Proceedings of the

Fourth *Starkfest* Conference on Vision and Movement in Men and Machines

School of Optometry
University of California, Minor Hall, Berkeley
May, 26 and 27, 2005

Foreword

The papers in this publication stem from the Fourth International Conference for Vision in Movement in Man and Machines, informally known as "Starkfest Memorial." This conference series began with a meeting at the Univeristy of California, Berkeley in 1994: a conference designed to honor the scientific achievements of Larry Stark on the occasion of his "official retirement." This first conference brought together many of Larry's former students and colleagues and proved to be both intellectually stimulating and socially gratifying. A smaller version of the original, termed "Starkfest Reunion" was held several years later as a satellite of the annual IEEE Engineering in Medicine and Biology meeting held in Chicago. A special feature of this conference was a group of presentations by young investigators that were students of students of Larry. This group of presentations was labeled "Starkfest, the Next Generation," reflecting the whimsy and humor carried within Larry and infected into many of his students. The third conference benefited from an outstanding setting, the Vieux Port in Marseille, France. This conference was the largest and most comprehensive of the series and included presentations in higher level and adaptive control including internal models, motor coordination, telemanipulation and virtual reality, and neruoinformatics. An effort was made to include presenters outside of Larry's vast circle: those not fortunate enough to have had Larry as a colleague or mentor.

The current conference reflects Larry's varied interests as well as many of his seminal ideas. Several papers are based on concepts originally developed by Larry such as the scanpath or "dry dissection." Other papers are derivative of Larry's pioneering work in the areas of virtual reality, neural modeling, optimal control, and the relationship between perception, eye movements and internal models (with particular relevance to pitching a baseball). One paper, on Larry's recent interest in visual processing, features Larry as a co-author and reflects the interest he continued to maintain in scientific study. A series of eclectic presentations featured research in neruoinformatics, cultured neural networks, territorial biology, and the study of early Korean water clocks. While some of these topics may have been beyond Larry's traditional area of study, there were not beyond his areas of interest!

In addition to the exchange of important scientific ideas, this conference was designed to honor the lifelong achievements of Larry Stark and his unique qualities as human being and friend. The conference began with a "Remembrances session" where

conference participants, family and friends informally traded stories and photos sharing their interactions with this remarkable individual. Several social activities facilitated interaction between conference participants who held a common bond with Larry.

While Larry was a great scientist, he was foremost an inspiring colleague, mentor, and friend. The library in the School of Optometry at Berkeley has a multi-volume series of the "Collected Works of Lawrence Stark." But these static works on paper do not represent his true contribution to science: the many students he has mentored and colleagues he has inspired. As with the previous conferences, Starkfest Memorial was held with the intention of "collecting" the truly most important works of Lawrence Stark: his students and colleagues.

John L. Semmlow, Ph.D., Program Chair
Robert Wood Johnson Medical School and Rutgers University, New Jersey, USA
June 14, 2005

Index:

What the Batter's Eye Tells the Batter's Brain

A. Terry Bahill and Dave Baldwin, University of Arizona

Abstract— **The batter must identify the speed and spin of the baseball, and use this information to calculate when and where it will cross the plate**

***Index Terms*—Visual theories, perception, science of baseball**

I. Introduction

For most pitches in baseball (excepting the knuckleball and other low friction pitches) the pitcher uses one of two grips: the two-seam grip or the four-seam grip [3]. Pitchers often say that using a two or four-seam grip causes the ball to move differently. When explaining baseball phenomena, we first use principles of Physics, then Physiology and finally Psychology. OK, Physics first. Most wind tunnel tests have shown little difference in forces between the two-seam and the four-seam orientation [6]. Simulations of 74 mph pitches needed differences in order to match the experimental data [1]. However, [5] summarize the literature with "the effect of seam orientation … is only present at low spin."

Physiological differences manifest in the grip. If two-seam and four-seam fastballs move through the air positioning of the fingers and thumb before release that differently, it is not due to the orientation of the seams of the ball in flight. However, it may be due to the fact that different orientations could produce different spin rates and spin axes. For example, with the two-seam fastball, the thumb is often on a smooth leather part of the ball. Whereas with the four-seam fastball, the two fingers and the thumb can all be on seams; therefore, the pitcher may be able to get more spin on the ball and consequently it will not drop as far as a two-seam fastball. However, this is a very small effect. Pitchers also often apply different finger pressures to the two-seam and four-seam fastballs, which could also produce small differences in spin axes and spin rates.

II. Methods and Materials

Psychological differences are based on the batters perception of the ball. The ball's appearance to the batter is different for the two-seam and the four-seam fastballs. We skewered baseballs on bolts in two-seam and four-seam orientations. The bolts were chucked in electric drills and were rotated at 1200 rpm. The speed of rotation was measured with a stroboscope. This setup is shown in Fig. 1.

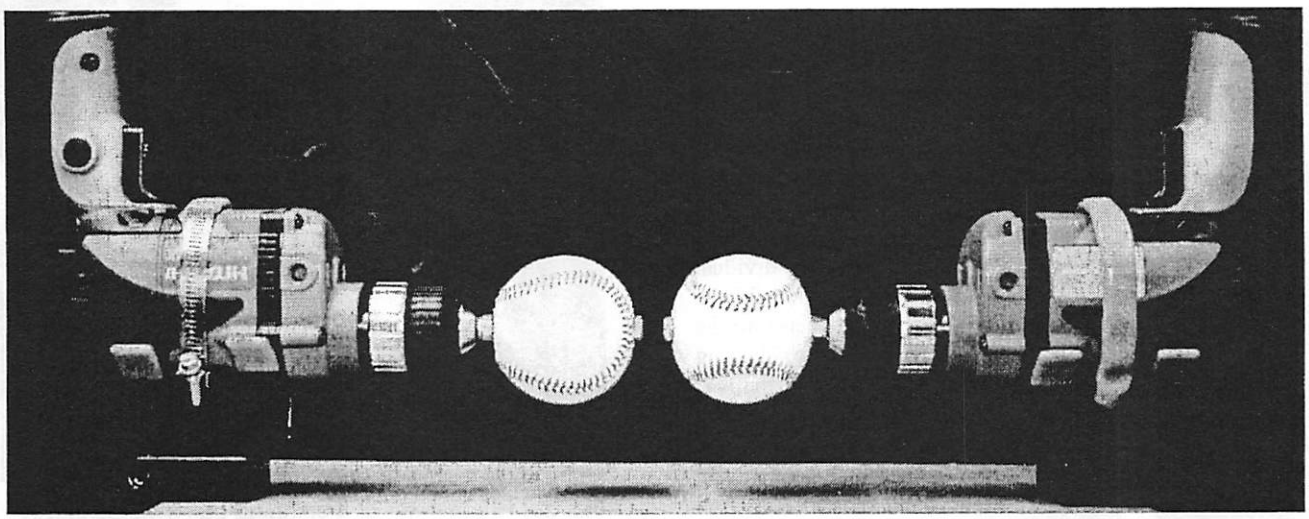


Fig. 1. Experimental setup for the four-seam (left) and the two-seam orientation (right). Photo credit Zach Bahill.

Corresponding author: Terry Bahill, Systems and Industrial Engineering, University of Arizona, Tucson Arizona, 85721-0020. terry@sie.arizona.edu.

It was difficult to put bolts through the exact centers of the balls. We restrained the balls with a jig made of two 5-inch floor flanges for 2-inch pipe and four 5-inch bolts. We drilled holes with a dual point drill bit. The professional league balls (with a rubber coating around the cork center) were frozen before drilling. Then 4-inch bolts were inserted through the balls and the nuts were secured with lock washers.

The following table shows the speed and spin rates for several popular pitches as thrown by major-league baseball players. The speed and spin rates decrease about 10% in route to the plate. We assume the pitcher's release point is five feet in front of the pitcher's rubber and the bat contacts the ball one and a half feet in front of home plate.

Numerical values for typical major league pitches				
Type of pitch	Initial Speed (mph)	Typical spin rates (rpm)	Typical spin rates (revolutions per second)	Rotations in route to plate
Fastball	85 to 95	1200	20	8
Slider	80 to 85	1400	23	10
Curveball	70 to 80	2000	33	17
Change-up	60 to 70	400	7	4
Knuckleball	60 to 70	60	1	½

III. EXPERIMENTAL RESULTS

By visual observation and in photos shown in Fig. 2, the four-seam fastball appears to be a grey blur with thin vertical red lines about 1/7 of an inch apart running perpendicular to the spin axis. These lines are the individual stitches of the baseball. However, we do not suggest that the batter sees these individual stitches. In contrast, the two-seam fastball exhibits two big red stripes, each about 3/8 of an inch wide. These stripes are evident because they represent seams rather than individual stitches. They provide easily perceived information for the batter to determine the angle of the spin and therefore the direction of the resultant deflection. Therefore, the big difference between the two-seam and four-seam fastballs is that (because of the visibility of vertical red stripes) the batter may be able to perceive the spin direction on the two-seam fastball quicker. For simplicity, the simulated fastballs of this paper were thrown with an overhand delivery. The more common three-quarter arm delivery would tilt the axis of rotation by 45 degrees.

The pitcher releases the ball about five feet in front of the rubber and the batter hits the ball about 1.5 feet in front of the plate. So a 95 mph fastball is in the air about 450 msec. The pitch is divided into thirds. In the first third, the batter gathers sensory information. In the middle third, he computes where and when the ball will be at the predicted collision point and he decides whether or not to swing. The last third is the swing. Because the batter has about one-eighth of a second to gather visual information, the photographs of Figure 2 were shot with a shutter speed of one-eighth of a second.

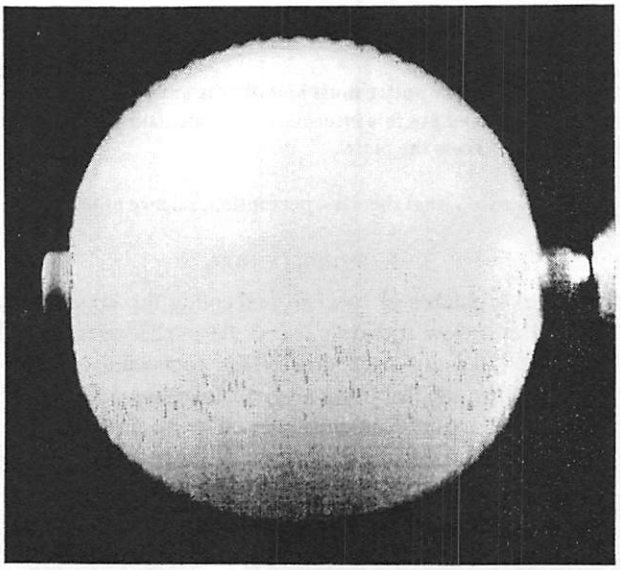


Fig. 2a. Photograph of a spinning ball simulating a fastball thrown with a four-seam grip. The ball is being rotated at 1200 rpm (20 times per second). The exposure time was 1/8 second. Photo credit Zach Bahill.

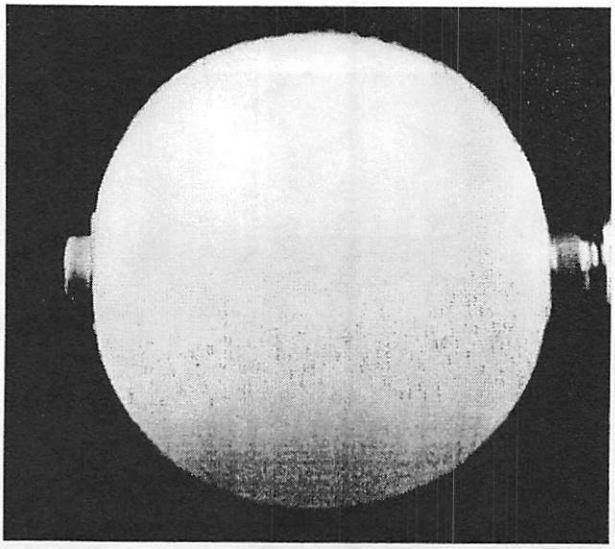


Fig. 2b. Photograph of a spinning ball simulating a fastball thrown with a two-seam grip. The ball is being rotated at 1200 rpm. The exposure time was 1/8 second. Photo credit Zach Bahill.

However, the difference in appearance of the two-seam and four-seam orientations is more dramatic for video motion than it is in still photographs. See

http://sie.arizona.edu/sysengr/baseball/2Seam-4Seam-Video.AVI. Moreover, the difference in appearance of the two-seam and four-seam orientations is more dramatic for spinning baseballs in our laboratory than it is in this video. Therefore, we think that there might be something in addition to the big red stripes that distinguishes the two-seam from the four-seam fastball. We hypothesize that this additional clue might be related to the critical flicker fusion frequency.

As the frequency of a blinking light increases, the light appears to flicker and then, at a certain frequency, it appears to be continuously illuminated. This transition point is the critical flicker fusion frequency (measured in Hertz, Hz: one Hz equals one pulse per second). For a person in a baseball park this frequency is probably between 40 and 50 Hz. Television screens present a new frame sixty times per second (60 Hz); therefore, their pictures do not flicker, whereas the time indicator on your VCR probably blinks once a second. This pulse rate clearly produces what is perceived as a blinking sequence. (At the beginning of the twentieth century, movies were called *the flicks*, because their 24 Hz frame rate produced a flickering of the images.)

A typical major league fastball rotates about 1200 rpm, or 20 times per second. For a two-seam fastball, the pair of seams that straddles the narrow isthmus of the ball and is parallel to the spin axis would cross the field of view once on each rotation. These seams lie so close together that they probably are seen as a single item. Therefore, the frequency of this pulse would be around 20 Hz, which is below the critical flicker fusion frequency, and perhaps the ball would appear to flicker, giving the batter a clue about the spin. He does not have to compute the spin rate to determine whether the pitch is a fastball or curveball: he just has to determine if it has topspin (a curveball) or backspin (a fastball). This helps the batter predict the movement of the ball. In contrast, each of the seams of a four-seam fastball would cross the field of view symmetrically for each rotation. This would produce a frequency of 80 Hz, which is above the critical flicker fusion frequency. Therefore, the ball would not flicker and the batter would not have this extra clue about the spin. The batter would have to guess if it were a curveball or a fastball. If he guesses incorrectly that a fastball is a curveball, he will think the pitch is slower than it actually is (and therefore will fall farther than it actually will). Consequently, when he discovers that the pitch is higher than he predicted, he might perceive a rising fastball [4]. This could be the reason that pitchers often say the four-seam fastball rises.

For the simulated fastballs in our laboratory, the stripes of the two-seam fastball were visible farther away than the thin red lines of the four-seam fastball, roughly 16 feet versus 10 feet. These data are non-athletes with normal vision. Professional baseball players undoubtedly have better dynamic visual acuity and can probably see the red stripes much farther away. For a professional fastball, the batter's swing starts when the ball is about 19 feet away from the batter. Information that the batter gains after the swing begins cannot help him during that swing.

The batter's visual acuity is vital for gathering information about the spin on the ball. Visual acuity can be classified as static (SVA) or dynamic (DVA) – two visual characteristics that are not directly correlated. SVA is used by an organism to obtain information about static objects (such as a lion observing grazing zebras). DVA is the ability to perceive moving objects (important to a lion ambushing a herd of stampeding gazelles). DVA is required to track and predict the flight of a baseball. Experienced athletes have better than average DVA, in part because of societal selection for this ability, and in part because DVA can be improved with training, but our survey of major league hitters indicated considerable variability among these players in their ability to see the spin on a pitch. Batters with good DVA can see the spin on the ball; those with poor DVA cannot. Most of us can read the label on a phonograph record turning at 33 rpm – this would be about the limit of our capabilities. The great hitter, Ted Williams, could read one turning at 78 rpm – far beyond the DVA of the average person.

IV. CONCLUSIONS

Physics and Physiology show negligible differences between the two-seam and the four-seam fastballs. The big differences seem to be Psychological, specifically perceptual. The batter can see the two red stripes of the two-seam fastball and the ball will seem to flicker. Both of these clues alert the batter to the type of spin on the ball and help him to predict its movement.

Our generalizations are that the pitcher should use a four-seam grip for fastballs and curveballs, to remove the perceptual clue of the two red stripes.

REFERENCES

[1] Alaways, L. W. and Hubbard, M., Experimental determination of baseball spin and lift, *Journal of Sports Sciences*, 19, 349-358, 2001.

[2] Bahill, A.T. and D. G. Baldwin, The Rising Fastball and Other Perceptual Illusions of Batters. In: *Biomedical Engineering Principles in Sports*. (Ed G. Hung and J Pallis), Kluwer Academic, New York, 2004.

[3] Bahill, A. T., Baldwin, D.G. and J. Venkateswaran, Predicting a Baseball's Path, *American Scientist*, 93: 218-225, May-June 2005.

[4] Baldwin, D.G. and A. T. Bahill, A Model of the Bat's Vertical Sweetness Gradient. In: *Proceedings of the 5th Conference of Engineering of Sport*, Davis, California, September 2004, ISEA, Sheffield, UK, 2004.

[5] Sawicki, G. S., M. Hubbard, and W. J. Stronge, How to hit home runs: Optimum baseball bat swing parameters for maximum range trajectories, *Am. J. Phy.* 71(11), pp. 1152-1162, 2003.

[6] Watts R.G. and Bahill, A.T. *Keep Your Eye on the Ball: Curve Balls, Knuckleballs and Fallacies of Baseball*. New York: W. H. Freeman, 2000, page 75.

Terry Bahill (S'66-M'68-SM'81-F'92) is a Professor of Systems Engineering at the University of Arizona in Tucson. He received his Ph.D. in electrical engineering and computer science from the University of California, Berkeley, in 1975. Bahill has worked with BAE SYSTEMS in San Diego, Hughes Missile Systems in Tucson, Sandia Laboratories in Albuquerque, Lockheed Martin Tactical Defense Systems in Eagan MN, Boeing Information, Space and Defense Systems in Kent WA, Idaho

National Engineering and Environmental Laboratory in Idaho Falls and Raytheon Missile Systems in Tucson. For these companies he presented seminars on Systems Engineering, worked on system development teams and helped them describe their Systems Engineering Process. He holds a U.S. patent for the Bat Chooser, a system that computes the Ideal Bat Weight for individual baseball and softball batters. He is a Fellow of the Institute of Electrical and Electronics Engineers (IEEE) and of the International Council on Systems Engineering (INCOSE). He is the Founding Chair Emeritus of the INCOSE Fellows Selection Committee.

His picture is in the Baseball Hall of Fame's exhibition *Baseball As America.*

Dave Baldwin, as a relief pitcher for the Washington Senators, Milwaukee Brewers and Chicago White Sox during the 1960s and 1970s, compiled a lifetime ERA of 3.08 with 6 wins, 11 loses and 22 saves in 176 games. Following a 16-year professional baseball career, he earned a Ph.D. in genetics and an M.S. in systems engineering from the University of Arizona.

Foveal versus Parafoveal Scanpaths of Visual Imagery in Virtual Hemianopic Normal Subjects viewing Artful Pictures

Wolfgang H. Zangemeister and Thomas Liman

Department of Neurology, Clinical Neuroscience Unit, University Hamburg, FRG

Abstract

Applying a sophisticated string analysis such as compressed Regional String Analysis [cRSE] permits to show differentially the many significant differences that follow therapeutical masking of the foveal region during a [virtual, model] hemianopia. The visual imagery scanpath is done over a compressed mental image that needs longer fixation duration but fewer saccades than the real image. The combination of different viewing tasks with different types of pictures permits to show how top down strategies of the scanpath can be enforced or decreased by a proper combination of task and picture; this is basically influenced by the mask of fovea versus no-mask of fovea difference, with bottom up mechanisms becoming more important with the additional loss of foveal viewing strategies in the mask condition.

1 INTRODUCTION

Scanpath theory: Eye Fixations shown as connected sequence resulting in a sequential string of visited ROIs. Alternating perceptual ROIs (lettered squares)and saccadic eye movements (circles with arrows) shown by solid arrows form the „feature ring" of the scanpath theory in the noniconic model of this example. The variability of ROI-sequences is represented by the dashed arrows

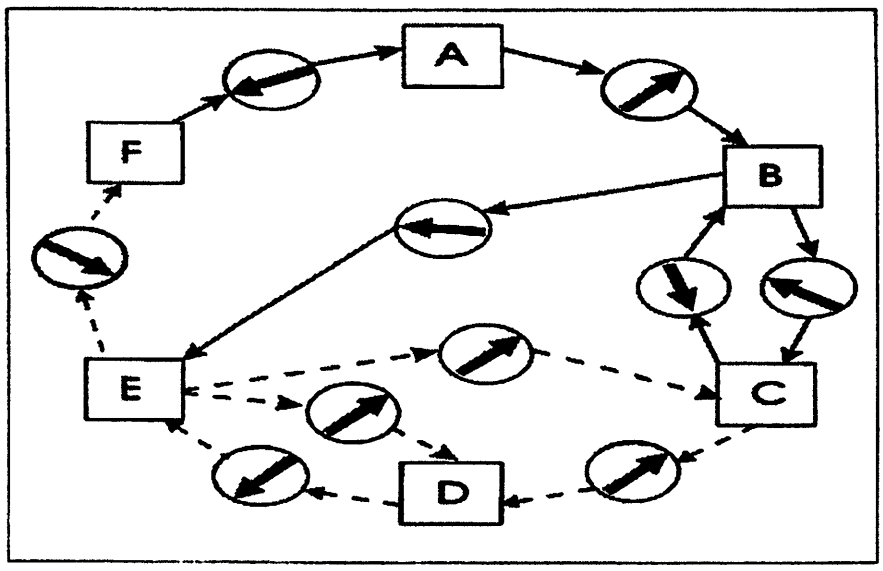


Fig.1: Scanpath Theory (Noton and Stark, 1971)

Several methods to analyse the eye scanpath (13,17,21,25,27) have been tested in the past. Besides the standard analysis like calculation of fixation durations, number of saccades in certain regions of the picture and the global-local ratio of scanpath saccades (25), there are statistical, more regionally weighting methods as the **Markov Analysis [MA]** (5,20) of zero and first order; and methods of **String Editing [SE]** primarily used in linguistics and genetics,

that have been introduced by Stark and others (4,9,19,28,29) since the early nineties into the field of scanpath analysis.

The different editing operations will be weighted in different ways, like pay expense. So, for inserting or deleting one label you have to *pay* 2, for changing a label you *pay* 1. The maximum distance of two strings with n^a, respectively n^b labels result in a similarity of range from 0 to 1 as shown in the following formula

$$D^{ab}_{SE,\max} = n^a \chi + (n^b - n^a)\delta$$

In this formula χ represents the **cost of changing** and δ stands for the **cost of deleting or inserting** a label. In the following formula we get a dimension of similarity :

$$S^{ab}_{SE} = 1 - \frac{D^{ab}_{SE}}{D^{ab}_{SE\max}}$$

Similarity Index

Markov Analysis

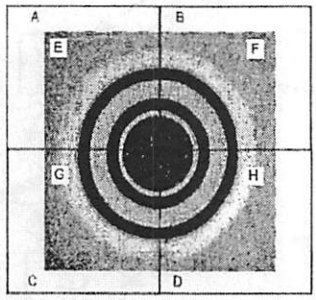
The Markov analysis of zero ordinal calculates the probability that one special ROI will be fixed during image viewing . Markov-analysis of first ordinal calculates the transitional probability that ROI *i* will be fixed when ROI *j* was fixed before. These transitional probabilities p_{ij} can be visualized in matrixes:

$$M = \begin{pmatrix} p_{11} & \cdots & p_{1N} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ p_{N1} & \cdots & p_{NN} \end{pmatrix}$$

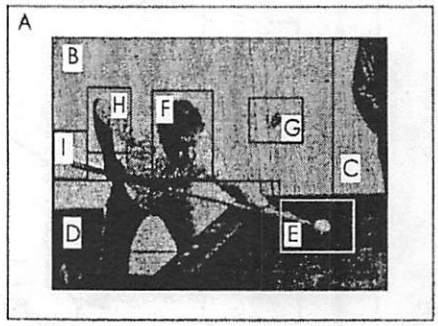
Fig.2: SE (left) and MA (right) algorithms used.

We have used and compared regional **SE (RSE)** and compressed RSE **(cRSE)** where direct repetitions of string-contents are deleted. Further, we have developed a vectorial representation of scanpath strings, Vector SE [VSE], and the weighted VSE [wVSE], where the string vector is weighted according to the frequency of occurrence of similar directions within strings. This latter method has the advantage of working independently of the setting of **a priori** (geometrical) **regions of interest (ROIs)** or **a posteriori** (intelligent, depending on picture content) ROIs **(Fig.3)**. In the case that ROIs are used, it is obvious that an a priori setting of ROIs favours the analysis of **bottom up [BU]** behaviour , and a posteriori setting a **top down [TD]** behaviour of the scanpath string.

Fig.3 a&b: a priori and a posteriori analysis.



Geometric *a priori* ROI definition. "Spring cool," Ken Noland, 1962.



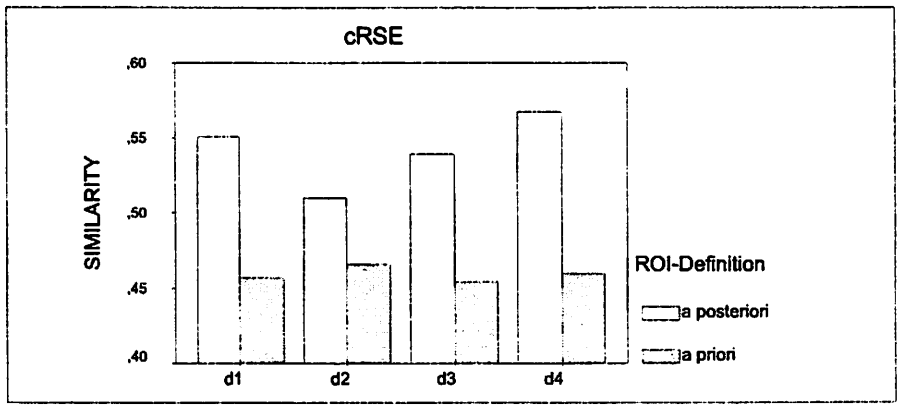
Semantic *a posteriori* ROI definition. "Deadeye," Lane Terry, 1971.

Visual Field Defects. Patients that are blind on one half of their visual field i.e. hemianopic (2,6,15,22,34), usually due to occipital stroke or tumor, show specific oculomotor adaptations due to spontaneous and therapeutic changes that have been reported over the years (12,18,23,26,29). Using specific training regimes, they may be able to adapt and circumvent this half-field blindness, and may also show specific changes in their visual imageries (1,3,7,810,11,18,33). Using a previously (24,30,31,32) described computer simulation of such a visual field defect on a visual display terminal (VDT) we posed the question:
Is it possible to enhance and record the effect of this adaptation through "training" of healthy subjects with a virtual hemianopia by enlargement of the 50% horizontal field defect towards the „healthy" seeing side (SHF) by +5 deg, in analogy to the "forced used therapy" developed

by Taub et al.(35) in hemiparetic patients where patients were forced to use only their paretic limb?

II METHODS

Comparing the different analytical options we decided to use the cRSE for analysis of our recordings in twenty normal subjects experiencing a complete dense 50% virtual homonymous hemianopia,- with ten of them undergoing an additional training of 10 minutes with enforced virtual hemianopia of 55%, whereas the other ten underwent a sham training of the same duration. The technique of inducing virtual hemianopia has been described by us before (24,30,31,32). In short, it consists of an IR-reflection eye movement monitor (0.1° resolution; overall bandwidth 0 – 250 Hz) that has an active link to the 23`` screen that is viewed by the subject and where a dark grey blinding of 50% or 55% respectively of the viewed image occurs directly related to the eye movements (overall latency of the generation of the mask 8msec).
Compressed RSE was obviously superior to the other methods in showing particular string similarity differences between Set1(d1) to Set4 (d4).
[Set1=initial recording, Set2=with mask training or sham training, Set3=control recording after "training" – all on the same day sequentially done; Set4= control recording 14 days later, no training]. For statistical comparisons we used the Mann-Whitney U test between all tasks.



Similarity as a function of ROI definition and set. Highly significantly differences of the ROI Definition by cRSE. A posteriori differentiates better than A priori.

Fig.4

Using a posteriori ROIs had the advantage of showing more sophisticated differences between single Sets. This is demonstrated differentially in **Fig.4,** where the extra-mask-training shows up with a significant ($p<0.001$) similarity decrease at Set2, the time when this training was performed. The *sequence* of pictures that our subjects viewed *[5 sec, 2 sec pause, 5sec visual imagery viewing the blank screen]* consisted of three realistic, three abstract, two search pictures *(see Appendix A&B showing examples of one complete run.)*. The *task* how to view these pictures was: first easy, then detailed, and finally recollection viewing (25).

III RESULTS

Interestingly, an *a priori cRSE* comparison of string similarities,- that favours analytically the BU view-, of the MaskGroup with the NO-MaskGroup between Set1 and Set3,- shows the mask effect more differentially **(Fig.5a):** Mask training is followed by **decreased** similarities, NO-Mask sham training by an **increase** of string similarities . That these significant effects hold true also with the *a priori cMA* analysis, where more regional aspects are weighted using

this method, was additionally demonstrated. This is a demonstration that the extra 5% mask generated more BU dependence of scanning and searching.

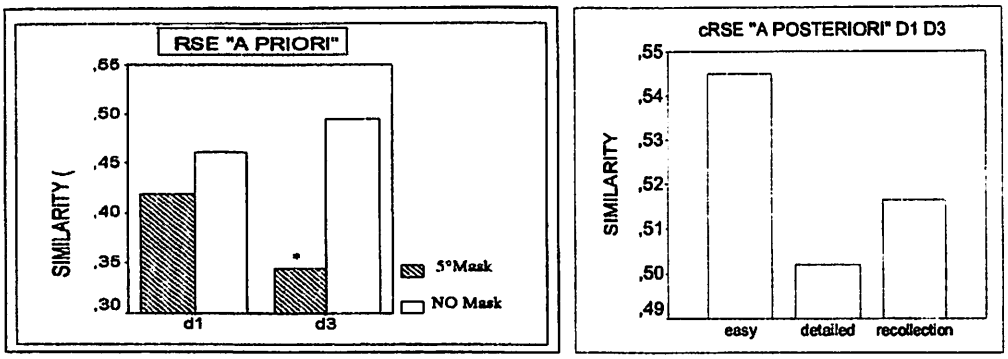
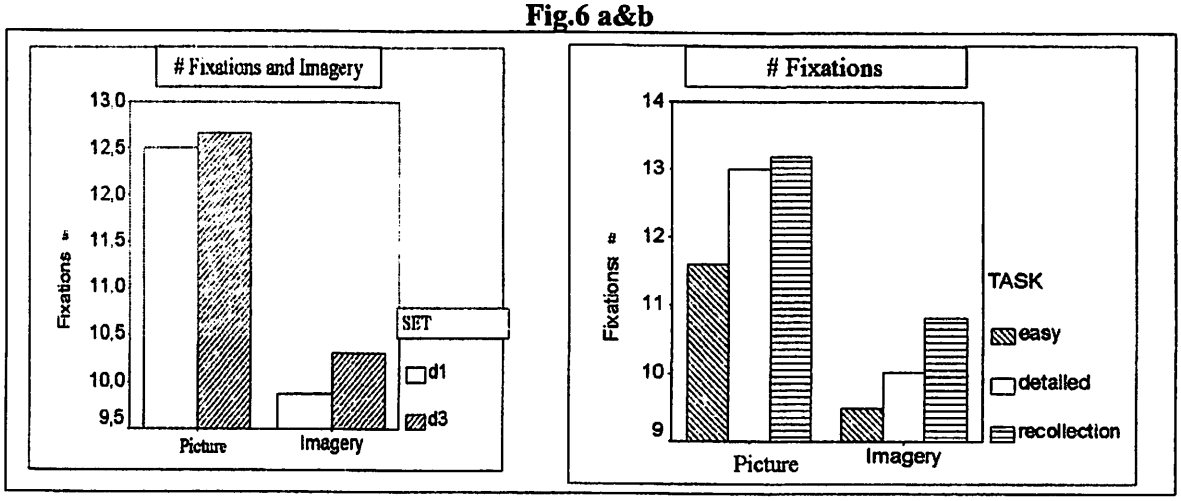


Fig.5 a&b

Results of the cRSE comparison between set1 and set3 with a Posteriori ROIs (**Fig.5b**) showed significant differences ($p<0.01$) between the above task as a function of the intervening mask-, compared to sham-training, whereas the other conditions did not show such similarity differences. Obviously, *detailed* viewing , where foveal fixations are important, was decreased ; whereas *easy* viewing showed increased similarities, where para- and extra-foveal viewing is more often contributing to the similarities of scanpaths.

