Sensitivity to Natural 3D Image Transformations during Eye Movements

Maryam Keyvanara
Department of Electrical Engineering
and Computer Science
York University
Toronto, Canada
maryam@cse.yorku.ca

Robert S. Allison
Department of Electrical Engineering
and Computer Science
York University
Toronto, Canada
allison@eecs.yorku.ca

ABSTRACT

The saccadic suppression effect, in which visual sensitivity is reduced significantly during saccades, has been suggested as a mechanism for masking graphic updates in a 3D virtual environment. In this study, we investigate whether the degree of saccadic suppression depends on the type of image change, particularly between different natural 3D scene transformations. The user observed 3D scenes and made a horizontal saccade in response to the displacement of a target object in the scene. During this saccade the entire scene translated or rotated. We studied six directions of transformation corresponding to the canonical directions for the six degrees of freedom. Following each trial, the user made a forced-choice indication of direction of the scene change. Results show that during horizontal saccades, the most recognizable changes were rotations along the roll axis.

CCS CONCEPTS

- **Information Interfaces and Presentation→ Multimedia Information Systems**; *Artificial, augmented, and virtual realities*
- **User Interfaces →** Graphical user interfaces (GUI); Three-dimensional graphics and realism- virtual reality.

KEYWORDS

Eye Tracking, Virtual Environments, Image Transformations, Saccadic Suppression

1 INTRODUCTION

Eye gaze has compelling features for interaction and can reveal a lot about a person's interests, intentions, and actions. Eye tracking in virtual environments can be used to track the user's eye movements, analyze how they observe the scenes and respond to the stimuli, and provide a substrate for high-quality interaction with the environment. Saccades are one of the most common types of human eye movements. Saccades are fast, ballistic eye movements that change fixation from one location to another [Westheimer 1954]. Saccades are used several times a second to move the fovea to different points of interest and gain an understanding of the visual environment. Although during every saccade there is a large movement of the image of the environment on our retina, our perception of motion and visual stimuli is attenuated and our visual acuity is suppressed. Hence, we have very limited capability to obtain visual information during this short time [Ibbotson and Krekelberg 2011]. This perceptual phenomenon is known as the saccadic suppression effect. Saccadic suppression is not apparent as the brain combines information from successive eye fixations to create a subjective impression of continuous view of the visual field and hence masks the motion-blurred images obtained during saccades [Leigh and Zee 2015].

In interactive computer graphics, there may be the need to perform different modifications to the images on the display to keep the system up to date. Immediate and abrupt changes in the displayed images could cause disturbing effects in 3D scenes for the viewers. The saccadic suppression effect has been suggested as a technique for masking extensive graphic updates in a 3D virtual environment [Herpers et al. 2004; Schumacher et al. 2004; Franke et al. 2014]. It can be used to achieve various manipulations with a computer display without the viewer being aware. This allows creation of displays with stimuli that change properties during a users' saccadic eye movements. It can also be useful in studies of visual representation and memory, change blindness and virtual environments [Triesch et al. 2002].

Horizontal translations which occur during a saccade while the user is viewing natural 2D images, are perceived as much smaller than similar translations during fixations [Allison et al. 2010; Herpers et al. 2004]. Schumacher et al explored saccadic suppression of image displacement by studying detectability of scene changes and masking of interactive graphical updates. Small horizontal translations were not noticeable when they occurred during large saccades averaging at least 58ms in duration. Even when noticeable, they were not very disturbing for the viewers [Schumacher et al. 2004]. Saccadic suppression of image displacement has also been suggested for natural locomotion and masking redirected walking manipulations in virtual environments (VE). By detecting and tracking the type of eye movement the user makes in a virtual environment, it is possible to subliminally reposition the users. This can happen during a blink or during a saccade [Bolte and Lappe 2015; Langbehn et al. 2016]. Bolte and Lappe studied saccadic suppression of image displacement in an immersive virtual environment by rotating or translating the camera during saccadic eye movements. They used saccadic suppression as a way of making orientation adjustments less noticeable while the user was viewing a virtual environment. Their results showed that participants were less sensitive to transitions during saccades than during fixations. They also showed that during saccades, transitions were less noticeable for rotations of size 5 degrees on the yaw axis and translations of size 0.5 m along the line of gaze [Bolte and Lappe 2015]. Their study only looked at this problem for two degrees of freedom. In a similar work, Langbehn et al investigated reorienting and repositioning users subliminally in a VE during eye blinks. The virtual scene that participants viewed was rotated or translated when they were asked to blink and subjects then indicated the direction in which they were virtually moved. Their results suggested that imperceptible position movements are possible during blinks [Langbehn et al. 2016]. Another study looked at change blindness during eye blinks while the user wore an HMD and showed that detectability of scene changes was not only dependent on the angle of rotation, but also on the layout of the scenes [Bruder and Langbehn 2017]. Saccadic suppression has also been used to reduce gaze-contingent foveated rendering latency in virtual reality applications. Gaze-contingent rendering can enhance the perceived quality of the graphics by focusing rendering effort on regions of high acuity [Albert et al. 2017]. In a recent study, the images of the display were updated according to the new predicted fixation location before the saccade ended, which made delays less visible for the users [Arabadzhiyska et al. 2017].

