Wearable Tetherless Computer-Mediated Reality: WearCam as a wearable face-recognizer, and other applications for the disabled

Steve Mann, N1NLF
steve@media.mit.edu
MIT Media Lab, Building E15-389, 20 Ames Street, Cambridge, MA02139

Abstract

'WearCam', a wearable multimedia system with video processing capability and wireless Internet connection, has recently been proposed(Mann 1994b). In this paper, WearCam is presented as a prosthetic device. In particular, two example applications: the '**personal visual assistant**'; and the '**visual memory prosthetic**' are described. The '**personal visual assistant**' embodies a *spatial visual filter*(Mann 1994a) that reconfigures the human visual system, providing a *coordinate transformation* (remapping of spatial coordinates). Such coordinate transformations, it is hoped, might someday be of use to the partially sighted. The '**visual memory prosthetic(?)**' embodies a *temporal visual filter* that provides computer-induced *flashbacks* (possibly together with annotation). These 'flashbacks' currently help the author overcome *visual amnesia*. It is hoped that, with further research, the apparatus and approach might someday lead to perceptual intelligence that we can wear, and be of great benefit to the disabled.

Introduction

'WearCam' is a portable, and in fact wearable, system that has many of the features found on a desktop multimedia computer: one or more cameras, microphones, high speed video processing and communications link, as well as its own Internet address. These items are miniaturized and placed within the wearer's clothing and the like, rather than on the desktop. The connectivity is wireless (the details of the communications channel are described in(N1NLF 1996)), using hat-mounted antennas, and battery power, so that the user has tetherless online connectivity. The apparatus is worn in a natural fashion, so that it can be used, at times, without (much) conscious thought or effort, and while performing other activities.

Historical notes

Ivan Sutherland, a pioneer in computer graphics, described a head-mounted display with half-silvered mirrors so that the wearer could see a virtual world superimposed on reality (Earnshaw, Gigante, & Jones 1993) (Sutherland 1968). Sutherland's work, as well as more recent related work (Feiner, MacIntyre, & Seligmann Jul 1993)(Feiner, MacIntyre, & Seligmann 1993)(Fuchs, Bajura, & Ohbuchi)(Drascic 1993)(Milgram 1994)is characterized by its tethered nature. The wearer is tethered to a workstation which is generally powered from an AC outlet.

The tetherless 'wearable multimedia'system (Fig 1(1980)) was designed and built by the author as a tool for 'personal imaging', with the goal of attaining an enhanced sense of visual awareness and producing visual art (Ryals 1995).

The early apparatus, housed in a heavy welded-steel-frame backpack, weighed more than fifty pounds, and could only run for a short time per charge. However, 'wearable multimedia' evolved to the extent of being worn comfortably in a small waist bag (Fig 1(1990)), or even sewn into a vest or the like (Fig 1(1995)). Currently a version of the 'wearable multimedia' apparatus is being built into a normal pair of eyeglasses, running from a Walkman-sized belt pack. The current apparatus is worn during normal day-to-day activities, such as walking to/from the office, waiting in line at the bank, shopping, etc..

Recently, the system became better known as the 'Wearable Wireless Webcam', when, with the advent of the World Wide Web, experiments in visual connectivity and shared visual memory were started (Mann 1994b), although these experiments were not central to the 'wearable multimedia' effort.

Social acceptance is important in the design of any prosthetic device; the hope is that miniaturization capabilities of modern technology will downsize it to the same order as devices such as hearing aids and regular eyeglasses. Attitudes toward various forms of the author's 'wearable multimedia' systems have significantly changed over the last fifteen years. In particular, it is now possible to wear the apparatus in many everyday situations where it would have been completely out of place just a few years ago. Through a combination of changes in the apparatus (its having become much less obtrusive, thanks to improvements in technology allowing for miniaturization), and changes in society (increase in society's acceptance of technology), it is

