes (slow animal motion; fast animal motion)
9 subjects (Ss)

Each subject (S) was presented all twelve cell conditions. Five subjects received scene 1 first while the other four received scene 2 first. Likewise, four subjects received increasing frame rate trials first per pair while the other five received decreasing frame rate trials first. Frame rates from 1.5 to 30 fps were explored.

The method of limits (Woodworth and Schlosberg, 1965) was used to quantify the effect of video frame rate on perceived image quality. This method employs alternating series of

decreasing and increasing frame rates where S indicated the frame rate at which he or she could no longer accept the quality of the moving imagery and then gave a numeric rating of image quality at each frame rate presented. Each series of trials was conducted at progressively smaller frame rate steps: Initial trials varied in five fps steps in order to quickly identify the approximate frame rate separating an acceptable from an unacceptable image. Subsequent trials varied in 2 fps and 1 fps steps. Thus, S was progressively exposed to finer and finer frame rate steps. Means of the 2 fps and 1 fps trials were combined to determine the final threshold frame rate for each subject.
Two separate judgments were made immediately following each 25 second-long scene finished:

(1) Was the scene of acceptable quality to make useful scientific judgments in their own scientific discipline (yes, no)?

(2) What was the image quality? A five point scale of whole numbers was used: (1) = image clarity completely unacceptable relative to 30 fps, (3) = image clarity is of average acceptability relative to 30 fps, and (5) image clarity is completely acceptable relative to 30 fps.

Video Tape Scene Description. The so-called "slow scene" showed two white rats within a small enclosure. Almost all of the scene showed the animals performing typical grooming activity (e.g., licking their fur, scratching with a hind leg at about 6 - 10 Hz, playfully biting each other). Neither animal walked around very much during the scene but exhibited typical slow limb and body movement, exploratory behavior such as sniffing, etc. The so-called "fast scene" showed the same white rats inside the same enclosure but they were engaged in playful behavior such as tumbling, chasing and rolling over each other, and mock fighting during most of the scene. The angular rates of some of their movements were so great that they appeared to be almost at the edge of blurring, viewed at 30 fps.

Procedure. A training and familiarization period was provided where the scene to be evaluated was presented many times (typically five to seven) on an 18" color standard television monitor at 30 fps so that the subject could become very familiar with it. An experimenter discussed the objective of the study and answered questions during this time. The subject was also asked to write down what scene details were of importance and which would be used to evaluate the scene. The objective was to try to ensure that the same scene-judgement criteria would be used throughout the study. This objective was also emphasized verbally prior to data collection.

A decreasing frame rate test run began with a twenty five second-long scene at 30 fps followed by another identical twenty five second-long scene at 25 fps, etc. Judgements were made immediately following each scene presentation. This procedure continued until the subject indicated that the scene details were no longer acceptable to them to make useful scientific judgments in their scientific discipline. This was followed immediately by an ascending series of trials beginning with the smallest frame rate. A ten second-long period of gray screen occurred between each scene presentation during which S looked away from the screen and verbalized his or her ratings and the experimenter changed the conditions for the next trial and recorded S's ratings. Another increasing and decreasing series of trials followed immediately in which frame rate was varied in 2 fps steps. A final series of increasing and decreasing trials then followed in 1 fps steps. The starting fps for the 2 and 1 fps step trials were estimated on the basis of each S's judgments made during the earlier trials.

Subjects. Nine volunteers took place, 5 male (minimum = 38 yrs; maximum = 56 yrs;

mean age = 50) and 4 female (minimum = 28 yrs; maximum = 42 yrs; mean age = 33.5). All possessed 20:20 corrected or uncorrected distance acuity and normal color perception. Two had taken part in previous video compression studies conducted by the authors.

Apparatus. All imagery was presented on a 16" (diagonal) VGA screen of the IBM computer. This PS/2 Model 80-321 computer has 10 megabytes (MB) of RAM and a 320 MB hard disk. The video imaging hardware installed in it consisted of Intel's "ActionMedia II" board set; an Action Media II Capture module attaches to the ActionMedia II Delivery Board as a daughter board. (FN-1) The prerecorded analog video segments (scenes) described above were played on a four-head, Heliquad II Model JR4500 VHS video cassette recorder whose video output was connected to the composite RS170 input connector of the ActionMedia II boards. They were displayed in a small inset video window measured 5.25" (h) by 3.75" (w) on the larger computer monitor and subtended 12.5 degrees horizontally and 9 degrees vertically (of the observer's visual field).