Number of fixations (**Fig.6a**) did not differ between picture and image, but between Set1 and Set3 – as an indirect sign of the scanpath's decrease of magnitude. However, number of fixations as function of different tasks demonstrated significant differences ($p<0.01$) between easy and detailed or recollection viewing,- "easy" showing the lowest number of fixations during set3 (**Fig.6b**). This demonstrates that with easy looking the simulated hemianopic subjects viewed the stimuli quite similarly than normal subjects (25).

Fig.6 a&b



The **mental image** or **visual imagery** showed the tendency to be miniaturized as described previously by Gbadamosi and Zangemeister (7), **Fig. 7, statistics(a), and example scanpths(b).** The statistical effect shown in the amplitude distribution as a function of mask training and imagery is shown in **Fig.7a**: In general, with visual imagery the amplitudes become significantly smaller. The special effect of the training (set3 versus set1) is that particularly after the training amplitudes decrease.

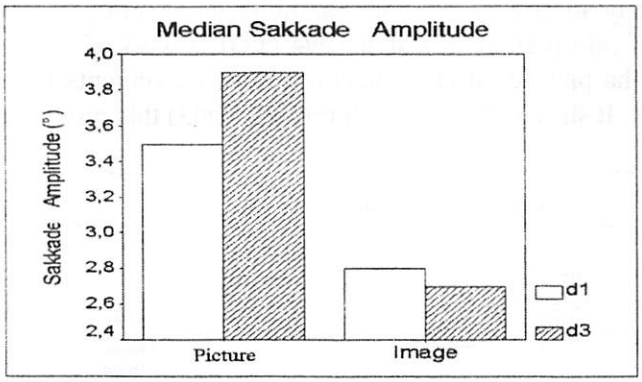


Fig.7a

An example of progressive miniaturisation of the scanpath with imagery 5sec, 30sec, 60sec after the real picture viewing (7) is shown in Fig.7 .

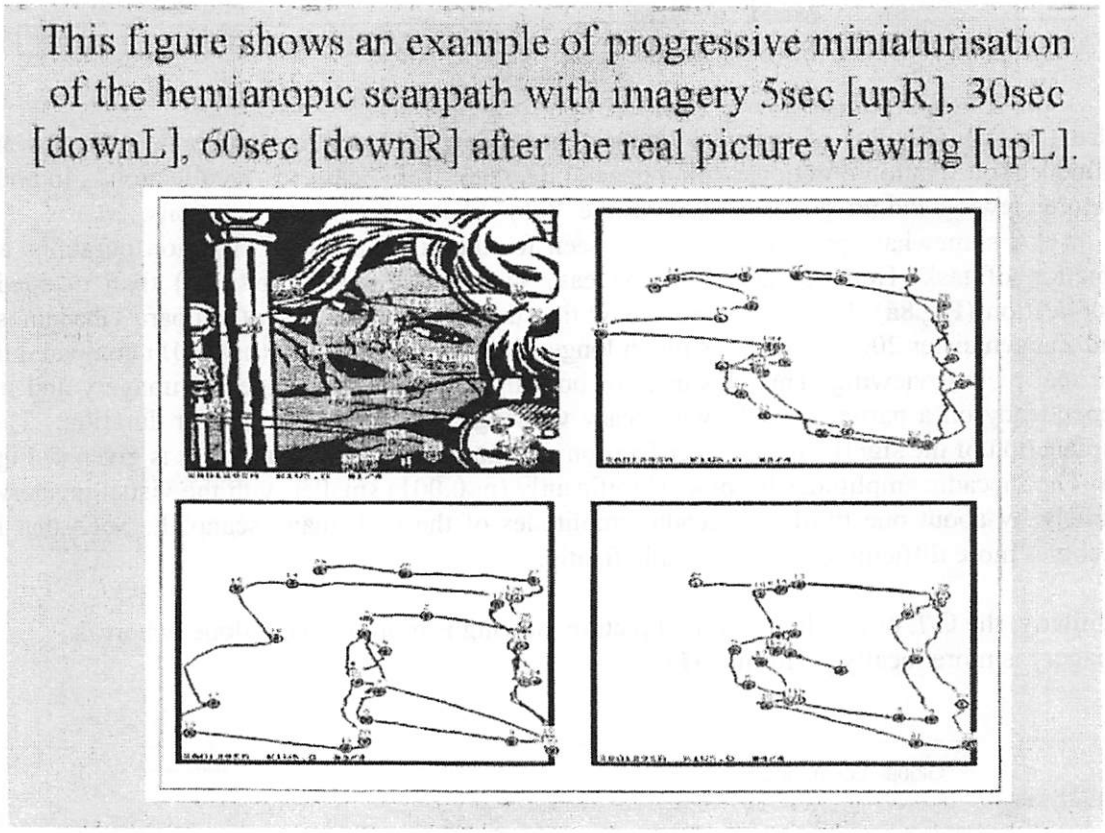


Fig. 7b

The significantly increased fixation durations in d3 (**Fig.8b**) demonstrate the "facilitated mobilisation of information (Kosslyn 1994 (11)) at work. This means, that with imagery the time to recall the picture through successive eye movements is significantly longer than the "real" scanpath. It shows also that with training (set3) this process can be facilitated.

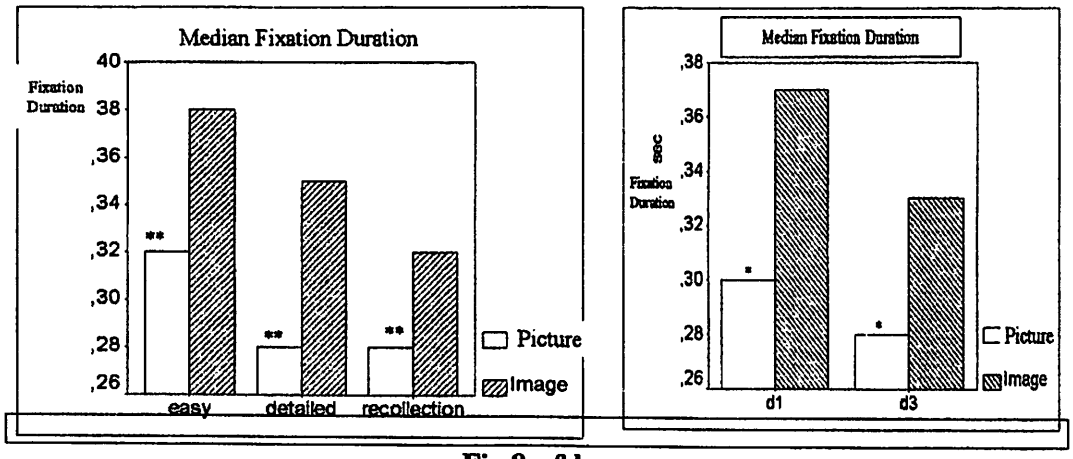


Fig.8 a&b.

Also the task (**Fig.8a**) of "easy" looking versus "detailed" or "recollection" indicates an influence on fixation durations with a general decrease from "easy to "recollection" , in both picture viewing and imagery scanpath, where "easy" shows the longest durations.

A somewhat opposite effect was seen for the comparison of fixation durations as function of task: Here, durations in the easy task where longer ($p<0.05$) than detailed/ recollection (**Fig.8a**). Fig. 8a and 8b show that the *imagery* scanpath (compare Gbadamosi and Zangemeister 2001 (7)) needs much longer duration of fixations ($p<0.001$) compared to the real picture viewing. This was true for both the duration differences of imagery and its dependency on a particular task, with "easy viewing" showing the longest duration. The explanation of the significantly longer fixation durations for imagined pictures is given in **Fig. 7a.** The saccadic amplitudes become significantly ($p<0.001$) smaller with the visual imagery, namely by about one third of saccadic amplitudes of the real image scanpath: Such that it becomes more difficult to generate stable fixations.

Similarly, the G/L ratios show that real picture viewing is more globally done, whereas Imagery is more locally performed (**Fig.9**).

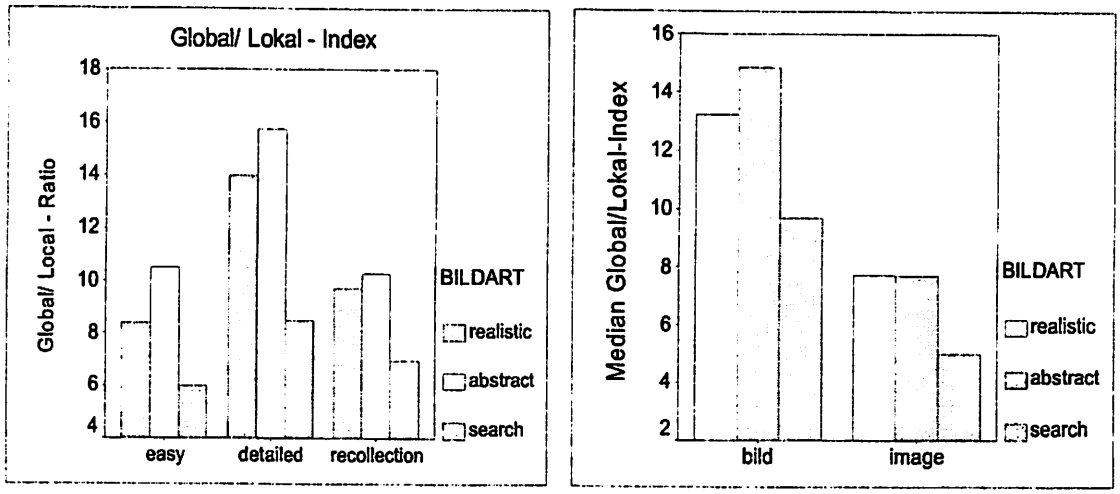


Fig.9a&b: Global/Local Ratios as a function of Task and Stimulus Type

When we use the global/local ratio, i.e. the ratio between saccades of >1.1° / < 1.1° [1.1° threshhold] as a measure to distinguish the three *picture types* during different tasks of viewing, it apperars that abstract pictures represent the most global viewing with the detailed task having the highest G/L ratio (**Fig.9a**); whereas the search pictures represent the most local viewing with easy viewing at the low end of G/L ratios. A similar G/L result can be noted when we compare real viewing with imagery (**Fig.9b**): realistic and abstract G/L ratios are generally higher than G/L ratios of the search task; and since saccadic amplitudes become smaller within imageries, the G/L ratio also becomes significantly smaller.All these differences are highly significant ($p<0.01$).

IV DISCUSSION

Homonymous hemianopia Simulation and related rehabilitative aspects. Hemianopia, defined as a complete visual field defect is divided into different forms according to the site of the lesion. The homonymous form more often shows macular sparing, leaving the foveal center of the visual field undamaged, which is responsible for the often occurring relatively adapted spontenuous recovery; a complete full hemianopia – such as we simulated it – often leads to enduring more pronounced visual disabilities: This is an obvious reason why a specific therapy in these cases is of great importance. Common etiologies of this disorder are the ischemic cerebrovascular event, followed in frequency by bleedings and trauma. Often hemianopia is associated with other cognitive dysfunctions like aphasia and visual hemineglect. Rossi et al. (17) found that more than 20 % of patients with stroke treated in rehabilitation centers expose hemianopic symptoms. The impact of this sensory deficit depends on size and localization of the lesion, impairing patients in visual information processing in many ways. Hemianopia usually causes problems exploring the blind hemifield causing patients to perform hypometric, slow down head movements and generate low amplitude saccades, therefore handicapping them more or less severely in orientation and safety in everyday living. Prospective studies of the natural course of vascular retrogenicular visual field defects showed that spontaneous restitution - e.g. axon-sprouting - in the blind hemifield takes place within the first 6 months after the event and that the average visual field gain may be up to 16 % in perimeter. To some degree oculomotor training strategies can compensate the sensory deficit.

Pommerenke et al. (**14**) found that a specific systematic exploration practice through perimetric saccade training improves visuo-spatial orientation in these patients. Furthermore, Zangemeister et al. (29,30) investigated the influence of cognitive motor gaze control strategies on the rehabilitation of visual field defects in hemianopic patients and found significant improvement in their visual behavior after taking part in a special cognitive training of gaze control.

A study of Butter et al. (2) tested visual imagery in hemianopic patients with occipital lesions using special imagery and perceptual control tasks. On the basis of their results, they postulated an impaired visual imagery in these patients because compared to a control group they performed worse in the imagery task when perceiving the stimulus ipsilateral to their visual field defect. They concluded that this findig supports the view that visual imagery involves topographically organized visual areas of the occipital lobe.

Scanpaths in hemianopic patients when compared to normal subjects. In a previous study we discovered (7) distinct characteristics of scanpaths in hemianopic patients when compared to normal subjects viewing the same stimuli suggesting a reduced extent of the image within the cognitive representation. Differential similarity measures demonstrated that

the gaze sequences of the picture exploration phase exhibited less (but non random) similarity with each other and a reduced field of view in the hemianopic patients than in normal subjects. This finding suggested a strong top–down component in picture exploration: In both groups, healthy subjects and hemianopic patients, a mental model of the viewed picture evolved very soon, which substantially determined the eye movements. As hemianopic patients showed analogous results to the normal subjects we concluded that well adapted patients have a preserved cognitive representation despite their perceptual defect, which follows the same top down vision strategies in the process of viewing.

In an earlier quanitative evaluation of human scanpaths (**30, 32**), we compared eye movements of hemianopic patients with normal subjects' eye movements while viewing abstract and realistic pictures before and after a special training. This training included special advice concerning eye movement strategies as in the present study but also several different settings of training; also, it lasted for 14 and 28 days respectively with control recordings after each time span: After the training the viewed pictures were presented again. The correlated scanpaths were divided into *a priori* geometric regions of interest (ROI's) and subjective *a posteriori* ROI's. The evaluation of the scan paths was done using Markov analysis and String-Editing. The results showed, despite long latencies between the time of lesion and the beginning of the training (6 –12 months), that our specific rehabilitative training was significantly successful when quantified by probabilistic and sequential measurements of the resulting scanpaths. Especially the comparison of a priori and a posteriori measurements permitted a differentiation of the training effects more closely.

This result correlated to a *facilitated mobilisation of information* of extrastriatal high level information :The visual information was transmitted trough a quick fill of the striatal visual buffer and an inside this visual buffer located *attention window* (Kosslyn et al. (10,11)) , that appeard to be opened [or enlarged] through the training. We took this as an expression of the subjective enlargement of interest trough stimulation of the associative cortex. After the training, there was a strong top-down-component during the picture-view, that alternated with a non-adapted bottom-up-component before the training. This demonstrated that it is particularly significant that the cognitive aspect of human vision in a rehabilitation of hemianopic patients has to consider special training methods.

Our here reported results in virtual hemianopic subjects are consistent with these earlier observations: Healthy subjects show quite similar oculomotor behaviour when confronted with the hemianopia simulation. This is true for simple pursuit and saccadic stimuli as well as search and scanpath tasks. Therefore, top-down mechanisms and the feature ring hypothesis explain the consistency and convergence of the oculomotor behaviour also in normal subjects confronted with the same sensory deficit as homonymous hemianopic patients as well as their similar response to our special training.

V CONCLUSION

In previous studies we tried to obtain information about the consistency and reproducibility of internal visual image representations in hemianopic patients in the context of active „high-level-vision" and about the quantitative effect of training in these patients even up to 12 months after the time of lesion. The mental image of the hemianopic patients was found to mirror their visuocortical sensory deficit. Special training evidently helped in this aspect.
In this study we analyzed the degree of adaptation that normal healthy subjects show when suddenly exposed to a virtual complete homonymous hemianopia under different training conditions.
Their saccadic, scan and search path strategies showed the same consistency and reproducibility of internal visual representations as hemianopic patients in the context of

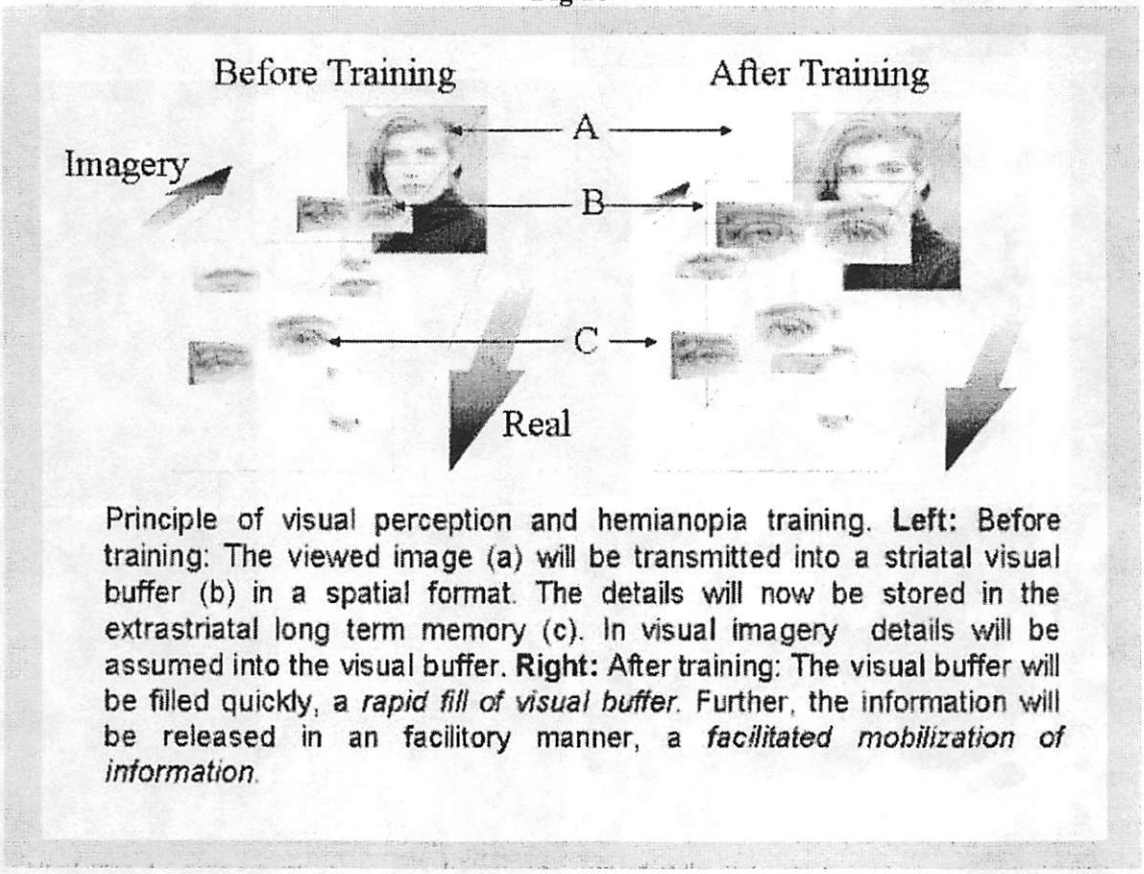
active „high-level-vision". Particularly those simulated hemianopia subjects that where well adapted with the aid of our special training did show this effect using a top down strategy when viewing targets and pictures on the side of their blind hemifield. Non-adapted subjects failed to show this effect. They stayed with a bottom up strategy of viewing targets on the side of their blind hemifield.
We conclude that – in addition to earlier described therapeutic tasks – hemianopic *patients* should apply our here in *healthy* subjects described additional increase of their field defect: Similarly as a "forced use therapy" as suggested by Taub 1993, so that they become faster aware of the increased gaze efficiency when shifting the zero meridian of their visual field towards the side of the BHF.

From these findings we can draw three conclusions that are illustrated in **Fig.10.:**

- Firstly, applying a sophisticated string analysis such as cRSE permits to show differentially the many significant differences that follow therapeutical masking of the foveal region during a [virtual, model] hemianopia.
- Secondly, the visual imagery scanpath is done over a compressed mental image that needs longer fixation duration but fewer saccades than the real image.
- Thirdly, the combination of different viewing tasks with different types of pictures permits to show how top down strategies of the scanpath can be enforced or decreased by a proper combination of task and picture; this is basically influencd by the mask versus no-mask

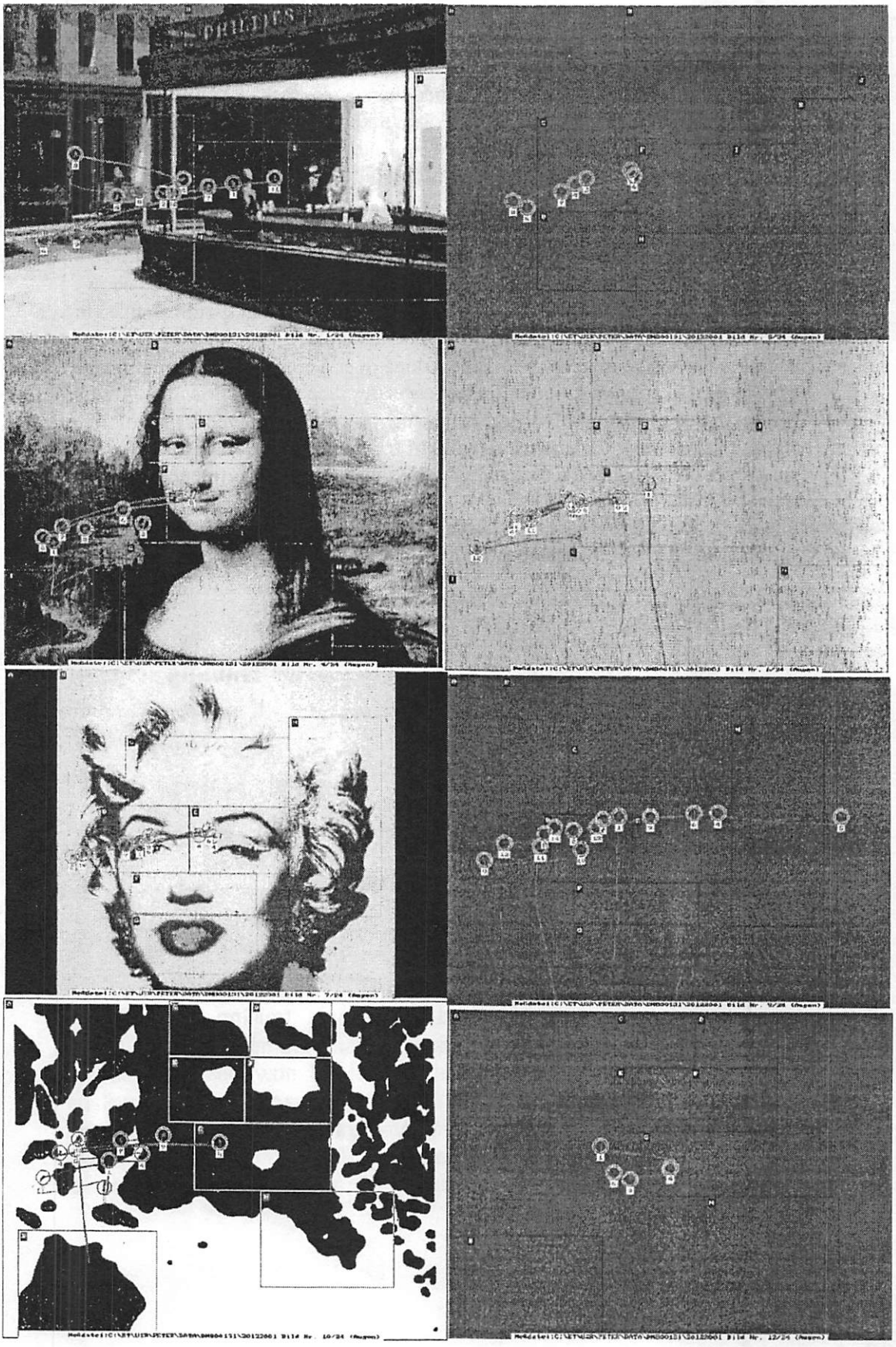
difference, with bottom up mechanisms becoming more important with the additional loss of foveal viewing strategies in the mask condition.

Fig.10

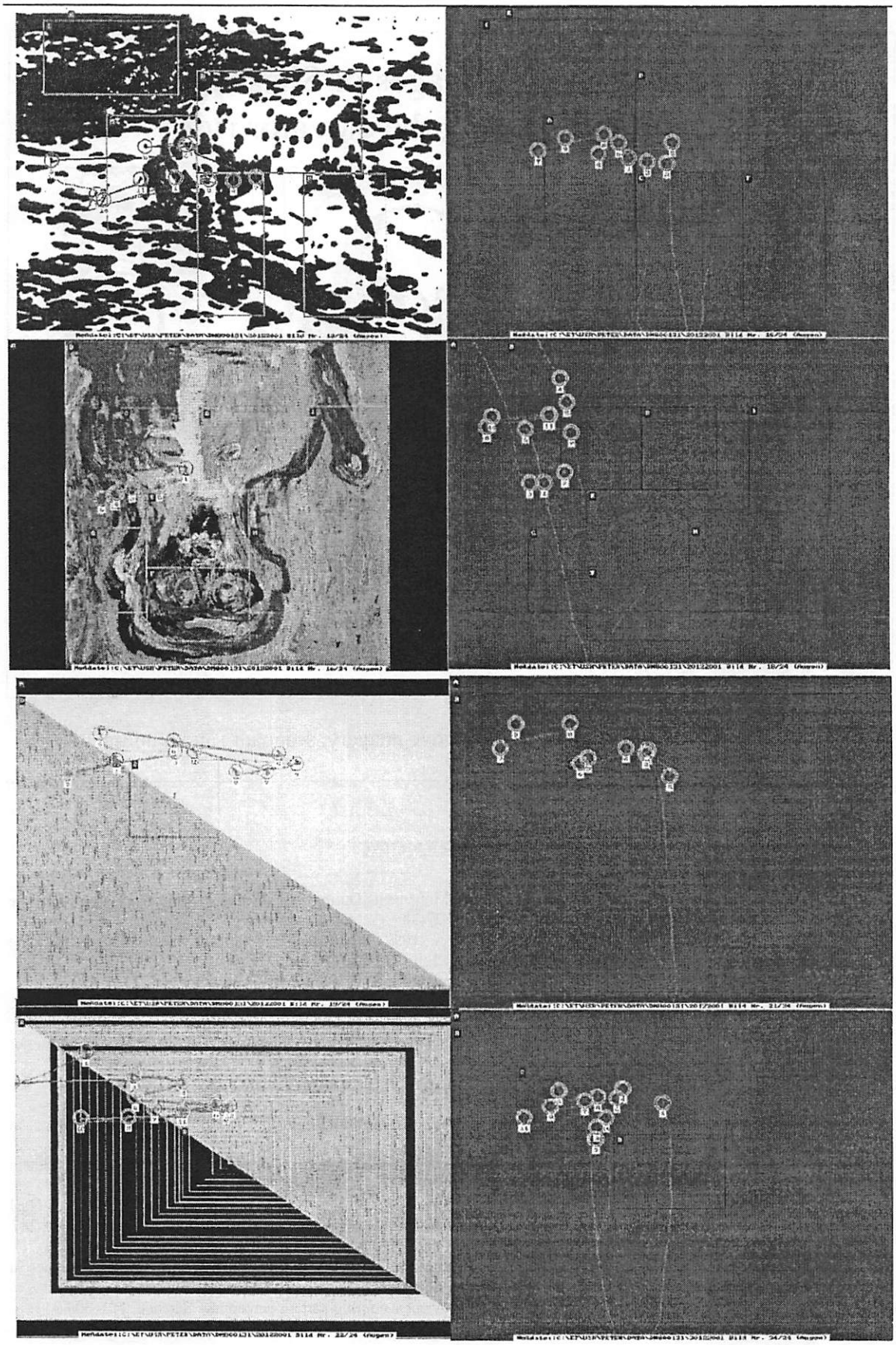


Principle of visual perception and hemianopia training. **Left:** Before training: The viewed image (a) will be transmitted into a striatal visual buffer (b) in a spatial format. The details will now be stored in the extrastriatal long term memory (c). In visual imagery details will be assumed into the visual buffer. **Right:** After training: The visual buffer will be filled quickly, a *rapid fill of visual buffer*. Further, the information will be released in an facilitory manner, a *facilitated mobilization of information*.

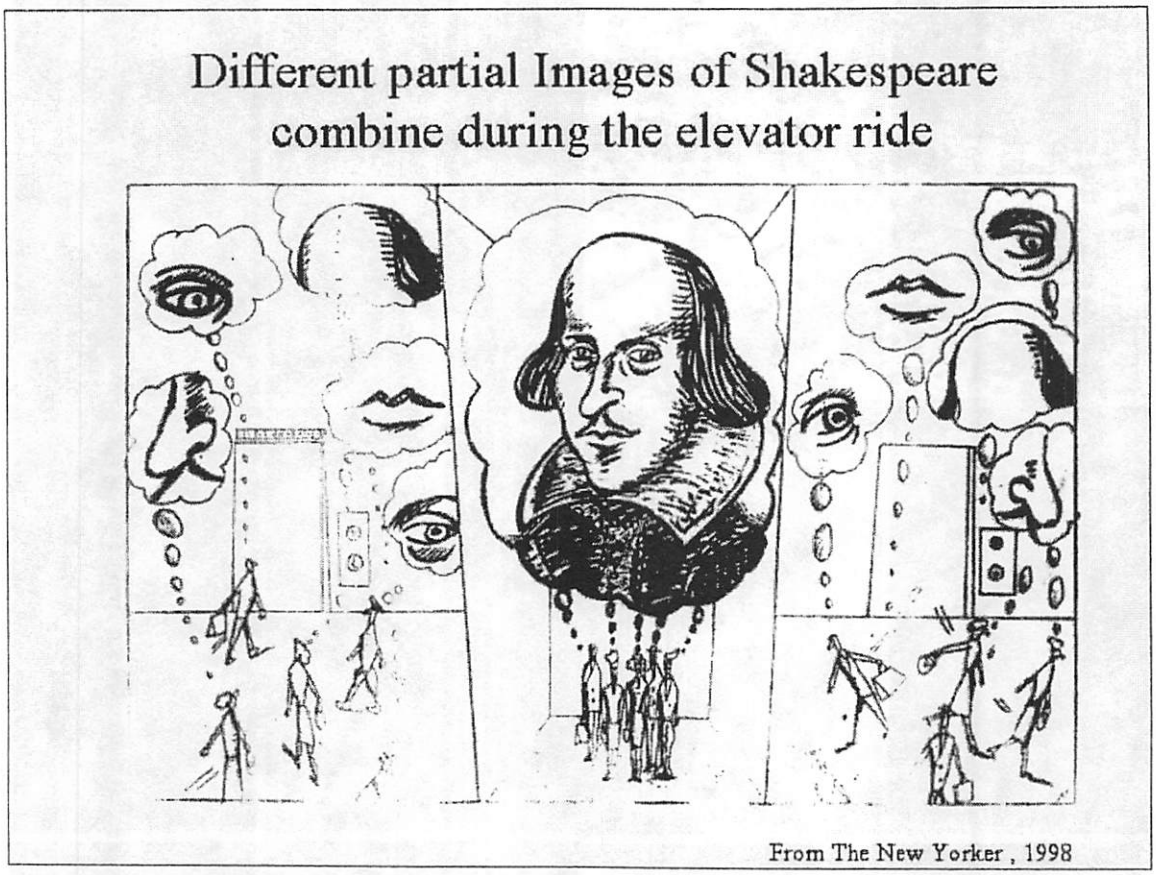
Appendix A: Scanpaths of Pictures 1-4 left versus Image right. 50% Mask to right.



Appendix B: Scanpaths of Pictures 5-8 left versus Image right. 50% Mask to right.



FIN



One of Larry Stark´s literary examples of collective imagery, scanpath , and binding.

V REFERENCES

1. Brandt, S. & Stark, L. W. (1997). Spontaneous Eye Movements During Visual Imagery Reflect the Content of the Visual Scene, Journal of Cognitive Neuroscience, 9: 27-38
2. Butter, C.M., Kosslyn, S., Mijovic-Prelec, D., Riffle, A. (1997). Field-specific deficits in visual imagery following hemianopia due to unilateral occipital infarcts, Brain: 120, 217-228
3. Chatterjee, A., Southwood, M.H. (1995). Cortical blindness and visual imagery. Neurology 45: 2189-95.
4. Chernyak, D.A., Stark, L.W. (2001). Top-Down guided eye movements. IEEE Trans Sys Man Cybern. 31: 514-522
5. Ellis, S.R. & Stark, L. (1978). Eye movements during viewing of Necker cubes. Perception 7: 575-581.
6. Gassel, M.M., Wiliams, D. (1963). Visual function in patients with homonymous hemianopia. Part 2. Oculomotor mechanisms. Brain, 86: 1-36
7. Gbadamosi, J., W.H. Zangemeister (2001). Visual Imagery in Hemianopic Patients . J Cogn. Neurosci. 13(7): 45-56
8. Goldenberg, G., Poderka, I., Steiner, M., Suess, E., Deecke, L. (1989). Regional cerebral blood flow patterns in visual imagery. Neuropsychologia 27: 641-664
9. Hacisalihzade, S. S., Stark, L. W. & Allen, J. S. (1992). Visual perception and sequences of eye movement fixations: A stochastic modeling approach. IEEE Trans. Systems, Man, Cyb., 22: 474-481.
10. Kosslyn, S.M., Alpert, L.M., Thopson, W.L., Maljkovic, V., Weise, S.B., Chabris, C.F. et al. (1993). Visual mental imagery activates topographically organized visual cortex: PET investigations. J Cogn Neurosci, 5: 263-287
11. Kosslyn, S. M. (1994). Image and Brain. MIT Press, Cambridge, Mass.
12. Meienberg, O., Zangemeister, W. H., Hoyt, W. F. & Stark, L. W. (1981). Saccadic eye movement strategies in patients with homonymous hemianopia. Ann. Neurol., 9, 537 - 544.
13. Noton, D. & Stark, L. W. (1971). Scanpaths in eye movements during pattern perception. Science, 171, 308-311.

