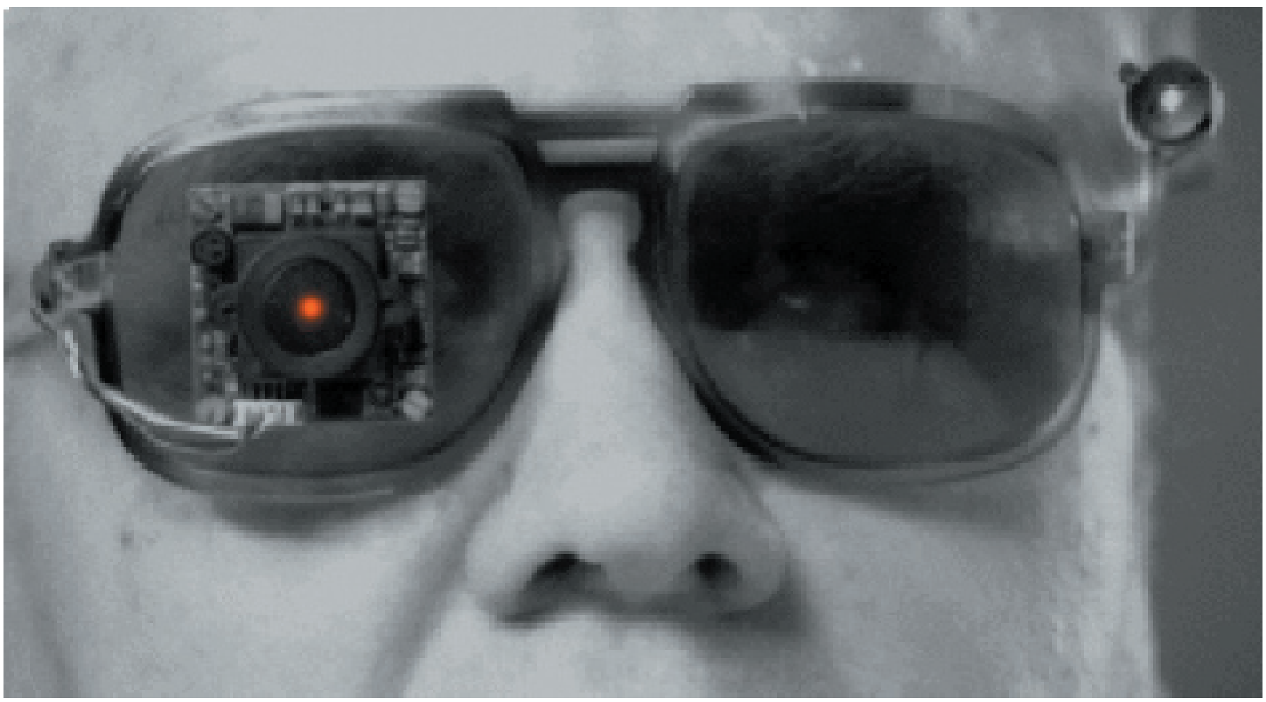
The Design and Feasibility of a Multimodal Prosthesis for the Blind



Gina Burkhart
Josh Schoenwald
Martin Voshell

June 2003

Introduction

What does it mean to be blind? Blindness refers to the loss (or severe degradation) of the ability to intake visual information through the eyes. This, in most cases has no bearing on the individual's ability to process visual information, just that they have a deficiency in getting that information to their visual cortex. Because the eyes are merely vessels of information and not processors, their loss does not signal a complete lack of ability to process visual cues. It therefore seems logical that blind individuals could take advantage of visual cues, as do sighted individuals, if such information was properly channeled to their brain. Because most blind people do not have such luxuries (not yet anyway), the remaining intact senses, principally the somatosensory and auditory senses, are built up to a level of heightened acuity to compensate for their lack of vision (Goldstein, 1999). We wish to explore the feasibility of introducing a multimodal device that can supplement certain limitations of the blind without interfering with the sensory systems that have already been developed.

In considering blindness, one must make the distinction between congenital blindness and late-blindness. In the former case, people exhibit a phenomenon known as sensory substitution, in which unused portions of cortex are mapped to the processing of stimuli from other areas. For example, when a congenitally blind person reads braille, her visual cortex is used to process the information (Burton, Snyder, Conturo, Akbudak, Ollinger, and Richle, 2002). Late blinded persons also exhibit this phenomenon, but to a far lesser extent. However, although their physiological compensation is not as developed (Heller, 1989), studies have shown that late-blinded people can integrate tactile information with prior visual experience enabling them subconsciously a better percept than blindfolded sighted individuals (Kennedy, 1993, 1997, Heller, 1989).

