Continuous Lifelong Capture of Personal Experience with EyeTap

Steve Mann
Dept. of Electrical and Computer Engineering
University of Toronto
Toronto, Canada
mann@eecg.utoronto.ca

ABSTRACT

I begin with the argument that continuous archival of personal experience requires certain criteria to be met. In particular, for continuous usage, it is essential that each ray of light entering the eye be collinear with a corresponding ray of light entering the device, in at least one mode of operation. This is called the EyeTap criterion, and devices meeting this criterion are called EyeTap devices. Secondly, I outline Mediated Reality as a necessary framework for continuous archival and retrieval of personal experience. Thirdly, I show some examples of personalized experience capture (i.e. visual art). Finally, I outline the social issues of such devices, in particular, the accidentally discovered inverse to surveillance that I call "sosuveillance". It is argued that an equilibrium between surveillance and sousveillance is implicit in the archival of personal experiences.

Categories and Subject Descriptors

I.4.0 [**Image Processing and Computer Vision**]: General; C.3 [**Computer Systems Organization**]: Special–Purpose and Applications–Based Systems—*signal processing systems*; J.5 [**Computer Applications**]: ARTS AND HUMANITIES—*Fine arts*

General Terms

Design, Experimentation, Performance, Theory, Verification, Algorithms

Keywords

Computer vision, computer mediated reality, surveillance, inverse surveillance, sousveillance, oversight, undersight, survey, sousvey, equiveillance, perveillance, weblog, cyborglog, concomitant cover activity, eyetap, terrorism, guerrorism, audit, vidit, auditor, viditor

CARPE'04, October 15, 2004, New York, New York, USA.

1. THE AUDITOR/VIDITOR ANALOGY

Imagine if government tax auditors prohibited you from keeping your own records of your own financial activities. And suppose that the auditors themselves had very accurate records of your activities. Now if the only records of your life's activities were in the hands of someone else, would it be reasonable for you to be held accountable to remember your actions, without being permitted the ability to collect or keep evidence in support of your own case or legal argument?

This is exactly what's at issue when Starbucks clerks, or New York City Transit Authorities tell you that you're not allowed to take pictures. I argue that in the future we will likely have not only a right, but possibly even a responsibility to keep our own lifelong cyborg log ("lifeglog") of our personal experiences. And in the same way we've moved from the aural tradition to the alphabit/alphabyte of written literary records, giving rise to audits and auditors, we may soon see the emergence of "vidits" and **au/vi**/ditors where it will be seen as highly unusual if we have any breaks in the continuity of our lifeglogs.

In this paper, I will describe not only the technologies I have invented for lifelong cyborglogging, over the past 30 years, but also what it means to actually use these inventions in the real world. In particular, these inventions are part of an iterative process of invention, design, building, using, and then back to reinvention, redesign, rebuilding, and re-using, and so on....

2. EYETAP DESIGNS FOR CONTINUOUS LIFELONG CAPTURE OF PERSONAL EXPERIENCE

EyeTap is an experience capturing system. EyeTap devices cause the eye to, in effect, function as if it were both a camera and display, by mapping an effective camera and display inside the eye.

I first discuss the evolution of this technology over the last 30 years, together with new EyeTap designs. I also discuss some of the various applications that immediately arise from the use of such technology and the various new practices that are made possible with EyeTap technology.

Driven by a personal desire to explore new ways of see-

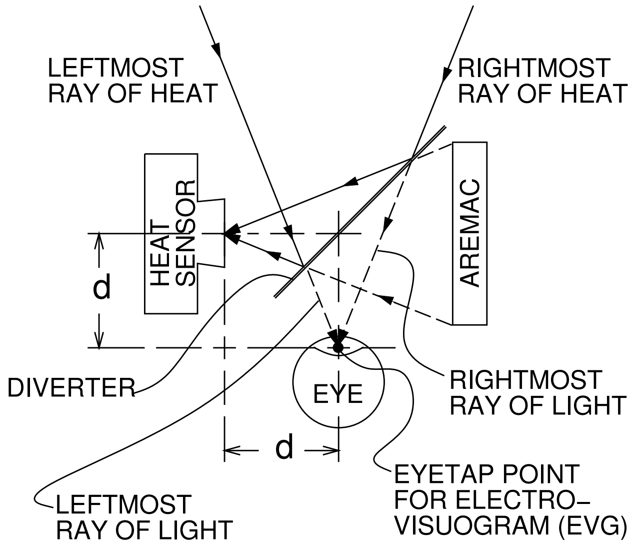

Figure 1: This unique arrangement allows EyeTap to capture and analyze rays of light passing through the eye, modify them, and resynthesize them.

ing, EyeTap devices have been invented, designed, built and used over the last 25 years. Long term usage in day-to-day life has led to greater insights into some of the issues that would normally have not been discovered in a controlled lab setting. It is this **continuous lifelong** usage that has required the development of devices that go beyond merely a wearable camera and display.

2.1 Invention of electric eyeglasses

Traditional (optical) eyeglasses are limited to modifying light by refraction, whereas next generation eyeglasses, called "EyeTap" devices, can also modify light computationally. With electric eyeglasses, in the future, instead of having to get new lenses ground, our eyeglass prescriptions might be downloaded over the Internet. There will also be new forms of visual correction not possible with optics.

EyeTap devices work by co-locating three things:

1. the center of projection of a scene analysis device such as a camera;
2. a view into a scene synthesis device such as an aremac;
3. the center of projection of the eye itself.

EyeTap devices allow for the creation of a "mediated reality", because of the ability to modify the light that passes through the eye. Such modification can be to change visual content (such as filtering out advertising) or more simply to see better in various ways. For example, such devices may be used to see in complete darkness, by simply mapping each ray of eyeward bound heat into a collinear and corresponding ray of eyeward bound light, as shown in Fig 1.

Traditional lens technology has allowed various optical disorders to be corrected and has allowed some of our visual abilities to be extended. Sunglasses and welding glasses, for example, allow us to see very bright scenes. Magnifying lenses such as a jeweler might use, allow us to see very small objects. More recently, devices have been made to allow us to see in the dark although they are often bulky and dangerous to wear because they cause unnecessary disorientation (e.g. because they consist of a separate camera and display, rather than the eye-centered design like EyeTap).

EyeTap eyeglasses merge all of these different visual aids into a single device, and consequently perform in situations where each one of these seeing aids used alone will not suffice. Consider driving on a country road at night. In this scenerio, one may benefit from using night-vision technology, however the headlights on an occasional oncomming could be blinding. The use of an EyeTap device allows a person to see in the dark without the danger of being blinded by the occasional flash of light, since the world is partially mediated by the EyeTap, but also can remain visible unmediated if desired.

Also, for those who suffer from more complicated visual deficiencies, EyeTap devices allow for the use of computational methods to correct the deficiency.

The more important features of EyeTap devices however, have no analog in traditional eyewear. EyeTap eyeglasses can help us remember better, through what is called a lifeglog (**life**long cyborg**glog**) or simply **'glog**, for short. A 'glog uses lifelong video capture to record what our eyes see over our entire lifetime. By using the data management capabilities of modern computers, we will be able to recall things that we have seen with perfect clarity in a natural and intuitive way. Having an on-demand photographic memory can help all of us by offloading, to a wearable computer, the task of memorizing now-mundane details that might only later become important. This kind of **visual memory prosthetic** is very beneficial to all of us, since our environments have become so overloaded with information. This lifeglog can also be used to increase personal safety and crime reduction by providing visual evidence for criminal acts, and to allow for trusted third party **inverse surveillance** ("sousveillance") in situations where the user may feel threatened. Moreover, in settings where surveillance already exists, sousveillance (the recording of an activity by a participant in the activity) can help prevent the surveillance recordings from being taken out of context. It is this contextual integrity of the evidence, combined with a personal right and responsibility of individuals to preserve evidence, that sets forth an equilibrium between surveillance and sousveillance.

EyeTap devices can also take an active role in helping us to filter our visual world so salient information stands out from the background clutter of visual advertising detritus. Imagine a world without advertisements! By using advanced computational methods, EyeTap technology has been used to remove the annoying visual propoganda that plagues our urban environments.

Since EyeTap devices can function as a computer display, they allow users to merge cyberspace with the real world. This feature is tremendously important since we are becoming increasingly bound to our computers due to our reliance on the internet as both an information and a social resource. As long as we are forced to focus our attention on a single physical object (such as a computer terminal, PDA or cellular phone), we will only feel more encumbered by computer and telecommunication technology as time passes. Therefore the EyeTap represents a liberating tool, to free us from the confines of attention-demanding computer terminals.

2.2 History of eyetap designs

Since my childhood in the 1970s I have been inventing, designing, building, and wearing computer systems for the creation of electronically mediated vision. The creation of

EyeTap followed the practice of designing these wearable systems. The EyeTap and wearable telephone/computer was and still is an important wearable input/output device satisfying the two most used senses of experiencing life as well as interacting with our surroundings, i.e. sight and sound.

The evolution of these inventions is shown in Fig 2

The drive to miniaturization led me to create, in 1995, devices having a completely normal appearance. However, subsequent realization that covertness stigmatizes the activity, led me to no longer see covertness as essential.

To explore new concepts in imaging and lighting, I designed and built the wearable personal imaging system. Originally, a Cathode Ray Tube (CRT, 1.5in, 5kV) on the helmet presented both text and graphics (including images), and a wearable light source helped me find my way around in the dark. I also carried an electronic flash lamp that let me explore, as a form of visual art, how subject matter responded to light.

Over the years development of smaller and smaller personal imaging systems has allowed for the EyeTap to shrink to an acceptable size: acceptable in terms of wearability and weight.

