Vision Research 48 (2008) 1705–1718

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Simulation of artificial vision: IV. Visual information required to achieve simple pointing and manipulation tasks

Angélica Pérez Fornos *, Jörg Sommerhalder, Alexandre Pittard, Avinoam B. Safran, Marco Pelizzone

Ophthalmology Clinic, Department of Clinical Neurosciences, Geneva University Hospitals, Rue Micheli-du-Crest 24, 1205 Geneva, Switzerland

ARTICLE INFO

Article history:
Received 20 November 2007
Received in revised form 17 March 2008

Keywords:
Retinal prosthesis
Blindness
Visuomotor performance
Target localization
Shape recognition

ABSTRACT

Retinal prostheses attempt to restore some amount of vision to totally blind patients. Vision evoked this way will be however severely constrained because of several factors (e.g., size of the implanted device, number of stimulating contacts, etc.). We used simulations of artificial vision to study how such restrictions of the amount of visual information provided would affect performance on simple pointing and manipulation tasks. Five normal subjects participated in the study. Two tasks were used: pointing on random targets (LEDs task) and arranging wooden chips according to a given model (CHIPs task). Both tasks had to be completed while the amount of visual information was limited by reducing the resolution (number of pixels) and modifying the size of the effective field of view. All images were projected on a 10° × 7° viewing area, stabilised at a given position on the retina. In central vision, the time required to accomplish the tasks remained systematically slower than with normal vision. Accuracy was close to normal at high image resolutions and decreased at 500 pixels or below, depending on the field of view used. Subjects adapted quite rapidly (in less than 15 sessions) to performing both tasks in eccentric vision (15° in the lower visual field), achieving after adaptation performances close to those observed in central vision. These results demonstrate that, if vision is restricted to a small visual area stabilised on the retina (as would be the case in a retinal prosthesis), the perception of several hundreds of retinotopically arranged phosphenes is still needed to restore accurate but slow performance on pointing and manipulation tasks. Considering that present prototypes afford less than 100 stimulation contacts and that our simulations represent the most favourable visual input conditions that the user might experience, further development is required to achieve optimal rehabilitation prospects.

1. Introduction

Technological advances have opened new perspectives and, today, it is possible to envision neural prostheses to restore some amount of vision to totally blind patients. Several research groups have initiated projects aiming at the development of such prostheses and the implantation of the first prototypes has boosted interest in this field (Chow et al., 2004; Dobelle, 2000; Richard, Hornig, Keseru, & Feucht, 2007; Veraart, Wanet-Defalque, Gerard, Vanlierde, & Delbeke, 2003; Yanai et al., 2007; Zrenner et al., 2007). Such devices aim to restore function by direct electrical stimulation of surviving neural tissue. Yet, the "artificial" visual percepts generated this way will be limited by different factors. According to their origin, these constraints can be classified into: Those due to the technical characteristics of the stimulating device, those resulting from the intrinsic properties of the electrode–nerve interface, and those due to the functional characteristics of the remaining visual pathway. While the latter two are still largely unknown at present (e.g., spatial selectivity of retinotopic activation, type of cells stimulated, pattern of firing elicited by electrical stimulation, etc.), the most important constraints due to the technical characteristics of the device can be identified. First, image resolution will be limited by the number of discrete stimulation contacts available in the implant. Second, visual percepts will be restricted to a limited and stabilised fraction of the visual field, depending on the size and implantation site of the electrode array.

Our research group is part of a multidisciplinary effort aiming at developing a subretinal implant; i.e., a device transforming *in situ* light reaching the eye into a pattern of stimulation currents (Lecchi et al., 2004; Mazza, Renaud, Bertrand, & Ionescu, 2005; Salzmann et al., 2006; Ziegler et al., 2004). In this context, we developed a series of simulation experiments mimicking the basic visual limitations related to the technical design of a retinal implant (Pérez Fornos, Sommerhalder, Rappaz, Safran, & Pelizzone, 2005; Sommerhalder et al., 2003, 2004). While these simulations do not pretend to actually mimic percepts experienced by a blind subject using a retinal prosthesis, they represent the most favourable visual input conditions that the user may experience (Dagnelie,

---

* Corresponding author. Fax: +41 223828382.
E-mail addresses: Angelica.Perez-Fornos@hcuge.ch, Angelica.Perez@medecine.unige.ch (A. Pérez Fornos).

0042-6989/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2008.04.027

Walter, & Liancheng, 2006). Therefore such simulations, presented to subjects with normal vision, allow determining how performance on various every-day tasks changes versus the amount of visual information provided. This approach is very instructive because performance versus amount of visual information follows in most cases a "step-like" (i.e., psychometric) function, which allows to determine a critical threshold information necessary to achieve the task. Therefore, for retinal prostheses to restore function, at least this critical "minimum" amount of information will have to reach the brain. The rationale is the following: Even if information transmission at the electrode–nerve interface is optimal and considering that brain plasticity may help to adapt to new unnatural percepts, the brain cannot invent visual information that is not provided.

1.1. Vision and eye–hand coordination

A variety of daily life and leisure activities involve gathering information from the environment and using it to visually guide movements towards a certain target (e.g., operating a telephone, locating and taking items from a crowded shelf, locating and using items on a dinner table, ...). A considerable amount of work has been carried out to understand the processes involved in such tasks (for a comprehensive review, see Desmurget, Pelisson, Rossetti, & Prablanc, 1998). Briefly, it has been demonstrated that both visual and non-visual information are used in conjunction during such tasks. Continuous visual monitoring of the motor apparatus significantly contributes to the accuracy of goal-directed movements. In addition, when visual and non-visual sensory signals diverge, visual input appears to be privileged.

The first step in tasks involving eye–hand coordination is the identification/localization of the potential target in space. There is evidence from psychophysical experiments that foveal and peripheral vision assume different roles during visual search. On one hand, detailed object information appears to be primarily coded in the fovea and its close surroundings. Parker (1978) explored eye movement behaviour during a picture recognition task. His results showed that most objects in a visual scene had to be directly fixated so that changes could be detected. Nelson and Loftus (1980) demonstrated that a particular feature of a scene is more likely to be detected when it has been closely (within 2.6°) fixated. These findings have been confirmed by others (De Graef, Christiaens, & d'Ydewalle, 1990; Henderson & Hollingworth, 2003; Hollingworth, Schrock, & Henderson, 2001; Nodine, Carmody, & Herman, 1979). On the other hand, experimental observations suggest that more peripheral areas play a major role in identifying potentially "informative" regions of visual field. Parker (1978) noted that some scene changes could be detected without directly fixating the object that was changed. In addition, objects that were changed were fixated sooner than unchanged objects. Other studies also suggest that subjects tend to fixate areas of the visual scene containing meaningful information based on information gathered from the periphery of the visual field (Antes, 1974; Loftus & Mackworth, 1978). This is consistent with the idea that the perceptual span in scene perception is larger than it is for reading (Henderson, McClure, Pierce, & Schrock, 1997; Rayner & Pollatsek, 1992). These observations indicate that, while visual information extracted from the fovea and its surroundings is crucial for object identification/recognition, useful information about changes of the visual environment can be gathered in more eccentric areas of the visual field, and that such information can be used to elicit a perceptual response (such as redirecting fixation).

These findings fit well with the anatomical and physiological characteristics of the different areas of the visual field. Central vision is functionally specialised in high-resolution sampling of spatial information, while eccentric vision mainly contributes to encoding dynamic and relative distance cues. It can be therefore concluded that central vision plays an important role in target identification and in fine position adjustments, while eccentric vision is mainly responsible of redirecting attention as well as controlling eye (Cornelissen, Bruin, & Kooijman, 2005; Hooge & Erkelens, 1999) and hand (Paillard, 1982; Sivak & Mackenzie, 1992) movements towards regions of interest.

Clinical and epidemiological studies corroborate these findings, revealing that different visual factors have differentiated impact on vision-related daily activities (Nelson, Aspinall, Papasouliotis, Worton, & O'Brien, 2003; Owsley, McGwin, Sloane, Stalvey, & Wells, 2001; West et al., 2002). Visual acuity deficits (i.e. disorders of the central part of the visual field) affect tasks requiring detailed vision, such as those involving object identification. Visual field defects affect localization and orientation abilities, critical for eye–hand coordination.

1.2. Eye–hand coordination in the context of artificial vision

Only a limited number of qualitative experiments have been carried out to explore eye–hand coordination in the context of artificial vision. In a first set of experiments, simulations were achieved using a head mounted video display and pixelising software (Humayun, 2001; Hayes et al., 2003). Almost all subjects were able to pour candies from one cup to another using a grid of 16 × 16 pixels. Under the same conditions, subjects were able to cut a black square drawn on a white paper sheet with approximately 50% accuracy. However, in these studies, some important aspects of prosthetic vision such as stabilised retinal projection and possible eccentric implantation[1] were not considered. A more recent study attempted to evaluate the issue of retinal stabilisation on a checker placing task (Dagnelie et al., 2006). Artificial vision was simulated as a 6 × 10 array of Gaussian pixels in two conditions: Free-viewing and stabilised viewing. Errors were rare, but augmented when task difficulty was increased and in stabilised-viewing conditions. During short periods of practice, performance improved slightly and the time required for completing the task in stabilised-viewing decreased, becoming similar to that achieved in free-viewing. Yet, in these experiments the eye tracking system used was relatively slow (30 Hz). This resulted in delayed stimulus presentation/update to subjects which impeded appropriate image stabilisation. Furthermore, some parameters of prosthetic vision were either not considered (e.g., non-foveal implantation) or not varied over a sufficient range (e.g., image resolution) to unambiguously determine optimum performance.

Psychophysical testing on blind subjects participating in a chronic implantation trial has also been presented (Humayun et al., 2003; Yanai et al., 2007). Subjects used an epiretinal prosthesis with an array of 4 × 4 stimulating electrodes and had to perform elementary tasks (e.g., locating a moving flash light in a dark room, determine the orientation of a capital L). While very interesting, actual performance of these subjects on such tasks does not allow extrapolation to more complex and realistic every-day situations.

---

[1] Due to the anatomo-physiology of the retina it might be preferable to implant retinal prostheses in eccentric areas of the visual field (>10°) to better preserve retinotopic activation. Refer to our previous publications (Sommerhalder et al., 2003, 2004) for details. This prediction, based exclusively on anatomical/histological observations, has not been tested yet. However, it could be verified in the first human clinical trials of retinal prostheses, which have recently begun and include patients suffering from Retinitis Pigmentosa. Other pathologies in which parafoveal areas of the retina are better preserved than central areas (e.g., age related macular degeneration) will offer further opportunities of testing.