Blindness can be defined as a compromise where light cannot enter the eye and be processed. Legal blindness is defined simply in terms of visual acuity. An individual that needs a distance of 200 feet to resolve images that people with "normal" vision ("normal" meaning ability to correctly identify a set of letters at 20 feet) can resolve at 20 feet constitutes blindness (Goldstein, 1999). However this is not a very robust definition, for there are many types of visual deficiencies. A more functional definition of blindness is when visual information is too impoverished (or even altogether absent) to be processed to a degree that allows individuals to easily operate in society. Blindness may manifest in many forms. In an optical sense concerning the eye itself , there are a number of issues that can arise. These include: inadequate foveal receptors such as acuity and color-blindness, damaged peripheral vision (tunnel vision), corneal damage not letting in light, cataracts that cloud the lens, or any manner of retina damage. Effects of these can sever or greatly disrupt information leaving the eye. Beyond the eye, visual damage may occur in the processing path as well. The optic nerve can become damaged through glaucoma, and further the optic chiasm the entire visual cortex are subject to brain damage, aneurysm, stroke, etc. Poor foveal vision is the largest cause of legal blindness, however, retinal blindness (retinopathy and disease) is the most common form of severe blindness (Goldstein, 1999).

Review of Visual Perception

In class discussions we have covered the normal visual sensory input perception cycle in sighted individuals. Information in the form of light enters the eyes, the analog light signal is projected onto the retina and converted into chemical signals traveling back to the optic nerve, across the optic chiasm, and into the visual cortex, or occipital lobe. The hierarchical arrangement of the visual cortex is a massively parallel system where information is sent to specialized areas. In the blind, this is not entirely the case. In most congenitally blind individuals the visual cortex develops normally and consists of the same structure as in sighted individuals. Depending on the type of blindness and the extent of the damage information could still be coming into the cortex but not making its way out (blind sight is a case where we can respond to visual stimuli without having the feeling of actually seeing them) and with the advent of functional brain imaging, we can now make some rather spectacular insights into what might be going on (See figure 1).

Traditional approaches in aiding the blind generally involve forms of sensory substitution (Bach-y-Rita, 1972). Touch, and more specifically tactile feedback, are used extensively in place of vision. The success of the Braille alphabet is a testament not just to just how well tactile feedback can be transformed and interpreted as readable text, but also that it can be done at such significant speed (100 words per minute on average, as opposed to 200-300 for a typical sighted person) (Goldstein, 1999). The ability to associatively couple touch information with phonics and lexical information is quite bewildering. In sighted individuals, vision and touch exhibit what is known as intersensory dominance (Rock and Victor, 1964). When both vision and tactile information are giving shape information, vision tends to dominate in its processing. In the blind, there is no such conflict, and they can rely on touch quite readily. Different senses can be trained to interpret information, and many studies speak of heightened tactile and auditory awareness. For example, Bach-y-Rita (1972) used a device that mapped light and dark areas of a scene to a tactile device that simulated the scene on the individual's back, taking advantage of the large somatosensory area of feedback. A portable sensory substitution system based on Bach-y-Rita's work was developed to help the blind read text and perceive graphic displays (Goldstein, 1999). The Optacon (See figure 2) transforms printed letters or graphic images into patterns of vibrations that can be sensed through the fingers.

What is more striking though is what is going on in the brain of a non-sighted individual at such times (See figure 1). Many recent functional imaging investigations have shown that large areas of visual cortex are active when the blind are involved in tactile tasks (Giudice, Madison, Costello, Zhuang, Legge, Hu, He 2002). Brain re-organization and neural plasticity are quite evident. This phenomena has received significant attention more recently and investigators are curious as to whether tactile information is taking over areas of vision or if there are parts of visual cortex actively contributing to the blinds' cognition. Some argue that since Braille literacy is more of a lexical task, then it is not surprising that activity shows up in the small language portion of the left occipital lobe, so is this merely that area taking over more unused space around it? Giudice et. al. (2002) found that activation in the visual cortex differed significantly between sighted and blind individuals. When performing tactile tasks the blind participants showed activity in primary visual, extrastriate, and high level cortical areas, however sighted participants did not show specific patterns (re-emphasizing intersensory dominance). Another compelling result occurred when investigating tactile feedback: visual cortex responded to shapes and letters (roman and Braille) which elicited very specific high level activity (salience), but random tactile noise elicited no visual activity. The visual cortex in the blind is quite developed and inherently plays quite an active role in cognition and perception.

Handicapped Aids

Those people unfortunate enough to be stricken with blindness have a limited but somewhat powerful set of tools to help them function in our society. First is the "long cane" which is merely as it sounds – a long cane that aids in mobility by helping to locate structures and obstacles within range of the person. Blind people become surprisingly adept at maneuvering with these canes, after being trained to use them. The other principal aid is that of the guide dog. Guide dogs are trained at infancy to understand a multitude of commands and lead their owner safely through the world (Heyes, 2003). Commonly referred to as "seeing eye" dogs, these animals are truly intelligent cognitive agents. The dogs show remarkable abilities and adaptation, providing an amazing array of activities, such as:

- Keeping on a direct route, ignoring distractions such as smells, other animals and people

- Maintaining a steady pace, to the left and just ahead of the handler

- Stopping at all curbs until told to proceed

- Recognizing and avoiding obstacles that the handler won't be able to fit through (narrow passages and low overheads)

- Bringing the handler to elevator buttons

- Helping the handler to board and move around buses, subways, and other forms of public transportation

The guide dog has many obvious benefits to the simple 'long cane,' as the dog provides real-time feedback giving some information about the state of the world, specifically, such information that is relevant to the handler's tasks. Both handler and dog become so intimately related that the dog can often begin to infer the handler's intentions about where to go (through repetition) and respond accordingly.