Our approach in this research uses the fact that during saccades visual sensitivity is reduced and that changes cannot be detected well. We study six directions of transformations corresponding to the canonical directions for the six degrees of freedom. Previous studies only looked at translational and rotational eye movements on one axis, but have not studied user sensitivity to the changes along other rotational and translational axes. By exploring how saccadic suppression depends on the size and direction of the rotational and translational camera changes, we can design applications that imperceptibly manipulate the graphics of a display according to human visual features. These applications include gaze-contingent displays, foveated rendering in virtual reality headsets and manipulation of users' position in a virtual environment [Bolte and Lappe 2015; Bruder and Langbehn 2017].

2 SACCADE DETECTION

Human eyes do not move in a deterministic manner. To detect a saccade, we need eye position estimates obtained through an eye tracking device as well as a robust real-time saccade detection algorithm. There have been different approaches proposed, such as dispersion-based, velocity-based, acceleration-based and area-based algorithms for detection of saccades [Salvucci and Goldberg 2000; Duchowski 2007]. However, some of the methods are not suitable for online saccade detection. Although they may be very accurate, they require that the entire saccades be recorded before performing the classification [Andersson et al. 2017; Nyström and Holmqvist 2010]. In velocity-based algorithms, eye samples are classified based on their point-to-point velocity. Using velocity thresholds, eye movements are classified as low-velocities for fixations and high-velocities for saccades. The time series of eye positions are converted into velocity values by using FIR (Finite Impulse Response) differentiator of velocities over five eye data samples, as in Equation (1). This velocity calculation suppresses the noise in the eye tracking data [Engbert and Kliegl 2003].

$$v_n = \frac{x_{n+2} + x_{n+1} - x_{n-1} - x_{n-2}}{6\Delta t} \tag{1}$$

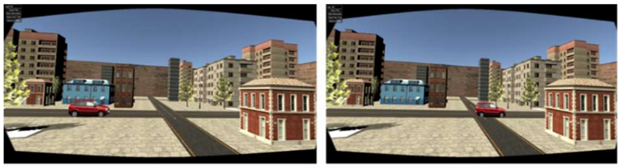
In this study, we used a velocity based saccade detection algorithm. We use an EyeLink 1000 [EyeLink Research Ltd, Oakville, Ontario, CA] to sample the user's eyes with a sampling frequency of 1000Hz. The display update rate is 120 Hz, implicating one sample every 8.33 ms. This means that we receive one eye tracker sample every millisecond (Δt) and therefore we have 8 samples in every refresh update of the display. Hence we are able to perform a five point FIR differentiator for velocity computation. We computed velocity from horizontal and vertical components of eye sample positions. Then we scaled them by the instantaneous pixel per degree, which we calculated for each current eye sample. We applied elliptic thresholds to the horizontal and vertical velocity components using a median estimator [Engbert and Mergenthaler 2006]. We applied the scene change when more than three subsequent eye samples were detected outside the ellipse determined by the horizontal and vertical thresholds. By carefully adjusting the parameters, we suppressed noise and reduced false positives. We verified the reliability of our algorithm by running a pilot session and further adjusting the parameters.

3 GENERAL METHODOLOGY

The main goal of this study is to discover user sensitivity to the type of image change, particularly between different natural 3D scene transformations.