With the advent of consumer camcorders, miniature CRTs became available, making possible the early to mid 1980s electric eyeglass design as shown in Fig 3.

Later I used a 0.6-inch 6kV CRT facing down (angled back to stay close to the forehead). This apparatus was later transferred to optics salvaged from an early 1990s television set (Fig 2, early 1990s). Though still somewhat cumbersome, the unit could be worn comfortably for several hours at a time. An Internet connection through the small hat-based whip antenna used TCP/IP with AX25 (the standard packet protocol for ham radio operators).

These wearable computer reality mediators have evolved from headsets of the 1970s, to EyeTaps with optics outside the glasses in the 1980s, to EyeTaps with the optics built inside the glasses in the 1990s to EyeTaps with mediation zones built into the frames, lens edges, or the cut lines of bifocal lenses in the year 2000 (e.g. exit pupil and associated optics concealed by the transition regions between glass and frame, or within the frame). For example, in one such design, the computational element of the EyeTaps is incorporated into the eyeglass frames, as shown in Fig 4. This system functions as both an electric seeing aid, as well as a wearable cameraphone, using a sleek and slender boom microphone as illustrated in Fig 5.

In view of such a concealment opportunity, I envisioned a new kind of EyeTap design in which the frames come right through the center of the visual field. With materials and assistance provided by Rapp optical, eyeglass frames were assembled using standard photo-chromatic prescription lenses drilled in two places on the left eye, and four places on the right eye, to accommodate a break in the eyeglass frame along the right eye (the right lens being held on with two miniature bolts on either side of the break). I then bonded fiber optic bundles concealed by the frames, to locate a camera and aremac in back of the head, for being concealed by the hair of the wearer.

This fully functional research prototype proves the viability of using eyeglass frames as a mediating element. The frames being slender enough (e.g. two millimeters wide) do not appreciably interfere with normal vision (especially when the apparatus is in operation) being close enough to the eye to be out of focus.

Figure 3: An early to mid 1980s eyeglass-based lifeglogging and computer mediated reality system, recently shown at the Smithsonian Institute, National Museum of American History, as well as at SIGGRAPH. This system was developed for the production of visual art in a computer mediated reality environment. The picture itself was captured using another similar system, and similar methodology.

This brings about a reversal of the roles of eyeglass frames and eyeglass lenses, in which the eyeglass lenses are a decorative design element, whereas the frames are what enables the seeing.

Although, in this design, one might think that the frames would block vision, the fact that they are so thin makes them quite tolerable. But even if the frames were wider, they can be made out of a see-through material, and of course they can be seen through by way of the illusory transparency afforded by the EyeTap principle. Therefore, there is definite merit in seeing the world trough eyeglass frames; the frames do not block vision because of illusory transparency, because they are computationally transparent as part of the compuational process of seeing.

Fig 6 illustrates the concept of illusory transparency. Subject matter blocked by the display device is still visible because the device exactly resynthesizes every corresponding ray of light that was absorbed in the analysis.

Additionally, beyond merely an illusory transparency, the computer screen is replaced by a computationally processed version of reality.

In eyeglasses, the long-term adaptation to seeing through a reality mediation device provides a unique opportunity to capture, process, store, and recall visual memories. Unlike

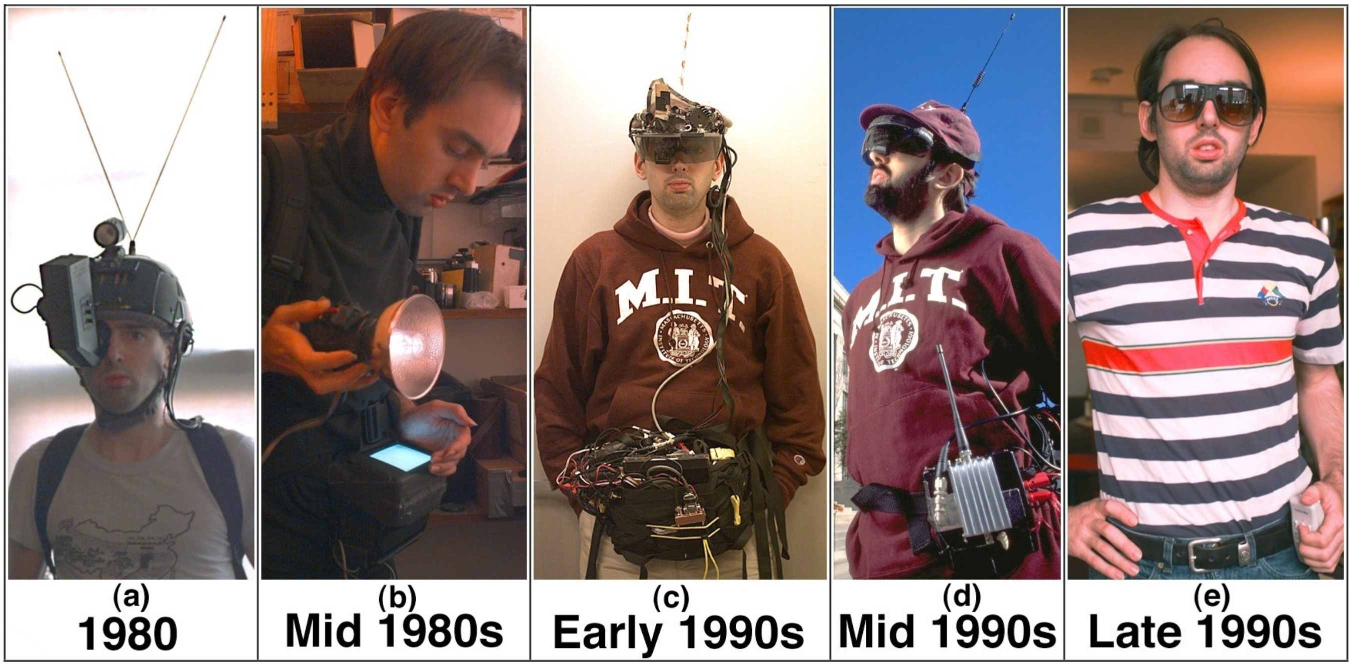

Figure 2: Evolution of lifeglogging (**life**long cybor**glogging**) for continuous archival and retrieval of personal experience.

Figure 5: Eyeglasses, having illusory transparent frames, form the basis of the wearable cameraphone, for the continuous archival and retrieval of personal video.

a mere wearable camera, the EyeTap, because it becomes a manner of seeing, captures exactly what the bearer does see. This results in a new kind of EyeTap cinematographic vision, which involves long-term adaptation to the new way of seeing.

A custom made injection moulded version of the EyeTap is shown in Fig 7. Such a design is suitable for mass production and commercialisation.

2.3 Design Principles: Basis of New EyeTap Designs

The design of the EyeTap has evolved considerably over the past 30 years. With components shrinking from day to day the EyeTap design has become more feasible for commercialisation and mass production. There are already many displays on the market but the unique configuration of an EyeTap with a camera and display simultaneously co-located effectively inside the eye, allows for more possibilities. It not only allows the merging of real world with computer mediated world but also opens new avenues for lifelong cyborglogging, and new modes of visual communication using computer mediated reality (Fig 8).

This configuration however has some critical points that allow the interface to be intuitive, non obtrusive and truly serves the possibility of having a computer as wearable as ones wrist watch and spectacles. In particular, the design is such that, after time, the user forgets he or she is wearing it.

While many displays exist on the market, none are suitable for long term usage such as lifelong capture. Therefore the EyeTap fills a necessary role in making lifelong continous capture possible. The collinearity is the most important factor that is necessary to achieve this lifetime synergy: **A very important factor is the EyeTap distance. The distance from the diverter to the eye should be exactly equal to that between the diverter and the camera. This calibration is vital to meet an EyeTap criterion in which the eye effectively becomes the camera.**

Social Aspects The early prototypes were quite obtrusive and often made people ill at ease, but more recently the apparatus has been gaining social acceptance. This can be attributed partly to miniaturization, which has allowed the construction of much smaller units, and partly to dramatic changes in people's attitudes toward personal electronics. With the advent of cellular phones, pagers, and so forth, such devices may even be considered fashionable.

When equipped with truly portable computing, including a wireless Internet connection and an input/output device like the modern style EyeTaps. the author found that people were not distracted by the device. In fact, they could not discern whether the wearer was looking at the screen or at the other party, because the two are aligned on exactly the same axis. This possibility only arises from the arrangement possible in an EyeTap of the camera, diverter and the

Figure 4: Eyeglasses having illusory transparent frames. When worn, the frames pass directly over the eyes, and the wearer sees "through" the frames by way of computer controlled laser light that resynthesizes rays of eyeward bound light. This fully functional but crude prototype leaves the workings visible, but in actual manufacture, the workings may be completely concealed within the frames.

aremac. The problem of focal length can generally be managed by setting it so that the display or aremac and anyone the wearer is talking with are in the same depth plane. This compared to all other systems that require complete attention into a display is far more intuitive in the user behavior that will develop. Thus the aremac, combined with camera, is preferable to a display.

3. APPLICATIONS, SECURITY AND PRIVACY

Just as computers have come to serve as organizational and personal information repositories, "smart clothing", when worn regularly, also becomes a multi purpose device.

3.1 Cyborglogging ('glogging)

After wearing the EyeTap for a number of years, one adapts to seeing the world that way, and this provides a new kind of personal experience capture, in day-to-day life. The hands-free nature of the EyeTap therefore helps document everyday personal interaction.

