Personal Imaging and *lookpainting* as tools for personal documentary and investigative photojournalism

Steve Mann

Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Room 2001, Toronto, Ontario, Canada M5S 3G4

A means and apparatus for covert capture of extremely high-resolution photorealistic images is presented. The apparatus embodies a new form of user-interface – instead of the traditional "point and click" metaphor which was thought to be the simplest photography had to offer, what is proposed is a "look" metaphor in which images are generated through the natural process of looking around, in a manner that does not require conscious thought or effort. These "lookpaintings" become photographic/videographic memories that may, at times, exceed the quality attainable with even large and cumbersome professional photographic film cameras, yet they may be captured through a device that resembles ordinary sunglasses. The method is based on long-term psychophysical adaptation using a covert sunglass-based reality-mediating apparatus, together with two new results in image processing. The first new result is a means of estimating the true projective coordinate transformation between successive pairs of images, and the second is that of estimating, to within a single unknown scalar constant, the quantity of light arriving at the image plane. Furthermore, what is captured is more than just a picture. The resulting environment map may be explored by one or more remote participants who may also correspond and interact with the wearer during the actual shooting process, giving rise to computer supported collaborative (collective) photography, videography, shared photographic/videographic memory, etc.

1. Introduction

A central goal of *Personal Imaging* is best captured in the following quote:

> The technologies for recording events lead to a curious result... Vicarious experience, even for those who were there. In this context "vicarious" means to experience an event through the eyes (or the recording device) of another. Yet here we have the real experiencer and the vicarious experiencer being the same person, except that the real experiencer didn't have the original experience because of all the activity involved in recording it for the latter, vicarious experience... we are so busy manipulating, pointing, adjusting, framing, balancing, and preparing that the event disappears... But there is a positive side to the use of recording devices: situations where the device intensifies the experience. Most of the time this takes place only with less sophisticated artifacts: the sketch pad, the painter's canvas... Those who benefit from these intensifying artifacts are usually artists... with these artifacts, the act of recording forces us to look and experience with more intensity and enjoyment than might otherwise be the case (Don Norman, [1]).

1.1. The photographic origins of the WearComp/WearCam project

The original motivation behind the WearComp project [2] was an attempt to define a new genre of imaging characterized by unprecedented control over lighting, and, to create a tool that could allow reality to be experienced with greater intensity and enjoyment than might otherwise be the case. This tool (figure 1) functioned much more like the *sketch pad* or the *painter's canvas* to which Norman refers, than like the underlying cameras that it embodied.

Figure 2 depicts two early 1980s attempts at creating expressive images using the personal imaging system developed by the author in the 1970s and early 1980s.

1.2. Photographic/videographic memory system architecture

The architecture of the author's early WearCam systems typically involved a mixture of digital and analog video signals, along with a hybrid (digital and analog) communications network, as depicted in figure 3. There were three possible video signal paths that could be switched, selected, or mixed in various proportions, as desired, by way of the computer-controlled analog NTSC video switcher/mixer. These three signal paths comprised:

- Direct path from video camera to display.
- Locally mediated path from camera, through processor, to display.
- Remotely mediated path from camera, through remotely located processor (by way of communications system), to display.

This architecture also provided a mixture of local and remote video recording/archival capabilities which proved useful for the acquisition, production, editing, and dissemination of documentary videos, etc.

(a) (b)

Figure 1. Early embodiments of the author's original "photographer's assistant" application of Personal Imaging. (a) 1970s "lightpainting pushbroom" system with 1980 CRT display. The linear array of lamps, controlled by a body-worn processor (WearComp), operated much like a dot-matrix printer to sweep out spatial patterns of structured light. (b) As this project evolved from the 1970s into the early 1980s, the components were typically spread out on clothing rather than located in a backpack. Separate 0.6 inch cathode ray tubes attachable/detachable to/from ordinary safetyglasses, as well as waist-worn television sets were typically used instead of the earlier and more cumbersome helmet-based screens of the 1970s. Note also the change from the two antennas in (a) to the single antenna in (b), which provided wireless communication of video, voice, and data to a remote base station. The use of black clothing (made of black velvet or black suede), a black hood, etc., sometimes together with black makeup, was typical of the optimum photographic function (turning the whole body into an optimal camera/imaging system). Early on, the author's body evolved this way in the same way that camera bodies are typically black so as to minimize light scatter or reflection.

