MENTATION PAGE

Form Approved
OMB No. 0704-0188

2

...s estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, ...ing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this ...ing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson ...d to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (*Leave blank*)

2. REPORT DATE
28 Feb92

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE

UAV Imagery Frame Rate and Resolution Requirements Study

5. FUNDING NUMBERS

C-N00019-90-D-0056

6. AUTHOR(S)

Merryanna L. Swartz, Ph.D.; Daniel F. Wallace, Ph.D.;
John M. Libert; Sharon Tkacz, Ph.D; Daniel Soloman

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Advanced Technology Department
Vitro Corporation
14000 Georgia Avenue
Silver Spring, MD 20906-2972

8. PERFORMING ORGANIZATION REPORT NUMBER

Job 02962-DO-6

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Unmanned Aerial Vehicle (UAV)
Joint Projects Office (JPO)
Washington, D.C. 20361-1014

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

A Human Factors performance assessment of various tasks under limited bandwidth video transmission. Conducted with assistence from the Navel Air Test Center, and Cambridge Research Associates.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION IS UNLIMITED

12b. DISTRIBUTION CODE

11908

13. ABSTRACT (*Maximum 200 words*)

Unmanned Aerial Vehicles (UAVs) are used to conduct a variety of reconnaissance and surveillance missions with human operators interpreting the transmitted imagery at ground stations. Current UAV datalink designs require a limited capacity which are expected to result in some cost to the operator. Two common techniques by which video data rates can be reduced exist, data compression and simple data reduction such as frame rate and resolution reduction. This study utilized real mission imagery to assess the influence of experience and training, compression, frame rate reduction, and spatial resolution reduction on operator detection, recognition and tracking performance for military targets. The comparisons between the control group and other conditions under both frame rate and resolution manipulations demonstrated how data reduction techniques adversely influence operator performance. Additional analyses demonstrated that for some tasks, frame rate has a greater influence on some performance measures, and resolution has a greater influence on other performance measures. In addition to the real imagery study, an analysis of operator performance under simulated tracking conditions was also performed.

14. SUBJECT TERMS

Human Factors; Unmanned Aerial Vehicles; UAV; video imagery; limited capacity datalink; video data compression; frame rate and resolution reduction; target detection, recognition and tracking performance; simulation; JTIDS; JPEG.

15. NUMBER OF PAGES
122

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Standard Form 298 (Rev. 2-89)

93 1 28 018

Vitro CORPORATION

14000 Georgia Avenue
Silver Spring, MD 20609

Final Report

UAV Imagery Frame Rate and Resolution Requirements Study

Merryanna Swartz, Ph.D.
Daniel Wallace, Ph.D.
John M. Libert
Sharon Tkacz, Ph.D.
Daniel Solomon

Advanced Technology Department
Vitro Corporation

February 28, 1992

UAV JPO contract N00019-90-D-0056
Job 02962, Delivery Order Six

Executive Summary

Study Objective

The Unmanned Aerial Vehicle (UAV) Joint Program Office (JPO) is developing a UAV system with which to conduct various reconnaissance and surveillance missions. A central component of this system is the data link that will download imagery data to the Mission Payload Operator. UAVs provide operational commanders with real-time video of opposing forces, terrain factors, and own-force disposition. Bandwidth reduction in a digital data link can guard against jamming and provide data link security; however, this reduction may result in an impact to the human operator. The objective of this research task is to determine the degree to which data volume can be reduced in terms of frame rate, spatial and grey-scale resolution, while retaining sufficient information to support human performance of mission tasks. This final report presents the results of these investigations.

Background

Two common techniques by which video data rates can be reduced exist: data compression and data truncation. The application of both methods may result in sufficient data reduction that existing digital data links with low to moderate data rates will be suitable for the downlinking of video imagery.

Data compression processes the video data into a more efficient form. All or most of the information is retained and may be recovered for use by applying the inverse of the compression process. In practice some information is lost due to inefficiencies in the compression and decompression processes.

Data truncation cuts out and discards some data to reduce the overall data rate. Truncated data is permanently lost and cannot be recovered. Data truncation includes techniques such as frame rate and resolution reduction.

Conventional video is transmitted at 25 to 30 frames per second. The result of reducing frame rate is that the operator is presented with only a subset of the frames sampled by the sensor. The human performance research literature reviewed supports the use of frame rates at 1.88 - 2 frames per second (fps) for static operator tasks (target detection and recognition) and 3.75 - 4 frames per second for more dynamic tasks (target tracking and designation). Resolution can be reduced across the total display or for the number of TV lines across some target dimension that are needed to resolve the target. Baseline values that support human performance for each type of resolution reduction method were identified in the literature and examined in the experiments conducted as part of this effort. None of the TV line resolution values were tested in designation or tracking tasks in the empirical research reviewed. Additionally, the studies reviewed were carried out with non-mission imagery using experimentally derived target scenarios. The work documented in this report contributes to the human performance literature by using mission-realistic scenarios and by evaluating operator performance with identified baseline levels of resolution (2, 8, and 12 TV lines) derived from manipulating mission parameters (sensor altitude, field of view, look-down angle) in target designation and tracking tasks and frame rate (2, 4, and 7.5 fps).

Subsequent to an extensive survey of related literature, experiments were conducted to determine the minimum video presentation requirements such that the operator could still perform the necessary tasks required by the mission. Four basic operator tasks were identified from the literature search: detection, recognition/identification, tracking and designation.

The experiments conducted evaluated the effect of minimum frame rate and resolution values on operator performance. The values chosen were identified from the literature. Data compression was implemented using a Joint Photographic Experts Group (JPEG) compression algorithm operating at a 50:1 compression ratio. In Experiment One, actual UAV imagery was obtained and used in order to evaluate operator performance in realistic mission scenarios. Pioneer mission footage was used to create two sets of simulations in which the effects of frame rate, resolution, and compression were evaluated. In Experiment Two, a pilot study of two dynamic tasks (designation and tracking) was conducted at the Joint Development Facility (JDF) in collaboration with Cambridge Research Associates, Inc., McLean, VA. The goal was to identify those minimum values that would support adequate performance in an interactive scenario with the operator in the control loop. Participants for the experiments consisted of VC-6 personnel and Vitro personnel with previous military experience in target acquisition tasks.

Results

Experiment One:

Frame rate was found to be a much more critical variable than spatial resolution. In both Experiment One studies, faster frame rates (4 fps) are associated with faster reaction times, higher confidence, and faster confidence ratings. The effect of frame rate on error performance, however, is less consistent and less easily interpreted. Higher frame rates resulted in a decreased number of time-outs (inability to complete the task in the allotted time) and a decreased error rate for designation tasks. An increased error rate in recognition was observed that was counter-intuitive and could not be explained in the context of the experiment.

Spatial resolution had no measurable effect on reaction times or confidence measures for any task. The only dependent variable affected by resolution across all three tasks was image quality rating. Resolution had a marginal effect on error rates for the recognition task. Experience was found to affect an operators' confidence in decision making. Experience also resulted in fewer time-outs which indicates better decision making ability. Thus experienced personnel were able to complete tasks more often and felt more confident about their performance capability. Further studies are appropriate to more completely evaluate the influence of experience on performance.

Experiment Two:

Consistent with Experiment One, frame rate was again found to produce more of an effect on performance than spatial resolution. A similar pattern was observed with higher frame rates associated with faster acquisition, faster designation time, smaller designation error, and smaller tracking error. In many tasks, no difference was observed between 4 and 7.5 fps which validates previous human performance results in RPV programs. A rate of 4 fps was sufficient to produce acceptable operator performance in both dynamic tasks.

Spatial resolution also had some effect on operator performance in Experiment Two, but these results were again not as consistent as was the frame rate effect. Whereas frame rate affects performance overall, spatial resolution affects only specific tasks. For example, the learning rate for the task improves only at the lowest resolution (2 lines). Designation time is faster at the highest resolution (12 lines), but resolution had no effect on designation error. Completion rate, the percentage of completed trials (which is analogous to time-outs in Experiment One) is better at lower resolution, presumably since the target was was always visible on the display.

The frame rate and spatial resolution interactions are of particular interest to the trade-offs considered. If higher resolution is needed for a task, then either 4 or 7.5 fps can be used and similar operator performance can be expected. Since 7.5 fps is not supported by the JTIDS data link at a 50:1 compression ratio, and since performance is the same at 4 or 7.5 fps, it is recommended that values of 4 fps and 8 or 12 lines across the target be adopted for tasks that require designation speed and acquisition accuracy if JTIDS is selected as the UAV data link. It is noted that further investigation of resolution values around 8 lines is needed to clarify some of the inconsistencies found. Examining human performance in similar tasks with 6, 8, and 10 TV lines of resolution should clarify any ambiguity. The interaction effects of frame rate and spatial resolution on percentage of trials completed suggests that operators need higher frame rates (4 fps) if higher resolution (12 lines) is available. A 2 fps/12 lines combination is to be avoided. As noted, the best completion rate performance was at 2 lines of resolution across the target. This is presumably because the target was sometimes lost from the display at higher resolution resolution levels.

Conclusions

The control group in Experiment One served to define operator performance under normal conditions. They obtained a 90% performance level for the three tasks evaluated (detection, recognition, designation) with reaction times ranging from 3.5 to 4.6 seconds. This performance criterion meets those suggested in the literature. However, none of the bandwidth trade-off conditions met this performance requirement. The best performance was observed in the groups that had 4 fps. As the Bandwidth Trade-Off Table shows (Section 7 of this report), the 4 frames, full resolution condition is not compatible with the JTIDS data rate at a 50:1 compression ratio. Performance comparisons between full and half resolution with 4 fps did not differ in ways that would affect operational performance. The Summary Human Performance Table - Experiment One (Section 7 of this report) shows performance levels of 70 to 78% with reaction times ranging from 3.9 to 5.9 secs with 4 fps at half resolution (full display). This could be used as a starting point for defining the digital data link requirements.

The 8 lines/4 fps and 12 lines/4 fps conditions for Experiment Two , shown in the Summary Human Performance Table - Experiment Two (Section 7 of this report) have similar performance in tracking and designation tasks. The results suggest that the lower resolution level of 2 lines can help operators re-acquire a target that moves off the display. No real differences are noticed at 4 fps with 2, 8, or 12 lines for designation task performance. In tracking tasks, the 8 lines/4 fps demonstrated the best performance. The results observed at 8 lines were less consistent than other data analyzed, and may be an appropriate subject for additional study.

Recommendations

A frame rate of 4 frames per second is sufficient to support the operator tasks of detection, recognition, designation, and tracking for the various UAV missions.

The adequate performance observed at half resolution across the total display suggests that reduced resolution does not effect performance markedly. A recommendation is made, however, for display tools to enhance operator performance and increase target detection sensitivity and recognition capabilities. These tools, such as those noted below, can enhance situation awareness in detection, recognition, designation and tracking tasks.

a. changeable FOV
b. selectable compression ratios
c. selectable frame rates
d. windowing at different resolutions or compression ratios

Performance can be enhanced by providing training for different data presentation trade-off combinations. It was found that practice is also beneficial in improving joystick control technique.

Reliable operator performance levels can be maintained up to a 50:1 compression ratio when using the JPEG DCT algorithm. Higher compression ratios may obtainable for video imagery using the Motion Pictures Experts Group (MPEG) compression standard, possibly up to a 200:1 ratio for UAV video imagery. This level of compression could yield an 8Hz frame rate for JTIDS type data rates. An MPEG type of compression algorithm was not available during the experiments. but its suitability for UAV video imagery should be investigated.

Further investigations of the dynamic tasks are recommended in order to identify more precise performance recommendations. While these research results provide preliminary data link design requirements. more information is needed to clarify performance around 8 lines of resolution (e. g., examine 6, 8, and 10 lines). Additionally, the method of joystick control influences performance in dynamic tasks and should be examined further. Finally, comparisons of different ratios and different algorithms in combination with different frame rate and resolution trade-offs can provide further insights into compression effects on human performance.

Acknowledgements. We would like to thank several key people who helped us implement the simulation and carry out the experiments. Chris Norloff and Adele Pierce, Naval Air Test Center, were kind enough to completely organize and set up our military test site. Cambridge Research Associates and the Joint Development Facility were used to carry out Experiment Two. Special thanks are given to Dan Brockway for creating the dynamic simulation software and to Mike Vertal for implementing the continuous control joystick. Finally, the VC-6 and NATC personnel, especially LCDR McDonald, is to be thanked for providing mission footage and participating in the experiments.

Table of Contents

1. Introduction

This report supports early planning leading to the specification and design of Unmanned Aerial Vehicle (UAV) data link and mission planning subsystems. The report covers a number of topics impacting the study of bandwidth reduction/compression options for the UAV in terms of operator performance.

a. Section 1 introduces the problem, objective, and technical approach of the study.

b. Section 2 briefly describes the mission requirements under consideration and the analysis of the missions in terms of operator tasks. These task requirements are the basis for examining human performance in the experiments conducted.

c. Sections 3 and 4 provide technical background on the bandwidth compression problem, data link requirements, and image analysis methodologies applicable to image interpretation and visual performance.

d. Section 5 discusses the results of previous human performance studies and assessments of the effects of bandwidth reduction on operator performance. This review identifies minimal values for frame rate and resolution reduction as a baseline in the experiments.

e. Section 6 describes the human factors experiments conducted and presents the results.

f. Section 7 summarizes results and presents specific design recommendations.

g. Appendices contain supplementary information and lists of references and acronyms used in this report.

1.1 Statement of the Problem

A primary use of the UAV is to provide operational forces with real-time imagery of opposing forces, terrain factors, targets and own-force disposition. Imagery may be collected with various devices including TV, Charge Coupled Device (CCD) cameras, Forward Looking Infrared (FLIR) devices, Infrared Line Scanners (IRLSs), and Synthetic Aperture Radar (SAR). The imagery is transmitted to a surface station via a data link system. For battlefield applications, digital data links are often considered more secure than analog links. In order to transmit the imagery at the full frame rate (e.g., 30 frames/sec) and at 6-to-8-bit grey scale resolution, the data rate must be on the order of 45 - 70 Megabits/sec. Currently, data links capable of meeting such requirements are too costly for the comparatively low cost UAV systems. Consequently, it is desirable to determine the extent to which sensor information density, hence bandwidth, can be reduced while maintaining the human operators' performance at high levels for specified UAV missions. This information serves both as a design guideline for the UAV systems and as a guide for mission task requirements.

1.2 Objective

The objective of this study is to determine the degree to which data volume can be reduced in terms of frame rate, spatial and grey-scale resolution, while retaining sufficient information for the Mission Payload Operator to perform mission tasks.

1.3 Technical Approach

Image processing technology was combined with human factors experimentation to design simulations of realistic tasks required of the Mission Payload Operator in performing the UAV missions. These simulations allowed us to assess human performance in terms of the operator tasks being performed with different combinations of frame rate, resolution, and compression.

An extensive body of experimental literature was surveyed and analyzed concerning human performance in imagery related tasks with various bandwidth reduction techniques. In parallel, image processing literature on bandwidth compression and reduction techniques was evaluated. As a result of these analyses, baseline frame rates and resolution values that support human performance in target acquisition tasks were identified. Similarly, a compression algorithm was identified that could compress video imagery at higher ratios in order to meet narrow bandwidth limits (119 kilobits per second). Two human factors experiments were then designed and conducted with military personnel to reassess previously identified performance results with 1) real mission imagery and 2) higher compression ratios. Two types of experiments were conducted: one that examined performance in detection, recognition, and designation tasks with real UAV mission imagery (called static tasks as none of the system parameters could be manipulated), and another that investigated the dynamic tasks (designation and tracking) in a simulated mission scenario with sensor flight parameter manipulations that resulted in specific ground resolved distances for targets.

In order to interpret the effects of different bandwidth reduction techniques on operator performance, a quantitative image analysis was also conducted on the imagery used in the experiments. This quantitative measure of imagery quality served as a baseline for defining the intelligibility of the imagery that was viewed by operators, and for making informed data link design recommendations.

2. Mission Functions and Operator Tasks

Mission definitions were taken from references from the UAV Joint Program Office (JPO) and Project Group 35 [1], [2]. The mission definitions provided are further characterized in terms of expected UAV operational parameters such as altitude, speed, etc. (or ranges of parameters), to specify the context of the imagery in relation to the operator performance analysis. Based on review of these mission definitions it was decided that, for missions utilizing imaging payloads, there are three basic mission functions that must be performed:

a. Reconnaissance, Surveillance and Target Acquisition
b. Gunfire/Artillery Spotting
c. Bomb/Battle Damage Assessment

Each UAV mission function has associated with it tasks that the Mission Payload Operator must perform in order to achieve mission objectives. Inasmuch as the present study concerns the evaluation of human performance, these missions were analyzed to determine the primary operator tasks necessary to perform the mission. As a result of this analysis, it was determined that the primary operator tasks of interest to the study were:

a. Detection
b. Identification or recognition
c. Designation
d. Tracking

Such tasks will be specified in the context of each mission function and under the conditions, such as UAV flight profile, for each mission discussed below.

2.1 Primary Missions

2.1.1 Reconnaissance, Surveillance and Target Acquisition (RSTA)

As a highly mobile, cued sensor, the UAV system will complement manned aircraft in performing RSTA missions in high-risk areas. Recent experience in Operation Desert Storm demonstrated the value of the UAV for such roles. For the purpose of this study, RSTA includes those activities that lead to targeting, including the detection, localization, identification, and classification of sea and shore targets. An electro-optic/infrared (EO/IR) sensor is used as the UAV payload in the RSTA missions. The UAV system provides imagery to the operator for the performance of the necessary operator tasks. The present study deals only with so-called framing camera imagery such as that from EO/IR sensors.

RSTA Operator Tasks. In RSTA missions the operator will view UAV imagery to detect, localize, identify (recognize), and classify targets. In some cases the operator may be required to slew the sensor to gain a different viewing angle. Limited target tracking may be required in order to keep a moving target in the field of view (FOV) during classification or identification. Tracking in this case need only involve keeping the sensor pointed roughly at the target area rather than the more difficult task of keeping the target positioned under a cross-hair.

UAV Mission Parameters. During the search mode of the typical RSTA mission, it is expected that the UAV will fly at an altitude of 1000 to 3000 meters, at an air speed of 90 knots. The sensor package will maintain a look down angle of 35° and a FOV of 30° horizontal -x- 40° vertical. Once a potential target is detected, the sensor may be slewed to position the target approximately at the center of the FOV and a longer focal length lens may be switched into position for localization, identification, and classification. The FOV in this case will be around 3°-x-4°. If necessary, the UAV may drop to an altitude between 100 - 1000 meters for target identification.

2.1.2 Gunfire/Artillery Spotting

The objective of the gunfire/artillery spotting mission is to detect, localize, and identify targets for naval guns and field artillery and to provide adjustments to the fall of shot for land and sea targets. For the at-sea mission, the UAV transits to the mission area and commences an imagery search along with Electronic Support Measures (ESM), when available. For the land mission, the UAV transits to the designated geographic position and provides imagery and navigation data to locate the desired target(s). The accuracy of the navigation data, when combined with other UAV capabilities, will enable the first fall of shot to be within the FOV of the modular mission payload. If available, ESM data may also be used to confirm the target location and identification. Adjustments in the fall of shot are made relative to the designated target by measurement enabled through the UAV control equipment. Onboard recording of the imagery data by the UAV may be helpful in the post-mission reconstruction, but plays no part in the real-time gun fire adjustment.

Gunfire/Artillery Spotting Operator Tasks. Initially the operator will perform tasks similar to those described above under RSTA. The principal additional task will be the measurement of the exact position of the shot fall relative to that of the target. The UAV system may be equipped with a semi-automated coordinate designation or shot correction system. In this case, the specification of targeting correction could involve the operator positioning a cross-hair or touching the screen in order to specify target position and shot fall position. The designation time need only be long enough for the system to register the appropriate screen coordinate. In this case, adjustment distances and direction would be computed by the system automatically. In the case that UAV navigation information is not sufficiently accurate for automated computation of correction information, the operator will have to specify approximate distance and directional information in much the same manner as would a conventionally deployed artillery spotter. Additional dwell time might be required to allow for estimates or measurements to be made from screen display.

UAV Mission Parameters. During the typical gunfire/artillery spotting mission, it is expected that the UAV will fly at a nominal altitude of 1000 - 3000 meters at an air speed of 90 knots. While in search mode, the sensor package will maintain a look angle of elevation 35° and a wide-angle FOV of 30° x 40°. Once a target is detected, the sensor will be slewed to position the target approximately at the center of the FOV and a longer focal length lens, FOV = 3° will be switched into position for localization, identification, and classification. The wide angle view will be selected for actual spotting such that both target and shot fall positions may be viewed simultaneously. The UAV will transition into an orbit mode such that the target remains in the sensor FOV.

2.1.3 Bomb/Battle Damage Assessment (BDA)

The objective of the BDA mission is to detect, locate and identify the extent of damage to ships or shore targets. The BDA mission will be performed in a mission area that is out of the line of sight of the weapons system and its sensors. If endurance permits, the UAV system could perform both pre-strike and post-strike support. The UAV system provides imagery to confirm the extent of damage on the desired target(s). The UAV system provides high resolution images of the desired targets. Onboard recording of the imagery data by the UAV may be necessary in order to allow for autonomous missions outside data link line of sight.

BDA Operator Tasks. The operator will have to localize the target in the sensor FOV by controlling the sensor position (slewing) and zoom state so as to locate and determine/measure damage to the target. This task will overlap with the RSTA and the Gunfire/Artillery Spotting missions to the extent that these also involve search and recognition functions. Ground resolution demands may be somewhat greater for the BDA mission, however. Damage may be subtle for certain targets and the operator may need to make estimates, for example, of the size of a hole in a hull and its distance above the waterline or may need to determine damage to operational parts of a tank, or other vehicle, or artillery piece. The UAV may need to orbit a target to obtain views from several different aspects in order for BDA mission to be accomplished. In this case, the sensor will have to be slewed to keep the target in the FOV.

UAV Mission Parameters. During the typical BDA mission it is expected that the UAV will fly at an altitude of 100 - 3000 meters and at an air speed of 90 knots in search mode with a FOV of 30° x 40°. Once the target is detected, the sensor will be slewed to position the target approximately at the center of the FOV and a longer focal length lens will be switched into position for BDA. The UAV will be maneuvered to a lower altitude, e.g. 100 - 1000 meters, so as to view the target from a lower angle to assess damage. The mission profile for this mission may simply be an extension of the RSTA or Gunfire/Artillery Spotting missions; in which case, the UAV will already be in position.

3. UAV Data Link Requirements

A single frame of a typical video system consists of, for example, 512 x 512 pixels. Within the dynamic range of most sensors, each pixel would be associated with a monochrome intensity value represented by a 6 to 8 bit word. That is, typical sensors and display systems can easily record 2^6 (64) to 2^8 (256) levels of grey. Typical Charge Coupled Device (CCD) cameras can produce 512 x 512 pixel frames at 30 frames per second (fps). A single frame quantized at a 6 bit grey-scale resolution would amount to 512 x 512 x 6 = 1,572,864 bits. Thus, approximately 1.6 Megabits (MBits) would be required to transmit the single image at maximum fidelity. Framing cameras generate such imagery at up to 30 fps. To transmit standard framing camera imagery at this resolution and frame rate requires a data rate of 1.6 Mbit/frame x 30 fps, or approximately 50 Mbit/second.

From this example, it becomes obvious that the bandwidth requirement could be reduced in one of three ways. One can reduce the size of the image being transmitted, i.e., reduce the number of pixels; reduce the number of bits with which to represent the intensity information, i.e., reduce the number of quantization levels; or one can reduce the rate at which frames are transmitted per unit time.

No bandwidth specification was provided *a priori* to be considered as a goal for the current study. However, recent tri-service efforts to conform to the Joint Tactical Information Distribution System (JTIDS) offer guidance in recommending bandwidth limitations. In contrast to a data rate of 50 Mbit described above, the current JTIDS would allow data rates on the order of 100 - 200 Kbits/sec. Significant data reduction/compression techniques would be required to fit the UAV imagery within this standard.

3.1 Bandwidth Reduction

Bandwidth reduction in a digital data link can guard against jamming and enhance communications security. This reduction, however, may result in a performance cost to the human operator. Two common techniques by which video data rates can be reduced exist: data compression and bandwidth reduction. Simple data bandwidth reduction methods involve manipulating frame rate and spatial resolution. Data compression involves the reduction of bits of picture elements (pixels). These methods are discussed briefly below.

Spatial Resolution Reduction. An image can be physically reduced by mapping multiple pixels in the input image to a single pixel in a smaller output image. For example, a 512 x 512 pixel image may be reduced into a 256 x 256 pixel image by replacing a block of four pixels with a single pixel representing the average value of four neighboring pixels in the original image. The resultant pixel value may be represented at any desired level of precision. The spatial resolution of the resultant image is reduced, however.

As spatial resolution is reduced, the maximum ranges for detecting targets by a given sensor will be reduced by the same factor. That is, combining two pixels each in the horizontal and vertical image dimensions merely increases the dimension of the ground resolution cell by the same factor in each dimension [16]. Thus, resolution reduction is unlikely to yield a net gain over simply clipping a smaller section of the frame for transmission and display at maximum resolution. i.e., image truncation.

Frame Rate Reduction. A significant reduction in bandwidth can be achieved by reducing the number of image frames transmitted in a unit of time, i.e., the frame rate. As indicated above, at 30 fps, conventional video or FLIR sensors generate data at an adequate rate for supporting human performance in target acquisition tasks. In order to preserve temporal integrity of the sensor system, i.e., the temporal correspondence between the scene sampling and the display, frame rate reduction must be accomplished by discarding image frames at the point of acquisition. From the perspective of saving computational time, it is wisest to discard frames prior to other processing, e.g., image compression. Regardless of where in the process the frames are discarded, the result is that the operator is presented with only a subset of the frames sampled by the sensor. The reduction of frame rate, however, generally implies that the frames be displayed on the video monitor at the full cycling rate of the cathode ray tube (CRT). That is, even though only one frame per second is presented to the observer, it is important that the frame be refreshed on the screen at the full 30 fps. Otherwise, objectionable flicker of the screen will result. This means that each frame must be buffered or stored and displayed repeatedly during the interframe interval.

Image Compression. Image compression techniques reduce the number of bits required to represent the image. Image compression methods are based on the premise that much image information is redundant or otherwise expendable. Thus, some compression methods reduce redundancy by transforming the original image to a more compact mathematical expression. Other methods discard image information that is beyond or near the limits of human visual perception and, therefore, is not missed in the decompressed image. Some use a combination of techniques. A survey of image compression techniques may be found in [3] and image processing in general in [4].

Many compression methods reduce the image size by reducing the length of the computer words used to represent the quantized level of energy intensity associated with each pixel location in the image. As indicated above, video imagery displayed on standard video cathode ray tube monitors is encoded in a 6-bit word. Six bits permits designation of 64 intensity levels by numbers from 0 to 63. Significant compression can be achieved by representing the intensity values in fewer than 6 bits. Below we discuss available compression standards and the choice of an algorithm to compress the experimental imagery used in Experiment One.

3.2 JPEG Standard Compression

"JPEG" stands for Joint Photographic Experts Group, a committee that has been involved in proposing a standard for compressing high-quality still images. The JPEG standardization activity in the U.S. is coordinated by the American National Standards Institute (ANSI) and internationally by the ISO (International Standards Organization). It is only one of several ongoing standardization activities attempting to impose some interim order onto an extremely dynamic field of endeavor. The Motion Picture Experts Group (MPEG) is developing another standard for full-motion, color video. While MPEG looks promising for compressing full-motion video, it was not available for testing at the time the experiments in this study were designed and conducted.

The main reason for emphasizing the JPEG/ISO compression standard for the present study is that it is widely recognized and used, is well documented and available commercially in both hardware and software implementations, and is an internationally recognized standard for still image compression. It is not optimal for motion video compression, but the methods employed are illustrative of compression techniques in general and thus useful for assessing human performance.

JPEG standards consist of a group of compression techniques that can be selected and used in various ways to achieve varying levels of compression depending upon the particular application. Three such techniques include: 1) a hybrid discrete cosine transform, 2) Huffman coding, and 3) differential pulse code modulation. For a more detailed, but still relatively high level, discussion of JPEG compression, the reader is referred to [5]. For the purposes of this study, we chose the discrete cosine transform (DCT) as the compression technique. The DCT has been well supported in the human performance literature and is a JPEG standard compression technique. In addition, this standard is widely recognized and readily available in hardware and software implementations.