While a camera may not be used as a record of truth, it can be used as a visual memory prosthetic. Our own family photographs bring back childhood memories for each of us, but mean little or nothing when shown to someone else. Similarly, We can relive our Christmas vacation by scrolling through an EyeTap image sequence, even if it is severely downsampled (e.g. throwing away every 100 frames or so). Even a low resolution movie brings the memory back to us clearly in our own mind, even when the images are barely discernable to others who were not present at the event.

Cyborg**glogging** ('glogging) is the (usually continous) recording of an activity by a participant in the activity, which often results in the serendipitous capture of precious moments, such as the birth of a newborn baby, as shown, for example, in Fig9.

Such a living and permanently installed/instilled photographic perspective allows the bearer to capture the birth of a newborn, or to capture baby's first steps.

3.2 Vicarious Soliloquy

I am a document camera.

The EyeTap also opens a new possibility: seeing another person's point of view from within, i.e. not just a point of view camera, but to actually see exactly what the other person sees. The use of this technique has been made on numerous occasions, including remotely delivered Opening Keynotes at conference and symposia, that some have said to be more compelling than the actual physical presence of keynote speakers in previous years, and thus the remote intervention is made interesting in and of itself. An example of a Visual Vicarious Soliloquy, was the author's keynote address at DEFCON 7, as illustrated in Fig 11.

This excerpt from the realtime 'glog depicts how the audience was, in effect, able to "be me", rather than "see me", in the sense that a first-person perspective was offered by

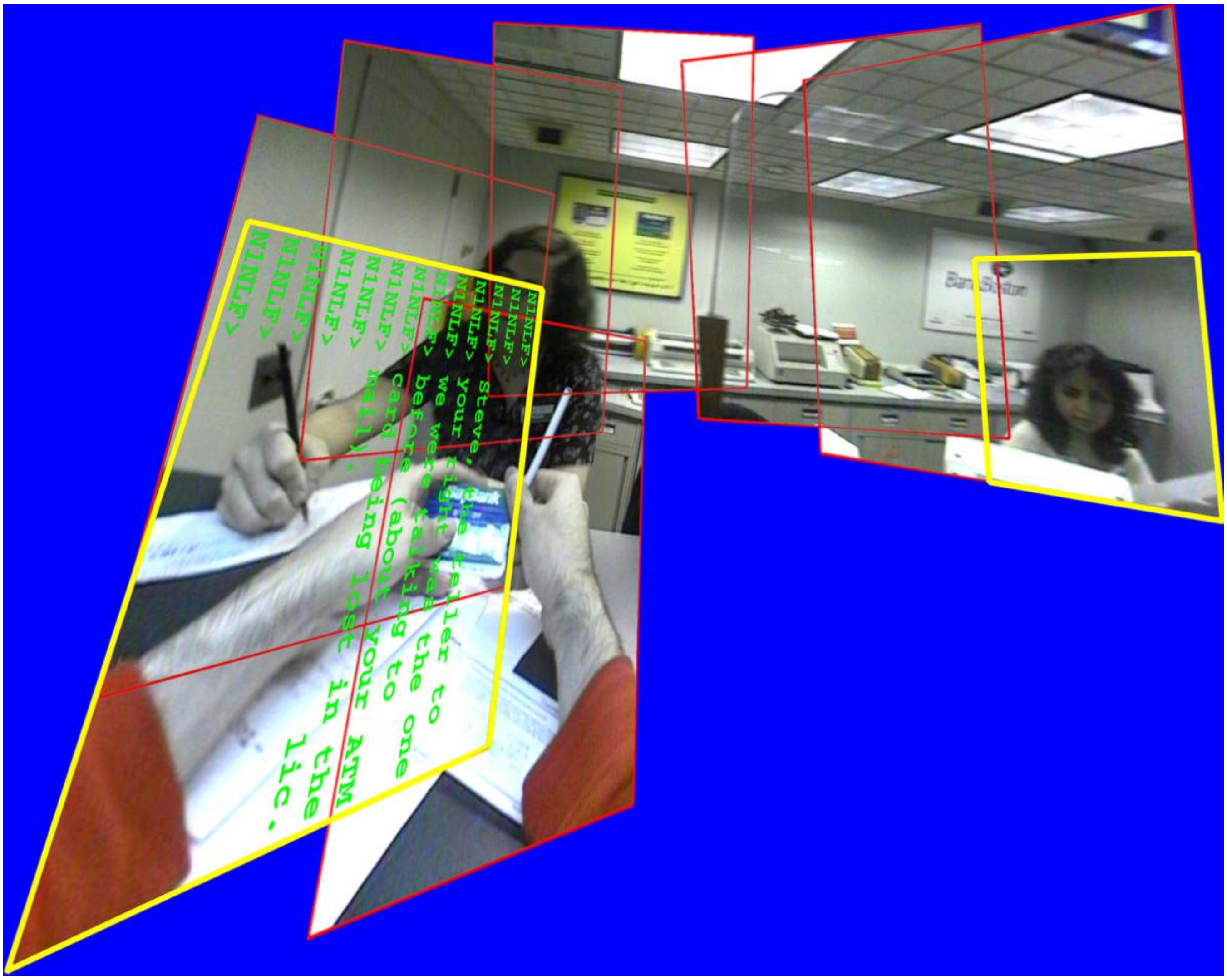

Figure 8: Computer mediated reality as a new form of interactive communication. A remote spouse may, for example, provide interactive annotation of visual space, to assist in memory, communication, etc..

the apparatus of the invention serving as an existential document camera.

The existential aspect of the apparatus of the invention puts the audience, in effect, inside the wearer's head to share a truly first person perspective. I presented the Keynote Address of DEFCON 7 as a lecture to myself, which I gave while walking around, while writing on a notepad. The EyeTap causes the eye itself to, in effect, function much like a document camera, but also, through long term adaptation, brings the wearer at one with his or her surroundings.

This capability adds a new dimension to videoconferencing.

4. PERSONALIZED EXPERIENCE CAPTURE: THE EVERYDAY WORLD IN A DIFFERENT LIGHT

One of the original goals behind the invention of the wearable computer and EyeTap devices was not just for **personal experience capture** but also for **personalized experience capture**.

Thus the EyeTap device that allows the wearer to capture everyday life without conscious thought or effort, also allows the wearer to interpret the world in an expressive and artistic way.

4.1 Computer Mediated Reality as a form of visual art

Computer Mediated Reality is created when the user's visual perception of their environment is augmented, diminished or otherwise altered in some way. I achieve this by the use of an EyeTap device to change the wearer's perception, possibly by altering the visual appearance of the scene, or adding/removing/modifying visual content. In this way, computer–generated information may be added into the scene, or may replace/modify content in the scene. Similarly, real–world objects may be removed from the scene. Finally, computer generated information may be blended into the real–world scene. In particular, this computer generated information may consist of a view of the world, seen in a different light.

Mediated Reality differs from virtual reality (or augmented reality) in the sense that it allows users to filter out things they do not wish to have thrust upon them against their

Figure 6: Illusory transparency. Here a computer display screen (VGA television) is mounted on an easel to display the subject matter directly behind it. While at first one might be inclined to think that such a device, whether in eyeglasses or life size, is totally useless. However, it is the ability to insert computation in the reality stream that makes it useful.

will. Just as a Sony Walkman allows us to drown out ambient noise or undesirable music with our own choice of music, Mediated Reality allows us to implement a "visual filter".

4.2 Photographic Origins of Computer Mediated Reality in the 1970s and 1980s

The original Personal Imaging application[8] was an attempt to define a new genre of imaging and create a tool that could allow reality to be experienced with greater intensity and enjoyment than might otherwise be the case.

This effort also facilitated a new form of visual art called Lightspace Imaging (or Lightspace Rendering) in which a fixed point of view for a base station camera was selected, and then, once the camera was secured on a tripod, the artist walked around and used various sources of illumination to sequentially build up an image layer-upon-layer in a manner analogous to paint brushes upon canvas, and the cumulative effect embodied therein. An early 1980s attempt, at creating expressive images using the personal imaging system developed in the 1970s and early 1980s is depicted in Fig 12. Throughout the 1980s, a small number of other artists also used the apparatus to create various lightvector paintings. However, due to the cumbersome nature of the

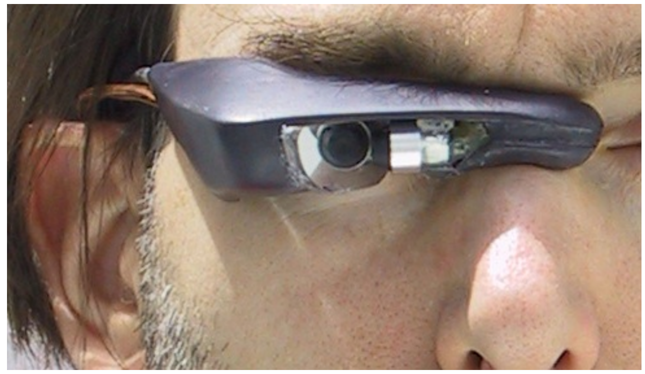

Figure 7: Injection molded EyeTap suitable for mass production. Note the illusory "eye is camera" (camera in the eye) appearance owing to the fact that rays of eyeward bound light are diverted into a camera (actually mounted in the nosebridge and pointing toward the wearer's right eye).

early WearComp hardware, etc., and the fact that much of the apparatus was custom made by, and to fit the author, it was not widely used over an extended period of time by others. However, the personal imaging system proved to be a new and useful invention for a variety of photographic imaging tasks.

My particular approach to creating this poetic narrative on reality, was to combine multiple exposures of the same subject matter.