1.3. Aim of this study

Our first studies focused on reading (Pérez Fornos et al., 2005; Sommerhalder et al., 2003, 2004). We investigated the effects of the limited resolution imposed by the finite number of stimulation contacts as well as the effects of the restricted visual field imposed by the limited size of the implant. Our results demonstrated that approximately 500 distinctly perceived phosphenes, presented on a restricted 10° × 7° window stabilised in the centre of the visual field, are required to achieve accurate reading at rates above 70 words/min. In eccentric vision (15° in the lower visual field), good reading accuracy could also be achieved, but only after a prolonged period of approximately 1½ months of daily practice and at significantly lower reading rates.

It is clear from the literature that important every-day life tasks require encoding spatial information and using it to direct a particular motor response, which might impose different visual requirements than reading. In this paper, we use simulations of artificial vision to explore how some tasks, requiring visual input and using motor output, are influenced by important technical limitations of a retinal prosthesis. We developed two “close to reality” pointing and manipulation tasks. A portable video system was built to project pixelised stimuli onto defined, stabilised visual field areas of normally sighted subjects. In a first experiment, conducted in central vision, we systematically investigated the influence of two important stimulus parameters (i.e., image resolution and size of the effective field of view) on performance. A second longitudinal experiment was then conducted to assess whether subjects could learn to perform the same tasks using eccentric vision.

2. Methods

2.1. Subjects

Five subjects (S1, female, 27-years-old; S2, male, 42-years-old; S3, male, 24-years-old; S4, male, 34-years-old; S5, male, 28-years-old), familiar with the purpose of the study, were recruited either from the staff of the Ophthalmology Clinic of the Geneva University Hospitals or from the staff of the University of Geneva. They all had visual acuity better than 16/20 on the tested eye, normal ophthalmologic status, and normal haptic perception.

All experiments were conducted according to the ethical recommendations of the Declaration of Helsinki, and were approved by local ethical authorities. All subjects signed appropriate consent forms.

2.2. Eye–hand coordination tasks

We designed the two tasks used in this study based on common clinical tests and on previous studies in natural tasks and settings (Humayun, 2001; Land, Mennie, & Rusted, 1999; Pelz, 1995; Pelz & Canosa, 2001; Pelz, Hayhoe, & Loeber, 2001; Purdy, Lederman, & Klatzky, 1999).

2.2.1. Pointing: The LEDs task

Subjects were sitting at a table facing a panel composed of a 6 × 4 array of red light emitting diodes (LEDs). The array of LEDs was covered with a red filter to avoid that potential targets were seen when not lit (Fig. 1a). Centre-to-centre distance between LEDs was 6 cm and when lit, the diameter of the circular bright spot of the LEDs was approximately 1 cm. A transparent 19.7″ touch screen (3M Touch Systems, Massachusetts, USA), subtending 39 × 31.5 cm^2, was placed over the LEDs panel for recording subjects’ responses.

In each experimental run, each of all 24 LEDs was successively lit in random order. Subjects had to point with the finger, as precisely as possible, on the bright target lighting up randomly on the panel. Pointing accuracy and pointing time for each target were recorded.

2.2.2. Manipulation and form recognition: The CHIPs task

Subjects were sitting at a table and facing a 5 × 4 template of square wooden chips, each representing one of 20 different black figures drawn on a white background (Fig. 1b). The figures appearing on the chips measured 3–6 cm along each axis. Chips measured 8 × 8 cm^2 and were covered with a smooth and transparent plastic sheet to remove tactile cues. The total working surface was 40 × 32 cm^2.

For each experimental run, a custom program randomly determined the position of the chips on the template (none of the subjects was presented twice with the same CHIPs arrangement). At the beginning of each run, the randomised template was placed in front of the subject, who also received a box containing a copy of each chip. The subject was instructed to pick up, one by one, chips from the box and to place them correctly (in the adequate position and orientation) on the template. Once the subject released a chip at a certain location, his performance was scored according to the following grading: 1 = correct position and orientation; 0.5 = correct position but wrong orientation; 0 = wrong position. Then, the examiner removed the chip to avoid the use of structural and tactile cues for identifying and positioning the remaining chips. The experiment ended once subjects had placed all chips. The total time to complete the task was measured from the time the task started (activation of the LCD display) until the moment the last chip was positioned on the template by the subject.

2.3. The artificial vision simulator

The stabilised projection of a 10° × 7° viewing area on the retina was achieved with a high-speed (250 Hz) video based eye tracking system (EyeLink I; SensoMotor Instruments GmbH, Berlin, Germany). A detailed description of the stationary setup used in our previous experiments can be found in our publications (Sommerhalder et al., 2003, 2004; Varsori, Pérez Fornos, Safran, & Whatham, 2004). Briefly, this system consists in three head-mounted cameras and two personal computers. The *operator PC*, dedicated to computing eye position, was a Compaq Deskpro EP (Celeron-400). The *subject PC* was connected to the stimulation screen and used for experiment control. It was a Dell Latitude C840 (P4-M, 2.2 GHz) notebook equipped with a nVidia GeForce4 440 Go 64-bit graphics card and running Windows XP SP1.

For this series of experiments, the stationary setup of the eye-tracker used in our previous experiments was modified to allow subjects to freely move their head and trunk during the tasks (see Fig. 2). A LCD display (NEC NL6448BC26-01), measuring 170.1 × 128.2 mm^2, was fixed to the headband, in front of the subject’s eyes, at an eye-to-screen distance of 23 cm. The LCD display was set to a refresh rate of 75 Hz and to a resolution of 640 × 480 pixels. It subtended therefore a visual field of 40° × 30° and 1 pixel on the screen corresponded to 0.06° of visual angle. All the

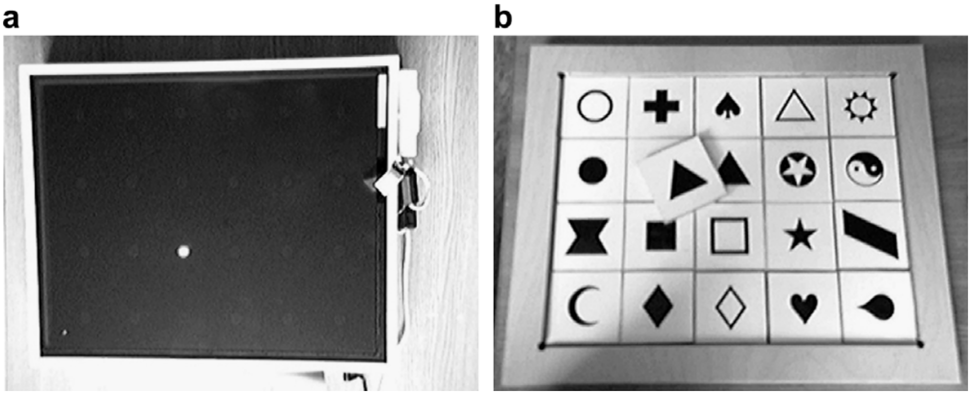

Fig. 1. Tasks used in this study: (a) the LEDs task consisted in pointing accurately on targets presented in random order, (b) the CHIPs task consisted in superimposing wooden chips with corresponding figures on a randomised pattern.

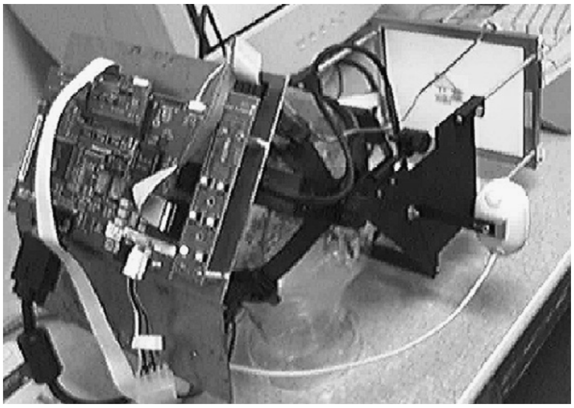

Fig. 2. The mobile artificial vision simulator. The stationary configuration of a SMI Eyelink I system was modified in order to allow for mobility. A LCD display was fixed in front of the subject's eyes and a webcam on the side of the system.

electronics needed for the LCD display were attached as counterweight in the back of the head. A cover of black cloth prevented the subject from seeing anything else than the screen. A bite-bar was used to provide the stability necessary for accurate eye position computation. A webcam (Philips ToUCam Pro), mounted on the side of the system (at eye-height), was connected to the subject PC and captured the visual environment at a frame rate of 30 Hz. Such a system guaranteed a maximum delay for stimulus update on the screen of 25 ms, which was enough to insure accurate image stabilisation on the retina and to be unperceived by subjects.[2]

Various image resolutions or "pixelisations" were used. This was achieved with basic image processing techniques that decomposed the image frames captured by the webcam into a given number of square, uniform pixels (real-time square pixelisation[3]). Details on the image processing algorithms can be found elsewhere (Pérez Fornos et al., 2005).

2.4. Testing procedure

Subjects were seated wearing the mobile setup (see Fig. 3a). All tests were performed monocularly. Each run started with a standard 9-point calibration of the eye tracking system. Pixelised portions of the images captured by the webcam were projected on a 10° × 7° viewing window (Fig. 3b). Gaze position compensation was used to project this viewing window onto defined and stabilised (central or eccentric) areas of the retina. The remaining screen surface was uniformly grey. At the end of each run the calibration was checked again for possible drifts. In rare cases when the average error obtained in the calibration check was $\geqslant 1°$, the results for the corresponding session were discarded from the analysis.[4] Test sessions included as many runs as possible, however never exceeding 30 min of testing to avoid subjects' fatigue.

Subjects could explore the environment (i.e., modify the content of the image displayed in the viewing window) in two ways: (1) by moving the viewing window on the screen with eye movements to scan the larger image captured by the webcam and/or (2) by moving their head and trunk to capture different portions of the visual environment with the head-mounted webcam. Eye and head movement data were recorded during the experiments to investigate how subjects used these movements to cope with different viewing conditions. The analysis of these recordings is presented in Appendix A.