A software application by IBM known as "Person-to-Person" was used in conjunction with the digital imaging hardware. This application runs with OS/2's Presentation Manager and permits live video to be displayed within an on-screen video window in the video conferencing mode. The following video settings were used: Tint = 50%, Saturation = 76%, Brightness = 66%, and Contrast = 50%, View = single, Effects = local, Large View. An on-screen frame rate control was used which allowed a frame rate to be selected between 30 frames per second and 1.5 frames per second.

All video imagery was compressed using a nine bit hardware-based compression technology developed jointly by IBM and Intel Corporation known as Digital Video Interactive (DVI). This compression approach divides each video frame into four by four pixel blocks and allocated one pixel representation. The pixel representation consists of eight bits for luminance and one bit for hue (color) and saturation. This algorithm is used within each frame i.e., no interframe encoding. Because the scenes presented here were repeated, identical twenty five second-long segments, the only perceptually relevant parameter that changed from trial to trial was frame rate.

Results

The results are presented in three sections: I. Mean image acceptance results, II. Highest Frame rate at which image quality was totally unacceptable, and III. Image evaluation criteria used.

I. Mean Image Acceptance Results. Table 1 presents the minimum acceptable frame rate (averaged across all trials per S) for each type of scene. Experience category, age and sex are also given for each S. The raw data are given in Appendix A and B. It can be seen that: (1) these Ss accepted image quality at frame rates between 1.5 to 8.5 fps. Indeed, the three most highly experienced Ss felt that they could obtain all needed information at rates *below* 1.5 fps which was the slowest rate possible from our hardware. (2) the slow versus fast animal scene did not yield a statistically significant difference in acceptable mean minimal frame rate across all Ss. However, four of the Ss did require a faster frame rate for the fast scene of about one fps, (3) when these data were grouped by general level of familiarity and experience with white rats, mean acceptance frame rate was not clearly different either for the slow or the fast scene across these experience levels, and (4) there was no significant difference between the male and female S's mean data.

Table 1

Mean Minimum Image Acceptance Results (fps)
for Each Subject Averaged Across 2 fps and 1 fps Trials

Experience Category (note 1)	Subj. No.	Age	Sex	Slow	Fast
A	7	45	M	<1.5 (2)	<1.5
A	1	55	M	<1.5	<1.5
A	8	56	M	<1.5	<1.5
B	4	28	F	4.9	6.0
B	2	34	F	3.0	4.3
B	5	38	M	3.9	4.9
B	9	42	F	8.5	6.4
B	3	56	M	5.6	3.7
C	6	30	F	5.0	5.1
			Mean =	3.9 (3)	3.9 (3)
			SD =	2.4	2.0

Footnotes:

1. A = 15 or more years of experience; B = 5 - 15 years; C = 0 - 5 years.
2. All values labelled < 1.5 were scored as 1.5.
3. Not statistically significantly different (t test).

II. Highest Frame rate at which image quality was totally unacceptable. This numeric rating provided a second response measure of the subjective usefulness or non-usefulness of low video frame rates. We are mainly concerned with the single highest frame rate that was judged to be of completely unacceptable image clarity. Table 2 and 3 provides these data.

Table 2

Highest Frame Rate Single Trial Judged to Provide a Totally
Unacceptable Image Quality for the *Slow* Scene
(Relative to 30 fps)

Subj. No.	Ascending Trials	Descending Trials
1	5.1	5.5
2	6.4	7.5
3	4.2	3.1
4	*	*
5	5.1	10.2
6	*	*
7	*	*
8	13.5	3.6
9	3.6	5.2
Mean =	6.3	6.0
	Grand Mean = 6.2	

* Indicates that subject's fastest unacceptable frame rate was <1.5.

In addition to the above results it was found that: (a) there were characteristic individual differences in these numeric ratings. Each S gave consistent numeric ratings throughout their viewing period and did not appear to change their judgment criteria. This was shown by the fact that the same numeric score tended to be assigned to the same frame rate over time even though they had viewed different frame rates in the meantime, and (b) the Ss appeared to have understood and followed these rating instructions.

The grand mean data of Table 2 and 3 reinforce the previous Table 1 data with regard to the frame rate - scene motion relationship, viz., the slower scenes required a higher frame rate in order to be judged as acceptable by these Ss.

Table 3

Highest Frame Rate Single Trial Judged to Provide a Totally Unacceptable Image Quality for the *Fast* Scene (Relative to 30 fps)

Subj. No.	Ascending Trials	Descending Trials
1	3.6	6.4
2	4.2	5.5
3	3.1	3.1
4	*	*
5	*	8.5
6	*	*
7	*	*
8	*	*
9	*	*
Mean =	3.6	5.9
	Grand Mean =	4.8

Footnotes:

* Indicates that subject's fastest unacceptable frame rate was <1.5.