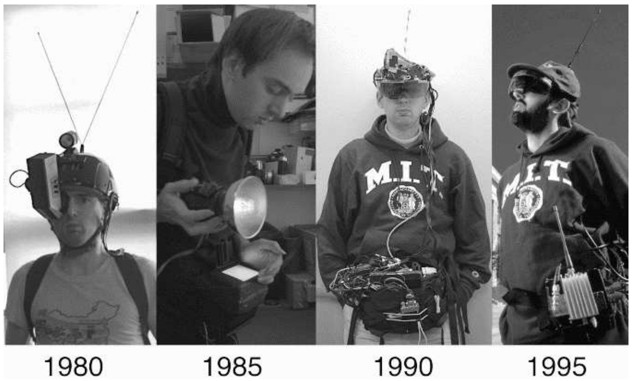


Figure 1: **'Wearable multimedia'** systems designed and built by author: Early (1980) apparatus comprised separate radios for inbound and outbound communications paths. A bulky 1.5 inch CRT, requiring a heavy bicycle helmet for support, was used for the display with 40 characters per line. The system was found to be too cumbersome, so a waist-mounted television was next (1985) adopted as the display, leaving both eyes unobstructed. With the advent of miniature CRTs in the late 1980s, a comfortable eyeglass-based system became practical, and was later transferred to a more modern visor. A single hat-mounted antenna provided communications in the ham bands. Presently (1985), cellular communications (antenna in hat) provide base-level (low-speed) data communications channels when the unit is too far from its home base to use the high speed ham radio unit (pictured here on waist).

not nearly as out-of-place as it was just a few years ago. Privacy issues associated with 'wearable multimedia/personal imaging' have also been addressed(?).

Efforts by others have been directed toward wearable computing(Bass 1985)(?)(Finger *et al.* 1996), but without the wearable, tetherless video capability.

The goal of this paper is to propose 'wearable multimedia/personal imaging' and its potential use for either a spatial visual filter or a temporal visual filter. The spatial visual filtering capability is presented as the Personal Visual Assistant, and then the temporal visual filter is presented as the Visual Memory Prosthetic.

'Personal Visual Assistant (PVA)'

The first of the two goals of this article is to propose use of the spatial filtering capability of a wearable-tetherless computer-mediated reality environment as an assistant to the partially sighted.

The apparatus is worn over the eyes, and has the capability to, in real time, computationally augment, diminish, or otherwise alter visual reality(Mann 1994a).

PVA Background: Transformation of the perceptual world

In his 1896 paper (Stratton 1896), George Stratton reported on experiments in which he wore eyeglasses that inverted his visual field of view. Upon first wearing the glasses, he reported seeing the world upside-down, but, after an adaptation period of several days, was able to function completely normally with the glasses on.

Figure 2: **Wearable-tetherless computer-mediated reality** as of late 1994, showing a color stereo head-mounted display (VR4) with two cameras mounted to it. The inter-camera distance and field of view match approximately my interocular distance and field of view with the apparatus removed. The components around the author's waist comprised primarily wireless communications equipment. Antennas, etc. are located at the back of the head-mount to balance the weight of the cameras, so that the unit is not front-heavy.

Dolezal (Dolezal 1982) (page 19) describes "various types of optical transformations", such as the *inversion* explored by Stratton, as well as *displacement*, *reversal*, *tilt*, *magnification*, and *scrambling*. Kohler (Kohler 1964) also discusses "transformation of the perceptual world".

Stratton, Dolezal, and Kohler explored the use of optics (lenses, prisms, and the like). Stuart Anstis was the first to explore, in detail, an electronically mediated world. Anstis (Anstis 1992), using a camcorder that had a "negation" switch on the viewfinder, experimented with living in a "negated" world. He walked around holding the camcorder up to one eye, looking through it, and observed that he was unable to learn to recognize faces in a negated world. His negation experiment bore a similarity to Stratton's inversion experiment mentioned in Sec , but the important difference within the context of this article is that Anstis electronically mediated his visual world — he experienced his world as a video signal.

Using a camcorder as a *reality mediator* has several drawbacks. Firstly, it is awkward (one hand is occupied constantly, and the apparatus protrudes far enough that it gets in the way of most day-to-day activities), and secondly, it makes people feel much more self-conscious Thirdly, it is not easy to insert arbitrary computational power between the camera and the viewfinder.