But even with the cane and Seeing Eye dog, the blind are still at a significant disadvantage to those people with sight. These disadvantages are manifested in two main problems (Heyes, 2003): orientation with respect to the environment, and the detection of hazards not found by the use of the primary mobility aid. Guide dogs, as adaptive and intelligent as they may be, lack the ability to fully comprehend the world in a matter useful to the handler. For the most part, guide dogs act as a safe means of transportation, not a true "guide" in the sense of help in getting somewhere. The task of getting from point A to B still lies in the brain of the handler, who must have some mental model of the space he is moving so that he can make the requisite moves. The dog is just there to make sure he gets there safely. Even with the limitation that they cannot spot every hazard, (dogs and humans have inherently different ideas of what constitutes a hazard), at this task, the guide dog is still king, and likely better than any automation that could be invented. However, many of these devices require a considerable amount of adaptation. Late-blinded individuals are at more of a disadvantage because they have prior knowledge that can interfere with the manner in which these devices are meant to be used. This is where our design comes into play: to aid late-blinded patients by providing both visual and auditory information.

Phosphene Based Vision

In 1968 it was found that phosphenes, the tiny dots of light best described as the stars one sees when they hit their head, could be elicited and reproduced by direct electrode stimulation in the brain (Brindley and Lewin, 1968). This phenomenal evoking of phosphenes has been reproduced numerous times and is best described as being a point of light about the size of a quarter at arm's length away. In the 1970's, WM. H. Dobelle, spurred on by further studies identifying specialized visual centers in the brain turned his focus directly onto the primary visual cortex (Dobelle 1977). This area, V1, receives relayed data from the retina and stimulation of this area was known to elicit single points of light. Dobelle figured that with enough stimulation and phosphene sensations, he could pattern them in a scoreboard manner (See figure 3) to reproduce a limited depthless, colorless, tunnel-based, visual field (Dobelle, Turkel, Hendoerson, Evans, 1979).

The potential of such a prosthesis has gained much acceptance (Schmidt et al, 1996) but most pertinent are the people who already employ such a system (Dobelle, 2000). The very nature of the technology changes rapidly, and therefore, it would be intuitive to establish some global normals and perceptual functions in such a realm. Two specific experiments (Cha, Horch, and Normann a,b 1992) have been conducted simulating mobility performance in such a field as well as simple acuity tests. Very little has looked specifically at such complex and cognitive issues as orientation and mobility. Mobility entails aiding in hazard avoidance, object detection, course navigation, and obstacle avoidance. Orientation devices are newer and emphasize environmental tagging that is related in some manner, usually verbal, to the blind individual to help locate landmarks, signs, and interest (Crandall, Brabyn, Bentzen, Myers, 1999).

Our Design - Apparatus

To partially compensate the needs of the blind that are not adequately met through other means, we propose a multimodal set of devices to enhance visual and auditory cues. These enhancements will enable greater independent mobility (with or without a seeing eye dog or cane), more precise navigation, and attentional awareness enabling faster reading of text and graphic displays. Our design is intended for late-blinded individuals, those who have prior visual information about the world.

We begin with a set of microeelctrodes which are implanted into the subject's visual cortex (See figure 4). Each electrode stimulates an acute area of the brain which in turn elicits a single phosphene.

These phosphenes can be integrated to produce a matrix of dots that display a low-resolution projection in visual space. Cha et. al. (1992) demonstrated that with such a pixelated display, normal walking was attainable with 625 or more total pixels. This work was done trying to recreate an entire visual scene. Dobelle's work (2000) shifted towards edge detection rather than complete scene recreation, and with a current resolution of 512 pixels, today's electrodes are sufficiently compact to produce a navigable matrix of dots representing edges.

These electrodes are connected to an integrated module somewhere in the skull that then transmits information wirelessly (via bluetooth or 802.11g protocols) to a microcomputer worn on the person. With advances in microtechnology, both of these modules could be incredibly small and easily portable. This module would then in turn be connected to a set of glasses fitted with microcameras that take in the visual input (Similar to figure 5).