14. Pommerenke, K., Markowitsch, H.J. (1989): Rehabilitation training of homonymous visual field defects in patients with postgeniculate damage of the visual system. Restor. Neurol. Neurosci. 1: 47-63
15. Poppelreuter, W. (1917): Die psychischen Schädigungen durch Kopfschuß im Kriege Vol. 1: Die Störungen der niederen und höheren Sehleistungen durch Verletzungen des Okzipitalhirns. Voss. Leipzig, 1917
16. Roland, P.E., Gulyas, B. (1994). Roland, P. E. & Gulyas, B. (1994). Visual imagery and visual representation. Trends Neuroscience, 17(7), 281-287. And: Visual representations of scenes and objects: retinotopical or non-retinotopical ? Trends Neurosci, 17(7): 294-297.
17. Rossi, P.W., Kheyfets, S, Reding, M.J. (1990). Fresnel prisms improve visual perception in stroke patients with homonymous hemianopia or unilateral visual neglect. Neurology 40: 1597-1599.
18. Schoepf, D. & Zangemeister, W. H. (1993). Correlation of coordinated gaze strategies to the status of adaptation in patients with hemianopic visual field defects. Ann. NY Acad. Sci., 682, 404-409.
19. Stark, L. W. and Choi, Y. S. (1996). Experimental metaphysics: The scanpath as an epistemological mechanism. In: Zangemeister, W. H. et al.: Visual Attention and Cognition, pp. 3-69. Elsevier Publishers, Amsterdam.
20. Stark, L. W. and Ellis, S. (1981). Scanpaths revisited: Cognitive models in active looking. In Fisher, B., Monty, C. & Sanders, M. (Eds.), Eye Movements, Cognition and Visual Perception, pp. 193-226. Erlbaum Press, New Jersey.
21. Stark, L.W. et al. (2001). Representation of human vision in the brain: How does human perception recognize images? J. Electronic Imaging 10: 123-151
22. Teuber H.L., Battersby W. und Bender M.B. (1960). Visual field defects after penetrating missile wounds of the brain. MIT Press, Cambridge.
23. Zangemeister, W.H., Meienberg, O., Stark, L., Hoyt, W.F. (1982): Eye-Head Coordination in Homonymous Hemianopia. J.Neurol. 225: 243-254.
24. Zangemeister, W.H., Dannheim, F., Kunze, K. (1986). Adaption of gaze to eccentric fixation in homonymous hemianopia. In E.L. Keller and D.S. Zee (Eds.): Adaptive processes in visual and oculomotor systems. Pergamon Press.
25. Zangemeister W.H., Sherman K., Stark L. (1995). Evidence for global Scanpath strategy in viewing abstract compared with reaistic images. Neuropsychologia 33: 1009 - 1025 (1995)
26. Zangemeister W.H., Stark L. (1982). Gaze latency: variable interactions of eye and head movements in gaze. Exp Neurol, 75: 389-406
27. Zangemeister, W. H., Oechsner, U. & Freksa, C. (1995). Short-term adaptation of eye movements in patients with visual hemifield defects indicates high level control of human scanpath. Optom. Vis. Sci., 72: 467-477.
28. Zangemeister, W. H. and Oechsner, U. (1996). Evidence for scanpaths in hemianopic patients shown through string editing methods. In Zangemester et al., Visual Attention and Cognition, pp. 197-221. Elsevier Publishers, Amsterdam.
29. Zangemeister, W.H., Poppensieker, K., Hoekendorf, H. (1999). Cognitive Neurovisual Rehabilitation using Coordinated Gaze Strategies, Shaker Publ., Aachen
30. Zangemeister, W.H. (1997). Virtual reality hemianopic scotomas induce eccentric fixation: Scanpath strategies to optimize high level vision in healthy subjects. IEEE 97CH36136EMBS: p.560 – 565
31. Zangemeister, W.H., U.Oechsner (1999). Adaptation to visual field defects with virtual reality scotoma in healthy subjects. Current Oculomotor Research, ed.by Becker at al., Plenum Press, New York, 89 – 95
32. Zangemeister W.H., P.Utz (2002). An increase in a virtual hemianopic field enhances the efficiency of secondary adaptive gaze strategies. Current Psychology of Cognition 21: 281- 303
33. Zeki, S. (1999) A vision of the Brain. Blackwell Scientific Publications. Oxford . p. 22-32
34. Zihl, J. (1981): Recovery of visual functions in patients with cerebral blindness: Effect of specific practice with saccadic localization. Experimental Brain Research, 44: 159-169 33.
35. Taub E. et al. (1993). Technique to improve chronic motor deficit after stroke. Arch Phys Med rehabil 74: 347-354

Scanpaths of motion sequences: where people look when watching movies

Eli Peli, Robert B. Goldstein, and Russell L. Woods,
The Schepens Eye Research Institute & Harvard Medical School, Boston, MA, USA

Abstract— **Magnification around the most important point of the scene (center of interest - COI) might be an effective aid for people with vision impairments that cause resolution loss. This requires that a COI exist for most video frames. Operationally, we defined the COI by recording the eye movements of normally-sighted subjects as they watched movies. Here we report the frequency that people looked at the same place during the movies and the spatial distribution of their COIs, and investigate age and gender differences.**

Index Terms— **Eye Movements, Magnification, Scanpath, Video, Visual Aids, Low Vision, Television.**

I. Introduction

People who suffer loss of visual resolution due to eye diseases could benefit from modified information displays. The most common modification used today is magnification. Magnification inherently restricts the field of view and thus may impede the acquisition of information attained in normal vision by the use of scanning eye movements. This problem may be addressed by dynamic control of the displayed information. Dynamic control of large text presentation is helpful for people with low vision [1-4]. We propose a similar approach to improve access to movies and television.

Magnifying moving images using electronic zoom [5] would enable users to select and vary the desired level of magnification from time to time. However, only part of the magnified scene can be presented on the screen. Consequently, large parts of the scene become invisible. Manual zoom-and-roam devices are available in commercial television systems (e.g. DVD players). However, the rapid changing of scenes in most movies may not allow for effective manual control of the magnified section of the image. We proposed pre-selecting the point in the scene on which to center the magnified view (the center of interest - COI) and providing that position with each frame [6, 7]. This selection should maintain the most relevant details in view.

Together with DigiVision (San Diego, CA), we have developed a computer controlled zoom-and-roam device for playback of movies on a television. The computer plays a DVD and simultaneously reads the COI. These coordinates are sent to the zoom and roam device so that the magnified image is centered on the COI coordinates. We proposed using eye movement recordings from normally-sighted observers watching the movie to determine the desired COI. Although other methods of determining the COI can be envisioned, eye movement recording is automatic and objective.

Choosing the COI using eye movements is akin to finding the scanpath for a movie sequence. Much work has been done regarding the scanpath of still images [8-11], but little is known about viewing moving images. With the exception of a few studies [12-14] most development that depends on knowing where the gaze is directed (e.g. compression schemes [15] and transmission of images for limited screen space [16]) assume that most people look at the same place all the time while watching movies. To our knowledge this assumption has not been verified experimentally. Here we quantify the proportion of the time multiple people look at the same place while watching a movie, and begin to examine the effects of age and gender on this behavior.

Film editors have used assumed knowledge of viewer's eye movements - and even blinks - to assemble movies [17]. Stelmach et al. [12] recorded 24 observers viewing 15 forty-five second clips to determine if viewing behavior can be incorporated into video coding schemes. They found that there was a substantial degree of agreement among viewers in terms of where they looked. In a follow-up experiment related to gaze-contingent processing techniques [13], recorded eye movements of subjects were used to create a "predicted gaze position". Tosi et al. [14] recorded the scanpaths of 10 subjects watching a variety of clips totaling about 1 hour and reported that, qualitatively, individual differences in scanpaths were relatively small. Theoretical saliency models [18, 19] predict where people will look and have made no assumptions regarding individual differences in predictions of regions of interest.

Here we address three specific questions relevant to our proposed low-vision aid for viewing television. (1) To what extent do people look at the same place when watching a movie? (2) Does that vary with age and gender? (3) Does the position of the COI differ from the center of the screen?

II. Methods

Six movie clips were selected to span a broad range of scene activity, from stationary newscasters to athletes in motion and

Supported in part by NIH Grants EY05975 and EY12890.

Corresponding author: Eli Peli is at The Schepens Eye Research Institute, 20 Staniford Street, Boston, MA, 02114 USA. (E-mail: eli@vision.eri.harvard.edu).

to appeal to both younger and older audiences (Table 1). The movie clips from DVDs were presented in a 16×9 movie format on a 26.5-inch diagonal NTSC (4×3) monitor as interlaced video at 30 frames per sec. (60 fields per sec.).

Table 1. Category, names, and length of movie clips used.

Category	Title	Time (min)
Sports	Any Given Sunday (1999)	4:12
Comedy	Big (1988)	6:29
Documentary	Blue Planet (2001)	8:14
News	Network (1976)	4:02
Game Show	Quiz Show (1994)	6:40
Drama	Shakespeare in Love (1998)	7:06
	Total per subject	37:29:00

26 normally-sighted subjects were seated 46 inches from the screen which spanned a 26.3°×14.8° visual angle. Subjects viewed movie clips while eye movements were recorded with an ISCAN model RK726PCI eye tracking system. The ISCAN had a nominal accuracy of 0.3° over a ±20° range and a sampling rate of 60Hz. Thus we could acquire two eye samples per video frame. The ISCAN compensated for head movements, permitting gaze monitoring without head restraint, and thus allowing a comfortable viewing situation. The ISCAN was calibrated using a 5-point calibration scheme. To optimize tracking [20, 21], we performed a pre-clip calibration (external to the ISCAN) which was repeated before each movie clip was viewed. Recording of the next clip did not proceed until all 5 calibration points had satisfactory (45 samples) data yield. A post-clip calibration was also recorded and the analysis program averaged the pre and post-clip results for the calibration equations.

During the recording phase, immediate feedback was available regarding the amount of valid data available. If less than 80% was valid, the subject's data were not considered for inclusion, and recording stopped. We did not repeat movie clips, as we wanted to record the subject's eye movements during their first viewing of the clip. This happened only with one subject. Of the remaining 25 subjects, the 5 subjects with the best eye movement data yields in each of 4 groups were selected. The 20 subjects were grouped by age and gender: Younger Female 18-29y, Male 16-36y; and Older Female 51-62y, Male 42-66y.

III. ANALYSIS

A. Preprocessing of individual records

The individual subjects' recordings were processed to apply the calibration to the data and remove recording artifacts caused by blinks and other failures. Recording could fail if the head moved too fast or when specular reflections, such as tear film menisci, were erroneously detected by the ISCAN as a cornea reflection. Blinks and loss of tracking were filtered from the file by removal of records containing zero value data or frames where the pupil diameter was out of a set range.

The DirectX 8.1 (Microsoft, Redmond, WA) DVD interface only interrupts the processor every 0.4 to 1.0 second with timing information. Frame numbers between such interrupts were calculated from the elapsed time, assuming a 30 frame per second rate. This procedure resulted, on occasion, in non-monotonic or duplicate frame numbers. Non-monotonic frames were discarded. Because the video and ISCAN data were recorded asynchronously, each assigned frame could be associated with one, two or three eye records. We have designated these multiple records per frame as "subframes" (note – these subframes are not video fields).

Table 2. Yield of acceptable eye samples data from each movie clip did not vary significantly between clips. Yield was greater for male subjects ($F_{1,16}$=7.8, p=0.01) and slightly greater for older subjects ($F_{1,16}$=3.2, p=0.09).

	OM	YM	OF	YF
Sunday	97.2%	95.2%	93.4%	93.0%
Big	96.9%	94.6%	94.8%	92.6%
Blue	97.7%	94.2%	94.6%	93.0%
Network	95.7%	95.6%	93.3%	92.6%
Quiz	96.3%	94.4%	93.6%	91.4%
Shakes	96.0%	95.1%	93.9%	91.8%

B. Merging of eye recordings of multiple subjects' records to find extent of overlap

The 120 subject data files (20 subjects × 6 clips) were processed to count how many of those subjects had valid data for each subframe (Fig. 1).

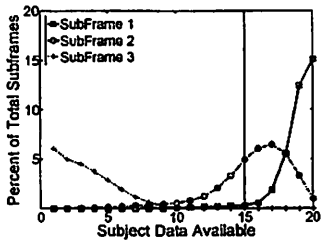


Fig. 1. The percentage of the total number of subframes for which the number of subjects had valid data. The percent of subframes where 15 or more subjects contributed data (vertical line) was 63%, which increased to 85% when data from subframe 3 were discarded.

For each subframe the calibrated (x, y) coordinates of individual subjects gaze points are distributed across the screen. Various methods have been applied to compute the level of coincidence between the gaze points of multiple subjects [11, 12, 22-25]. We chose to calculate the area of the best-fit bivariate contour ellipse (BVCEA) to quantify the degree of spatial coincidence of the eye positions of all the subjects with valid data points. This measure has been used in the past to quantify fixation eye movement stability [2, 26]. The k parameter of the BVCEA determines the degree of enclosure of the ellipse. We set k=1, for which 63% of the points would have been enclosed by the ellipse.

The cumulative distributions of the BCVEA found for each movie clip (Fig. 2) were fit with a logistic function,

$$y = c + (1-c)/(1+\exp(-(x-a)/b)) \quad (1)$$

where a is the mid-point of the function, b is related to the rate of rise, and c is the lower asymptote. Similarly, to quantify effects of age and gender, the BCVEA was calculated for every subframe for which there were eye position data for at least 4 subjects from each group of 5 subjects. In subsequent

data analyses we used the ½ point and *b*. Data were evaluated using analysis of variance, with movie clip treated as a within-subject factor.

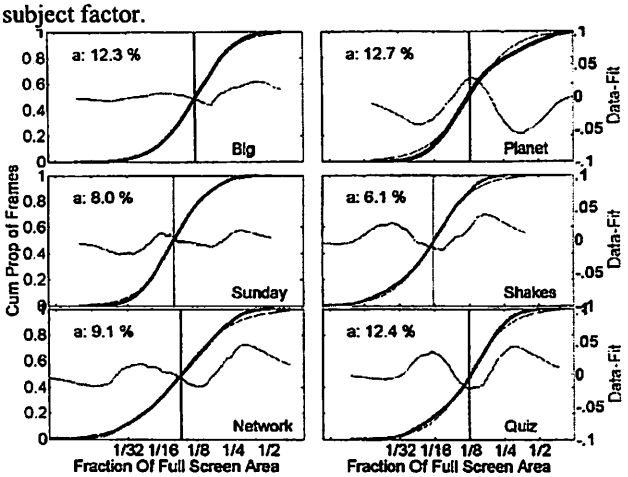


Fig. 2. For each movie-clip subframe, eye position coordinates were used to calculate the BVCEA as a fraction of the full screen area. Only those subframes where 15 or more subjects had data were used. The cumulative curves show the proportion of the total subframes for which the BVCEA was less than a given fraction of movie screen area. Logarithmic transforms of the distributions were fitted to a logistic function (with *c*=0) that were then used to calculate the screen fraction for which 1/2 of the samples had a smaller BVCEA (vertical line and the value indicated by the inset). The residuals of the fits are shown.

IV. RESULTS

As shown in Fig. 2, for all six movie clips, more than 1/2 of the time most of the subjects (15 to 20) looked within an area that was less than 13% of the movie scene. This represents an area equivalent to a circle with a diameter of about 8 deg.

To examine the effects of age and gender on COI we performed analyses of variance on the 1/2-point and *b* of the fits to the data shown in Fig. 3. As seen in Fig. 4, male and older subjects were more likely to look in the same direction (smaller 1/2-point) than female ($F_{1,20}$=6.3, p=0.02) and younger ($F_{1,20}$=22, p<0.001) subjects, respectively. Older subjects were slightly more variable (slower rise – larger *b*) than younger subjects ($F_{1,20}$=3.4, p=0.08). Between the movie clips, there were significant differences in *b* ($F_{5,18}$=4.4, p=0.009) but not of the 1/2 point, indicating that the COI was more variable for some movies. Subjects were more variable in *Network* than *Planet* (p=0.04), Big (p=0.03) and *Sunday* (p=0.004). This shows that, as might be expected, movies with high level of motion more tightly control the observer COI than movies with relatively static scenes.

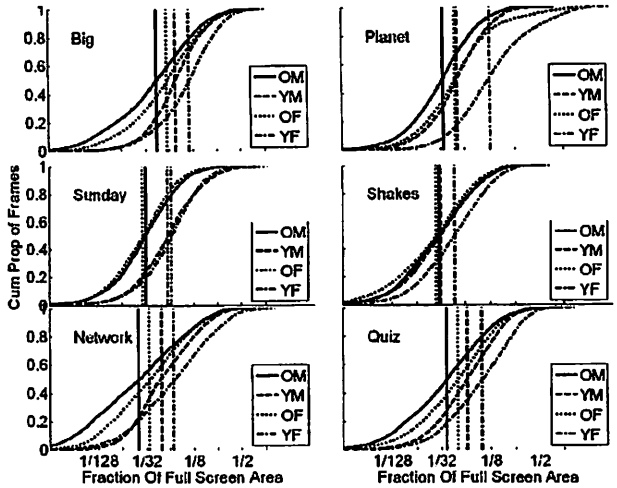


Fig. 3. The cumulative area distributions and positions of the 1/2-point of each clip for each age-gender group (where at least 4 out 5 subjects had useable data). Older and male groups had tighter positions of gaze points.

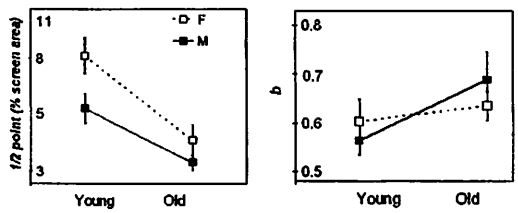


Fig. 4. There were significant effects of gender and age on the likelihood that the subjects in a group looked in the same direction (the 1/2-point); and there was a small, non-significant age effect on the variability of direction of gaze (*b*). Note that the scale labels for the 1/2-point are non-linear, since the fit was done in the logarithmic transform of the area. Error bars indicate SEM.

For those subframes for which there was eye position data for more than 15 of the 20 subjects, the position of the COI was determined as the mean x and y coordinates of the group. The COI distributions (32×24 bins for the letterbox area) are shown in Fig. 5. In general, the peak of the COI distributions were approximately in the center of the movie scene, though they varied by as much as 1/4 of the width or height from center, and the distributions varied between movie clips.

V. DISCUSSION

Measuring and providing the coordinates of the COI along with each frame may allow magnification to be used to its full potential as a low vision aid for watching movies (and other television programs). The eye movement method presented here is a natural and efficient way of determining these COIs. We envision that, just as programs are now being provided in "closed captioned" and described video formats, movies can be provided with these COIs encoded.

We have demonstrated that it is possible to determine the COI in a movie scene by recording the eye movements of normally-sighted observers while they watch a movie. Over

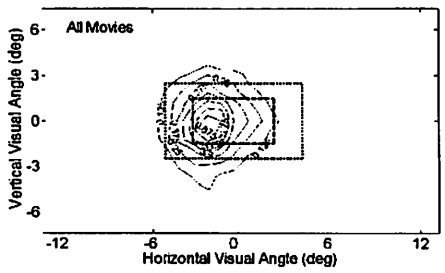


Fig. 5. COI distribution for all movie clips for all subjects (for subframes with eye position data for at least 15 subjects). Although the distribution peaks near the center of the screen, the spread indicates that a large proportion of the time, people did not look at the center of the movie scene. The dotted and dashed regions represent 1/9th (11%) (3X magnification) and 1/25th (4%) (5X magnification) of the screen area, respectively. 73% of COIs lay outside the 4% screen area.

1/2 of the time, the gaze of more than 15 subjects was contained within an area that was less than about 13%) of the movie scene (Fig. 2) or about 5% of the screen for 4 subjects (Fig.4). This is crucial for our application. We rarely expect to need to magnify the image by greater than a factor of 4 (showing 1/16th (6%) of the frame). Higher magnification might cause too much loss of context. The distribution of COIs (Fig. 5) illustrates that magnification centered on the COI would provide more information than magnification simply centered on the center of the movie scene. Also, we found that there are some significant differences in the observation behaviors between gender and age groups. The current analysis only found that the older and male observers' COIs were more tightly grouped than the younger and female observers (Fig. 4). We still need to determine if the COI locations varied with gender and age. Also, conditions or scenes that did not result in a tight COI (i.e. large BCVEA) might be just as interesting as the condition of tight COI.

In addition to our interest in the application of this technique to our movie (or television) magnification device, we see this work as a beginning of an interesting examination of the nature and characteristics of the motion scanpath of dynamic environments — the movie environment being one that is simpler to study — perhaps followed by the dynamic real world of a mobile observer.

ACKNOWLEDGMENT

Assistance was provided by Gang Luo and Shabtai Lerner.

REFERENCES

[1] G. Legge, G. Rubin, D. Pelli, and M. Schleske, "Psychophysics of reading II. Low vision," *Vision Research*, vol. 25, pp. 253-66, 1985.

[2] G. Rubin and K. Turano, "Reading without saccadic eye movements," *Vision Research*, vol. 32, pp. 895-902, 1992.

[3] E. Fine and E. Peli, "Computer display of dynamic text," *Vision 96: International Conference on Low Vision, Madrid, Spain*, pp. 259-67, 1996.

[4] E. Fine and E. Peli, "Benefits of rapid serial visual presentation (RSVP) over scrolled text vary with letter size," *Optometry and Vision Science*, vol. 75, pp. 191-6, 1998.

[5] S. Kazuo and J. Shimizu, "Image scaling at rational ratios for high-resolution LCD monitors," *Society for Information Display Digest of Technical Papers*, vol. 31, pp. 50-53, 2000.

[6] E. Peli, "Vision multiplexing - an engineering approach to vision rehabilitation device development," *Optometry and Vision Science*, vol. 78, pp. 304-315, 2001.

[7] R. B. Goldstein, H. Apfelbaum, G. Luo, and E. Peli, "Dynamic magnification of video for people with visual impairment," *Society for Information Display, Digest of Technical Papers*, pp. 1152-1155, 2003.

[8] L. Stark, "Abnormal patterns of normal eye movements in schizophrenia," *Schizophrenia Bulletin*, vol. 9, pp. 55-72, 1983.

[9] L. Stark, "New quantitative evidence for the scanpath theory: Top-down vision in humans and robotics," presented at First Meeting of the International Society of Theoretical Neurobiology, 1993.

[10]L. W. Stark, K. Ezumi, T. Nguyen, R. Paul, G. Tharp, and H. I. Yamashita, "Visual search in virtual environments," *Human Vision, Visual Processing, and Digital Display III/Human Perception, Performance, and Presence in Virtual Environments*, pp. 577-589, 1992.

[11]C. M. Privitera and L. W. Stark, "Algorithms for defining visual regions-of-interest: comparison with eye fixations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 970-982, 2000.

[12]L. Stelmach, W. J. Tam, and P. Hearty, "Static and dynamic spatial resolution in image coding: An investigation of eye movements," *Human Vision, Visual Processing and Digital Display II*, vol. 1453, pp. 147-152, 1991.

[13]L. Stelmach and W. J. Tam, "Processing image sequences based on eye movements," *Human Vision, Visual Processing, and Digital Display V*, vol. 2179, pp. 90-98, 1994.

[14]V. Tosi, L. Mecacci, and E. Pasquali, "Scanning eye movements made when viewing film: preliminary observations," *International Journal of Neuroscience*, vol. 92, pp. 47-52, 1997.

[15]W. S. Geisler and H. L. Webb, "A foveated imaging system to reduce transmission bandwidth of video images from remote camera systems," NASA, Austin, TX 19990025482; AD-A358811; AFRL-SR-BL-TR-98-0858, 1998.

[16]U. Rauschenbach and H. Schumann, "Demand-driven image transmission with levels of detail and regions of interest," *Computers and Graphics*, vol. 23, pp. 857-866, 1999.

[17]E. Dmytryk, *On Film Editing*. Boston, London: Focal Press, 1984.

[18]X. Marichal, T. Delmot, C.DeVleeschouwer, V. Warscotte, and B. Macq, "Automatic detection of interest area of an image or a sequence of images," *International Conference on Image Processing (ICIP '96)*, vol. 3, pp. 371-374, 1996.

[19]L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, pp. 1254-1259, 1998.

[20]A. J. Hornof and T. Halverson, "Cleaning up systematic error in eye-tracking data by using required fixation locations," *Behavior Research Methods, Instruments & Computers*, vol. 34, pp. 592-604, 2002.

[21]D. M. Stampe, "Heuristic filtering and reliable calibration methods for video-based pupil-tracking systems," *Behavior Research Methods, Instruments & Computers*, vol. 25, pp. 137-142, 1993.

[22]D. Salvucci and J. Goldberg, "Identifying fixations and saccades in eye-tracking protocols," in *Eye Tracking Research & Applications Symposium*. Palm Beach Gardens, FL: Association for Computing Machinery, 2000, pp. 71-78.

[23]J. Goldberg and J. Schryver, "Eye-gaze-contingent control of the computer interface: Methodology and example for zoom detection," *Behavior Research Methods, Instruments & Computers*, vol. 27, pp. 338-350, 1995.

[24]L. Stark and Y. Choi, "Experimental metaphysics: the scanpath as an epistemological mechanism," in *Visual Attention and Cognition*, W. H. Zangemeister, H. S. Stiehl, and C. Freska, Eds. Amsterdam; New York: Elsevier, 1996, pp. 3-69.

[25]W. Osberger and A. M. Rohaly, "Automatic detection of regions of interest in complex video sequences," *Proceedings of SPIE. Human Vision and Electronic Imaging VI*, vol. 4299, pp. 361-372, 2001.

[26]G. T. Timberlake, M. K. Sharma, S. A. Grose, D. V. Gobert, J. M. Gauch, and J. H. Maino, "Retinal location of the preferred retinal locus (PRL) relative to the fovea in scanning laser ophthalmoscope (SLO) images," *Optometry and Vision Science*, vol. 82, pp. 177-185, 2005.

Eye Movements and Smart Technology

Christian Freksa and Sven Bertel, Cognitive Systems Research Group, Universität Bremen, Germany

Abstract—**This paper describes the use and stepwise development of smart eye movement analysis technology from a cognitive science / artificial intelligence perspective. The scanpath concept is identified as the instantiation of a general sequencing principle that permeates the organization of spatial scene knowledge on most levels of mental processing. As such, it helps create methodologies to open up windows onto higher-level cognitive processes. The paper argues that these methodologies form a robust basis for smart applications that employ eye movements to assess and to assist in diagrammatic problem solving.**

Index Terms—**mental imagery, neighborhood, diagrammatic problem solving, scanpaths, spatial structure, visual attention**

I. Introduction

EYE movement technology has come a long way. In the beginning, ways to register eye movements were needed: copper coils were attached to the eye-balls to record induced electric currents in a magnetic field when the eyes were moved and cornea reflections of infrared light sources were measured by photo cells to compute eye positions. Determining eye position through image analysis from a camera image was first proposed and realized in the 1970s, but computer memory was expensive, processing times were too slow, and cameras were too large to use camera-based methods for real-time eye tracking in those days.

Due to the technological issues involved in tracking eye positions, much of the early eye movement studies were more determined by technological boundary conditions than by the cognitive issues of interest to the researchers. For example, the head movements of the participants in eye movement experiments had to be constrained through chin and head rests as well as through bite bars that restricted not only the head movements but also the comfort and relaxation of the participants. The quality of equipment calibration competed against the quality of the participants' response as calibration required full attention and tired out the participants.

This situation has changed considerably with modern camera-based mobile eye recording equipment that functions with mostly unconstrained head movements and can be carried around comfortably. We now can go beyond simple eye movement recording tasks and employ advanced technologies for new tasks that make even more use of the sophistication of the oculomotor system.

Acknowledgments: This work is supported by the German Science Foundation (DFG) in the framework of the project R1-[ImageSpace] of the Transregional Collaborative Research Center SFB/TR 8 Spatial Cognition: Reasoning, Action, Interaction.

Corresponding author: Christian Freksa is in the Department of Informatics, Universität Bremen; Bibliothekstr. 1, 28359 Bremen, Germany. (freksa@informatik.uni-bremen.de).

II. Spatial Structure and Visual Attention

Grasping the physical environment with our senses and giving it meaning provides a great challenge to cognitive systems – both natural and artificial. The entities in the environment and their relations span a huge space too complex to exhaustively search by computational approaches. This complexity is due to high dimensionality of the conceptual space we obtain when we relate each entity to every other entity; the problem is compounded when we take into account the wide range of scales and granularities at which we are interested in entities in the environment ranging from sub-atomic structures to astronomic scales.

In other words, there is no way that cognitive agents – specifically humans and autonomous robots – can understand the world around them by exhaustively analyzing the inter-relationships of the entities involved; they must be extremely selective to come to grasp with the world. Cognitive agents and their perceptual and conceptual mechanisms can establish certain relationships between perceptual entities easily and quickly while others require a lot of effort and cannot be used in real-time recognition processes.

For this reason, it is natural that the concepts and meaningful structures cognitive systems establish and employ closely follow their perceptual and conceptual abilities; such is a central idea of studying embodied cognitive systems. Conversely, perceptual and conceptual abilities have developed through evolutionary processes that provide support for establishing connections between related entities in the world.

A. Spatial structure

Perception makes use of and is biased by spatial structures in the physical environment extensively: if each entity in the world was equally related to each other entity, cognitive systems had to deal with a space whose dimensionality equals the number of entities in the world - 1. Physical space restricts the number of dimensions to three. This essentially means that not all entities can have the same distance to one another; some entities are closer to one another than others. Thus, by restricting the problem space to few dimensions it will be much less complex – at the expense of accessibility of more distant concepts.

As embodied cognitive agents are at the same time physical agents, they are also subject to the same restrictions of

physical space with regards to their physical movements: some locations in space will be closer than others. Additional restrictions apply: due to gravity and/or other physical constraints, agents cannot move equally easily in all three dimensions, thus, not all geometrically equidistant locations are equally reachable. For perceptual space other restrictions apply: vision, for example, can overcome egocentric distance to a certain degree, but it is confined to the two true spatial dimensions of the visual field; information about the third dimension (i.e. distance) can only be supplied indirectly. Thus, vision is partially constrained by spatial structure and partially overcomes the spatial restrictions.

B. Visual attention

Let us consider the two spatial dimensions that are maintained spatially on a retina. Spatial structure – specifically spatial neighborhoods – can be exploited by the homomorphic 'campotopic' organization of the retina: neighboring locations in the visual field are mapped to neighboring locations on the retina. As the retina is 'implemented' through physical neural networks, it also is subject to spatial constraints. In particular, neighboring neurons can interact more easily than distant neurons. For many visual tasks, the spatial organization of the retina (as well as of subsequent cortical areas) is very helpful; certain recognition tasks (e.g. edge detection, movement detection) can be performed at retinal processing levels. For other tasks, local neighborhood structures are too restrictive.

This is exactly where eye movements set in. Eye movements also act locally – but at a different scale than the retina. 'Strategic' attention-shifting eye movements depend on retinal information and are capable of abbreviating tedious local propagation of information at lower processing levels. Viewed in this way, eye movements generate spatial neighborhoods at a coarser level of granularity; these neighborhoods are induced by higher-level connections between larger-scale entities. As cognitive agents are equipped with world and domain knowledge, they do not have to analyze all details of their visual input. Strategic selection of specific locations to support (or refute?) a hypothesis is enough to build up a coherent image and understanding of a visual scene.