WearCam provided a practical solution to these three problems, by serving as a wearable, tetherless 'reality mediator'. A suitable realization of WearCam, made from a battery-powered color stereo display, having 480 lines of resolution, is depicted in Fig 2. Here, the author mounted the cameras the correct interocular distance apart, and used cameras that had the same field of view as the display devices. With the cameras connected directly to the displays, an *illusion of transparency* (Mann 1994a) will be realized to some degree, at least to the extent that each ray of light entering the apparatus (e.g. absorbed and quantified by the cameras) will appear to emerge at roughly the same angle (by virtue of the display).

Although the apparatus provided no depth-from-focus capability there was enough depth perception remaining on account of the stereo disparity for the author to function somewhat normally with the apparatus.

The use of head-mounted displays for helping the visually handicapped, using the contrast adjustments of the video display to increase apparent scene contrast has been recently explored(jhu 1995), but without the use of computational processing of the imagery. The approach described in this paper (Fig 2) also contains computational capability, and therefore extends and generalizes that recently described in (jhu 1995).

A first step in using the wearable-tetherless reality mediator was to wear it for a while to become accustomed to its characteristics. Unlike in typical beam-splitter implementations of *augmented reality*, transparency, if desired, is synthesized, and therefore only as good as the components used to make the apparatus.

The apparatus was worn in the identity map configuration (cameras connected directly to the displays) for several days, in order to adapt to its imperfections and irregularities (the identity map is never fully achieved). It was found that one could easily walk around, up and down stairs, through doorways, etc.. Some difficulties, however, were experienced, in scenes of high dynamic range, and also in reading fine print (for example, in a restaurant, where the menu was located behind a counter, preventing a close-up look, or in a department store, where the receipt was printed in faint ink).

The attempt to create an illusion of transparency was itself a useful experiment because it established some working knowledge of what can be performed when vision is *diminished* or *degraded* to RS170 resolution and field of view is somewhat limited by the apparatus.

Video mediation

The compute-power required to perform general-purpose manipulation of color video streams was, at the time of the experiments (1994) too unwieldy to be worn in comfortable clothing (although more recently, a clothing-based apparatus has been constructed to facilitate general-purpose reality mediation without relying on broadband wireless communications). In particular, a remotely-located compute-engine was used by establishing a full-duplex video communications channel between the wearable apparatus and the host computer(s). A high-quality communications link (called

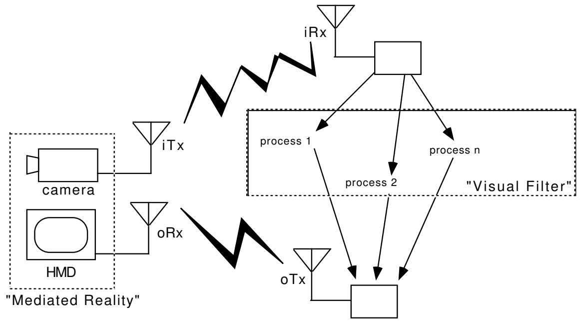


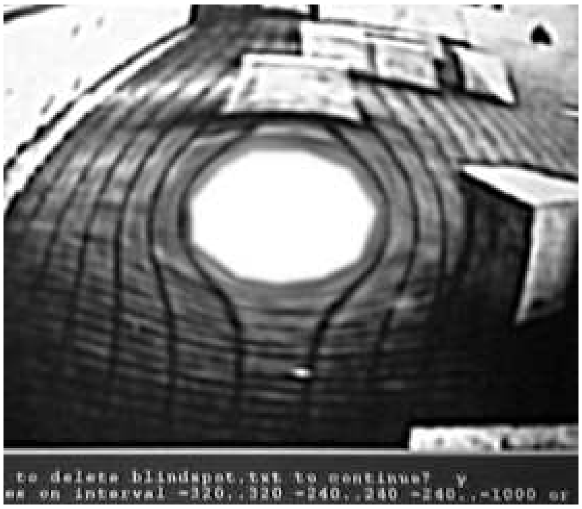
Figure 3: **Implementation of tetherless computer-mediated reality for use as a personal visual assistant.** The camera sends video to one or more more computer systems over a high-quality microwave communications link, called the 'inbound channel'. The computer system(s) send back the processed image over a UHF communications link called the 'outbound channel'. Note the designations "i" for inbound (e.g. iTx denotes inbound transmitter), and "o" for outbound. The term 'visual filter' refers to the process(es) that mediate(s) the visual reality and optionally insert *virtual* objects into the reality stream.

the 'inbound-channel') was used to send the video from the camera(s) to the remote computer(s), while a lower quality communications link (called the 'outbound channel') was used to carry the processed signal from the computer back to the head-mounted display (HMD). This apparatus is depicted in Fig 3. Ideally both channels would be of high-quality, but the machine-vision algorithms were found to be much more susceptible to noise than was the wearer's own vision. Originally, communication was based on antennas that the author had installed on various rooftops, but presently, the mediation may also be achieved completely on local (worn) processors.