The system would also transmit audio cues through a stereo headset connected to the computer (Similar to the I-Glass device, figure 6). This would deliver information (natural language) that would complement the visual stimuli. The audio is explict and linked to the visual scene being displayed by the phosphenes. This means that other auditory cues in the environment are still heard at normal levels; the information coming through the headset is completely new information that only pertains to the visual scene. The audio and visual are intended to work in parallel, indeed to be integrated seamlessly. The power in this method is to associate a visual element with an audio component, sort of like when you see a car coming, you can also hear the car. We take this sort of redundancy for granted, but blind people are at a distinct disadvantage without this collaboration. As we can all attest, sometimes it is insufficient to attend to something by merely hearing it. In addition, the user is in complete control of when the system is used. The audio, as well as the camera, can be turned off and on as desired. By having complete control over the audio and visual portions of the interface, we have the unique ability to control and coordinate the stimuli. Additionally, if the person should require greater visual detail (or less) this is easily accomplished. This is an advantage that even sighted individuals do not possess.

Our Design - Multimodal Perspective

What else is out there?

A variety of electronic devices have cropped up in this field. Electronic Travel Aids (ETAs) are examples of technology that attempt to compensate in this area and range from simple object detectors, to personal handheld ultrasonic tactile and auditory machinery, to laser augmented canes that detect objects and warn of depth-related drop-offs. Many such devices do not use appropriate modality cues easily to lead to confusion and data overload. With the tactile inputs aside, many of these devices rely upon audio information exclusively and commonly have quite adverse effects by overloading the auditory sense which the blind are quite dependent upon. We are attempting to take a multimodal approach to effectively complement existing processes that does not interfere with other resources. Two challenges await us. First and foremost, multimodal interfaces tend to start with the clutter inherent in visual displays and are using alternative modalities to complement and *take the load off* the visual channel. In our case, we start *without* a visual channel and thus are attempting to use the same principles but in a slightly different manner. Second, a well-designed multimodal system that fuses two or more information sources can be an effective means of reducing recognition uncertainty (Oviatt, 2002).

The value of the visual information we propose is that it can give the blind person an external visual representation of the world. Blind people do have mental representations of their surroundings, to the degree possible with the use of long canes, guide dogs, and their physical sensations. While these may be adequate in some situations (especially tactile feedback when handling objects), this information is not as helpful when walking down the street or navigating an unfamiliar environment. The visual information available in our system would provide crude (but effective) outlines of objects (See figure 7) which would enable users to have a general idea of the types of obstacles around them. While not providing detailed information, this would increase the users' sense of awareness, especially in familiar environments. This heightened awareness would only make it easier for someone to get a 'feel' for where they are, increasing mobility significantly. This information can be seamlessly integrated with the use of a guide dog or long cane as well. Guide dogs are wonderful for avoiding most hazards and ensuring the safety of the handler, but they do not have an intrinsic sense of their environment as it relates to the goals of the handler. By giving the handler (user) a greater awareness of his surroundings, he can therefore greatly increase his benefit from his dog, since the dog is adept at getting the handler somewhere safely as long as the handler knows how to get there. As with the long cane, it too is incapable of giving the user a sense of his environment beyond what is within his reach. The user has little or no idea of anything beyond his immediate path, which in the absence of sufficient audio cues makes it difficult to avoid, much less anticipate, hazards in the world.

The integration of audio signals in addition to the visual signals is a critical element. Expressing audio cues to blind persons can significantly increase their ability to walk in a straight line (Leonard). Alfred Leonard's study was to walk behind blind people and whisper in their ear information about where they're travelling. Doing this allowed the people to walk in a straight line with increased confidence. This study demonstrates an important point: that the audio information becomes important when it is presented in conjunction with visual information. When Leonard made the whispers to the blind people, he was doing so based on his visual assessment of the scene before him. The importance of integrating the audio and visual information should be realized in certain configurations. Spence and Driver (1997) found that audio cues can be anticipated and reacted to much quicker if they are presented in the same general area as complementary visual cues. Our design gives us the power to correlate such stimuli with incredible precision, just as would be the case if the invidual were able to see on his own. Another important justification for the audio cues comes from the finding that audio is the preferred medium for presenting time-critical warning and alerts, especially when such alerts require immediate attention (Spence and Driver, 1997). Finally, audio cues presented in concert with visual information can create anticipation for events, which increases response time. This would be especially useful in cases of unexpected emergencies, when quick decision making and action is crucial. The ability to anticipate is further increased due to the fact that auditory stimuli do not require attentional focus to be perceived. For instance, a person might miss an ambulance if they're not looking at it, but it's not likely they'll miss the siren, no matter what else they happen to be doing. This can also be enhanced if the stimuli are continually presented in the same medium i.e. if warnings are always presented acoustically (Spence and Driver, 1997).