3.1 Stimuli

The virtual environments that the users viewed consisted of two 3D scenes as shown in Fig. 1. In Fig. 1(a), the participants were asked to look at a car in the scene and follow it with their eyes when it jumped horizontally to the right. Similarly, in the second scene, they had to look at the flowerpot, as in Fig. 1(b).


(a)

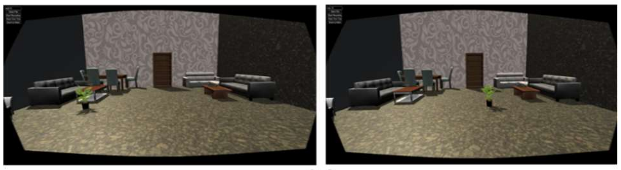

(b)

Figure 1: The scenes: (a) Outdoor Scene, (b) Indoor Scene

We used two scenes to add some variety to the scenes participants viewed. In both scenes, the participants made a 15-degree horizontal saccade. We were interested to see users' sensitivity to changes along the roll (clockwise and counter-clockwise camera rotations), the pitch (upward and downward camera rotations) and the yaw axis (right and left camera rotations) during a horizontal saccade. Moreover, we looked at translational movements in depth (forward and backward camera translations), the vertical (upward and downward camera translations) and the horizontal axis (right and left camera translations).

3.2 Apparatus

Stimuli were presented on a 27-inch 3D Samsung LCD monitor, with a resolution of 1920H*1080V pixels. All visual environments were created on a desktop computer with AMD FirePro W9000 FireGL, Windows 7, Intel Core CPU 3.50 GHz and 3.50 GB RAM. Eye movements were recorded with a video-based system EyeLink 1000 [SR Research Ltd, Oakville, ON, Canada] with a sampling rate of 1000 Hz. We used the tower-mount setup. In addition, a chinrest was used for stabilizing subjects' head and maintaining the 55 cm viewing distance. The experiment was designed in Unity3D and programmed in C# scripts.

3.3 Procedure

The overall experiment was designed to test users' sensitivity to natural image transformations when viewing a 3D scene. While the user viewed a 3D scene, they were asked to look at an object in the scene and follow it as it jumped from one point to another. We studied 6 translation directions for each of two sizes of 0.5 m and 1.5 m, as well as 6 rotation directions for both 2 degrees and 7 degrees rotation magnitudes. One session consisted of 72 trials, and in every session each condition was repeated three times in random order. Each participant attended three sessions. Every session started with a training task to remind the participant how each camera movement in the scene appeared. Once we ensured the participant was clear about their task, we performed a calibration which involved fixating 9 points displayed sequentially on the screen in pseudo-random order. Upon a successful calibration, we continued immediately to the validation step, in which another set of 9 points appeared randomly and sequentially on the display. Then the first trial began. The duration of each trial was 2.0 seconds and the object was displaced after 1.0 second. As the participant looked at the object being displaced, they performed a saccade of 15 degrees of visual angle. At the same time as this saccadic eye movement and upon detection of a start of a saccade, a translation or rotation in a specific direction was applied to the scene (i.e. to the virtual camera). Participants were asked to indicate in which direction they detected the camera change. They had to choose one of the eight directions in the forced-choice question and indicate their confidence level for their answer, on a spectrum of "Not Confident" to "Very Confident". They then proceeded to the next trial. In trials where they did not notice any changes, they were still required to guess the direction of camera change.

There were catch trials in each session. In such trials, no display updates occurred but the users were still asked to indicate the direction of camera change during that trial. In cases where a saccade was not detected in a trial, the current trial was repeated until a saccade was detected. However, the users did not notice this and answered these trials in the same manner. These trials were counted as additional catch trials.

3.4 Participants

Ten university students with normal or corrected to normal vision (5 female and 5 male, with average age of 25.4 ranging from 20 to 32) participated in this experiment. Participants had normal or corrected to normal vision; 6 participants had normal vision and 4 habitually wore glasses for myopia. However, they did not wear glasses nor contact lenses for the duration of the experiment as they reported the monitor was close enough to them to view it clearly. All participants signed a written informed consent form prior to the start of experiment in accordance with a protocol approved by the York University Ethics Board and received a compensation for their participation.