2.4.1. Experiment 1: Investigations in central vision

Previous research has demonstrated that the size of the field of view has a significant influence in performance for tasks involving visual search and orientation, such as eye–hand coordination and mobility tasks (Kerkhoff, 1999; Nelson, Aspinall, Papasouliotis, Worton, & O'Brien, 2003; Rubin et al., 2001; Szlyk et al., 2001). Large fields of view allow the exploration of a significant part of the whole visual scene at glance, and thus tend to facilitate visual search and orientation. As schematised in Fig. 4, it is clear that a field of view encompassing the whole visual scene at the highest possible image resolution would be the optimal and most natural condition. However, in the context of artificial vision where the number of stimulation contacts or "pixels" is limited, a large effective field of view would imply a deterioration of image resolution which will impact object recognition abilities (right-hand column in Fig. 4). A more pragmatic approach to cope with a limited number of pixels would be to consider using a smaller effective field of view (lower line in Fig. 4), but this would imply time-consuming "tunnel-scanning" of the scene to achieve the task. Actually, an almost infinite number of possibilities exist since the subject can smoothly optimize this field of view by varying the object-to-camera distance (i.e., approaching or retreating from the working plane) by head and trunk movements. How will these two counteracting constraints affect performance on the tasks considered in this study?

Experiment 1 was designed to investigate this issue quantitatively. We systematically varied: (1) the effective field of view (by changing the frame size of the image captured by the webcam) and (2) the resolution of the image (by changing the total number of pixels representing the image contained in the viewing window). Note that across all conditions, the actual viewing area into which this information was projected was always 10° × 7°. In other words, only the amount of information contained in the stimulus image changed across experimental conditions, not the size of the viewing window (which represented a retinal implant of fixed physical size; see also Fig. 4). Psychophysical experiments were conducted using three different effective fields of view (8.2° × 5.8°, 16.5° × 11.6°, and 33° × 23.1°) and five different pixelisation levels (17920, 1991, 498, 221, and 124 pixels). Three runs were performed in each of these 15 experimental conditions. The testing order of the effective fields of view per subject was determined using a Latin Square. For each field of view, subjects started with the easiest (highest) pixelisation level and progressed towards the most difficult (lowest) one. This experimental protocol aimed at avoiding possible learning effects favouring a particular effective field of view, but tended to favour performance at low pixelisation levels.

Before starting the actual experimental sequence, all subjects performed three control sessions for each task in normal viewing conditions (not wearing the mobile setup). These results were used as a baseline for "normal" performance on the tasks.

2.4.2. Experiment 2: Learning effects in eccentric vision

Due to the anatomo-physiology of the retina, it might be preferable to implant retinal prostheses in eccentric areas of the visual field (>10°; see Sommerhalder et al., 2003, 2004 for details). It is thus important to also test performance with a viewing window stabilised in an eccentric area of the visual field. This condition is unnatural for normally sighted subjects and, consequently, they have to adapt to eccentric viewing to get optimal results. Subjects participating in this experiment were naïve to eccentric viewing (i.e., they had no previous experience with eccentric viewing). Possible learning effects were investigated by repeatedly performing the tasks in the same experimental condition for more than 1 month. The experimental condition used for this experiment (498-pixel resolution with an effective field of view of 16.5° × 11.6°) was determined on the basis of the results of Experiment 1. The viewing window was stabilised at 15° eccentricity in the lower visual field.

Each experimental session consisted of one run of the CHIPs task followed by one run of the LEDs task. Two to three experimental sessions were conducted each working day of the week (5 days per week). The criterion used to stop the experiment was the stabilisation of temporal performance (i.e., time required to perform the tasks).

2.5. Data analysis and statistics

For the LEDs task, performance was measured on the basis of two variables: mean pointing time per target and mean pointing error. The first was calculated as the total time required for locating and pointing on all targets divided by the number of targets. This represented thus a global measure per target that included reaction time, visual search time (time required to locate the target), and movement time. Mean pointing error was calculated as the cumulative pointing error for all targets divided by the number of targets.

For the CHIPs task, performance was determined on the basis of the mean chip placement time (calculated as the total time required for placing all chips divided by the number of correctly placed chips) and on the basis of the percentage score

---

[2] The 25 ms delay represents a maximum for a saccade spanning the largest diagonal of the screen. Please note that the frame rate of the webcam has no influence on image stabilisation on the retina. It simply limits the rate at which changing (environment) images are updated. The 30 Hz rate used is fast enough for real-time video streaming.

[3] A number of studies have demonstrated that performance is considerably hampered when images are decomposed into uniform square pixels (Harmon & Julesz, 1973). We investigated this issue in a previous study (Pérez Fornos et al., 2005). Our results demonstrated that simulations of artificial vision should use real-time pixelisation algorithms where the grey level of each pixel varies dynamically according to gaze position, as we used in this study. Yet, in this condition the actual shape of the pixel (square versus Gaussian) did not have a significant influence on performance. We therefore decided to use square pixelisation in this study to avoid making any assumptions on the actual shape of the phosphenes to be evoked by a retinal prosthesis (which are still unknown at present).

[4] In practice, drifting was relatively rare. We did not have to discard any results for experiments conducted in central vision (Experiment 1 and adaptation to the tasks during Experiment 2). However, a small number of sessions had to be discarded for Experiment 2 in eccentric vision. These sessions were removed from the analysis but were not repeated since we can assume that some cognitive learning actually took place.

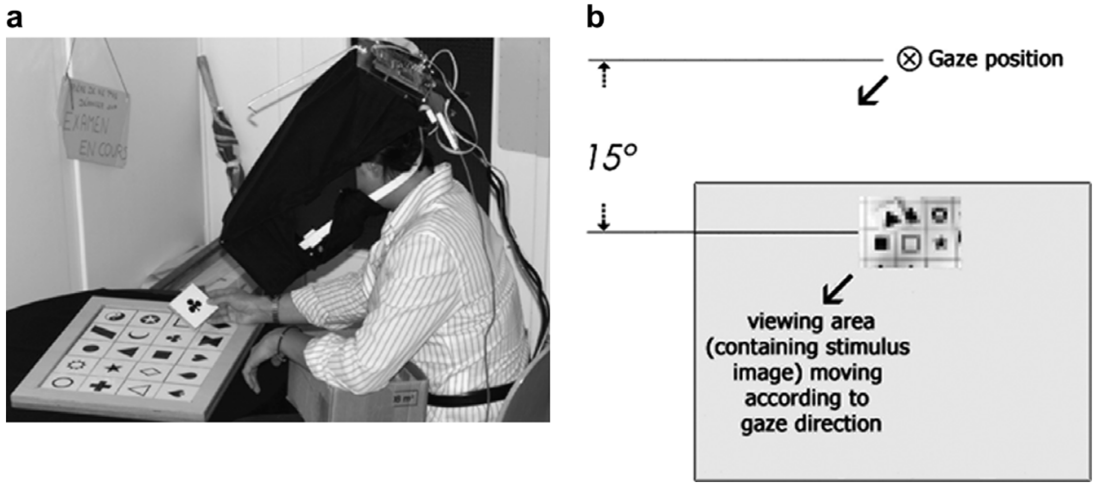

Fig. 3. (a) A subject wearing the mobile setup during the CHIPs task. (b) The stimulation screen as viewed by the subject. The viewing window, containing pixelised fragments of the environment images captured by the webcam, moves on the screen according to the direction of gaze and with a certain offset (in this example, 15° eccentricity).

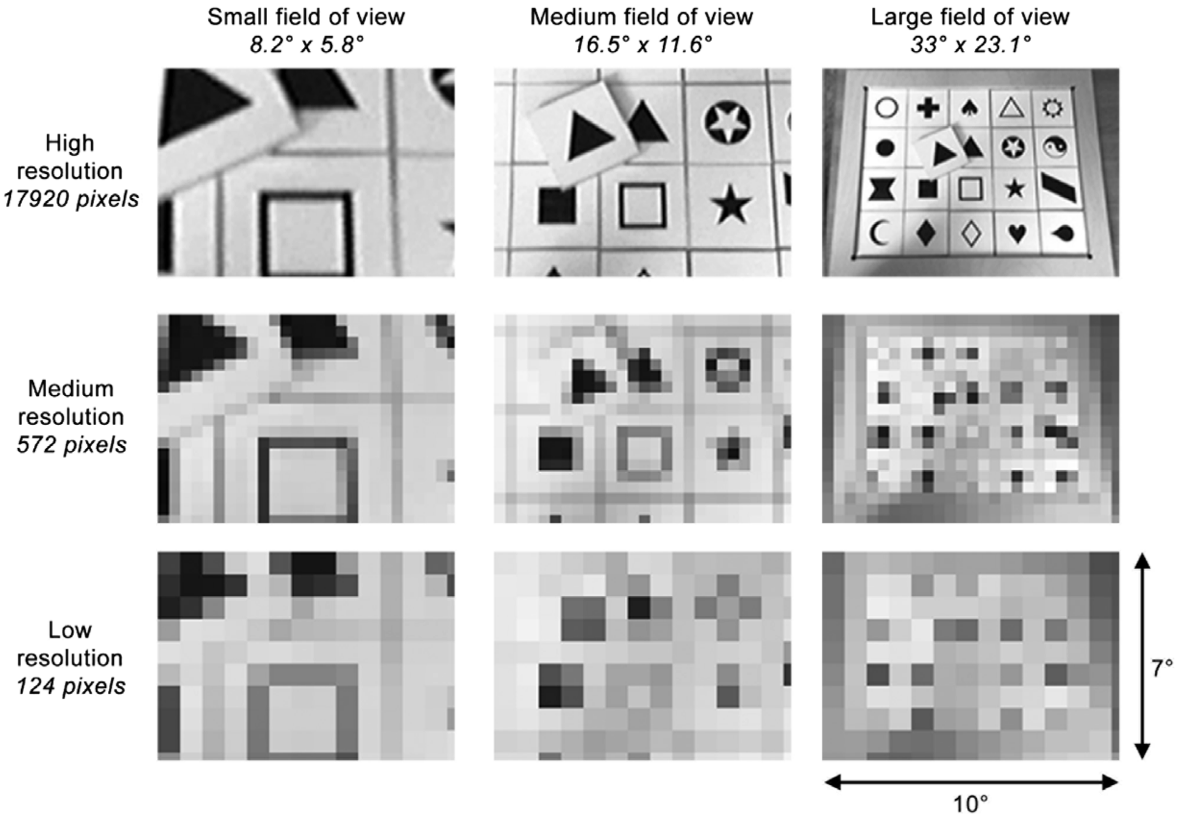

Fig. 4. Examples of images presented on the 10° × 7° viewing window at various experimental conditions during the CHIPs task. Lines illustrate the effect of varying the size of the effective field of view (8.2° × 5.8°, 16.5° × 11.6°, and 33° × 23.1°) at a given image resolution. Columns illustrate the effect of varying image resolution (17920, 498, and 124 pixels) at a given field of view.

of correctly placed chips. Percentage scores were transformed to rationalised arcsine units (RAU)[5] for statistical analyses. However, equivalent %-scales are shown on the right axes of the graphs for clarity.