Our Design - Limitations

Although the multimodal design has the potential to be a great aid to the blind, there are some limitations that accompany this new technology. For instance, much experimentation goes into mapping the electrodes to the correct parts of the brain. If the electrodes are not mapped correctly, the images will not appear in the correct shapes. In studies (Schmidt, et. al, 1996), many anomalies and strange interactions resulted when many surface electrodes were fired in the cortex. New evidence suggests that promise lies in stimulation at the level of thalamic input. It requires much less electricity to elicit a phosphene at this level and there is potentially a much greater area to stimulate (Schmidt, et. al, 1996). The type of blindness is also an issue. If the person is blind due to occipital lobe damage, phosphenes will not be elicited. As well, our design will not work with congenitally blind individuals, due to the necessary neural mapping that does not occur in these peopl. Another issue comes with the hardware. Previous research using phophene-based vision involved an incision in the back of the head. The electrodes were connected to the visual cortex, but the wires were exposed through the incision. Leaving the person to function with wiring coming out of his head raises the potential for infection, as well as the shifting of electrodes. The new design has tried to work around these problems by utilising a wireless transmission from the electrodes to the computer. Even though the mecahnism is wireless, it is still susceptible to software failures which could render the entire system temporarily inoperable. Despite the research that has been accomplished, the feasibility of making the system work correctly in a crowded wireless environment also poses potential for problems, many of which could manifest themselves in the future.

Another significant limitation deals with depth perception. The system is not able to present images in their true three-dimensional glory. The lack of depth perception is surely a handicap. Some depth can be percieved through the relative size of the images in comparison to each other, but true depth perception is not possible with this system. Assuming that the system can be successfully implemented and is in fact helpful to the blind, further research should be done to expand its capabilities. To improve depth perception, it might be feasible that a tactile addition could be utilized. Using tactile feedback to gauge depth instead of audio (which quickly becomes intrusive), might enable the blind to have a sense of space as well as objects. This would make it easier for the blind to navigate through environments.

Summary

Although someone is blind, it does not mean that information processing centers are useless. If the information centers are able to be used, phophene vision along with auditory cues and normal aids (long cane, dogs, etc.) may further assist the blind. Our design is not without limitations, however the potential advantages stand to aid the blind in a much *brighter* future.

References

Bach-y-Rita, P. (1972). *Brain Mechanisms in sensory substitution*. New York: Academic Press

Brindley GS, Lewin WS: The sensations produced by electrical stimulation of the visual cortex. J. Physiol., 1968, 196: 479-493.

Burton, H, Snyder, Z, Conturo, T. E., Akbudak, E, Ollinger, J. M., and Richle, M. E (2002). Adaptive Changes in Early and Late Blind: A fMRI Study of Braille Reading. J Neurophysiol, 87, 589-607.

Cha, K. Horch, K. Normann, R. (1992, Jan 15). Mobility Performance with a Pixelized Visual System. Vision Research, 1367-72

Crandall, W., Brabyn, J., Bentzen, B., & Myers, L. (1999). Remote infrared signage evaluation for transit stations and intersections. *Journal of Rehabilitation Research and Development*. 36, 341-355.

Dobelle, WH. (1977, Sept). Current Status of Research on Providing Sight to the Blind by Electrical Stimulation of the Brain. Visual Impairment and Blindness, 290-297

Dobelle, WH. Turkel, J. Henderson, D. Evans, J. (1979, Oct) Mapping the Representation of the Visual Field By Electrical Stimulation of Human Visual Cortex. American Journal of Ophthalmology, 727-735

Dobelle, WH. (2000) Artificial Vision for the Blind by Connecting a Television Camera to the Visual Cortex. American Society of Artificial Internal Organs, 3-11.

Giudice, N.. Madison, P., Costello, J. C., Zhuang, G. E., Legge, X., Hu, S., He, S. (2002) Tactile vision: neuroimaging and brain-reorganization in the blind; implications for learning and adaptive-technology. Proceedings of Technology and Persons with Disabilities Conference

Goldstein, E. Bruce. (1999). *Sensation & Perception*. Brooks Cole Publishing, Pacific Grove, CA.

Heller, M. A. (1989). Picture and pattern percetion in the sighted and the blind: The advantage of the late blind. Perception, 18, 379-389

Heyes, Tony. (2003). Electronic Travel Aids- Why Bother? [Online]. Available: http://users.bigpond.net.au/heyes/pa/quest_ge.htm (2003, May 29)

Kennedy, J. M. (1993). Drawing and the blind. New Haven, CT: Yale Univesity Press.

Kennedy, J. M. (1997, January). How the blind draw. Scientific American, 76-81

Oviatt, S. (2002). Breaking the Robustness Barrier: Recent Progress on the Design of Robust Multimodal Systems. In M. Zelkowitz (Ed.), Advances in Computers, Academic Press.

Rock, I., & Victor, J. (1964). Vision and touch: An experimentally created conflict between the senses. Science, 143, 594-596

Schmidt, E. Bak, M. Hambrecht, F. Kufta, C. O'Rourke, D. Vallabhanath, P. (1996) Feasibility of a visual prosthesis for the blind based on intracortical microstimulation of the visual cortex. Brain, 507-522

Spence, C. and Driver, J. (1997). Cross-modal links in attention between audition, vision, and touch: Implications for interface design. International Journal of Cognitive Ergonomics, 1(4), 351-373.