4 RESULTS

The users' answers for direction of camera change, and their level of confidence for each answer was recorded and analyzed. We were interested in users' sensitivity to different sizes of transitions and different directions of change. We only processed the answers in which a saccade was detected and hence a change applied. In some cases, users were not able to distinguish camera changes of different signs along the same axes from each other. For instance, they knew the scene change was horizontal, but they were not sure if it was to the right or left. Therefore, we considered user answers which indicate the correct axis, as correct responses regardless of sign. In many cases, the users were not able to determine the scene change correctly and were not sure if there was a change. As the question was a forced choice one, they selected a direction for those trials. A t-test on all participants confidence levels for correct direction guesses and wrong direction guesses shows that they were significantly more confident about their correct guesses ($M=69.28$, $SD=16.38$) than incorrect guesses ($M=24.36$, $SD=10.54$) conditions, $t(9)=14.5$, $p<0.0001$, $r=0.8$. In total, participants had a confidence level of 72.42% for their correct direction guesses for the translational scene changes and 67.29% for rotational changes. In translations of size 0.5 m and in rotations of size 2 degrees, users only detected 17.03% of display updates correctly. This number rises to 39.81% for large transitions including translations of size 1.5 m and rotations of 7 degrees. Fig. 2 shows the percentages of correctly detected changes for each direction for both translational and rotational display changes, as well as the average level of users' confidences about their answers. This diagram shows the detection rates for all sizes. The most detected camera changes were rotations along the roll axis (clockwise and counter-clockwise rotations). Upward and downward rotations and forward and backward translations were more detectable for the users than horizontal translations and rotations. Horizontal camera shifts are movements on the same axis as user's saccade during each trial. Participants had lower confidence levels for these camera changes.

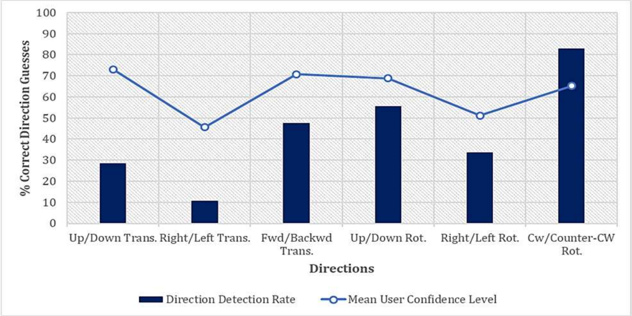


Figure 2: Percentage of correct responses for the six directions of translational and rotational eye movements, and the average level of confidence for correct responses of each direction.

The results show that for translation changes, the larger forward and backward updates were more detectable for the users, while the horizontal changes were detected the least. For rotational display shifts, changes in the yaw were the hardest to detect. Rotations of 7 degrees were easier for users to detect. The users' levels of confidence increased as the size of translation and rotation increased. A Chi-Square test revealed that there was a significant difference in detection rate in the two different sizes of translation scene changes along each of the three axes, ($\chi^2(2)=8.68$, $p<0.05$). In addition, there was a significant difference in detection rate between different sizes of rotation scene changes along each axes ($\chi^2(2)=48.4$, $p<0.001$).

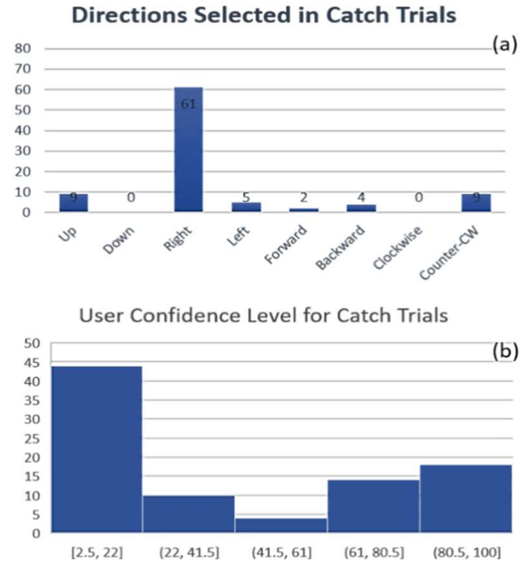


Figure 3: (a) Frequency of directions selected in catch trials. (b) Percentage of user confidence levels for their answers during the catch trials.