In Experiment 1, results are presented as the mean of the cumulative data for each subject (three sessions per subject) ± standard error of the mean (SEM). Statistically significant effects were determined using two-factor, repeated measures analysis of variance (ANOVA) with a significance level of 0.05. In Experiment 2, significant learning effects were determined using simple linear correlation (Pearson's correlation).

3. Results

3.1. Experiment 1: Investigations in central vision

The goal of Experiment 1 was to determine which parameter set, amongst the 15 different experimental conditions, allowed to

---

[5] Percentage scores are not adequate for statistical analyses since these data are not normally distributed around the mean and values are not linear in relation to test variability (refer to Pérez Fornos et al., 2005; Sommerhalder et al., 2003, 2004; Studebaker, 1985 for details).

achieve “optimum” performance on the tasks. Three subjects (S1, S2, and S3) participated in this experiment.

3.1.1. Results for the LEDs task

Fig. 5 compares mean performance for the LEDs task versus image resolution for each effective field of view. Mean pointing error was significantly influenced by both image resolution in the viewing window (ANOVA: $F_{(4,30)} = 8.77$, $p < .0001$) and by the size of the effective field of view projected in the viewing window (ANOVA: $F_{(2,30)} = 7.26$, $p < .01$). At the two highest resolution levels tested, pointing errors were approximately 1 cm, independent of the effective visual field tested. With the 8.2° × 5.8° and 16.5° × 11.6° fields of view, pointing errors tended to increase only at the lowest resolution tested (124 pixels). With the larger 33° × 23.1° field of view, pointing errors increased already at 498 pixels.

Mean pointing time was significantly influenced by the size of the effective field of view projected in the viewing window (ANOVA: $F_{(2,30)} = 38.42$, $p < .0001$), but no significant effect of image resolution was observed. The longest pointing times were obtained with the 8.2° × 5.8° field of view (~9.1 s). With the 16.5° × 11.6° field of view, mean pointing times were approximately 6.0 s. The shortest pointing times were obtained with the 33° × 23.1° field of view, which yielded values of about 4.0 s at the highest image resolutions.

Comparison of these results with performance obtained with normal viewing (solid grey lines in Fig. 5) reveals that pointing errors were 2–3 times larger and pointing times were 3–7 times longer than normal in our particular experimental conditions.

During the LEDs task, subjects had to visually detect the position of the target and then use this information to coordinate the pointing movement. Yet, this task did not require any form recognition abilities. It is therefore not surprising that reducing image resolution had only a relatively small effect on pointing accuracy and did not significantly affect pointing time. In contrast, pointing time was quite sensitive to the size of the field of view, probably reflecting the increased difficulty of scanning the visual environment with a restricted viewing window (“tunnel vision”).

3.1.2. Results for the CHIPs task

Fig. 6 compares mean performance for the CHIPs task versus image resolution in the viewing window, for each effective field of view. Chip placement scores were significantly influenced both by image resolution (ANOVA: $F_{(4,30)} = 5.30$, $p < .01$) and by the size of the effective field of view (ANOVA: $F_{(2,30)} = 4.55$, $p < .05$). Close to perfect scores (>95%) were achieved in most viewing conditions. Placement accuracy dropped below 95% correct only at 124 pixels with the 16.5° × 11.6° field of view, and at 221 and 124 pixels with the larger 33° × 23.1° field of view.

Overall, chip placement time was significantly influenced only by image resolution (ANOVA: $F_{(4,30)} = 7.36$, $p < .001$). Mean chip placement time tended to increase at low image resolutions for the three effective fields of view. With the 8.2° × 5.8° and with the 16.5° × 11.6° fields of view, values increased at 221 pixels and below. With the 33° × 23.1° field of view chip placement time increased at 498 pixels already. Similar to what was observed in the LEDs task, the effect of the size of the field of view was clear but only at the highest image resolution tested. However, opposite to what was observed in the LEDs task, the advantage of using the largest field of view was not preserved at lower resolutions where the effects of image resolution and effective field of view intermingled (the curves crossed at resolutions below 1000 pixels; compare Figs. 5b and 6b). At the highest resolutions chip placement time was approximately 2–6 times longer than in normal viewing conditions (~2 s; solid grey lines in Fig. 6).

3.2. Experiment 2: Learning effects in eccentric vision

The objective of this experiment was to determine whether normal subjects could adapt to performing eye–hand coordination tasks with a low-resolution, restricted viewing area stabilised in eccentric vision. Three subjects (S1, S4, and S5) participated on these experiments. Subject S1 was already familiar with the tasks since she also participated in Experiment 1.

Based on the results of Experiment 1, a resolution of 498 pixels and an effective field of view of 16.5° × 11.6° were selected as a good resolution/visual span compromise for performance. Subjects also spontaneously reported preferring this viewing condition to the others.

3.2.1. Adaptation to the tasks

In order to separate possible learning effects to the experimental setup *per-se* from those due to adaptation to eccentric viewing, we first had subjects perform several test sessions using a viewing window stabilised in central vision. Minor adaptation effects were observed, especially within the very first sessions. Testing continued however until subjects’ performance systematically asymptoted (~20 sessions). Stabilised performance (Table 1) was computed

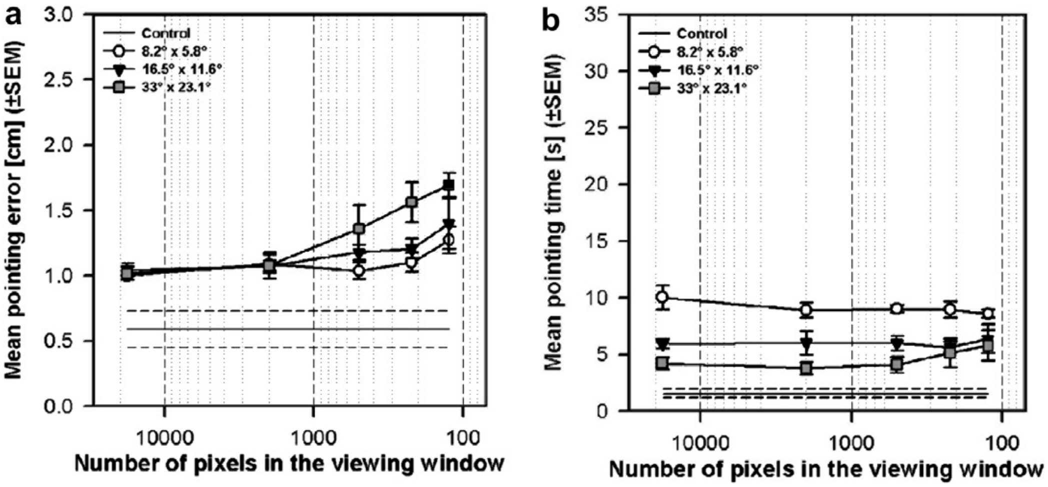

Fig. 5. Performance versus image resolution for 3 normal subjects performing the LEDs task. Three effective fields of view projected in the 10° × 7° viewing window are compared in central vision: 8.2° × 5.8° (empty circles), 16.5° × 11.6° (black triangles), and 33° × 23.1° (grey squares). (a) Mean pointing error [cm] ± SEM. (b) Mean pointing time (average time required for finding and pointing on a target) [s] ± SEM. The grey lines indicate mean performance results ± SEM using normal viewing (control condition).

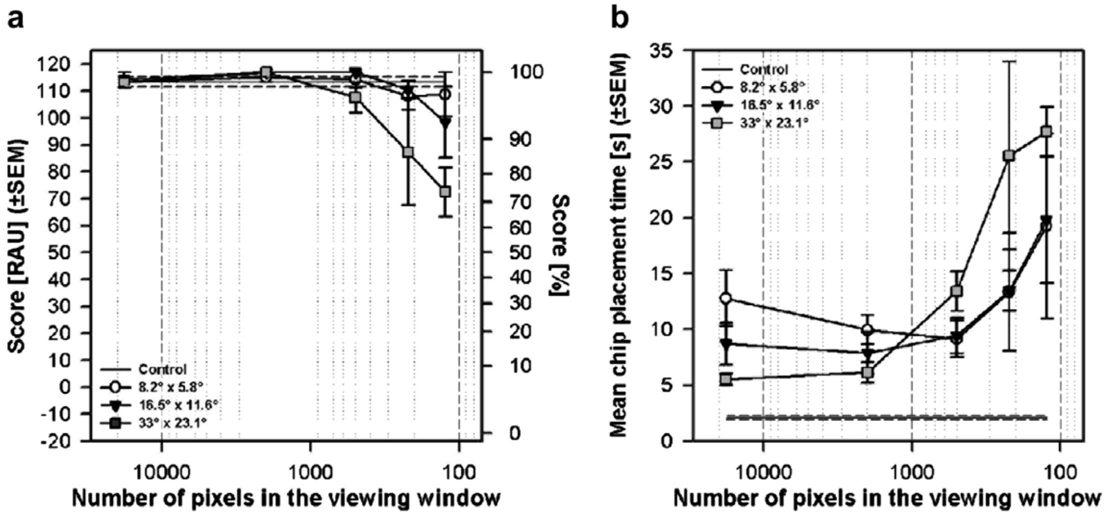

Fig. 6. Performance versus image resolution for three normal subjects performing the CHIPs task. Three effective fields of view projected in the 10° × 7° viewing window are compared in central vision: 8.2° × 5.8° (empty circles), 16.5° × 11.6° (black triangles), and 33° × 23.1° (grey squares). (a) Mean correct scores expressed in RAU ± SEM (left scale) and in % (right scale). (b) Mean chip placement time (mean time required to identify and correctly place a chip) [s] ± SEM. The grey lines indicate mean performance results ± SEM during control sessions (normal viewing conditions).

as the mean ± SEM of the last five sessions. These data collected in central vision serve as reference to those collected in eccentric vision.

3.2.2. Learning in eccentric vision

Once subjects had adapted to the tasks in central vision, data collection in eccentric viewing (using a viewing window stabilised at 15° eccentricity in the lower visual field) began.