Figure Captions

FIgure 1: Excerpted fMRI image from Giudice, et. al. showing significant activity in a blind subjects' visual cortex

Figure 2: Optacon incorporates OCR with tactile Braile feedback

Figure 3: Early letter mappng in a subject's phosphene field

Figure 4: X-Ray image of microelectrode array

Figure 5: Dobelle's apparatus on an individual

Figure 6: Digital I-Glasses

Figure 7: Comparison of pre-processed video and and post-processed edge detection

Figures

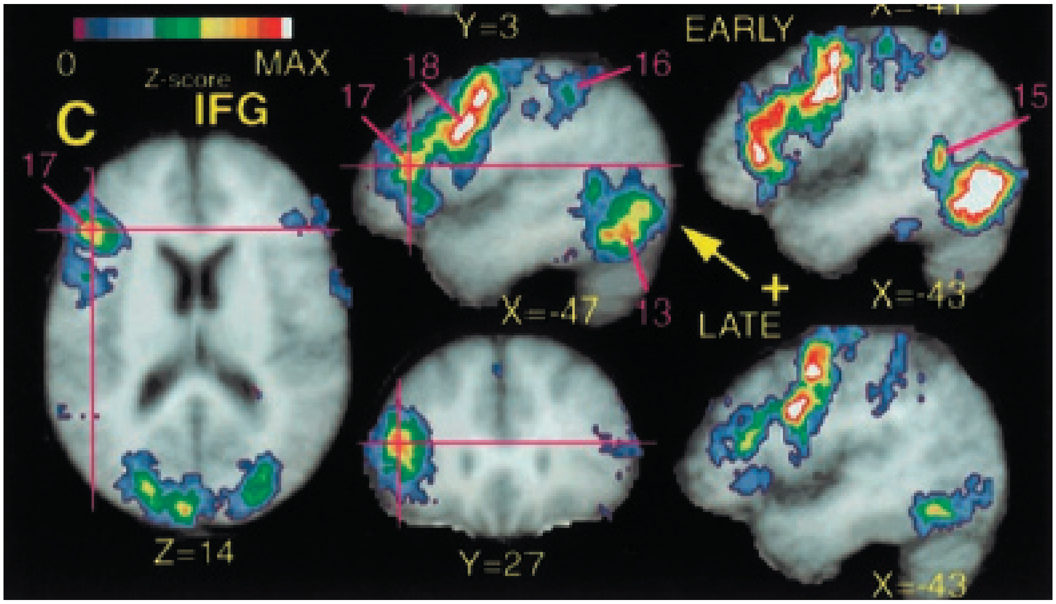


FIgure 1: Excerpted fMRI image from Giudice, et. al. showing significant activity in a blind subjects' visual cortex