Typically exposures are maintained as separate image files overlaid on the artist's screen (EyeTap) together with the current view through the camera. The exposures being in separate image files allows the artist to selectively delete the most recent exposure, or any of the other exposures previously combined into a running "sum" on the EyeTap. ("Sum" is used in quotes here because the actual entity, a summation in an antihomomorhic vectorspace [10].) Additional graphic information is also overlaid to assist the artist in choice of weighting for manipulation of this "sum". This capability is quite useful, compared to the process of painting on canvas, where one must paint over mistakes rather than simply being able to turn off the brushstrokes that were mistakes. Moreover, the ability to adjust the intensity of brushstrokes after they are made, is also useful, beyond merely turning them on and off. Furthermore, exposures to light can be adjusted either during the shooting or afterwards, and then re-combined. The capability of doing this during the shooting is an important aspect of the personal imaging invention, because it allows the artist to capture additional exposures if necessary, and thus to remain at the site until a final picture is produced. The final picture as well as the underlying dataset of separately adjustable exposures is typically sent wirelessly to other sites so that others can manipulate the various exposures and combine them in different ways, and wirelessly send comments back to the artist (e.g. by email), as well as by overlaying graphics onto the artist's head mounted display which then becomes a collaborative space.

Figure 9: Video still from 'glog showing Christina Mann, immediately after birth.

4.3 Lightstrokes and Lightvectors

Each of a collection of differently illuminated exposures of the same scene or object is called a lightstroke. In the context of Personal Imaging, a lightstroke is analogous to an artist's brushstroke, and it is the plurality of lightstrokes that are combined together that give the invention described here it's unique ability to capture the way that a scene or object responds to various forms of light. From each exposure, an estimate can be made of the actual quantity of light falling on the image sensor, by applying the inverse transfer function of the camera. Such an estimate is called a lightvector [5].

Furthermore, a particular lightstroke may be repeated (e.g. the same exposure may be repeated in almost exactly the same way, holding the light in the same position, each time a new lightstroke is acquired). These seemingly identical lightstrokes may collectively be used to obtain a better estimate of a lightvector, by averaging each of the lightvectors together to obtain a single lightvector of improved signal to noise ratio. This signal averaging technique may also be generalized to the extent that the lamp may be activated at various strengths, but otherwise held in the same position and pointed in the same direction at the scene. The result is to produce a lightvector that captures a broad dynamic range by using separate images that differ only in exposure level[12][6][7].

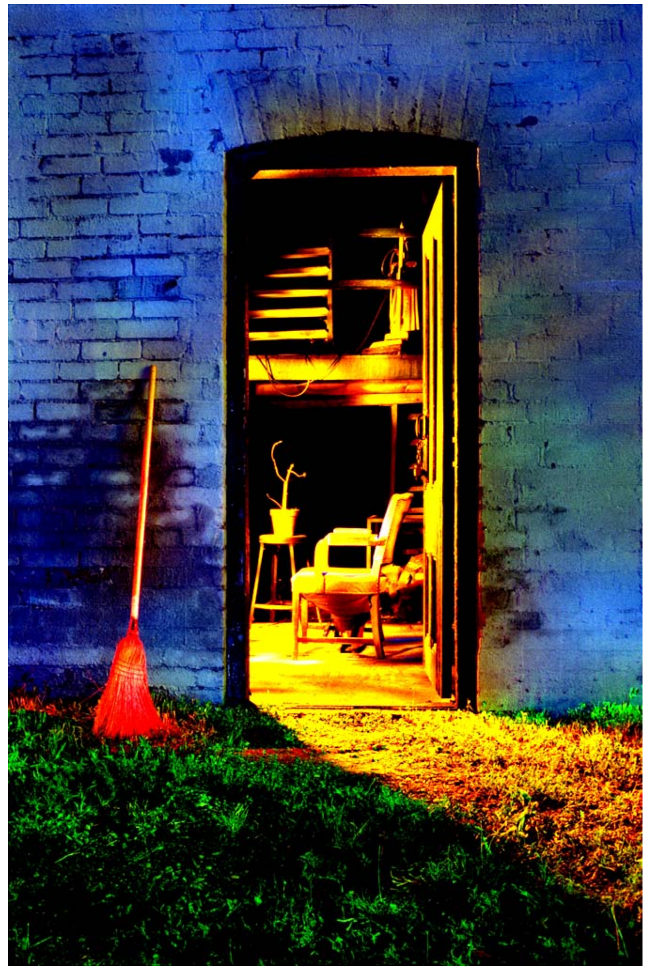

Figure 12: The goal of Personal Imaging [8], is to create something much more like the sketch pad or artist's canvas than like the camera in its usual context. This kind of Personal Experience Capture produced as artifacts, results that are somewhere at the intersection of painting, computer graphics, and photography, with an emphasis on the personal interpretation of reality. Notice how the broom appears to be its own light source (e.g. self-illuminated), while the open doorway appears to contain a light source emanating from within. The rich tonal range and details of the door itself, although only visible at a grazing viewing angle, are indicative of the affordances of the Lightspace Rendering [5][13] method.

4.4 Computer Mediated Reality as a tool for transforming everyday life into visual art

Stepping beyond the obvious practical uses of Computer Mediated Reality, there is a more existential motivation regarding how we, as humans, are able to choose the manner in which we define ourselves [14]. The lifelong cyborglog recorder is more than just a visual memory prosthetic. It is also a new tool for the visual arts.

One of the original goals of Computer Mediated Reality was to create a body–borne wireless sensory environment which, although technically sophisticated, would function more in the spirit of an artist's personal notes or a painter's canvas.

An example of the combination of different exposures of the same subject matter, to generate lightvector paintings, is shown in Figure 3 of an accompanying publication in ACM Multimedia 2004, entitled "Sousveillance".

This process of "painting with lightvectors" was also possible with a group of people wearing computerized seeing aids that were tuned to the same virtual channel, so that there was a shared computer-mediated visual reality. In this way, the team experienced a collectively modified view of the world, in the production of visual art. See Fig 13.

4.5 Lightvectoring: Blending in lightspace

It has been shown that cameras respond non–linearly to light [10]. Because Mediated Reality typically uses a camera to view the user's environment it may be considered a photographic process. Thus, the non–linear response of the camera can be taken into account. This allows the use of techniques dealing with camera response functions to be applied towards creating a convincing and visually accurate Mediated Reality. In successful implementations of mediated reality, computer–generated information is blended into a real–world scene while taking into account the inherent non–linearity of the camera response.

Just as computer–generated information can be spatially registered with the real–world, so to must it be tonally aligned. Because the user's environment is percieved through video taken from a camera, the photographic qualities of the camera become important considerations when analyzing images, and synthesizing output images.

This is important for Mediated Reality for a variety of reasons. Firstly, from a user–interface standpoint, some form of blending is required. If computer–generated information is simply overlayed on the input image, then objects appear or disapper behind the computer–generate information. While this creates a more convincing appearance of the computer–generated information as being attached to the real–world, the user can no longer see the occluded portions, which, in some applications, may be undesirable. Thus, displaying a blended version of the real–world and the computer–generated information is often desirable.

For tonally accurate Mediated Reality, the computer generated information should appear to be appropriatedly lightened or darkened to match the real–world image in some sense. For instance, in a dark environment, it may be desirable to dim the computer generated information to naturally match the video image, and vice versa in a well-lit environment. Additionally, this also presents the eye with a more uniform amount of light, allowing it to better adjust. The computer generated information will be appropriately bright in a bright image, and not be lost or difficult to see. Likewise, computer–generated information will be appropriately dim in a dim image, and not overpower the image.

Finally, typical blending techniques may not produce desirable results when applied to photographic material. The blending applied for computer graphics applications does not take into account the non–linear nature of the images. Photoquantigraphic technique (lightspace rendering) produce superior blending.

Fig 14 shows a lightvector painting for being exhibited to replace billboards and advertising. By replacing offensive billboards, advertising, and other visual detritus with fine art, we can reduce the distracting and unpleasant visual clutter, to allow the wearer of the electric eyeglasses to focus more on what is important.

I found that when billboards and other ads are filtered out, there is an abrupt and startling transition that is distracting — sometimes even more distracting than the ad itself, and that this effect could be mitigated by making the ad dissolve, much like the dissolve units used to crossfade between two 35mm slide projectors.

Fig 15 shows the results of an advertisement dissolve, as part of ongoing collaborative work with my PhD student James Fung.

The need to diminish existing architecture from a user's view is discussed in [4]. The removal of billboards via a projective plane tracking algorithm is discussed in [11]. Our approach differs in the use of a photographic response based blending techniques (lightspace rendering), rather than complete removal or replacement of scene content.

Our approach is based upon comparametric methods of determining the response function of the cameras [10, 2, 3, 1]. We assume that the camera's response function is known, and use this information in the blending process. Candocia [2, 3, 1] discusses use of the response function for blending images of the same scene to create image panoramas and composites from differently exposed images registered both spatially and tonally. Our blending differs in that it uses computer–generated content from another scene to be blended into the real–world scene.

Computer vision and image processing is achieved on the graphics hardware by treating the input image as a texture. This image texture is mapped onto a quadtrilateral, and this quadrilateral is displayed with the desired geometry. The display geometry allows for scaled images to be processed (i.e. by displaying the image larger or smaller than its original size), or a projection of the image to be displayed. The latter is useful when combined with a projective plane tracker since images are naturally defined in a projective space in both the tracking solution and the graphics hardware. This also moves the task of image element interpolation onto the graphics hardware which uses interpolation techniques on the texture.