There is little comparative data available to support selection of the "best" image compression technique for the UAV. Most of the literature reported in connection with development of the Remotely Piloted Vehicle (RPV) uses some variation on the JPEG methods described above. Human performance results from these studies (see Section 5) are reported in the context of compression ratios or in terms of bits per pixel. Previously studied ratios were no greater than 30:1, however. Performance effects due to peculiarities of the particular compression implementation are not well studied, nor are dependencies of results on the imagery used. The DCT hybrid technique is reasonably well supported in the performance literature, and is currently available in hardware for use in the present study. The DCT [6] was used to compress the experimental imagery at a 50:1 ratio. This ratio has yet to be tested with human operators and will help us assess the feasibility of restricted bandwidth limits.

4. Compressed Imagery Analysis and Visual Performance

One would like to achieve the maximum compression possible while satisfying the constraint that "suitable" fidelity be maintained. The fidelity requirement depends upon the application and the characteristics of the input image.

Quantitative fidelity measures for the compressed imagery are available. For example. the root-mean-square error between corresponding pixels of input and output images provides a good objective measure by which to evaluate fidelity. Military handbooks [7] and research methodology [8], [9] for other quantitative image fidelity measures are also useful for objectively measuring image quality. Given a set of targets and mission parameters (e.g., altitude. sensor look-down angle, field of view, resolution. speed, range to target) for the TV imaging sensor. it is possible to calculate the ground-resolved distances required for interpreting specific targets. The National Imagery Interpretability Rating Scale (NIIRS) is a standard used to convert the Ground Resolved Distances (GRDs) into a quality rating for a particular image sample. Figure 1 illustrates the different ratings and definitions for each point on the NIIRS scale.

Unfortunately, objective measures may not provide an adequate measure of suitability of a compression technique for images to be interpreted or used by people. For some applications, such as medical imaging, quantitative measures are adequate because only perfect recovery of the imagery is acceptable regardless of subjective assessment of the quality. For many applications, however, including the UAV, objective fidelity measures may not provide sufficient assessment of the suitability of the decoded image for viewing and interpretation by a human observer. The human visual system is capable of extracting usable information from imagery that when objectively measured, is seriously degraded. Thus, two pictures having the same amount of assessed error may have profoundly different visual qualities when judged subjectively by humans.

Most image compression studies either avoid the issue of evaluation of imagery entirely or merely display input and output pictures and comment on a vague impression of the preservation of image quality. These comparisons focus primarily upon the aesthetic appeal of the image and fall short of addressing the issue of its intelligibility by humans. That is, relatively few compression studies evaluate the effects of the compression/decompression on subsequent visual performance.

Rating Level 0

- Interpretability of the imagery is precluded by obscuration, degradation, or very poor resolution.

Rating Level 1

- Detect a medium size port facility and/or distinguish between taxiways and runways at a large airfield.

Rating Level 2

- Detect large hangers at airfields.
- Detect large static radars (e.g., AN/FPS-85, COBRA DANE, PECHORA/KRASNOYARSK, HENHOUSE).
- Detect military training areas.
- Identify an SA-5 site based on road pattern and overall site configuration.
- Detect large buildings at a naval facility (e.g., warehouses, construction halls).
- Detect large buildings (e.g. hospitals, factories).

Rating Level 3

- Identify the wing configuration (e.g., straight, swept, delta) of all large aircraft (e.g., 707, Concord, BEAR, BLACKJACK).
- Identify radar and guidance areas at a SAM site by the configuration, mounds, and presence of concrete aprons.
- Detect a helipad by the configuration and markings.
- Detect the presence/absence of support vehicles at a mobile missile base.
- Identify a large surface ship in port by type (e.g., cruiser, auxiliary ship, non-combatant/merchant).
- Detect trains or strings of standard rolling stock on railroad tracks (not individual cars).

Rating Level 4

- Identify all large fighters by type (e.g., FOXBAT, FULCRUM, F-15, F-14).
- Detect the presence of large individual radar antennas (e.g., TALL KING).
- Identify, by general type, tracked vehicles, field artillery, large river crossing equipment, wheeled vehicles, when in groups.
- Detect an open silo door.
- Determine the shape of the bow (pointed or blunt/rounded) on a medium size submarine (e.g., ROMEO, HAN, Type 209).
- Identify individual tracks, rail pairs, control towers, switching points in rail yards.

Rating Level 5

- Distinguish between a MIDAS and CANDID by the presence of refueling equipment (e.g., pedestal and wing pod).
- Identify radar as vehicle mounted or trailer mounted.
- Identify, by type, deployed tactical SAM systems (e.g., FROG, SS-21, SCUD, LANCE).
- Distinguish between SS-20/SS-25 mobile missile TELs and missile support vans (MSVs) in a known support base, when not covered by camouflage.
- Identify TOP STEER or TOP SAIL air surveillance radar on KIROV, SOVREMENNY, KIEV, SLAVA, MOSKVA, KARA, or KRESTA-II class vehicles.
- Identify individual rail cars by type (e.g., cattle, enclosed box) and/or locomotive by type (e.g., steam, diesel).

Rating Level 6

- Distinguish between models of small/medium helicopters (e.g., HELIX A from HELIX B from HELIX C, HIND D from HIND E, HAZE A from HAZE B from HAZE C).
- Identify the shape of antennas on EW/GCI/ACQ radars as parabolic, parabolic with clipped corners, or rectangular.
- Identify the spare tire on a medium sized truck.
- Distinguish between SA-6, SA-11, and SA-17 missile airframes.
- Identify individual hatch covers (8) of vertically launched SA-N-6 on SLAVA class vessels.
- Identify automobiles as sedans or station wagons.

Rating Level 7

- Identify fitments and fairings on a fighter sized aircraft (e.g., FULCRUM, FOXHOUND).
- Identify ports, ladders, vents on electronics vans.
- Detect the mount for anti-tank guided missiles (e.g., SAGGER on BMP-1).
- Distinguish between the inner and outer liner in a missile silo when the door is open.
- Identify the individual tubes of the RBU on KIROV, KARA, KRIVAK class vehicles.
- Identify individual railroad ties.

Rating Level 8

- Identify rivet lines on bomber aircraft.
- Detect horn-shaped and W-shaped antennas mounted atop BLACK TRAP and BLACK NET radars.
- Identify a hand held SAM (e.g., SA-7/14, REDEYE, STINGER).
- Identify joints and welds on a TEL or TELAR.
- Detect winch cables on deck mounted cranes.
- Identify windshield wipers on a vehicle.

Rating Level 9

- Differentiate cross-slot from single slot heads on aircraft skin panel fasteners.
- Identify small light-toned ceramic insulators which connect wires of an antenna canopy.
- Identify vehicle registration numbers (VRN) on trucks.
- Identify screws and bolts on missile components.
- Identify braid of rope (1-3 inches in diameter).
- Detect individual spikes in railroad ties.

Figure 1. National Imagery Interpretability Rating Scale (NIIRS)

Quantitative imagery analysis is one method of defining the baseline image quality of the video footage used for the experiments in this study. This type of analysis, coupled with operators' subjective image quality ratings and examined within the context of operator performance in selected mission tasks, should be useful for understanding the effects of compressed imagery of a specified quality on the UAV operator.

5. Human Performance in Target Acquisition Tasks With Different Data Reduction Techniques

This section contains a brief review of the extant human performance literature on frame rate, resolution, and compression trade-offs and costs to the operator. Most of this data was collected and analyzed from the 1960s to the early 1980s and was based on wideband video data links used in connection with the development of remotely piloted vehicles (RPVs). Some of the later work reviewed addresses the bandwidth limitation and jamming impacts for sensors such as TV and FLIR.

All of the studies reviewed were carefully designed empirical investigations using sound control methods and statistical analyses. As is the case with much experimental laboratory work, the imagery used was largely simulated mission imagery that frequently lacked realism. That is, the target sets were usually single items (tanks, jeeps, ships, APCs) placed on a plain background. Only two studies reviewed used terrain imagery [10], [11], and both of these scenarios were prepared specifically for the purposes of experimentation.

Previous research has demonstrated that a combination of data transformation techniques with bandwidth reduction, through reduced frame rate or resolution, can result in a reduced data rate [12]. Considerable research on human performance has been conducted when these two techniques are used to manipulate frame rate and resolution [10], [13], [14].

Results from these studies also show that the manner in which the reduction or compression is implemented affects the ultimate data reduction [9]. Some of the empirical results are presented below as background for the experiments conducted in thus study and to provide a theoretical framework in which to discuss the human performance results observed in the experiments conducted as part of the current task.

5.1 Frame Rate Studies

Human performance has been investigated using frame rates as low as 0.12 frames per second and increasing to a normal 30 frames per second. In general, errors increased dramatically as frame rates dropped below 2 fps. These data indicate that lower frame rates (0.94, 1.88) are extremely difficult to use while higher frame rates (15, 30) are relatively easy to use depending on the nature of the task. For example, more dynamic, complex cognitive and motor tasks including tracking, slewing, and designation are performed better with higher frame rates.

Studies in target detection and recognition found that operators could perform their tasks with frame rates lower than 2 frames per second [10] [14], [15]. Frame rates less than 1 fps, however, resulted in an initial frame delay inherent in the transmission of sensor imaging to the ground station display that affected performance. Hershberger and Vanderkolk [14] found that the 1 to 2 frame transmission delay resulted in an initial range time penalty for operators (e. g., a 45 feet range penalty with data transmitted at 7.5 fps). When transmission delays are eliminated, operator performance was found to be proportional to frame rate.

Once a target is detected, the Mission Payload Officer (MPO) may be required to position the sensor so that the target will be near the center of the search field prior to switching to a stronger viewing lens. This will permit higher magnification (but at a reduced FOV) for target recognition. This requirement involves target slewing. In a study on the effect of frame rate on precision sensor slewing the greatest reduction in slewing time occurred at the 3.75 frame rate [14], [16]. No significant performance changes were noted when frame rate increased to 7.5 frames per second. The method of sensor control typically used in previous UAV systems (image motion compensation, continuous, bang-bang) was also found to interact with frame rate on the required time to slew a target, but no differences in performance between the three control modes were noted at the 3.75 frame rate. The UAV planners should be made aware of the effects of these methods in order to specify sensor packages.

In general, UAV operators were able to successfully detect and recognize targets with minimal errors at 1.88 fps [13], [16]. These studies show that once the target has been detected and classified, higher frame rates are needed so that an adequate sampling of imagery is presented to the operator during slewing or tracking tasks such as those mentioned above. This research on frame-rate reduction in a tracking task indicates that 3.75 fps supports adequate operator performance.

These findings suggest that data reduction trade-offs depend on the type of task the operator is performing. Control modes will affect performance in more dynamic tasks and should be considered when defining the bandwidth requirements. More specifically, (1) the MPO display console designers should provide the operator with variable frame rate capabilities for specific tasks. (See Vitro report on Human Engineering Guidelines for the UAV Mission Planning Console System. December 1991), and (2) the lowest acceptable frame rate value for dynamic task performance should be investigated in order to assure adequate performance for all tasks specified in the missions. This study will address item (2).

5.2 Resolution Studies

Seminal research on human performance requirements for sensor resolution was carried out by Johnson [17] over 30 years ago. He used TV lines as the measure of resolution and established performance baselines for target detection and recognition that are still the standard used today. He found that 2 TV lines are required for detection at a 0.50 probability criterion level and 8 TV lines at a 0.90 probability criterion level. Similarly, he found that 3 TV lines are required for recognition at a 0.50 probability criterion level and 14 TV lines at a 0.90 probability criterion level. Later work by Erikson [18] validated and extended Johnson's work to include more operator tasks and targets. These criteria define the minimum resolution requirements of interest to current UAV bandwidth limitations.

Hershberger [12] reviews several resolution studies that examined target detection. In some of his studies he examined the effects of resolution on single and multiple target detection and recognition [13] using the number of TV lines across a single dimension (length or width for example) of the target as the performance measure for various levels of bandwidth compression. Detection of multiple targets was easily accomplished with less than 5 TV lines while similar performance (0.80 probability) for detecting a single target required 12 TV lines. Recognition of targets requires greater resolution, as one would expect, since this task requires that more information be processed to identify the target. These results suggest task dependencies for resolution reduction similar to the frame rate limits discussed above.

In summary, 2 TV lines across the target height are required for multiple target detection when 2 bits-per-picture-element compression is used. Four TV lines are required for single target detection.

5.3 Bandwidth Compression Studies

Various human factors studies have examined effects of bandwidth compression on operator performance. Performance on a range of operator tasks (detection. recognition. slewing) has been tested with imagery quantized at 0.4. 0.8, 1.6. 3.2. and 6.0 bits per picture element. Over this range of compression levels no significant difference was measured in an operator's ability to detect multiple targets (target numerosity) [13]. Single targets proved more difficult to detect under data compression and the imagery degradation at the higher compression levels made the task more difficult to perform. A second study in the same report looked at several compression levels (5-, 2-, and 1-bit per picture element) using the DCT/DPCM technique in combination with zero, 10^{-3}, and 10^{-2} error bit-rate jamming. A noncompressed imagery condition (6-bits per picture element) was also used in the study to collect baseline performance data. Results showed there was no performance degradation in a recognition task until a 6:1 compression (1-bit per picture element) was reached. Bit error rate jamming had no effect on performance. The author concluded that operators can perform recognition tasks with minimal effects on performance at 1.5-bits per picture element and 10^{-2} bit-rate error jamming levels.

Mills, et. al. [11] compared two different compression techniques (1-dimensional Hadamard transform, with digital pulse code modulation, and a 2-dimensional cosine transform with frame sampling) in a human performance study. They used a 2 X 2 X 2 factorial design manipulating different levels of frame rate (1 and 7.5 fps), resolution (128- and 256- pixels), and compression methods (listed above). Operators were required to find a pre-briefed target, lock onto the target, perform lock-on adjustments, and control the weapon target with a hand control device. Results suggest that video symbology (display cross-hairs, messages) should not be processed with imagery at transmission rates of 300 kilobits or less, and display resolution of 128 lines (or pixels in the vertical dimension) or less. This result should be considered by the UAV display designers. Additionally, results suggest that display resolution may be more important relative to frame rate and the number of bits per picture element. However, the subject pool was small (N = 8) and none of the results were statistically significant.

A summary of the essential bandwidth studies was reviewed in Hershberger and Farnochi [12]. The general consensus among the studies surveyed is that operator performance on UAV tasks can be supported with compression at 2 bits per pixel. This value was further verified in related human performance studies of frame rate, resolution, and grey scale manipulation in a teleoperation task [19].

5.4 Summary

The general recommendations that resulted from these studies suggest adequate human performance at a 300-400 kilobits per second transmission rate. This is considerably more than is available in the UAV program if JTIDS is selected (119 kilobits per second). The data support the use of a 1.88-2 frame rate for target detection and recognition and an increase to 3.75-4 for target tracking and slewing. These frame rate values are dependent upon the sensor resolution that can be displayed to the operator [12], [13]. Baseline total display resolution values (full and half) have been identified from the literature review that can be used in combination with the minimum frame rates in our experiments. The compression algorithms used in these studies could achieve only a 30:1 maximum ratio which does not reduce the imagery to fit within narrower bandwidths. Furthermore, while these studies provide valid and reliable results, they did not utilize real mission imagery. This is another factor this study addresses.

In order to meet the required bandwidth limitation, the use of compression to reduce the imagery data to within these limits is proposed. What must be verified is whether the identified minimum frame rate values (1.88 and 3.75) enable the human operator to perform at an acceptable performance level. In the simulation prepared for Experiment One, 2 and 4 frames per second were used. This was a result of the compression technique used (See Equipment section). In addition, actual mission imagery was used instead of experimentally prepared scenarios so as to provide more reasonable fidelity with actual mission performance. Full versus half resolution (total display) given the constraint of using existing imagery footage was compared. The bandwidth for the specified variables of interest are given in Figure 2 below. Human performance for each matrix cell in the two experiments described in the next section.

	2	4	30
640 X 480	98 kb/sec	196 kb/sec	73 Mb/sec
640 X 240	44 kb/sec	98 kb/sec	

Figure 2. Bandwidth Limits for Two Levels of Frame Rate (2, 4), Resolution (Full-640 X 480, Half-640 X 240) and One Compression Level at a 50:1 ratio with the DCT method (Control imagery- 30 fps, full resolution - was not compressed).

6. Human Factors Experiments

The primary purpose of these experiments was to verify and validate that data reduction of video imagery through frame rate and resolution trade-offs is sufficient to meet the needs of the human operator in fulfilling mission objectives. For study purposes, a JTIDS-based data rate of 119-200 kilobits per second was used as a benchmark. Psychophysiological assessment was made of operator performance in three basic tasks that support the selected UAV mission tasks discussed earlier (i.e., detection, recognition, designation and tracking) using actual military video imagery. We examined both static (detection, recognition) and dynamic (designation, tracking) tasks in order to establish the appropriate frame rate and resolution combination that we will provide as guidance for the UAV data link design.

The experiments validated many of the human performance results and image quality assessment measures as described earlier. One condition (3.75 frames, full resolution) does not meet the target JTIDS data rate limitation of 119 kb/sec. It was examined nonetheless since the best performance was predicted for that condition. It is recognized that only one JPEG algorithm, the DCT, was tested in these experiments. The DCT algorithm is currently available in hardware as a non-developmental item and is well documented as an international standard. The state of the technology is developing rapidly and in the near term other algorithms will be available for testing. Future work assessing these algorithms can be performed as the technology matures.

Two types of experiments were conducted: one that examined performance in the static tasks with real UAV mission imagery, and another that investigated the dynamic tasks in a simulated mission scenario with sensor flight parameter manipulations.

6.1 Experiment One

The first experiment evaluated operator visual performance in target detection, recognition, and designation tasks using actual Desert Storm Pioneer imagery. The missions supported in this experiment cover Reconnaissance, Surveillance, and Target Acquisition, Gunfire/Artillery Spotting, and Bomb/Battle Damage Assessment. Several conditions were prepared that manipulated two of the three bandwidth reduction variables of interest: frame rate and resolution. Compression was held constant at a 50:1 ratio. A simulation of the three target acquisition tasks was developed utilizing imagery clips taken from actual mission footage. The assumption about the Pioneer footage is that the imagery and missions used represent reliable real world UAV scenarios. While the sensor mission characteristics for these missions were not under experimental control, it was assumed that UAV operators conduct their tasks with imagery of similar quality. In order to account for the differences in image content in the experiment, the images were categorized into groups which characterized the complexity and nature of the different scenes and target sizes.

6.1.1 Methodology

6.1.1.1 Research Design

A 2 X 2 X 3 mixed factorial design was used to present the two different levels of frame rate and spatial resolution (total display) in three target acquisition tasks (detection, recognition, designation) to 57 participants. The imagery was compressed at a 50:1 ratio using a DCT method. A control group that viewed the imagery at full frame rate (30 fps), full resolution, and no compression was added to the design. In this way, five experimental conditions were prepared: 1) 4 frames, full (640 X 480) resolution, 50:1 compression, 2) 2 frames, full (640 X 480) resolution, 50:1 compression, 3) 4 frames, half (640 X 240) resolution, 50:1 compression, 4) 2 frames, half (640 X 240) resolution, 50:1 compression, and 5) 30 frames, full (640 X 480) resolution, no compression. The order of presentation of the three tasks was balanced across all participants to control for any sequence or learning effect. (Readers not familiar with experimental design and related technical terminology should consult "Design and Analysis, A Researcher's Handbook," by Geoffrey Keppel, Prentice Hall, 1982.)

A baseline study was conducted first with experienced servicemen to serve as the performance baseline. A subsequent experimental study was then conducted following the same design and procedures except that experience was evaluated by comparing active-duty military personnel with non active-duty personnel who had prior, related military experience.

Independent Variables. Frame rate was set at levels of 2 and 4 fps (between subjects manipulation). Spatial resolution was set at display levels of 640 x 480 (full resolution) and 640 x 240 pixels (half resolution) (between subjects manipulation). The video compression ratio was fixed at 50:1. A control group was also included at full resolution, full frame rate (30 fps), and no compression. Operator tasks were target detection, target recognition/identification, and target designation. These tasks were a within subjects manipulation (i.e., all participants performed each of these tasks).

Dependent Variables. Several measures of operator performance were collected to include: reaction time, errors in detection of targets, errors in designation of targets, errors in recognition of targets, confidence ratings, and video image quality rating.

To compensate for the variability in image quality inherent in "real mission footage", dependent measures for each participant were taken as deviation scores from the grand mean for each clip.

6.1.1.2 Participants

Twelve naval personnel from the Naval Air Test Center (NATC), Patuxent River, served in the baseline study. Forty-five subjects participated in the experimental study. Fifteen were NATC personnel, and 30 were Vitro employees who were screened and selected on the basis of prior military experience with target acquisition tasks.

6.1.1.3 Imagery and Target Scenarios

Representative video clips were selected for each mission (over-the-horizon target detection and classification, naval gunfire support, and battle damage assessment over land). The imagery quality was analyzed using a quantification methodology based on the Rome Air Development Center (RADC) technique [8], [9]. This technique uses the National Image Interpretation Rating Scale (NIIRS).

Image analysis consisted of determining the minimum ground resolved distance (GRD) for each clip on the experimentally compressed tape. For ease of use, the GRD is converted into a rating based on the NIIRS, reproduced from a Rome Laboratory report [8] and shown below.

NIIRS Rating	1	2	3	4	5	6	7	8	9
Ground Resolved Distance	30 ft	15 ft	8 ft.	4 ft.	30 in	16 in	8 in.	4 in.	2 in.

NIIRS Rating Scale

The assumption is that representative video clips serve as the foundation for assessing image quality in general for the mission imagery used in this experiment.

Target Scenarios. TV video imagery taken from actual Pioneer Remotely Piloted Vehicle missions was used to prepare the experimental simulation. Scenarios were developed after reviewing several hours of footage. Clips were selected that reflected actual examples of the three types of missions listed above. In contrast to other experimental work where all dimensions of the imagery (scene complexity, background, targets, contrast) are controlled, the footage used here varied across these dimensions. Furthermore, the footage included real-time mission flight parameters (vehicle speed, altitude, field of view, ground range to target) that changed the sensor altitude as the Pioneer flew its course. In order to control these variables for data analysis, the imagery was categorized according to several relevant image dimensions. Once categorized, a sequence of 10-second clips were developed for each of the three experimental tasks.

The dimensions selected included scene type (land or sea), scene complexity (feature - no features), target size (large-small), and overall noise level for the footage (yes - no). This last category was used since some inherent signal interference was captured when the footage was originally obtained. These categorized groupings resulted in 12 types of general imagery that were utilized in the experiment. Different target types were available: vehicles (tanks, ships, trucks, planes, helicopter), personnel, military installations (communications centers, observation towers), military artifacts (bunkers, revetments), and shell fire (artillery bursts).

6.1.1.3.1 Quantitative and Subjective Imagery Analysis

Unedited video footage was used in the quantitative analysis, making sure the frames evaluated were identical to those the participants viewed by stepping frame by frame to the appropriate clip. Most of the original clips were longer than the ten to twenty seconds used in Experiment One. This allowed recognition of more and smaller. objects. The additional time was reviewed for this analysis. The smallest distance visible on the experimental clip and measurable on the original clip served as the GRD. In many of the clips it was possible to estimate the GRD by measuring in the original clip the smallest object visible in the experimental clip. In other cases one could estimate the size of an object in the original clip and use it as a 'yardstick' in the experimental clip.

In most of the clips, the video images have superimposed telemetry information. One key datum is the scale (marked 'SCL XXXX-YYYY' in the lower right hand corner) representing computed ground distances in meters per centimeter on the UAV screen in the X and Y direction. The scale parameters become less accurate as the object of interest moves away from the center of the screen. The parameters are integers. so that a '5' represents a value between 4.5 and 5.5 m/cm. The UAV monitor has an image size of 4.5 X 6 inches. For the image analysis. the clips were viewed on a television having an image size 2.5 times as large, so the scale parameters are in units of meters per inch. One representative clip from each of the scenario complexity and target categories described above was analyzed. The results are summarized below. The table gives the NIIRS rating for the selected clip in each division.

Target / Conditions Terrain	Small		Large	
	Noise	Non-noise	Noise	Non-noise
Sea	1	5	6	9
Land No Features	6	1	6	7
Land Features	5	7-8	4	7-8

NIIRS Ratings for Representative Video Clips

The descriptions of the terrain scenarios in each category are given below with a description of each rating. Following this quantitative analysis, we present a similar rating table with the participants' subjective ratings of image quality.

Sea scenarios, small targets

Noise. This clip depicts TV imagery of a single ship in a vast sea scenario with no other objects visible. The overall clip is foggy which introduces general noise in the clip. The ship is viewed from a great distance making it a small target with respect to the rest of the scene. In the compressed clip, one can barely see the outline of a ship, obliquely, and only when told it is present. In the control clip, the ship is more discernible and can be detected. No people or objects are visible on the ship, but the scale can be roughly estimated. The GRD is approximately 30 feet. NIIRS 1

Non-noise. This clip is similar to the above clip, but shows a different ship at a close-up range. Here the target is very large; the obvious object of interest. The clip is clear and details of the ship's surface and name are easily visible. In the original clip the letter 'I' presents itself near the cross-hairs. It is 1/4 of an inch wide, and the scaling factor from the Pioneer telemetry window is 4. Therefore, the resolved letter is 1 meter wide. It is likely that something 30 inches would also be resolvable. NIIRS 5

Sea scenarios, large targets

Noise. This clip provides FLIR imagery of a warship. The contrast for most ship elements is low. This is largely due to a few selected high contrast areas resulting from the heat signatures of the exhaust stacks. Objects about 1/8" thick can be resolved on the compressed tape. On the original tape, the scaling factor from the Pioneer telemetry window can be read as 3, resulting in a GRD of 15". NIIRS 6

Non-noise. This clip provides daylight TV imagery of a ship loaded with a number of vehicles. There are many covered stake bed trucks on the deck of the ship. The trucks include cabs and are smaller than American semis. The windshields and doorposts are clearly visible. Substantially narrower cables are also visible on the ship. NIIRS 9

Land scenarios-no features, small targets

Noise. This clip depicts a FLIR image of a cluster of about 6 or 8 people with a few other personnel nearby. Image is quite grainy. Forms can be distinguished moving around, but they are only easily identified as personnel on the noncompressed tape. NIIRS 6

Non-noise. This clip shows an artillery burst in a relatively featureless desert using FLIR imaging. The burst appears as a high contrast black spot. The smallest object resolvable in the compressed clip is about 1/4" on the screen. Using the scaling factor from the pioneer telemetry window of 50, the GRD is calculated to be 12.5 meters. NIIRS 1

Land scenarios-no features, large targets

Noise. This clip provides daylight TV imagery of a large formation of surrendering troops. Troops are seated on the ground. The formation is closely packed in a rectangular array, making it difficult to resolve personnel within the formation as individuals. In the original clip people can be seen walking along the sides of the formation. Forms can also be seen moving in the compressed tape. NIIRS 6

Non-noise. This clip provides daylight TV imagery of a communications or radar installation. In the original clip the communications tower can be seen to have shafts or antennae sticking out at the top. They are less than 1/16" in thickness on the screen, and the scaling factor from the pioneer telemetry window is 3, indicating that they are less than 7.5" thick. The beams are distinguishable on the compressed tape. NIIRS 7.

Land scenarios-features, small targets

Noise. This clip depicts daylight TV imagery of a helicopter carrying a High Mobility Multi-Wheeled Vehicle (HMMWV). The HMMWV was the target of interest. In the original tape, the narrowest part of the helicopter fuselage is resolved at less than 1/4" in thickness on the screen, and the scaling factor from the pioneer telemetry window is 3, indicating that it is less than 3/4 of a meter thick. NIIRS 5

Non-noise. This clip gives daylight TV imagery of a small truck about the size of a land rover. The vehicle is moving over an empty field between two developed areas. The vertical door posts between the windows can be seen briefly. The bar seen should be less than 6" wide. NIIRS 7-8

Land scenarios-features, large targets

Noise. A vehicle can be seen moving along a paved roadway on this FLIR clip. The heat signature of the vehicle has high contrast with the surrounding terrain. In the original tape, resolvable vehicle features measure 1/4" x 1/2" with a scaling factor from the telemetry window equal to 5, so GRD is approximately 50". NIIRS 4

Non-noise. This clip gives daylight TV imagery of a small truck about the size of a land rover. The vehicle is moving over an empty field between two developed areas. The vertical doorposts between the windows can be seen briefly. The bar seen should be less than 6" wide. NIIRS 7-8

6.1.1.3.2 Subjective Ratings

A similar ratings table for the same representative clips analyzed above was prepared using the mean subjective ratings that participants recorded in Experiment One. The subjective rating scale differed from the NIIRS scale in two respects. First, a 5-point scale ranging from 1 'very poor quality' to 5 'very good quality' with 3 as 'average' was used. The table below shows the mean rating scores for the same representative clips analyzed with the NIIRS scale. When data from both scales was analyzed, a modest correlation between the two ratings was found (r = 0.35). Thus many of the ratings followed the same rating trend with a few exceptions.