III. Visual Recognition Strategies

Visual recognition strategies use spatial structure in two orthogonal ways: (1) region growing and boundary detection by relating visual input on a given level of granularity (horizontal processing) and (2) aggregation / refinement by relating visual input across various levels of granularity (vertical processing). While the first type of processes is a data-driven interaction between visual input and neural structures, the second type relies heavily on internal structures e.g. [1], including those retrieved from memory. Together, the two types represent an interchange between bottom-up and top-down visual processing. They result in a sequence of eye movements with which a subject scans a visual scene and organizes features in a linear order [2].

The fact that visual recognition is partly controlled by higher-level processes suggests that high-level knowledge can be strategically employed for active search [3] and for problem solving, as will be discussed later on. One has to keep in mind, however, that as with any feedback system bottom-up and top-down processing do not constitute two separate streams of activities, but rather interact and are codependent in many ways.

IV. Scene Analysis and Image Construction

The spatial ordering of a scene's content that is induced by the scanpath is not just a phenomenon of early levels of visual perception. Instead, on a representation theoretic level, the underlying sequentialization of information can be regarded as one of many principles of spatial organization that transcend the perception of physical space and serve to structure knowledge throughout a number of cognitive (e.g. memory) subsystems, cf. [4]. The concept 'scanpath' instantiates the sequentializing principle in the domain of visual perception; on the other hand, this principle stems from the dominance of linear ordering in visual scene analysis.

On a practical note, where physical space restricts the number of dimensions to three, sequencing the salient features of the space linearly further reduces representational complexity. In mental reasoning, linear representations can facilitate emergence of goal-directed behavior from distributed processes (i.e. by providing a thread).

A. Mental representation

The spatial ordering of visual information can be traced to many later stages of cognitive processing; for example, the sequence of eye movements has been suggested to be part of the mental representation of a visual scene, cf. [3]. As such, it is stored in long-term memory along with other perceived information, and also serves as a spatial key to retrieving memory content.

B. Spatial index in image construction

It is important to underscore that data-driven and top-down streams of visual processing rely in part on highly overlapping cognitive subsystems, such as for the shifting and zooming of the spatial focus of attention [5], and that the subsystems' actual states and outputs constitute combined effects of inputs from both streams. Eye movements can be seen as resulting from some of these outputs, and more, they are reflective of spatial shifts of focus in underlying attentional processes.

That being said, it becomes clear that the concerned output necessarily depends on a relation between input from both streams. Where, for example, input from top-down has to stand against a constant broad-band input through visual perception its overall effect on eye movements has to be less than it is in the absence of perceptual input. The latter situation characterizes the case of mental imagery.

Imagery provides good examples of the pervasiveness of the scanpath / sequentialization concept. Even in the absence

of visual perception, effects of spatial ordering of a scene can be found. [6] show that the scanpath during mental imagery reflects the content of the imagined scene. What is more, evidence suggests that eye movements during mental imagery are not epi-phenomenal but play a functional role for imagery processes [7]. Laeng and Teodorescu further argue that scanpaths might provide a spatial index to the parts of a mental image, a position that is also embraced by [8]. Under this notion, again, the scanpath can be viewed as a part of mental representations that is abstracted from the actual eye movements, and that plays a functional role in general relative spatial indexing.

C. *Scene analysis and image construction combined*

Pictorial representations (including diagrams, sketches, pictures, etc.) have been attributed with particular perceptual, cognitive, representational, and computational advantages over sentential representations [9]. This applies particularly to tasks that can capitalize on the spatio-analogical properties of these forms of representation. Specifically, diagrams have been related to visual mental images, either by representational or by computational metaphors. Cognitive mechanisms involved in the inspection of diagrams and those involved in the construction and inspection of mental images are found to interface at later [10] and earlier stages of mental processing [11], [12]. Such conclusions complement the findings on overlapping cognitive subsystems in both capacities, for example regarding attentional focus.

They also help explain why human reasoning, when it is based on mental images and external diagrams combined, is as effective as it is: mental and external representations are interfaced by the significant overlap in the mental processing of either one of them. Thus, their different representational and computational strengths complement each other rather than that their weaknesses would add up.

V. Mental / Visual Construction

While visual recognition aims at identifying entities and features that are really out there in the visual field, visual problem solving also has to deal with hypothetical views that are supplied by mental construction. The boundary between visual and mental input to attention control can be shifted even further towards the mental side: imagery-driven processes can control eye movements for creative construction tasks.

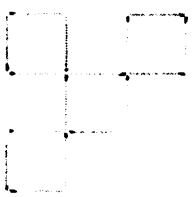
Visual imagery in working memory is subjected to more severe capacity constraints than vision; therefore imagery-based construction reaches a complexity threshold beyond which the consistency of a mental construction cannot be assured. At this point, we can reverse the relation between visual input and mental representation we have maintained in problem solving: whereas mental imagery assisted to fill gaps in the visual domain, we now can employ the visual domain to fill gaps in the mental realm. We do this by externalizing mental images on a visual medium and then employ vision processes to provide feedback about consistency and other properties of the mental construction.

VI. Diagrammatic Problem Solving

We have seen that the organizational principle of linearization, as it is instantiated in the scanpath-concept, among others, permeates the levels of mental processing. Along various lines of exploration, eye movements have been shown to be reflective of attentional shifts as well as of underlying functional organizations of mental representations and processes, both under conditions that involve visual perception or mental imagery. The combination of these two thoughts leads us to postulating that some telling relation holds between eye movements on a visual scene on the one hand and attentional shifts on mental representations of this scene, on the other. With it, robust (i.e. partial) mappings can be defined from eye movements to manipulations on higher-level, cognitive concepts which have been abstracted from a scene, and vice versa.

Clearly, there is no homomorphism between attentional shifts in the oculomotor system and moving foci in executive control of working memory. There are just too many abstraction levels in-between and other components involved to make a simple and direct mapping possible in all cases.

However, many tasks require close coupling between mental imagery and visual perception. For example, designers make extensive use of diagrammatic representations to visualize imagined configurations. With these tasks, attentional shifts are well coupled across abstraction levels, so that eye movement records taken during the problem solving can effectively help generating hypotheses as to what happens on higher (i.e. problem solving) levels at a given point in time.



Add 3 matches to get 6 squares of equal size

Fig. 1. A spatial matchstick configuration task.

A. *Relating eye movements and spatial problem solving*

With respect to mental reasoning about spatial configuration problems, we suggest that eye tracking during the mental construction phase of a solution model in a diagrammatic reasoning task can be employed to robustly assess a reasoner's individual preferences. Fig. 1 gives an

example of a spatial configuration task; in it, a certain number of matchsticks need to be added to construct a specific geometric pattern. The construction process is entirely mental while the reasoner has the problem in view. In the case of the example, four valid different solution models exist, as there are two ways each to create another square by adding two of the three matches, and an additional sixth square by adding the third match.

As a first step of the analysis, the eye movements that occur under problem solving are analyzed post-hoc for various quantitative and qualitative measures, leading to the postulation of actions, behaviors, and problem solving episodes. Based on an evaluation of these concepts, a hypothesis is generated on the constructed solution model, and consequently checked against the solution actually provided by the reasoner.

A simple relation that holds in some cases has the constructed solution as a function of fixation proportions over diagram parts (compare Fig. 2). Generally however, the relations that can be established are less straight-forward, but no less effective.

B. Smart assistants based on eye movement analysis

An analysis of eye movements, however, does not have to end with generating post-hoc hypotheses. As a second step, the developed classification and evaluation routines can be applied *during* problem solving to generate an any-time behavior in which the currently best hypothesis of the eventual result is respectively used to tailor the behavior of a computational assistant in the task. Applications can be found in human-computer collaboration scenarios, in particular in spatial reasoning, as for such collaboration to be satisfactory the computational side has to dynamically adapt to changes in the human reasoner's attentional focus as well as to the problem solving decisions that he takes along the way.

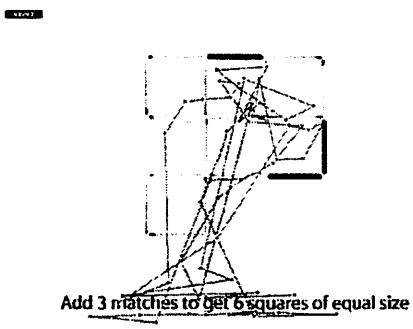


Fig. 2. Eye movements and actually constructed solution.

VII. Conclusion

The research discussed here relies on the hypothesis that attentional shifts in the human reasoner provide indications as to his mental processes and representations. Cognition thus guides attention, at least within the topics conversed here. Eye movements are seen as reflecting attentional shifts. The linear organization of features in the visual scene as scanpaths possess corresponding concepts on the various levels of cognitive processing. We argue that the combination of these properties helps to propagate attentional foci across different levels of processing.

Eye movement recording technology has advanced to a point where it can be used not only to register naturally occurring eye movements but to provide feedback about intentionally induced eye movements in connection with computer-driven visualization. In the context of human-computer collaboration scenarios, the discussed mental and visual construction processes have good potential to be extended towards true dynamic procedures which serve to coordinate concurrent problem solving activities of more than one agent. It is in combination with other human-computer interface technologies that eye movement recording will prove useful in yet more domains.

Acknowledgment

C. Freksa acknowledges countless stimulating discussions and inspiration on the cognitive control of eye movements with Larry Stark over a span of thirty years. The authors also acknowledge fruitful discussions and cooperation with Mary Hegarty and members of her spatial thinking lab.

References

[1] Stark, L., Ellis, S., Inoue, H., Freksa, C., Portnoy, Z., and Zeevi, J. Cognitive models direct scanpath eye movements: evidence obtained by means of computer processing of perceptual scanpath eye movements. in XII Intl. Conf. Medical & Biological Engineering. Jerusalem, 1979.

[2] Noton, D.and Stark, L. Eye movements and visual perception. *Scientific American, 224*(6). 1971.

[3] Stark, L. W. and Choi, Y. S. Experimental metaphysics: The scanpath as an epistemological mechanism. In W. H. Zangemeister, H. S. Stiehl, and C. Freksa, Eds., *Visual Attention and Cognition*. Amst.: Elsevier. 1996.

[4] Engel, D., Bertel, S., and Barkowsky, T. Spatial principles in control of focus in reasoning with mental representations, images, and diagrams. In C. Freksa et al. (Eds.), *Spatial Cognition IV. Reasoning, Action, and Interaction*, pp. 181-203, LNCS. Berlin: Springer. 2005.

[5] Nobre, A. C., Coull, J. T., Maquet, P., Frith, C. D., Vandenberghe, R., and Mesulam, M. M. Orienting attention to locations in perceptual versus mental representations. *J Cogn Neuroscience, 16* (3), 363-373.

[6] Brandt, S. A. and Stark, L. W. Spontaneous eye movements during visual imagery reflect the content of the visual scene. *J Cogn Neuroscience, 9*, 27–38. 1997.

[7] Laeng, B. and Teodorescu, D. Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26*, 207-231. 2002.

[8] Mast, F. W. and Kosslyn, S. M. Visual mental images can be ambiguous: Insights from individual differences in spatial transformation abilities. *Cognition, 86*, 57–70. 2002.

[9] Larkin, J. H. and Simon, H. A. Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99. 1987.

[10] Kosslyn, S. M. and Sussman, A. L. Roles of imagery in perception: Or, there is no such thing as immaculate perception. In M. S. Gazzaniga, Ed., *The cognitive neurosciences* (pp. 1035-1042). Cambridge, MA: MIT Press. 1995.

[11] Ishai, A. and Sagi, D. Common mechanisms of visual imagery and perception. *Science, 268*, 1772-1774. 1995.

[12] Ishai, A. and Sagi, D. Visual imagery facilitates visual perception: Psychophysical evidence. *Journal of Cognitive Neuroscience, 9* (4), 476-489. 1997.

Top-down and Bottom-up Visual Processing Revised

Toyomi Fujita, Claudio M. Privitera, Dimitri Chernyak and Lawrence W. Stark , School of Optometry, University of California at Berkeley.

Abstract—**Top-down informativeness elaboration and bottom-up conspicuity processing are intimately interconnected in visual perception. An internal cognitive model of the external world must necessarily control not only our visual recognition but also the sequence of eye movement jumps. We demonstrated however that a specific combination of conspicuity operators can predict the spatial distribution of hROIs , human visually selected Regions-of-Interest. In this paper we propose a new method for defining bottom-up image-processing algorithms able to predict hROIs based on a self-organizing principal components procedure. A general top-down/bottom-up Bayesian scanpath model will be finally formulated.**

Index Terms—**Image processing, scanpath theory, principal component analysis, top-down visual perception.**

I. Introduction

SACCADIC eye movements and the closely linked mechanisms of visual attention shifts have been attracting scientists for decades. These are very complex functions which utilize many different aspects of the human nervous system from early visual areas and brainstem motor competencies up to high-level semantic elaboration.

Human vision largely rests upon top-down, internal spatial-cognitive models. The Scanpath theory introduced by Noton and Stark [1],[2], embodies such a top-down, intelligent vision approach and relies upon studies of human eye movements. A principal objective of our study was to demonstrate that the loci of human scanpath can be characterized by bottom-up conspicuity and that, based on this conspicuity, it is possible to create algorithms that predict the spatial distribution of scanpath fixations over a picture.

The dichotomy regarding top-down (informativeness) or bottom-up (conspicuity) domination of human vision has continued since the Scanpath theory was introduced and remains an important source of debate. We want in the present paper to further explicate this link between conspicuity and informativeness and see how bottom-up, conspicuity-based, operators can serve to categorize top-down generated hROIs. This categorization is strictly connected with the ability of a given conspicuity operator to predict a general viewing human scanpath and is measured by a spatial metric, S_p, which has been defined on the basis of physiological and experimental observations.

The empirical studies on conspicuity will be finally integrate into a global, cognitive Bayesian framework of visual attention and scanpath eye movements control and discussed in the light of some recent visual attention and conspicuity maps neurophysiology findings.

Corresponding author: Claudio Privitera is with the School of Optometry, 582C Minor Hall, University of California, Berkeley, 94720-2020. (E-mail: claudio@scan.berkeley.edu).

II. Methods: Defining hROIs

Three different sets of images were used in this study; they were classified as natural (images of nature landscape, rivers, mountains), interior (internal environments like a theatre, office space, house kitchen) and geological (for example images of a rocky terrain or a desert). They all are mostly still-life or composition of generic objects, without any close-up human representations or strong color or texture discontinuity which might condition the viewer's looking.

All images are part of a wider collection and classification of images and they can be inspected in [3]. All images, ten for each set, have the same size of 400x400 pixels which correspond 3 to a subtended visual angle of twenty-four degree when displayed on a computer monitor during the scanpath experiments; this is also equivalent to a visual angle of 16F x 16F if we consider F, the foveal angle equal to 1.5 degree, as the basic unit. Standard procedures were used to extract hROIs, (*h*uman Regions-of-Interest), the scanpath sequence of fixational loci. Six different subjects participated in the experiment and no specific viewing tasks was specified.

Subjects repeated the experiment three times over different days; each image viewing session lasted about two seconds and generated an average of seven hROIs for each image (Figure 1, upper left small inset, one single experiment generated four different hROIs indicated by circles and temporally connected by arrows) which yielded to a total of about 3800 hROIs data points for the entire population of images, subjects and repetitions. For each image, clusters of hROIs from different subjects were finally retained as the most representative hROIs (Figure 1, clustered-hROIs, right panel) and used as the training set for the principal component procedure.

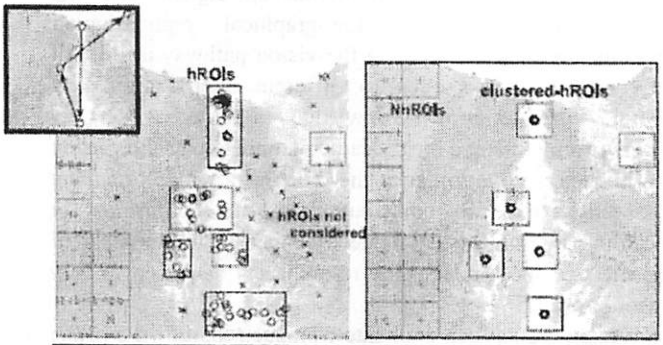


Figure 1: Loci of human visually selected Regions-of-Interest, hROIs, for six different subjects and experiment repetitions are shown with circles and x's in the left main panel for one image of the natural image class. A single experimental instance is shown for example in the upper-left inset, circles and arrows indicates the sequence of recorded hROIs for one subject viewing the image. Clusters of hROIs from different subjects are indicated by a rectangular (left main panel) and finally retained as the most representative hROIs (clustered-hROIs, right panel). The x's in the left panel indicate isolated hROIs and are not considered. Pluses indicate Regions-of-Not-Interest or NhROIs; they are the spatial complement of the hROIs and they likely correspond to not significant visual elements of the viewed image.

III. Visual Conspicuity and Eye Movements

A. *Visual conspicuity operators*

Visual conspicuity is in general recognized as a fundamental property of the visual system which actively influences the control of the selective attention of eye movements behavior [4]. Various forms of conspicuity are associated experimentally to eye fixation loci. The work by Reinagel and Zador for example [5], shows how subjects look at loci that have high spatial contrast and poor pair-wise pixel correlation, a definition of internal intensity non-uniformity. Similarly, higher-order statistics have been used by Krieger *et. al* [6] to characterize eye fixation loci and to propose features like corners, occlusions, and curved lines as conspicuity detectors. Various form of contrast and color and orientation opponency is also proposed by Parkhurst *et al.* [7].

A collection of bottom-up conspicuity image processing algorithms, together with an eccentricity clustering procedure was investigated by our group [8] to generate algorithmic Regions-of-Interest, aROIs and to predict the loci of eye fixations, hROIs. The proposed technology did not rely on our intuition in a priori selection of algorithms, but rather, derived from extensive similarity experiments, and therefore a posteriori. The evaluation of an algorithm was aided by having a clear criterion, that a successful algorithm must find Regions-of-Interest similar in location to those found experimentally by using human eye movements and their fixations as identifiers. Several sources supplied our algorithms. Early-stage vision yielded to center-surround operators or complex concentric features, spots or local projections and in general orientation elements. Canonical algorithms based upon mathematical algorithms such as fractals generative models or wavelet decomposition could also be used to identify hROIs (see also [9]).

B. *Principal component visual conspicuity operators*

Principal component analysis is an important compacting code in computer vision; high dimensional and redundant image representation can be projected onto a small dimensional space using a basis of orthogonal functions which are defined to maximize variance and to represent the fundamental structure of the input set of images. The entire collection of clustered-hROIs were used to train a principal component analysis engine and to define a set of Eigen-hROIs (Figure 2).

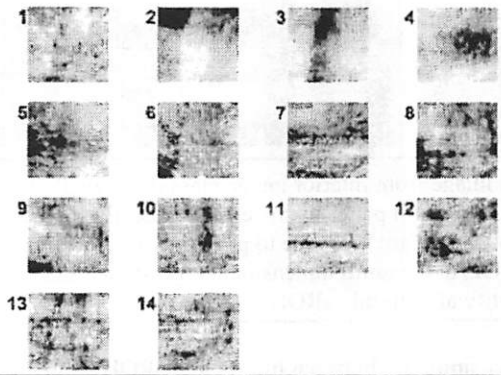


Figure 2: Eigen-hROIs are created from the entire collection of clustered-hROIs (see also Fig. 1) experimentally recorded for the ten images of the natural image class. Only the top fourteen eigen-hROIs are shown, accounting for more than 1% of the variability.

C. *Conspicuity map and aROIs (algorithmic Regions of Interest)*

The collection of the principal eigen-hROIs can be treated as a generic filter bank representing the fundamental structure of the experimental hROIs and they can be applied to any generic image. This is a convolution-like operation which first projects each pixel of the image into the n-dimensional eigen-space and then associate that specific eigen-space point to a scalar energy value defining the distance from the clustered-hROIs. We thus obtained an energy map which preserve the topology of the original input image and where peaks of energy represents proximity to the training clustered hROIs and serve to define aROIs, algorithmic selected Regions-of-Interest.

The procedure for defining aROIs is divided into two basic stages: the energy map is first evaluated to identify an initial large collection of conspicuity local maxima. Maxima are then clustered in a final number of eccentric clusters of maxima (seven, for example) and for each cluster, the highest maximum identifies the position of the aROI (indicated with

squares in Figure 3, right panel). This is similar to the classical winner-take-all sequential maximum conspicuity selection widely proposed in the literature, but it encapsulates a few important advantages as the clustering of the local maxima is a parameter-free process that does not depend, for example, on the size of the image or on the scale of aROIs.

An image from the interior class (Figure 3) was for example processed using the previous eigen-hROIs; the very peripheral surrounding of the image was not considered in the operation and left at zero energy.

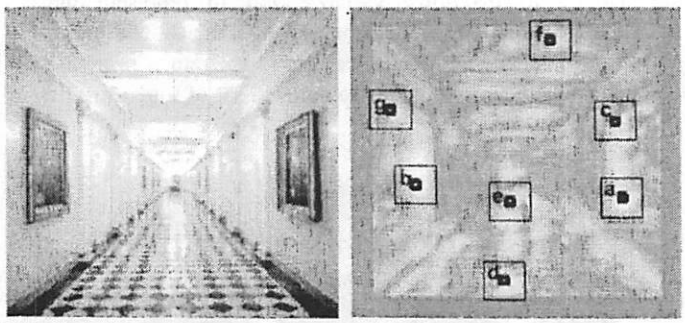


Figure 3: An image from interior image class (left panel) and its conspicuity map (right panel) in the eigen-space. Bright areas or maxima of energy are associate to proximity to the training clustered hROIs in the multi dimensional eigen-space and they serve to identify algorithmic aROIs.

The loci similarity between hROIs and aROIs is expressed by a spatial similarity index, S_p which represents the percent of experimental hROIs which are in close proximity to at least one aROI. An S_p value of 1 means that all hROIs are captured within the aROIs; zero means a complete spatial dissociation between the two vectors. The level of this S_p hROIs-aROIs correlation (0.6) is close the inter-subject Sp similarity (0.69) which represents the top-anchor; algorithms cannot be expected to predict human fixations better than the coherence among fixations of different persons.

IV. TOP-DOWN VS. BOTTOM-UP VISUAL PERCEPTION

A. Top-down feedback propagation

The interaction between visual attention and a hierarchical network of conspicuity maps along the vision pathways has been widely investigated. Conspicuity maps preserve the topography of the external viewed world and activity of each single unit measures levels of conspicuity in the corresponding unit's spatial receptive field; their interaction with the selective attention mechanism has been originally introduced by Didday and Arbib [11].

Cognition must influence these conspicuity representations in a top-down manner and hierarchical cortical top-down feedback have been reported in all levels of the visual pathways [12].

How top-down higher level cognitive computation is matched or synergizes with bottom-up signals and the corresponding conspicuity topographical representation hierarchically distributed along the vision pathway is still not well understood and open to different interpretations. In general, stored *a priori* knowledge is necessary for *recognizing* a picture and the saccadic planning control probably uses this knowledge to detect the most informative or *salient* loci, those that can contribute most to confirming or discarding internal visual scene hypotheses. The process of saccadic eye movements thus resembles a sequence of *best questions to ask.*

Bayesian models can explicate well such top-down *a priori* knowledge / bottom-up (conspicuity signals) interaction; given a specific scene hypothesis, Bayesian conditional probabilities can be used to evaluate the informative value of a region of the image based on the corresponding conspicuity distribution and thus to program the sequence of saccades as proposed by Chernyak and Stark [13].

B. A Bayesian framework for the TD/BU integration

The top-down cognitive-spatial model (Figure 4, top panel) contains a knowledge representation of the elements in the background setting of the picture, the corresponding conditional probability (What), of the loci (Where), and of the sequences (When). A scene hypothesis (Beach Scene) is for example formulated based on the matching of low-resolution (peripheral) conspicuity objects, O^{cs}. These objects can be identified with the conspicuity operators discussed in our conspicuity study, which represent specific spatial and geometric features belonging to the visual scene.

A priori knowledge in the form of conditional probabilities between the hypothesis and the complete foveated object is used to make saccades more accurate and sure in their generation in their attempt to capture, match and thus check on complete, most informative, foveal objects, O^f. Every time the fovea is placed on one of these objects O^f, bottom-up visual signals in log polar retinotopic iconic form arrive via the retina and lateral geniculate nuclei to layers 4 and 5 of the visual cortex (Figure 4, middle panel, Iconic matching). These icons are sequentially, and probably hierarchically, matched with the cognitive-spatial representations projected top-down from other cortical areas. Matching of low-resolution peripheral conspicuity regions (Figure 4, middle panel, Peripheral image acquisition) is important to inform the planning control and help manage internal representation. Matching confirmation results of these hypothesized icons thus either support the ongoing overall perception or contribute to the creation of a new hypothesis. Here is the entire confirmatory scanpath process in an algorithmic format:

Step 1) Various forms of conspicuity objects O^{cs} are analyzed from a low-resolution inspection of the visual stimulus. They can be specific spatial operators, such as the geometrical kernels discussed in this study, and/or other bottom-up conspicuities such as color;

Step 2) Create an initial scene hypothesis (What and Where) based on these conspicuity objects O^{cs}. For example

small triangular patterns on a blue background or vertical dark bars on a yellow background can identify a natural beach scene with sailboats on the ocean and people strolling on the sand;

Step 3) Assign a saliency to each conspicuity based on the contribution of the corresponding object to the current hypothesis (i.e. $P(\text{Boat}^f \mid \text{Beach Scene})$;

Step 4) Fixate the center of the object with the highest saliency that has not yet been fixated (When). The icon matching of complete foveal objects O^{fs} serves to strengthen - or weaken - the probability of the hypothesis;

Step 5) Repeat Step 4 until verification is complete or start again from Step 1 with a new hypothesis if verification is not achieved.

The same mechanism is applicable if the picture is only imagined; fixations made during scanpaths to imagined targets are longer, on average, than saccades to the same patterns when they are visible [14]. The lack of help from a low-resolution peripheral conspicuity guide. They increase the time occupied by computations before a saccade is generated.

V. Conclusions

A basis of eigen-hROIs conspicuity operators are automatically generated from the collection of experimental hROIs and serve to generate conspicuity maps. aROIs, algorithmic-defined Regions-of-Interest are extracted from these maps. We demonstrated that aROIs were able to predict the hROIs loci of two different classes of images used as a testing set, accordingly to S_p. A general Top-down/Bottom-up model was finally discussed.

Acknowledgment

We wish to thank Drs. Yeuk Ho, for his indispensable management of the laboratory computational resources and Michela Azzariti for supervising some of the experiments.

References

[1] D. Noton and L.W. Stark, "Eye movements and visual perception", *Scientific American*, vol. 224, no. 6, pp. 34-43, June 1971.

[2] D. Noton and L. W. Stark, "Scanpath in eye movements during pattern perception", *Science*, vol. 171, pp. 308-311, Jan. 1971.

[3] T. Fujita, http:scan.berkeley.edu/people/toyomi/pictures/.

[4] F. L. Engel, "Visual conspicuity, visual search and fixations tendencies of eye", *Vision Research*, vol. 17, no. 1, pp. 95-108, 1977.

[5] P. Reinagel and A. M. Zador, "Natural scene statistic at the center of gaze", *Network: Comput. Neural Syst.* vol. 10, pp. 341-350, 1999.

[6] G. Krieger, I. Rentschler, G. Hauske, K. Schill, C. Zetzsche, "Object and scene analysis by saccadic eye-movements: an investigation with higher-order statistics", *Spatial Vision*, vol. 13, no. 2,3, pp. 201-214, 2000.

[7] D. Parkhurst, K. Law, and E. Niebur, "Modeling the role of salience in the allocation of overt visual attention", *Vision Research*, vol. 42, pp. 107-123, 2002.

[8] C. M. Privitera and L. W. Stark, "Algorithms for Defining Visual Regions-of-Interest: Comparison with Eye Fixations", *IEEE Trans. PAMI*, vol. 22, no. 9, pp. 970-982, 2000.

[9] C. M. Privitera and L. W. Stark, "Human-vision-based selection of image processing algorithms for planetary exploration", *IEEE Trans. IP*, vol. 12, no. 8, pp. 917-923, 2003.

[10] M. Turk, and A. Pentland, "Eigenfaces for Recognition", *Journal of Cognitive Neuroscience*, vol. 3 no. 1, pp. 71–86, 1991.

[11] R. L. Didday and M. A. Arbib, "Eye movements and visual perception: a two visual system model", *International Journal of Man-Machine Studies*, vol. 7, no. 4, pp. 547-569, 1975.

[12] S. Kastner and L. G. Ungerleider, "Mechanisms of visual attention in the human cortex", *Annual Review Neuroscience*, vol. 23, pp. 315-341, 2000.

[13] D. A. Chernyak and L. W. Stark, "Top-down guided eye movements", *IEEE Trans. on SMC, part B*, vol. 31, no. 4, pp. 514-522, 2001.

[14] S. A. Brandt and L. W. Stark, "Spontaneous eye movements during visual imagery reflect the content of the visual scene", *J. Cognitive Neuroscience*, vol. 9, no. 1, pp. 27-38, 1997.

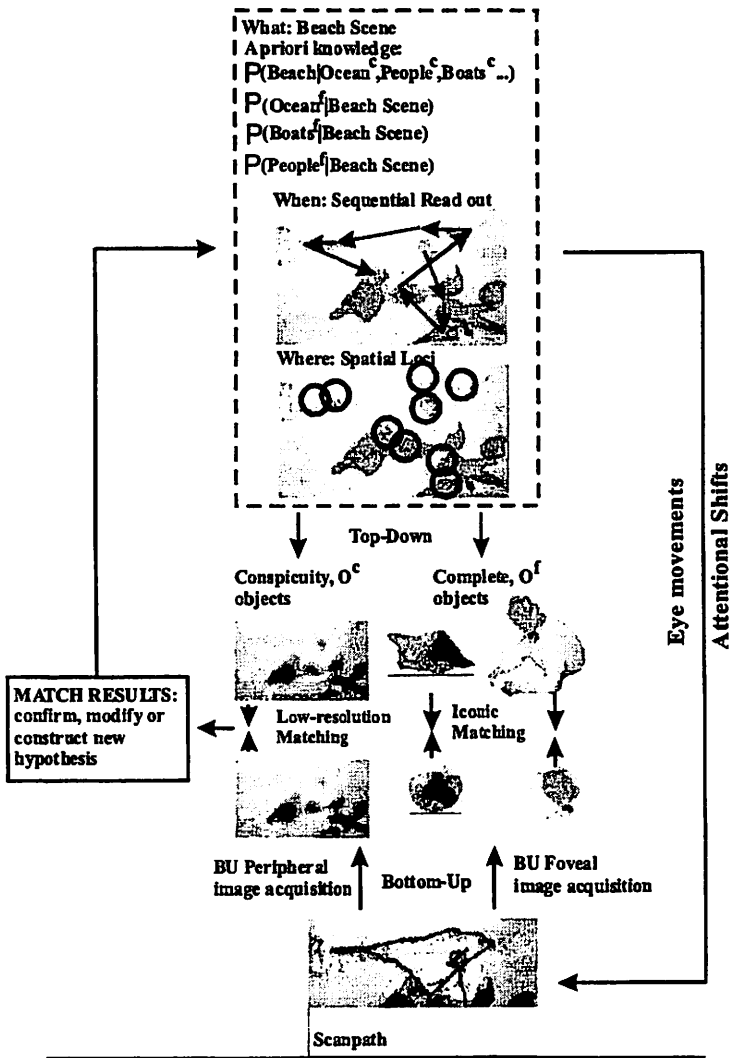


Figure 4: An *a-priori* top-down, cognitive model of the external world is composed of *What*, *Where* and *When* information which contains conditional probabilities of the elements of the picture. This model controls the scanpath shifts of attention (eye movement sequence, bottom panel) and the iconic matching in the visual cortex between foveated bottom-up visual signals and the top-down model cognitive representation of the foveated object O^f. Low resolution conspicuity matching, O^c is also necessary. The result of the matching is given as feedback to the top-down representation to confirm or correct the internal visual

Efference Copy and its Limitations

Bruce Bridgeman, University of California, Santa Cruz

***Abstract*—Efference copy, an internal brain signal informing the visual system of commands to move the eye, was the dominant explanation for visual space constancy for over a century. The explanation is not viable, however; the signal is to small, to slow, and too unreliable to support the perception of perfect constancy. Efference copy is a viable explanation of static position perception and sensorimotor interaction, but the rich, stable visual world is an illusion.**

***Index Terms*—Efference copy, outflow, inflow, corollary discharge, space constancy.**

I. Introduction

Perceiving a stable visual world establishes the basis for all other visual function. The perception seems paradoxical because all of the visual information that we acquire comes through retinal images that jump with each eye movement. A half-century ago, it seemed that the problem of space constancy had been solved with a signal emanating from motor areas of the brain to inform the visual system about when and where the eyes had moved. At the time of an eye movement this signal could be subtracted from the shift of the retinal image, achieving space constancy. The idea is now called efference copy, an extraretinal signal [1] affecting vision but not originating at the retina. It is also called 'outflow', because a signal flows out from the oculomotor centers [2].