The apparatus (Fig 2) permitted one to experience any coordinate transformation that could be expressed as a mapping from a 2D domain to a 2D range, and the apparatus could do this in real time (30frames/sec = 60fields/sec) in full color, because a full-size remote processing engine was used to perform the coordinate transformations. This apparatus allowed experiments with various computationally-generated coordinate transformations to be performed both indoors and outdoors, in a variety of different practical situations. Examples of some useful coordinate transformations appear in Fig 4.

In order for the 'personal visual assistant' to be useful, it will need to be small, lightweight, tetherless, and unobtrusive. Many of the design issues have already been dealt with, but much remains to be done. For example, although the current body-worn multimedia system has the capability to mediate the video stream locally, eye tracking capability built right into the head-mounted display would greatly advance the research toward a system that will hopefully someday be of widespread benefit to the visually challenged.

(a)



(b)

Figure 4: **Living in coordinate-transformed worlds:** Color video images are transmitted, coordinate-transformed, and then received back at 30 frames per second – the full frame-rate of the VR4 display device. (a) This 'visual filter' could someday allow a person with very poor vision to read (due to the central portion of the visual field being hyper-foveated for a very high degree of magnification in this area), yet still have good peripheral vision (due to a wide visual field of view arising from demagnified periphery). (b) This 'visual filter' could some day allow a person with a *scotoma* (a blind or dark spot in the visual field) to see more clearly, once having learned the mapping. Note the distortion in the cobblestones on the ground and the outdoor stone sculptures.

The 'visual memory prosthetic'

The 'visual memory prosthetic' is the second application of WearCam described in this article. While the PVA was based on spatial visual filtering, the 'visual memory prosthetic' is based on temporal visual filtering.

'Edgertonian' eyes

Early experiments with a variety of different *visual filters* were described in(Mann 1994a). Each of these filters provided a different visual reality. It was found that by applying a repeating freeze-frame effect to WearCam (with the cameras' own shutters set to 1/10000 second), the video *sample and hold*, caused nearly periodic patterns to appear to freeze at certain speeds. For example, while looking out the window of a fast-moving car, periodic railings that were a complete blur without the apparatus would snap into sharp focus with the apparatus, while slight differences in each strut of the railing would be highly visible as characteristic patterns that would move about rapidly. Looking at airplanes in flight, the number of blades on a spinning propeller would often be readily discernible, and, depending on the sampling rate of the apparatus, the blades would appear to rotate slowly backwards or forwards, in much the same way as objects do under the stroboscopic lights of Harold Edgerton(Edgerton 1979). By manually adjusting the processing parameters of the apparatus, one could see many things that would escape normal vision. It became evident that the temporal visual filter could function as a prosthetic device to improve vision.

Flashbacks and freeze-frames

Of greater interst than just being able to see things one would otherwise miss, was the fact that sometimes the effect would cause one to remember certain things much better. The author found, for example, that faces would be remembered much better with this *freeze-frame* effect. Often the frozen image would linger in visual memory much longer than the moving one. There is current evidence to support the existance of a *visual memory*(Kosslyn 1994), so it became apparent that one could build a 'visual memory prosthetic'.

Deja vu

A free-running visual memory prosthetic might in general indicate to the wearer that he or she has seen a particular object before. For example, it might alert the wearer wandering around in circles that he or she has been in that particular place before. In some sense it could function as the visual equivalent of Starner's rememberance agent(Starner 1993) (a text interface that constantly watches what the user types and automatically reminds the user of related text files).

Salient stack

We've all no doubt been lost at one time or another, but some suffer from 'visual amnesia' much more than others.