More complete results on image quality ratings are given in the results section. It is interesting to note is that subjective ratings appear to be task dependent. For example, participants rated the same clip used in two separate tasks differently. In a recognition task, which was more difficult for the participants to perform, a 2.8 rating was obtained whereas the same clip in the designation task was rated 3.4. The designation task was the easier task of the two in terms of performance requirements. It appears then that when operators feel confident in their ability to perform, they tend to rate the imagery as higher quality regardless of the type of degradation involved (compression, frame rate, resolution).

Target / Conditions — Terrain	Small		Large	
	Noise	Non-noise	Noise	Non-noise
Sea	2.0	3.4	2.6	3.9
Land No Features	2.0	3.2	2.4	2.8
Land Features	3.2	3.4	3.1	3.3

Subjective Mean Ratings for Representative Video Clips

6.1.1.4 Equipment

The equipment utilized in the implementation of the simulation included the following components: hardware, software control programs and video imagery compression. The bandwidth compression and frame-rate reduction were controlled through hardware components. Targets and mission scenarios were made available through existing military TV video tapes. Operator tasks, response measures, and experimental procedures and sequencing were defined under software control.

The Vitro Human-Computer Interaction Laboratory was used to carry out the experiments with Vitro personnel. A portable version of the equipment was taken to Naval Air Test Center (NATC), Patuxent River, MD, to collect data from the active-duty personnel. The equipment used to present the experimental imagery included an Apple Macintosh II FX with a 19" SuperMAC monitor and a Raster Ops 24XLTV video board. A NEC PC-VCR model PV-S98A was used to present the imagery under software control to the Macintosh system. The Macintosh had 8 Megabytes of memory, and a 300 Megabyte hard disk. A mouse was used in the target designation task of the experiment. The visual angle for FOV from operator to the monitor was fixed at 16° to 20°. The Macintosh was running MacOS System 6.0.5, and Finder version 6.1.5. The Raster Ops video board was set to display 24-bit color, and captures video from the PC-VCR at a rate of 30 frames per second (fps). The control tape was displayed at 30 fps and no compression. Another VHS tape was created using compression, described below, at a rate of 4 fps. The 2 fps configuration was done by capturing the video at the 2 fps rate on the Raster Ops video board, and only displaying those images as they were captured.

The source video was transmitted to the Raster Ops board by the PC-VCR by a composite National Television Standard Committee (NTSC) signal. The Raster Ops board displayed the image in a 640 X 480 window on the Macintosh's 19" monitor. The half-resolution screen used the same configuration except the screen size displayed was 640 X 240 pixels across and down respectively. The PC-VCR is capable of playing VHS, and SuperVHS tapes. The experiment used standard VHS tapes created from 3/4" tapes supplied by the NATC.

The experimental imagery was coded according to task and target type. A series of 10 - second clips for the detection, recognition, and designation tasks was time-stamped and prepared. This experimental videotape was then compressed using a hybrid DCT algorithm. The tape was slow scanned at 4 frames per second through professional video equipment in Vitro's Video Production Laboratory.

The compression algorithm used was obtained from GEC - Marconi. GEC-Marconi's video compression workstation accepts a standard interlaced NTSC 60 field/sec video signal (since two fields are needed for a frame, we are accepting 30 frames/sec). The system captures a single frame at 512 pixels horizontal, by 200 lines vertical resolution (slightly underscanned) and performs a DCT-based compression algorithm. The resulting compressed video frame is then decompressed and displayed on a second frame grabber at 512 X 200 resolution. Again, the second frame grabber produces an interlaced NTSC video signal. This system would normally produce a two fps video output which was not sufficient to perform testing at NATC. To solve this problem, the test video was played at 15 fps into the video digitizer (with the GEC compression algorithm operating at two fps) and the output recorded at 15 fps onto a final video tape. When the final video is played at full speed, a 4 fps video sequence is produced.

The degree of compression (called 'q' factor) can be controlled on a per frame basis, but a given 'q' does not yield a fixed compression ratio. In addition, there is no way of predicting (with much accuracy) the amount of compression you will get on a particular frame for a given 'q' value. To solve this problem, several test runs were made on the test video sequence to pick a 'q' factor that resulted in an average compression ratio of 50:1 over the 20 minute video tape. We used this 'q' factor to compress the experimental video imagery for use in the experiment.

6.1.1.5 Procedures

Upon arrival at the laboratory, the experimenter greeted each participant and had the participant perform the following activities: complete a demographic background questionnaire; read the task description; and read and sign informed consent and nondisclosure documents as appropriate. The experimenter encouraged questions, and answered them as necessary.

Participants were then assigned to one of the five conditions (including the Control group) in the experimental design. Each participant performed each of the three representative target acquisition tasks whose order was counterbalanced. Procedures for each task were as follows:

Target Detection. A series of 2 practice video clips and 30 test video clips was presented to each participant. Of the 30 test video clips, 15 contained a target, and 15 did not. Each video clip was 10 seconds in length. The stimuli and automated instructions for each target detection trial were presented sequentially on-line. Instructions given to the participants at the beginning of the session identified the types of targets considered to be military targets for the purposes of the current experiment. At the start of each task, the participant was given two practice trials. Each trial started with an instruction for the participant to respond "yes" or "no" (by pressing appropriately labeled keys) as soon as he was "reasonably confident" whether a military target was present or not. For certain clips a specified target or distracting feature was specifically excluded for that particular trial. Instructions remained on the screen until the participant pressed a key to continue. In this way, the onset of each trial was self-paced. The clip ran until the participant responded "yes" or "no" or for 10 seconds, whichever was shorter. Time interval was collected from the start of the clip until a keypress or the end of the clip. If the participant responded "yes", the participant was asked to enter the quadrant containing the target. This response was recorded. A rating of the participant's confidence that a target was present was also obtained. The participant was asked to rate the quality of the imagery on a scale from 1 to 5. If the participant responded "no" or allowed the clip to "time out," the participant was only asked to make a video image quality rating.

Target Recognition. A series of 2 practice video clips and 20 test video clips was presented to the participant. Each video clip was 10 seconds in length. Instructions were given on the monitor to provide some detailed information regarding a particular feature in the scene or the type of target present in the video clip to be recognized. The participant was instructed to press a key when he was reasonably confident that he could identify the feature or target present in the clip. These instructions remained until the participant pressed a key indicating readiness to go on.

The clip ran until the participant responded that he could identify the feature or target in the scene or for 10 seconds, whichever was shorter. Regardless of whether he responded or let the clip "time out", the participant was directed to rate his confidence of the identification made on a scale from 1 to 5. Next, a four option forced-choice question related to identification of the target or feature was presented on the screen and the participant was asked to respond by selecting a number key (1 to 4) that mapped to a particular choice. Lastly, the participant was asked to rate the quality of the video image on a scale from 1 to 5.

Target Designation. A series of 2 practice video clips and 20 test video clips was presented to the participant. Each video clip was 10 seconds in length. Instructions were given on the monitor to designate a particular target. The participant was instructed to use a mouse and position the cursor over the target described, and then press the mouse button to "designate" it. These instructions remained until the participant pressed a key indicating readiness to continue. The clip ran until the participant designated the target in the scene or for 10 seconds, whichever was shorter. Latency of designation and accuracy were recorded. Lastly, the participant was asked to rate the quality of the video image on a scale from 1 to 5. Upon completion of all trials of all tasks the participant was debriefed and dismissed.

Next results from the baseline study for Experiment One and then the complete Experiment One results are presented. This is followed by a general discussion of all the results between tasks and the dependent variables.

6.1.2 Baseline Study Results

Data from the baseline study were collected from NATC personnel so that a performance baseline with experienced personnel could be established. Twelve participants engaged in the 3 target acquisition tasks described above: Detection, Recognition, and Designation. Five dependent variables were examined that include objective measures of reaction time (RT), error rates, subjective ratings of confidence and of image quality, and RT for the confidence rating decision.

The following sections describe the analysis and results for each task.

6.1.2.1 Target Detection

Reaction Time. The effect of frame rate (FR) on RT was highly significant ($F = 42.48$, $p < .0001$), indicating that higher FR influences response time. RT for the 2 fps clips was 5.866 secs. compared to 3.900 secs for the 4 fps clips. Performance in the control group, who viewed full resolution clips at 30 fps, was 3.567 secs. In post-hoc comparisons, this group was not significantly different from the 4 fps group, but was significantly different when compared to the 2 fps group.

Spatial resolution (SR) had no effect on RT for either full or half resolution conditions (4.950 secs and 4.717 secs, respectively). There was no interaction effect between resolution and FR.

Errors. There were no significant effects for either FR or SR on the percent of correct detections made by the participants despite an apparent difference in the mean percentage of correct detections (4 fps had 50%; 2 fps had 36%).

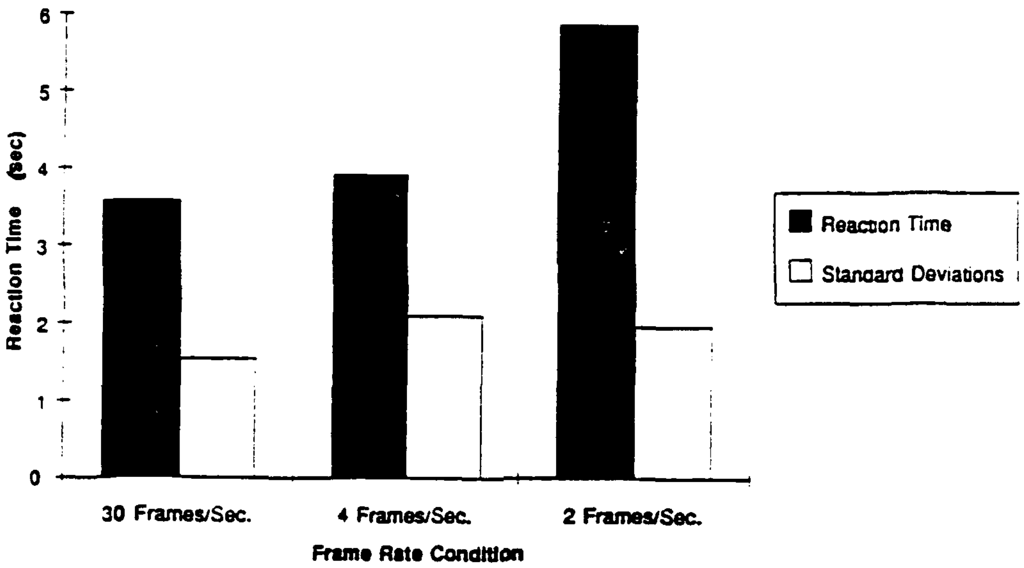
The control group, when included in the analysis, performed no better than any of the experimental groups, with 59% of correct detections obtained. Figure 3 illustrates the means and standard deviations for RT and errors.

Another factor related to error performance is how often participants "timed-out." We arbitrarily chose 10 second clips for use during the experiment, to put a boundary around the RT measure. This measure helped us to interpret operator performance requirements for real-time, mission-critical situations when time is a factor. These data describe a participant's inability to make a decision either "YES" or "NO" within the 10 sec. time frame. Time-out responses were not included in the data analysis of reaction time. By looking at these data separately, it is seen to what degree participants were uncertain about their decision making in the different experimental conditions.

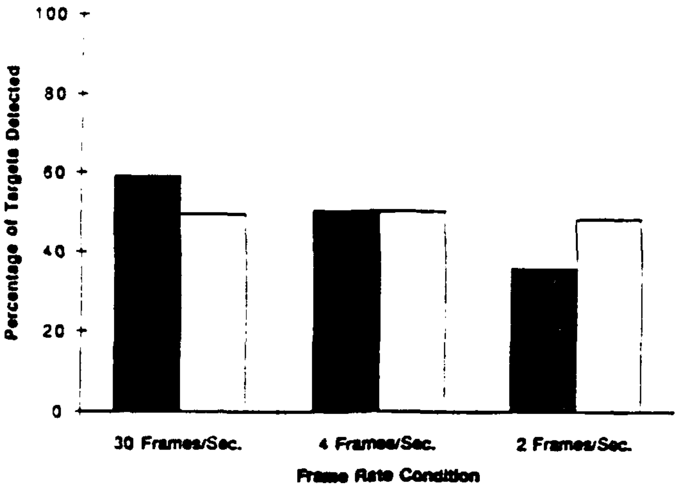
These percentages are quite high indicating that the participants were unable to detect targets 40% to 50% of the time within 10 seconds. While this seems to suggest that experienced NATC personnel have difficulty detecting targets quickly, we must consider the nature of the task. Target detection usually occurs in reconnaissance and surveillance missions. In these missions, operators are scanning terrain for possible targets of interest and may view imagery for several minutes or

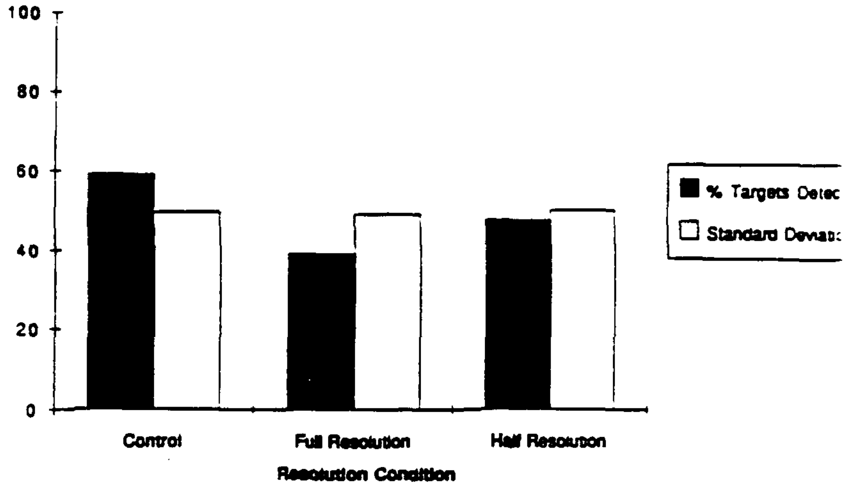
Figure 3. Means and Standard Deviations for Reaction Times and Errors in Target Detection - Baseline

Reaction Time



Errors





hours at a time rather than a few seconds. This was reported verbally by NATC personnel. Thus our 10-second time frame may not be realistic for this type of task. Nevertheless, timeliness may be a critical factor in real-time missions. There were no significant effects for either FR or SR on the percentage of clips where participants were unable to decide whether a target was present or not. Table 1 illustrates the mean percentages of time-outs during the detection task.

Table 1. Mean Percent of Time-Outs During Target Detection

	Percent
Frame Rate	
30	45.56
4	43.33
2	50.00
Resolution	
Control	45.56
Full	50.67
Half	42.67

Ratings. The influence of FR was evident in both confidence measures. Confidence ratings (CR) for the presence of a target was significant ($F = 6.77$, $p < .01$) with means for 2 and 4 fps equal to 3.05 and 3.35, respectively. Another indirect measure, assumed to be related to this subjective rating, was also collected. This measure, confidence rating time (CRT), indicates the speed with which a participant made his confidence rating. It was assumed that higher confidence would result in faster responses. The effect of FR on CRT was highly significant ($F = 13.641$, $p < .0003$). The 4 fps groups were faster (2.467 secs) than the 2 fps groups (3.367 secs). This suggests that confidence of a response is related to higher FR, thus less degraded imagery. There was no effect for SR on the rating measures.

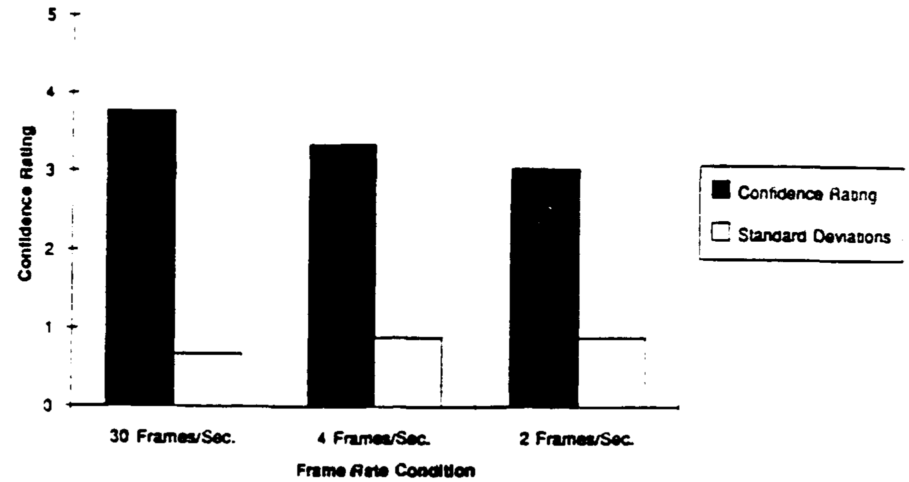
Both CR and CRT results were similar when the control group was included in the analysis. The effect of FR on both CR ($F = 6.77$, $p < .01$) and CRT ($F = 13.64$, $p < .0003$) was significant. The mean control CR was 3.77 and the mean control CRT was 2.900 secs. Post hoc analyses showed that the control group differed from both FR conditions for CR but not CRT. Thus, while the control CRT was actually slower than in the 4 fps groups (2.467 secs), it was not significantly so. There was no effect for SR on the rating measures when the control group was added into the analysis.

Image quality ratings were influenced by SR ($F = 6.99$, $p < .009$) and FR ($F = 6.05$, $p < .01$). There was no interaction. The SR results were counter-intuitive, however. Full resolution groups rated imagery lower (2.2) than those in half resolution groups (2.7). The FR effect was in the expected direction, with a mean 2.3 rating for 2 fps and 2.6 for 4 fps.

There was also a significant effect for SR ($F = 7.90$, $p < .005$) and FR ($F = 6.83$, $p < .009$) when the control group was added into the analysis. Controls rated the image quality at 3.1. Post-hoc analyses showed that the control group differed from both the full and half resolution conditions. For FR, the controls differed from the 2 fps condition (2.3), but not the 4 fps condition (2.6). The means and standard deviations for these data are shown in Figure 4.

Figure 4. Means and Standard Deviations for Confidence and Imagery Quality Ratings is Target Detection - Baseline

Confidence Rating



Confidence Rating Time

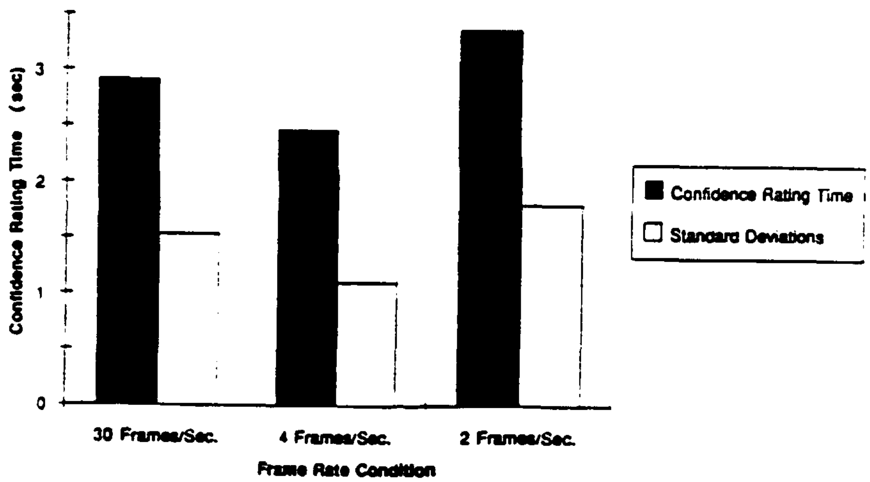
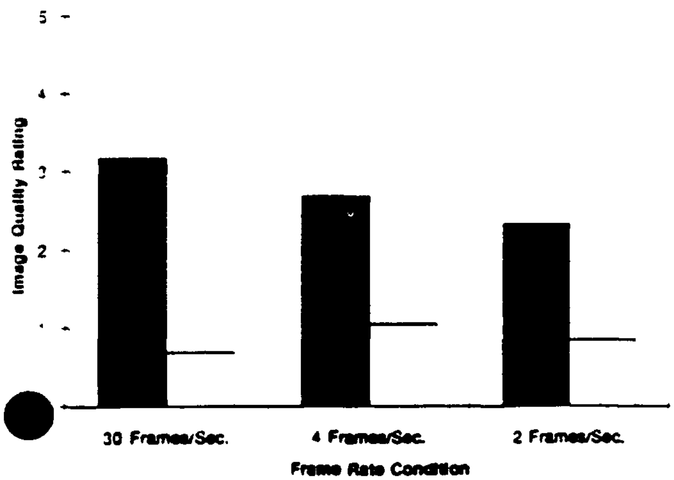
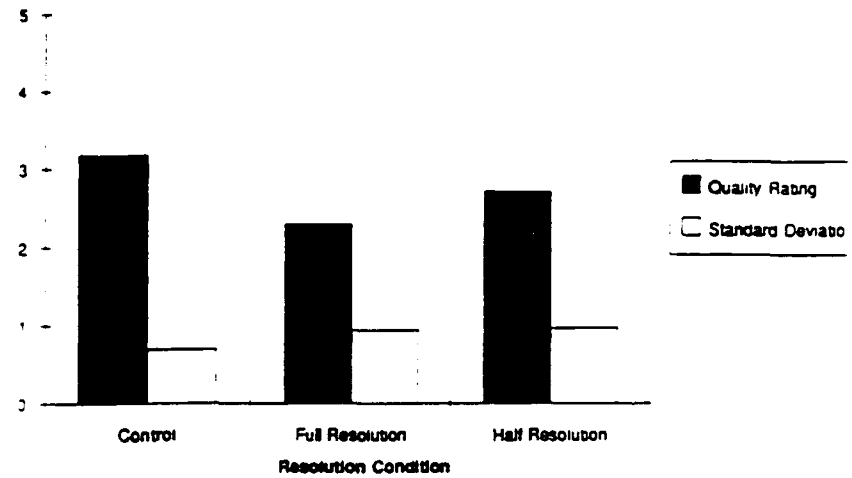


Image Quality





6.1.2.2 Target Recognition

Reaction Time. The effect of FR on RT was highly significant ($F = 19.73$, $p < .001$) providing support for the influence of higher FR on faster responses. RT for the 4 fps clips was 4.650 secs: 5.883 secs for 2 fps clips. These results follow the same pattern found in target detection. SR had no effect on RT (4.950 secs for full resolution. 5.317 secs for half resolution.

The RT results were also significant when the control group was added to the analysis ($F = 12.93$, $p < .0004$). This supports the trend for higher FRs resulting in quicker response times. RT for the control group was 4.133 secs.

Errors. The effect of FR on the percentage of targets correctly recognized was significant ($F = 4.22$, $p < .04$). The results were not what was expected. however. The 2 fps groups made fewer errors, recognizing targets more frequently (92%) than the 4 fps groups (77%). There was no effect for resolution.

There was no significant difference in performance when the control group was added to the analysis. Controls were able to recognize targets 75% of the time. The means and standard deviations for these data are shown in Figure 5. Again. there was also no effect for resolution.

The effect of FR on the number of time-outs was significant ($F = 6.49$, $p < .01$). The percentage of time-outs was highest in the 2 fps groups (49%). A post-hoc comparison showed that this condition differed significantly from both 4 fps and control (30 fps) groups. Table 2 illustrates the mean percentages of time-outs for FR and SR.

Table 2. Mean Percent of Time-Outs During Target Recognition

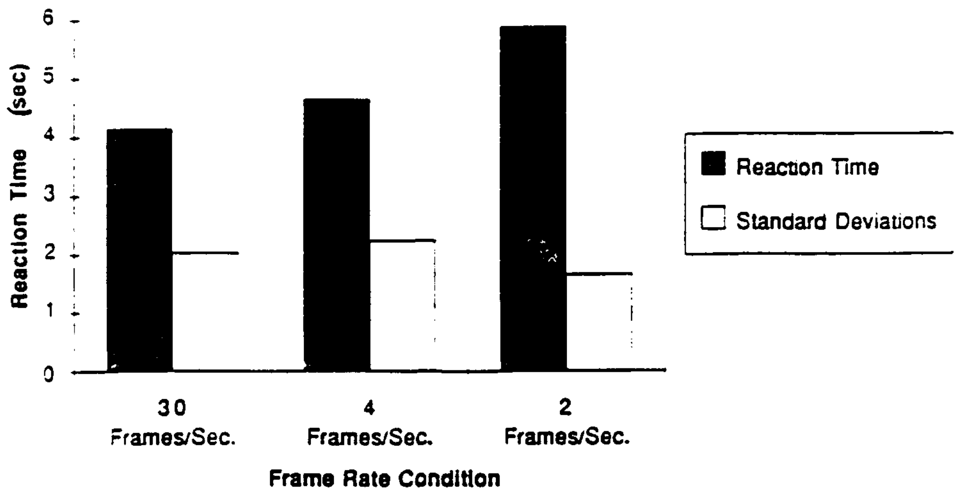
	Percent
Frame Rate	
30	21.67
4	23.00
2	49.00
Spatial Resolution	
Control	21.67
Full	37.00
Half	35.00

With the exception of 2 fps, these percentages are much lower than in the detection task. This reduction is probably due to the nature of the task, in that recognition was a more defined task. Operators knew a target of interest was there. They had to make an identification decision about that target, given a plausible set of target candidates.

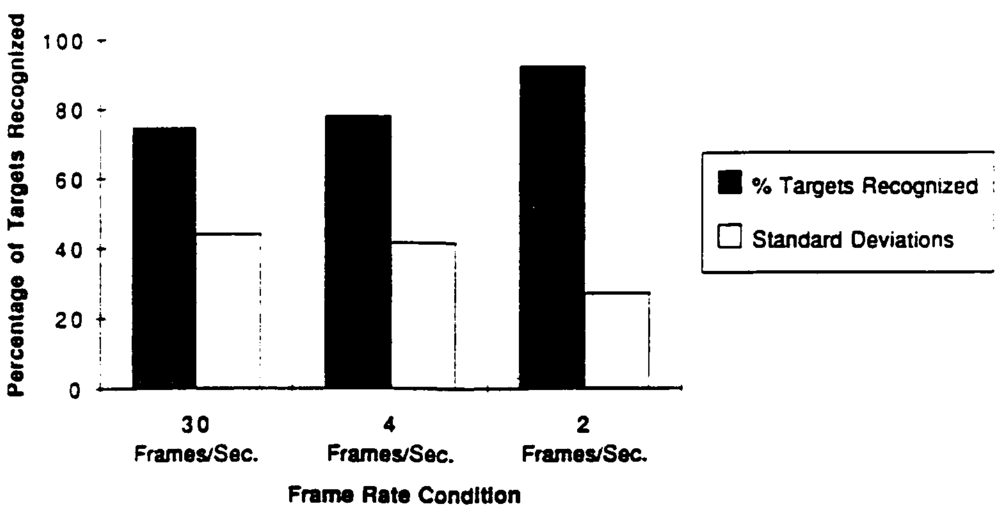
Ratings. The influence of FR was observed in both confidence measures. CR was highly significant ($F = 13.83$, $p < .0003$). Mean ratings for 2 fps and 4 fps were 3.1 and 3.4, respectively. The FR and resolution interaction was also significant ($F = 6.54$ $p < .01$). There was no difference in ratings between half resolution clips at 2 fps (3.2) and 4 fps (3.2). However, for full resolution clips, FR is directly related to CR, with lower FR (2 fps) associated with lower confidence ratings (3.1) and higher FR (4 fps) associated with higher confidence rating (3.7). CRT was also significant ($F = 5.36$, $p < .02$) for FR, indicating that higher FR (4 fps) is associated with faster CRT (3.00 secs) and lower FR (2 fps) is associated with slower CRT (4.617 secs).

Figure 5. Means and Standard Deviations for Reaction Times and Errors in Target Recognition - Baseline

Reaction Time



Errors



FR also influenced rating results when the control group was compared with the experimental groups. There was a significant effect for FR on CR ($F = 12.49$, $p < .0005$) with the mean CR for the control group at 3.3. The interaction between FR and SR was also significant ($F = 5.91$, $p < .01$). However, the pattern was not linear. From the results on CR above, we should expect a higher CR for the control groups, but this was not the case. The mean CR was 3.3, compared to 3.4 for the 4 fps group. The effect of FR on CRT in the control group was also significant ($F = 6.98$, $p < .009$) and linear. As expected, the control group, which had the highest FR (30 fps), had the quickest CRT (2.750 secs).