II. HISTORY

The first explicit formulations of the efference copy idea originated with Bell in 1823 [3] and Purkinje in 1825 [4] who apparently described the idea independently. Both descriptions are based in part on perceptions that occur when the side of an eye is pressed with a finger. If the eye is pressed in darkness with an afterimage on the retina, no motion of the afterimage is perceived. An active eye movement, though, will result in apparent movement of the afterimage. Experience with a real image is just the reverse – it appears to move when the eye is pressed, but does not move with a voluntary eye movement.

These observations could be explained if an active eye movement elicited an extraretinal signal to compensate for the eye movement, but the eyepress did not. The afterimage fails to move with the eyepress in darkness because the afterimage would remain fixed on the retina while the eyepress did not elicit an extraretinal signal. The movement of the afterimage with an eye movement in otherwise dark surroundings could be explained only by an efference copy, for only the efference copy is changing in this condition. A normal eye movement in a normal environment would not elicit apparent motion because the retinal image motion would be matched by the efference copy, but the eyepress in a normal environment would elicit apparent motion because the resulting retinal image motion would not be compensated by the extraretinal signal. The four conditions are neatly explained with a single theory, as summarized in table 1.

The two conditions in bold type result in space constancy. In the other two conditions, space constancy fails because of a mismatch between efference copy and image movement.

	Retinal image motion	No retinal image motion
Efference copy	**A. Normal eye movement**	Afterimage with saccade
No efference copy	Eypress in normal field	**B. Eyepress in darkness**

Both Bell and Purkinje went further to conclude that gaze movement signals cancelled retinal image displacements to achieve space constancy.

For more than a century after this, efference copy was the major mechanism assumed to mediate space constancy. The dominance of outflow mechanisms in explaining space constancy was assured in 1866 by Helmholtz [5]. He expanded the empirical base for outflow theories with observations of neurological patients who had muscle pareses, so that they could not use part of their oculomotor fields.

Helmholtz extended the efference copy idea to include sensorimotor coordination as well as perception. The patient has two facts to evaluate, for example in pointing with a gaze that is paralyzed for movements to the right:

1) I am looking toward the right.
2) There is an image on my fovea.

His conclusion is that there is an image to his right, though due to the failure of the eyes to move, the gaze has actually remained straight ahead. Helmholtz called this reasoning an unconscious inference, analogous to the processes of formal logic but executed effortlessly and without training. Pointing too far in the direction of the paretic field ('past pointing'), to the right in this case, shows that the patient has no information from eye muscle proprioception or any other source that might inform him of the actual gaze position. It is only the intended

Acknowledgements: Supported by a UCSC faculty research grant to BB.

Corresponding author: Bruce Bridgeman is in the Department of Psychology, University of California, Santa Cruz, Santa Cruz, Ca. 95064 USA. (E-mail: bruceb@ucsc.edu).

gaze position that affects perception and action. Helmholtz called the intention to change gaze position a 'Willensanstrengung', an effort of will.

For the perception of a jump of the world in the direction of an intended eye movement, the explanation is similar. Before the intended jump there is an image on the fovea and a Willensanstrengung straight ahead. After the intended jump the eyes have not moved because of the paresis, but the Willensanstrengung is now directed toward the right, and the same image is still on the fovea. The conclusion is that the image has now jumped to the right, because eye position (as reported by Willensanstrengung) has changed but the retinal image position has not.

Perhaps because Helmholtz saw his eye movement signal as related to the will, he did not analyze it mathematically. The first step in that direction was by Ernst Mach [6], another physicist-physiologist, who explicitly proposed that a neuronal copy of oculomotor efference sums algebraically with the retinal signal.

III. Quantitative Theories

Building on advances in control theory made during World War II, two papers appeared independently in 1950 that defined efference copy theory for the next generation. In fact the phrase 'efference copy' first appeared in a paper in German by Erich von Holst and Horst Mittelstaedt [7] as 'Efferenzkopie'. This was an empirical paper, describing the results of inverting the head of the blowfly *Eristalis* by rotating its neck 180 deg. Von Holst and Mittelstaedt observed that the fly would circle continually. When the fly was in darkness, though, its locomotion seemed normal. With light restored, the fly would circle either in the original direction or in the opposite direction at random. The results were explained by assuming that the fly monitored the output of its locomotor system and compared that output with the retinal flow field. The copy of locomotor efference, the 'Efferenzkopie,' would be subtracted from the retinal signal to stabilize locomotion by negative feedback. Inverting the head converted the negative feedback to a positive feedback – a random nudge in one direction would feed back a signal to 'correct' in the same direction. That would result in a further deviation in the same direction, and continuous circling would result. Von Holst and Mittelstaedt formalized their idea with an engineering flow diagram and algebraic analysis, the efference copy exactly cancelling the afferent signal. This seminal paper also introduced the terms exafference, a change in retinal motion signals resulting from motion of objects in the world, and reafference, a change in retinal motion signals resulting from movements of the organism.

Sperry [8] made similar observations in a fish whose eye he inverted surgically. He concluded that his fish's normal swimming in the dark excluded the possibility of brain or nerve damage, and introduced the term 'corollary discharge' to identify the efferent signal. These papers formalized the quantitative compensation idea that had dominated physiology and psychology for more than a century. The new evidence offered for the idea was motor rather than sensory in nature, an emphasis that would prove important in the coming decades, though some speculations about perception were made.

A. Limitations of compensation theories

Though compensation theories completely dominated thinking about space constancy up to this point, there had always been some problems with them. Considerations from control theory, which had developed rapidly during World War II, made these problems clear.

First, the efference copy is a feedforward, a signal the informs the brain of where the eyes ought to be rather than where they actually are. As such it cannot be exact – it should drift with time, and is not corrected when it is in error. Yet the perception of space constancy is perfect – the world does not appear to jump in the slightest when the eyes move. To the average person, the idea that the world should jump with each saccade seems bizarre at best. If perception is rock solid, but the efference copy is not, something else must be supplementing the feedforward signal, and that something else might be all that is necessary to do the job.

Recognizing that efference copy could not be perfect, E. Matin [9] proposed that saccadic suppression could mask the inevitable errors. It was known then that displacements of the entire visual world would not be detected perceptually if they occurred during saccadic eye movements [10, 11]; if the imprecision of efference copy was less than the displacement thresholds during saccades, space constancy could be maintained despite small mismatches of efference copy and retinal displacement. Matin's solution was the best idea available at the time, but it didn't last long. The first parametric description of saccadic suppression of displacement showed that at the optimal timing of image displacement and saccade, the perceptual threshold was about 1/3 as large as the saccade itself [12]. Clearly, any visual orientation mechanism that tolerated an error of one part in three had no idea where the visual world was, and could support neither perceptual space constancy nor a reasonable visual-motor calibration.

This result along with similar observations should have led to a capitulation of the efference copy theory, but it did not. The reason why is that a theory cannot be abandoned because of evidence; it can only be replaced by another theory.

Other problems with the efference copy theory soon emerged. One of them began with the technique of reverse modeling, applying an output (behavior) to a linear model and running the equations backward to read the input (nerve signals to the muscles) that must have driven the behavior. Applied to the oculomotor system, reverse modeling clarified the motor signals that drive voluntary nystagmus, a rapid oscillation of the eyes that can be performed by a small proportion of the population. The oscillations are small in amplitude, usually 3 deg. or less, but high in frequency, up to 20 Hz. The resulting rotatory accelerations of the eye are so great that the oculomotor driving signals can be generated only by the saccadic controller. Even though the movements have a nearly sinusoidal profile, they must be elicited by the pulse-step mechanism of saccades; the sinusoidal appearance is a result of temporal filtering by the oculomotor plant.

All of this is relevant to the space constancy question because subjects experience oscillopsia, a back-and-forth fluttering of the visual world, during voluntary nystagmus. In short, space constancy breaks down. But normally, space constancy survives saccades, which are accompanied by saccadic suppression. What is going on? The possibility that small saccades do not elicit saccadic suppression was disproved by Nagle, Bridgeman & stark [13], who compared suppression during voluntary nystagmus to suppression during single voluntary saccades matching the amplitude of nystagmus in the corresponding subject. The suppression was virtually identical in both cases; saccadic suppression did not necessarily accompany space constancy. Further, an afterimage remained motionless during voluntary nystagmus (Table 2), showing that the changes of eye position failed to elicit changes in apparent position. The mechanism of constancy must have some other basis.

	Retinal image motion	No retinal image motion
Space constancy	***B. Normal saccade***	***C. Nystagmus with afterimage***
No space constancy	Nystagmus in normal field	***D. Saccade with afterimage***

The voluntary nystagmus experiment showed that single isolated saccades were accompanied by space constancy, while rapidly alternating saccades of the same size were not. Perhaps the space constancy mechanism was still operating, but could not keep up with the rapidly alternating saccades of voluntary nystagmus. A better temporal resolution of the constancy/frequency relationship was achieved in studies of the apparent movement of an afterimage with saccades in darkness. Grüsser, Krizic & Weiss [13] asked subjects to make saccades from one loudspeaker to another, cued by tones from each speaker. After a bright light gave a lasting afterimage, saccades were performed in darkness. Grüsser et al. measured the subjects' estimates of the spatial separation of the afterimages when the eye was aimed at the left versus the right speaker. As saccades became more frequent, the subjective separation of the afterimages became smaller and smaller, until at the highest saccade frequency (about 3.8 saccades/sec) the afterimage appeared to remain fixed in front of the subject. Space constancy had failed completely. The result confirmed that perceptual compensation is quite slow. Even for saccade intervals well within the temporal range of saccades accompanying normal perception, the compensation was much smaller than the saccade amplitudes.

The temporal properties of space constancy were linked directly with efference copy a few years later, in experiments exploiting the deceptively simple maneuver of pressing on the outer canthus of the eye. Explaining the method in these experiments requires a brief diversion into methodology.

The eyepress has been misunderstood for centuries, since Purkinje's 1825 assumption that it resulted in a passive eye movement, as reviewed above. Helmholtz [5] made the same assumption, that pressing on the eye moves it passively, and that the resulting apparent motion originates from retinal image movement without an efference copy. Two observations support this interpretation: First, the entire visual world appears to move in the direction opposite the eyepress; and second, the eye of another person appears to move when it is observed during their eyepress. The two observations are consistent with one another, but both are misinterpretations.

The observation of image movement is based on the inference of motion of the retina, but the two kinds of motion are not necessarily linked. The real situation is easily demonstrated – simply pick a fixation target, then slowly press on the outer canthus of one eye with the other closed. You will find that you continue to hold your gaze at your fixation target, even while the entire visual world, fixation target and all, appears to move. This means that the retina is not moving at all with respect the visual world. The motion originates with the successful effort of oculomotor tracking mechanisms, which cannot be turned off, in keeping the eye on the fixation target. The effort requires oculomotor innervation, and with it an efference copy. So, far from demonstrating an effect of passive eye movement, the eyepress demonstrates the effects of active compensation.

The second misinterpreted observation is that the eye of another person performing an eyepress appears to move. What the observer sees, however, is not an eye rotation but a lateral translation of the eye in the orbit. The eye is influenced by two rotational forces in opposite directions; one originates from the pressing finger producing a nasalward rotational force, and the other is an equal and opposite force generated by the lateral rectus muscle. The oculomotor innervation is driven by a retinal slip initiated by the finger, but compensated by an involuntary optokinetic tracking system. Thus the two rotatory forces cancel. But each of these forces also introduces a translation in the medial direction, the finger pushing the anterior part of the eye nasalward and the lateral rectus pulling the posterior part of the eye nasalward. The translational forces sum to move the eye several mm in the orbit [14]. Because the cue that humans use to perceive movements of the eyes of others is the amount of sclera visible on the two sides of the iris, observers misinterpret the translational motion as rotation. The rotation of the occluded fellow eye, which is not cancelled by the eyepress, provides an objective measure of the forces applied.

Now the eypress technique, which causes a deviation in the efference copy without a change in the retinal image position, can be applied to the problem of measuring the temporal aspects of efference copy. It was possible to use scleral search coils in both eyes simultaneously, and also to press on the eye without popping out the required scleral contact lens. Again the non-pressed eye is occluded, so that its movements are measured in darkness. In this experiment, rather than replicating the static experiments of Stark & Bridgeman [14], we pressed repeatedly on the viewing eye in a roughly sinusoidal pattern [15]. A force transducer on the fingertip provided an objective record of eyepress frequency.

We found that at low temporal frequencies the viewing eye does not rotate, replicating Stark & Bridgeman [14]. Rather, the occluded eye rotates under its occluder, revealing

the compensatory oculomotor innervation; according to Hering's law, that innervation affects both eyes equally. When we began pressing more rapidly on the eye, however, the compensation was no longer complete. At a rate of less than 1 Hz the occluded eye still rotated, but in addition the viewing eye rotated passively as it was repeatedly pressed and released. At the surprisingly low rate of 2 Hz, the occluded eye ceased its rotation completely, and only the viewing eye rotated, in the passive manner that Purkinje and Helmholtz would have predicted. Interpolation of our data implied that the oculomotor compensation system ceases to function at about 1 - 1.5 Hz. The implication is that any efference-copy based system that normally contributes to space constancy must cease to function at these relatively low rates, well within the bandpass of normal perceptual events.

By 1989, then, evidence was converging on the idea that efference copy could not be responsible for space constancy. It was too slow, and its gain too low to support a perceptual compensation for saccades. There was also evidence of a more qualitative sort that should have eliminated compensation theories, but did not, again because of the lack of an alternative. One bit of evidence came from an experiment on saccadic suppression by Brune & Lücking [16], who fed an eye movement signal into a mirror that moved an image with the eyes, but at variable gain. At low gains, when the image was moving one-tenth as far as the eye, the image appeared always to be stable, replicating Bridgeman et al. [17]. But at a slightly higher gain, when the world as a whole continued to appear stable, 'prominent objects' would seem to jump or jiggle with each saccade. Efference copy theories, however, do not allow parts of the image to move relative to one another. The observation would seem to eliminate all efference copy and related theories in a single stroke.

The Brune & Lücking experiment might have resulted in dissociations for uninteresting reasons, however. The prominent objects might have been brighter than the background, for example, so that signals coding them would move through the visual system faster than the signals from dimmer parts of the image. In a continuously moving environment, this might result in prominent objects being perceived in different locations than the background. Another possible artifact is that the prominent objects might have been fixated, so that signals from them would course through slower, high-acuity channels, with the same result.

All of these possibilities were eliminated in a replication and extension of the study that used tesselations of a plane by the Dutch artist Maurits Escher as the stimulus materials [18]. Escher used two repeated shapes that interlocked to completely cover a surface. For instance, devils and angels might tesselate a plane. Some subjects could selectively concentrate on just the angels, or just the devils, at will. All of those subjects saw slight movement of the attended figure while the 'background' figure remained stable, at a near-threshold feedback gain from eye movement to image movement. Because this perception occurred for both figures, without any change in the stimulus, all image variables were controlled. No compensation theory can account for this result.

IV. The Alternative – No Compensation

Efference copy was finally discarded as a space constancy mechanism with a new theory, centered on visual search and on a reanalysis of what information is carried over from one fixation to the next. According to the visual search interpretation, attention shifts to the saccade target before the saccade is executed [19]. Due to the attention shift, location and visual attributes of the target and of surrounding objects are stored in transsaccadic memory. After the saccade, the visual system searches for the previous object of attention within a restricted spatiotemporal "constancy window" which is about 50 ms in duration, and is confined to a few degrees around the saccade target. If the object is found, the world is assumed to be stable. Spatial information from the previous fixation is discarded or ignored, and localization proceeds using the currently available information. Only if the object is not found do outflow and other information sources come to bear.

The more radical part of this reanalysis posits that little is carried over from one fixation to the next; we do not build a visual world by pasting together samples calibrated with efference copy, but simply use what is currently available, plus a gist and a few previously attended objects. The stable, rich visual world of our perception is more promise than physiological reality. Extraretinal signals are used in static conditions, though, especially for controlling motor behavior [20].

References

[1] Matin, L., Eye movements and perceived visual direction. In *Handbook of Sensory Physiology,* vol. 7 part 3, D. Jameson and L. Hurvich, Eds. New York: Springer, 1972, pp. 331-380.

[2] Teuber, H.-L., Perception. In *Handbook of Physiology,* sect. 1; *Neurophysiology,* vol. 3, J. Field and H. Magoun, Eds. Washington, DC: American Physiological Society, 1960, pp. 1595-1668.

[3] Bell, C., Idea of a new anatomy of the brain (1823). In *Francois Magendie, Charles Bell and the Course of the spinal Nerves,* P. Cranefield, Ed. Mt. Kisco, NY: Futura, 1974.

[4] Purkinje, J., Über die Scheinbewegungen, welche im subjectiven Umfang des Gesichtsinnes vorkommen, *Bulletin der naturwessenschaftlichen Sektion der Schlesischen Gesellschaft,* vol. 4, pp. 9-10, 1825.

[5] Helmholtz, H. von, *Handbuch der physiologischen Optik.* Leipzig, Voss, 1866.

[6] Mach, E. *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen,* 5th edition, Jena, Fischer, 1906.

[7] von Holst, E., and Mittelstaedt, H., Das Reafferenzprinzip. Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften,* vol. 27, pp. 464-476, 1950.

[8] Sperry, R., Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative Psychology and Physiology,* vol. 43, pp. 482-489, 1950.

[9] Matin, E., Saccadic suppression: A review and an analysis. *Psychological Bulletin,* vol. 81, pp. 899-917, 1974.

[10] Mack, A., An investigation of the relationship between eye and retinal image movement in the perception of movement. *Perception & Psychophysics,* vol. 8, pp. 291-298, 1970.

[11] Wallach, H., and Lewis, C. the effect of abnormal displacement of the retinal image during eye movements, *Perception & Psychophysics,* vol. 81 pp. 25-29, 1965.

[12] Nagle, M., Bridgeman, B., and Stark, L., Voluntary nystagmus, saccadic suppression, and stabilization of the visual world, *Vision Research,* vol. 20, pp. 1195-1198, 1980.

[13] Grüsser, O.-J., Krizic, A. and Weiss, L.-R., Afterimage movement during saccades in the dark. *Vision Research*, vol. 27, pp. 215-226, 1984.
[14] Stark, L. and Bridgeman, B., Role of corollary discharge in space constancy, *Perception and Psychophysics*, vol. 34, pp. 371-380, 1983.
[15] Ilg, U., Bridgeman, B., and Hoffman, K.-P., Influence of mechanical disturbance on oculomotor behavior, *Vision Research*, vol. 29, pp. 545-551, 1989.
[16] Brune, F. and Lücking, C. Okolomotorik, Bewegungswahrnehmung und Raumkonstanz der Sehdinge. *Der Nerverarzt*, vol. 240, pp. 692-700, 1969.
[17] Bridgeman, B., Hendry, D., and Stark, L., Failure to detect displacement of the visual world during saccadic eye movements, *Vision Research*, vol. 15, pp. 719-722, 1975.
[18] Bridgeman, B. Cognitive Factors in Subjective Stabilization of the Visual World, *Acta Psychologia*, vol. 48, pp. 111-121, 1981.
[19] Deubel, H. Bridgeman, B. & Schneider W. X. Different effects of eyelid blinks and target blanking on saccadic suppression of displacement. *Perception & Psychophysics*. vol. *66*, pp. 772-778, 2004.
[20] Bridgeman, B. and Stark, L., Ocular proprioception and efference copy in registering visual direction, *Vision Research*, vol. 31, pp. 1903-1913, 1991.

Error! Bookmark not defined. (PhD Stanford, 1971) was a NIH postdoctoral fellow with Lawrence Stark at the UC Berkeley School of Optometry in 1973.

Accommodation and Size-Constancy of Virtual Objects

Robert V. Kenyon, Moses Phenany, and Daniel Sandin, Electronic Visualization Lab, University of Illinois at Chicago, IL 60607

Abstract—**Accommodation as been suspected as a contributor to size illusions in virtual environments (VE) due to the lack of appropriate accommodative stimuli. Simple experiments examining size-constancy in VE have shown that 2D cues to depth are a major contributor to correct size perception. When these cues are removed a visual angle performance occurs. An open-loop accommodation situation using a Pinhole viewing system did not restore size-constancy performance in our subjects as would be expected, thus signifying that accommodation does not play a dominate role in this process.**

Index Terms— **accommodation, pinhole, size-constancy, virtual environment**

I. Introduction

Users of the CAVE® (CAVE Automatic Virtual Environment) and other VE systems have reported that they perceive virtual objects to be incorrectly sized. This effect can be attributed to a variety of factors including hardware errors, software errors, and perception errors. This study examines the role of perception errors that might be induced by inappropriate accommodative stimuli.

Many studies on the perceived size of objects in the physical world have been performed. Decartes [1] first described the phenomenon known as "size-constancy" where an object is perceived as being the same size regardless of its distance from the observer even though the retinal size of the object gets smaller with increasing distance from the observer. Holaday [2] showed that removal of various cues would change this behavior to one relying on the physical optics of the situation. She showed that as the number of two-dimensional (2D) cues to depth [eg. shadows, motion parallax, etc] is reduced performance suffers and subjects adopt a size judgment that is based on the visual size of the object on the retina also know as visual angle (VA) size judgments. Holaday [3] confirmed these findings for objects from 10-40ft from the observer. Harvey and Leibowitz [4] showed similar results at distances of 1-9 ft from the observer. Furthermore, they and [5] showed that removal of 3D cues to depth (i.e. Stereovision) had little to no effect on performance and that performance was only affected by the removal of 2D depth cues. Kenyon et al. [6] have shown this same effect in VE; i.e., 2D cues to depth play a major role in size-constancy performance.

However, previous experiments have established that accommodation plays a role in a person's estimate of distance to an object [7] (an important component in object size judgments). Interestingly, in virtual systems like the CAVE, accommodation is stimulated by the distance the user is from the CAVE wall regardless of where the virtual objects appear in space. Hence, the accommodative information does not necessarily correspond to the distance that the virtual object is drawn from the user. Because of this lack of congruity between the accommodation and other visual information, it has been suggested that accommodation could be an important factor in the lack of size-constancy when many 2D cues are absent from virtual scene. Our experiments examine whether accommodation is a factor in making size judgments of virtual objects when few 2D cues to depth are available. To determine the influence of accommodation, a pinhole viewing system is used to open the feedback loop of accommodation eliminating blur and the associated accommodative stimulus.

II. Methods

A. The CAVE

The CAVE is a projection-based VE system [8] where 4 screens are arranged in a 10ft cube composed of three rear-projection screens for walls and the fourth projector (overhead) is directed by a mirror to the floor. A viewer is supplied stereovision through stereo shutter glasses (Stereographics, Inc). Proper perspective images are drawn for each eye using head position from a six-degrees-of-freedom head-tracking device (Intersense 900) calibrated to an accuracy of ±0.5 inch. Subject's interpupillary distance (IPD) was measured (R.H. Burton Digital P.D. Meter) and used by the CAVE program to generate the stereo images for that subject. A joystick and buttons in a hand-held wand provided the needed interaction with the VE to change the size of a virtual object.

Acknowledgements: This research was supported by General Motors Research contract #2-5-37457.

Corresponding author: Robert Kenyon is in the Department of Computer Science, 851 S. Morgan; University of Illinois at Chicago; Chicago, IL, 60607 USA. (E-mail: kenyon@uic.edu).

[8] The CAVE is a registered trademark of the Board of Trustees of the University of Illinois.

An SGI Onyx with two Infinite Reality graphics pipelines, each split into two channels controlled the projected images. The image resolution was 1024 X 1024 pixels (Marque 5000 projector) with a refresh rate of 120 Hz, i.e., an effective stereo refresh rate of 60Hz.

B. Pinhole Glasses

Pinhole glasses worn by the subjects, open-looped accommodation and reduced the effect of the CAVE wall accommodative stimulus during the size judgments. The pinhole glasses provided a single ≅ 1 mm diameter aperture for each eye that resulted in an approximate 35° field of view (FOV). Users could adjust the distance between the pinholes (horizontal direction) for their specific IPD. They could also adjust the pinholes in the vertical direction. Users were asked to adjust the pinholes so that the image for the two views had a 20% binocular overlap along the horizontal axis.

C. Neutral Density & Reduced Field of View Filter

To provide the same viewing conditions as with the pinhole glasses but without open-loop accommodation affects, an aperture and a neutral density filter were fixed to the outer surface of the shutter glasses to equal the light intensity and FOV produced by the pinhole glasses. The circular aperture placed over the shutter glasses approximated the 35° FOV provided by the pinhole glasses. The neutral density filters (3f stops) placed over this circular opening was used to approximate the reduced light intensity encountered when using the pinhole glasses. The circular opening was 0.25 inches in diameter and could be adjusted along the horizontal and vertical axes. Users were again asked to adjust the filters as described above.

D. The Physical World

A 2-liter coke bottle was placed on a black plastic table at a height of 4 ft. The table was positioned at the front right hand side of the CAVE at an approximate distance of 4 ft from the subject. The height of the coke bottle was 12 inches tall and 5.5 inches (maximum) wide. The coke bottle was lit by a spotlight mounted on the top left hand wall of the CAVE at a height of 10 ft and at a distance of 8 ft from the front CAVE wall.

E. The Virtual World

The virtual world simulated 2 different scenes: (a) (ENV) The scene consisted of a gray green checkered floor with a wood textured table and a coke bottle on top of the table (Fig. 1). The coke bottle was textured with an image of the real coke bottle. The height and appearance of the table was changed for different sets of measurements, as was the distance of the coke bottle from a subject (2, 3.5, 5.0, 6.5, and 8.0 ft from subject). Subjects used the wand joystick to increase and decrease the size of the coke bottle and a wand button to continue once they had finished sizing the virtual coke bottle. (b) (NoENV) In this scene, subjects were presented with only a gray background and few 2D cues to depth. The Coke bottle appeared suspended in mid air at the same distances and heights (from the floor) from the user as in ENV condition.

F. Procedure

Three viewing conditions were tested for each subject: (1) (REG) Subjects vision was unobstructed and only the shutter glasses were worn (FOV: 100°H X 50°V). (2) (PIN) Subjects wore pinhole glasses and wore the shutter glasses over the pinhole glasses. (3) (ND) Subjects wore the aperture and neutral density filter over the shutter glasses.

Fig. 1. The ENV visual scene contained a number of 2D cues to depth. The patterned floor and textured table provided information for the subject to make size judgments. The floor and table were absent in the NoENV visual scene.

For each viewing condition described above, subjects were tested using 2 environments, ENV and NoENV, under which they had to size the virtual coke bottle. Furthermore, for each of the conditions, subjects sat on a chair facing the front wall of the CAVE and were placed at 3 different view distances from the front screen: 6.5 ft (Far), 5 ft (Mid), and 3.5 ft (Near). These conditions were randomly sequenced for each subject. The physical coke bottle was visible to the subjects by simply turning their heads. The initial size of the virtual coke bottle was randomly varied from 0.2–2.0 of its correct size and the subject was required to size the virtual coke bottle to match the size of the real coke bottle placed at that distance from the subject. Ten coke bottle sizing operations were performed for each virtual bottle location.

The first run in each experiment was a trial run using the ENV scene. This allowed the subjects to familiarize themselves with sizing the virtual coke bottle. Data was collected but not used in the analysis. Subjects were encouraged to take 5 minute breaks as often as they needed to avoid eye fatigue. The total experiment time varied from 45 minutes to 75 minutes.

G. Subjects

All subjects were healthy with normal oculomotor function, visual and stereo acuities. Each subject was familiar with the operation and environment of the CAVE.

H. Analysis

The subject's setting of the virtual object size was stored by the computer and then compared to the true size of the object. To evaluate subject performance, we developed a measure called size-ratio that represents the size of the perceived bottle set by subject divided by the correct bottle size. Consequently, when the subject is performing according to size-constancy the Size-Ratio values will be 1 at each bottle location. Otherwise, virtual bottles that appeared too large would be reduced by the subject resulting in size-ratios less

than 1. Coke bottles perceived too small would be enlarged and result in size-ratios greater than 1.

Because of the samall number of subjects, data was analyzed separately, i.e., no group statistics were calculated. Statistical tools used were Microsoft Excel and SPSS for Windows. Statistical methods used included single factor ANOVA, and multiple factors ANOVA and post hoc t-tests. The test for similar slopes of regression lines was done using confidence intervals. The mean size-ratio (for 10 sizing operations) for each bottle location served to create regression slopes using the least squares method. Bottle distances were grouped hierarchically by case, then by subject, then by condition and finally by view distance.

III. Results

Subjects viewing the scene under the REG condition, showed sizing behavior that was consistent with size-constancy when the ENV scene was presented (Fig. 2). Within each subject, we found no statistical difference in the size-ratio settings as a function of bottle distance from the subject. The regression slopes relation (as a percentage of the expected VA slope) shown in Table I are very small (REG-ENV) and speak to the consistency of the responses from these subjects for bottles at different distances [6].

The removal of most of the 2D cues to depth, NoENV scene, showed a performance that was closer to VA (Fig. 2). The data from this condition fell well below the size-ratio of 1 when the virtual bottle was close to the subject. There was a significant difference between size-ratio values in the NoENV scene compared to the ENV scene at 2.0 and 3.5 ft from the subject. Furthermore, the bottles at 2.0, 3.5, and 5.0 ft from the subject were significantly different from each other viewing the NoENV scene. The percent VA slope for all subjects' REG-NoENV regression slopes were significantly different from REG-ENV ($p < 0.045$) slope at all viewing distances except for Subject 1 at the Mid viewing distance (Table I).

Table I
The percent agreement between subjects' size-ratio regression slopes and that predicted by VA regression slopes (%)

Subject	View Dist.	Regression Slope Percent of VA			
		REG (ENV)	*REG* (NoENV)	*PIN* (NoENV)	*ND* (NoENV)
Sub 1	Near	0.6	34.4	55.2	53.3
	Mid	-2.5	-3.0[a]	41.0	61.5
	Far	2.4	9.1	15.4	30.4
Sub 2	Near	3.9	24.0	124.1	81.2
	Mid	8.5	14.5	51.0	45.5
	Far	2.8	9.4	36.7	32.9
Sub 3	Near	19.5	39.0	83.8	78.0
	Mid	17.5	26.0	62.0	58.5
	Far	8.7	32.5	37.1	36.7

The open-loop accommodative condition, PIN, showed the same VA sizing behavior while viewing the NoENV scene as was reported for the REG condition. The resulting data in Table I show a similar significant increase in the percent VA slopes compared to REG-ENV. Furthermore, reduction in the size-ratio for bottles close to the subject was seen in all subjects, e.g., near bottles looked larger than normal, despite the open-loop condition for accommodation. For example, examination of the data for subject 3 in Fig. 2 shows close adherence of these data using PIN condition to the expected VA slope.

Allowing accommodative feedback but retaining the same FOV and intensity, ND, showed a similar response as the PIN condition. The percent VA data in Table I remain distant from size-constancy behavior. Further examination of subject 3 data in Fig. 2 shows that the ND data is similar to that found in the PIN condition.

Using linear regression analysis (with distance as the predictor variable and size-ratio as the dependent variable) and comparing the regression slopes between the NoENV condition and the ENV condition in the REG case revealed that slopes are higher for the NoENV case. Also, calculating confidence intervals for the slopes and checking for intersections between the intervals revealed that the slopes (with one exception) are significantly different.

Comments from subjects indicated that the ENV scene under the REG condition was the easiest to perform while the NoENV scene under all conditions was substantially more difficult. Both the PIN and ND conditions were the most difficult to perform; subjects took longer to set the bottle size. Furthermore, we found that the standard deviation of the mean increased as bottles moved away from the subjects for all conditions tested.

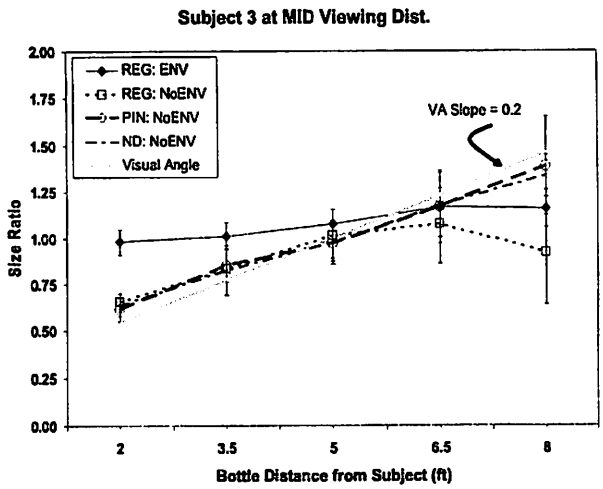


Fig. 2. Comparing the REG condition viewing the ENV scene with the REG, PIN, and ND conditions viewing the NoENV scene reveals that the absence of 2D cues shifts the sizing curve from size-constancy to visual angle. Also, open-loop accommodation (PIN) has no effect on performance this performance. The grey line is the expected slope for visual angle performance.