Fig. 7 shows performance in eccentric vision versus session number for the LEDs task. Pointing errors were relatively stable for all three subjects and comparable to the values they achieved in central vision. No systematic learning effect versus time could be observed. However, important and statistically significant learning effects were observed in pointing times for two subjects. Subject S1 improved from 21.7 to 6.0 s (Pearson's correlation: $r = .68$, $p < .0001$) and subject S5 from 22.3 to 6.5 s (Pearson's correlation: $r = .63$, $p < .001$). Only a slight but non-significant similar trend, from 7.4 to 5.1 s, was observed for subject S4.

Fig. 8 shows performance in eccentric vision versus session number for the CHIPs task. Scores for subject S1 improved rapidly and impressively: from initial scores below 10% correct, up to final scores above 97% correct. Subject S4 already achieved scores between 95% and 100% correct in the initial sessions, and achieved perfect scores at the end of the experiment. Subject S5 started the experiment with relatively high scores, between 83% and 98% correct, and consistently achieved perfect scores at the end of the experiment. Improvements in scores were statistically significant for subjects S1 and S5 (Pearson's correlation: $r = .47$, $p < .05$ and $r = .73$, $p < .001$, respectively). Improvements in chip placement time were more progressive and more consistent across subjects. Subject S1 showed an approximately 4-fold improvement: from above 30 s in the first sessions, down to around 6.9 s. For subject S4, chip placement time decreased from around 25 s in the initial sessions, down to approximately 8.9 s. Subject S5 started the experiment with values above 20 s, and stabilised around 9.3 s. Improvements in chip placement time were statistically significant for all subjects (Pearson's correlation: $r = .80$, $p < .0001$ for S1; $r = .61$, $p < .005$ for S4; and $r = .81$, $p < .0001$ for S5).

Comparison of final performance in eccentric vision with average performance in central vision (Table 1) reveals that, after training, subjects achieved similar accuracy (pointing precision for the LEDs task and chip placement scores for the CHIPs task) in both tasks. Only time performance in eccentric vision remained slightly slower for all subjects and on both tasks.

4. Discussion

In this paper we attempted to assess how performance on simple pointing and manipulation tasks was affected when visual information is artificially limited, as will be the case for future users of retinal prostheses even with optimal information transmission from the device to the brain. It is known from experimental research as well as from clinical practice that visual acuity and visual field deficits affect performance on such tasks (Antes, 1974;

Table 1
Mean performance of three subjects in central vision

	S1	S4	S5
LEDs task			
Pointing error [cm] ± SEM	1.1 ± 0.04	1.1 ± 0.08	1.4 ± 0.11
Pointing time [s] ± SEM	3.9 ± 0.22	3.4 ± 0.12	4.8 ± 0.45
	CHIPs task		
Score [RAU] ± SEM (%)	112.8 ± 0.00 (100%)	109.9 ± 2.92 (98%)	110.9 ± 1.91 (99%)
Chip placement time [s] ± SEM	4.5 ± 0.15	6.6 ± 0.22	6.8 ± 0.36

Experimental condition: 10° × 7° viewing window containing 498 pixels and subtending a 16.5° × 11.6° effective field of view. Results were calculated on the basis of performance measured during the last five experimental sessions of a short adaptation period (~20 sessions).

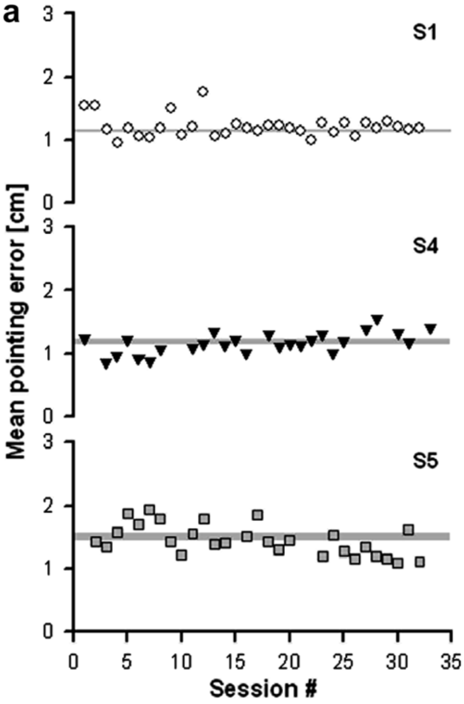

Fig. 7. Performance versus session number obtained for three normal subjects performing the LEDs task in eccentric vision (15° in the lower visual field). Experimental condition: 10° × 7° viewing window containing 498 pixels and subtending a 16.5° × 11.6° effective field of view. Results expressed as: (a) Mean pointing error [cm] and (b) mean pointing time [s]. The grey bars indicate mean performances ± SEM in central vision (see Table 1). Some sessions had to be discarded from the analysis due to large calibration drifts during the experiment. These can be identified as the missing points in the graphs (refer Section 2.4 for details).

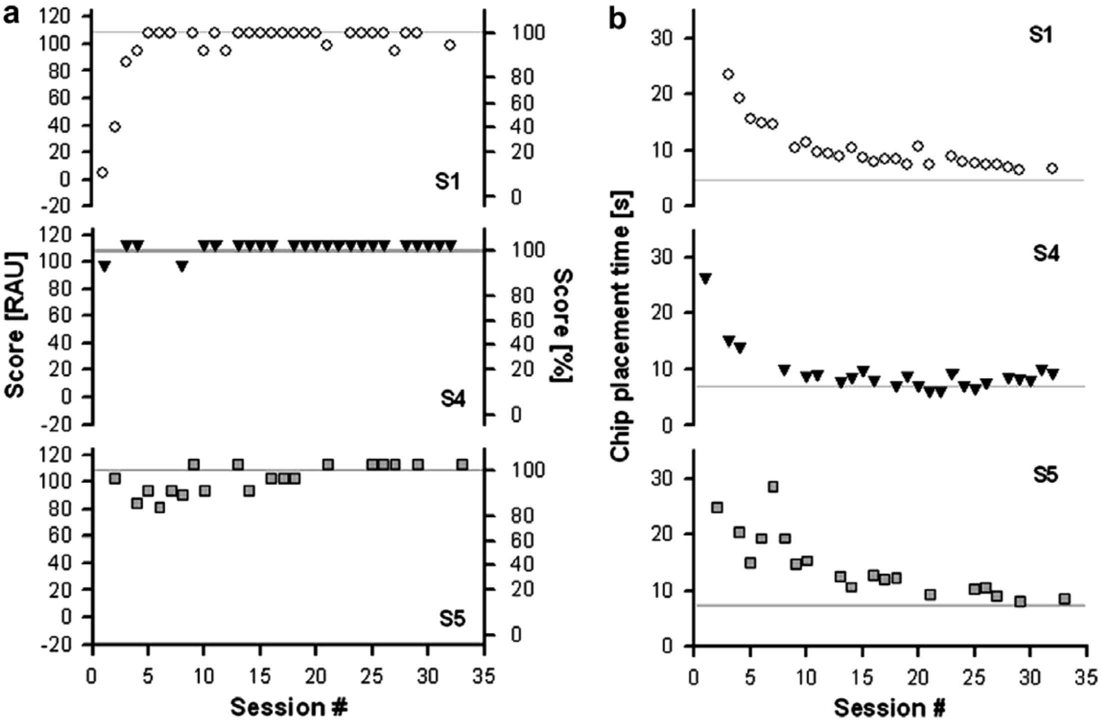

Fig. 8. Performance versus session number obtained for three normal subjects performing the CHIPs task in eccentric vision (15° in the lower visual field). Experimental condition: 10° × 7° viewing window containing 498 pixels and subtending a 16.5° × 11.6° effective field of view. Results expressed as: (a) Correct scores expressed in RAU (left scale) and in % (right scale). (b) Chip placement time [s]. The grey bars indicate mean performances ± SEM in central vision (see Table 1). Some sessions had to be discarded from the analysis due to large calibration drifts during the experiment. These can be identified as the missing points in the graphs (refer Section 2.4 for details).

Cornelissen et al., 2005; Loftus & Mackworth, 1978; Nelson, Aspinall, Papasouliotis, Worton, & O'Brien, 2003; Owsley et al., 2001; West et al., 2002). This is due to the fact that image resolution determines the capacity of discriminating details in a visual image (e.g., to identify potential targets), while the size of the field of view determines the amount of information available for redirect-

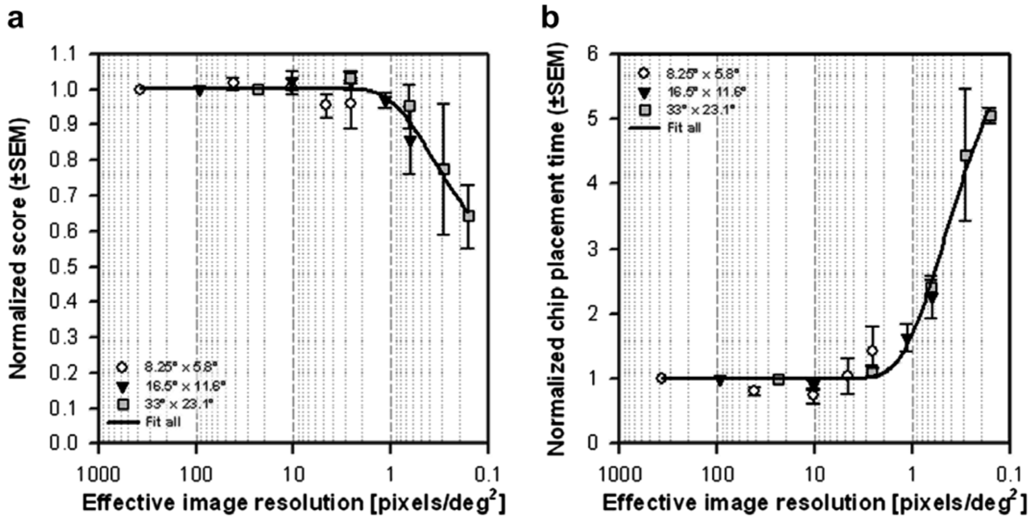

Fig. 9. Normalised performance versus effective image resolution for three normal subjects performing the CHIPs task. Three effective fields of view projected in the 10° × 7° viewing window are compared in central vision: 8.2° × 5.8° (empty circles), 16.5° × 11.6° (black triangles), and 33° × 23.1° (grey squares). (a) Normalised score ± SEM. (b) Normalised chip placement time ± SEM. The solid black lines indicate the best fit to all data.

ing eye/head movements towards "informative" regions of the environment and to plan efficient visual search strategies (e.g., to locate potential targets). In this study, we used two simple tasks (pointing on randomly flashed targets; identifying and arranging geometric forms according to a randomised model) that could not be achieved without relying on visual information. Furthermore, we took special care to eliminate all possible tactile cues and experiments were done in poorly structured visual backgrounds (Loftus, Murphy, McKenna, & Mon-Williams, 2004; Magne & Coello, 2002).