4.6 Projective Image Registration

The most computationally intensive part of our method for Computer Mediated Reality is the calculation of projective image registration parameters P which spatially align the coordinate systems $[x,y]^T$ and $[x',y']^T$ of two input images [9]. Specifilly, the algorithm calculates:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}}{\begin{bmatrix} c_1 & c_2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} + 1} \tag{1}$$

which solves for the 3 parameters (8 scalar parameters)

$$P = [a_{11}, a_{12}, b_1, a_{21}, a_{22}, b_2, c_1, c_2] \tag{2}$$

to spatially register two images.

It has been shown that this algorithm is capable of tracking a user's head motion, as well as tracking planar objects in the scene [11], using a wearable computer system and EyeTap eyeglasses. In order to create a convincing computer mediated reality for the user, we apply this algorithm to register frames of video input from the EyeTap so the sequential images can be registered, and computer generated information placed into each frame so that it appears affixed to the real–world.

We have developed our system such that new image processing and computer vision algorithms can be quickly implemented, and then run as independent processes on multiple parallel graphics devices, so that the system runs in real time.

5. LIGHTVECTORING MADE EASY: "DUSTING" FOR MAINSTREAM SOCIETY

Having worked on mediated reality and lightvectoring for some 30 years, it is my desire to bring this technology to the masses, and make it readily available for anyone to use. Thus, in part, this paper expands on aspects of the companion paper I have published in ACM Multimedia 2004, in particular, to make the art of lightvectoring easy, we designed a simple hand grip for a light source and a simple display (Fig 16).

Such a device is also easy for children to learn how to use it (Fig 17). Once familiar with the operation, the device is easy to actuate as a lightvectoring medium (Fig 18).

An example of an expressive/creative lightvector painting by Christina appears in Fig 19.

Moreover, an archival of personal experience may help a child remember a past that would be otherwise forgotten, at an age when memory is not so well formed.

6. "SOUSVEILLANCE": PERSONAL EXPERIENCE CAPTURE IS INVERSE SURVEILLANCE

Early on, in my pursuit of the invention, design, building, and using electric seeing aids, and tools for archival and recall of personal experiences, visual art, etc., I observed a strong opposition from certain organizations.

Over the past 30 years I have noticed an increased peer acceptance, over time. Over the same time period, organizational acceptance has decreased. By organizational acceptance, I refer to the acceptance by persons acting in a role, or official capacity.

Most notably, I noticed that the places where I suffered the strongest form of discrimination where places that had the greatest degree of surveillance cameras. It seemed as though the more surveillance cameras an organization had within it, the more opposition I would encounter. Thus I came to think of **personal imaging** (i.e. the archival of personal experience) as an inverse to surveillance, for which I coined the term "sousveillance". Surveillance is French for "veiller" ("to watch") and "sur" ("above"). Thus the word "surveillance" means "to watch from above". This notion of a God's eye view, watching down on us from on-high (Fig 20), has been pervasive, implicit, but unspoken, and unquestioned. In fact a recent series of posters put up by the London Underground incorporated the imagery of eyes in the sky, designed around the London Underground logo (Fig 21).

6.1 The Sousveillance Industry

There is now a growing sousveillance industry. For example, the Hitachi Design Center in Milano recently sponsored an event entitled "*Applied Dreams Workshop 3: 'Surveillance and Sousveillance'*". (See excerpt in Fig 22.) Other companies such as Microsoft (this workshop, i.e. CARPE, SenseCam, etc.), Nokia ("lifeblog" which is quite similar to the author's lifeglog project), and Hewlett Packard (the author worked at HP Labs on many related ideas in the early 1990s) with their "casual capture" project, are doing similar work.

Sousveillance differs from counterveillance ("counter surveillance is the practice of avoiding surveillance or making it difficult" http://en.wikipedia.org/wiki/Surveillance) in that it's not necessarily aimed at avoiding or eliminating surveillance, but, rather, at creating a separate view in the other direction.

7. EQUIVEILLANCE: THE EQUILIBRIUM BETWEEN SURVEILLANCE AND SOUSVEILLANCE

Proponents of ubiquitous surveillance (or "perveillance" – pervasive surveillance, ubiqcomp, pervcomp, etc.) might be inclined to propose an increase in transparency as a good thing, or if not a good thing, as something inevitible. However, what people seem to not notice, is the one-sided nature of surveillance.

I noticed, for example, that taxicab drivers began to become upset when photographed, around the same time that surveillance cameras began to appear in taxicabs. Shopkeepers became upset at being photographed around the same time that surveillance cameras began to appear in shops, and so on. More recently, around the same time that new surveillance cameras were installed in the parks in Toronto, signs were also put up warning people not to take pictures of their own (Fig 23). One wonders if the surveillance cameras were put up to enforce the no photography policy.

By simply trying to live my own life, without bothering anyone, I found myself discriminated against, and I found that this discrimination was correlated to the amount of surveillance in an immediate environment. I do not try to make the claim that the discrimination was caused by the surveillance, but the correlation certainly was very evident, and consistent over a 20 year period in many different countries around the world. It seemed to depend very little on culture, i.e. it was consistent across many different cultural norms.

Ironically, the prohibition on photography, and the like, was allegedly for privacy reasons, e.g. the officials would typically say that they were trying to protect the privacy of other patrons, citizens, or the like.

Protect privacy by installing surveillance cameras?

Well, even if the surveillance cameras were, in fact, installed to help enforce rules that prevent people from photographing each other, one must ask: What definition of privacy is being protected?

The answer, is of course, the very same kind of privacy that prisoners of Bentham's Panopticon enjoy: they have absolute privacy from those in other cells, and zero privacy from the guards. Thus if privacy is the condition of being photographed by a central authority, who protects us from photographing each other, then we must be living on a prison planet — a Panoptic prison to be exact.

This led me to the notion of inequiveillance (equiveillance being the equilibrium between surveillance and sousveillance), and in particular, to the following equiveillance doctrine:

> **Equiveillance 1:** Although there may well be situations where sousveillance might be inappropriate, sousveillance must never be prohibited in situations where surveillance exists.

The existence of surveillance takes away a reasonable expectation of privacy, and therefore creates, of a space, a free-fire zone. But more importantly, a subject of surveillance has a right (and perhaps also a responsibility) to contextual integrity of the surveillance data. Therefore, to prevent the said surveillance data from being taken out of context, the subject may wish to capture his or her own personal record of his or her own life, to provide proper context for what the surveillance data might show.

This leads me to a second, although weaker requirement:

> **Equiveillance 2:** One who has placed a person, or persons under surveillance invalidates the surveillance data by prohibiting persons under surveillance from also keeping their own record (sousveillance) of their actions.

Thus, for example, it is my opinion that High Park should not be able to use their surveillance data in a court of law, because they have a policy and practice of prohibiting patrons from constructing their own case. By prohibiting subjects from collecting data that could be used in their own defense, the surveillers should be seen as tampering with evidence. Not that they have actually edited the recorded data, but the manner in which the data is recorded must be seen as suspect at best, when denying another party from a reasonable opportunity to collect data in his or her own exoneration. Moreover, for a person who has a memory impairment, one who prevents that person from reasonably remembering what happened, should not be taken as credible. Indeed, when, for example, a contract is signed, between two parties, each party has both a right and a responsibilty to keep a copy of the contract. Were one party would prevent another from keeping a copy of the contract, the preventing party should lose the force of the contract.

Thus in much the same way that a radio station keeps a logfile of what they transmit, perhaps a society under surveillance should keep a logfile of what it says. Accordingly, the art of sousveillance, i.e. the recording an activity by a participant in the activity, may very well be what is needed to tame the monster of surveillance with a piece of itself.

8. CHALLENGING INEQUIVEILLANCE

I have come to believe that secrecy, rather than privacy (true privacy, not panoptic privacy), is much to blame for crime, terrorism, and the like. In particular, it could be said that crime is pervasive, at all levels, whereas surveillance provides only a one-sided top-down "oversight" without a corresponding "undersight".

But we are at a pivotal era in which it is possible to turning the tables toward correcting the imbalance that was recently introduced with the proliferation of a surveillance-only paradigm.

Information is power, seeing is believing, and organizations believe in power - power over individuals. But the very miniaturization (Fig 24) that has made it possible for police to hide cameras in shopping-mall washrooms has also made camcorders small and light enough for average citizens to carry around and capture events like the Rodney King beating, and similar human rights abuses in other countries as well. As with many problems, the problem of surveillance contains the seeds of its own solution, namely sousveillance.

The camcorder represents a highly portable system, that exists at the end of a 3 step hierarchy of portability:

- Fixed: (also known as "base stations): fixed devices on or in buildings, homes, offices, or other "fixtures such as a ham shack, post, outpost, or attached to a tree, or other fixture;
- Mobile: Vehicular or ship-based systems, wireless systems in trucks, vans, cars, boats, or motorcycles. A wireless station on a bicycle such as N4RVE's behemoth would also be categorized as Mobile. People who use mobile communications devices are often called mobileers (http://en.wikipedia.org/wiki/Mobileers);
- Portable: Handheld or wearable systems. Systems borne by (e.g. worn or carried upon) the human body. An implantable system such as a wireless communicator injected beneath the skin would also fall under the Portable category.

There are three weaknesses in the camcorder technology, however:

1. inconvenience and obtruseveness;
2. destructibility of the evidence; and
3. insufficient protection from forced disclosure

(the latter two pertaining to siezure by authorities). Firstly a camcorder's use requires an active role. Despite names like "handycam", it does require thought and effort to pull it out and begin recording with it, whereas its big brother (the surveillance camera upon the lamp post or ceiling) requires zero effort to engage - it is always on. If the officers had seen the witness pulling out a camcorder (pulling it out does attract considerable attention) they would have probably confiscated the recording. In fact there have been numerous situations in which persons recording police misconduct have themselves become targets of misconduct. This brings us to the second weakness of the camcorder: local storage.