There was a significant effect for SR on image quality rating ($F = 4.98$, $p < .02$). However, the ratings were contrary to what was expected. Half resolution resulted in higher ratings (3.1) when compared to full resolution (2.8).

SR also had a significant effect ($F = 5.58$, $p < .01$) when the control group was added to the analysis. The mean rating for the controls was 3.3, and the highest rating as predicted when compared to half (3.1) and full (2.8) resolution. These results are shown in Figure 6.

6.1.2.3 Target Designation

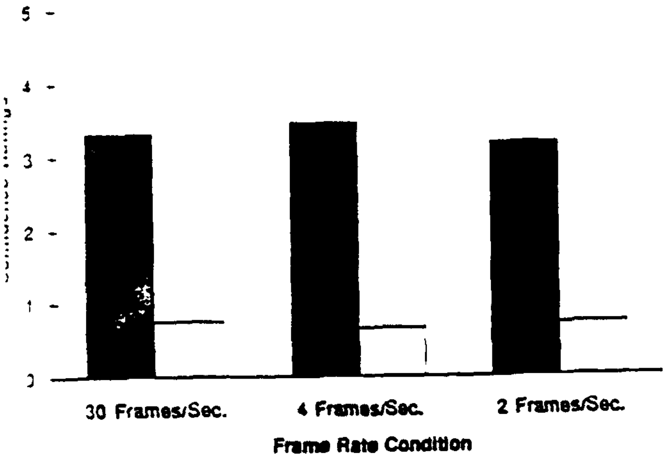
Reaction Time. The effect of FR on RT is highly significant ($F = 68.37$, $p < .0001$), following the same pattern that occurred in the other two tasks. Higher FR (4 fps) resulted in faster responses (3.400 secs) and lower FR (2 fps) resulted in slower responses (5.150 secs). SR had no effect on RT, and there was no interaction.

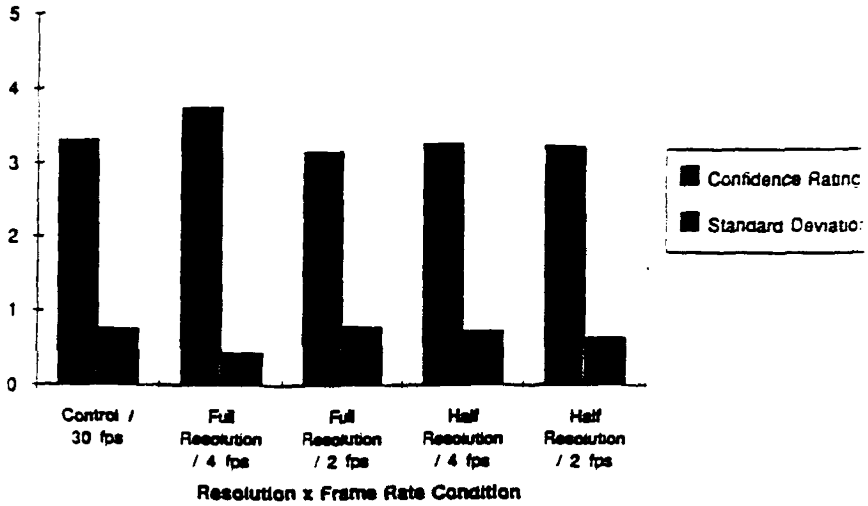
The effect of FR on RT when the control group was added into the analysis was also significant ($F =- 69.13$, $p < .0001$) with the mean RT for the control group at 3.567 secs. Posthoc comparisons indicated that all 3 groups (2, 4, and 30 fps) differed significantly.

Errors. Neither FR nor SR had any effect on the percent of correct designations. Higher FRs (4 fps) had 76% correct designations; lower FRs (2 fps) had 80%. Full resolution resulted in correct designations 82% of the time; 73% for half resolution. Neither variable was significant when the control group was analyzed, although the mean percentage (97%) was very high. Figure 7 shows RT and error data.

Figure 6. Means and Standard Deviations for Confidence and Imagery Quality Ratings in Target Recognition - Baseline

Confidence Rating





Confidence Rating Time

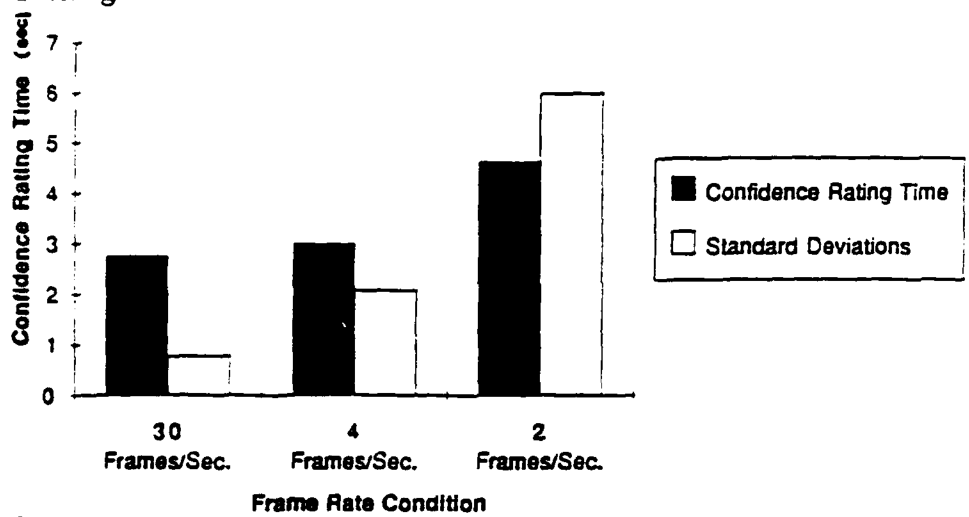


Image Quality

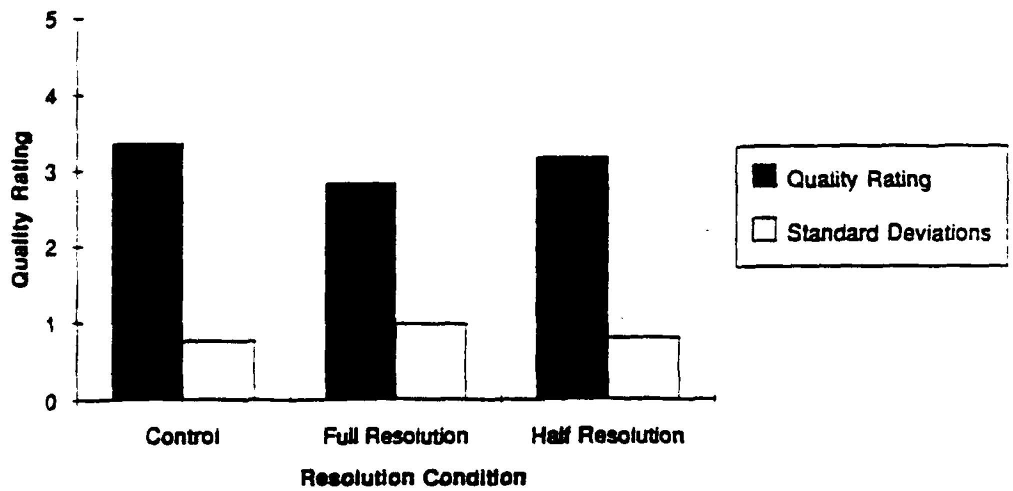
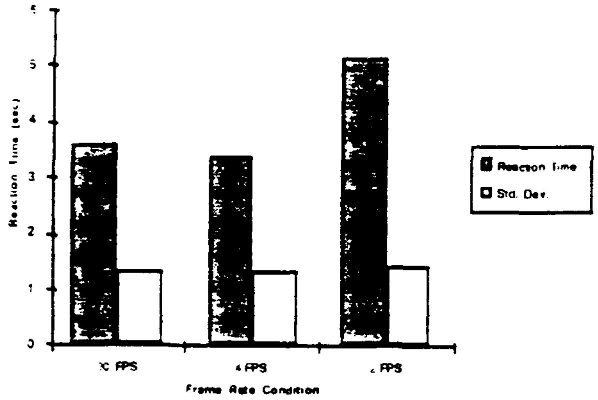


Figure 7. Means and Standard Deviations for Reaction Times, Errors and Image Quality Target Designation - Baseline

Reaction Time



Errors

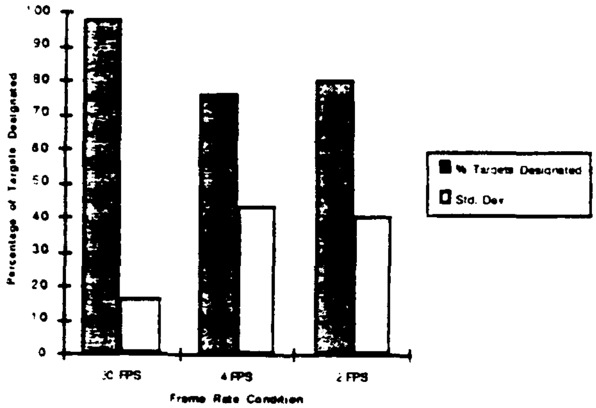
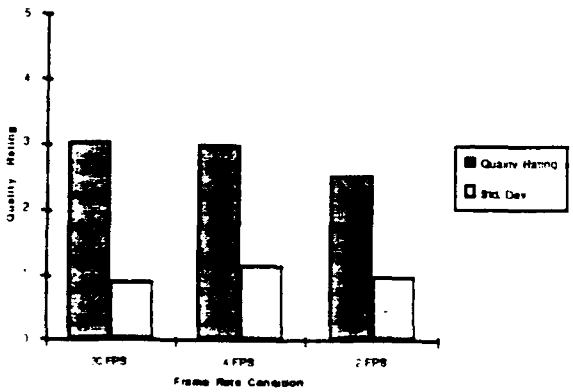


Image Quality



There were no significant effects for either FR or SR on the percentage of time-outs during the designation task. Table 3 illustrates the mean percentages of time-outs for FR and SR.

In this task, the percentages of time-outs are considerably lower than in the detection or recognition tasks. This result shows a somewhat linear pattern across tasks, with the fewest timeouts in the designation task. This task was the most constrained and defined of the 3 tasks examined. To designate a target, participants were told what the target was (recognition decision provided). By extension, we assume detection. Therefore, participant performance involved a relatively simple visual - motor coordination task which was relatively easy to perform within the 10 second parameter. This was due to the fact that mission parameters were not dynamic, thus making the control mapping between the control device and the display a 1:1 relationship.

Table 3. Mean Percent of Time-Outs During Target Designation

	Percent
Frame Rate	
30	2.50
4	5.00
2	18.00
Spatial Resolution	
Control	2.50
Full	9.00
Half	16.25

Ratings. Both FR and SR had an effect on image quality rating for this task. FR was significant ($F = 4.4$, $p < .03$) with higher FR (4 fps) associated with higher image quality ratings (2.9), and lower FR (2 fps) associated with lower ratings (2.5). SR was significant ($F = 9.67$, p

< .002) but with unusual results, similar to those obtained for recognition and detection. Higher resolution (full) was associated with lower ratings (2.4) and lower resolution (half) associated with a higher rating (3.0).

A similar pattern was found for FR when the control group was included in the analysis. FR was significant ($F = 4.7$, $p < .03$) with the mean rating for controls at 3.0 following the same linear relationships of FR to rating as above. A post-hoc comparison showed that the control group differed from both levels of FR. SR was also significant ($F = 10.62$, $p < .001$), but the mean rating for controls (3.0) was not different from the half resolution condition (3.0). This was confirmed in post-hoc comparisons with only full resolution (2.4) differing from both the control and half resolution groups. Figure 7 shows the means and standard deviations for these data.

6.1.3 Experimental Study Results

Data from this study were collected on the same three tasks described in the baseline study: Detection, Recognition, and Designation. In this study we also evaluated experience effects on performance. Each of the tasks was examined with respect to the same independent and dependent variables as in the baseline study. The five dependent variables include objective measures of reaction time (RT) and error rates, subjective ratings of confidence and image quality, and experience with target acquisition tasks.

The following section first describes the sensitivity analysis with regard to tactical decision making skill for target detection.. A brief review of Signal Detection Theory and the Receiver's Operating Characteristic (ROC) curve as the theoretical framework for this type of performance and the implications for data link design requirements is provided. This is followed by a discussion of the results for each task, first summarized over all participants. The effect of experience is examined by comparing NATC personnel to Vitro employees. Finally, comparisons among tasks on the five dependent variables are discussed.

6.1.3.1 Detection

Detection Sensitivity. The analysis of an operator's ability to detect objects of interest in real-time environments often includes the ability to discriminate a signal (target) from a noisy environment. This is especially true when the image quality of sensor imagery is degraded due to data reduction. In such tactical decision tasks, the operator must decide between two discrete states of the world: 1) a signal (target) is present, or 2) a signal (target) is not present. Signal detection theory (SDT) provides a framework in which to analyze how effectively such choices are made. The decision made between the two alternatives listed above have been found to be affected by the decision-maker's bias or criterion level for selecting one alternative over the other [20], [21]. SDT research has shown that an individual's response can be influenced by such factors as expectation and motivation with regard to the objects of interest, and the probability and utility of an occurrence. A graphical technique for plotting such probabilities against a defined decision criterion (called beta) is called the Receiver's Operating Characteristic (ROC) curve. A ROC analysis depicts an operator's sensitivity for a discrimination task and can illustrate the trade offs when the probability of a target being present is plotted against the probability of an incorrect detection when no target is present (a false alarm). Figures 8 and 9 illustrate ROC curves.

Detection sensitivity (d') plots the probability of a correct detection or Hit (P(Hit)) against the probability of an incorrect detection or False Alarm (P(FA)) in order to better understand an operator's ability to discriminate true targets from false ones. It was assumed that FR and SR, when degraded, would affect detection sensitivity performance.

As expected, significant effects for both FR and SR were observed when examining d' differences. The source of this difference was significant for the P(Hit) for both FR ($F = 11.97$, $p < .0001$) and SR ($F = 11.15$, $p < .0001$). There were no significant differences in P(FA) performance. Regardless of imagery condition, participants' biases toward false alarm rates were consistent. The mean percentages for Hits and FAs and the corresponding ROC curve are shown in Figure 8 for FR and Figure 9 for SR.

Post-hoc comparisons showed that the control group (90%) differed from 4 fps (73%) and 2 fps (68%). There was no significant difference between 2 and 4 fps. Similarly for SR, only the control group (90%) differed significantly from either full (69%) or half (71%) resolution. Controls exhibited more sensitivity with higher hit and lower false alarm rates. The controls' 90%

criterion level in both conditions is equivalent to the level identified as an expected mission performance level [1]. The experimental groups had Hit rates ranging from 68 - 73%.

There were no significant differences in d' for experience.

Figure 8. Sensitivity for Detecting Targets with Frame Rate as a Factor

Control Group (30 fps)

	"YES" Response	"NO" Response
Target Present	(Hit) .91	(Miss) .09
Target Absent	.19 (FA)	.81 (CR)

4 Frames/Sec

	"YES" Response	"NO" Response
Target Present	(Hit) .73	(Miss) .27
Target Absent	.24 (FA)	.76 (CR)

2 Frames/Sec

	"YES" Response	"NO" Response
Target Present	(Hit) .68	(Miss) .32
Target Absent	.21 (FA)	.79 (CR)

ROC Sensitivity Curve For Frame Rate

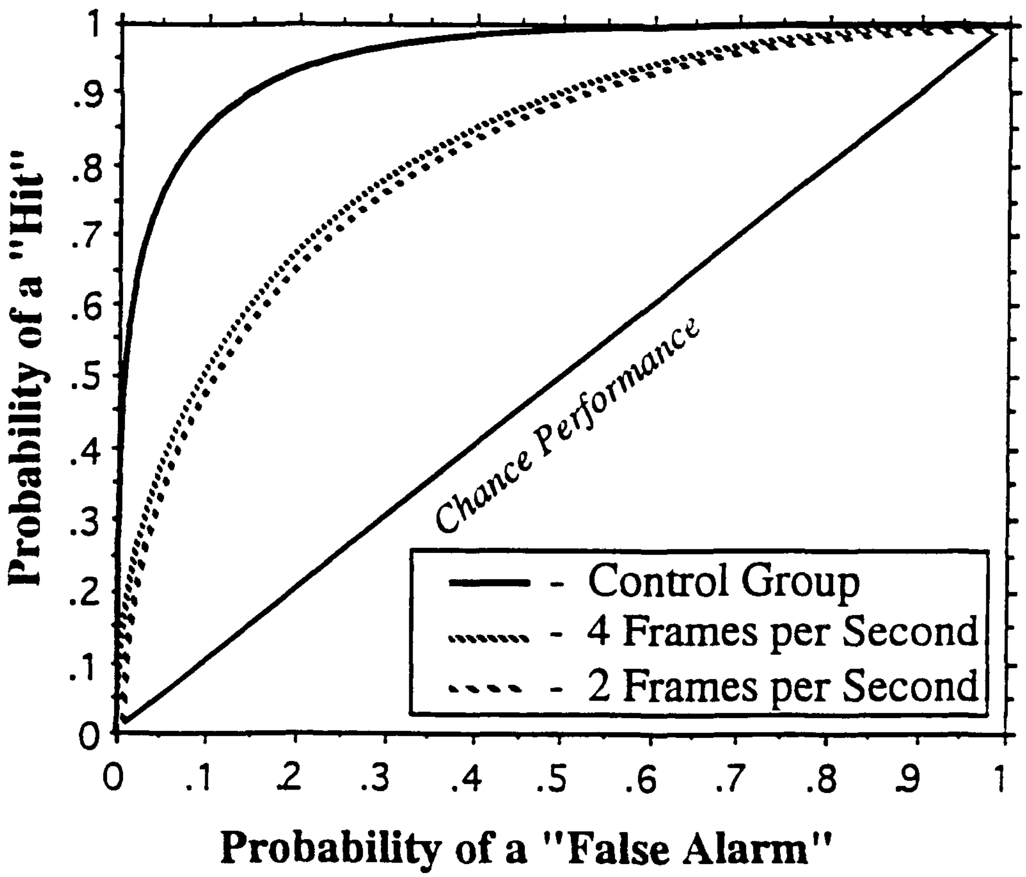


Figure 9. Sensitivity for Detecting Targets with Resolution as a Factor

Control Group (Full/Uncompressed)

	"YES" Response	"NO" Response
Target Present	(Hit) .91	(Miss) .09
Target Absent	.19 (FA)	.81 (CR)

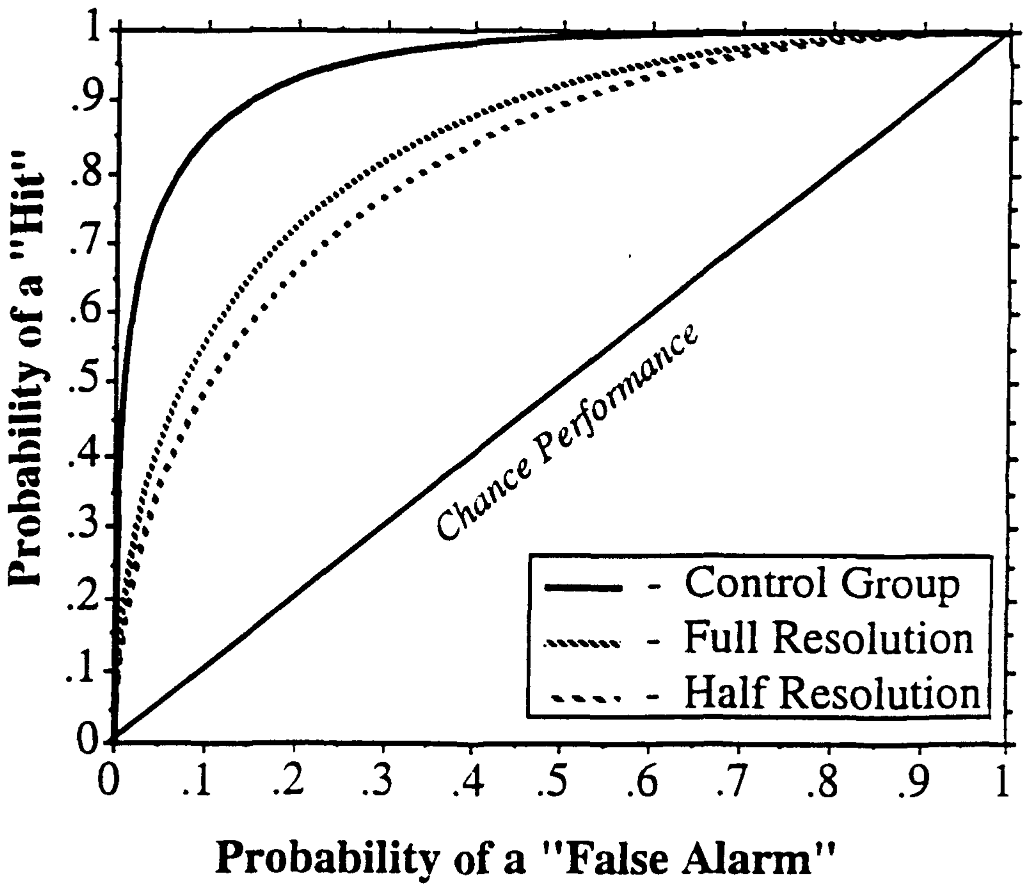
Full Resolution

	"YES" Response	"NO" Response
Target Present	(Hit) .70	(Miss) .30
Target Absent	.24 (FA)	.76 (CR)

Half Resolution

	"YES" Response	"NO" Response
Target Present	(Hit) .71	(Miss) .29
Target Absent	.22 (FA)	.78 (CR)

ROC Sensitivity Curve For Resolution



Reaction Time. The effect of FR on RT was highly significant ($F = 13.3$, $p < .0003$), indicating that a higher FR results in faster responses. RT for the 2 fps clips was 5.333 secs, compared with 4.750 secs for the 4 fps clips.

The performance of the control group, who viewed full resolution clips at 30 fps, was 3.700 secs. In post-hoc comparisons, this group was found to be significantly faster than both the 2 and 4 fps groups.

SR had no effect on RT (4.950 secs and 5.100 secs for full and half resolution, respectively), and there was no interaction between SR and FR.

Errors. Neither FR nor SR had any effect on the percent of correct detections, which was 75% for both conditions. There was a small interaction between FR and SR ($F = 4.0$, $p < .04$). For half resolution clips, correct detection rates for 2 fps and 4 fps, respectively, are 73% and 79%, indicating better detection rates at the higher FR. However, the rates of correct detections for full resolution clips were 79% and 71%, for 2 fps and 4 fps, respectively. In this case, the higher FR is associated with poorer performance.

This interaction was also obtained with the control group ($F = 4.2$, $p < .0401$). The control group detected 83% of the targets. Means for RTs and errors are shown in Figure 10.

There were no significant effects for either FR or SR on the percentage of clips where participants were unable to decide whether a target was present or not. Table 4 shows the mean percentages of time-outs during the detection task. The trend in these percentages is similar to what was observed in the baseline data. Participants timed-out about half of the time, with the least amount of time-outs in the control group.

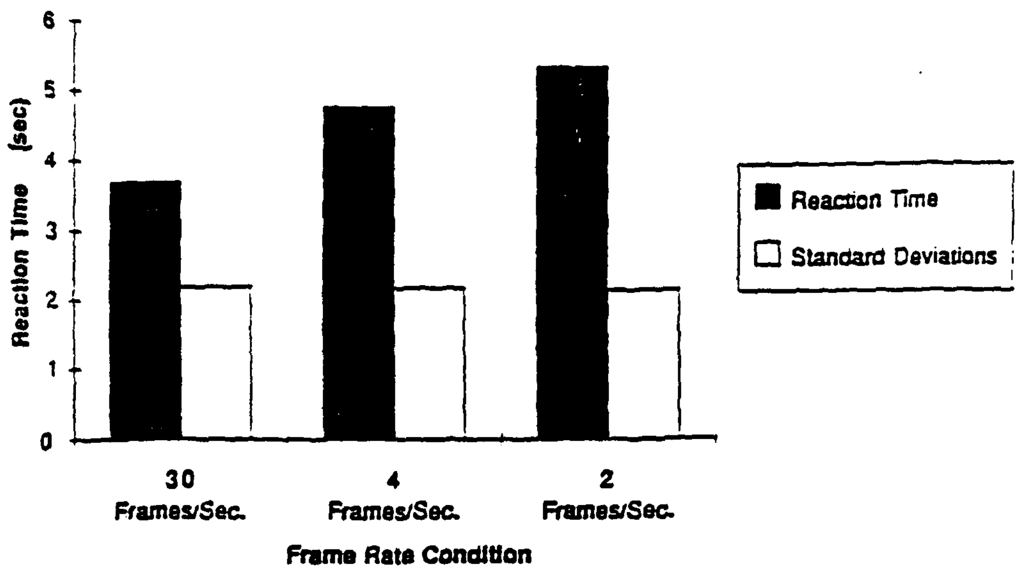
Table 4. Mean Percentage of Time-Outs in Target Detection

	Percent
Frame Rate	
30	45.33
4	51.50
2	55.17
Resolution	
Control	45.33
Full	53.33
Half	53.33

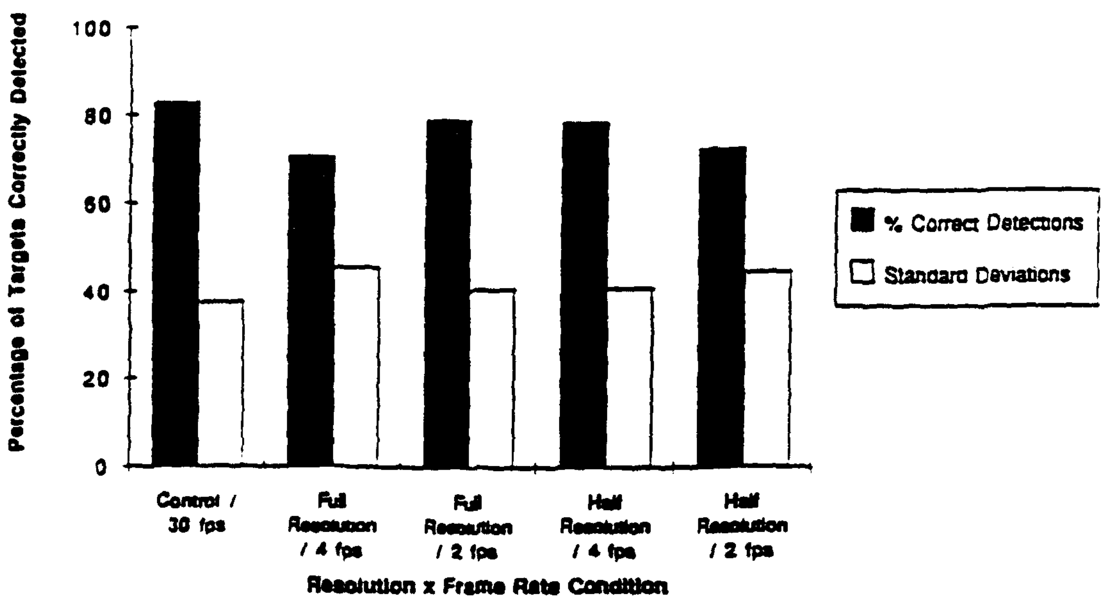
Ratings. There was no effect of either FR or SR on confidence rating (CR). There was an effect of FR on CRT, with the higher FR (4 fps) associated with faster CRT (2.633 secs). The lower FR (2 fps) resulted in slower ratings (3.200 secs).

Figure 10. Means and Standard Deviations for Reaction Times and Errors in Target Detection

Reaction Time



Errors



This same pattern was obtained when the control group was included in the analyses. There were no effects on CR, but a significant effect of FR on CRT ($F = 6.0$, $p < .0147$). The mean CRT for the control group was 3.200 secs.

There was no effect of FR or SR on image quality rating. These data are shown in Figure 11.

Experience. There was no effect of experience on RT, error rate or image quality rating. However, both confidence measures indicate that NATC personnel are more confident. CR was highly significant ($F = 17.1$, $p < .0001$); the mean rating for NATC personnel was 3.7, compared to 3.3 for Vitro employees. Results for the CRT measure were similar ($F = 9.7$, $p < .002$). Response times for NATC and Vitro personnel were 2.617 secs and 3.083 secs, respectively.