Results from experiment 1 indicate that the fact of exploring the environment with a small and stabilised viewing window presented monocularly strongly limited performance by itself. Even with the largest field of view and at the highest image resolution levels, a condition in which all necessary visual information seems to be available (see Fig. 4), led pointing and chip placement times were at least 2–3 times slower than in normal viewing conditions (see Figs. 5b and 6b). Restricting the size of the effective field of view further limited performance. In the LEDs task, independent of the resolution level, mean pointing times increased significantly and systematically when reducing the effective field of view. This observation suggests lengthier visual search due to a "tunnel vision" effect, as observed in other studies (Bertera, 1988; Bertera & Rayner, 2000; Cornelissen et al., 2005; Henderson et al., 1997; Rayner & Bertera, 1979). To test this hypothesis, we examined possible correlations between eye and head movements recorded during the task (see Appendix A) and pointing times. Interestingly, pointing times were highly and significantly correlated to both eye and head movements (Pearson's correlation: $r = .936$, $p < .0005$ for horizontal eye movements; $r = .932$, $p < .0005$ for vertical eye movements; $r = .825$, $p < .0005$ for horizontal head movements; $r = .782$, $p < .0005$ for vertical head movements; $r = .705$, $p < .0005$ for transversal head movements). A similar observation could be made for chip placement times. First, at the highest resolution tested, chip placement times increased clearly when reducing the field of view. Second, eye and head movements were also significantly correlated with chip placement times (Pearson's correlation: $r = .771$, $p < .0005$ for horizontal eye movements; $r = .643$, $p < .0005$ for vertical eye movements; $r = .628$, $p < .0005$ for horizontal head movements; $r = .580$, $p < .0005$ for vertical head movements; $r = .611$, $p < .0005$ for transversal head movements). This series of significant correlations demonstrates that reducing the effective field of view induces slower performance because time is mainly spent in visual search, therefore supporting our "tunnel vision" hypothesis.

Results of experiment 1 also show an effect of image resolution. This can be seen as a decrease in pointing precision as well as chip placement accuracy when image resolution fell below 500 pixels (Figs. 5a and 6a). These decreases in performance were more pronounced for larger fields of view, consistent with results of previous studies (Bingham & Pagano, 1998; Watt, Bradshaw, & Rushton, 2000). In addition, in the case of chip placement times (Fig. 6b), the effects of image resolution intermingled with those of the size of the field of view. The interaction between these two parameters can be examined by computing the "effective image resolution" (i.e., the number of pixels necessary to code a 1° × 1° effective field of view in each condition). Fig. 9 shows a plot of normalised performance[6] versus this "effective image resolution" for the CHIPs task. Interestingly, all experimental data tend to fall on the same curve (solid black lines in Fig. 9). Fitting this general trend with a single exponential function reveals that best performance is sustained down to resolutions of about 1.5 pixels/deg^2. A similar analysis of the data obtained for the LEDs task (not shown) yields a threshold "effective image resolution" of about 2 pixels/deg^2. Consequently, best performance is obtained using the largest field of view that still provides that "effective image resolution". If not, the advantage of using a larger field of view is counteracted by insufficient resolution. Note that since subjects were able to freely move their head/trunk, they were able to adapt their visual search strategy to optimize this "effective image resolution" according to the task and the viewing condition in use. In practical terms, this meant that in case of insufficient resolution, subjects would instinctively approach the camera towards the target and thus dynamically modify the effective field of view subtended by the object of interest. Head movements recorded during the experiments (see Appendix A) are consistent with this observation.

The previously mentioned results demonstrate that the experimental conditions chosen to simulate artificial vision (limited resolution, restricted field of view, and monocular viewing) increased

---

[6] Performance results for each subject were normalised to values obtained at the highest resolution (17920 pixels) to compensate for the "tunnel vision" effect.

the difficulty of the tasks and were the main reasons for the concomitant decrease in performance. It is also worth mentioning that some additional factors might also be involved, such as those related to the overload caused by the use of the experimental setup itself (e.g., weight of the artificial simulator, lack of depth perception due to stimuli projection on an LCD screen), but these are difficult to evaluate.

The second goal of this study was to investigate whether subjects with normal vision could adapt to the unnatural condition of performing our simple tasks using eccentric viewing. The condition tested in Experiment 2 (498 pixels, 16.5° × 11.6° field of view) corresponded to 2.6 pixels/deg^2, and therefore fulfilled the minimum "effective image resolution" criterion for both tasks. Results from Experiment 2 clearly show that a learning process occurred when our naïve subjects used eccentric vision. The effect of learning was best seen in significant reductions of the time necessary to achieve the tasks. Pointing errors and chip placement scores were only moderately improved by learning. In less than 15 training sessions, normal subjects achieved performances with eccentric vision that were extremely close to those observed with central vision in the same condition. This experiment demonstrates, therefore, that after a short adaptation period simple pointing and manipulation tasks can be achieved relatively easily with eccentric vision, even by normal subjects for whom this condition is disturbing and unnatural.

A potential caveat of this study is that results are based on a small number of subjects. This choice was based on the fact that experiments were extremely time-consuming and extended over a relatively long period of time. However, all subjects showed the very same trends and intersubject variability was much smaller than the observed effects. In these conditions, adding more subjects to the study would certainly give more statistical power to the results, but it would not fundamentally change the observed trends. It is also important to mention the limitations of experiments attempting to simulate artificial vision. It is true that the exact characteristics of the percepts elicited by retinal implants remain still largely unknown at present, and our simulations do not pretend to mimic them exactly. We are however convinced that this type of approach is useful and pertinent. Our experiments attempted to determine the amount of visual information needed to perform a given task. Although the transmission of visual information by a real retinal prosthesis will most certainly not be perfect, this does not change the fact that a sufficient amount of visual information should reach the brain to allow for performance. Our approach is very similar to what has been done with acoustic simulations in the field of cochlear implants (Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). For example, the effect of the limited number of frequency channels in speech recognition was studied this way. Furthermore, it has been demonstrated that "star" cochlear implant users perform in real life at the same level as normal subjects do when listening to acoustic simulations (Friesen, Shannon, Baskent, & Wang, 2001). Thus, such simulation experiments can provide knowledge of primary importance in the design and understanding of prosthetic devices.

5. Conclusions

If blind subjects were able to perform the two tasks used in these experiments, we believe this would represent a significant improvement of their professional and social reinsertion prospects, even if performance remains slower than in normal viewing conditions. In this context, the main outcome from this study is that, if vision is restricted to a small visual area, simple pointing and manipulation tasks can be achieved at relatively low image resolutions. Performance decreased only at image resolutions below 500 pixels,[7] and essentially when large fields of view were used. Our experiments suggest that all necessary information could be transmitted by about 400 pixels coding an effective field of view of approximately 16° × 12°. This pixel density might still seem very elevated, especially when considering that present prototypes of retinal prostheses contain less than 100 stimulation contacts. However, devices containing several hundreds of stimulation contacts seem to be technically feasible with present technology and are, as matter of fact, already under development (see e.g., Loudin et al., 2007; Ohta et al., 2007). Our results demonstrate that the technical efforts required to go beyond present prototypes are justified and important to improve the rehabilitation prospects of future users.

Acknowledgments

This work was supported by the Swiss National Foundation for Scientific Research (Grants 3100-61956.00 and 3152-063915.00). We thank Prof. A. Schnider for his valuable advice in the conception of the experiments, especially the development of the tasks. We also thank Prof. J.M. Meyer and M. Bertossa for their help in conceiving the bite-bar of our mobile setup.

Appendix A

Similar to what people do in everyday situations, our experimental subjects used head/trunk movements to optimize the viewing conditions in order to achieve the eye–hand coordination tasks. We recorded eye and head movements during the experiments to quantify these effects.

A.1. Measurement of eye and head movements

Eye movements on the screen were measured with the eye tracker (EyeLink I; SensoMotor Instruments GmbH, Berlin, Germany). Head/trunk movements were recorded with a 3D tracker (Head Tracker; Logitech Inc., California, USA). This device consists of a mobile receiver (attached to the back of the head-mounted setup) and a static transmitter. 3D head position (cm) and orientation (°) data were sent to the subject PC at rate of 50 Hz.

A.2. Analysis of eye and head movements during Experiment 1

Eye and head movements were analyzed by computing their total length along each degree of freedom[8] during each experimental trial. Results were calculated as the mean of the cumulative data for each subject ± SEM. Statistically significant effects were determined using two-factor (image resolution and effective field of view), repeated measures ANOVA with a significance level of 0.05.

A.3. Results for the LEDs task

Fig. A1 displays the mean cumulative length of eye (left panel) and head (right panel) movements versus image resolution with each effective field of view, for the LEDs task. The total length of eye movements along both axes were significantly influenced by the size of the effective field of view (ANOVA: $F_{(2,30)} = 119.85$, $p < .0001$ for the horizontal coordinate and $F_{(2,30)} = 59.35$, $p < .0001$ for the vertical coordinate), but no significant effect of image resolution was observed. Largest values were observed with the 8.25° × 5.8° field of view (about 38 m horizontally and 25 m

---

[7] Equivalent to a visual acuity of 20/400 (100 μm pixel width).

[8] Horizontal and vertical for eye movements; horizontal, vertical, and transversal for head movements.