III. *Image Evaluation Criteria Used* (Professional Discipline, Experience Level, and Minimal Frame Rate). It was expected that each subject might use a somewhat different set of criteria for evaluating the moving imagery of each scene. Such differences might reflect differences in one's disciplinary training and professional experience. This was found to be the case. In fact, large individual differences were found in the minimum acceptable frame rate people selected during their scene evaluations. Having a lot of prior experience seemed to play an important role in making these judgements, perhaps by improving one's capability to extract subtle image cues or ignoring distracting cues that are present. For example, the three Ss who possessed the most research experience also had prior experience in viewing one (1) fps images of rats in micro-gravity. They judged all scenes at 1.5 fps and higher as being entirely adequate for making their judgments of grooming behavior, general weight and health of the animals, evidence of edema and porphorin (exudate) build-up around the nose, ears and eyes, reaction to allergies, fecal matter build-up around the tail, and leg extension movements. Apparently, their prior experience permitted them to notice these details regardless of how quickly and discontinuously the image shifted across the screen. However, it must be noted that this particular list of image characteristics is made up mostly of static cues. Less experienced subjects generally required higher frame rates to make their judgments. This finding argues in favor of allowing each user to set his or her own frame rate, if possible, to support their own scientific requirements.

Discussion

A minimum frame rate was identified in the present study where experienced subjects judged the quality of image motion (and other details) as being acceptable to them to adequately judge the overall status and behavior of rats. The minimal frame rate averaged across all subjects was approximately four fps for both the slow and fast scene. Minimal acceptable frame rates varied from 1.5 fps to 5.1 fps for both the slow and the fast scene. It is clear from this study that what is an acceptable minimal frame rate is directly related to at least three complex factors: (1) the type of visual discriminations that must be made from the frames, (2) the nature of the moving images to be examined, and (3) the level of experience one has in making the judgments. These visual-cognitive discriminations range from being very general (e.g., is the animal alive?) to highly specific (e.g., is the animal displaying specific signs of allergic reactions or vestibular dysfunction?).

More than one third of all of the judging criteria cited by these Ss were static in nature (e.g., nasal discharge, hair texture, signs of blood, posture). It is possible that the presentation of multiple frames per second actually impeded visual judgments of these specific kinds of image features. Thus, there is probably a class of static image details of importance to the S, a class of dynamic image details of importance to the S, and a third class in which both are relevant in varying degrees. This possibility suggests the need for further experimentation in which various mixes of cues ranging from static only to dynamic only be presented at different frame rates to see if it is possible to identify minimal frame rates within each class of image details.

Visual Integration of Object Motion. The perception of a moving image on a TV screen is actually the result of visually smoothing a series of time sampled (strobed) still image frames into an apparently continuous movement. As individual picture elements (pixels) making up the full frame each change in intensity and color the eye attempts to integrate them and to identify the meaning of this constantly changing array of luminous dots.

Image details may or may not appear to move across the screen depending upon many variables. For instance, the combination of visual angle and duration over which adjacently illuminated pixels appear to change determines whether the image is seen as a strobed (jumping) or continuously (smooth) moving image. Watson et al. (1983) has found that image sampling frequency (Hz) increases almost linearly with an increase in the angular velocity of an image seen on a screen in order to produce smooth motion rather than strobe motion. Images translating at about one degree arc per second must be sampled at about 30 Hz in order to appear to be moving smoothly.

It is interesting to speculate whether minimal acceptable frame rate may be related somehow to the time required for the visual system to extract information from a scene during a single glance. For instance, Senders et al. (1964) reported that the mean visual dwell time (FN-5) for visual informational displays having information bandwidths from 0.05 to 0.48 Hz was 0.4 sec. Interestingly, several other studies of eye fixation dwell time on displays also have shown a mean duration of about 0.4 second across a wide range of display bandwidths (Harris and Christhilf, 1980; Carbonell et al., 1968).

Subject Variables. There is little doubt that the human visual system is remarkably adept at extracting useful information from relatively degraded video imagery. If resolution is degraded, for example, perception probably shifts to lower spatial frequencies which incorporate slightly higher visual contrasts in order to perceive image translation across the scene.