The same pattern was also obtained with the control group data. There were no differences in RT, error rate or image quality rating, but both confidence measures were significant. CR was highly significant ($F = 29.4$, $p < .0001$). The mean rating for NATC personnel in the control group was 4.2, compared with 3.7 for the Vitro group. Significant differences in CRT were also obtained between the NATC and Vitro controls ($F = 11.0$, $p < .0009$). Response times for NATC personnel was 2.483 secs, compared to 3.783 secs for Vitro controls. These data are shown in Figure 12.

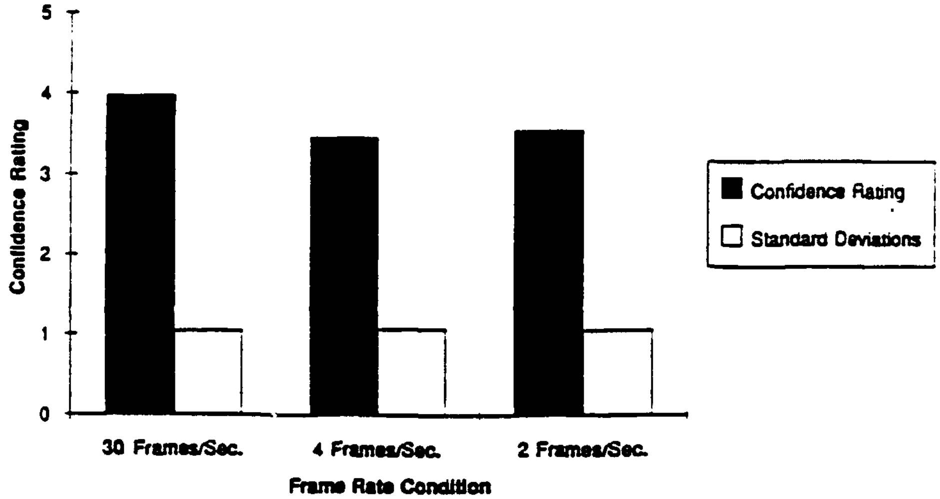
6.1.3.2 Recognition

Reaction Time. The effect of FR on RT was significant ($F = 6.9$, $p < .0086$), indicating again that a higher FR results in faster responses. RT for the 2 fps clips was 6.417 secs, compared with 5.983 secs for the 4 fps condition. SR had no effect on RT (6.283 secs and 6.083 secs for full and half resolution, respectively), and there was no interaction between SR and FR. These results parallel those for target detection.

The results are similar when the control group is included in the analyses. The effect of FR on RT was significant ($F = 6.3$, $p < .0122$). Again, the trend is that higher FRs result in faster responses. RT for the 2 fps, 4 fps, and 30 fps clips were 6.417, 5.983, and 4.650 secs, respectively. There was no effect of resolution, and there was no interaction.

Figure 11. Means and Standard Deviations for Confidence and Imagery Quality Ratings in Target Detection

Confidence Rating



Confidence Rating Time

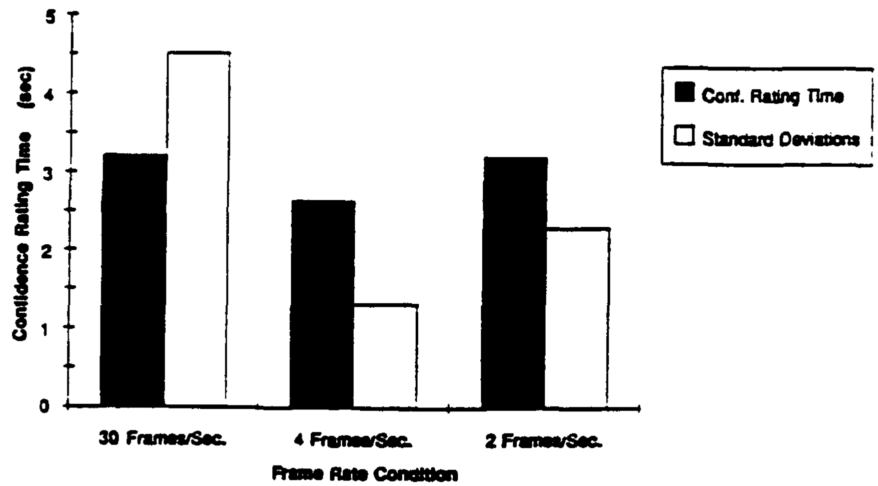
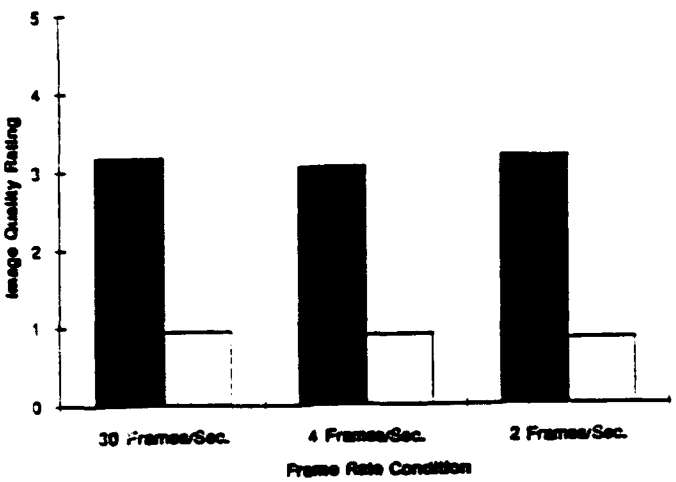


Image Quality



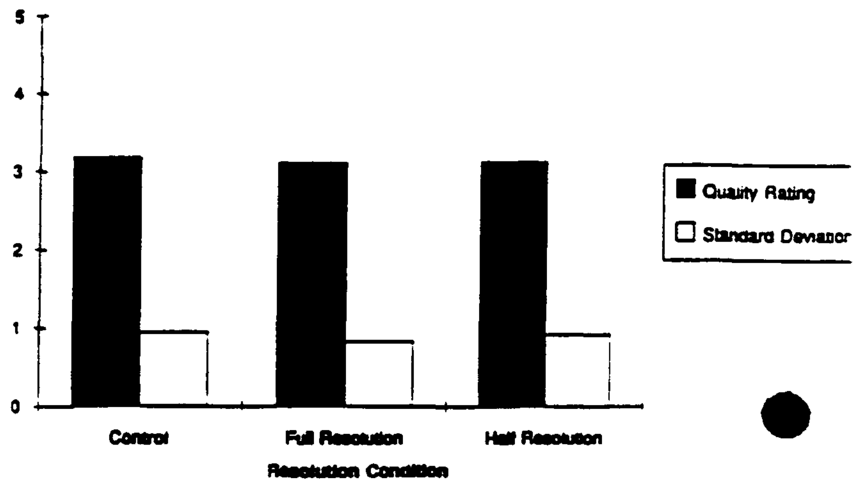
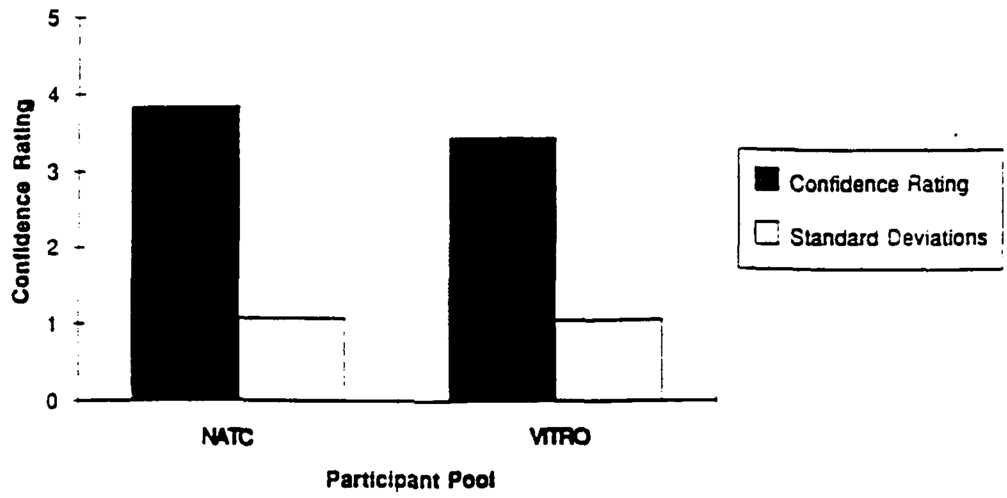
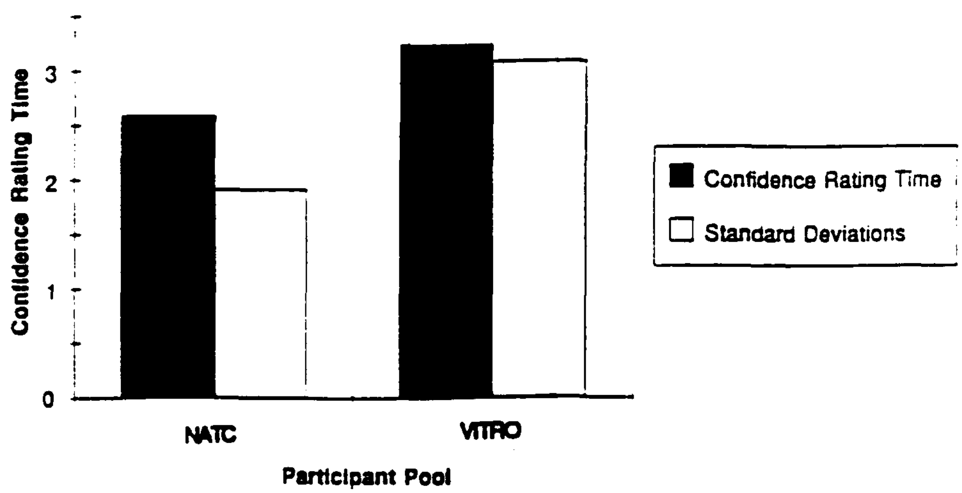


Figure 12. Means and Standard Deviations for Experience in Target Detection

Confidence Rating



Confidence Rating Time



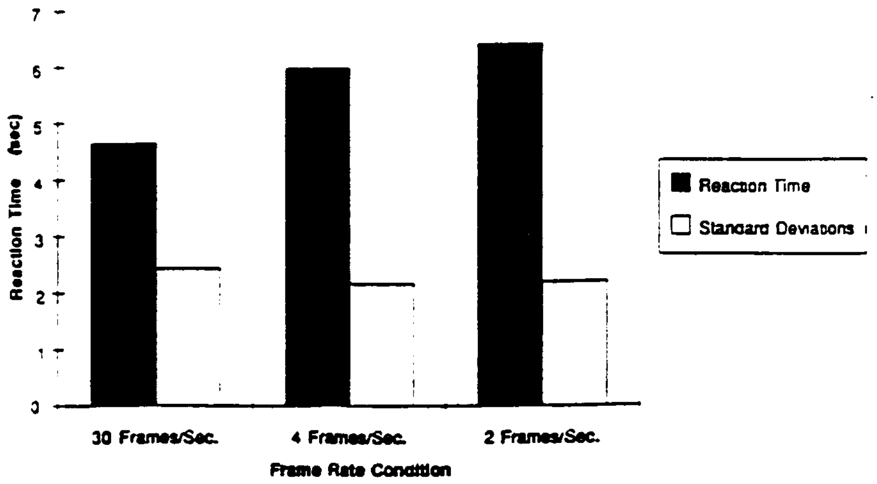
Errors. The effect of SR on error rate was marginally significant. ($F = 3.9$, $p < .0502$). The percent of correct detections was 75% at full resolution and 67% at half resolution. As in the detection task, FR had no effect on error rate.

The effect of SR on error rate was also significant when the control group data was included in the analysis. ($F = 4.2$, $p < .0421$). The percent of correct detections was highest when clips were not compressed (84%). Post-hoc comparisons indicate that the difference between half resolution/compressed clips and full resolution/no compression clips was significant. Means for RTs and errors are shown in Figure 13.

There were no significant effects for either FR or SR on the percentage of clips where participants were unable to decide whether a target was present or not. Table 5 shows the mean percentage of time outs during the recognition task. These data are consistent with the pattern found in the baseline study.

Figure 13. Means and Standard Deviations for Reaction Times and Errors in Target Recognition

Reaction Time



Errors

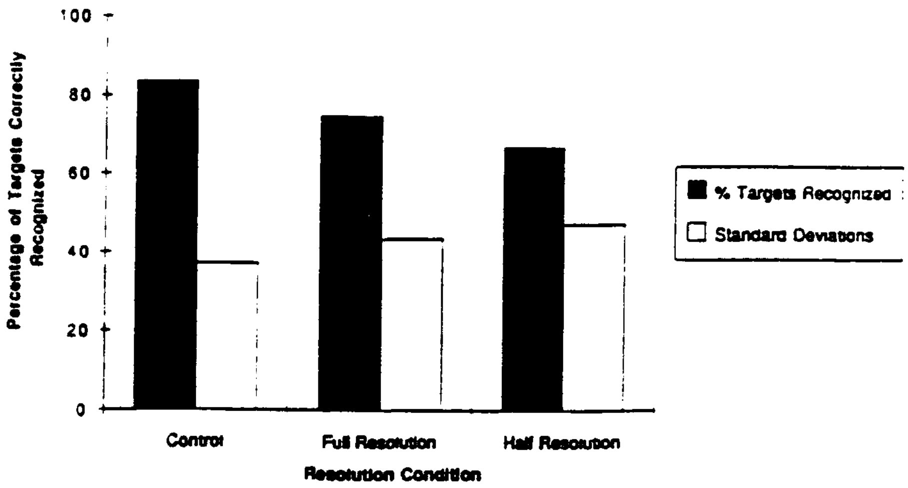


Table 5. Mean Percentage of Time-Outs in Target Recognition

	Percent
Frame Rate	
30	32.50
4	34.00
2	42.00
Resolution	
Control	32.50
Full	37.25
Half	38.75

Ratings. The influence of FR was seen in both confidence measures. CR was marginally significant ($F = 3.8$, $p < .0513$). Mean ratings for 2 fps and 4 fps were 3.5 and 3.3, respectively. The interaction of FR and SR was also significant. ($F = 10.6$, $p < .0012$). CRT was highly significant ($F = 24.9$, $p < .0001$). As in the detection task, the higher FR (4 fps) was associated with faster CRT (2.033 secs) and lower FR (2 fps) with slower ratings (2.617 secs).

Similar results are obtained when control group data was included in the analysis. CR was marginally significant ($F = 3.5$, $p < .0603$). Mean rating for the 30 fps clips was 3.8. The interaction of FR and SR was also significant ($F = 9.8$, $p < .0018$). For half resolution clips, FR is directly related to CR, such that the lower FR (2 fps) results in lower confidence (3.3); the higher FR (4 fps) results in higher confidence. However, in the full resolution conditions, the relationship is not linear. Ratings for the 2, 4, and 30 fps conditions are 3.6, 3.2, and 3.8, respectively.

The effect of FR on CRT in the control group was highly significant ($F = 27.2$, $p < .0001$) and linear. Again, higher FRs were associated with lower CRTs. Mean CRTs for the 2, 4, and 30 fps clips were 2.606, 2.033, and 1.987 secs, respectively.

As in the target detection task, there was no effect of FR or SR on image quality rating. These data are shown in Figure 14.

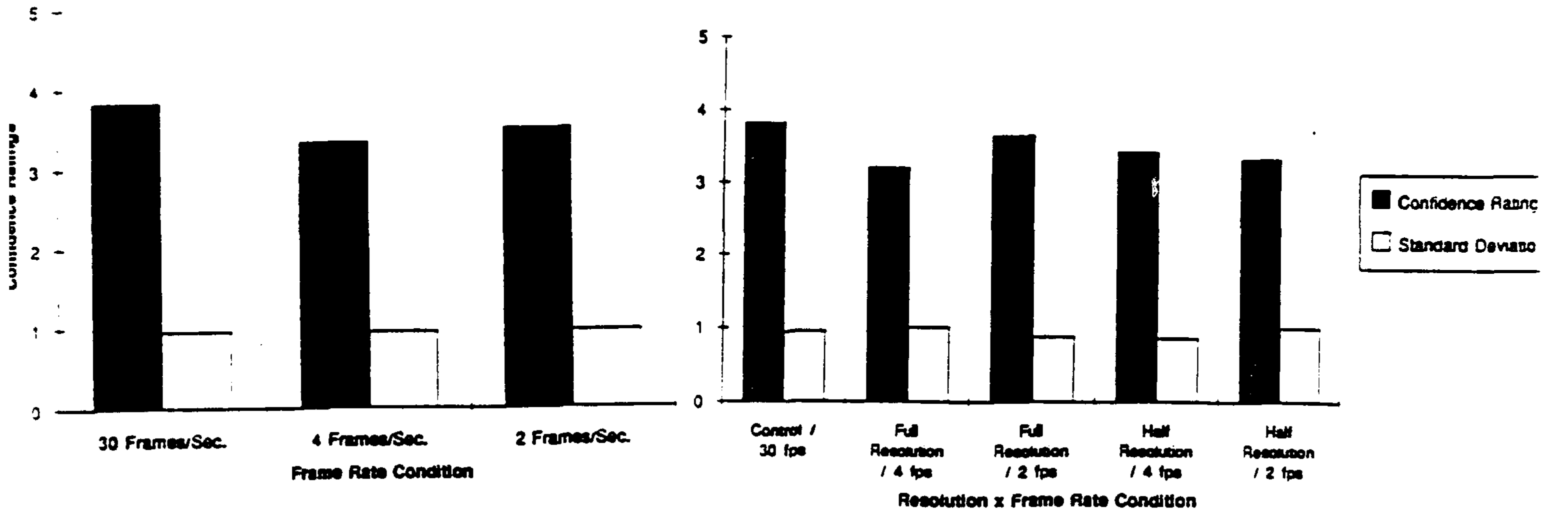
Experience. There was no effect of experience on RT, error rate or image quality rating. Both confidence measures again indicated that NATC personnel are more confident. CR was highly significant ($F = 21.7$, $p < .0001$); the mean rating for NATC personnel was 3.6, compared to 3.2 for Vitro employees. Results for the CRT measure were similar ($F = 12.4$, $p < .0005$). Response times for NATC and Vitro personnel were 2.050 secs and 2.483 secs, respectively.

There was no effect of experience on either error rate or image quality rating when control group data were analyzed. However, in contrast to the experimental groups, a significant difference was obtained between NATC and Vitro personnel on recognition RT ($F = 4.5$, $p < .0337$). NATC personnel were faster than Vitro employees (5.717 secs compared to 5.967 secs, respectively). Corresponding means for the experimental groups were 6.283 (NATC) and 6.133 secs (Vitro).

As in the analysis of the experimental groups, both confidence measures indicated that NATC control group personnel are more confident. CR was highly significant ($F = 41.6$, $p < .0001$); the mean rating for NATC personnel was 4.0, compared to 3.5 for the Vitro control group. Results for the CRT measure were comparable ($F = 17.6$, $p < .0001$). Response times for NATC and Vitro personnel were 1.833 secs and 2.133 secs, respectively, indicating participants with high confidence ratings are also faster at rating. These results are shown in Figure 15.

Figure 14. Means and Standard Deviations for Confidence and Imagery Quality Ratings in Target Recognition

Confidence Rating



Confidence Rating Time

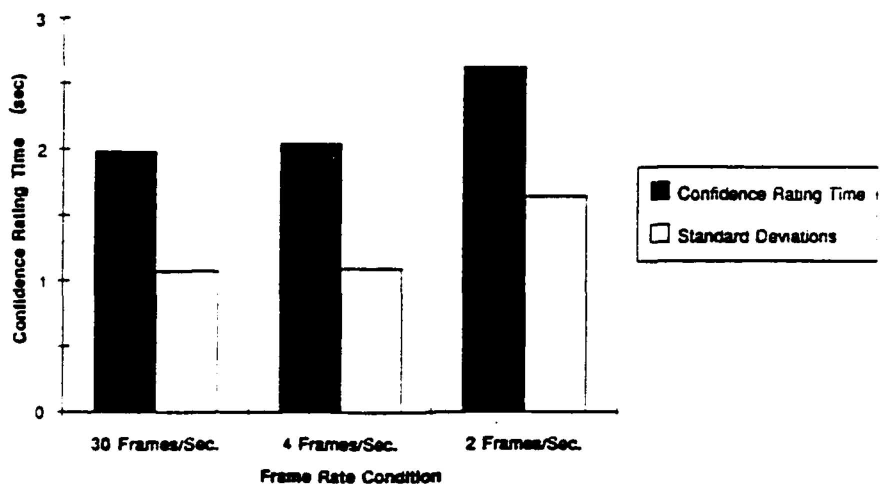
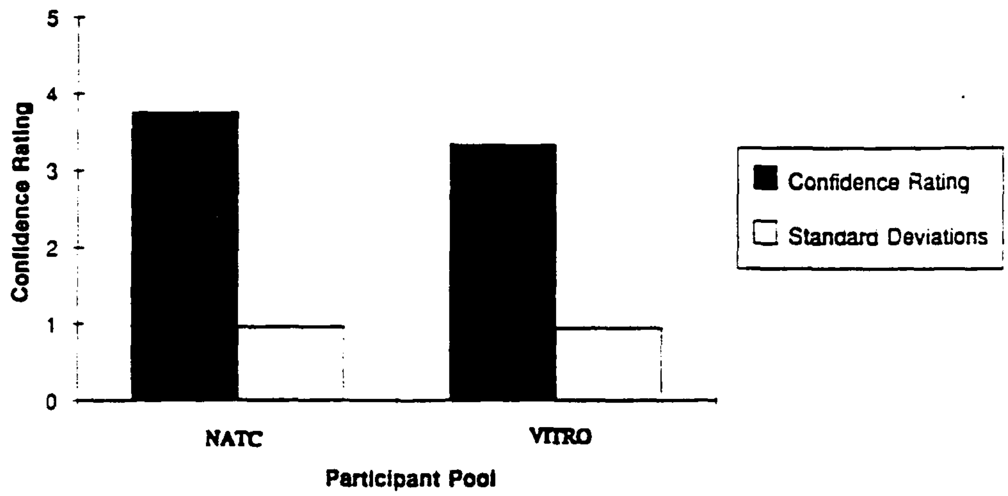
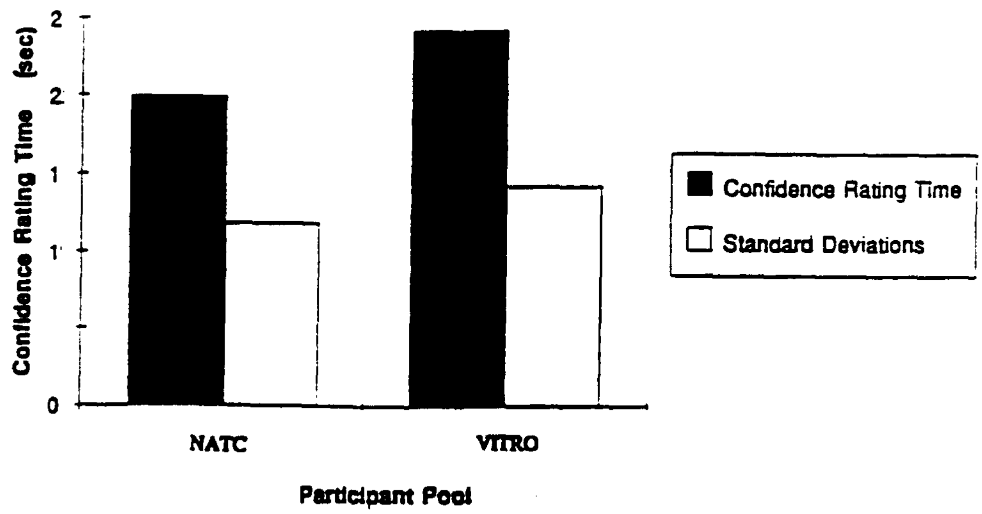


Figure 15. Means and Standard Deviations for Experience in Target Recognition

Confidence Rating



Confidence Rating Time



6.1.3.3 Target Designation

Reaction Time. Consistent with both detection and recognition tasks, the effect of FR on designation RT was significant ($F = 28.2$, $p < .0001$). Higher FR (4 fps) resulted in faster responses (3.983 secs. compared with 4.533 secs for the 2 fps condition). SR had no effect on RT, and there was no interaction. These results parallel those for target detection and recognition. The RT for control group was 3.517 secs. Post-hoc comparisons indicated that all 3 groups (2, 4, and 30 fps) differed significantly.

Errors. The effect of FR on error rate was also significant ($F = 7.4$, $p < .0068$). The percent of correct detections was 75% for the higher (4 fps) rate and 66% for the lower (2 fps) rate. This result is in contrast to both the detection and recognition task results, where FR had no effect on error rate. Correct detection rate for the control group was 89%. As with the RT results, post-hoc comparisons indicated that all 3 groups (2, 4, and 30 fps) differed significantly. Means for RTs and errors are shown in Figure 16.

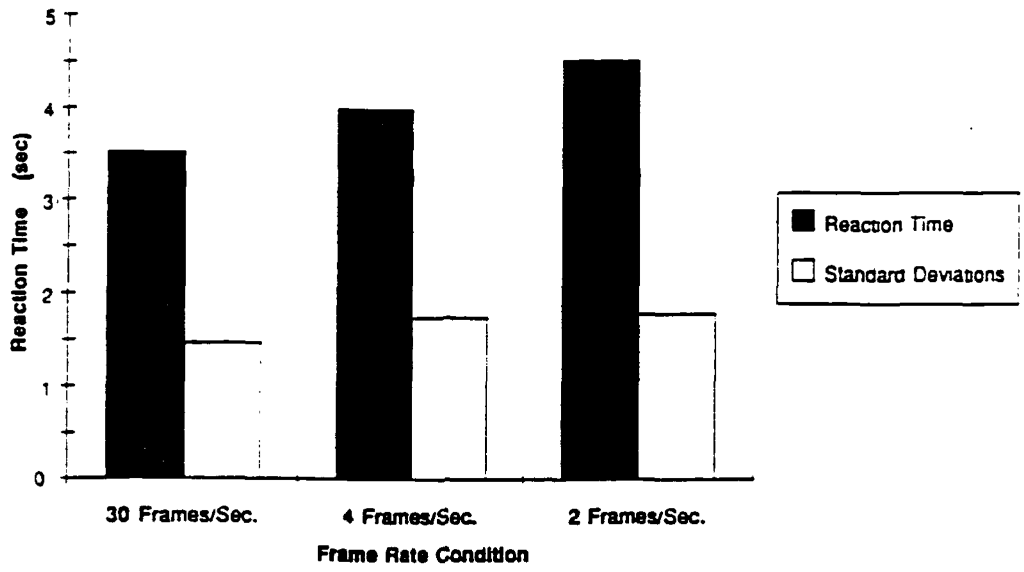
Again, the outcome and pattern of time-out performance in the designation task was consistent with what occurred in the baseline study. The percentage of time-outs during designation was greatly reduced when compared to the other two tasks. Table 6 shows the mean percentages.

Table 6. Mean Percentage of Time-Outs in Target Designation

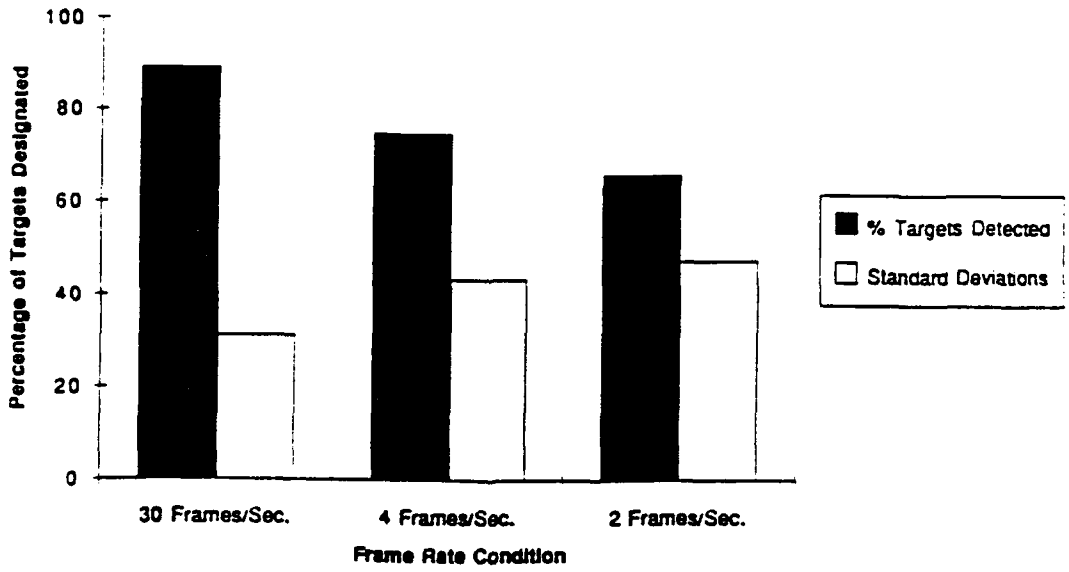
	Percent
Frame Rate	
30	5.25
4	3.00
2	6.75
Resolution	
Control	3.00
Full	6.00
Half	6.00

Figure 16. Means and Standard Deviations for Reaction Times and Errors in Target Designation

Reaction Time



Errors



Ratings. As in the detection and recognition tasks, there was no significant main effect of FR or SR on image quality rating. However, there was an interaction between these variables ($F = 8.5$, $p < .0036$). The ratings for full resolution clips were 3.03 and 3.24, for 2 fps and 4 fps, respectively. In this case, the higher FR is associated with higher quality ratings. However, for half resolution clips, ratings for 2 fps and 4 fps, respectively, are 3.26 and 3.05, indicating perceived better quality at the lower FR. The data for the control group is consistent with the results obtained for the full resolution experimental groups, with a mean rating of 3.3. That is, the highest FR is associated with the highest quality rating. These data are shown in Figure 17.