The EyeTap combined with the WearCam allows one to put images and video onto his World Wide Web home page with near zero delay. The EyeTap points ahead, matching the view of the wearer, and it sends images over the internet, so that they can be backed up in one or more remote locations, perhaps in different countries around the world. Basically an EyeTap has the capability of producing an indestructible visual record.

Although other forms of sousveillance, such as neckworn cameras, and the like, are possible (e.g. Fig 25), the EyeTap has been found to provide the best and most useful data, owing to the notion of an "eye-centered design".

8.1 The end of video Surveillance

The distributed nature of the EyeTap mediated memory data would make it less subject to a totalitarian control than video surveillance. Video surveillance will always be upon us. Quite likely, the establishment, with its use of video surveillance, will have the upper hand, for they have the advantage of fixed camera geometry calibrated within the environment, the ability to do motion detection (e.g. when nobody is present, all pixels remain the same), and better communications (hard-wired closed-circuit). However, the ubiquitous use of wearable EyeTap will tip the balance a little toward the center, i.e. towards a little bit of fairness on the surveillance superhighway. While the taxi drivers, law enforcement officers, shopkeepers, and government will continue to have surveillance, now the passengers, suspects, shoppers, and citizens will be able to look back at the former on a more fair and equal footing.

Privacy advocates are often either ignored, or focused on the wrong issues (e.g. worrying about ways to reduce junk mail). Another approach that might be worth considering is shooting back.

But sousveillance is not merely a 20th century "us versus them" concept. For example, a cab driver on day may be a passenger in someone else's cab the next. A shopkeeper may at times be a shopper in someone else's store. Thus sousveillance, when it becomes a form of inverse surveillance, is not about shooting clerks. It's about creating balance. Moreover, a more general notion of sousveillance is the recording of an activity by a participant in the activity, which need not necessarily have anything to do with political hierarchy. (See also côtéveillance, also known as coveillance, meaning "peer to peer" watching to the side, http://www.surveillance-and-society.org/articles1(3)/sousveillance.pdf in Surveillance and Society, 1(3), pp 332-355.)

8.2 Citizens as mainstream culture, and guards as activits for counter culture

The sousveillance industry has created and will continue to create a mainstream cultural force that is unstoppable. In this sense, it is not the users of sousveillance products who are activits in a counter culture. These users are merely becoming part of the mainstream culture. Instead, it is the security guards who use, or threaten to use violence to stop sousveillance who are the activists fighting for their political cause of surveillism and surveillist monopoly.

A guard or garrison that threatens to use violence in order to achieve its political goal (such as a protest against lifeglogs) is known as a "guerrorist" (à la guerre). In this sense a security guard becomes an activist, protesting against inevitible technologies for the Continuous Archival and Recall of Personal Experience that would record the guard's (inverse) civil disobedience. It is as if those of us wearing the cameras are inside the walls of the city, having a summit to discuss the future of C.A.R.P.E. while the guards have gathered outside to throw stones at us in protest. In this sense, there has been a role reversal, in which sousveillance has become or will soon become the cultural "Strong Force" and Panopticism has become or will soon become (relatively speaking) the cultural "Weak Force".

9. CONCLUSIONS

EyeTap devices facilitate the continuous archival and retrieval of personal experiences, by way of lifelong video capture. As a form of electric seeing aid, and wearable cameraphone, such devices function as natural extensions of the mind and body. More generally, a visually mediated reality facilitates new forms of visual art.

Unfortunately, the greatest difficulties to be overcome are not technical ones – these have largely been solved over the past 20 or 30 years. What remains to be solved is the problem of inequiveillance, i.e. the imbalance between surveillance and sousveillance.

10. ACKNOWLEDGEMENTS

I greatfully acknowledge the support of Nikon, and the assistance of Daymen Photo (Metz). I'd also like to thank Chris Aimone (Camere distortion parameters), Anurag Sehgal who worked with me on the original draft part of a similar article on EyeTap designs, and James Fung who worked with me on a similar paper on realtime wearable hardware implementation of advertisement filters with lightvectoring.

11. REFERENCES

[1] A. Barros and F. M. Candocia. Image registration in range using a constrained piecewise linear model. *IEEE ICASSP*, IV:3345–3348, May 13-17 2002. avail. at http://iul.eng.fiu.edu/candocia/Publications/Publications.htm.

[2] F. M. Candocia. A least squares approach for the joint domain and range registration of images. *IEEE ICASSP*, IV:3237–3240, May 13-17 2002. avail. at http://iul.eng.fiu.edu/candocia/Publications/Publications.htm.

[3] F. M. Candocia. Synthesizing a panoramic scene with a common exposure via the simultaneous registration of images. *FCRAR*, May 23-24 2002. avail. at http://iul.eng.fiu.edu/candocia/Publications/Publications.htm.

[4] D. R. Gudrun Klinker, Didier Stricker. Augmented reality for exterior construction applications. In W. Barfield and T. Caudell, editors, *Fundamentals of wearable computers and augmented reality*, chapter 12, pages 397–427. Lawrence Erlbaum Press, New Jersey, 2001.

[5] S. Mann. Lightspace. Unpublished report (Paper available from author). Submitted to SIGGRAPH 92. Also see example images in http://wearcam.org/lightspace, July 1992.

[6] S. Mann. Compositing multiple pictures of the same scene. In *Proceedings of the 46th Annual IS&T Conference*, pages 50–52, Cambridge, Massachusetts, May 9-14 1993. The Society of Imaging Science and Technology. ISBN: 0-89208-171-6.

[7] S. Mann. Joint parameter estimation in both domain and range of functions in same orbit of the projective-Wyckoff group. pages III–193–196, Lausanne, Switzerland, December 1996. Also appears in: M.I.T. M.L. T.R. 384, 1994.

[8] S. Mann. Wearable computing: A first step toward personal imaging. *IEEE Computer; http://wearcam.org/ieeecomputer.htm*, 30(2):25–32, Feb 1997.

[9] S. Mann. Humanistic intelligence/humanistic computing: 'wearcomp' as a new framework for intelligent signal processing. *Proceedings of the IEEE*, 86(11):2123–2151+cover, Nov 1998. http://wearcam.org/procieee.htm.

[10] S. Mann. Comparametric equations with practical applications in quantigraphic image processing. *IEEE Trans. Image Proc.*, 9(8):1389–1406, August 2000. ISSN 1057-7149.

[11] S. Mann and J. Fung. EyeTap devices for augmented, deliberately diminished, or otherwise altered visual perception of rigid planar patches of real world scenes. *PRESENCE*, 11(2):158–175, 2002. MIT Press.

[12] S. Mann and R. Picard. Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures. In *Proc. IS&T's 48th annual conference*, pages 422–428, Washington, D.C., May 7–11 1995. Also appears, M.I.T. M.L. T.R. 323, 1994, http://wearcam.org/ist95.htm.

[13] C. Ryals. Lightspace: A new language of imaging. *PHOTO Electronic Imaging*, 38(2):14–16, 1995. http://www.peimag.com/ltspace.htm.

[14] S. M. (with Hal Niedzviecki). *Cyborg: Digital Destiny and Human Possibility in the Age of the Wearable Computer*. Randomhouse (Doubleday), November 6 2001. ISBN: 0-385-65825-7.

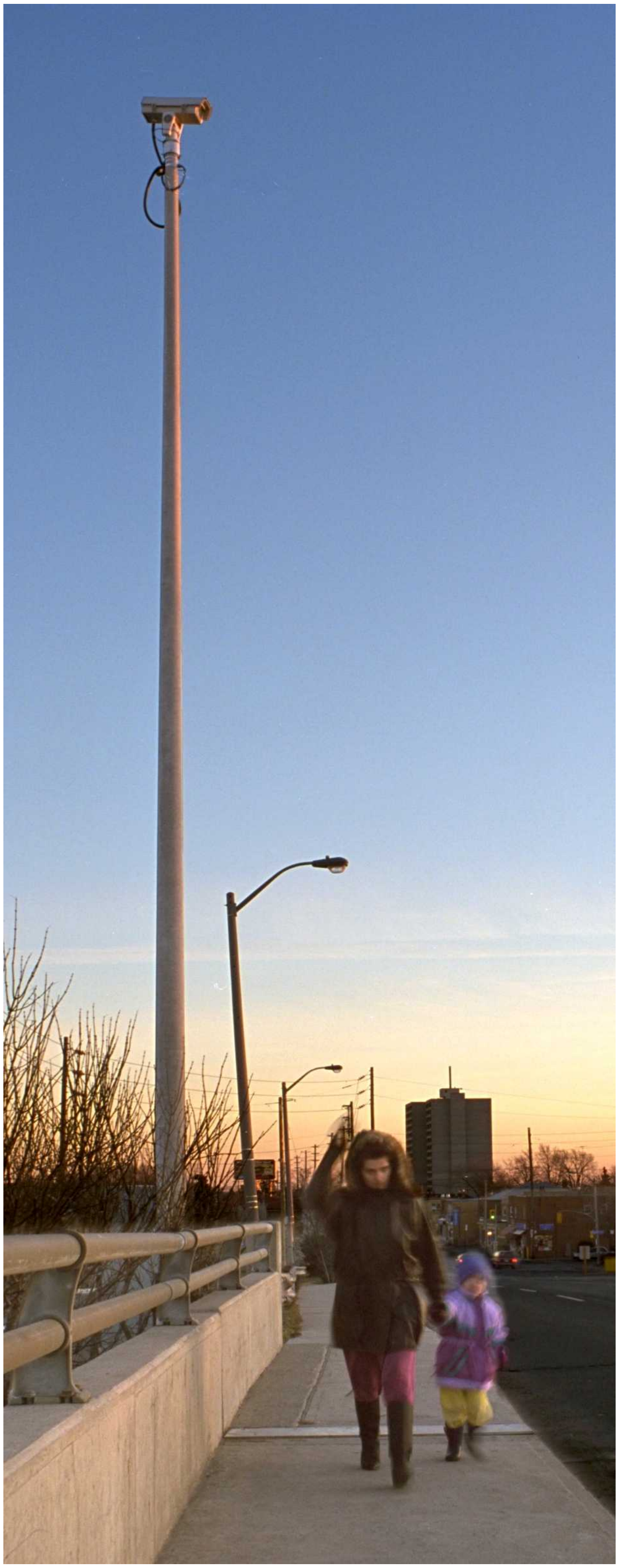

Figure 20: The word "**sur**veillance" is French for "to watch from above". This God's eye view is typical of the notion of watching from on-high.