Like unraveling a thread to avoid getting lost in a maze, one way the author overcomes visual amnesia and avoids getting lost in large shopping complexes, airports, or the like, is by accumulating a stack of images, consciously deciding to capture an image at each branch point. A current implementation uses the Wearable Wireless Webcam: each image is appended to a World Wide Web page, and then when it is time to find the way back, the Web browser (Mosaic) is invoked.

Alternate implementations have been built using a local image stack but the Web stack allows for shared applications (e.g. another person such as a spouse can also start up the the Web browser and quickly catch up, by observing the path taken from an agreed-upon or otherwise known starting point).

The 'wearable face-recognizer'

Many people (author included) have difficulty remembering faces. However, the author has found that faces can be remembered much better by using computer-induced 'flashbacks' (short visual stimuli — perhaps as short as a single frame — that either prime the memory subconsciously, or consciously but as a background or secondary task).

The wearable multimedia apparatus is easily and quickly reconfigurable, under program control. One of the many modes of operation of WearCam was such that after meeting someone for which there is a desire to form an associative memory between name and face, a picture of a person's face was presented to 'flashback' 1 minute later, 2, 4, 8, 16, etc. minutes later. The author found that this repetition greatly improved the ability to remember the face. Furthermore, it was also found that it was much easier to learn the names associated with faces that had been 'flashed back' periodically, even though neither the names of those faces, nor those of a control-group that were not 'flashed back' were known until after the exposure phase of the experiment was completed.

Humans are quite good at recognizing faces, but computers are also good at recognizing faces. Previous work in the computer-face-recognition community is based on using a fixed camera (Turk)(Pentland *et al.* 1993). Face recognition work has not previously been directed toward wearable cameras, but, instead seems more aimed at video surveillance with a fixed camera and people moving through its field of view.

An object of this paper is to propose the 'wearable face recognizer'. The use of a wearable face-recognizer suggests the possibility of turning the tables on the traditional third-person perspective (such as a ceiling-mounted surveillance camera), and, instead, using face recognition from a first-person perspective. In particular, the apparatus may be used as a prosthetic device for those suffering from visual amnesia, or even those with visual impairment, who are unable to see the face (or see it clearly enough to recognize it).

In researching the best form of the 'wearable face recognizer' the author tried a variety of implementations. These ranged from storing the database of candidate faces on the body-worn apparatus, to connecting remotely from the apparatus to the database (because WearCam is connected to the Internet, any database accessable from the Internet may, at least in principle, be used).

In one implementation, faces were captured using an early version of the 'wearable multimedia' system, running under KA9Q's Network Operating System (NOS). Images of candidate faces were transmitted to a workstation-class computer over the inbound channel, while the name associated with the closest match was received back over the outbound channel.

In order to display the identity to the wearer, an enlarged font was used (this would enable a visually handicapped person to still be able to read it).

The simplest implementation involved assigning a new filename to the image of the candidate face to convey the identity of the subject. This required no additional software development because it just so happened that the image display program created a very large font display of the filename while the image was being "painted" down the screen (the early 'wearable multimedia' system was quite slow at displaying pictures, providing plenty of time for the wearer to read the filename — the identity of the face). Two successive frames of the video going into the author's right eye are shown in Fig 5.

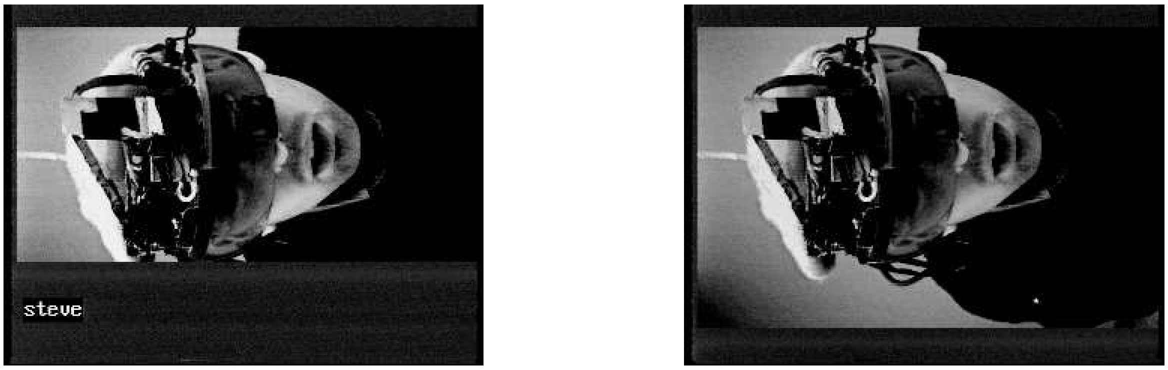
The 'wearable face recognizer' may be either 'free running', or in 'query mode'. In the former, the system captures images continuously and attempts to prompt the wearer, (inserting a name into the reality stream) whenever it can. In the latter, the wearer initiates each query.