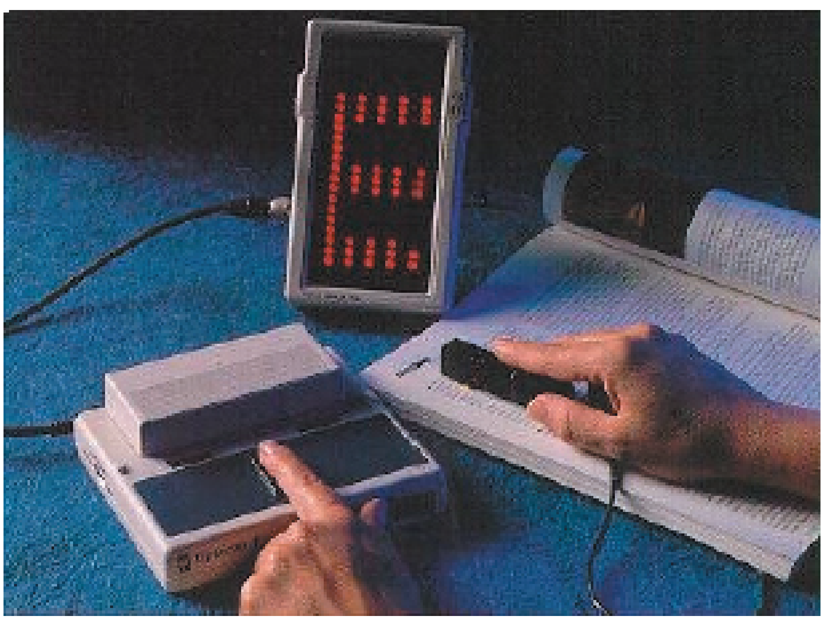


Figure 2: Optacon incorporates OCR with tactile Braile feedback

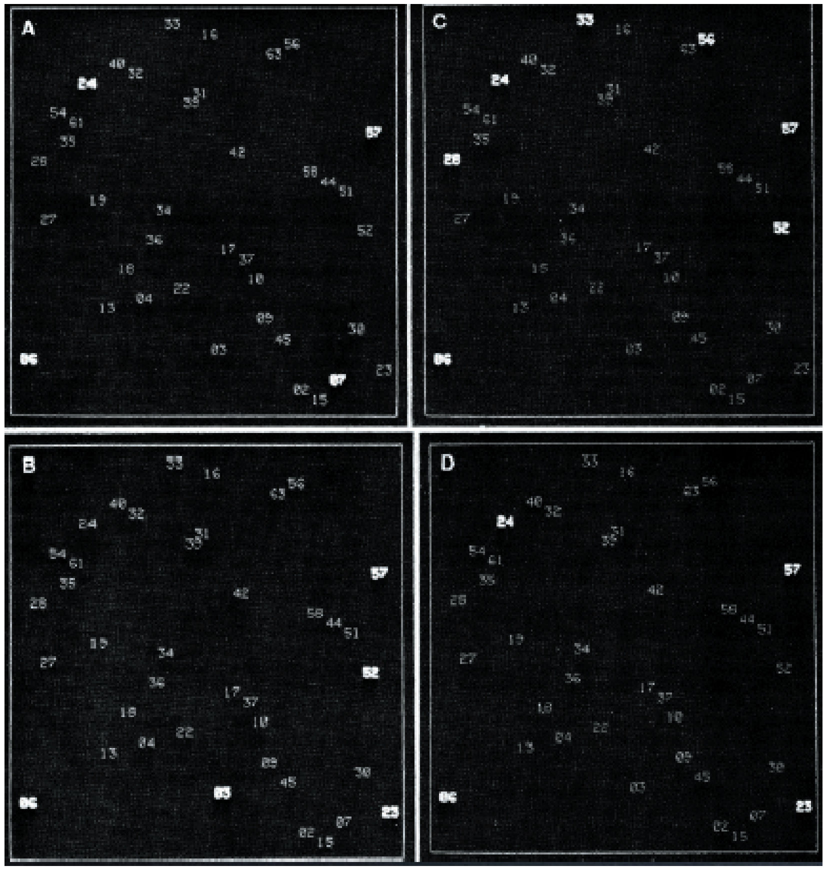


Figure 3: Early letter mappng in a subject’s phosphene field

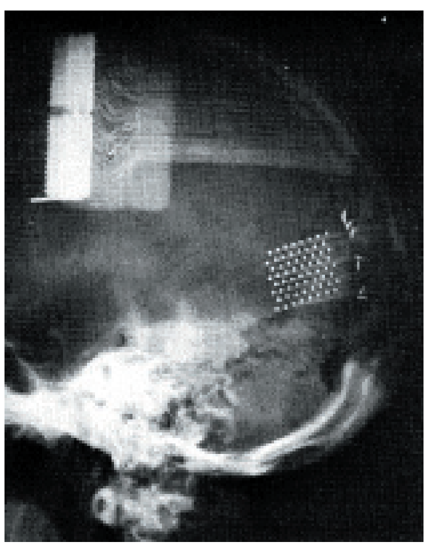