Application of Data to Space Station Freedom Operations. The planned video downlink rate capacity for Space Station Freedom will be variable in the following five steps (Corder, 1992):

60 (full frame) fields per second	41.1 MB/s
30 (1/2 frame) fields per second	20.8 MB/s
15 (1/2 frame) fields per second	10.5 MB/s
7.5 (1/2 frame) fields per second	5.3 MB/s
1.875 (1/2 frame) fields per second	1.5 MB/s

Assuming a full frame video image format of 500 x 400 x 8 bits and 30 fps the required data rate would be 6 MB/s. Even without digital image compression, use of 7.5 fps (which is a higher frame rate than almost all of the present Ss accepted) would reduce the downlink data rate by a factor of 4 relative to the 30 fields per second data rate given here. If the Tracking/Data Relay Satellite's (TDRSS) Ku band maximum downlink rate is 43 Mb/s (5.37 MB/s) then without video compression it would support only one (NTSC) video channel. Clearly, the downlink bandwidth of all channels must be reduced significantly in order to be able to support all of the required control and monitoring functions planned. Reducing frame rate appears to be an acceptable means of accomplishing this objective in some research situations.

Conclusions

We conclude from these findings that video bandwidth may be reduced from SSF to the ground by a factor of more than 4 times the normal 30 fields per second (approx. 4 fps) and still provide an acceptable image to the majority of scientists and animal care personnel. Observer prior experience plays a central role in determining minimal acceptable frame rate. It is not yet clear whether these data can be extrapolated to other life science animal specimens.

References

Carbonell, J.R., J.L. Ward, and J.W. Senders, 1968. A queuing model of visual sampling: Experimental validation. *IEEE Transactions on Man-Machine Systems*, MMS-9, Pp. 82-87.

Connor, D.J., and J.E. Berrang, 1974. Resolution loss in video images. *NTC 74 Record,* (IEEE Publ. 74, CHO 902-7 CSCB, Pp. 54-60), San Diego, Calif.: Institute of Electrical and Electronics Engineers.

Corder, E.L., 1992. Internal video subsystem overview. Presentation given at Payload Data Services Workshop, Huntsville, Alabama, August 5-6.

Deghuee, B.J., 1980. Operator-adjustable frame rate, resolution, and gray scale trade-off in fixed-bandwidth remote manipulator control. Mass. Inst. of Technol., *M.S. Thesis*, (Department of Aeronautics and Astronautics), Boston, Massachusetts.

Farrell, R.J., and J.M. Booth, 1984. *Design handbook for imagery interpretation equipment*. Boeing Aerospace Co, Document D180-19063-1, Seattle, Washington.

Gonzalez, R.C., and P. Wintz, 1987. *Digital Image Processing*. 2nd Ed., Addison - Wesley Publ. Co., Menlo Park, Calif.

Haines, R.F., and S.L. Chuang, 1992. The effects of video compression on acceptability of images for monitoring life sciences' experiments. *NASA Technical Paper* 3239.

Harris, R.L., and D.M. Christhilf, 1980. What do pilots see in displays? *Proceedings of the Human Factors Society, 24th Annual Meeting*, Los Angeles, CA, Human Factors Society, Pp. 22-26.

Haskell, B.G., and R. Steele, 1981. Audio and video bit-rate reduction. *Proceedings of the IEEE*, Vol. 69, No. 2, Pp. 252-26.

Ranadive, V., 1979. Video resolution, frame rate and gray scale tradeoffs under limited bandwidth for undersea teleoperation, Mass. Inst. of Technol., *M.S. Thesis* (Department of Aeronautics and Astronautics), Boston, Massachusetts.

Senders, J.W., J.E. Elkind, M.C. Grignetti, and R.P. Smallwood, 1964. An investigation of the visual sampling behavior of human ovservers. *NASA CR-434*, Bolt, Beranek & Newman, Cambridge, Mass.

Watson, A.B., 1987. Efficiency of a model human image code. *J. Optical Society of America*, Series A, Vol. 4, No. 12, Pp. 2401-2417.

Watson, A.B., A. Ahumada, Jr., and J.E. Farrell, 1983. The window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. NASA Technical Paper 2211.

Wood, C.B.B., J.R. Sanders, and D.T. Wright, 1971. Image unsteadiness in 16mm film for television. *Journal of the Society of Motion Picture and Television Engineers*, Vol. 80, Pp. 812-818.

Woodworth, R.S., and H. Schlosberg, 1965. *Experimental Psychology*. Holt, Reinhart and Winston, Inc., New York.

Footnotes

1. One of the present test subjects served as an investigator on the SL-3 project and had a great deal of experience viewing 1 fps scenes.

2. ActionMedia II boards digitize and compress a video signal for display on a monitor

and/or storage on a hard disk. The boards used here employed a dual-chip, B-series i750 Video Display Processor.

3. Integration here refers to performing content associations and storing this information in visual memory.

4. The Nyquist theorem states that it is necessary and sufficient to visually sample signal at two times its bandwidth.

5. Visual dwell time refers to the duration over which no eye movement occurs.