Several different implementations of the capture/display configuration (e.g. having both the camera and display rotated 90 degrees, having the camera rotated 90 degrees with the display still in *landscape* orientation, etc) were tried and tested. It was found that the best overall configuration was to have the camera rotated 90 degrees (*portrait*) but with the display still in *landscape* orientation.

Improvements to the 'wearable face recognizer' included providing means of alignment, using a *registration template* (Fig 6). This made use in the 'query mode' much simpler and more precise: The author would would wait until the candidate happened to be facing toward the camera, then center the face on the computer screen by tilting the apparatus (the orientation of the apparatus, of course, can be controlled by head movements) and then press the "trigger". The trigger was a pushbutton switch connected to one of the eight lines on the parallel port of the computer (actually a full chording keyboard can be easily implemented by using 7 such switches).

Such a switch may, for example, be shoe-mounted (as in the roulette computers described by Thomas Bass (Bass 1985)) for unobtrusive use, or attached to a belt or waist bag within easy reach.

The author experimented with a lenses having a variety of different focal lengths, and found that the focal



(a) (b)

Figure 5: **Annotated computer-induced flashbacks** Two frames of video sequence entering author's right eye. The identity of the face is found to be the author (in actual fact, standing in front of a mirror hence this figure also depicts the apparatus used). Note the 90 degree rotation of the image which serves two purposes: (1) to match the aspect ratio of the human face to the camera, and (2) to create a distinct dual-adaptation space. (a) The very large font is quite readable. (b) The text remains on the screen until covered by the last few rasters of the image which displays slowly enough that the name can be easily read before being replaced by the facial *freeze-frame*.

length of 11 millimeters with a 1/3inch CCD provided the most optimum tradeoff between image stability and reach. The longer focal lengths were found to be harder to aim, while the shorter focal lengths were found to require that the apparatus be so close to the candidate as to invade the personal space of the candidate. In particular, in situations where the candidate was someone behind a physical barrier, such as the deli counter at the grocery store or the returns desk in a department store, the 11mm lens provided enough *throw* to reach across such a barrier.

Why are the faces rotated 90 degrees? (dual adaptation spaces)

The 'personal imaging' apparatus allowed the author to repeat the classic experiments like those of Stratton and Anstis (e.g. living in an upside-down or negated world), as well as some new experiments, such as learning to live in a world rotated 90 degrees.

It was observed that visual filters differing slightly from the identity (e.g. rotation by a few degrees) had a more lasting 'relative-aftereffect' (after removal of the apparatus) than visual filters that were far from the identity (e.g. rotation by 180 degrees = "upside-down"). By 'relative-aftereffect', It is meant the degree to which one is incapacitated upon removal of the apparatus, compared to the severity of the mapping. Obviously a rotation by a degree or two does not have much of an effect, while rotating the image 180 degrees has a much more profound effect on one's ability to perform tasks. Furthermore, the visual filters close to the identity tended to leave a more pronounced (relative) opposite aftereffect. (For example, one would consistently reach too high after taking off the apparatus where the images had been translated down slightly, or reach too far 'clockwise' after removing the apparatus that had been rotating images a few degrees counter-clockwise.)

Visual filters far from the identity (such as reversal or upside-down mappings) did not leave an opposite aftereffect, confirming other research with respect to the upside-down glasses(Stratton 1896)(Stratton 1897).

The 'visual memory prosthetic' was based on a partially mediated reality(Mann 1994a), that is, only part of the visual field of view was mediated, in this case, with the computer-induced (and sometimes annotated) flashbacks.