No confidence measures were collected for this task.

Experience. In contrast to both detection and recognition tasks, the effect of experience on RT was significant for the designation task ($F = 6.9$, $p < .009$). NATC personnel were faster (4.100 secs) than Vitro employees (4.350 secs). This result also obtained within the control groups ($F = 8.7$, $p < .0033$). NATC controls were faster (3.950 secs) than Vitro controls (4.200 secs).

Consistent with both detection and recognition tasks, the effect of experience on error rate and image quality rating was not significant. These results are shown in Figure 18.

There was a significant effect for experience on the percentage of time-outs during the designation task ($F = 5.13$, $p < .02$). Table 7 shows the mean percentage of time-outs for all 3 tasks, across experience.

Table 7. Percentage of Time Outs in Each Condition

Condition	Detection	Recognition	Designation
NATC	49.82	30.27	2.9
Vitro	52.90	40.97	6.94

While the pattern of these data is consistent with the time-out performance reported earlier. in designation tasks we see that less-experienced personnel timed-out over twice as much as the more-experienced personnel.

Figure 17. Means and Standard Deviations Imagery Quality Ratings in Target Designation

Quality Rating

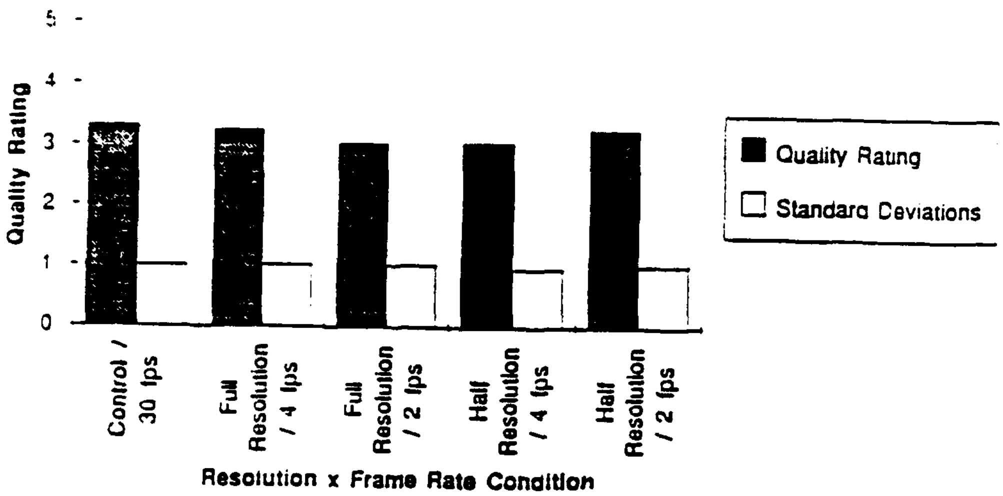
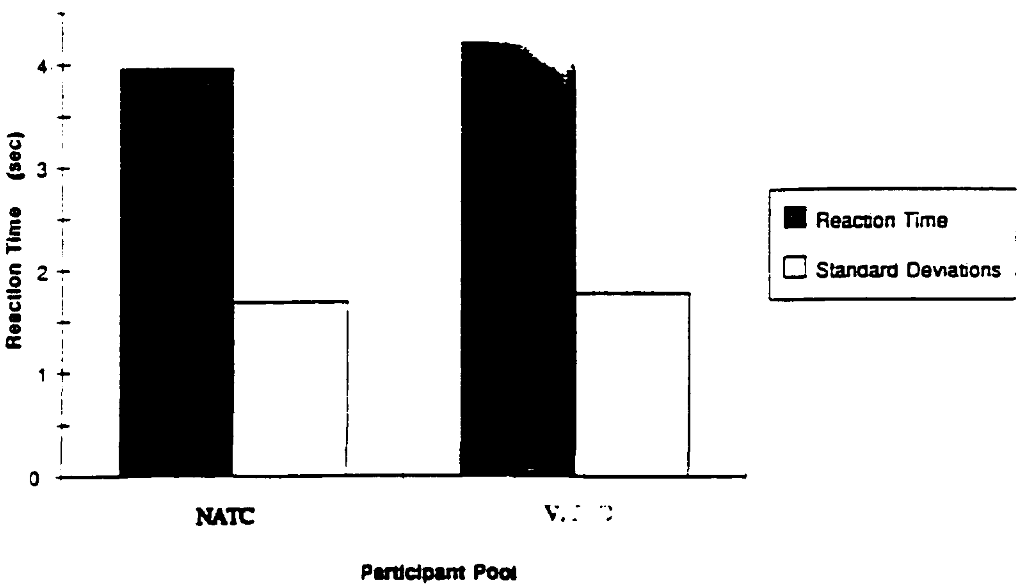


Figure 18. Means and Standard Deviations for Experience in Target Designation

Reaction Time



6.1.4 Experiment One Discussion

6.1.4.1 Baseline Study

The pattern of results across all three tasks examined in the baseline study (detection, recognition, and designation) was similar. In most cases, there were no differences in significance tests due to the inclusion or omission of the control groups.

The only independent variable affecting RT was frame rate. Higher frame rates significantly reduced RT in all 3 tasks. This result suggests that increases in FR will significantly increase task efficiency and, consequently, mission performance.

The effect of frame rate was also significant for the confidence measures examined in both the detection and recognition tasks. (Confidence measures were not collected for designation.) An expected relationship was found, with higher frame rates associated with higher confidence ratings and quicker response times, and lower frame rates associated with lower confidence ratings and slower response times.

Imagery ratings were affected by resolution for all 3 tasks; the effect of frame rate on these ratings was found only for detection and designation.

Unique results that occurred only in the recognition task involved frame rate. This variable was related to both the percentage of errors and the percentage of time-outs. Unexpectedly, fewer errors occurred with slower frame rates. However, the effect of frame rate on time-outs was in the expected direction, with the percentage of time-outs for the slowest group (2 fps) more than twice as high as in the other 2 groups.

6.1.4.2 Experimental Study

Overall, the pattern of results across all three tasks examined in the experimental study and both participant groups (NATC and Vitro personnel) was very consistent. In addition, in most cases, there were no differences in significance tests between analyses that included or omitted the control groups.

For both the detection and designation tasks, identical results were obtained with or without the control groups. For the recognition task, the only exception to this pattern was obtained when the effect of experience on reaction time was examined. In this case, the comparison between NATC and Vitro personnel was not significant within the experimental group, but was significant when the control group was included.

The biggest discrepancy among tasks appears in the percentage of time-outs. The percentages of time-outs are considerably lower than in the detection or recognition tasks. A somewhat linear pattern is observed, with the fewest time-outs in the designation task. This task was the most constrained and defined of the 3 tasks examined. To designate a target, participants were told what the target was (recognition) and by extension, we assume detection. Therefore, participants' performance involved a visual - motor coordination task which was relatively easy to perform within the 10 second parameter.

When examining the pattern of results across dependent measures, the effect of frame rate on reaction time appears to be the most consistent. Higher frame rates significantly reduced reaction times in all three tasks.

The effect of the independent variables on error rates was less consistent. Higher frame rates improved the rate of correct performance only in the designation task. Higher resolution improved correct performance only in the recognition task. These two variables interacted in the detection task, but there were no main effects.

Confidence ratings and confidence rating times, collected only for the detection and recognition tasks, were both influenced by frame rate. Higher frame rates produced higher confidence ratings only for recognition. However, higher frame rates produced faster confidence ratings for both detection and recognition. Imagery ratings were least affected by FR or resolution. The only significant effect was in the designation task, where there was an interaction of FR and resolution, but no main effects.

Experience was a significant factor for several dependent measures. First considering participant confidence, NATC personnel were more confident and faster at making confidence judgments for both detection and recognition tasks.

The effect of experience on reaction times was less consistent. NATC personnel were faster than Vitro employees on the designation task, for both experimental and control groups. However, the experience effect for the recognition task was only significant when the analyses included the control group.

This kind of result is not unexpected, since the overall response times of the control groups are faster than that of the experimental groups. For the detection task, the reaction time means for control and experimental groups, respectively, are 3.70 and 5.07 secs. For the recognition task, corresponding means are 4.67 secs (control) and 6.27 secs (experimental). For the designation task, corresponding means are 3.50 secs (control) and 4.22 secs (experimental). The control groups viewed clips that were not degraded or compressed. While they performed the same tasks as the experimental groups, the difficulty level was not equivalent. Consequently, with less perceptually difficult tasks to perform, the control groups had better reaction time performance.

Finally, experienced NATC personnel had less than half the number of time-outs than less experienced participants.

6.2. Experiment Two

A second dynamic experiment was conducted at the Joint Development Facility (JDF) in collaboration with Cambridge Research Associates for the UAV Program. This study was a dynamic simulation to evaluate human performance in a target designation and tracking task when mission parameters for the sensor are specified. This experiment allowed us to use the SIMNET capability in the JDF in order to create the experimental mission scenarios. The current implementation of the JDF SIMNET allows only land-based scenarios. The simulation task evaluated effects of two bandwidth reduction variables and one task variable on operator performance.

The bandwidth reduction variables were frame rate (FR) and spatial resolution (SR). The task variable is angle of incidence (AOI). In order to evaluate human performance in this dynamic task, we varied the mission parameter characteristics (e. g., sensor aspect to the target, direction of target) so that different AOIs to the target are simulated. We also used mission/payload characteristics in the scenarios (e. g., land targets of 3 meters or larger) that are provided for the missions in the NATO PG/35 Ad Hoc Technical Working Group Mission Analyses Working Papers. April 1991.

The assumption is that these scenarios simulate typical UAV target designation and tracking operations. Further, it is assumed that the simulated imagery will be of lesser fidelity than that used in Experiment One. Nevertheless, the simulation environment will permit the evaluation of operator performance with respect to various dynamic behaviors (e. g., tracking error and tracking time). Further, the design will permit assessment of the interaction between frame rate and spatial resolution on operator performance. The hypothesis is that operators will have better performance in general with the higher frame rates and higher spatial resolution as measured by tracking and designation errors and reaction times.

6.2.1 Methodology

6.2.1.1 Research Design

A single group was used. Within the group subjects designed with FR and SR as repeated measures to provide data on tracking performance. Three levels of SR were nested under each of the three levels of FR. In order to eliminate the effects of practice, this design is completely counterbalanced, so that all levels of both FR and SR occur in every order. FR is counterbalanced between subjects. SR is counterbalanced within subjects. This design results in nine blocks of trials (3 FR levels x 3 SR levels) for each subject.

Within each block of trials, five AOI values were used. Two trials of each of the AOIs were used, resulting in ten trials, presented in a different random order for each block. Accordingly, the total number of experimental trials for each subject is 90 (9 blocks of trials x 10 trials per block). An additional 30 practice trials, described below, were presented but not included in the data analysis.

Independent Variables. Frame rate was set at levels of 2, 4, and 7.5 frames per second. Spatial resolution levels were 2, 8, or 12 resolution lines across the target. AOI includes vertical, horizontal, and diagonal directions : 0°, 90°, 180°, 225°, and 315°. All of these variables are within-subjects manipulations. That is, all participants performed target designation and tracking under all combinations of conditions.

Dependent Variables. Several measures of operator performance were calculated from the raw data, which consisted of 30 observations per second of the positions of the cross-hairs and of the target. These measures are designation time, designation error, error rate, tracking error, tracking slope, acquisition error, and acquisition slope.

Designation time (DT) is the time, in seconds, collected from the onset of a trial until a response was made indicating designation. Designation error (DE) is the spatial displacement, in meters, between the location of the cross-hairs and the true center of the target when designated. Completion Rate (CR) is the percent of successfully completed trials. These three measures will describe the speed and accuracy with which the participant performed the designation part of the task.

Tracking error (TE) may be considered a "cumulative DE". That is, tracking error is the mean displacement over the entire 25 second tracking task, collected 30 times per second. Tracking slope (TS) is the slope of the best-fit straight line that describes TE over the 25 second tracking interval. This was calculated by averaging TE each second, and plotting a function of each of these 25 TE averages over time. The TS measure may be considered a "continuous TE", and is an indication of how tracking performance changes over time. The slope of this function should indicate the rate of learning or improvement in performance for each participant. These two measures will describe the accuracy with which the participant performed the tracking part of the task.

Measures corresponding to TE and TS were also calculated to describe accuracy and performance changes <u>before</u> designation. These are acquisition error (AE) and acquisition slope (AS). AE corresponds to DE, and is the mean displacement over the entire target acquisition period, up to the point of designation. Again, raw position data were collected 30 times per second. Acquisition slope (AS) is the slope of the best-fit straight line that describes AE over the target acquisition interval. This was calculated by averaging AE each second, and plotting a function of each of these averages. Unlike the tracking interval, however, this function is variable length, since it starts at the beginning of each trial and ends at the point of designation. Consequently, AS is based on averages of 30 observations per second, for a variable number of seconds. The AS measure may be considered a "continuous AE", and is an indication of how acquisition performance changes over time. The slope of this function should indicate the rate of learning or improvement in performance before designation for each participant. Comparisons of

TE and TS measures with AE and AS measures will describe the same kind of task performance before and after designation.

Finally, initial distance from the cross-hairs to the target was determined for each trial. This measure was calculated to estimate the initial conditions or difficulty level of each trial in terms of target detection, since the initial position of the target was determined randomly. This is not a measure of participant performance per se. However, it was felt that it might affect performance.

6.2.1.2 Participants

Ten volunteers having prior military experience with imaging displays and 20/20 visual acuity or better (corrected or uncorrected) served as participants. Nine participants were Vitro employees; one Navy pilot from Cecil Field became available during the week of data collection, and was included in order to serve as a comparison or baseline, similar to Experiment One. All participants were right-handed.

6.2.1.3 Equipment

Imagery was presented to participants at the JDF facility at Cambridge Research Associates. The AAI Mission Planning and Control Station (MPCS) presented the graphics on an Silicon Graphics, Inc. (SGI) monitor with a screen size of 640 x 480 pixels. The SGI monitor was set to NTSC video mode and a *Panasonic UTP-2 Universal Transcoder* converted the RGB signals to SuperVHS. The SVHS signal was carried to the AAI workstation where it was digitized and presented to the participants. A joystick on the AAI's flight control box controlled the payload.

System Architecture. Payload rate commands were sent from the participant at the AAI MPCS to a real-time processor which computed payload responses. Payload positions, both azimuth and elevation, were sent to a SGI workstation, which displayed the scene based on air vehicle position and payload orientation.

Air Vehicle and Payload Control. Straight line trajectories were computed for the air vehicle by the SGI; the air vehicle flew straight and level at a constant velocity of 60 knots. Targets were also driven along straight and level paths at a constant velocity of 5 knots. Direction of the targets relative to the path of the air vehicle is one of five specified angles-of-incidence (0°, 90°, 180°, 225°, 315°). The imaging payload, however, was under direct control of the participant. Specifically, continuous rate control of payload azimuth and elevation were utilized. Rate commands were given using an x-y deflection joystick, which was weighted with a squared shaping function. The commands were sampled and sent to the payload dynamics simulation at a 10 Hz rate. Payload motor dynamics were modeled as critically damped, second-order systems with a bandwidth of 60 radians per second. Further, the maximum slew rate was set to 20 degrees per second. Motor dynamics and maximum slew rates were the same for both azimuth and elevation. Payload dynamics were computed by the real-time processor at a 100 Hz rate; azimuth and elevation rate responses were also integrated at 100 Hz to yield azimuth and elevation positions. Payload elevation ranged from 0° (straight down) to 90° (forward); payload azimuth ranged a full 360°, with 0° due north, 90° due east, etc. Positions were sent at a 30 Hz rate to the SGI graphics process which coupled payload orientation with air vehicle position to generate the appropriate sensor view.

Graphics Process. Sensor positions were read from the payload dynamics process at 30 Hz. A single target was displayed for each test sequence at one of three specified frame rates (2, 4, 7.5). Further, the size of the target was based on one of the three specified resolutions (2, 8, 12 TV lines or vertical pixels). The video compression ratio was fixed at 50:1. Look-down angle was fixed at 35°. The field of view (FOV) was determined by each resolution (SR) level in order to maintain an altitude of approximately 1000 meters. FOV for 2, 8, and 12 TV lines of resolution was 3°, 4°, and 17°, respectively. Drift rate of the target over ground on the display was 50 knots. To achieve this effect, the UAV flew in a straight line at 62 to 65 knots. The simulated background was plain desert terrain, with occasional features such as trees or roads to provide perspective. Other mission parameters, such as the initial position of the target, direction of the target, and flight path, were combined in order to create the five AOIs. Operator tasks, response measures, and experimental procedures were under software control.

6.2.1.4 Procedures

Each participant followed the same procedure, with the exception of the unique order of task performance. They read the task description, completed a background questionnaire, and read and signed informed consent and non-disclosure documents as appropriate. The experimenter answered questions as necessary, and explained that initially the task would seem difficult until they became familiar with the sensitivity and range of motion of the joystick.

When they were first seated at the simulator screen, they read instructions explaining the designation and tracking tasks. They then initiated a short sequence of 10 practice trials at a fixed frame rate (FR) level to acquaint them with the task demands and the equipment. After the practice trials were completed, they were told to begin the experimental trials when ready.

The participant's task incorporated both target designation and target tracking on each trial. The joystick controlled the position of the cross-hairs on the screen, and was manipulated by the participant's right hand. One button, manipulated by the left hand, was used to designate the target and to initiate each trial. The button was dark between trials. When a new trial was initiated, the button was lighted and remained so until designation. When the target was designated, the button was darkened and remained so until the next trial was initiated. When the trial began, the participant was required to designate the center of the target as quickly and accurately as possible. He was also required to track the target for 25 seconds after it was designated. After 25 seconds, the screen went blank, and the participant initiated the next trial.

After the first block of experimental trials was completed (approximately 10 minutes in length), a screen appeared to provide a short break. At this time, the SR level was changed. The participant was not explicitly notified of this change, except to expect some change in the screen parameters after each break. After the second block of experimental trials was completed, another screen appeared to provide another short break. The SR level was changed again. When the third block of experimental trials was completed at the first FR level, a new screen appeared to inform the participant that the screen parameters would change significantly, and that additional practice trials were required.

At this time, the second FR level was used, and remained constant for the next three blocks of experimental trials. Before the experimental trials began, another sequence of 10 practice trials was presented at the new FR level. Each experimental block had different SR levels, and a short break was provided after each block. Again, the participant was not informed of the change in SR level between blocks.

After these three blocks were completed, another screen appeared to inform participants that the screen parameters would change significantly again, and that additional practice trials would follow. Another sequence of 10 practice trials was presented at the third FR level. The final FR level was used, and remained constant for the last three blocks of trials. Again, each of these blocks had different SR levels, and a short break was provided after each block. Upon completion of all 9 blocks of trials, the final screen announced the end of the experiment and thanked the participant for his cooperation.

The participant was debriefed concerning the different experimental conditions, and was given a briefing as to how much information regarding the study may be discussed freely. Any questions that the participants may have had were answered at this time.

6.2.2 Experiment Two Results

Data from Experiment Two was collected on one continuous task, as described above. Participants were instructed to acquire, designate, and track a single target. Although the behavior measured on each trial may be considered continuous, the results are organized according to these 3 "subtasks", for each independent variable.

The effect of each independent variable (FR, SR, and AOI) on the 8 dependent variables is discussed below. These dependent variables are designation time (DT), completion rate (CR), designation error (DE), tracking error (TE), tracking slope (TS), acquisition error (AE), and acquisition slope (AS).

6.2.2.1 Acquisition

Frame Rate. Both AE ($F = 33.5$, $p < .0001$) and AS ($F = 5.97$, $p < .0027$) were influenced by FR. Higher FRs resulted in small AEs. The slowest FR (2 fps) had the largest AE (99.5 m), which was significantly different from the other 2 conditions. As will be seen repeatedly below, this pattern emerges in several of the other dependent measures. Means for the 4 fps and 7.5 fps conditions were 63.4 m and 52.8 m, respectively.

With respect to AS, post-hoc comparisons show some tendency for the lower FR to be associated with less accuracy across the target acquisition interval. The only significant comparison, however, is that between 2 fps (-6.5 m/sec) and 4 fps (-8.7 m/sec). The highest FR (7.5 fps) did not differ (-8.3 m/sec) from the other 2 conditions. Means and standard deviations for both AE and AS data are shown in Figure 19.

Spatial Resolution. SR had a significant effect on both AE ($F = 69.5$, $p < .0001$) and AS ($F = 125.27$, $p < .0001$), as was the case with FR. Post-hoc comparisons showed that all 3 resolution conditions differed for AE, with lower resolution associated with greater error. AE means for 2, 8, and 12 lines of resolution are 102.7 m, 55.9 m, and 42.1 m, respectively.

For the AS measure, the lowest SR condition (2 lines) was significantly different from the 8 and 12 line conditions, which were essentially the same. AS means for 2, 8, and 12 lines are 15.2 m/sec, -4.1 m/sec, and -3.1 m/sec. These data are shown in Figure 20.

Angle of Incidence. There was a very marginal effect of AOI on AE ($F = 2.4$, $p < .0499$), and no effect on AS. These data, shown in Figure 21, were very consistent. AE means ranged from 58.3 m to 74.3 m. AS means ranged from -6.6 to -9.5 m/sec.

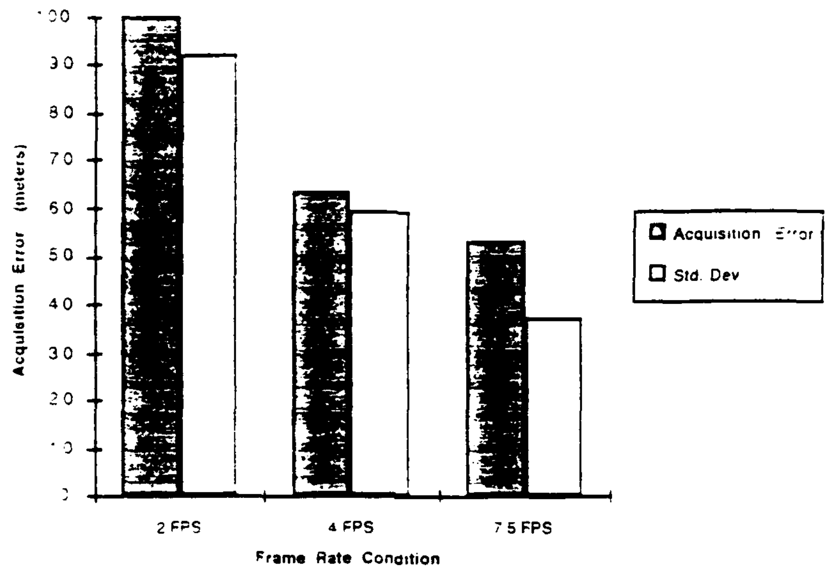
6.2.2.2 Designation

Frame Rate. The effect of FR on DE was highly significant ($F = 105.06$, $p < .0001$) and in the expected direction. Higher FRs produced small DEs, and post-hoc comparisons indicate that all 3 conditions differ significantly. The 2 fps condition has the greatest error (65.1 m) compared to the 4 fps (32.6 m) or 7.5 fps (18.7 m) conditions. These data are shown in Figure 22.

The effect of FR on DT was also highly significant ($F = 51.82$, $p < .0001$) and is similar to the AE effect. That is, higher FRs produced better performance, in the form of faster DTs, as shown in Figure 22. Post-hoc comparisons indicate than the 2 fps condition is slower (22.68 sec) than either 4 fps (15.05 sec) or 7.5 fps (14.12 sec). Again, the 7.5 fps condition is not significantly different than 4 fps.

With respect to CR, the faster DTs, produced by higher FRs, resulted in fewer timed-out trials ($F = 36.40$, $p < .0001$). Participants were given an average of 43 seconds (s.d. = 11.4 sec) to designate before a trial was interrupted with a blank screen. CRs for 2, 4, and 7.5 fps were 61.5%, 79.6%, and 90.4%, respectively. Post-hoc comparisons indicate that all 3 of these conditions differ significantly from each other.