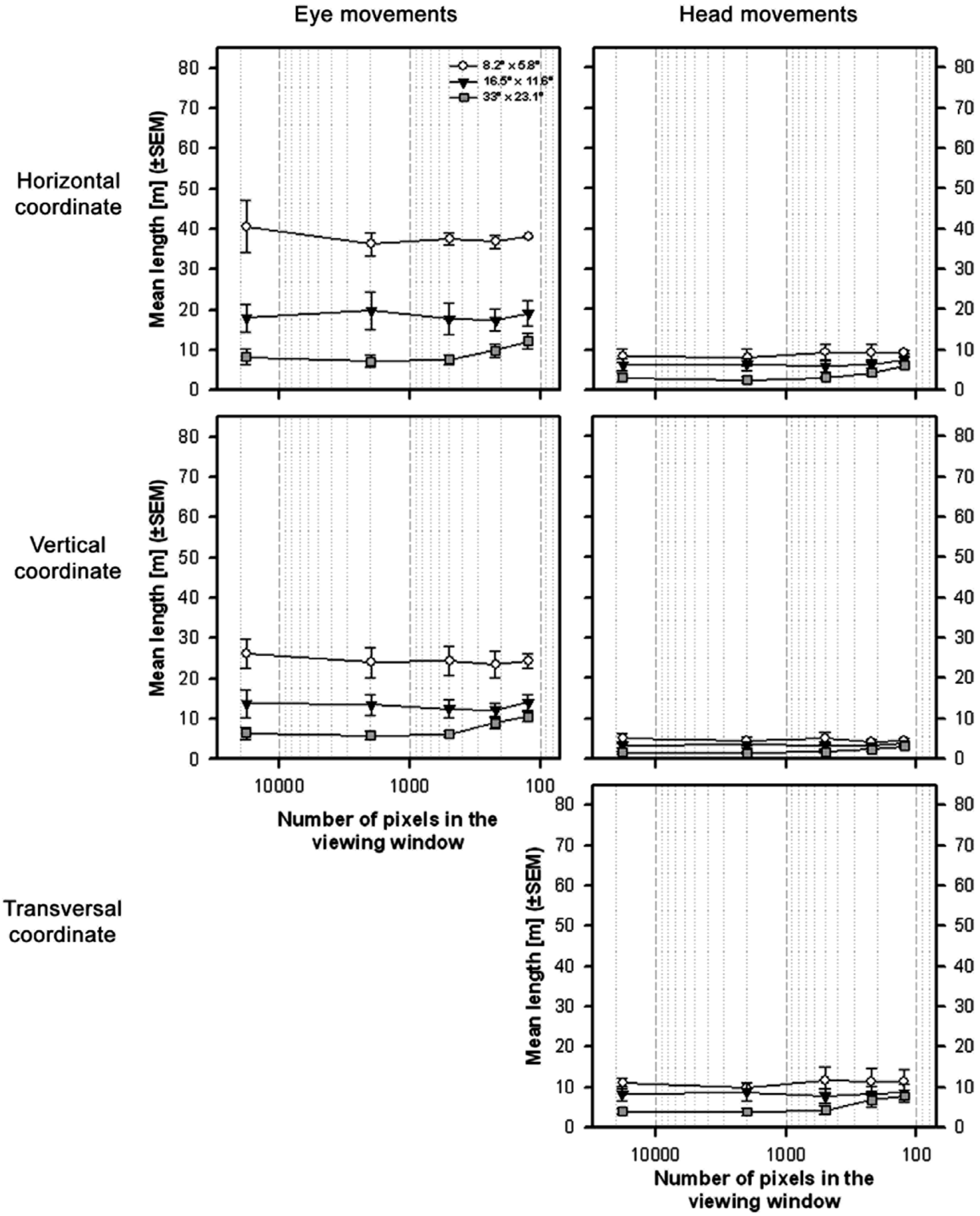

Fig. A1. Mean length of eye and head movements [m] ± SEM versus image resolution for three normal subjects performing the LEDs task. Three effective fields of view projected in the 10° × 7° viewing window are compared in central vision: 8.25° × 5.8° (empty circles), 16.5° × 11.6° (black triangles), and 33° × 23.1° (grey squares).

vertically) while the shortest values were obtained with the 33° × 23.1° field of view (about 9 m horizontally and 8 m vertically). Head movements for this task were relatively small. However, they were still significantly influenced by the size of the effective field of view (ANOVA: $F_{(2,30)} = 18.72$, $p < .0001$ for the horizontal coordinate; $F_{(2,30)} = 13.43$, $p < .0001$ for the vertical coordinate; and $F_{(2,30)} = 10.45$, $p < .001$ for the transversal coordinate), increasing as the effective field of view decreased. Image resolution did not have a significant effect on head movements, similar to what was observed for eye movements on this task.

A.4. Results for the CHIPs task

Fig. A2 displays the mean cumulative length of eye (left panel) and head (right panel) movements versus image resolution with each effective field of view, for the CHIPs task. Horizontal eye movements were significantly influenced by both image resolution (ANOVA: $F_{(4,30)} = 9.68$, $p < .0001$) and by the size of the effective field of view (ANOVA: $F_{(2,30)} = 4.21$, $p < .05$). Vertical eye movements were significantly influenced by image resolution (ANOVA: $F_{(4,30)} = 6.55$, $p < .001$), but the influence of the

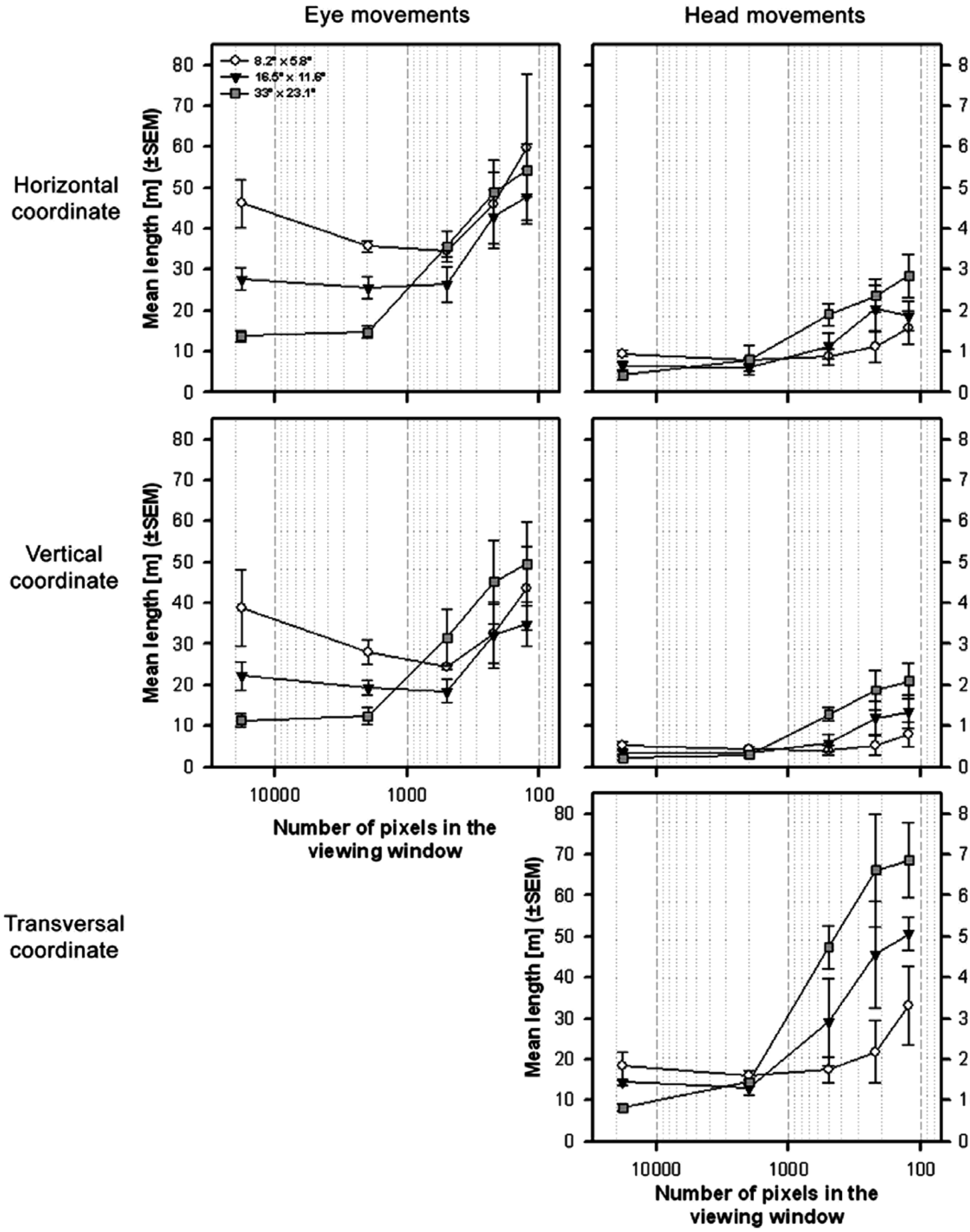

Fig. A2. Mean length of eye and head movements [m] ± SEM versus image resolution for three normal subjects performing the CHIPs task. Three effective fields of view projected in the 10° × 7° viewing window are compared in central vision: 8.25° × 5.8° (empty circles), 16.5° × 11.6° (black triangles), and 33° × 23.1° (grey squares).

effective field of view was not statistically significant. At the highest image resolution tested (17920 pixels), the largest eye movements were obtained with the 8.25° × 5.8° field of view (approximately 45 m horizontally and 40 m vertically) and the smallest with the 33° × 23.1° field of view (about 15 m horizontally and vertically). Head movements along the three axes were significantly influenced by image resolution (ANOVA: $F_{(4,30)} = 11.45$, $p < .0001$ for the horizontal coordinate; $F_{(4,30)} = 10.13$, $p < .0001$ for the vertical coordinate; and $F_{(4,30)} = 16.50$, $p < .0001$ for the transversal coordinate) and by the size of the effective field of view (ANOVA: $F_{(2,30)} = 4.45$, $p < .05$ for the horizontal coordinate; $F_{(2,30)} = 7.19$, $p < .01$ for the vertical coordinate; and $F_{(2,30)} = 9.24$, $p < .001$ for the transversal coordinate). At 17920 pixels, the largest head movements were obtained with the 8.25° × 5.8° field of view (approximately 9 m horizontally, 5 m vertically, and 18 m transversally), and the smallest with the 33° × 23.1° field of view (approximately 4 m horizontally, 2 m vertically, and 8 m transversally). The total length of head movements systematically increased as image resolution decreased. The most dramatic increases were observed for the 33° × 23.1° field of view, especially for the transversal coordinate. This clearly indicates that when resolution in the viewing window was artificially reduced, subjects compensated by moving their heads closer to the working area. This strategy also led to larger horizontal and vertical head movements, but to a minor extent.

A.5. Summary and conclusion

Altogether, quantitative measurements of eye and head movements during both tasks demonstrate that the visual search strategy was significantly influenced by the viewing conditions used. Three main observations can be drawn from these measurements:

1. With small effective fields of view, subjects explored the environment by using eye and head movements, similar to patients suffering from conditions resulting in "tunnel vision".
2. With large fields of view, subjects used fewer eye and head movements as long as resolution was sufficient to identify the target of interest.
3. At low resolution and when detailed form recognition was required, subjects moved their heads closer to the target (essentially with transversal head movements) to compensate for the lack of resolution.

References

Antes, J. R. (1974). The time course of picture viewing. *Journal of Experimental Psychology, 103*, 62–70.