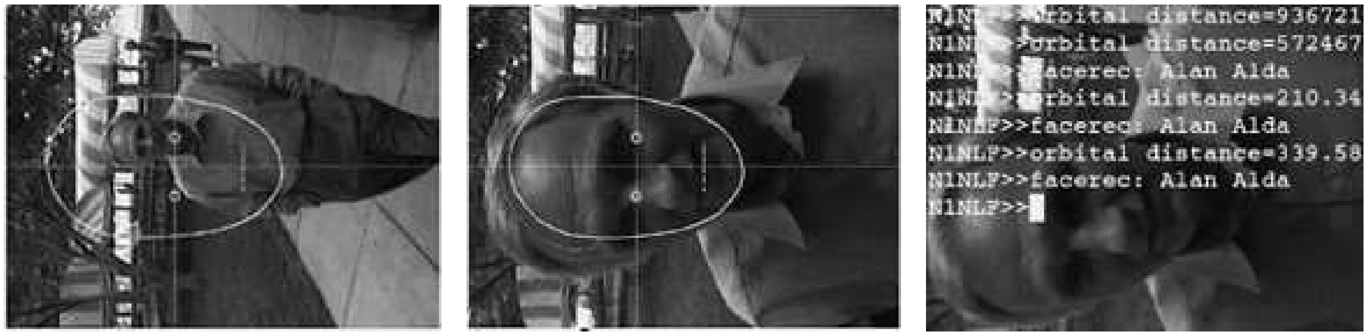
The reason for using the 'rot90' (rotate 90 degrees) arrangement was twofold: firstly this matched the aspect ratio of the face (which is generally taller than it is wide, in fact it is generally about 4 units high and 3 wide which exactly matches 'rot90' video), and this created a distinct dual adaptation space.

When two (or more) adaptation spaces were distinct, for example, in the case of the identity map (unmediated zones of the glasses) and the rotation operation ('rot 90'), It was possible to sustain a dual adaptation space and switch back and forth between the 'portrait' orientation of the identity operator and and 'landscape' orientation of the 'rot 90' operator without one causing lasting aftereffects in the other.

Conclusions on the visual memory prosthetic

Deliberate computer–induced *flashbacks* were explored as a means of assisting those (author included) with visual amnesia. Two modes of operation were presented, free-running flashbacks (requiring no input or attention from the user), and user-controlled flashbacks. The use of annotated flashbacks was also explored, in particular, through the implementation of a wearable face-recognition apparatus.

The 'visual memory prosthetic' begins to enlarge the scope of the concept of 'memory', for it is now possible to 'remember' something that one never knew in the first place. One might 'remember' the name of someone



(a) (b) (c)

Figure 6: **Template-based wearable face-recognizer** (a) As candidate approaches, an effort is made to orient the apparatus (by turning of the head) so that the candidate is centered. This is easy because the full-motion color video input stream appears on the computer screen together with the template. (b) At some point, the distance to the candidate will be such that the scale (size of the face on the image plane) will be appropriate, and, while still keeping the orientation appropriate, the match is made. (c) After the match is made, the template image drops away, revealing a radioteletype (RTTY) window behind it, upon which is displayed the desired information (for example, the name of the candidate, "Alan Alda", and possibly additional parameters or other relevant information).

one has never met before, provided that someone else has enrolled that person into a face database, and the boundaries between seeing and viewing, and between remembering and recording will begin to diminish.

Conclusions

With further research, the wearable, tetherless, computer-mediated reality apparatus (WearCam) will hopefully be of assistance to the visually handicapped, as well as those suffering from visual amnesia. These two applications were presented in the form of a personal visual assistant (spatial visual filter) and a visual memory prosthetic (temporal visual filter). The former reconfigures vision through a completely mediated reality, while the latter reconfigures vision through a partially mediated reality.