Figure 19. Means and Standard Deviations for FR Effects on AE and AS During Acquisition



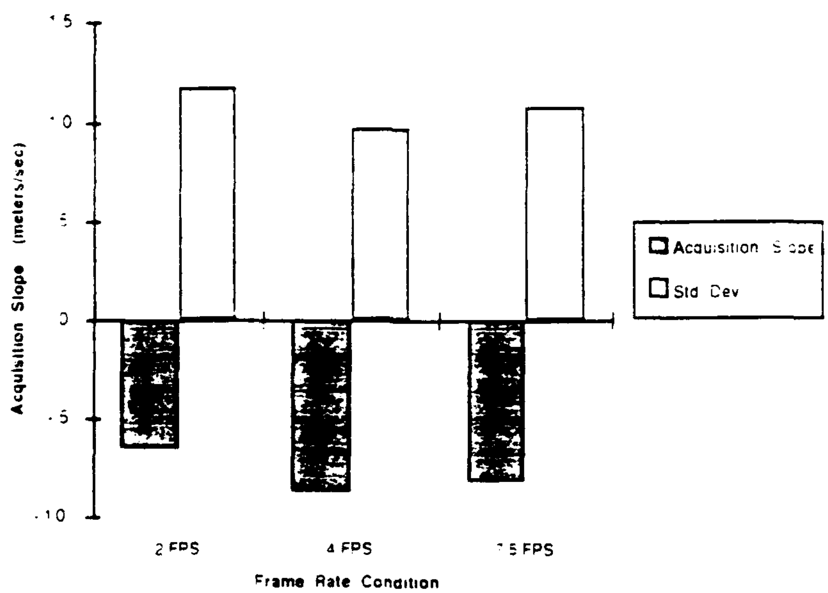
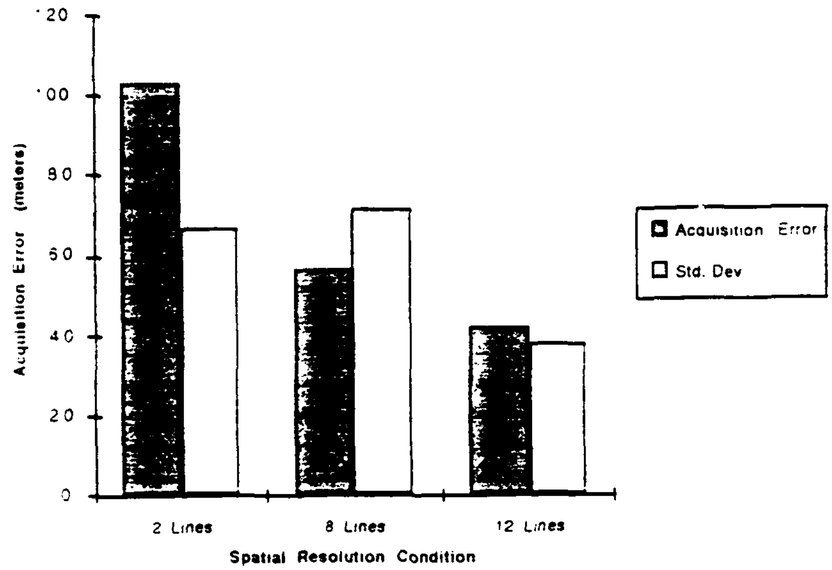


Figure 20. Means and Standard Deviations for SR Effects on AE and AS During Acquisition



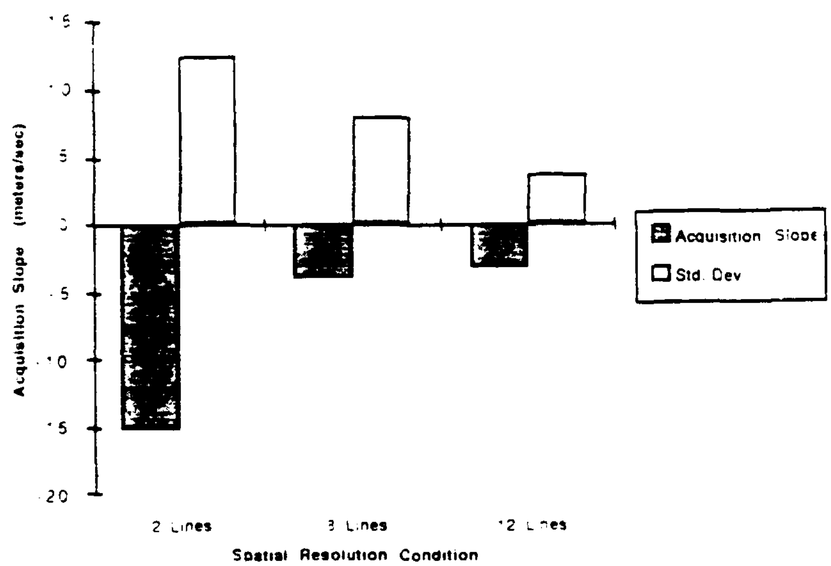


Figure 21. Means and Standard Deviations for AOI Effects on AE During Acquisition

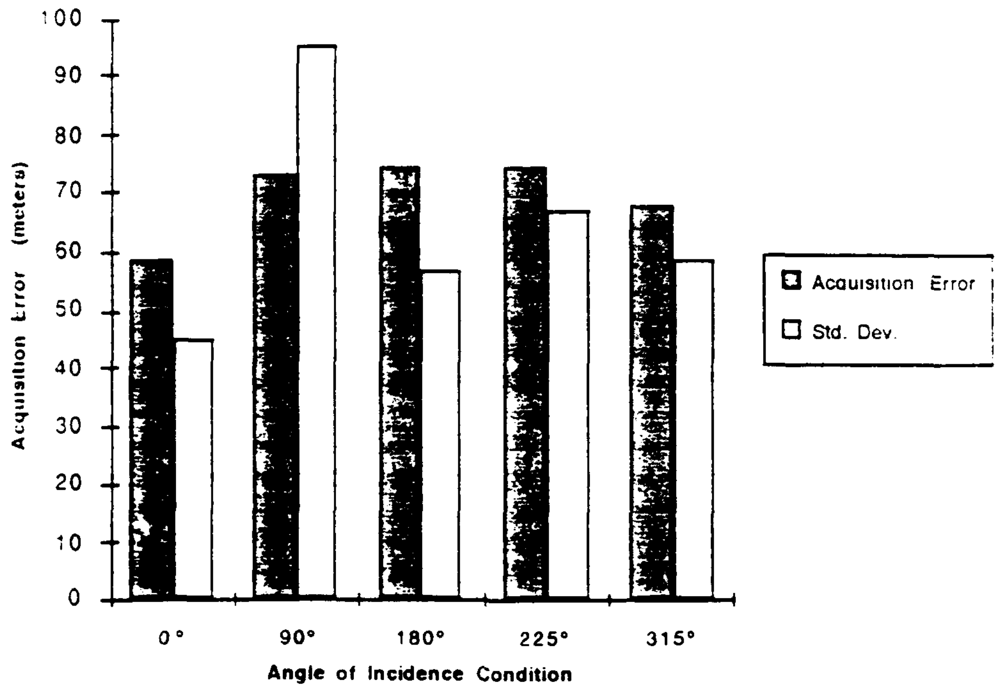
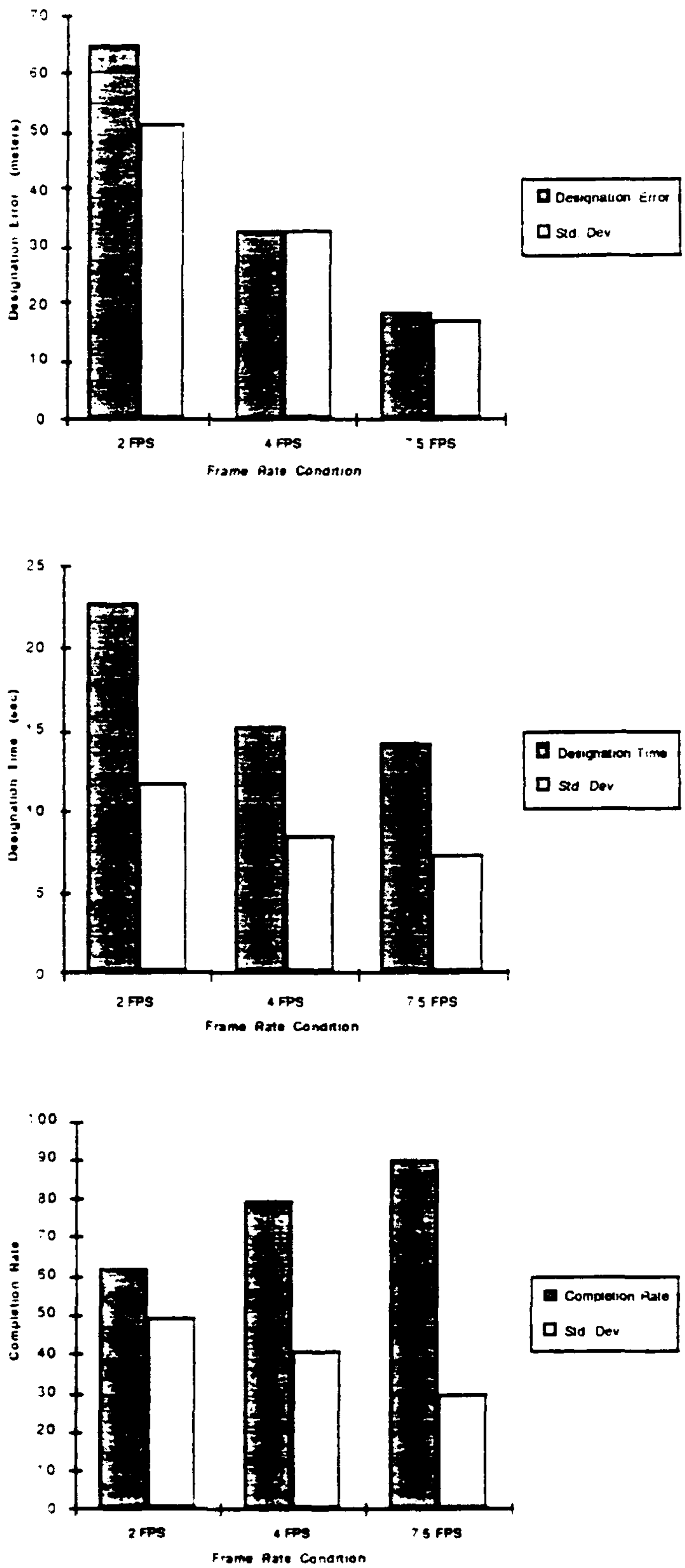


Figure 22. Means and Standard Deviations for FR Effects on DE, DT and CR During Designation



Spatial Resolution. As was the case with FR, both DT and CR were influenced by SR ($F = 7.01$, $p < .0010$ and $F = 11.41$, $p < .0001$, respectively). As shown in Figure 23, higher resolution resulted in faster responses. When SR was equal to 12 lines, DT was fastest (14.42 sec). Post-hoc comparisons show that this SR level produced faster DTs that the other two conditions, and that the 8 line and 2 line conditions did not differ from each other (18.01 sec and 17.46 sec, respectively). These data are presented in Figure 23.

In contrast to the DT data, however, high resolution produced more time-out trials. CR for the highest SR conditions was only 70.4% for 12 lines and 74.8% for 8 lines. These means do not differ significantly. The lowest SR condition (2 lines) had the highest CR (86.2%), and differs significantly from the other two conditions.

There was no effect of SR on DE. Means for 2, 8, and 12 lines were 37.2, 33.3, and 37.4 m, respectively.

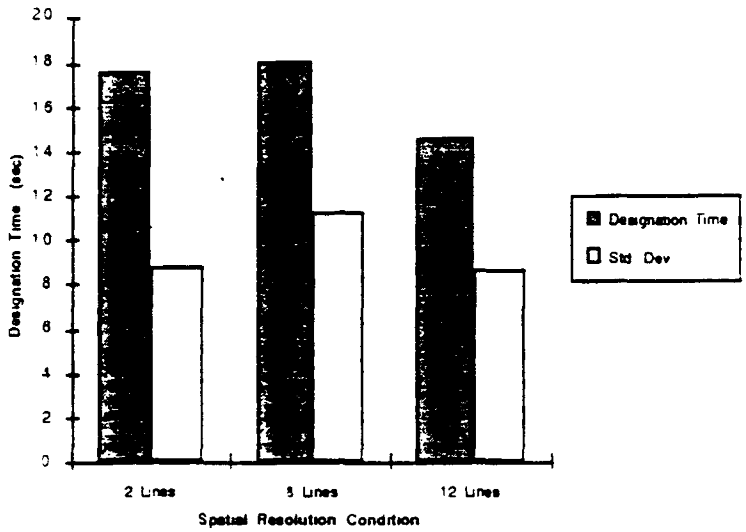
Angle of Incidence. The influence of AOI on DT was significant ($F = 7.46$, $p < .0001$). Fastest responses (13.72 sec) occurred at 0° (i. e., target and aircraft moving in the same direction) and slowest responses (18.98 sec) occurred at 180° (i. e., target and aircraft moving in opposite directions). Post-hoc comparisons reveal that 0° is a unique condition, in that it is significantly faster than 3 (90°, 180°, and 225°) of the 4 other conditions. Mean DTs for 90°, 180°, 225°, and 315° are 17.28, 18.98, 17.55, and 16.47 sec, respectively. No other comparisons were significant. These data are shown in Figure 24.

There was no effect of AOI on CR. However, as shown in Figure 24, the results parallel those obtained with DT. That is, optimal performance occurs at 0° (83.3%) and minimal performance occurs at 180° (71.6%). AOI did not affect DE either. Means ranged from 33.1 m (at 225°) to 37.8 m (at 315°).

6.2.2.3 Tracking

Frame Rate. The effect of FR on TE ($F = 72.4$, $p < .0001$) and TS ($F = 5.6$, p , $.0037$) was significant. The pattern of significant differences for the TE data is the same as that for DE and CR data. That is, higher FRs resulted in superior performance, and all post-hoc comparisons

Figure 23. Means and Standard Deviations for SR Effects on DT and CR During Designation



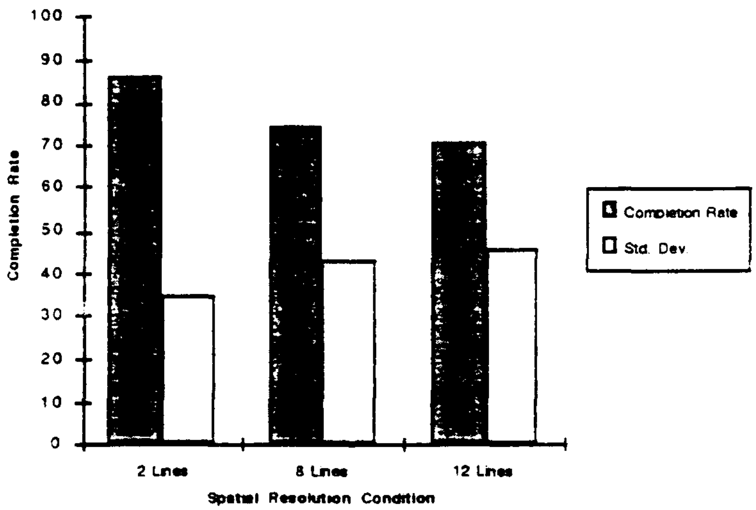
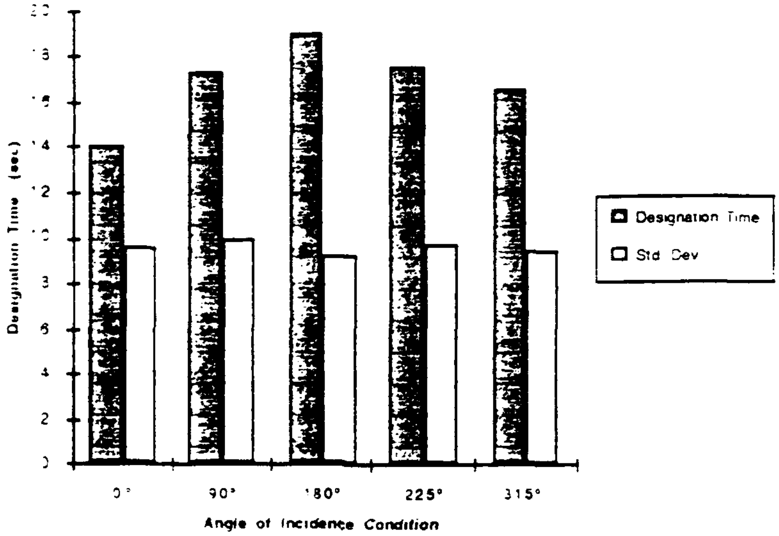
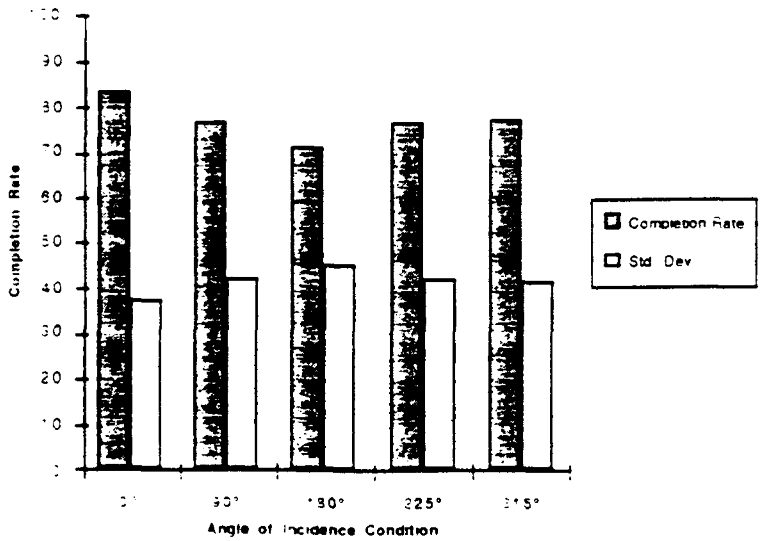


Figure 24. Means and Standard Deviations for AOI Effects on DT During Designation





were significantly different. As shown in Figure 25, large differences were obtained as FR increased. Mean TE for 2, 4, and 7.5 fps are 116.8 m, 55.7 m, and 30.0 m, respectively.

TS means followed the same post-hoc trend as that for DT and AE data. More specifically, the slowest condition (2 fps) differs from both other conditions (4 and 7.5 fps), which do not differ from each other. Mean TS for 2, 4, and 7.5 fps are equal to 3.5 m/sec, 1.0 m/sec, and .3 m/sec.

Spatial Resolution. The effect of SR on TE was very marginally significant ($F = 3.02$, $p < .049$) and differences were very small, with means of 70.0 m, 51.0 m, and 62.4 m obtained for 2, 8, and 12 lines, respectively. These data shown in Figure 26. There was no effect of SR on TS. TS means ranged from 0.6 m/sec to 1.9 m/sec.

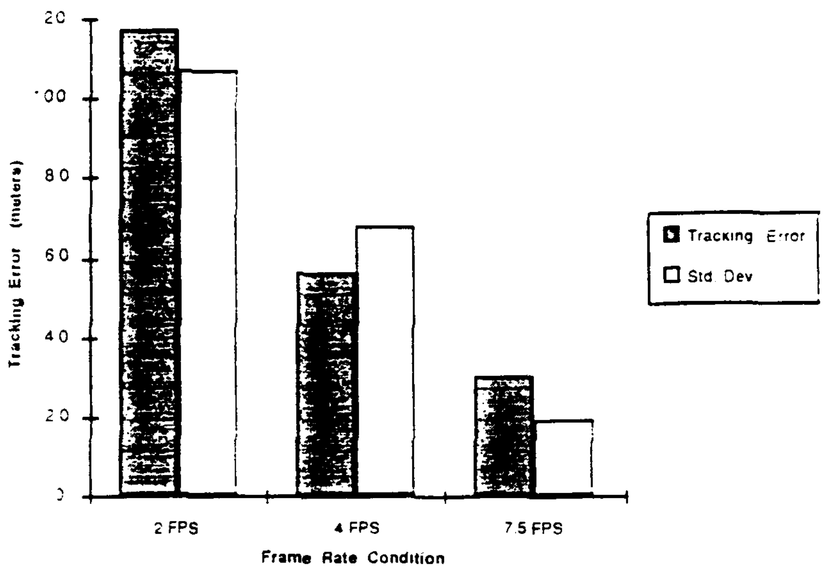
Angle of Incidence. AOI did not affect TE or TS. These data were also very consistent, ranging from 59.5 to 68.2 m for TE. For TS, the range was 1.1 m/sec to 1.8 m/sec.

6.2.2.4 Additional Comparisons

Experience. As mentioned above, 1 Navy pilot became available during the week of data collection, and was included to provide baseline data. The results described above <u>do not</u> include this data. All statistical analyses were performed on the 9 Vitro personnel. While it is statistically inappropriate to conduct analyses between the Vitro group and the Navy individual, these data are included here for comparison purposes.

Table 8 provides overall means for the Vitro group and the Navy individual for each dependent variable. Since Vitro data are based on a group mean of 9 participants and the Navy data are based on an individual, these data should be interpreted cautiously. Nevertheless, it is interesting to note that no differences in acquisition performance (AE and AS) are apparent. However, both designation performance (DE, DT, and CR) and tracking performance (TE and TS) for the Navy participant seem both faster and more accurate.

Figure 25. Means and Standard Deviations for FR Effects on TE and TS During Tracking



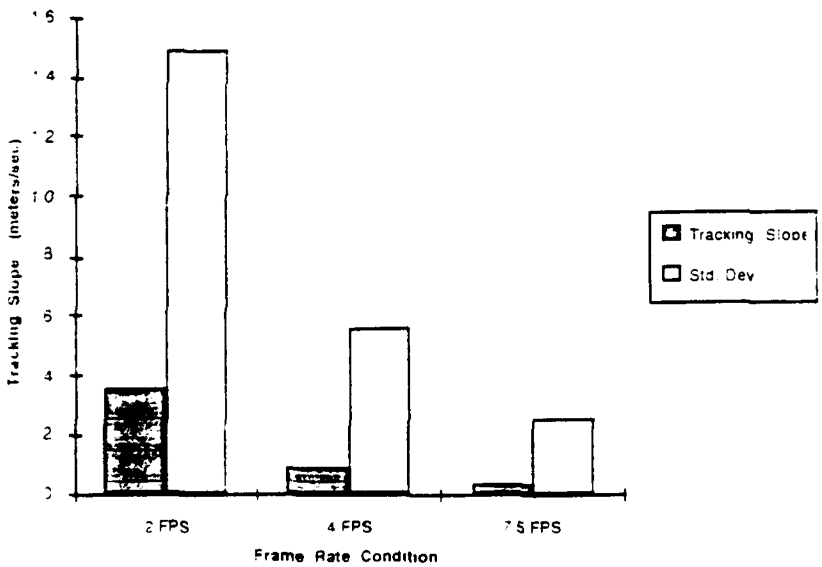
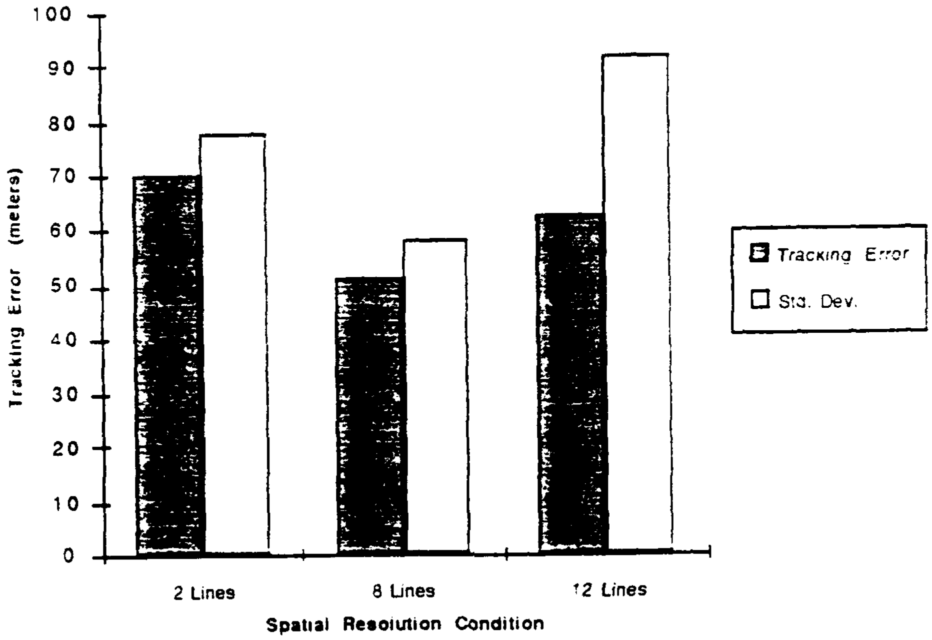


Figure 26. Means and Standard Deviations for SR Effects on TE During Tracking



It should also be kept in mind that AOI was not a variable of interest comparable to FR or SR. Instead, it was used to provide realism and some variety in the scenarios. However, the relationships between AOI, SR, and ID should be considered when interpreting the interactions described below.

Interactions. Three different interaction patterns were obtained among the independent variables. The two-way interaction between FR and SR was observed in the AE ($F = 2.78$, $p < .0259$), DT ($F = 2.50$, $p < .0415$), and CR ($F = 3.50$, $p < .0076$) measures. For all of these measures, main effects for both FR and SR were also significant. The FR x SR interactions are meaningful, because they suggest that the detrimental effect of lower FR is more pronounced at some resolution levels than others. In fact, the interactions are observed because this detrimental effect occurs at different SR for different dependent variables.

The top panel of Figure 27 describes the interaction effect on DT, showing that best overall performance occurs at SR12. Further, it suggests that performance at FR4 is comparable to that at FR7.5. The middle panel describes the FR x SR effect on CR. Here, the detrimental effect of low FR is most clearly observed at SR12. Optimal performance occurs at SR2. Since DT and CR are complementary measures of performance (i. e., low DT scores and high CR scores both indicate faster responses), this result is not unexpected. For both DT and CR, optimal FR and SR combinations could be predicted from the main effects of FR and SR. These interactions indicate that minimal performance occurs at FR2, under varying SR conditions. Finally, the FR x SR interaction on AE (lower panel of Figure 27) shows that optimal performance occurs again at SR12, where the detrimental effect of FR is least pronounced.

The only other two-way interaction, between SR and AOI, was observed in 5 dependent measures. These are AE ($F = 2.78$, $p < .0259$), TE ($F = 2.93$, $p < .0032$), TS ($F = 3.17$, $p < .0016$), DT ($F = 2.50$, $p < .0415$), and DE ($F = 2.58$, $p < .0090$). This interaction is most easily observed in the AE data at the top of Figure 28. SR2 results in extremely poor performance, while there appears to be little difference between SR8 and SR12. However, the SR2 pattern may be accounted for by the differences in ID across AOI described above. Also shown in Figure 28 is the effect of the interaction on DT, where the results parallel those obtained with AE. The same mapping of ID across AOI occurs here, too, at SR2. More specifically, means for ID across AOI are 77.9, 92.3, 114.8, 111.5, and 103.3 m, while means for AE across the same AOI are 78.4, 96.7, 111.4, 118.2, and 119.9 m. Corresponding means for DT are 13.6, 16.1, 21.8, 19.2, and 18.2 sec.

Figure 27. Means and Standard Deviations for FR X SR Interactions in DT, CR and AE During Tracking

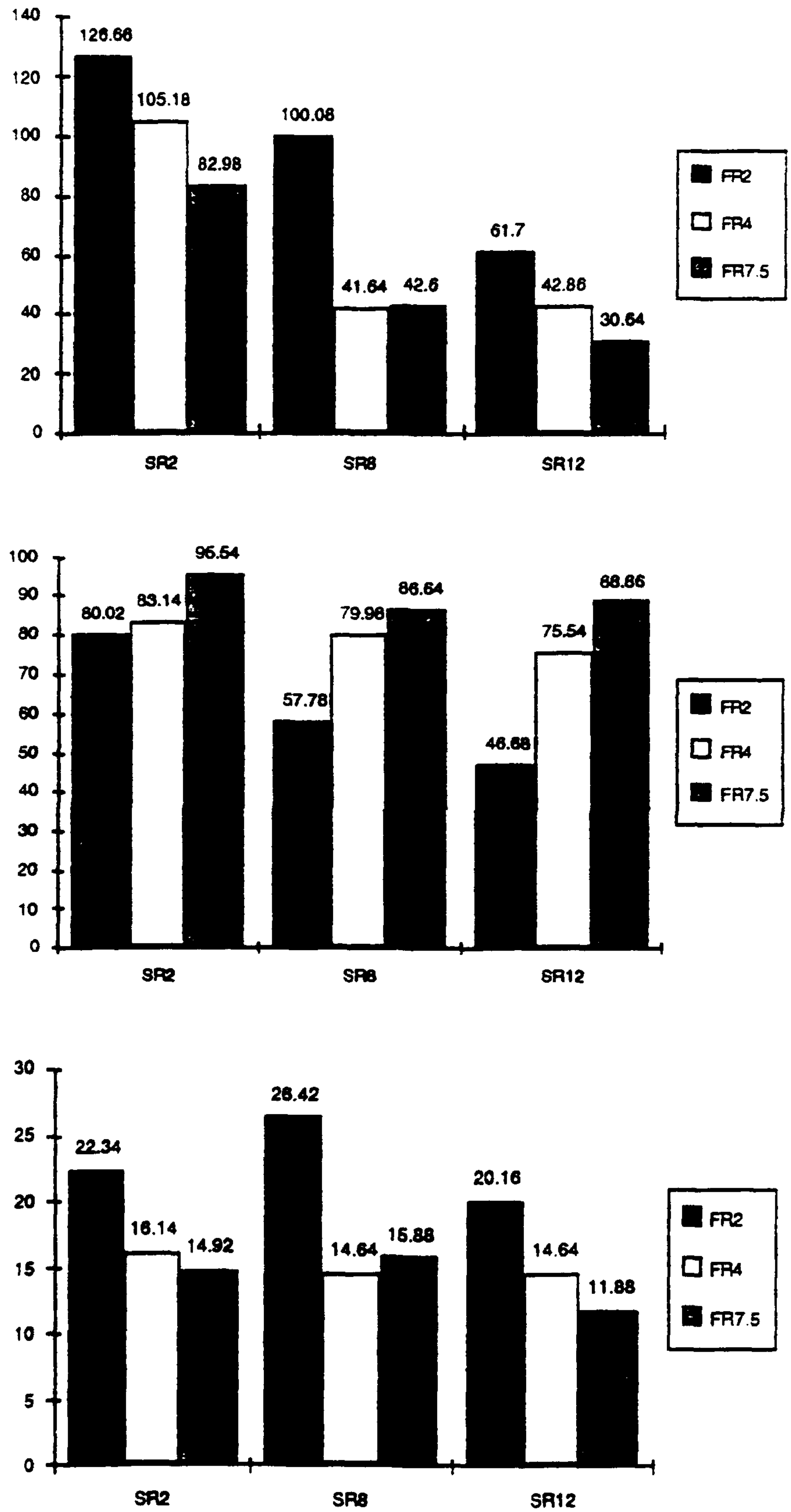
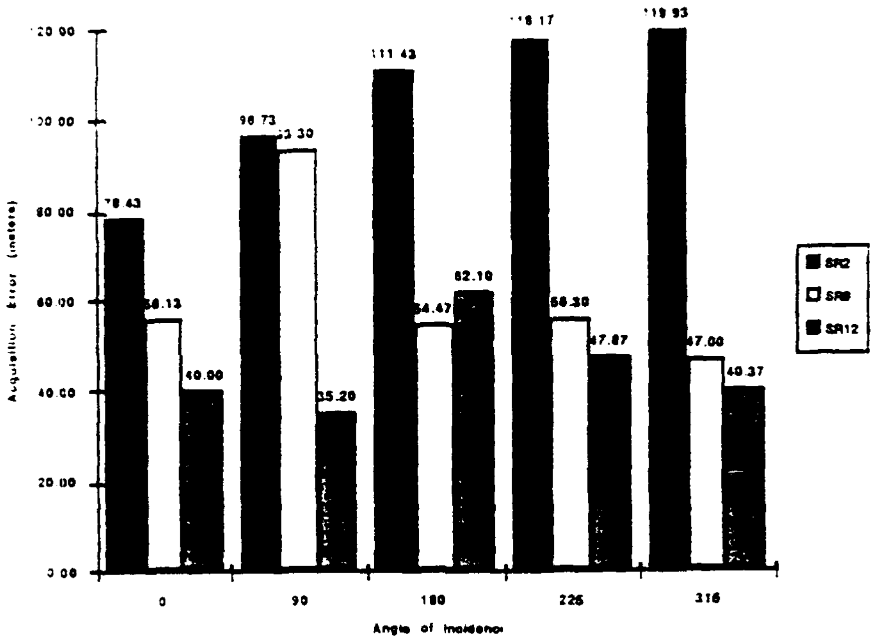
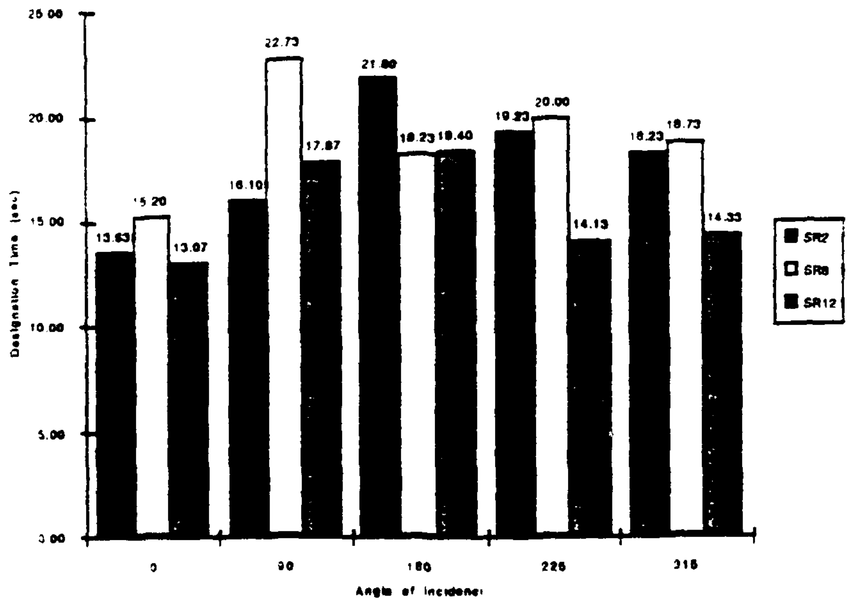


Figure 28. Means and Standard Deviations for SR X AOI Effects on AE and DT During Tracking





In view of the fact that the AOI effect on ID is largely responsible for these 2 interactions, the other 3 SR x AOI interactions are ambiguous and very difficult to interpret. Further, for all 3 of these dependent measures (DE, TE, and TS), there were no significant main effects for either SR or AOI. Accordingly, we will not discuss them further in this report.