There were catch trials in the experiment, in which there were no changes to the display graphics, but the user was still asked to indicate scene change direction. Results show that in 90 catch trials, participants answered with confidence levels of larger than zero and thought they had seen a camera change. Participants' average level of confidence for their responses in catch trials was 39.83%. As demonstrated in Fig. 3(a) users selected the rightward translational and rotational scene change in 61 trials which is 4.29% of all the catch trials. But they never indicated downward and clockwise transitions. Fig. 3(b) shows the user confidence levels for their answers in catch trials. Most users were not very confident about their guess of the scene change. However, in more than 35 catch trials, users made a guess with more than 60% confidence about their answers.

5 DISCUSSION AND CONCLUSIONS

The results of this research show that change detectability depends on rotation angle and translation size, as well as type of transformation. We found that users are less sensitive to certain image transformations while they make saccadic eye movements. These results are consistent with the previous research [Bolte and Lappe 2015; Allison et al. 2010]. Saccadic suppression of image displacement is larger with bigger saccades and smaller target displacements [Stark et al. 1976]. Results of the current experiment show that when viewing a 3D scene, users are sensitive to scene transitions that occur during a saccade; and can recognize direction of changes in which the camera makes large angles of rotation (7 degrees) or sizes of translation (1.5 m) as compared to smaller ones of 2 degrees rotations and 0.5 m translations.

We translated or rotated the user's viewpoint during saccadic eye movements. When participants were viewing the scene, they did not notice most of the reorientations (rotations) of 2 degrees and translations of 0.5 meters during saccades. However, it was different for clockwise and counter-clockwise rotations of the camera along the roll axis, as these were more obvious, changed the simulated standing angle of the user, and felt very unnatural. These camera shifts have larger peripheral motions. It is worth noting that the participants may not have noticed the camera shift during the saccade itself, but guessed the direction of camera change correctly after the saccade had ended. This is because when the camera moves clockwise or counter-clockwise it is much easier to detect its change of position as the image rotated relative to the display. This could also be the case for large rotational changes by noticing parts of the image shifting on or off the display. The users mentioned that even though they did not see the image shift, they could guess the direction of its change correctly according to the image they saw after their saccade had ended. This might be different with a larger field of view or when wearing an HMD where frame cues are typically weaker, and these before and after images will be much less detectable, especially when the users are not aware that there are any display updates. This could be explored in a future experiment.

In the catch trials participants believed they saw horizontal changes as the object in the scene moved horizontally from left to right. This could be because users expected some amount of shift in the image with their eye movement, or because they were less certain of small image shifts in this direction on top of the large retinal image shift produced by the eye movement. As the object in the scene and users' eyes moved horizontally to the right, users may have interpreted the motion of the object as image shift. When a viewer and a fixation stimulus experience simultaneous acceleration, the fixated stimuli seem to move in the direction of acceleration despite of no physical movement relative to the observer. So during self-motion, fixation on a stationary environment results in perceived object motion [Whiteside et al. 1963; Post and Leibowitz 1982]. In fact these may be "illusional movements" which are a part of autokinetic effect caused by a short-term imbalance of the neural systems directly concerned with the visual registration of movement [Gregory and Zangwill 1963]. Overall, since the participants were given a forced-choice question to guess a direction, they acted very cautiously and tried to detect any movements in the scene and make as many correct direction guesses as possible and attributed them to the noisy horizontal direction when uncertain or unseen.

Gaze-contingent hiding of graphics updates is an application of eye tracking that can be used to improve the design of 3D virtual environments. Accurate eye tracking technology like that employed in the current study is increasingly being integrated into VR displays and provides a framework for producing such displays as well as other interactive virtual reality and gaze-contingent applications. Based on the results of this research, saccadic suppression is an applicable tool for hiding graphics updates when users view a 3D virtual setting through an eye tracker. There are image transformations in certain sizes that are more apparent and recognizable for the viewers, such as rotations along the roll axis.

In future research it would be interesting to investigate user sensitivity in less controlled environments with more natural viewing behaviors and saccades in arbitrary sizes and directions. Also, we can look at effect of visual depth cues, such as occlusion, shadows, straight lines or arrangement of the objects in the scene on the detection of the scene transitions.