Figure 4: X-Ray image of microelectrode array

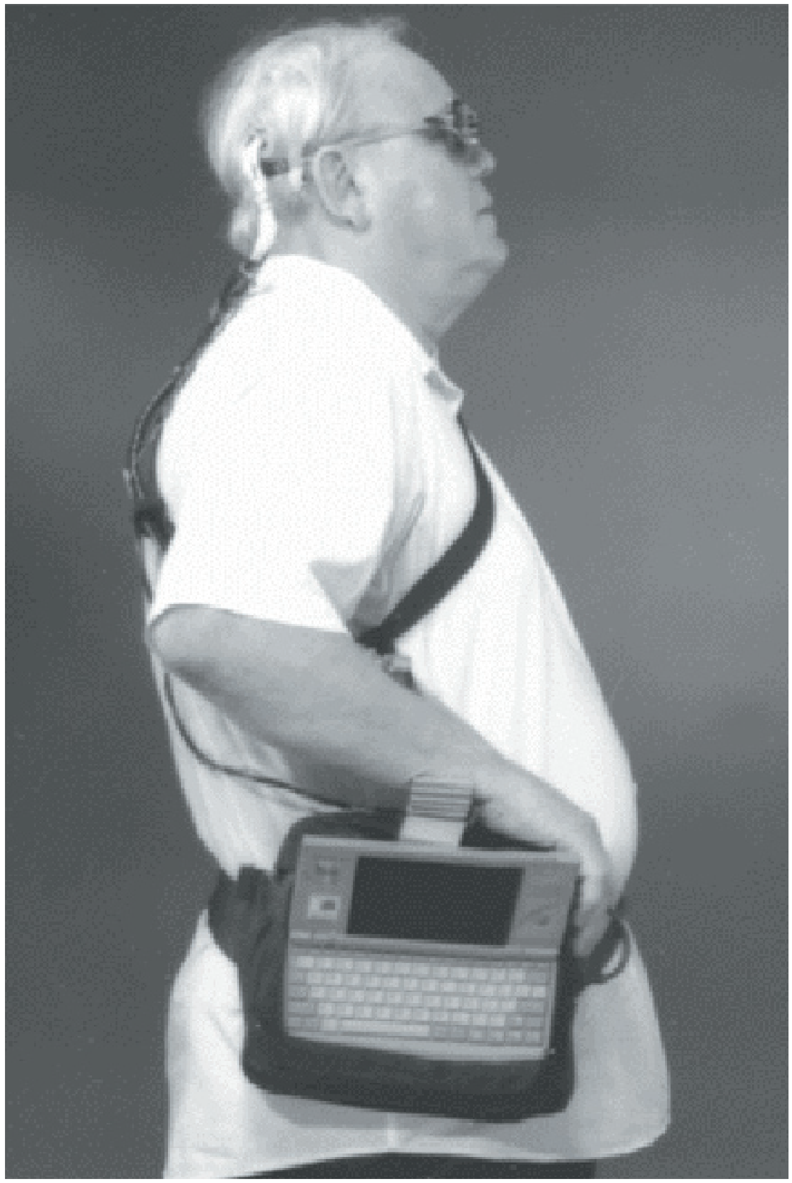


Figure 5: Dobelle's apparatus on an individual

Figure 6: Digital I-Glasses

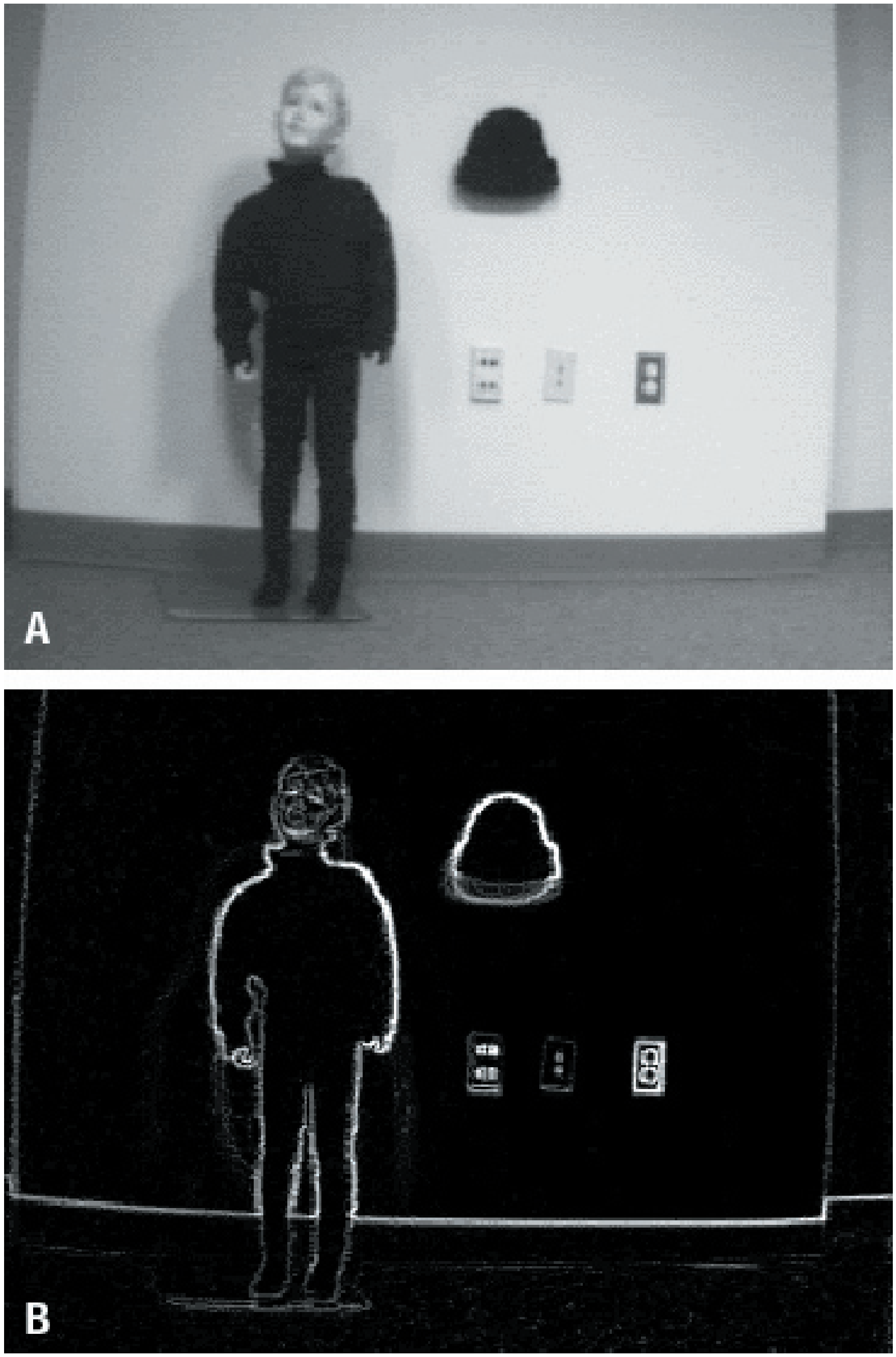


Figure 7: Comparison of pre-processed video and and post-processed edge detection