A three-way interaction of FR x SR x AOI was observed in the AE ($F = 2.68$, $p < .0004$), AS ($F = 3.03$, $p < .0001$), TS ($F = 3.17$, $p < .0016$), and DT ($F = 1.67$, $p < .0491$) data. It should be noted that the SR x AOI interaction was also obtained for 3 of these 4 dependent variables. Therefore, these interactions are also noninterpretable at the present time.

6.2.3 Experiment Two Discussion

When examining the effects of all three independent variables on the seven dependent variables, the importance of frame rate is evident. Higher frame rates were associated with better performance for six of the seven dependent variables. For three of these, DE, CR, and TS, posthoc comparisons showed that all levels of frame rate differed, and superior performance increases directly with frame rate. These data suggest that operator performance will continue to improve as frame rate increases.

For AE, DT, and TS, post-hoc comparisons showed that the slowest frame rate is significantly different from both 4 fps and 7.5 fps, which are essentially the same. These data suggest that operator performance differences between 4 and 7.5 fps will be minimal. This result is supported by similar investigations we reviewed [12], [13], [14], [16]. Comments from participants also indicate that 2 fps should be avoided. Even more important, the FR x SR interaction results suggest that the equivalence of 4 and 7.5 fps holds across higher resolution conditions. In other words, for 4 fps, acquisition accuracy (AE) at SR8 is 41.6 m, while accuracy at SR12 is 42.9 m. Conversely, for SR8, accuracy at FR4 is 41.6 m, while accuracy at FR 7.5 is 42.6 m (see bottom panel of Figure 28). Parallel results are obtained for designation speed. For 4 fps, responses at SR8 and SR12 are both 14.6 sec. Similarly, for SR8, responses at FR4 are 14.6 sec. while responses at FR7.5 is 15.88 sec.

The effects of spatial resolution on the dependent measures is less pervasive and less consistent than the frame rate effects. Four measures showed better performance to be associated with higher resolution, and post-hoc comparison patterns were unique for each of these measures.

For example. AE performance at all three resolution levels differed significantly, suggesting that operator performance should continue to improve as spatial resolution increases. AS data, however, indicates that the lowest resolution is significantly different from the other levels, but that 8 and 12 lines are not different. It is somewhat surprising that DE was not related to SR performance. It was expected that AE, DE, and TE variables would have analogous results, since they are similar measures of accuracy. The lack of relationship between DE and SR suggests that SR is not critical for designation accuracy.

DT data shows a somewhat opposite pattern, in that the two lowest resolution levels do not differ, but 12 lines of resolution are significantly faster and should produce superior performance. The TE effect was very marginally significant, and indicates that the most accurate performance occurs at 8 lines of resolution. Together, these results suggest that 8 lines of resolution is a "borderline" value. Under some circumstances, performance is equivalent to that at higher resolutions; in other circumstances, performance at 8 lines is inferior to that at 12 lines.

Finally, the CR effect indicates that better performance in terms of trials are completed occurs at the <u>lowest</u> resolution. This result seems to be counter-intuitive, but is easily interpreted in light of the participant's debriefing comments. At higher resolution levels, the FOV was only 3° or 4°. When higher resolution was paired with slow frame rates, participants frequently "lost" the target, and were unable to re-acquire it. In particular, the combination of FR2 and SR12 was described as "difficult" or "extremely irritating". They preferred the low resolution, for which the FOV was 17°, because the target rarely left the screen.

As mentioned above, AOI was used as an independent variable to create a variety of scenarios, and was not expected to have large effects on performance. In fact, only two measures showed any main effect of AOI. AE had a very marginal effect. DT data showed that optimal performance occurs at 0°, and that participants designated faster at this AOI than all other AOIs except 315°. Most of them also expressed a definite preference for the 0° scenarios.

7. Summary of Experiments One and Two and Recommendations

7.1 Summary

Design requirements for the UAV digital data link must be discussed in light of Government-specified performance criteria and results of operator performance in realistic degraded imagery conditions. Research in this regard shows that the proposed bandwidth limitations on sensor imagery data transmission rates impose constraints on operator performance. These constraints, (a result of imagery compression, frame rate, and spatial resolution trade-offs), are assumed to be factors that will function in real-time mission scenarios. The analysis contained in this report provides data with which certain basic requirements and trade-offs are recommended to support human performance under these conditions. The data is summarized first and then the essential requirements are listed.

Results from Experiment One indicate that frame rate is a much more critical variable than spatial resolution. In both Experiment One studies, faster frame rates are associated with faster reaction times, higher confidence, and faster confidence ratings. The effect of frame rate on error performance is less consistent and less easily interpreted. In the baseline study, percentages of recognition errors and time-outs were both influenced by frame rate, although in opposite directions: faster frame rates increased errors but decreased time-outs. In the experimental study, designation errors were reduced by faster frame rates. What this means is that high frame ratestend to decrease designation errors at the expense of increasing recognition errors.

In contrast, spatial resolution had no effect on reaction times or confidence measures for any task. The only dependent variable affected by resolution across all three tasks was image quality rating. This effect was difficult to interpret because full resolution clips were consistently judged to be of lower quality than half resolution clips. Finally, resolution had a marginal effect on error rates, but only for the recognition task. The influence of experience within the context of Experiment One appears to affect operators' confidence with tactical decision making. Experience also resulted in fewer time-outs which indicates better decision-making ability. Further studies are needed to more completely evaluate the influence of experience on performance.

Consistent with the first experiment, frame rate has more of an effect on performance than spatial resolution in Experiment Two. A similar pattern was observed with higher frame rates associated with faster acquisition, faster designation time, smaller designation error, and smaller tracking error. In many tasks, no difference was observed between 4 and 7.5 fps. This validates previous human performance results in RPV programs [12], [13], [14] and [16]. Therefore 4 fps was shown to be sufficient for operator performance in both static and dynamic tasks.

Spatial resolution had some effect on performance in Experiment Two, but these results were not as consistent as the frame rate effect. Whereas frame rate affects performance overall, spatial resolution affects only specific tasks. For example, acquisition error is reduced as resolution increases; but acquisition slope, which indicates learning rate, improves only at the lowest resolution (2 lines). Designation time is faster at the highest resolution (12 lines), but there is no resolution effect on designation error. Completion rate, or the percentage of completed trials, is better at lower resolution, presumably since the target was always visible on the display.

In Experiment Two, the frame rate and spatial resolution interactions are of particular interest to the trade-offs considered. If higher resolution is needed for a task, then either 4 or 7.5 fps can be used and similar operator performance can be expected. Since 7.5 fps won't meet the 119 kb/sec criterion of interest to the UAV JPO, and performance is the same at 4 or 7.5, then 4 fps and 8 or 12 lines of resolution is recommended for tasks that require designation speed and acquisition accuracy. It is noted that further investigation of resolution values around 8 lines is needed to clarify some of the inconsistencies found. Examining human performance in similar tasks with 6, 8, and 10 TV lines of resolution should clarify any ambiguity. The interaction effects of frame rate and spatial resolution on percentage of trials completed suggests that operators need higher frame rates (4) if higher resolution (12 lines) is available. A 2 fps/12 lines combination is to be avoided. The best completion rate performance was at 2 lines of resolution across the target.

The data reduction trade off conditions for all groups compared in Experiment One and mean performance scores for each group are shown below.

Summary Bandwidth Trade Off Table - Experiment One

Study Condition	Frame Rate	Resolution	Compression Ratio	Data Rate
Control Group	30 fps	640x480 pixels (Full)	1:1	73 Mbits/sec
Full Res. / 4fps	4 fps	640x480 pixels (Full)	50:1	196 Kbits/sec
Half Res. / 4fps	4 fps	640x240 pixels (Half)	50:1	98 Kbits/sec
Full Res. / 2fps	2 fps	640x480 pixels (Full)	50:1	98 Kbits/sec
Half Res. / 2fps	2 fps	640x240 pixels (Half)	50:1	49 Kbits/sec

Summary Human Performance Table - Experiment One

Study Condition	TASKS					
	Detection		Recognition		Designation	
	RT	% Correct	RT	% Correct	RT	% Correct
Control Group	3.70	82.9	4.65	83.7	3.52	89.2
Full Res. / 4fps	4.71	70.7	6.19	75.8	3.82	76.3
Half Res. / 4fps	4.80	78.7	5.77	68.2	4.16	73.6
Full Res. / 2fps	5.24	79.2	6.38	73.9	4.48	65.8
Half Res. / 2fps	5.42	72.7	6.47	65.5	4.57	66.1

The control group in Experiment One served to define operator performance under normal, nondegraded conditions. They obtained a 90% performance level for the three tasks evaluated (detection, recognition, designation) with reaction times ranging from 3.5 to 4.6 secs across tasks. This performance criterion meets those suggested in [2]. However, none of the bandwidth trade off conditions met this performance requirement. In general, the best performance was observed in the groups that had 4 fps. As the Summary Bandwidth Table shows, the 4 frames, full resolution condition does not meet the JTIDS throughput limit of 119 kb/sec. Performance comparisons between full and half resolution with 4 fps did not differ in ways that would affect operational performance. The Summary Table above shows performance levels of 70 to 78% with reaction times ranging from 3.9 to 5.9 secs with 4 fps at half resolution (full display). This combination could be used as a staring point for the digital data link design although performance did not reach a 90% level.

A similar table of summary performance data is provided for Experiment Two as shown below.

Summary Human Performance Table - Experiment Two

Study Condition	TASKS					
	Acquisition		Designation		Tracking	
	Error	Slope	Error	Time	Error	Slope
2 Lines / 2 FPS	138m	-12.51	64m	22.16 sec	116m	4.49
2 Lines / 4 FPS	117m	-16.89	37m	15.89 sec	65m	0.62
2 Lines / 7.5 FPS	97m	-15.96	22m	14.81 sec	39m	0.54
8 Lines / 2 FPS	94m	-2.95	69m	26.30 sec	100m	3.36
8 Lines / 4 FPS	52m	-5.18	33m	15.05 sec	46m	0.28
8 Lines / 7.5 FPS	48m	-4.78	20m	15.68 sec	28m	0.29
12 Lines / 2 FPS	76m	-2.21	66m	19.61 sec	141m	6.72
12 Lines / 4 FPS	48m	-3.41	38m	15.00 sec	55m	1.52
12 Lines / 7.5 FPS	35m	-3.40	22m	11.96 sec	26m	-0.14

No defined performance levels were identified for this experiment. It must be noted that Experiment Two is a pilot study that provides preliminary guidance for data link design with regard to dynamic designation and tracking. The 8 lines/4 fps and 12 lines/4 fps conditions have similar performance in acquisition tasks, although the results suggest that 2 lines can help operators re-acquire a target that moves off the display. As shown in the table, no real differences are noticed at 4 fps with 2 lines, 8 lines, or 12 lines for designation performance. In tracking tasks, the 8 lines/4 fps has the best performance. A recommendation of 8 lines/4 fps results from these comparisons. Results observed at 8 lines were less consistent than other data analyzed.

7.2 Recommendations

It was stated earlier that certain missions such as BDA may require higher resolution in order to precisely identify the type and extent of damage. Thus an operator - selectable tool that changes the sensor FOV to increase resolution could enhance recognition performance in this mission. In order to meet the bandwidth available in the target JTIDS, a trade-off with frame rate (reduce 4 frames to 2) during the time the FOV is narrowed should be made. After the critical identification decision is made, FOV and frame rate would revert of the their original values.

The completion rate effects in Experiment Two at higher resolution values were caused by participants "losing" the target. The size of the target, a function of the different FOVs (3°, 4°, and 17°) in relation to the total display, is important for maintaining situation awareness. If 3° or 4° FOVs are used and the target moves off the display, operators should be able to choose a 17° FOV in order to re-acquire the target. This could influence designation time in mission-critical situations. Moreover, mission requirements for Naval Gun Fire Support state that designation time need only be supported until the system can pick up the appropriate display coordinates for the target. Automatic lock-on and tracking can then be initiated. A FOV tool should aid performance in this mission task.

Reduced acquisition and tracking error at higher resolution has implications for performance in certain missions. For example, in Over-the-Horizon targeting, the operator requirement involves course tracking. Thus, operators need only be able to keep the sensor pointed roughly over the target and larger tracking error can be tolerated. The experimental results suggest 8 lines of resolution as the best level for such performance.

A continuous control joystick was used as the sensor slewing device. Although the slewing control device was not a variable of interest in these experiments, a learning effect was observed to be associated with control of the joystick. Participants expressed frustration with operating the joystick even with ten practice trials before each condition change. Therefore, it is recommended that adequate pre-mission training be introduced so that joystick control will not inhibit operator performance.

The results presented are consistent for certain requirements that are also supported by previous investigations [12]. Therefore, a recommendation of

- **4 frames per second**

to support detection, recognition, designation, and tracking tasks is appropriate.

A recommendation is made for variable

- **resolution tools to enhance performance**

and increase detection sensitivity and recognition capabilities, and

- **changeable FOV tools to enhance situation awareness**

in designation and tracking tasks.

Overall, performance can be enhanced by providing

- **training for different trade-off combinations and joystick control**

It should be noted that this suite of experiments did not evaluate compression per se. Further work is needed in order to address issues related to compression versus no compression in each of the frame rate - resolution trade off conditions that were examined. Nevertheless, the experimental data do indicate reliable performance levels at an averaged 50:1 compression ratio using the DCT algorithm.

Appendix A: References

[1] Concept of Operations for Maritime UAV, 1991.

[2] NATO PG/35 Ad hoc technical working group mission analysis working pagers. 5 April 1991.

[3] Rabbani, Majid, P. W. Jones, Digital Image Compression Techniques, SPIE Optical Engineering Press, Bellingham, WA, SPIE Technical Publication, Vol. TT07, 1991, 221 pages.

[4] W. K. Pratt, Digital Image Processing, John Wiley & Sons: New York, 1978, 750 pages.

[5] C. Caviogli, "JPEG Compression: Spelling Out The Options." Advance Imaging, Vol. 6, No. 3, pp. 44-48, March 1991.

[6] N. Amed, T. Natarjan, and K. R. Rao, "Discrete Cosine Transform, "IEEE Trans. Compute., Vol. C-23, pp. 90-93, Jan 1974.

[7] Image Interpretation Handbook, Naval Air Systems Command, TM30 - 245A, Vols. 1 and 2, December 1967.

[8] S. J. Anthony, J. W. Buffington, S. W. Havens "Unmanned Aerial Reconnaissance Vehicle (UARV) Imagery Interpretation Study," Rome Laboratory Technical Report, Image Systems Division, Griffiss AFB, NY, June 1991.

[9] C. R. Walling, "Imagery Interpretation Requirements for Reconnaissance Systems", Rome Laboratory Technical Report, RADC-TR-90-370, Image Systems Division, Griffiss AFB, NY, December 1990.

[10] J. E. Swistak, "Effect of spatial and temporal video image compression on military target detection," U.S. Army Night Vision & Electro-optics Laboratory, Fort Belvoir, VA, Report No. DELNV-TR-0010, April, 1980.

[11] R. G. Mills, F. T. Hutson, & W. B. Hartman, "Evaluation of alternative video imagery processors in unjammed and jammed environments in terms or operator performance in a weapon

delivery system." SDS Tech Report TR-81-45, 1981.

[12] M. L. Hershberger, & A. J. Farnochi. "Application of operator video bandwidth compression/reduction research to RPV design," Hughes Aircraft Tech Report TR-78-2415-F. 1981.

[13] K., Agin, M. L. Hershberger, & V. Lukosevicius, "Video bandwidth reduction/compression research for the Army remotely piloted vehicle system," U. S. Army Electronics Research & Development Command Technical Report TR-770-2415-1, Fort Monmouth, N.J., 1980.

[14] M. L. Hershberger, & R. J. Vanderkolk. "Video image bandwidth reduction/compression for remotely piloted vehicles," Hughes Aircraft Company, Hughes Report No. P76-243R. Contract No. F33657-75-C-0532, October 1976.

[15] H. C. Self, and S. A. Heckart, "TV target acquisition at various frame rates," Technical Report AMRL-TR-73-111, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Dayton, OH, 1973.

[16] M. L. Hershberger, "Operator Performance Evaluation of Mini-RPV Video Image Bandwidth Reduction/Compression Techniques," Display Systems & Human Factors Department, Hughes Aircraft Co., Culver City, CA, Final Report for Period Jun 1975 - Feb 1976, Contract N66001-75C-0228, 1978.

[17] J. Johnson, "Analysis of image forming systems," in Image Intensifier Symposium, #AD220160, Ft. Belvoir, VA, 1958.

[18] R. A. Erickson, "Line criteria in target acquisition with television," Human Factors, Vol. 20, pp. 573-588, 1978.

[19] T. B. Sheridan. "Supervisory control: Problems, theory, and experiment for application to human-computer interaction in undersea remote systems," MIT Tech Report AD-A116-236, 1982.

[20] G. A. Swets, "The relative operating characteristic is psychology," Science, Vaol. 182, pp. 990-994, 1973.

[21] G. A. Swets & D. M. Green. "Applications of signal detection theory," in Psychology: from research to practice. H. L. Pick. Jr., et al (eds.) Plenum Publishing Corporation.

A.1 References Consulted

[1] Advisory Group for Aerospace R&D, "Image and sensor data processing for target acquisition and recognition", 1980.

[2] A. V. Barber, "Visual mechanisms and predictors of far field visual task performance," Human Factors, Vol. 32, pp. 217-233, 1990.

[3] F. W. Campbell, & L. Maffei, "The influence of spatial frequency and contrast on the perception of moving patterns," Vision Research, Vol. 21, pp. 713-721, 1980.

[4] D. Cohen, "Filling or outlining shapes with color: The effects on a visual search task." Naval Air Development Tech Report NADC-89004-60, 1988.

[5] H. C. Diener, E. R. Wist, J. Dichgans, & T. Brandt, "The spatial frequency effect on perceived velocity," Vision Research, Vol. 16, pp. 169-176, 1976.

[6] P. S. Eisen, & P. O. Passenier, "Human-machine interface issues of unmanned air vehicles: Proposal," Netherlands, NTO draft document, 1990.

[7] J. L. Sites, "AMIDARS tactical target acquisition analysis," BDM Technical Report, 1986.

[8] A. P. Ginsburg, "Direct performance assessment of HUD display systems," Air Force Aerospace Medical Research Lab Tech Report, 1983.

[9] T. J. Gleason, "Data Link Trade-offs For Unmanned Aerial Vehicles," Gleason Research Associates, Inc., GRA Report 128, Prepared for Missile Research, Development, and Engineering Center, U. S. Army Missile Command, Contract DAAH01-88-P-0503.

[10] C. F. Greening, "Mathematical modeling of air-to-ground target acquisition." Human Factors. Vol. 18. pp. 111-148. 1976.

[11] C. F. Greening and M. J. Wyman. "Visual detection model." Human Factors. Vol. 13. pp. 435-445. 1970.

[12] R. B. Huntoon. B. Schohan, & U. Shvern, "Visual search performance in simulated remotely piloted vehicle utilization as a function of auxiliary task loading on the observer," U. S. Army Research Institute Technical Report 357, Alexandria, VA., 1979.

[13] P. L. Jacobs, D. F. Williams. & H. Y. Fry, "MITLS: a man-in-the-loop simulation of remotely piloted homing missiles." Army Redstone Arsenal Tech Report RG-823, 1982.

[14] G. Kuperman, J. Kettlewell. W. N. Kama. & J. Fragiotti, "Research and simulation in support of near real time/real time reconnaissance RPV systems," Aerospace Medical Research Laboratory Technical Report AMRL-TR-77-33: Wright Patterson Air Force Base, OH., 1977.

[15] N. Mitchell, "Terrain travel simulation: Data and application," U. S. Army Research Institute Technical Report, Alexandria, VA., 1984.

[16] B. Moulden, & F. Kingdom. "Effect of pixel height, display height, and vertical resolution on the detection of a simple vertical line signal in visual noise." Human Factors, Vol. 29 pp. 433445. 1987.

[17] P. N. Narendra, & B. L. Westover. "Advanced target tracker concepts.," Honeywell Tech Report AD-A094-008, 1980.

[18] L. E. Reed, "Visual proprioceptive cue conflicts in control of remotely piloted vehicles," U. S. Air Force Human Resources Laboratory Technical Report TR-77-57, 1977.

[19] Remotely piloted vehicle, Air Force Development Test and Experimentation Technical Report, Fort Hood, TX., 1986.

[20] S. N. Roscoe, "Ground-referenced visual orientation with imaging displays," New Mexico State Tech Report TR-81-0080, 1980.

[21] H. G. Sage et al., "Investigations into automatic target cuing techniques for man-in-the-loop," U. S. Army Command Redstone Arsenal Technical Report - GC-86-10, 1987.

[22] T. B. Sheridan. "Supervisory control of remote manipulators vehicles and dynamic processes," MIT Tech Report, 1983.

[23] B. L. Silbernagel, "Using realistic sensor, target, and scene characteristics to develop a target acquisition model," Human Factors, Vol. 22, 1982.

[24] S. R. Sims, R. H. Farmer. "An automated human factors analysis system for imaging data," Army Redstone Arsenal Tech Report RE-83-24. 1983.

[25] C. F. Weiman, "Robot vehicle image compression," SBIR report #13432. 1989.

[26] H. Frendendall, and W. L. Behrend, "Picture Quality - Procedures for Evaluating Subjective Effects of Interference," Proc. IRE, Vol. 48, 1960, 30-1034.

[27] E. M. Lowry, and J. J. Depalma, "Sine Wave Response of the Visual System, I. The Mach Phenomenon," J. Opt. Soc. Am., Vol. 51, No. 7, July 1961, pp. 740-746.

[28] E. M. Lowry, and J. J. Depalma, "Sine Wave Response of the Visual System, II. Sine Wave and Square Wave Contrast Sensitivity," J. Opt. Soc. Am., Vol. 52, No. 3, March 1962, pp. 328335.

[29] M. B. Sachs, J. Nachmias, and J. G. Robson, "Spatial Frequency Channels in Human Vision," J. Opt. Soc. Am., Vol. 61, No. 9, Sept. 1971, 1176-186.

[30] C. F. Hall and E. L. Hall, "A Non-Linear Model for the Spatial Characteristics of the Human Visual System," IEEE Trans. Syst., Man, Cyber., SMC-7, No. 3, March 1977, pp. 161-170.

Appendix B

Discrete Cosine Transform

$$F(u,v) = \frac{4C(u)C(v)}{N^2} \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} f(j,k) \cdot \cos\left[\frac{(2j+1)u\pi}{2N}\right] \cos\left[\frac{(2k+1)v\pi}{2N}\right]$$

for $u,v = 0,1,\ldots,N-1$, where

$$C(w) = \begin{cases} \frac{1}{\sqrt{2}} & \text{for } w = 0 \\ 1 & \text{for } w = 1, 2, \ldots, N-1. \end{cases}$$

The inverse transform is given by

$$f(j,k) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u)C(v)\, F(u,v) \cdot \cos\left[\frac{(2j+1)u\pi}{2N}\right] \cos\left[\frac{(2k+1)v\pi}{2N}\right]$$

for $j,k = 0,1,\ldots,N-1$.

The DCT by itself does not result in compression. The original array of spatially distributed intensity information is merely replaced by an array of coefficients corresponding to the weights to be applied to various spatial frequency components in order to recover the original image.

Compression is effected in the frequency domain by selecting only the subset of spatial frequency components having magnitudes above a specified threshold. After thresholding, the range of the coefficients is rescaled and quantized into a number of levels that can be coded using a desired number of bits. At this point, the original image could be recovered only approximately as the discarded information is not recoverable. Typically, apparent visual degradation is small as the procedure preserves the most significant spatial frequency information. Of course, the selection of thresholds and the number of quantization levels for the DCT can affect the quality of the results.

Appendix C

Acronyms & Abbreviations

AE	acquisition error
ANSI	American National Standards Institute
AOI	angle of incidence
APC	armored personnel carrier
BDA	Bomb/Battle Damage Assessment
CCD	Charge Coupled Device
cm	centimeter
CR	completion rate
	confidence rating
CRT	cathode ray tube
	confidence rating time
d'	detection sensitivity
DCT	discrete cosine transform
DE	designation error
DPCM	differential pulse code modulation
DT	designation time
EO	electro-optic
ESM	Electronic Support Measures
FA	false alarm
FLIR	Forward Looking Infrared
FOV	field of view
fps	frames per second
FR	frame rate
GRD	Ground Resolve Distance
HMMWV	High Mobility Multi-Wheeled Vehicle
Hz	Hertz
ID	identification
IR	infrared
IRLS	Infrared Line Scanner
ISO	International Standards Organization
JDF	Joint Development Facility

JPEG	Joint Photographic Experts Group
JPO	Joint Program Office
JTIDS	Joint Tactical Information Distribution System
kBit	kilobit
m	meter
MBit	Megabit
MPCS	Mission Planning and Control Station
MPEG	Motion Picture Experts Group
MPO	Mission Payload Operator
N	sample size variable
NATC	Naval Air Test Center
NIIRS	National Imagery Interpretability Rating Scale
NTSC	National Television Standard Committee
P(FA)	probability of false alarm
PG	Planning Group
RADC	Rome Air Development Center
RGB	red-green-blue
ROC	Receiver Operating Characteristic
RPV	Remotely Piloted Vehicle
RSTA	Reconnaissance, Surveillance and Target Acquisition
RT	reaction time
SAR	Synthetic Aperture Radar
s.d.	standard deviation
SDT	signal detection theory
sec	second
SGI	Silicon Graphics, Inc.
SVHS	Super VHS
TE	tracking error
TS	tracking slope
UAV	Unmanned Aerial Vehicle
VCR	video cassette recorder