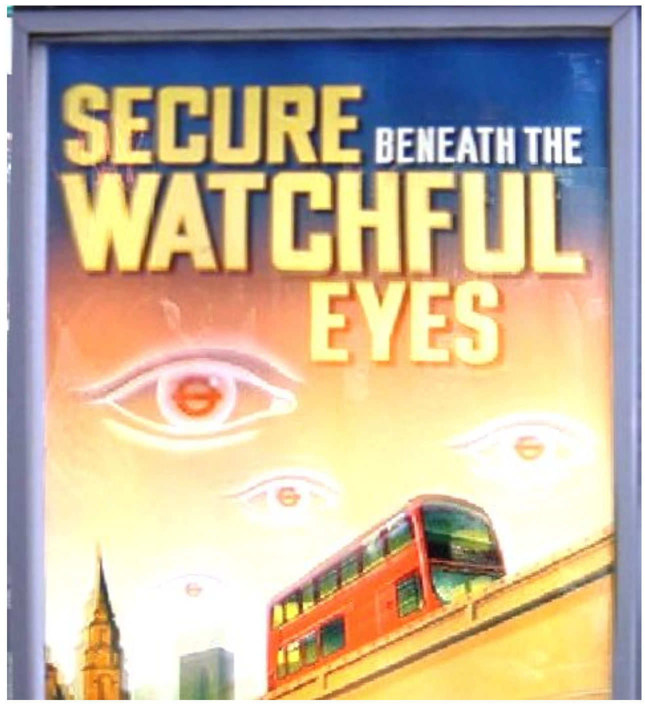

Figure 21: Eye-in-the-sky: London Underground posters to let people know they are being watched from above.

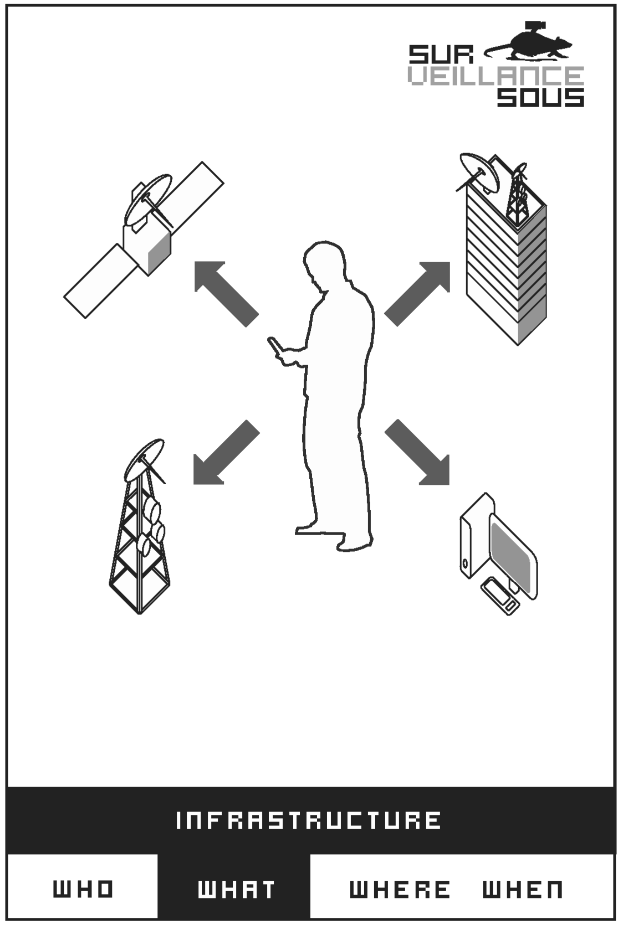

Figure 22: Excerpt from Hitachi Design Center's recent event (Milano), entitled: "Applied Dreams Workshop 3: 'Surveillance and Sousveillance"'. Many other companies are also looking at sousveillance in various ways.

Figure 24: Mobile surveillance: Miniaturization of surveillance cameras has made it possible to put passengers in many taxicabs under surveillance.

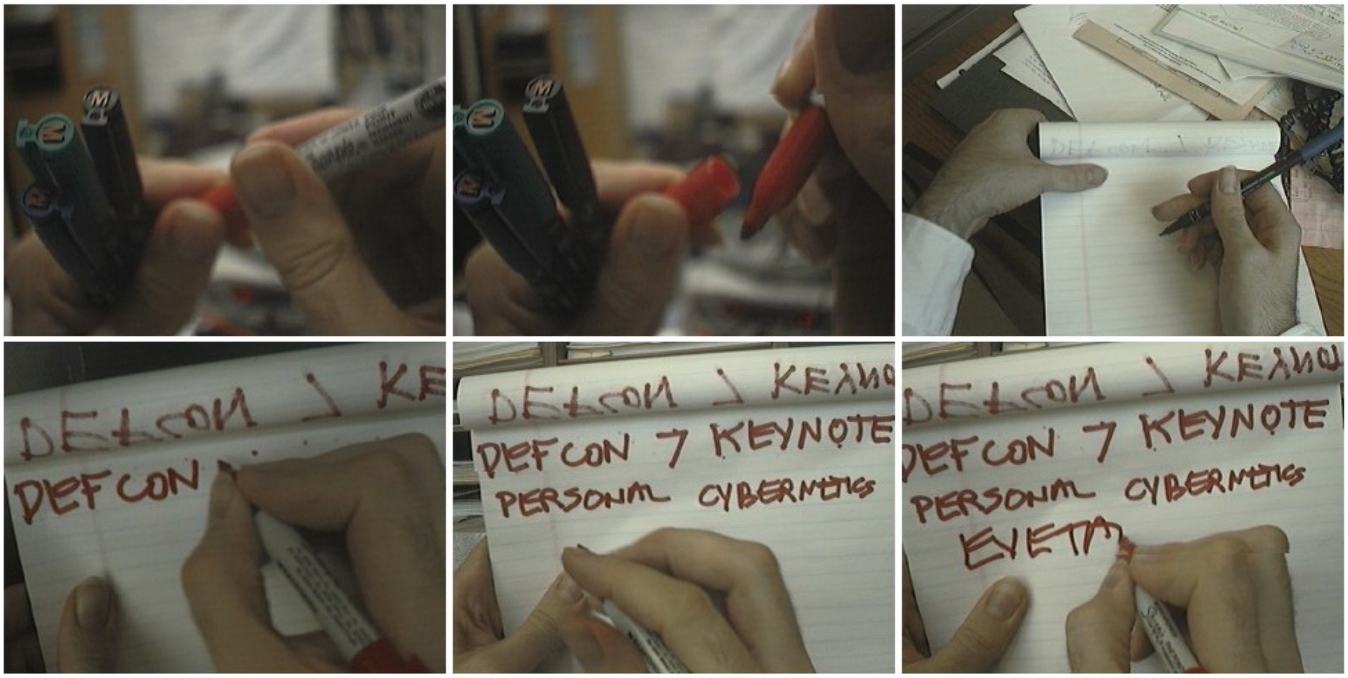

Figure 10: The Vicarious Soliloquy genre as Keynote Address at DEFCON 7: Six still frames from a realtime 'glog used as a new communications medium.