Acknowledgements

Research supported, in part, by HP labs, Palo Alto. The author would like to thank Thad Starner for helping getting the generalized spatial 'visual filter' software running on the SGI Reality Engine. Thanks also to: Roz Picard; Neil Gershenfeld; Sandy Pentland; Ted Adelson; Jennifer Healey; and many others, too numerous to mention here, for many interesting and important discussions; and Matt Reynolds, KB2ACE for help in upgrading the outbound ATV system. Thanks to Chris Barnhart for processing hardware, and to HP labs, VirtualVision, Compaq, Ed Gritz, BelTronics, M/A-Com, ShieldWorks, Addenda Electronics, and Virtual Research for lending or donating additional equipment that made the experiments possible.

References

Anstis, S. 1992. Visual adaptation to a negative, brightness-reversed world: some preliminary observations. In Carpenter, G., and Grossberg, S., eds., *Neural Networks for Vision and Image Processing.* MIT Press. 1–15.

Bass, T. 1985. *The Eudaemonic Pie.* Boston: Houghton Mifflin Company.

Dolezal, H. 1982. *Living in a world transformed.* Academic press series in cognition and perception. Chicago, Illinois: Academic press.

Drascic, D. 1993. David drascic's papers and presentations. http://vered.rose.utoronto.ca/people/david_dir/ Bibliography.html.

Earnshaw, R. A.; Gigante, M. A.; and Jones, H. 1993. *Virtual reality systems.* Academic press.

Edgerton, H. E. 1979. *Electronic flash, strobe.* Cambridge, Massachusetts: MIT Press.

Feiner;
MacIntyre; and Seligmann. 1993. Karma (knowledge-based augmented reality for maintenance assistance). http://www.cs.columbia.edu/graphics/projects/karma/ karma.html.

Feiner; MacIntyre; and Seligmann. Jul 1993. Knowledge-based augmented reality. Communications of the ACM, 36(7).

Finger, S.; Terk, M.; Subrahmanian, E.; Kasabach, C.; Prinz, F.; Siewiorek, D.; Smailagic, A.; Stivorek, J.; and Weiss, L. 1996. Rapid design and manufacture of wearable computers. *Communications of the ACM* 63–70. issue on Computer Science in Manufacturing.

Fuchs, H.; Bajura, M.; and Ohbuchi, R. Teaming ultrasound data with virtual reality in obstetrics. http://www.ncsa.uiuc.edu/Pubs/MetaCenter/SciHi93/ 1c.Highlights-BiologyC.html.

1995. Lions vision research and rehabilitation center. http://www.wilmer.jhu.edu/low_vis/low_vis.htm.

Kohler, I. 1964. *The formation and transformation of the perceptual world*, volume 3 of *Pshchological issues*. 227 West 13 Street: International university press. monograph 12.

Kosslyn, S. M. 1994. *Image and Brain The Resolution of the Imagery Debate*. M.I.T. Press.

Mann, S. 1994a. 'mediated reality'. TR 260, M.I.T. Media Lab Perceptual Computing Section, Cambridge, Ma.

Mann, S. 1994b. Personal WWW page: wearable wireless webcam. http://wearcam.org.

Milgram, P. 1994. Augmented reality: A class of displays on the reality-virtuality continuum. http://vered.rose.toronto.edu/people/paul_dir/SPIE94/SPIE94.full.html.

N1NLF. 1996. 'wearstation': With today's technology, it is now possible to build a fully equipped ham radio station, complete with internet connection, into your clothing. To Appear, CQ-VHF.

Pentland, A.; Starner, T.; Etcoff, N.; Masoiu, A.; Oliyide, O.; and Turk, M. 1993. Experiments with eigenfaces. In *Appears: Looking At People Workshop, IJCAI'93*.

Ryals, C. 1995. Lightspace: A new language of imaging. *PHOTO Electronic Imaging* 38(2):14–16. http://www.novalink.com/pei/mann2.html.

Starner, T. 1993. The remembrance agent. Class project for intelligent software agents class of Patti Maes.

Stratton, G. M. 1896. Some preliminary experiements on vision. *Psychological Review*.

Stratton, G. M. 1897. Vision without inversion of the retinal image. *Psychological Review*.

Sutherland, I. 1968. A head-mounted three dimensional display. In *Proc. Fall Joint Computer Conference*, 757–764.

Turk, M. A. *Interactive-Time Vision: Face Recognition as a Visual Behavior*. PhD dissertation, MIT, Media Arts and Sciences Section. Advisor: A.P. Pentland.