IV. Discussion

Our results, while not conclusive, strongly suggest that accommodation does not clearly influence size judgments made by the subjects in our experiments. Comparing the PIN, ND, and REG conditions when viewing the NoENV scene shows higher slopes than that expected under size-constancy performance. Furthermore, these slopes were more like those expected for VA performance and were significantly different than those obtain viewing the ENV scene. The performance using PIN and ND conditions caused the behavior to appear more like VA than with the REG-NoENV condition. In any event, the PIN condition did not restore the behavior expected for size-constancy seen in the REG-ENV condition. If accommodation was playing a dominant role we would expect the subject's performance to return to or approach size-constancy rather than VA. Since accommodation is open-loop when viewing the image through the pinhole glasses, inappropriate accommodative signals should be reduced and other visual information such as convergence would dominate size and distance judgments about virtual objects. However, this was not the case in our subjects.

V. Conclusion

The conditions of virtual environment limitation on accommodative stimuli and its inappropriateness to many of the positions of the virtual objects does not appear to be the main reason for perceptual illusions related to inappropriate size of virtual objects. As in the physical world where convergence and accommodation are coincident, two dimensional cues to depth [lighting, perspective, shadows] promote the behavior of size-constancy. Perhaps a larger population of subjects would illuminate this tendency further. Studying other aspects of VE visual condition such as the quantization of the image necessitated by the use of computer graphics might better instruct us on the factors that are important for size-constancy behavior to be maintained in VE.

Acknowledgment

The authors thank our subjects for their diligence and tireless interest in these experiments. We also thank Rajesh Balasubramanian, and Tim Portlock for their contributions to these experiments.

References

[1] R. Descartes, *La Dioptrique, discours VI: de la vision*, vol. 1. Cambridge, England: Cambridge Univ. Press,1985, 1637.

[2] B. E. Holaday, "Die Grössenkonstanz der Sehdinge bei Variation der inneren und äusseren Wahrnehmungsbedingungen," *Arch. ges. Psychol.*, vol. 88, pp. 419-486, 1933.

[3] A. H. Holaday "Determinants of Apparent Visual Size with Distance Variant," *American Journal of Psychology*, vol. 54, pp. 121-151, 1941.

[4] L. O. Harvey, Jr. and H. W. Leibowitz, "Effects of exposure duration, cue reduction, and temporary monocularity on size matching at short distances," *J Opt Soc Am*, vol. 57, pp. 249-53, 1967.

[5] H. W. Leibowitz and R. A. Dato, "Visual size-constancy as a function of distance for temporarily and permanently monocular observers," *Am J Psychol*, vol. 79, pp. 279-84, 1966.

[6] R. Kenyon, D. Sandin, R. C. Smith, and R. Pawlicki, "Size constancy in the CAVE," *Presence*, Submitted for publication.

[7] H. Leibowitz and D. Moore, "The role of changes in accommodation and convergence in the perception of size," *Journal of the Optical Society of America*, vol. *56*, pp. 1120-1123, 1966.

[8] C. Cruz-Neira, D. Sandin, T. Defanti, R. Kenyon, and J. Hart, "The CAVE Audio-Visual Environment.," *A.C.M. Trans. on Graphics*, vol. 35, pp. 65-72, 1992.

Robert V. Kenyon (M'72) received his BS in electrical engineering from the University of Rhode Island, Kingston, RI, 1971, MS in bioengineering from the University of Illinois at Chicago, IL, 1972, and Ph.D. at University of California, Berkeley, CA, 1978. He was a student of Professor Lawrence Stark from 1972 to 1978. He was a faculty member at MIT's Department of Aeronautics and Astronautics from 1979-1986. He moved to University of Illinois at Chicago in 1986 where he is an associate professor of Computer Science. His work has concentrated on studying visual-vestibular interaction through measurements of motor responses. Of particular interest is how animals and humans adapt to unusual environments. Recent work has applied VE coupled with robot manipulators to assess and aid in the recovery of motor function in stroke patients.

Characteristics of Eye Movements in Age Related Macular Disease: Preliminary Results

Vasudevan Lakshminarayanan
College of Optometry and Dept. of Physics and Astronomy
University of Missouri
One University Blvd., Saint Louis, MO 63121, USA

Abstract
Eye movements of normal subjects as well as subjects with age related macular disease who were fixating on a stationary target were analyzed. It appears as though these two groups exhibit different characteristics. A degree of control appears to be missing in the AMD data allowing the movements to follow a simple model. Movements of normal subjects have a more complicated structure.

Introduction

The eye while fixating at a stationary point undergoes small eye movements. These small eye movements have three different components: tremor, drift and microsaccades. Because of these eye movements, the retinal image moves around in an irregular manner and traces out a path whose dimensions are very large compared to the foveal photoreceptor. The amplitude of the tremor is substantially smaller than other micromovements and is of the order of the diameter of the cones. The measurements of these eye movements have been reported by a number of investigators (e.g., 1,2). Sophisticated models of the oculomotor movements have been produced (e.g., 3). Although these drifts and jerks seem random, and are not always goal directed, they do stay within a small region about the fixation point, and therefore imply some external position control element. These movements may be modeled by a 2-dimensional Langevin equation with noise. The first stochastic model of these eye movements was reported by Vasudevan et al (4). This was extended to two dimensions by Lakshminarayanan (5). The presence of these eye movements have clinical implications (e.g., 6).This brief paper investigates the characteristics of eye movements in both normal subjects and in two subjects with AMD.

Methods

Eye movement records were collected and kindly supplied by Dr. Ronald Schuchard using a scanning laser ophthalmoscope. Sets of normal and diseased data were plotted to get a general idea of the movements. These included plots of position and the path taken (Figure 1), normal distribution plots (Figure 2) and plots representing changes in successive movements (Figure 3). The plots of position and path suggested that the data may be following a Langevin equation of the form given below with a shift of origin if necessary (5,7):

$K_1(\mathbf{q}) = -\alpha q_1; K_2(\mathbf{q}) = -\alpha_2 q_2;$
with
(x,y)=(q_1, q_2) = $\mathbf{q}$

V(q) = $\frac{\alpha_1}{2} q_1^2 + \frac{\alpha_2}{2} q_2^2$ plays the part of a potential function and

$\dot{q}_i = -\alpha_i q_i + F_i(t); i = 1,2$

The above equation is simply the Langevin equation and the F's follow a white noise process. This allows one to use the Fokker Planck equation in two dimensions as:

$$\dot{f} + \sum_{k=1,2} \frac{\partial}{\partial q_k} \left(\left(K_k f - \frac{1}{2} \sum_{l=1,2} Q_{k,l} \frac{\partial f}{\partial q_l} \right) \right) = 0$$

which can be shown to have a stationary solution,

$$f(\mathbf{q}) = N \exp\left\{ \frac{-2V(\mathbf{q})}{Q} \right\}$$

provided that

$$K_k = -\frac{\partial}{\partial q_k} V(\mathbf{q}); Q_{k,l} = \delta_{k,l} Q$$

Note that irrespective of whether the data follows the distribution f, the above equations do not tell us anything about the way in which the movements occur, notably considering the large and small drifts and jerks (see the histograms of the distance of each jump for normal and AMD subjects; figure 3).

Results

Figure 1 shows the typical eye movements during fixation. The range of the fluctuations, particularly in the Y direction differ largely between the normals and the AMD subjects. The scale for the normal data in the Y direction is strectched, and there is much more variation in the X direction than in the Y direction. Groups of points are present in both cases, though there is far more variation in the AMD data. The jerks are still present in the normal data though (see figure 3a,b). In Figure 2 the normal plots for each case in X and Y are displayed. Note the departure from the normal distribution for the *normal* data. Even with the small groups of points outside the main group (see Figure 1), removed, the remaining points still fail to follow the normal distribution rigorously.

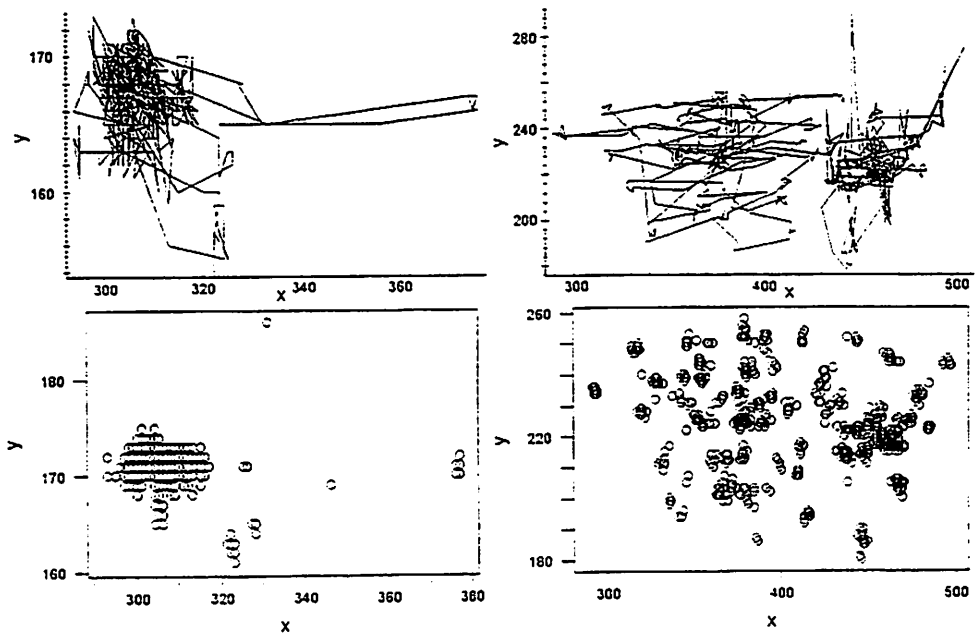


Figure 1: Upper panels show typical paths for normal (left) and AMD(right0 subjects. Note the grouping of points clearly seein in the lower panels.

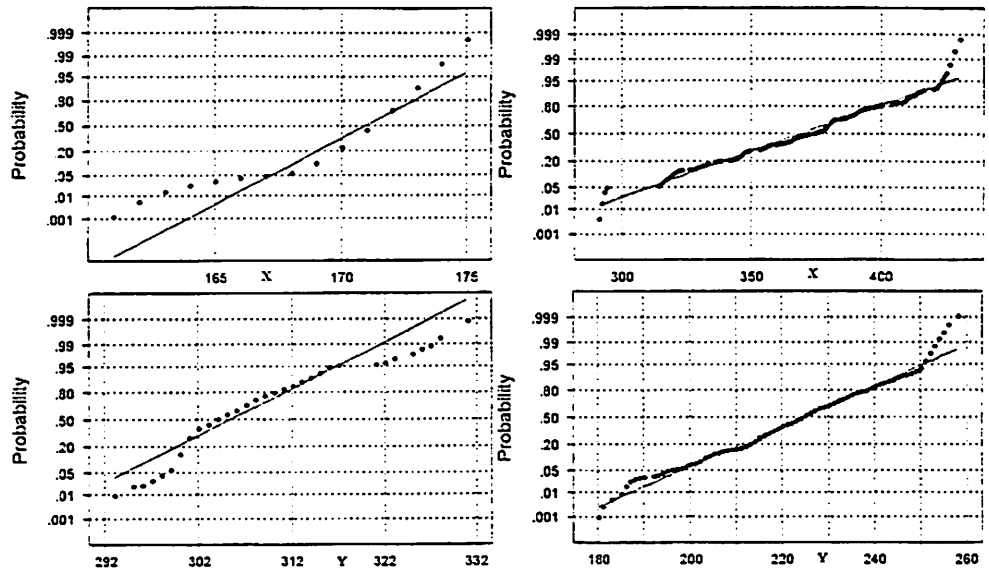


Figure 2: Normal probability plots for X (upper panels) and Y (lower panels) in normal (left) and AMD (right)subjects.

Figures 3 a and 3b show the distribution of the size of each jump. This distribution does not fall away like may be expected in an exponential. This is due to the superposition of the distributions for the drifts and jerks. When looking at the consecutive jumps, there seems little difference between the data, but the plots do show that jerks rarely occur close in time.

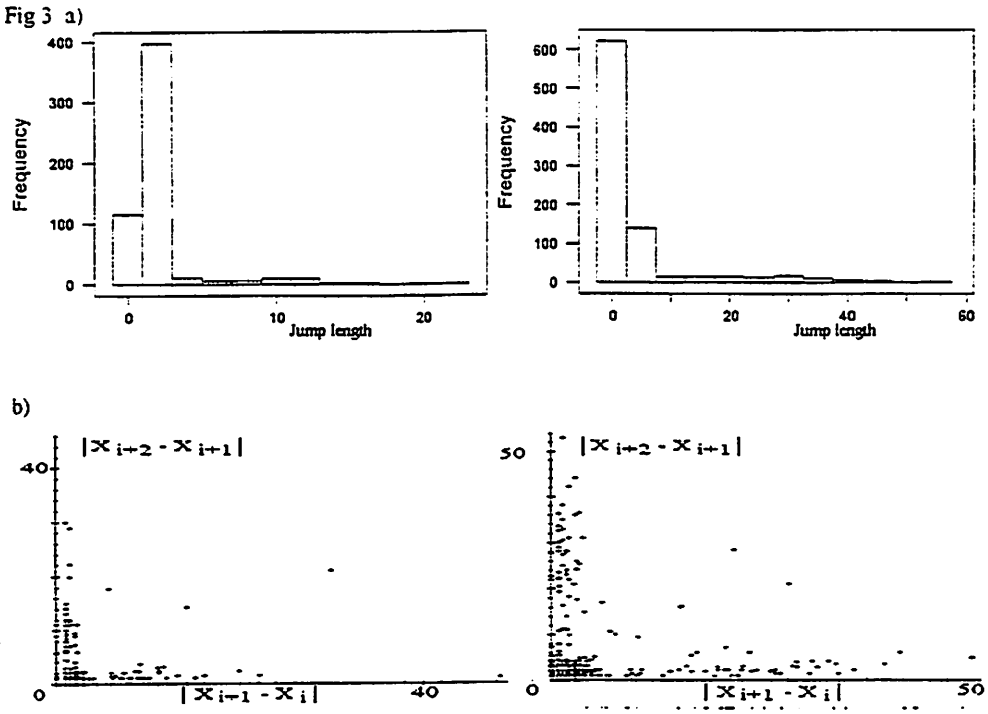


Figure 3: (a) Distribution for the size of each jump for normal (left) and AMD (right) subjects. Note the scale difference for the length of the jumps. (b). Size of jump vs. size of next jump. Most points appear near the axes and very few appear in the central regions. Very few jerks happen consecutively.

Discussion

By considering the probability density function, we find it is equivalent to a bivariate normal distribution. Hence the marginal distributions f_x and f_y are normal distributions. In Figure 2, we see that this is the case only for the AMD data. the normal data does not follow the normal distribution at all (even with all the extreme and spurious data removed). There is also a difference in the X and Y distirbutions for the normal data. In both cases though, the spread in the Y direction is much less than that in the X direction.

When taking into account the different scales of the jumps and the way in which they occur, it seems as though the jerks obey the Langevin equation. This suggests that in Langevin equation, the F should be replaced by two separate components, one for the drifts and one for the jerks. In the normal case, the F's for the jerks may still be random, but F(t) becomes F(q,t). The solution of the Fokker Planck equation then becomes an approximation for the distribution in normal data if the F(q,t) is not symmetric and the same everywhere.

In Figure 3a, for the normal data, the large jumps are very approximately of about

size 10 arbitrary units. In Figure 1 we see that this is about the size of the central group. Also, in Figure 4a it seems that these jumps are much better aimed toward the center than for AMD movements. Therefore, it seems that they are more controlled in normal eyes and tend to jump across the center. The path in Figure 1 and the angles in Figure 4a show this is not the case for AMD eyes.

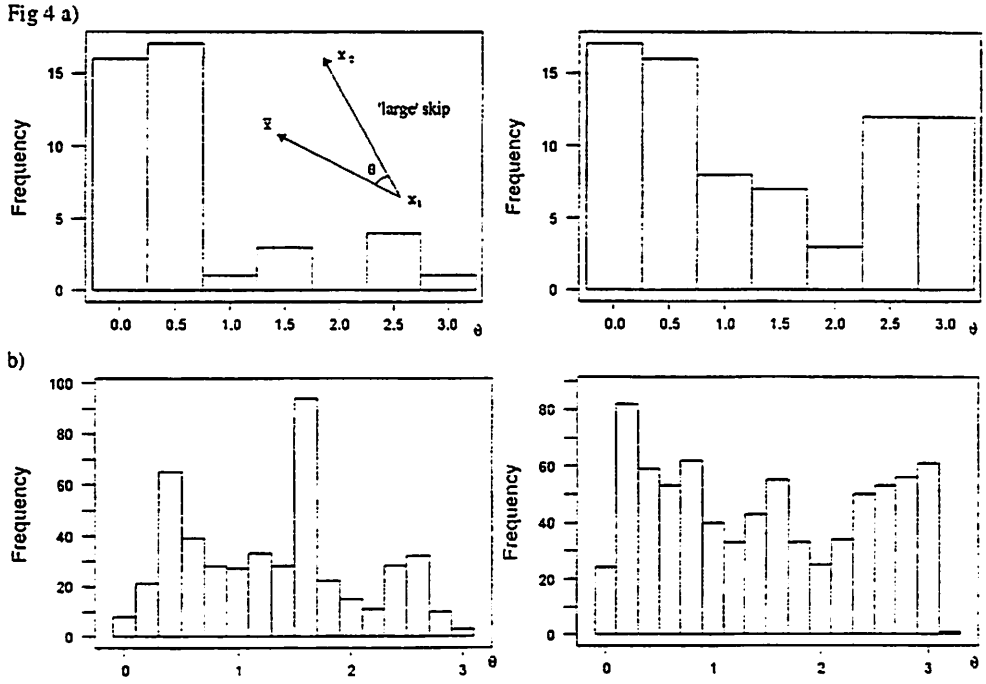


Figure 4: (a) Angle of jerks relative to the overall center. Skips greater than (mean+standard deviation of the length of jumps) were considered large. (b) as in (a) for all jumps, i.e., mostly drifts.

Since the angle of each jerk is quite random in the AMD data, F(q,t), is approximated quite well by F(q), and the probability density given above can be taken to be correct. For the normal data, it is not just the restoring forces in the Langevin equation, that are responsible for the keeping the eye looking toward the center. The jerks are directed more toward the center which is probably responsible for the spread (Table 1). The probability density is a poor approximation in this case.

	Normal	AMD
Standard deviation, X	10.943	50.288
Standard deviation, Y	2.136	15.722
Mean distance from center	6.5632	47.838
Mean length of jump	1.8775	4.0421
Std. deviation of jump	3.728	7.669

Table 1: Summary statistics. Arbitrary units

Between each jerk, which occur on a relatively short timescale, the eye does not drift back toward the center in either case (Figure 1 and the the distribution of the angles of drifts 4b). It appears as that it is solely the action of the jerks that keep the eye from wandering too far away. Thus, the restoring forces in the Langevin equation seems neglible. In fact, a fundamental question is if using the Langevin equation in this way is appropriate at all. This should be a topic for further research. As the F depends on spatial position, this suggests future investigations into the characteristics of microsaccades. Do the directions and size depend on distance from say, the center? Does the average distance from the center increase with time between blinks? Errors occcur in the data when the eye blinks, and the system has to reposition "from scratch". The presence of the AMD condition in the data could easily be seen in the normal probability and path plots. Drifts and jerks are present in both sets of data, however, the way in which the jerks occur appears to differ . This simplistic model needs further improvement and analysis. Especially needed are studies regarding the characterisitics of the jerks.

This type of analysis would have delighted and stimulated Larry Stark. His deep insight into the oculomotor system in particular, and biomedical control in general is deeply missed.

References

1. R.W.Ditchburn & B.L.Ginsborg, Involuntary eye movements during fixation, Nature, **170**:36, 1952.
2. P.R.Boyce, Monocular fixation in human eye movement, Proc.R.Soc.Lond.B, **167**:293-315, 1967.
3. S.L.Lehman & L.Stark, Simulation of linear and non-linear eye movement models: sensitivity analysis and enumeration studies of time optimal control. Cybern.Info.Sci., **4**:21-43,1979.
4.R.Vasudevan, A.V.Phatak & J.D.Smith, A stochastic model for eye movements during fixation on a stationary target. Kybernetik, 11:24-31, 1967.
5. V. Lakshminarayanan, Stochastic eye movements while fixating on a stationary target, in *Stochastic Processes and Their Applications*, ed. A.Vijayakumar and M. Sreenivasan, Springer-Narosa, New Delhi, Pages 39-49, 1999.
6. V. Lakshminarayanan,R.A.Knowles, J.M.Enoch & R.Vasudevan, Measurement of fixational stability while performing a hyperacuity task using a scanning laser ophthalmoscope: preliminary results, Clin. Vis. Sci., 7:557-663, 1992.
7. H.Haken, *Synergetics: Introduction and Advanced Topics*, Springer, Heidelberg, 2004.

Technological developments in studying neural information processing in patterned neuronal networks *in vitro*

Yoonkey Nam[1], Bruce C. Wheeler[2], Moon-Hyon Nam[3]
[1]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign
[2] Department of Bioengineering, University of Illinois at Urbana-Champaign, USA
[3] Department of Electrical Engineering, Konkuk University, Korea

***Abstract*—Spatially defined biological neural networks were created from dissociated rat brain tissues and grown on culture dishes with embedded metal electrodes (planar multi-electrode array). Neural networks matured in two weeks and were spontaneously active. Electrical stimulations evoked phase-locked responses. Multichannel recordings provided a chance to study the neural coding developed in live neural networks *in vitro*.**

***Index Terms*—cell culture, microelectrode array, multichannel recording, patterned neuronal network**

I. Introduction

PLANAR multielectrode arrays (MEAs) have been used to study the properties of electrogenic tissues including acute brain slices, brain cells (cortical neurons, spinal cord neurons, hippocampal neurons, etc.), and heart cells. One could either detect small extracellular potentials created by action potentials or deliver currents to generate electrical fields to stimulate nearby tissues through the electrodes (usually 60 electrodes) [1-4]. There have been efforts to develop cell-based biosensors for screening neurotoxins using this type of system [5, 6]. It has also been demonstrated that one can study learning related phenomena in cultured cortical neural networks [7, 8].

In our laboratory, we have been developing technologies to design neural circuits *in vitro* by controlling the growth of neurons from rat embryos [9-11]. Advanced lithography, surface chemistry, cell culture techniques and multichannel data processing capability made it possible to consider designing and characterizing the patterned neuronal network *in vitro*.

Here we overview the recent progress on the implementation of patterned neuronal circuits on MEAs. Dissociated hippocampal neurons derived from rat embryos were successfully grown in pattern for a month. Multichannel recordings allowed us to monitor the spatially distributed or correlated network activities. Time-locked evoked responses were induced by electrical stimulation through theses electrodes. Statistical analyses such as rate histogram, interspike interval histogram, auto-correlogram, cross-correlogram, perievent histogram were routinely used to characterize the neural data. Devising the proper measure to represent the multidimensional data was very difficult task in general.

Acknowledgements: This work was supported by the National Science Foundation under Grant, EIA 0130828 and by the National Institutes of Health under Grant R01 EB000786

Corresponding author: Yoonkey Nam is in the Beckman Institute, 405 N. Mathews Ave.; Urbana, IL 61801 USA. (E-mail: ynam1@uiuc.edu).

II. Methods

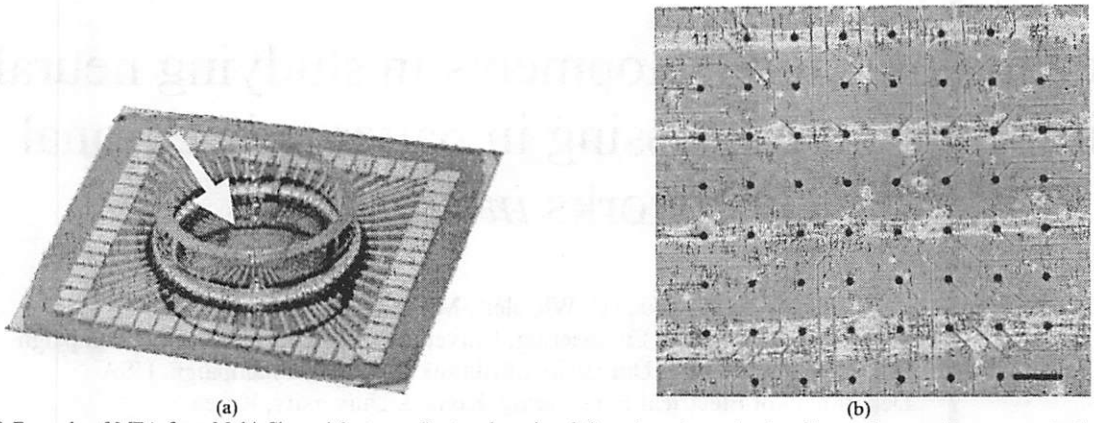
A. Cell culture

Dissected hippocampi (18-day gestation Sprague-Dawley rat embryos) were purchased from Brain Bits™ (www.BrainBitsLLC.com). Tissues were mechanically dissociated and cultured at 37°C, 5% CO_2, 9% O_2, in serum free B27/Neurobasal medium (Invitrogen, Gaithersburg, MD) with 0.5 mM glutamine and 25 μM glutamate. Each dish gets about 4,000 ~ 16,000 cells depending on the area and plating density (50 ~ 200 cells/mm^2). Cultures were maintained by changing half of the medium weekly, without glutamate.

B. Multi electrode arrays

MEAs (Fig. 1(a)) serve as a culture dish as well as a metal electrode array. The metal layer pattern is photolithographically defined on the substrate (glass) and coated by an insulation layer. Then via holes are opened to be microelectrodes (size of 10 ~ 30 μm in diameter). Contact pads are formed at the edge of the substrate to make external connections.

There are a few commercially available MEAs. We purchased 60 channel MEAs from Multi Channel Systems (Fig. 1(a), Reutlingen, Gemany). They had 60 electrodes (8×8 configuration with four corners empty) and each electrode was made of TiN (titanium nitride, 10 μm in diameter) and AC impedance was 300 kΩ ~ 500 kΩ at 1 kHz. A metal layer is insulated by silicon nitride (0.5 μm in thickness). We also used custom designed MEAs from Elume Inc. (Semi Valley, CA). These MEAs were insulated with polyimide (4 ~ 5 μm in thickness) layer and electrodes were made of transparent indium-tin oxide conductor. The tip of the electrodes was platinized to reduce the AC impedance for low noise recording



(a) (b)

Fig. 1. (a) Example of MEA from Multi Channel Systems. Center glass ring defines the culture chamber. External connections are made though edge contact pads. The arrowed area is magnified in part (b). Substrate size: 5 cm × 5 cm, (b) Center of MEA with live cultures at 24 days *in vitro*. This culture was linearly patterned (see text), scale bar: 200 μm.

and large charge storage for electrical stimulation.

C. Surface chemistry

The surface of the substrates (here, the insulator of MEAs) needs to be treated to permit attachment and growth of neurons for a long term (at least two to three weeks in culture). Usually, this is done by coating the whole surface with cell adhesive biomoelcules such as poly-D-lysine (PDL, applied 0.1 mg/ml in deionized water).

To create the surface pattern that would control the growth of neurons, we imprint PDL pattern in micrometer scale by inking the PDL solution onto silicone rubber microstamp and transferring to the surface by micro-contact printing (stamping) procedure [12]. The procedure was often assisted by the protein linking chemistry to form a strong covalent bond between the surface and biomolecules, which also proved to extend the longevity of the patterned cultures up to a month [9]. Detailed procedures could be found elsewhere [10, 11]

D. Neural recording and stimulation

During the experimental sessions, a culture containing MEA dish was pulled out of the incubator and transferred to the recording chamber, where the MEA was connected with the 60-channel amplifier unit (Gain 1200, bandwidth 10 – 3000 Hz, MEA 1060 system, Multi channel Systems). The temperature of the MEA was maintained at 37 °C with a heating stage, and controlled gas (5% CO_2, 10% O_2, backfilled with nitrogen) was fed to the recording chamber to maintain the pH level of culture medium near 7.4. A silver-silver chloride wire or a large ITO electrode was used as a reference electrode.

The continuous stream of 60-channel data were digitized with a 64-channel AD board (MC Card, f_s = 40 kHz per channel, Multi channel Systems) and displayed and stored using dedicated software (MC Rack, Multi channel Systems). Post analyses were usually performed by replaying the recorded data and identifying the spikes with a thresholding method and classifying the spikes ('spike sorting') if there were multiples units in a same channel.

For real time experiments, the Multichannel Acquisition Processor developed by Plexon Inc. (Dallas, TX) was used to do tasks such as spike detection, spike sorting, spike rate histogram, and interspike interval histogram. Signals larger than 5 times the standard deviation of the background noise were regarded as neural spikes. The spike detection window was 1.4 ms in total with 0.6 ms for pre-threshold. When a certain number of spikes (usually more than 100 spikes) were collected, they were sorted by template matching or principal component analysis. Each channel could have four different units that were processed individually.

Biphasic or monophasic voltage pulses were delivered directed to one of the electrodes on the MEA. The magnitude of the stimulus varied from 0.1 to 2.2 V. The pulse duration was 0.2 ms for each phase. Each stimulus was delivered every 1 second (1 Hz) or 2 seconds (0.5 Hz). Both recording and stimulating channels suffered from large stimulation artifacts. It was not possible to access the stimulating channel for more than a few hundred milliseconds after the stimulation. For neighboring electrodes, early responses were readily accessible after digitally adjusting the baseline as long as the amplifier outputs are not saturated by the stimulation artifacts. In this way, spikes that appeared as early as 3 – 5 ms could be detected easily by simple threshold crossing method.

E. Data analysis

After extracting spikes from raw data, time stamp series (spike trains) were processed using Neuroexplorer (NEX, Nex Technologies, Littleton, MA). Rate histograms, auto correlograms, cross correlograms and perievent histograms were calculated. The mean frequency and confidence limits were calculated under the assumption that the spike train has a Poisson distribution. These limits were used to tell whether the calculated histogram shows any statistically significant events. More complicated analysis such as burst analysis or neural population vector analysis (principal component analysis) could be carried out by this software.

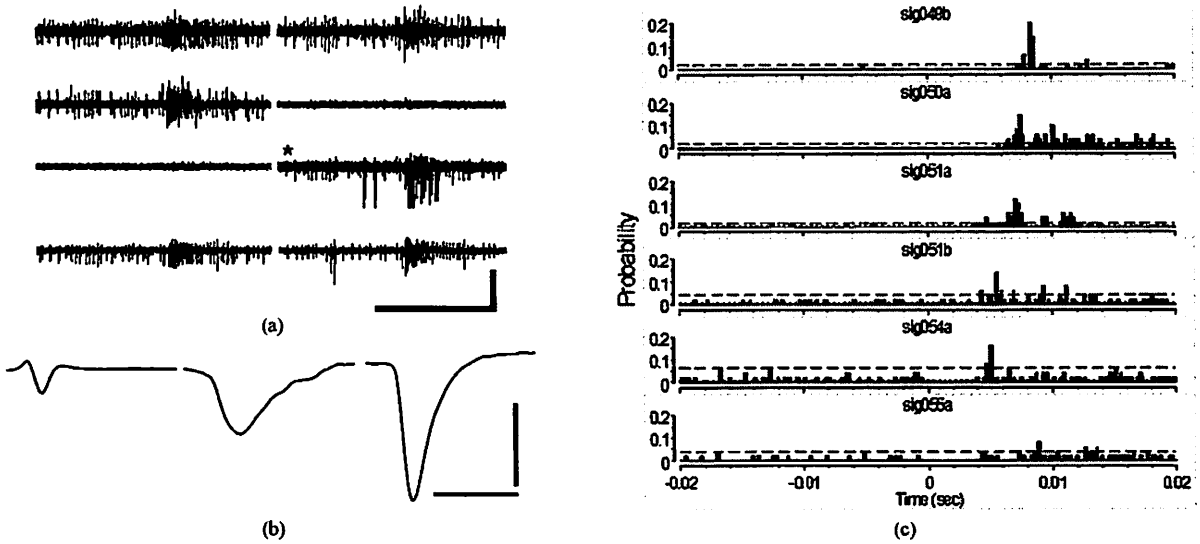


Fig. 2. (a) Spontaneous recording from patterned cultured in fig. 1(b), scale bar: 1 s, 200 μV. (b) Three different units sorted from the marked channel in part (a), scale bar: 1 ms, 200 μV. (c) Perievent histogram constructed after recording evoked responses from a line pattern similar to fig.1(b). Sig055a was stimulated by biphasic voltage pulses.