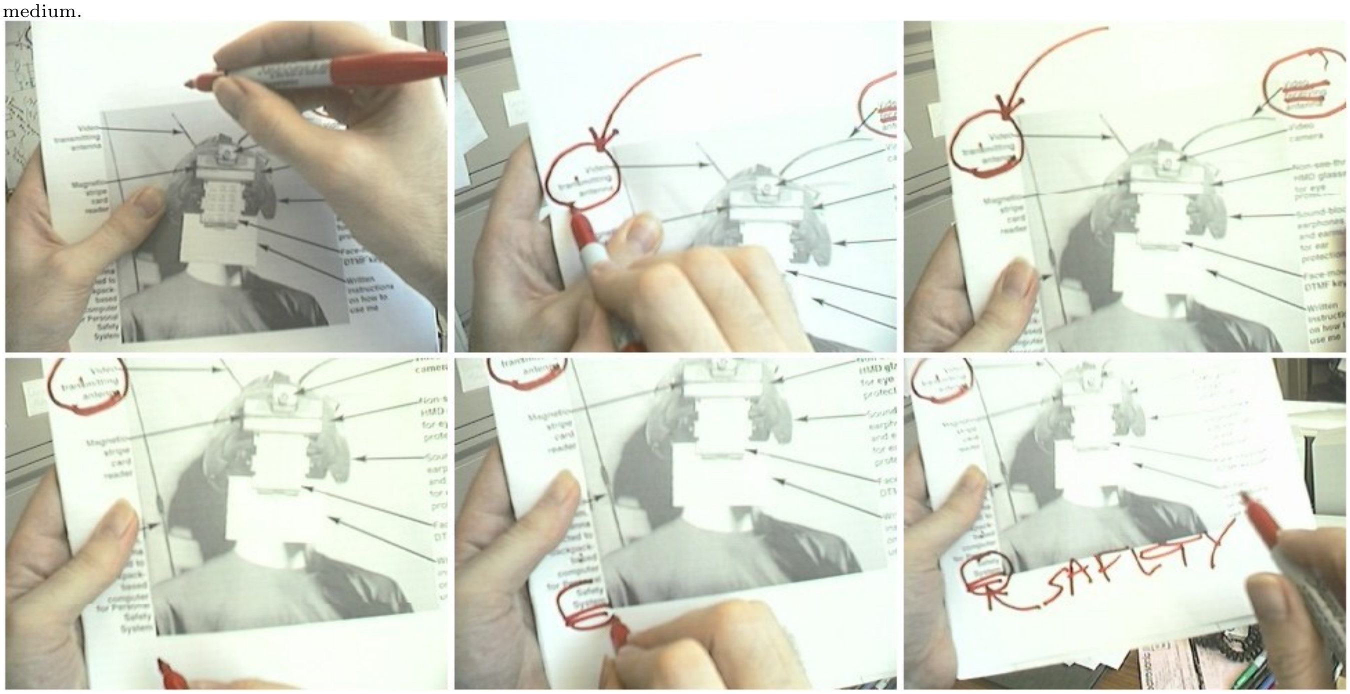

Figure 11: The eye itself as a document camera: Annotation of existing media: eye as document camera.

Figure 13: Shared computer mediated reality makes possible the personalized experience capture of a “cyborg” collective.

Figure 14: Lightvector painting made in Harold Edgerton's Strobe Lab, exactly as it was left after his death. This image is from the Microseconds and Millennia exhibit of the author's work at Olga Korper Gallery. The beautiful glow in a closeup detail of the flash tube, located above the 30,000 volt potentiometer control, was extracted, magnified, and used to replace a cigarette advertisement.

Figure 15: "Adissolve" project: Advertisements dissolve into fine art. Here a cigarrette ad dissolves into a closeup view of the lamp glow in a selection from the Microseconds and Millennia exhibit at Olga Korper Gallery. The top row shows video frames from a linear dissolve (image-based rendering). The bottom row shows lightspace rendering. While the end points look very similar, midway along the dissolve, the lightspace-based rendering looks much more realistic, and is more vibrant than the image-based rendering.

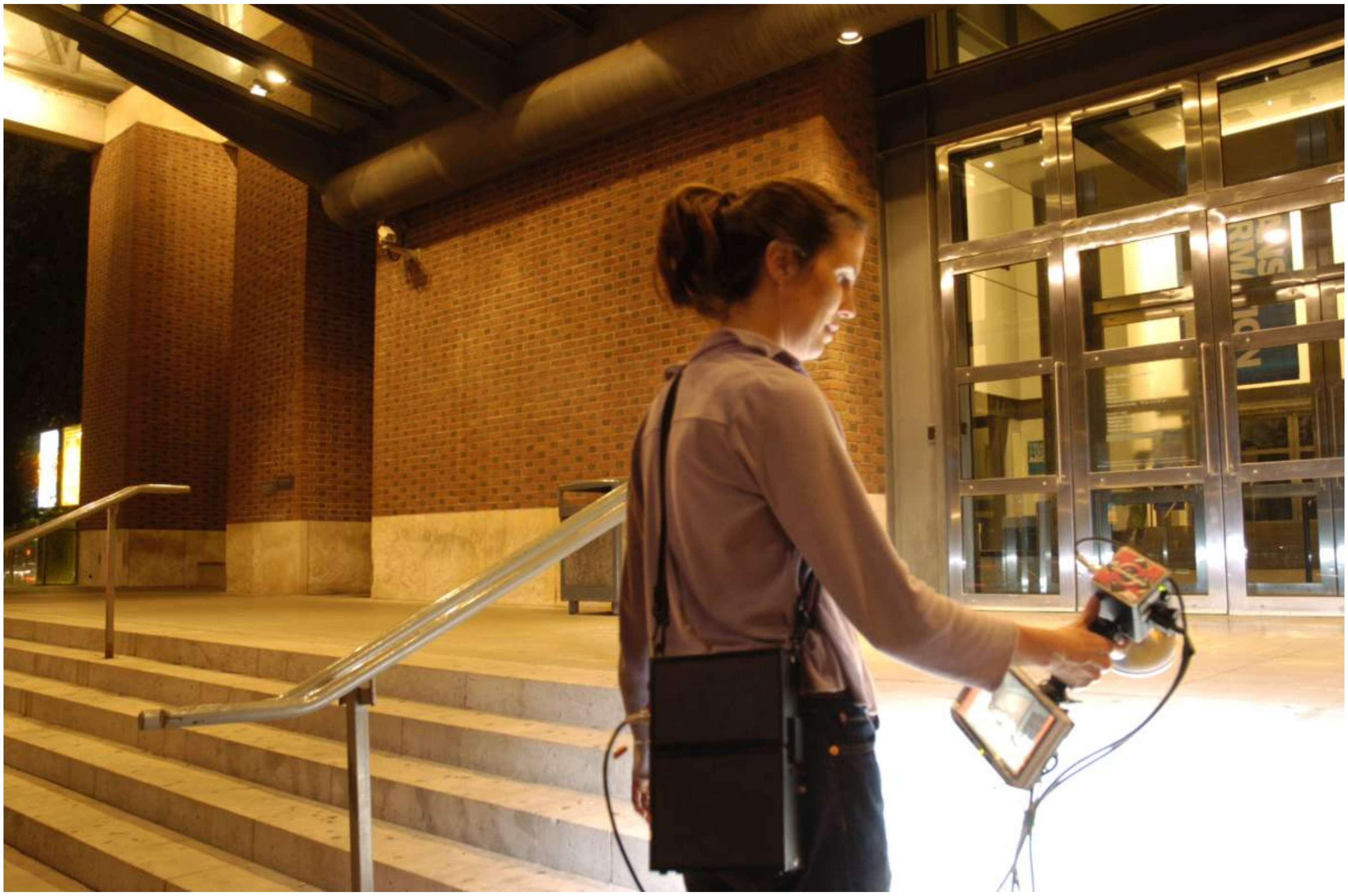

Figure 16: A new and simple keyer grip interface for lightvectoring.

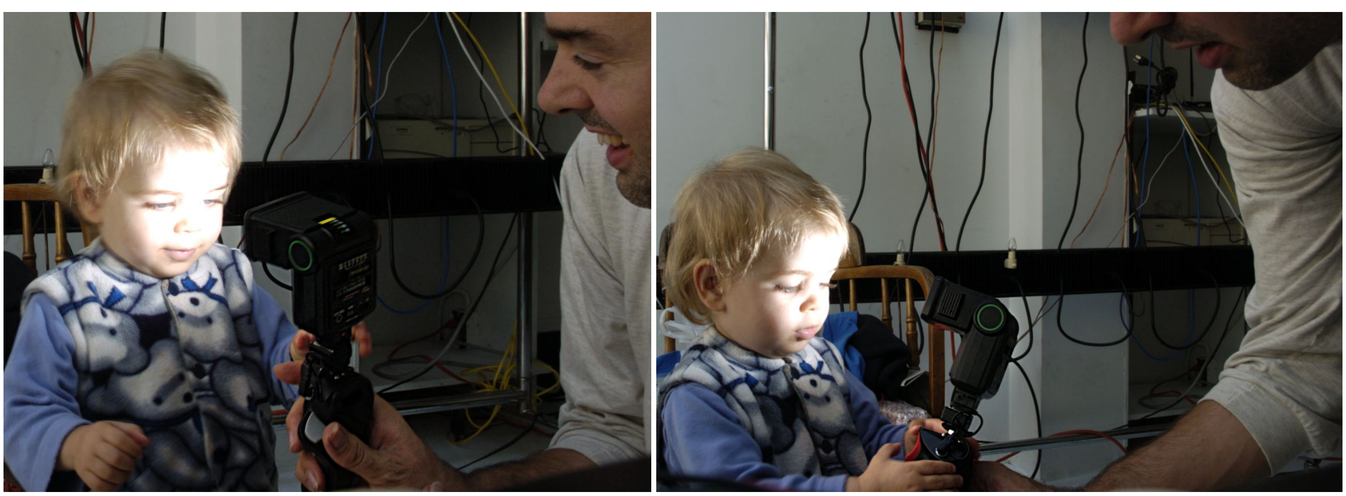

Figure 17: Christina familiarizing herself with the keyer.

Figure 18: Self portrait by Christina, single lightvector.

Figure 19: Lightvector painting by Christina Mann.

Figure 23: These signs were put up around the same time as surveillance cameras were installed in High Park, in Toronto.

Figure 25: A more literal reversal of surveillance brings the familiar ceiling domes of wine-dark opacity from the heavens down to earth (i.e. from the lamp posts and ceilings down to human level). This form of sousveillance fashion re-situates the surveillance dome as an accessory to a wearable computer system, as a Sousveillance Situationist challenge to our pre-conveived notions of surveillance. Although such human-centered sousveillance maintains the familiar appearance of surveillance (i.e. matches the decor of just about any gambling casino or department store), an eye-centered design (i.e. EyeTap) was found to provide much more useful data.