Bertera, J. H. (1988). The effect of simulated scotomas on visual search in normal subjects. *Investigative Ophthalmology & Visual Science, 29*, 470–475.

Bertera, J. H., & Rayner, K. (2000). Eye movements and the span of the effective stimulus in visual search. *Perception & Psychophysics, 62*, 576–585.

Bingham, G. P., & Pagano, C. C. (1998). The necessity of a perception–action approach to definite distance perception: Monocular distance perception to guide reaching. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 145–168.

Chow, A. Y., Chow, V. Y., Packo, K. H., Pollack, J. S., Peyman, G. A., & Schuchard, R. (2004). The artificial silicon retina microchip for the treatment of vision loss from retinitis pigmentosa. *Archives of Ophthalmology, 122*, 460–469.

Cornelissen, F. W., Bruin, K. J., & Kooijman, A. C. (2005). The influence of artificial scotomas on eye movements during visual search. *Optometry & Vision Science, 82*, 27–35.

Dagnelie, G., Walter, M., & Liancheng, Y. (2006). Playing checkers: Detection and eye–hand coordination in simulated prosthetic vision. *Journal of Modern Optics, 53*, 1325–1342.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research, 52*, 317–329.

Desmurget, M., Pelisson, D., Rossetti, Y., & Prablanc, C. (1998). From eye to hand: Planning goal-directed movements. *Neuroscience and Biobehavioral Reviews, 22*, 761–788.

Dobelle, W. H. (2000). Artificial vision for the blind by connecting a television camera to the visual cortex. *ASAIO Journal, 46*, 3–9.

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America, 110*, 1150–1163.

Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science, 180*, 1194–1197.

Hayes, J. S., Yin, V. T., Piyathaisere, D., Weiland, J. D., Humayun, M. S., & Dagnelie, G. (2003). Visually guided performance of simple tasks using simulated prosthetic vision. *Artificial Organs, 27*, 1016–1028.

Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception & Psychophysics, 65*, 58–71.

Henderson, J. M., McClure, K. K., Pierce, S., & Schrock, G. (1997). Object identification without foveal vision: Evidence from an artificial scotoma paradigm. *Perception & Psychophysics, 59*, 323–346.

Hollingworth, A., Schrock, G., & Henderson, J. M. (2001). Change detection in the flicker paradigm: The role of fixation position within the scene. *Memory & Cognition, 29*, 296–304.

Hooge, I. T., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research, 39*, 1567–1575.

Humayun, M. S. (2001). Intraocular retinal prosthesis. *Transactions of the American Ophthalmological Society, 99*, 271–300.

Humayun, M. S., Weiland, J. D., Fujii, G. Y., Greenberg, R., Williamson, R., Little, J., et al. (2003). Visual perception in a blind subject with a chronic microelectronic retinal prosthesis. *Vision Research, 43*, 2573–2581.

Kerkhoff, G. (1999). Restorative and compensatory therapy approaches in cerebral blindness—A review. *Restorative Neurology and Neuroscience, 15*, 255–271.

Land, M., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception, 28*, 1311–1328.

Lecchi, M., Marguerat, A., Ionescu, A., Pelizzone, M., Renaud, P., Sommerhalder, J., et al. (2004). Ganglion cells from chick retina display multiple functional nAChR subtypes. *Neuroreport, 15*, 307–311.

Loftus, A., Murphy, S., McKenna, I., & Mon-Williams, M. (2004). Reduced fields of view are neither necessary nor sufficient for distance underestimation but reduce precision and may cause calibration problems. *Experimental Brain Research, 158*, 328–335.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 565–572.

Loudin, J. D., Simanovskii, D. M., Vijayraghavan, K., Sramek, C. K., Butterwick, A. F., Huie, P., et al. (2007). Optoelectronic retinal prosthesis: System design and performance. *Journal of Neural Engineering, 4*, S72–S84.

Magne, P., & Coello, Y. (2002). Retinal and extra-retinal contribution to position coding. *Behavioural Brain Research, 136*, 277–287.

Mazza, M., Renaud, P., Bertrand, D., & Ionescu, A. (2005). CMOS pixels for subretinal implant prosthesis. *IEEE Sensors Journal, 5*, 32–37.

Nelson, P., Aspinall, P., Papasouliotis, O., Worton, B., & O'Brien, C. (2003). Quality of life in glaucoma and its relationship with visual function. *Journal of Glaucoma, 12*, 139–150.

Nelson, W. W., & Loftus, G. R. (1980). The functional visual field during picture viewing. *Journal of Experimental Psychology: Human Learning and Memory, 6*, 391–399.

Nodine, C. F., Carmody, D. P., & Herman, E. (1979). Eye movements during visual search for artistically embedded targets. *Bulletin of the Psychonomic Society, 13*, 371–374.

Ohta, J., Tokuda, T., Kagawa, K., Sugitani, S., Taniyama, M., Uehara, A., et al. (2007). Laboratory investigation of microelectronics-based stimulators for large-scale suprachoroidal transretinal stimulation (STS). *Journal of Neural Engineering, 4*, S85–S91.

Owsley, C., McGwin, G., Jr., Sloane, M. E., Stalvey, B. T., & Wells, J. (2001). Timed instrumental activities of daily living tasks: Relationship to visual function in older adults. *Optometry & Vision Science, 78*, 350–359.

Paillard, J. (1982). The contribution of peripheral and central vision to visually guided reaching. In D. J. Ingle, M. A. Goodale, & J. W. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 367–385). Cambridge, MA: The MIT Press.

Parker, R. E. (1978). Picture processing during recognition. *Journal of Experimental Psychology: Human Perception and Performance, 4*, 284–293.

Pelz, J.B. (1995). Visual representations in a Natural Visuo-motor Task. PhD thesis. University of Rochester, New York, United States of America.

Pelz, J. B., & Canosa, R. (2001). Oculomotor behavior and perceptual strategies in complex tasks. *Vision Research, 41*, 3587–3596.

Pelz, J. B., Hayhoe, M., & Loeber, R. (2001). The coordination of eye, head, and hand movements in a natural task. *Experimental Brain Research, 139*, 266–277.

Pérez Fornos, A., Sommerhalder, J., Rappaz, B., Safran, A. B., & Pelizzone, M. (2005). Simulation of artificial vision, III: Do the spatial or temporal characteristics of stimulus pixelization really matter? *Investigative Ophthalmology & Visual Science, 46*, 3906–3912.

Purdy, K. A., Lederman, S. J., & Klatzky, R. L. (1999). Manipulation with no or partial vision. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 755–774.

Rayner, K., & Bertera, J. H. (1979). Reading without a Fovea. *Science, 206*, 468–469.

Rayner, K., & Pollatsek, A. (1992). Eye movements and scene perception. *Canadian Journal of Psychology, 46*, 342–376.

Richard, G., Hornig, R., Keseru, M., & Feucht, M. (2007). Chronic epiretinal chip implant in blind patients with retinitis pigmentosa: Long-term clinical results. *Investigative Ophthalmology & Visual Science, 48*, 666.

Rubin, G. S., Bandeen-Roche, K., Huang, G. H., Munoz, B., Schein, O. D., Fried, L. P., et al. (2001). The association of multiple visual impairments with self-reported visual disability: SEE project. *Investigative Ophthalmology & Visual Science, 42*, 64–72.

Salzmann, J., Linderholm, O. P., Guyomard, J. L., Paques, M., Simonutti, M., Lecchi, M., et al. (2006). Subretinal electrode implantation in the P23H rat for chronic stimulations. *The British Journal of Ophthalmology, 90*, 1183–1187.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303–304.

Sivak, B., & Mackenzie, C. L. (1992). The contributions of peripheral vision and central vision to prehension. In L. Proteau & D. Elliot (Eds.), *Vision and Motor Control* (pp. 233–259). Amsterdam: Elsevier Science Publishers.

Sommerhalder, J., Oueghlani, E., Bagnoud, M., Leonards, U., Safran, A. B., & Pelizzone, M. (2003). Simulation of artificial vision: I. Eccentric reading of isolated words, and perceptual learning. *Vision Research, 43*, 269–283.

Sommerhalder, J., Rappaz, B., de Haller, R., Perez Fornos, A., Safran, A. B., & Pelizzone, M. (2004). Simulation of artificial vision: II. Eccentric reading of full-page text and the learning of this task. *Vision Research, 44*, 1693–1706.

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *Journal of Speech and Hearing Research, 28*, 455–462.

Szlyk, J. P., Seiple, W., Fishman, G. A., Alexander, K. R., Grover, S., & Mahler, C. L. (2001). Perceived and actual performance of daily tasks: Relationship to visual function tests in individuals with retinitis pigmentosa. *Ophthalmology, 108*, 65–75.

Varsori, M., Pérez Fornos, A., Safran, A. B., & Whatham, A. R. (2004). Development of a viewing strategy during adaptation to an artificial central scotoma. *Vision Research, 44*, 2691–2705.

Veraart, C., Wanet-Defalque, M. C., Gerard, B., Vanlierde, A., & Delbeke, J. (2003). Pattern recognition with the optic nerve visual prosthesis. *Artificial Organs, 27*, 996–1004.

Watt, S. J., Bradshaw, M. F., & Rushton, S. K. (2000). Field of view affects reaching, not grasping. *Experimental Brain Research, 135*, 411–416.

West, S. K., Rubin, G. S., Broman, A. T., Munoz, B., Bandeen-Roche, K., & Turano, K. (2002). How does visual impairment affect performance on tasks of everyday life? The SEE project. Salisbury eye evaluation. *Archives of Ophthalmology, 120*, 774–780.

Yanai, D., Weiland, J. D., Mahadevappa, M., Greenberg, R. J., Fine, I., & Humayun, M. S. (2007). Visual performance using a retinal prosthesis in three subjects with retinitis pigmentosa. *American Journal of Ophthalmology, 143*, 820–827.

Ziegler, D., Linderholm, P., Mazza, M., Ferazzutti, S., Bertrand, D., Ionescu, A. M., & Renaud, P. (2004). An active microphotodiode array of oscillating pixels for retinal stimulation. *Sensors and Actuators A: Physical, 110*, 11–17.

Zrenner, E., Wilke, R., Zabel, T., Sachs, H., Bartz-Schmidt, K., Gekeler, F., Wilhelm, B., Greppmaier, U., & Stett, A.SUBRET Study Group. (2007). Psychometric analysis of visual sensations mediated by subretinal microelectrode arrays implanted into blind retinitis pigmentosa patients. *Investigative Ophthalmology & Visual Science, 48*, 659.