III. Results and Discussion

A. Formation of patterned neuronal network

We can form a patterned network on MEA and maintain it up to a month. So far, most of the patterned cultures were dense with clusters of cell bodies and bundles of neural fibers. Low density high resolution cultures were also possible with relative low yield due to the poor viability. We are actively pursing the techniques to improve the yield of the low density high resolution cultures.

The longevity and the compliance of the patterns were dependent upon the physical dimensions of the pattern. For example, 40 μm wide linear pattern produced high yield cultures. 83% of the linearly patterned cultures (10 out of 12 cultures) showed good compliance to the surface pattern. When grid patterns were tested, only 50% (6 out of 12 cultures) were evaluated as being in a good compliance.

B. Spike detection and sorting

The background noise level was 3 ~ 4 $\mu V_{(rms)}$. Spikes larger than 15 ~ 20 $\mu V_{(0\text{-}peak)}$ were readily detected by the thresholding method. Often, when cells were tightly coupled to the electrode, we were able to record spikes as large as 500 ~ 700 $\mu V_{(peak\text{-}peak)}$. As an example, Fig.2(a) shows the spontaneously bursting activities recorded from linearly patterned network (Fig.1(b)). Multiple units were detected from a single channel (see Fig.2(b)). Various spike waveforms spikes were detected.

C. Spontaneous activity of patterned cultures

Neurons in patterned cultures showed spontaneous activities as early as 7 days in culture (Fig.2(a)). However, robust and reproducible recordings could be found only after 2 weeks in culture. Most of the electrodes showed correlated, periodic bursting. Often a few channels displayed tonic firing behaviors. We constructed 64 lines from 16 cultures and 54 lines were intact and healthy. Each line was overlaid with 8 electrodes which could report activities. Forty one lines were spontaneously active at the time of recording. Mean firing rate was distributed from 0.1 Hz to 30 Hz. There was no significant difference in mean firing rate when we compared different patterns of cultures.

These spontaneous activities could be modulated by applying drugs such as bicuculline or Mg^{2+}. Bicuculline increased the spike rate and made the bursting rate more regular. High level of Mg^{2+} could reversibly block the activities suggesting the involvement of NMDA receptors in observed activities.

D. Electrical stimulation and evoked responses

Voltage stimulation was used since this was easier to control the electrode voltage than current stimulation. Evoked responses showed all or none behaviors near the threshold level. Positive first biphasic voltage pulses were more effective than negative first biphasic pulses.

More electrodes recorded evoked responses as the stimulus intensity increased. Since we were stimulating through extracellular fields, a higher stimulation level depolarized more units near the electrode.

The latency of the evoked spikes was within 10 ms, which is likely to be synaptically or axonally conducted responses. Estimated conduction velocity ranged from 0.1 – 0.3 m/s. Fig.2(c) [13] is an example of axonally propagated direct response in linearly patterned cultures. Estimated conduction velocity was 0.3 m/s.

E. Connectivity of neuronal network

Patterned cultures had a fan-in and fan-out structure. We performed a detailed analysis of one culture with a grid pattern culture and found that each electrode connected with 5.0 ± 4.3 electrodes (N = 21, mean ± s.d.) and more than one electrode (3.9 ± 3.3 electrodes, N = 21, mean ± s.d.) evoked responses

from the same electrode. There were some pairs of electrodes that showed reciprocal connectivity with slightly different latencies.

More elaborate analysis paradigm is required to efficiently investigate multichannel recording and stimulation data set.

IV. Conclusion

Designing the neural circuit in dish is in general a very challenging task. Our works show that it would be possible to create a patterned network *in vitro* by combining the state-of-the-art technologies in various fields including micro-fabrication, surface chemistry, and cell cultures. Multichannel data analysis techniques showed that artificially created neural circuits had desired characteristics to be minimally functional. We are currently using these model networks to study the relationship between the form and the function of the network.

Acknowledgment

Yoonkey Nam thanks Rudi Scharnweber and Elizabeth Ujhelyi for the support in cell culturing, and Kathleen Motsegood for the technical assistance in microfabrication.

References

[1] C. A. J. Thomas, P. A. Springer, G. E. Loeb, Y. Berwald-Netter, and L. M. Okun, "A miniature micoelectrode array to monitor the bioelectric activity of cultured cells," *Experimental Cell Research*, vol. 74, pp. 61-66, 1972.

[2] G. W. Gross, "Simultaneous single unit recording *in vitro* with a photoetched laser deinsulated gold multimicroelectrode surface," *IEEE Transactions on Biomedical Engineering*, vol. 26, pp. 273-279, 1979.

[3] J. L. Novak and B. C. Wheeler, "Multisite hippocampal slice recording and stimulation using a 32 element microelectrode array," *Journal of Neuroscience Methods*, vol. 23, pp. 149-159, 1988.

[4] J. Pine, "Recording action potentials from cultured neurons with extracellular microcircuit electrodes," *Journal of Neuroscience Methods*, vol. 2, pp. 19-31, 1980.

[5] J. V. Selinger, J. J. Pancrazio, and G. W. Gross, "Measuring synchronization in neuronal networks for biosensor applications," *Biosensors and Bioelectronics*, vol. 19, pp. 675-683, 2004.

[6] S. I. Morefield, E. W. Keefer, K. D. Chapman, and G. W. Gross, "Drug evaluations using neuronal networks cultured on microelectrode arrays," *Biosensors and Bioelectronics*, vol. 15, pp. 383-396, 2000.

[7] Y. Jimbo, H. P. C. Ronbinson, and A. Kawana, "Strengthening of synchronized activity by tetanic stimulation in cortical cultures: Application of planar electrode arrays," *IEEE Transactions on Biomedical Engineering*, vol. 45, pp. 1297-1304, 1998.

[8] G. Shahaf and S. Marom, "Learning in networks of cortical neurons," *Journal of Neuroscience*, vol. 21, pp. 8782-8788, 2001.

[9] D. W. Branch, B. C. Wheeler, G. J. Brewer, and D. E. Leckband, "Long-term maintenance of patterns of hippocampal pyramidal cells on substrates of polyethylene glycol and microstamped polylysine," *IEEE Transactions on Biomedical Engineering*, vol. 47, pp. 290-300, 2000.

[10] Y. Nam, J. C. Chang, B. C. Wheeler, and G. J. Brewer, "Gold-coated microelectrode array with thiol linked self-assembled monolayers for engineering neuronal cultures," *IEEE Transactions on Biomedical Engineering*, vol. 51, pp. 158-165, 2004.

[11] J. C. Chang, G. J. Brewer, and B. C. Wheeler, "A modified microstamping technique enhances polylysine transfer and neuronal cell patterning," *Biomaterials*, vol. 24, pp. 2863-2870, 2003.

[12] B. C. Wheeler, J. M. Corey, G. J. Brewer, and D. W. Branch, "Microcontact printing for precise control of nerve cell growth in culture," *J Biomech Eng*, vol. 121, pp. 73-8, 1999.

[13] Y. Nam and B. C. Wheeler, "Multichannel recording and stimulation of neuronal cultures grown on microstamped poly-D-lysine," presented at The 26th Annual International Conference of the IEEE Engineering In Medicine And Biology Society, San Francisco, CA, 2004.

Yoonkey Nam (S'01) received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 1997 and the M.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, in 2003. He is currently working towards the Ph.D. degree in the Department of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign.

His current research interests include cell-electrode coupling problem in neuro-electronic interface, micropatterning, learning and memory in cultured neural networks.

Bruce C. Wheeler (S'75–M'80–SM'02) received the B.S. degree from the Massachusetts Institute of Technology, Cambridge, in 1971 and the M.S. and Ph.D. degrees in electrical engineering from Cornell University, Ithaca, NY, in 1977 and 1981, respectively. Since 1980, he has been with the University of Illinois at Urbana-Champaign, where he is now Interim Head of the Department of Bioengineering and Professor of Bioengineering, Electrical and Computer Engineering, and the Beckman Institute, and a member of the neuroscience program.

His research interests include the technology of micropatterning of neuronal networks, microelectrode arrays, and neural signal processing, all aimed at better scientific understanding of the behavior of small networks of neurons.

Moon-Hyon Nam (M'82) received the B.E. degree, the M.E. degree, and the Ph.D. degree from Yonsei University, Seoul, Korea, in 1970, 1972, and 1975, respectively. Since 1976, he has been Professor with Department of Electrical Engineering, Konkuk University, Seoul, Korea. He was a Visiting Professor in University of California at Berkeley from 1980 to 1982. He served as Dean of the Graduate School of Konkuk University (2002 – 2004).

He was the Founding President of the Korean Society of the History of Technology and Industry (1999 – 2005) and the Member of the History Committee, IEEE (2001 – 2003). He is now the Member of Culture Properties Committee of the Agency of Culture Heritage in Korea and Chairman of JAGYEONGNOO Research Institute, Inc.

His research interests include the history of science and technology, digital control, biomedical instrumentation, and physiological modeling.

A Proposed Neuroinformatics Platform for the Pupil

Shiro Usui, Kazutsuna Yamaji, Hiroyuki Sakai and Yoshihiro Okumura

Laboratory for Neuroinformatics, RIKEN Brain Science Institute,

2-1 Hirosawa, Wako, Saitama 351-0198 Japan

Abstract—**Neuroinformatics is a novel research paradigm that promotes the organization of neuroscience data and knowledge bases by means of Internet technology such as open web databases. The XooNiPs is a database management system for neuro- informatics we are developing. In this paper, we propose a neuroinformatics platform on pupil based on the XooNiPs.**

Index Terms—**Database, Neuroinformatics, Platform, Pupil**

I. Introduction

In a recent report by the Neuroinformatics Working Group of the OECD (Organization for Economic Cooperation and Development) Global Science Forum, a pressing need was expressed to utilize information technology to improve our understanding of brain mechanisms and functions [1]. This group proposed the establishment of an International Neuroinformatics Coordinating Facility to facilitate this goal.

The task of understanding any functional brain system is hindered by the inevitable need for tight focus and specialization. This fragmentation makes the synthesis and integration of disparate lines of evidence exceptionally difficult. In order to address this difficulty, an organized framework is needed to facilitate integration and to provide a fertile ground for sharing information.

Neuroinformatics is a novel research paradigm that promotes the organization of neuroscience data and knowledge bases by means of Internet technology such as open web databases (i.e., platforms). A neuroinformatics platform, at the same time, provides a rich environment for the integration of experimental evidence with mathematical models. The mathematical model serves as a framework for the comprehension of brain functions and mechanisms and also provides a common language for systematic interdisciplinary study.

In this paper, we propose the schematic design of a neuroinformatics platform for the pupil. The pupil (iris) has been the object of study in a widely variety of research areas; e.g., anatomy, physiology, ophthalmology, pharmacology, and psychology [2]. In addition, a number of researches, notably Larry Stark and his colleagues, have already developed various mathematical models for the pupil [2-9]. We anticipate that a pupil platform will integrate the vast amount of existing evidence as well as verify and improve the models. This will lead us to a better, more systematic understanding of the pupil.

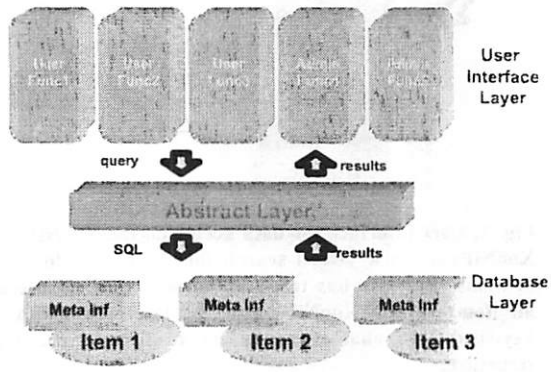


Fig. 1: Schematic diagram of the XooNiPs. The abstract layer encapsulates the database layer, and provides the application interface to the user interface layer for the secure database access. Each item is stored with its own meta-information in the database layer.

II. Neuroinformatics Base Platform: XooNiPs

A neuroinformatics platform on the pupil will be constructed using XooNiPs (Xoops based Neuroinformatics Platform System). XooNiPs is a content management system we are developing and it is composed of three functional layers as follows (see Fig. 1):

1) *Database layer*
   This layer manages all manner of digital information such as documents, experimental data, analysis tools, mathematical models, etc. All items are saved in association with their own meta-information such as item type, authors, title, keywords, comment, etc.
2) *Abstract layer*
   This layer serves as the application interface to the user interface layer for secure database access. Meta information search is also provided.
3) *User interface layer*
   This layer provides a web interface to access the items. Users can browse the items utilizing a keyword tree in

addition to a popular keyword search (Fig. 2). This layer also offers an administrator interface to set the site policy.

The functions of XooNiPs are highly modular. Therefore their combination allows us to easily customize the site policy. Furthermore, even if desired functions are absent in the current version, one can create those missing functions because XooNiPs is an open source project.

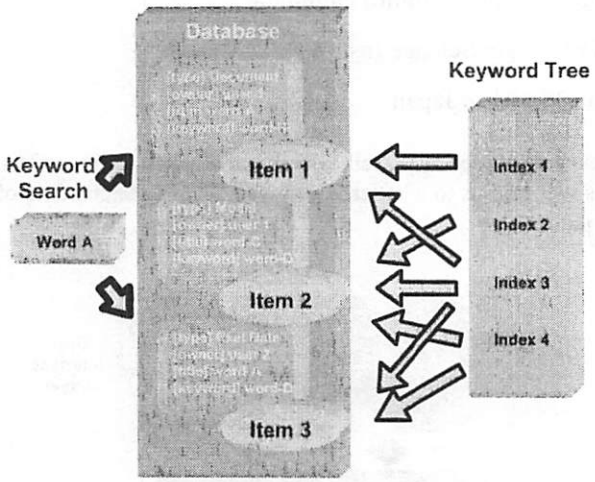


Fig. 2: User interface for data access on the XooNiPs. The XooNiPs has a keyword search for the meta-information of items, and also has tree-structured keywords. Because all items in the XooNiPs have at least one index, the keyword tree enables the users to browse the entire structure.

III. Pupil Platform (Scheme)

To construct a neuroinformatics platform for the pupil using the XooNiPs, several steps are required. The first is the development of a tree-structured keyword list. This must cover a wide range of research areas and present a systematic view of the various types of evidence. According to a precedent setting neuroinformatics project on vision science begun in Japan [10,11], the most difficult task is the creation of a keyword list representing the broad array of pupil research. In the case of the pupil, however, an amazing handbook written by Loewenfeld [2] is available. This handbook includes a considerably wide range of research topics. Thus, this handbook's table of contents can be used to supply the keyword tree for the pupil platform, at least for the initial stage.

The next step is to prepare the initial set of items. Eventually, the pupil platform will be open to the public on the Internet and research information and resources will be voluntary exchanged by an organized user community via the platform. At the beginning, however, attractive items might be necessary to draw users to the platform. Thus, we will prepare items from three different categories as follows.

1) *Literature Information*
The extensive bibliography of the handbook (over 15,000 references) will also be available as the literature component of the platform. The bibliography provides an excellent research resource because each reference is automatically associated with the keyword tree (table of contents).

2) *Mathematical Models*
Several mathematical models of the pupillary light reflex system will be downloadable as executable files which can be implemented on several different simulators. Furthermore, we plan to provide an online simulation environment for those models in collaboration with simulation servers opened in the Internet; e.g. Visiome Simulation Server [12]. Also, experimental measurements of the pupillary light reflex will be available to test model behavior.

3) *Conference Information*
The Pupil Colloquium (http://jismail.ac.uk/files/pupil/) is a meeting of researchers from different backgrounds who have a common interest on the pupil. We expect to cooperate closely with this group. For instance, people who register at the meeting will simultaneously get an account for the Pupil Platform and his/her abstract might be submitted to the platform with related research resources.

IV. Conclusion

We here propose a schematic design for a neuroinformatics platform on the pupil. The platform will be based on the content management system XooNiPs. We anticipate that the platform will promote knowledge sharing and advance systematic understanding of the pupil. Finally, it is an honor to be involved with the development of a tool that will facilitate the transfer of research efforts by pioneers such as Larry Stark and Irene Loewenfeld. This will enable future generations of researchers to better understand and appreciate their great achievements.

References

[1] S.-I. Amari, F. Beltrame, J. G. Bjaalie, T. Dalkara, E. D. Schutter, G. F. Egan, N. H. Goddard, C. Gonzalez, S. Grillner, A. Herz, K.-P. Hoffmann, I. Jaaskelainen, S. H. Koslow, S.-Y. Lee, L. Matthiessen, P. L. Miller, F. M. D. Silva, M. Novak, V. Ravindranath, R. Ritz, U. Ruotsalainen, V. Sebestra, S. Subramaniam, Y. Tang, A. W. Toga, S. Usui, J. V. Pelt, P. Verschure, D. Willshaw, and A. Wrobel, "Neuroinformatics: the integration of shared databases and tools towards integrative neuroscience," *J. Integr. Neurosci.*, 1, 2002, pp. 117-128.

[2] *The Pupil: Anatomy, Physiology, and Clinical Applications*, I. E. Loewenfeld (Ed.), Oxford: Butterworth-Heinemann, 1999.

[3] J. Semmlow and L. Stark, "Simulation of a biomechanical model of the human pupil," *Math. Biosci.*, 11, 1971, pp. 109-128.

[4] JL Semmlow and DC Chen A simulation model of the human pupil light reflex. Math. Biosci. 33:5-24, 1977.

[5] S. Usui and L. Stark, "A model for nonlinear stochastic behavior of the pupil," *Biol. Cybern.*, 45, 1982, pp. 13-21.

[6] F. Sun, W. C. Krenz, and L. Stark, "A system model for the pupil size effect," *Biol. Cybern.*, 48, 1983, pp. 101-108.

[7] W. Krenz and L. Stark, "Neuronal population model for the pupil-size effect," *Math. Biosci.*, 68, 1983, pp. 247-265.

[8] S. Usui and Y. Hirata, "Estimation of autonomic nervous activity using the inverse dynamic model of the pupil muscle plant," *Ann. Biomed. Eng.*, 23, 1995, pp. 375-387.

[9] W. Sun, F. Sun, and G. Hung, "Neural network model for pupillary responses to spatial stimuli," *Biol. Cybern.*, 79, 1998, pp. 131-138.

[10] S. Usui, "Visiome: neuroinformatics research in vision project," *Neural Networks.*, 16, 2003, pp. 1293-1300.

[11] *Visiome Platform*, Available at: http://platform.visiome.org
[12] *Visiome Simulation Sever*, Available at: http://150.12.207.51/XOOPSss/modules/visiomeSS/

Self-Striking Water Clock *Jagyeongnoo* – Early Korean Cybernetics

Nam Moon-Hyon
Konkuk University, Seoul 143-701, Korea
and JAGYEONGNOO Research Institute, Inc.
monroe@konkuk.ac.kr

Timekeeping was both a royal duty and a royal prerogative in Korea. Throughout the Joseon dynasty (1392-1910) astronomical, calendrical, meteorological, and clepsydral affairs were entrusted to the royal observatory bureau. King Taejo (r. 1392-98), the founder of the Joseon dynasty, built in central Seoul a Bell Tower which announced the time provided by a newly cast night-watch* clepsydra in 1398. The third king, Taejong (r. 1400-18), was a scientifically inclined ruler who worked with his son, the King Sejong (r. 1418-50), to improve printing technology, astronomical instruments and clock-making. Assisting them was the Chief Court Engineer Jang Yeong-Sil who cast a new night-watch clepsydra (an instrument with night-watches and [their] points)* in 1424. King Sejong became particularly interested in astronomy and, in 1432, initiated a project to re-equip the royal observatory. While the project was underway, fifteen kinds of astronomical instruments and clocks were made, including two monumental water clocks and four kinds of sundials. How the clocks were made is recorded in the *Sejong Sillok* (Factual Record of Sejong).

The *Jagyeongnoo* is an elaborate timekeeper which announced the twelve "double-hours" (two-hour blocks) that corresponded to the animals of the Chinese zodiac – the rat (which straddled midnight), ox, tiger, rabbit, dragon, snake, horse(which straddled noon), sheep, monkey, rooster, dog, and pig – as well as the five night-watches that covered the hours from 5 p.m. to 7 a.m. (time-length is seasonally variable and corresponded with the monkey, rooster, dog, pig, rat, ox, tiger, rabbit double-hours) and their five "points" automatically.* Having been completed in 1434 under the commission of King Sejong to Engineer Jang, it was assigned as national standard clock from the first day of July in 1434.

The clock is divided into three parts (Fig. 1). The inflow part of the clepsydra has four water-supplying vessels and two measuring vessels for everyday changeover. A floating indicator rod, graduated for the twelve double-hours and the points of the five night-watches, was in the measuring vessels, while above these vessels were wooden vertical ball racks with twelve (for the double-hours) and twenty five (for the five points of the five night-watches) small bronze balls. The rod would rise and release horizontal latches that let the

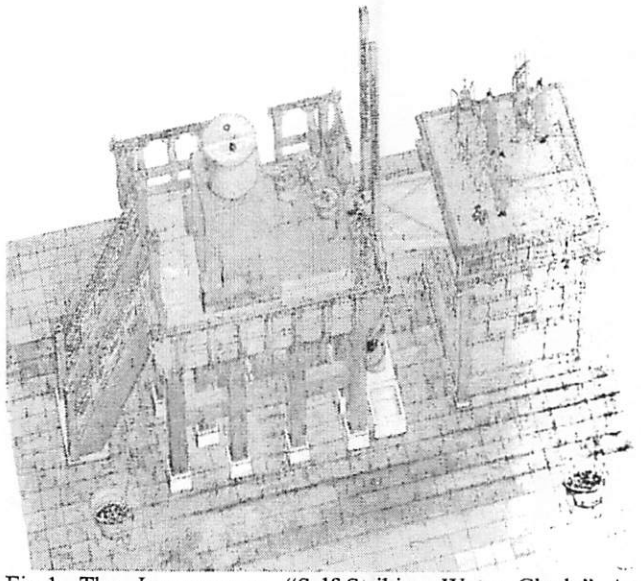


Fig.1. The *Jagyeongnoo*, "Self-Striking Water Clock." An artist's concept, after a design made by the author. Left: The inflow part of the clepsydra consists of a reservoir, compensator, regulator, and overflow receptacle; two measuring vessels are on the floor bed. Middle: Two ball racks are implanted on the measuring vessels, which release small bronze balls through the holes that correspond to the twelve double-hours and the five night-watches. Right: Devices to announce the time are in the box, and the three jacks (biological automata) on the box strike a bell, drum, or gong; the jack in the middle of the box displays the time by showing the double-hour tablet through the window as the bell is stricken. Twelve-zodiac wooden animals show twelve double-hours tablet in turn.

bronze balls out; the balls then fell and ran along an inclined channel to cause a set of twelve and twenty five iron balls about as big as a hen's egg to trigger the time-announcing mechanisms. These mechanisms consists of two ball-relays and an audible time signal for the double-hours and the night-watches, a visible time-indicator for the double-hours, three wooden jacks which sounded a bell for the double-hours, a

drum for the night-watches, and a gong for the night-watch points.

The ball-relaying mechanisms were separated by two large channels for the balls to move along. One channel allowed the balls to cause a jack in the form of a god to ring its bell for the double-hours; it also caused a horizontal wheel with twelve small gods to turn so that the appropriate god would display its double-hour tablet, an example having the character午 for the hours of the horse (11 a.m. to 1 p.m.). The second channel caused another jack in the form of a god to announce the night-watches by striking a drum the appropriate number of times (e.g. three for the third night-watch), while a third god sounded a gong once for the beginning of a night-watch as well as the appropriate number for each night-watch point (e.g. five strikes of the gong for a fifth point). Striking mechanisms are mechanical digital counters controlling time-announcing biological automata.

During the Japanese invasion of Korea in 1592, this clepsydra was destroyed. Presently a reconstruction project for the *Jagyeongnoo* is undergoing by the author for the National Royal Museum of Korea.

In 1536, by order of King Jungjong (r. 1506-44), the horologist Park Se-Ryong made a new self-striking clepsydra similar to the King Sejong. The parts which announced time have been dismantled and three water supplying vessels and two measuring vessels including a indicating rod with floats were collectively entered as "National Treasure Number 229 Jagyeongnoo of the Borugak".

NOTE

* "Night-watch" and "night-watch point" are technical terms from ancient Chinese timekeeping. The expression "night-watch" is used for the Korean word *gyeong* (更, *geng* in Chinese), a two-hour period in the nighttime. There were five *gyeong* each night, corresponding with the two-hour blocks or "double-hours" from 5 p.m. to 7 a.m., and each *gyeong* was divided into five *jeom* (點, *dian* in Chinese), an expression translated as (night-watch) "point."

REFERENCES

Nam, Moon-Hyon. Jang Yeong-Shil and Jagyeongnoo – Reconstruction of Time Measuring History of Joseon Period. Seoul National University Press, 2002. (in Korean)

Needham, Joseph, Lu G.-D., Combridge, John H., and Major, John S. *The Hall of Heavenly Records: Korean Astronomical Instruments and Clocks, 1380-1780*. Cambridge: Cambridge University Press, 1986.

Can signals issued from the vestibular output stabilize the forelimb in space to compensate for body movements as it stabilizes gaze?

Gabriel M. Gauthier, Jean Blouin, Christophe Bourdin and Jean-Louis Vercher

UMR CNRS/Université de la Méditerranée "Mouvement et Perception", Faculté des Sciences du Sport
163 Avenue de Luminy, 13288 Marseille, cedex 9, France
gauthier@laps.univ-mrs.fr

Abstract—**An attempt is made to demonstrate that when the human body moves around, the brain has the option of using signals originating in the vestibular system to stabilize the hand in space to "automatically" compensate for body movements. Such behavior is extremely useful in situations such as picking and placing objects while the body is in motion. Besides, vestibular control can develop in the absence of vision and does not require cognitive contribution. A parallel is proposed between the vestibulo-ocular reflex system which is designed to stabilize gaze and the vestibulo-manual system in its contribution to stabilization of the hand in space while the body moves. Both systems have short latency, high velocity response and show similar adaptive ability when their input-output is visually altered. We suggest that a model describing eye and hand stabilization in space through the vestibular system might become extremely useful to engineers facing the problem of independently or conjointly controlling arm-held tools and visual monitoring devices as the robot main body moves around.**

Index terms—**Adaptive control, hand stabilization in space, vestibulo-manual system, vestibulo-ocular reflex.**

I. INTRODUCTION

The idea about this study originated in Marseilles about 10 years ago, in the course of a discussion led by Lawrence Stark between colleagues from the Sensorimotor Lab and from the Robotics Lab headed by Jean-Claude Bertrand. Roboticians were pointing out the easiness with which the position of a robot arm can be controlled in space when the robot base is immobile as compared to the troublesome problem of maintaining the robot arm still in space while moving the robot body around. A striking observation of the ability of Man to stabilize a hand-held tray in space as the body rotates around a vertical axis suggests that decrypting the way the brain accomplishes such prowess might be useful as a model for robot hand stabilization control. The way the brain stabilizes gaze through the activation of the vestibulo-ocular reflex (VOR) has for long inspired roboticians to solve the problem of controlling for example the position of a camera in space while the robot moves around. In fact, the VOR with its apparent simplicity, its basic neuronal network being made of three neurons, was also used as a model for motor control learning.

We have investigated hand stabilization in space during body movement with in mind the question : Can a parallel be made between gaze stabilization and hand stabilization in space? Indeed, both functions share equivalent purposes, which are to prevent movement of a motor structure with respect to earth as the body evolves. Hence, the experiments we carried out were designed to demonstrate that similar processes, actuated by visual and vestibular information, control both functions.

Stabilization of the eyes in space

During body movement, the visuo-vestibulo-ocular system stabilizes the eyes in space by means of the VOR, thus allowing appropriate retinal discrimination. This reflex, which may operate in total darkness is based on the activation of a three-neuron path. Head rotation is sensed by the semi-circular canals of the labyrinths and further central processing produces an activation of the ocular muscles which finally drives the eyes. If the reflex input-output relationship is well set, an eye rotation with a very short latency is produced in the direction opposite to the head movement and with the same amplitude, so that the fixated visual target remains stationary on the fovea. With an overall reflex latency of the order of 15-20 ms, and a velocity and acceleration of up to 300°/s and 5000°/s/s, respectively, the VOR is the most reactive sensorimotor system in man. The VOR is used daily to stabilize our eyes when gaze shifts are produced by a combination of eye and head movements (Fig. 1). For example, if an observer is fixating straight-ahead, the occurrence of a target of interest at 20° to the right will trigger a fast eye rotation (saccade) after about 150 ms and slightly later the head will start its rotation in the

Acknowledgment : This work was partly supported by grants from CNRS and MAIF Fondation

Corresponding author : Gabriel M. Gauthier, UMR CNRS/Université de la Méditerranée "Mouvement et Perception", Faculté des Sciences du Sport, 163 Avenue de Luminy, 13288 Marseille, cedex 9, France. (E-mail : gauthier@laps.univ-mrs.fr

direction of the target. Mainly because of their smaller inertia, the eyes 'reach' the target while the head is still moving. Therefore, to keep gaze on the target while the head rotates, the eyes must counter-rotate in their orbits to compensate for the head rotation. This counter-rotation is driven by the fast-acting VOR. The eye-head coordination process described above is the most commonly observed but as described by [8], at least three other types of coordination can be observed depending on the task and the instructions provided to the observer. Complex vestibulo-ocular interactions also develop when the visual field is restricted as in progressive lens wearers [7].

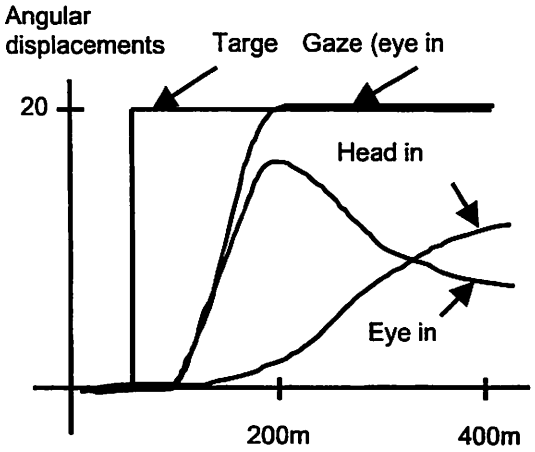


Figure 1. Vestibulo-ocular reflex. Typical ocular and gaze saccades. Target, eye and head signals in response to the attempt by the observer to shift gaze (visual interest) from straight ahead to a target situated 20° to the right. The VOR inducing a counter-rotation of the eyes in the orbit contributes to the perfect stability of gaze on the reached target while the head terminates its rotation.

VOR adaptation

Because the VOR is open-loop (the output of the system, that is the eye movement, has no direct effect on the vestibular input signal), a good stability of the eyes on a stationary target during head rotation can only be achieved if the reflex gain (i.e. eye movement/head movement) is close to unity. Nature has provided the VOR with an adaptation control that compensates for internal and external changes.

In 1976, Gonshor and Melvill-Jones [4] showed that the VOR will adaptively change to progressively produce an appropriate response in an observer fitted with reversing prisms. One year earlier, we had shown similar adaptive changes in an observer fitted with X2 magnifying lenses [2]. With such an optical device, for gaze to be stable when the head rotates, the visuo-vestibulo-ocular system must produce eye movements in the direction opposite to the head movement, twice as large as the normal response. After four days of adaptive exposure the gain reached 1.65 (Fig. 2). Concomitantly, the object size and depth perception, as well as the visual world stability perception, considerably affected at the early stage of the experiment, were recovering. Other types of adaptive processes have been observed in the visuo-vestibulo cephalic system such as the one which develops fairly rapidly in response to changes the dynamics of head [3].

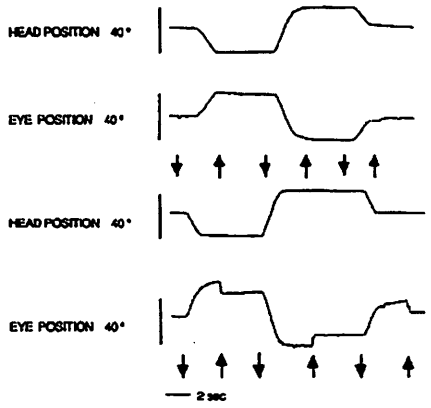


Figure 2. Adaptation of the VOR gain in response to the use of X2 lenses. Before adaptive exposure, the gain of the VOR in response to rotations about the vertical axis while the observer fixated a remembered target stationary with respect to earth was close to 1 (top recordings) and reached 1.35 after 4 days of permanent wearing of the lenses. The arrows indicate the time the fixation target was turned off (head down arrows) and on (head up arrows) just before and after the rotations, respectively. The amplitude of the saccades following target onset, at the end of the rotations, is an estimate of the gain error with respect to the appropriate unity value. (from Gauthier and Robinson, 1975)

Hand stabilization in space : The visuo-vestibulo-manual system and vestibulo-manual-like reflex

Involvement of the upper limbs in object grasping handling and manipulating is the appanage of higher primates and man. We have tested the contribution of vestibular signals in stabilizing the hand during body motion [1]. More specifically, we evaluated subjects' capacity to correct unseen arm movements when they were submitted to whole-body rotations.