ACKNOWLEDGMENTS

The authors of this work would like to acknowledge Natural Sciences and Engineering Research Council of Canada (NSERC) grant RPGIN-2015-06732 to R. Allison.

REFERENCES

Albert, R., Patney, A., Luebke, D., and Kim, J. 2017. Latency Requirements for Foveated Rendering in Virtual Reality. *ACM Transactions on Applied Perception 14*, 4, 1–13.

Allison, R.S., Schumacher, J., Sadr, S., and Herpers, R. 2010. Apparent motion during saccadic suppression periods. *Experimental Brain Research 202*, 1, 155–169.

Andersson, R., Larsson, L., Holmqvist, K., Stridh, M., and Nyström, M. 2017. One algorithm to rule them all? An evaluation and discussion of ten eye movement event-detection algorithms. *Behavior Research Methods 49*, 2, 616–637.

Arabadzhiyska, E., Tursun, O.T., Myszkowski, K., Seidel, H.-P., and Didyk, P. 2017. Saccade Landing Position Prediction for Gaze-contingent Rendering. *ACM Trans. Graph. 36*, 4, 50:1–50:12.

Bolte, B. and Lappe, M. 2015. Subliminal reorientation and repositioning in immersive virtual environments using saccadic suppression. *IEEE transactions on visualization and computer graphics 21*, 4, 545–552.

Bruder, G. and Langbehn, E. 2017. Subliminal Rotations During Eye Blinks for Redirected Walking. *Journal of Vision 17*, 10, 1266–1266.

Duchowski, A.T. 2007. *Eye tracking methodology: theory and practice*. Springer, London.

Engbert, R. and Kliegl, R. 2003. Microsaccades uncover the orientation of covert attention. *Vision Research 43*, 9, 1035–1045.

Engbert, R. and Mergenthaler, K. 2006. Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences 103*, 18, 7192–7197.

Franke, I.S., Günther, T., and Groh, R. 2014. Saccade Detection and Processing for Enhancing 3D Visualizations in Real-Time. *International Conference on Human-Computer Interaction*, Springer, 317–322.

Gregory, R.L. and Zangwill, O.L. 1963. The Origin of the Autokinetic Effect. *Quarterly Journal of Experimental Psychology 15*, 4, 252–261.

Herpers, R., Schumacher, J., and Allison, R.S. 2004. Hiding graphic updates during long saccadic suppression periods. *Dyn. Percept*, 77–82.

Ibbotson, M. and Krekelberg, B. 2011. Visual Perception and Saccadic Eye Movements. *Current opinion in neurobiology 21*, 4, 553–558.

Langbehn, E., Bruder, G., and Steinicke, F. 2016. Subliminal Reorientation and Repositioning in Virtual Reality During Eye Blinks. ACM Press, 213–213.

Leigh, R.J. and Zee, D.S. 2015. *The Neurology of Eye Movements*. Oxford University Press.

Nyström, M. and Holmqvist, K. 2010. An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods 42*, 1, 188–204.

Post, R.B. and Leibowitz, H.W. 1982. The effect of convergence on the vestibulo-ocular reflex and implications for perceived movement. *Vision Research 22*, 4, 461–465.

Salvucci, D.D. and Goldberg, J.H. 2000. Identifying Fixations and Saccades in Eye-tracking Protocols. *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*, ACM, 71–78.

Schumacher, J., Allison, R.S., and Herpers, R. 2004. Using Saccadic Suppression to Hide Graphic Updates. *EGVE*, 17–24.

Stark, L., Kong, R., Schwartz, S., and Hendry, D. 1976. Saccadic suppression of image displacement. *Vision Research 16*, 10, 1185–1187.

Triesch, J., Sullivan, B.T., Hayhoe, M.M., and Ballard, D.H. 2002. Saccade Contingent Updating in Virtual Reality. *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, ACM, 95–102.

Westheimer, G. 1954. MECHANISM OF SACCADIC EYE MOVEMENTS. *A.M.A. Archives of Ophthalmology 52*, 5, 710–724.

Whiteside, T.C., Graybiel, A., and Niven, J.I. 1963. *VISUAL ILLUSIONS OF MOVEMENT.* NAVAL SCHOOL OF AVIATION MEDICINE PENSACOLA FL, NAVAL SCHOOL OF AVIATION MEDICINE PENSACOLA FL.