II. METHODS AND RESULTS

Subjects were seated in total darkness on a motorized platform made to rotate around a vertical axis (Figure 3). Subjects were required to reach for a memorized visual target presented 20 degrees on their right. At the onset of their arm moments, subjects were rotated to either direction, with various amplitudes (10°, 20°, 30° and 40°). Mean hand trajectories produced by one subject for trials without rotation, with leftward and rightward 40° rotations are shown in figure 4. Hand trajectories in space remained almost unaffected by body rotations. The preserved spatial constancy in hand trajectory resulted from egocentric hand paths that were markedly different depending on rotation direction. Hand trajectory regulation was smooth and continuous all along the movement.

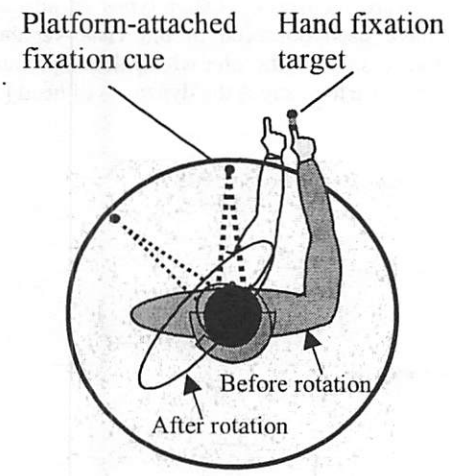


Figure 3. Experimental design to evaluate hand movement precision in hand stabilization during body rotation. Before rotation the observer positioned his index finger on a target while fixating a cue attached to the platform, while during rotation he attempts to keep his hand still in space.

With the following experiment, we tested the adaptive capacity of the vestibulo-manual system. More specifically, we tested subjects' ability to adapt to a new relationship between the arm and the body in subjects attempting to keep the extended arm still with respect to earth during body rotations. For this, we used the same experimental set-up as that described above, coupled to a learning paradigm derived from typical infographic or prism adaptation paradigms [5,6]. A laser projected a spot of red light on a horizontal board placed in front of the subject. The spot position was the position around which the subject was to stabilize his index finger. A LED was fixed at the tip of the finger. Because the spot position was servocontrolled by the chair angular position by means of a mirror galvanometer, platform rotation smoothly displaced the laser dot in the same direction and by a rotation-related amplitude. At the end of the rotation, the laser dot was shifted by 4.4°, 7.8° and 12° for 20°, 30° and 40° body rotations, respectively. Thus, to keep the finger aligned with the laser spot during body rotations, subjects had to move the arm in the direction opposite to the body but with a smaller magnitude. The subjects were not informed that the spot might move in space (in fact the subjects, when questioned after the experiment, never reported to have experienced spot displacement). Subjects were instructed to gaze at a straight-ahead chair-fixed light during body rotations. The experimental design consisted of three phases: (i) a pre-test session in which subjects' task was to keep the unseen extended arm stationary during leftward or backward 40°-body rotations (ii) the adaptation exposure described above and (iii) a post-test session in which the effect of the adaptation procedure on the arm movement stability was assessed in the same conditions as in the pre-test.

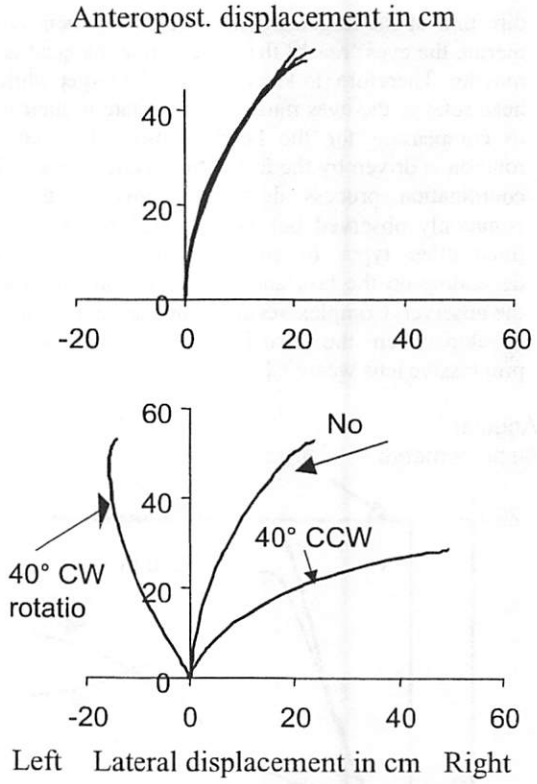


Figure 4 . Hand reaching trajectories to a visual memorized target. The plots show trajectories in space (above) and with respect to body (below) while the subject is either immobile or rotated 40° CW or CCW around the vertical axis, during the reaching movements.

Before adaptation, while the subject was in total darkness, the hand stabilogram (hand movement in space coordinates while the chair rotates) covered a small area. After adaptation, the hand stabilogram was markedly elongated in the predicted direction (Figure 5). The hand moved in space about 75% of the adaptive imposed displacement meaning that the hand stabilization system has changed its calibration, very much the same way the visual ocular relationship evolves when wearing magnifying lenses or prisms

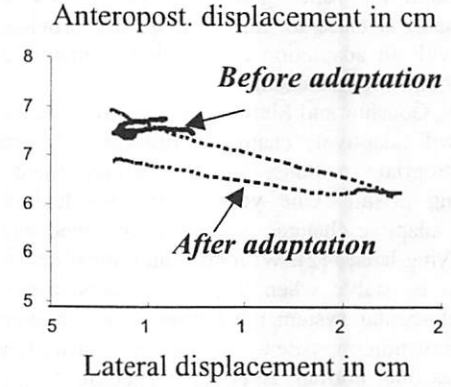


Figure 5 : Hand stabilograms (hand displacement as seen from above) in a subject before and after a period of adaptation to a rightward displacement of the target used as a cue to stabilize the hand in space during 40° rotations

Pointing trials during body rotation carried out after adaptation also demonstrated a change in the relationship between vestibular inputs and compensatory arm movements. The smooth correction of the arm trajectory during body rotation could therefore be based on a mere counteraction to the detected body rotation where subjects integrate the extent of self-motion and perform a continuous cancellation of rotation-induced effects on motor output.

III. DISCUSSION

The various oculomotor system outputs (saccades, smooth pursuit, VOR, OKN) serve vision. Dynamic analyses have demonstrated that the ocular motor system allowing vision to operate when the body is in motion is predominantly under vestibular control. Eye displacements likely to be induced by high body accelerations are efficiently compensated for, in a fairly automatic manner, via the VOR. Conversely, this reflex is unable to react to very low acceleration and to constant velocity displacements. Still, in the latter conditions, eye stabilization is achieved through visual feedback activation of the tracking systems. A large dynamical spectrum of motions is thus taken care of and vision is allowed to operate from high to low velocity body movements. In everyday life, most of the work is done in a fairly automatic manner by the vestibular system via the VOR (the VOR is activated some 10 000 to 40 000 times a day). Vision, even if not used in stabilizing processes *per se*, provides signals for higher control including adaptation to maintain in a long term, appropriate eye to head movement relationship. Obviously, and necessarily, the VOR is inhibited, suppressed or compensated for when its action would interact the ongoing eye movements such as in target tracking. If one considers the upper limb movement system in its function devoted to stabilization of the hand in space, one may establish a parallel with the ocular motor system in its gaze stabilization function. Although the oculo-manual system is essentially under visual control, in everyday life, we execute numerous actions without visual control of the arm and hand. Several experiments have demonstrated that the visual manual system is doted of an adaptive control to maintain a precise relationship between visual and motor spaces, -and incidentally to allow us to use corrective spectacles, microscopes, telemanipulators, etc. Most of the knowledge regarding visuo-manual relationship adaptation comes from experiments involving alteration of the relationship with optical prisms, currently referred to as the prismatic adaptation paradigm. One may assume that similar processes handle vestibulo-manual relationship integrity. Still, this has to be demonstrated along with the role of visual information (hand to space instability) as a teaching cue to activate the plastic changes. Further study must be carried out to better understand the vestibular involvement in stabilizing the hand in space during body movement and to describe the neural network sustaining the function. Still, the vestibular signal seems to be an ideal candidate to be at the origin of this function. What would be the other candidates? The arm stretch reflex would act, if it did, against stabilization in space. Indeed, this reflex tends to keep the hand stable with respect to the trunk, the overall result of which being a rotation of the arm in space. It follows that hand stabilization in space requires the stretch reflex to be turned off during rotation. Conversely, arm inertial forces act towards stabilization in space, and the fact the vestibulo-manual system is doted of an adaptive control tends to argue that hand stabilization in space during body rotation derives from "controlled" rather than "passive" processes.

Automatic-like control through vestibular information make the systems controlling the eyes and the hand less dependent from cognitive concern, but still fairly efficient in terms of reaction time and precision. It is tempting to suggest that these two systems may be integrated as a functional unit serving the purpose of manipulation under direct visual control in which body movements are automatically compensated by vestibule-issued signals at both hand and eye levels via the vestibulo-manual like reflex and the vestibulo-ocular reflex, respectively. A model describing eye and hand stabilization in space through the vestibular system might become extremely useful to engineers facing the problem of independently or conjointly controlling the arm tool and the visual monitoring device when the robot main body moves around.

References

[1] Bresciani, J-P., Blouin J., Sarlegna F., Bourdin C., Vercher J-L. and Gauthier G. M. (2002) On-line versus off-line vestibularly-evoked control of goal-directed arm movement. *Neuroreport.* 13: 1563-1566.

[2] Gauthier, G. M., and Robinson, D. A., (1975) Adaptation of the human vestibuloocular reflex to magnifying lenses. *Brain Res.*, 92: 331-335.

[3] Gauthier, G.M., Martin, B.J. and Stark, L. W. (1986) Adapted head- and head-eye movement responses to added-head inertia. *Aviat. Space Environ. Med.* 57 :336-342.,

[4] Gonshor, A., and Melvill-Jones, G., (1976) Extreme vestibulo-ocular reflex adaptation induced by prolonged optical reversal of vision. *J. Physiol. (London).* 256: 381-414.

[5] Ingram, H., van Donkelaar, P., Cole, J., Vercher, J.L., and Gauthier G.M. (2001) The role of proprioception and attention in a visuo-motor adaptation task. *Exp. Brain Res.*, 132 : 114-126.

[6] Roby-Brami, A. and Burnod, Y. (1995) Learning a new visuo-motor transformation: error correction and generalization. Cognitive Brain Res., 2: 229-242.

[7] Semmlow, J. L., Gauthier, G.M. and Vercher, J.L. (1990) Identification of peripheral visual images in a laterally restricted gaze field. *Perceptual and Motor Skills,* 70:175-194.

[8] Zangemeister, W.H. and Stark, L. W. (1989) Types of gaze movement : variable interactions of eye and head movements. *Exp. Neurol.* 77:563-577.

Models for Neurovestibular Adaptation

Lawrence R. Young
Man Vehicle Laboratory Massachusetts Institute of Technology,
Cambridge, MA USA.

Abstract: Dynamic models have played a more prominent role in the vestibular and oculomotor field than in any other branch of physiology. The ease of identification of input and output variables and the challenge of multi-loop, multi-axis adaptive control has attracted numerous modelers from engineering and shaped behavioral and neurophysiological experimental programs. In particular, the adaptive characteristics of the neurovestibular system and generated continuing speculation as to mechanisms. This treatment of adaptation and multi-sensor integration covers the development and application of such models, principally in the author=s laboratory. It emphasizes the continuing relevance of both Amodel reference@and Aerror pattern recognition@notions of adaptive control.

Keywords: distributor, oculomotror, adaptive, model.

Independent Component Analysis of Divergence Eye Movements

John L. Semmlow[2] and Tara L. Alvarez[1]
1 Department of Biomedical Engineering. New Jersey Institute of Technology, Newark, NJ
2 Dept. of Biomedical Engineering, Rutgers University and Depart of Surgery, Bioengineering, Robert Wood Johnson, Medical School- UMDNJ New Brunswick, NJ

Abstract- **A new application of Independent Component Analysis (ICA) can be used to isolate motor components from a coordinated response as long as multiple responses can be obtained to the same stimulus. This ensemble ICA is well suited to analyze eye movement behavior as these motor responses are mediated by a number of independent control processes. Here we apply ensemble ICA to vergence eye movements: the inward (convergence) or outward (divergence) turning of the eyes which allows us to view images at various distances. Previous studies on the dynamics of convergence and divergence eye movements have produced varied, sometimes contradictory, results. Subjects tracked four degree disparity step changes for convergence and divergence at different initial target positions. Results show that the dynamics of divergence movements not only differ from convergence movements, but depend on the initial vergence position. Unlike convergence movement whose dynamics are insensitive to initial position, the velocities of divergence eye movements in response to targets that were initially near to the subject were approximately twice that of responses to initially distant targets. To determine the underlying cause of this behavior, ensemble ICA was applied to selected divergence responses. Preliminary data show the divergence system contains a step and pulse component and the amplitude of the pulse is dependent on the initial stimulus position.**

I. INTRODUCTION

In the late 1950's, Lawrence Stark proposed a general strategy for utilizing multiple input and output channels to gain information about internal neural configurations, a strategy he termed "dry dissection." This approach can be used as long as the input channels can be independently stimulated and/or the output channels can be separately monitored. The dry dissection approach has been particularly successful in isolating the various overlaying neural control components that control human eye movements.

The disparity vergence system is one of five distinct control processes responsible for guiding or positioning (i.e., fixating) the eyes. Other eye movement systems include the saccadic, smooth pursuit, vestibular, and accommodative vergence. The disparity vergence system mediates inward and outward turning movements that are used rotate the eye until paired images project onto the foveas. By using only binocularly symmetrical stimuli, the motor responses of this system can be selectively isolated. However, classical dry dissection techniques are not able to isolate vergence system subcomponents that are thought to exist.

Based on significant behavioral data, the vergence system is thought to be composed of two control subcomponents The control structure of the disparity vergence system has been described by the "Dual Mode Theory" which states that the vergence response is guided by an open loop, feedforward or transient component which generates a pulse-like neural signal along with a feedback or sustained component that generates a step-like neural signal. The feedfoward component improves the motor response dynamics while the feedback component is responsible the accuracy of the final position. An ensemble application of principle component analysis (PCA) has confirmed that two components constitute the vast majority of the variability found in a vergence eye movement implying that two systems are responsible for the combined response.

In humans, it is not possible to record the component signals directly from neurons. To isolate the components of the disparity vergence system, the methods of dry dissection need to be extended to apply to single channel responses produced by single inputs. If multiple observations can be produced from the same stimulus, a modification of independent component analysis (ICA) can be used in lieu of separate output channels. Essentially, each component contributes to the response-to-response variability and ICA can resolve the components based on their variability.

Controversy exists regarding the relative dynamics of convergence and divergence. Several studies report that convergence is faster than divergence [1][2][3][4] by as much as double [4], while other studies report pure divergence and convergence to have approximately the same velocity characteristics [5]. He we show that the dynamics of divergence vary as a function of initial position. Depending on the initial location of the stimulus, the relationship between convergence and divergence dynamic properties can vary dramatically. Applying ICA to both divergent responses, we show that a pulse component is present in divergence movements, but the amplitude of this pulse is dependent on the initial position.

II. METHODOLOGY

Four subjects (18 to 60 years old) participated in this study. Two subjects were male, and two were female. All subjects signed informed consent forms before the experiments that were approved by the New Jersey Institute of Technology (NJIT) Institutional Review Board (IRB). Head position was stabilized using a custom chin rest to avoid any influence from the vestibular system. All subjects were able to perform the task easily. One subject was highly trained and aware of the goals of this study while the other three subjects were inexperienced and naïve to the goals of the study.

Experimental Design

Disparity vergence stimuli were presented using a dynamic haploscope constructed from two computer monitors that generated a symmetrical stimulus of paired vertical lines. Since the nonlinearities of our system were small (a maximum of 5% [7], we used a two-point calibration before and after each individual response. Only the targets produced by the stimulus displays were seen by the subject during the experiment, and no proximal cues associated with depth information related to the target distance were present. [8]

During an experimental session, a variety of convergent or divergent stimuli were presented: all 4 deg. step changes. For the divergence experiments, stimuli began at initial vergence positions of 20, 16, 12, and 8 deg. The four stimuli were randomly presented after a random delay of 0.5 to 2.0 seconds to avoid subject prediction which can alter vergence dynamics. [9][10] Convergence experiments also had four initial positions: 16, 12, 8 and 4 deg. (The 20 deg. divergent and 16 deg. convergent stimulus was omitted in one subject that could not fuse that close a target.)

Eye movements were recorded using an infrared limbus tracking system (λ = 950 nm) manufactured by Skalar Iris (model 6500). The manufacturer reports a resolution of 2 minutes of arc. All eye movements were well within the system's ± 25 degree linear range. Data acquisition was done at a sampling rate of 200 Hz, which is well above the Nyquist frequency for vergence eye movements.

Data Analysis

The main objective of this analysis was to investigate the dynamic change in the responses, so movements that contained small saccades during the final, steady-state portion were analyzed as long as the saccades in the two eyes cancelled in the net vergence response. The left and right eye responses were subtracted to yield the net disparity vergence movement. When displayed graphically, convergence was plotted as positive, and divergence was plotted as negative. The velocity response was computed using a two-point central difference algorithm [12]. Data were analyzed, in part, by measuring the magnitude of the peak velocity. Individual subject data were compared using an unpaired student t-test to determine if the dynamic changes were statistically significant.

ICA can be applied to solve the "neural cocktail party" problem blindly separating an individual neuronal signal from a composite assuming the sources are statistically independent sources. ICA solves the mixed source separation problem "blindly" in that it separates the signal sources without making any prior assumptions [14]. The ICA model is a generative model: it attempts to explain how the sources, in this case the underlying neural subcomponents, are mixed to generate the observed signals based on the linear mixing model: [14][15][16]

$$\mathbf{X} = \mathbf{As}$$

where **X** and **s** are vectors representing the signals and sources, respectively. In this application, the signals, **X**, are the recorded eye movements, and the sources, **s**, are the underlying transient and sustained components. The mixing matrix, **A**, consists of the component mixture and movement-to-movement variability.

This analysis used the "Fast ICA" program developed by the ICA group at Helsinki University. Figure 1 shows ensemble divergence data on the left and two components generated by the ICA algorithm depicting the underlying transient and sustained components of divergence.

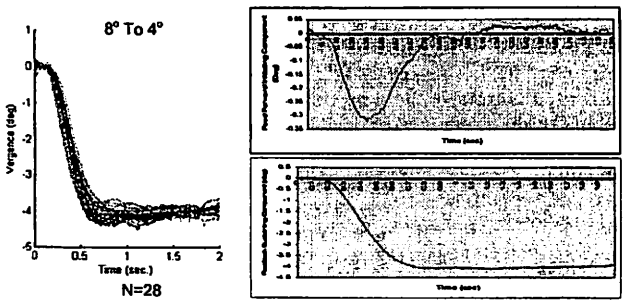


Figure 1: Raw divergence eye movements (left) and the two components found by the ICA algorithm (right)

III. RESULTS

Typical individual divergence eye movements are shown for each subject in Figure 2. The upper traces are velocity and the lower traces are vergence position. The near and far responses exhibit clear differences. The dashed line shows responses that began near to the subject while the solid line shows responses that had far initial positions. Qualitatively, the divergence responses to stimuli near to the subject occurred earlier and had a greater peak velocity than responses to stimuli far from the subject depicted in the solid red line. Conversely, qualitative inspection of convergence eye movements does not show a consistent change in dynamic properties with initial position, Figure 3.

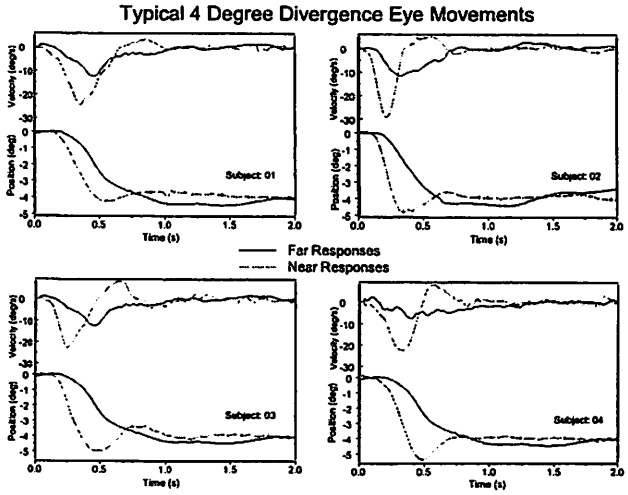


Figure 2: Divergence Responses for Stimuli near (dashed green line) and far (solid red line) from subject

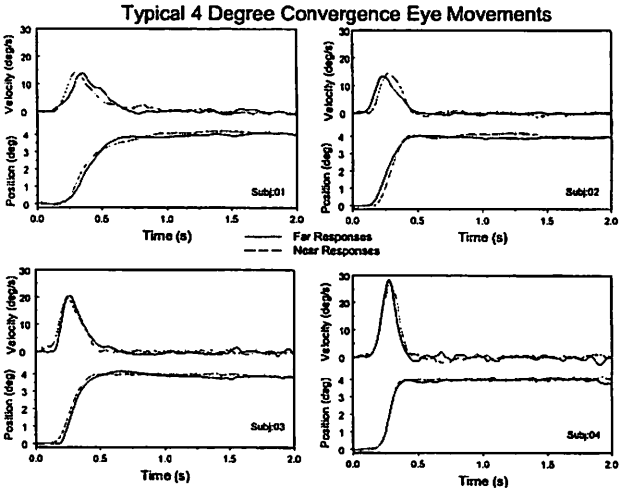


Figure 3: Convergence Responses for Stimuli near (dashed green line) and far (solid red line) from subject

Response dynamics were quantified using peak velocity (Table 1). For divergence movements to the same four degree step change in disparity, the magnitude of peak velocities for the near responses (initial position of 20 or 18 degrees) were approximately twice that of the far responses (initial position of 8 degrees). When comparing the responses to near and far stimuli, the dynamic differences were statistically significant ($P < 0.0001$) for all four subjects individually and combined. Conversely, convergence dynamics showed no consistent trend for the four subjects when combined. Subjects 002 and 003 did show some difference between near and far four degree responses ($P < 0.01$), but not when combined with all subjects ($P < 0.2$).

IV. DISCUSSION

Since divergence dynamics are dependent on initial position and convergence dynamics are not, any comparison between the two will depend on initial condition. This may account for the differences in convergence versus divergence velocities reported by different groups. For example, convergence and divergence peak velocities of subject 003 were approximately the same when the responses were between 16 and 20 degrees: 19.4 deg/sec for divergence and 18.1deg/sec for convergence. This is similar to the findings of Collewijn and colleagues [5]. This group did not report on the initial stimulus positions used in there experiments. When the responses ranged between 4 to 8 degrees, this same subject showed convergence dynamics that were more than twice as fast as divergence: 8.1 deg/sec for divergence and 20.9 deg/sec for convergence. This finding would be more along the lines reported by Hung and colleagues [1]. This group reported that their responses were recorded at an intermediate initial starting position of 8 degrees. Since the divergence system is dependent on the initial position of the stimulus, any comparison between convergence and divergence must take into account the stimulus range.

Divergence				
Stimulus (Deg)	Velocity (Deg/Sec) ± Standard Deviation			
	Subj: 001	Subject: 002	Subject: 003	Subject: 004
20 to 16	17.6±3.1 * N = 19	20.5 ± 5.5 N = 18	19.4 ± 4.6 N = 11	21.1 ± 4.8 N = 29
16 to 12	14.5 ± 2.8 N = 18	18.2 ± 5.9 N = 19	18.6 ± 6.0 N = 20	16.3 ± 2.7 N = 36
12 to 8	12.5 ± 4.0 N = 17	12.1 ± 2.0 N = 13	12.1 ± 2.7 N = 11	11.6 ± 2.6 N = 21
8 to 4	12.2 ± 1.4 N = 10	10.3 ± 2.3 N = 25	8.1 ± 1.4 N = 19	11.3 ± 2.4 N = 50
Convergence				
Stimulus (Deg)	Velocity (Deg/Sec) ± Standard Deviation			
	Subject: 001	Subject: 002	Subject: 003	Subject: 004
16 to 20	-----	7.5 ± 1.2 N = 18	18.1± 4.9 N = 38	31.2± 6.5 N = 22
12 to 16	9.1 ± 3.6 N = 13	11.3± 2.4 N = 40	18.3± 4.7 N = 35	31.4± 5.1 N = 24
8 to 12	9.8 ± 3.4 N = 35	12.6± 3.2 N = 14	18.5± 5.0 N = 31	33.6± 4.3 N = 20
4 to 8	11.4 ± 3.9 N = 50	11.8± 2.0 N = 16	20.8± 3.8 N= 35	34.8± 6.1 N = 14

Table 1: Velocity ± Standard Deviation for Divergence and Convergence Responses for Different Initial Stimulus Conditions

The position dependent dynamics could be due to nonlinearities in the neuromuscular plant on in the control processes. The fact that position dependence appear to occur only in divergence movement suggests the mechanism is of control origin, although this is not definitive since a 20 to 16 deg. divergence movements might be more influenced by plant nonlinearities than a 16 to 20 deg convergence movement which actually begins at a more distant initial condition.

A control-based explanation is supported by some primate neurophysiologic studies. These studies have shown that different cells exist for convergence and divergence movements [17][18] and that convergence cells are more prevalent than divergence cells in the midbrain specifically within the mesencephalic reticular formation [17]. Thus, divergence is not simply a negative convergence movement but a separate neurophysiological system. The neuro-control strategy of the two systems may be different. Evidence of burst and burst-tonic cells show that combinations of a transient and a sustained neural signals are used for both convergence and divergence movements. The divergence cells may fire with more synchronization for near stimuli creating a pulse with a greater magnitude and stronger kinematics compared to divergence responses to stimuli further from the subject. It is also possible that the divergence cell pool may be dependent on the initial position and thus different cells are responsible for stimuli at different distances from the subject.

The Dual Mode Theory was developed using mostly convergence data. Model simulations indicate that convergence is composed of a pulse-step control structure similar to that found in the saccadic system [3][6][7]. However, fewer studies exist to describe the control structure for divergence. Horng [7] postulates that divergence is composed mainly of a step response. He further speculates that if a pulse component is present, it would have to be

smaller than the pulse found in convergence. He does not report the initial vergence position of the stimuli.

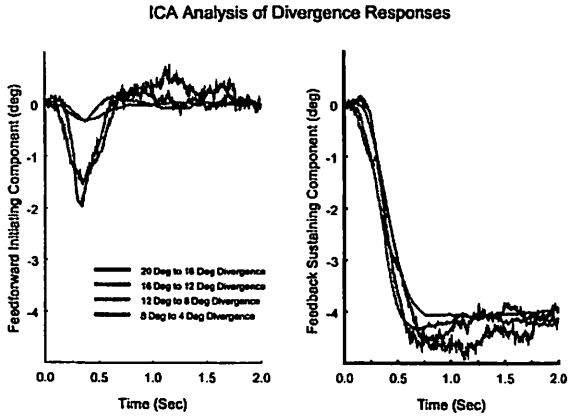


Figure 4: ICA of Subj 004 divergence responses were preprogrammed pulse (left) decreases in amplitude as the stimulus is place further from the subject. The feedback controlled step (right) does not appear to be dependent on initial stimulus position.

Our group has begun to apply ensemble ICA to the responses collected at different initial positions and the result form one subject is shown in Figure 4. These preliminary results suggest that the transient component is present during divergence responses stimulated by near targets, but the amplitude of this component decreases as the stimulus is moved further from the subject. For responses that occur far from the subject the pulse component becomes small to almost neglible. These preliminary findings agree with Horng's speculations about the control structure of divergence, but only for more distant stimuli. Further study is needed to determine if this trend is consistent in multiple subjects.

V. CONCLUSION

The dynamics of divergence eye movements are dependent on the initial stimulus position. The closer the stimulus is to the subject, the faster the responses as quantified by the maximum velocity. Convergence is not dependent on the initial stimulus position. An independent component analysis revealed that the preprogrammed pulse of divergence maybe responsible for difference in divergence dynamics where the magnitude of the pulse or the fusion initiating component is dependent on the initial stimulus position.

VI. ACKNOWLEDGEMENT

This work was funded in part by Essilor International.

VII. REFERENCES

[1] G. K. Hung, H. Zhu, and K. J. Ciuffreda, "Convergence and divergence exhibit different response characteristics to symmetric stimuli," *Vision Research*, vol. 37, pp. 1197-1205, 1997.

[2] D. S. Zee, E. J. Fitzgibbon, and L. M. Optican, "Saccade-vergence interactions in humans," *Journal of Neurophysiology*, vol. 68, pp. 1624-1641, 1992.

[3] G. K. Hung, J. L. Semmlow, and K. J. Ciuffreda, "A dual-mode dynamic model of the vergence eye movement system." *IEEE Trans on Biomedical Engineering*, vol. 33, pp. 1021-1028, 1986.

[4] G. K. Hung, K. J. Ciuffreda, J. L. Semmlow, and J., -L. Horng, "Vergence eye movements under natural viewing conditions," *Invest Ophthalmology and Visual Science*, vol. 35, pp. 3486-3492, 1994.

[5] H. Collewijn, C. J. Erkelens, and R. M. Steinman, "Voluntary binocular gaze-shifts in the plane of regard: Dynamics of version and vergence," *Vision Research*, vol. 35, pp. 3335-3358, 1995.

[6] J. -L. Horng, *Dynamic Model of Vergence Eye Movements*, Ph.D. Dissertation Rutgers University, pp. 123-132, 1994.

[7] J. -L. Horng, J. L. Semmlow, G. K. Hung, and K. J. Ciuffreda, "Dynamic asymmetric in disparity convergence eye movements," *Vision Research*, vol. 38, pp. 2761-2768, 1998.

[8] M. Rosenfield, and K. J. Ciuffreda, "The effect of surround propinquity on the open-loop accommodative response," *Investigative Ophthalmology & Visual Science*, vol. 32, pp. 142-147, 1991.

[9] W. Yuan, J. L. Semmlow, and P. Munoz, "Effects of prediction on timing and dynamics of vergence eye movements," *Ophthalmic Physiological Optics*, vol. 20, pp. 298-305, 2000.

[10] T. L. Alvarez, J. L. Semmlow, W. Yuan, and P. Munoz, "Comparison of disparity vergence system responses to predict and non-predictable stimulations," *Curr. Pscyh. Cog.*, vol. 21, pp. 243-261, 2002.

[11] G. K. Hung, "Saccade-vergence trajectories under free- & instrument-space environments" *Curr Eye Res.*, vol. 17, pp. 159-164. 1998.

[12] A. T. Bahill, J. S. Kallman, and J. E. Lieberman, "Frequency limitation of the two-point central difference differentiation algorithm," *Biol. Cybern.*, vol. 45, pp. 1-4, 1982.

[13] J. L. Semmlow, and W.Yuan, "Components of disparity vergence eye movements: application of independent component analysis." *IEEE Trans.Biomed.Eng.*, vol. 49, pp.805-811, 2002.

[14] P. Comon, "Independent Component Analysis -- a new concept?" *Signal Processing*, vol. 36, pp. 287-314, 1994.

[15] A. Hyvarinen, J. Karhunen , and E.
Oja., *Independent Component Analysis*. New York: John Wiley & Sons, Inc; 2001.

[16] V. F. Kiviniemi, J. H. Kantola, J. Jauhiainen, A. F. Hyvarinen, and O. Tervonen, "Independent component analysis of nondeterministic fMRI signal sources," *Neuroimage*, vol. 19, pp. 253-260, 2003.

[17] L. Mays, J. Porter, P. and Gamlin, "Neural Controal of Vergence Eye Movements: Neurons Encoding Vergence Velocity." Neurophysiology, Vol 56, No. 4, pp 1007 – 1021, 1986.

[18] L. Mays, and J. Porter, "Neural Control of Vergence Eye Movements: Activity of Abducens and Oculomotor Neurons." Neurophysiology, Vol. 52, No. 4, pp 743 – 759, 1984.