1.3. 'Underwearables': Covert embodiments of WearComp/WearCam

The early personal imaging systems were characterized by a heavy and obtrusive nature. Due to much criticism that the author received in wearing these systems in day-to-day life, a new generation of unobtrusive personal imaging systems was developed. The current unobtrusive visual mediation system is concealed in a pair of ordinary sunglasses which are connected to the underwearable computer. Typical embodiments of the underwearable computer resemble an athletic undershirt (tank top) made of durable mesh fabric, upon which a lattice of webbing is sewn. This facilitates quick reconfiguration in the layout of components, and re-routing of cabling.

(a) (b)

Figure 2. Norman's criticism of the camera arises from the fact that, in many ways, it diminishes our perception and enjoyment of reality. However, a goal of Personal Imaging [2], through the apparatus depicted in figure 1, is to create something much more like the sketch pad or artist's canvas than like the camera in its usual context. The images produced as artifacts of Personal Imaging are somewhere at the intersection of painting, computer graphics, and photography. (a) Notice how the broom, dilapidated chair, and flower pot (especially the dead plant inside) appear to be their own light sources (e.g., self-illuminated), and 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/lightpainting [3] method. (b) Hallways offer a unique perspective, which can also be illuminated expressively.

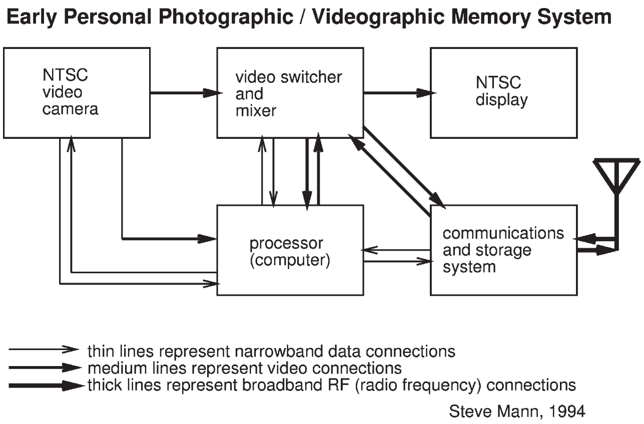


Figure 3. Block diagram of early Personal Imaging body-worn system hardware. The processor (computer) was connected to a communications system, and, therefore, to other computers at remote locations (either worn by other people, or fixed in the environment) by a digital data link. The computer also controlled parameters (settings) of an analog NTSC video camera as well as an NTSC video switcher/mixer, so that the analog video from the camera could be routed to the NTSC display, providing a viewfinder function. The processor (computer) also produced an NTSC output that was fed to the video mixer for overlays. Moreover, video from the camera could instead be directed first through the processor prior to display, so that the visual perception of reality could be altered (mediated). The computer-controlled video switcher facilitated either locally mediated reality (by way of the body-worn processor), or remotely mediated reality (simultaneously sending and receiving analog NTSC video over the communications link for remote processing) or a combination of both of these.

An example of the underwearable multimedia computer video editing facility and wireless communication system, as it normally appeared when worn under clothing, is depicted in figure 4, where the normal appearance is quite evident. Covert data-entry devices typically comprised switches located on the underwearable (undergarment) itself. These switches were easily actuated by pressing through clothing worn over the apparatus. Alternatively, a belt-mounted input device was sometimes used (figure 5).

It should be noted that this class of system is more than just a wearable computer as might send and receive email, but, rather, it is a complete WearComp/WearCam personal imaging system, as defined in [2]: it contains special-purpose image processing hardware [4] and a complete video editing facility developed for the creation of the investigative metadocumentary ShootingBack (a documentary about making a documentary about video surveillance). The unobtrusive nature was necessary because it was found that the locations selected in which to shoot the documentary – establishments where video surveillance is used extensively (e.g., gambling casinos, department stores, banks, etc.) – also prohibit photography and video other than by their own cameras. Part of the purpose in constructing this apparatus was to challenge/investigate the nature of these one-sided (e.g., as satisfy the formal definition of "totalitarian") establishments, and shoot in establishments where photography/video was strictly prohibited.

Due to its ordinary appearance, the unobtrusive personal imaging system also suggests practical utility in everyday matters of personal life, such as personal documentary, personal safety, crime reduction, as well as investigative photojournalism.

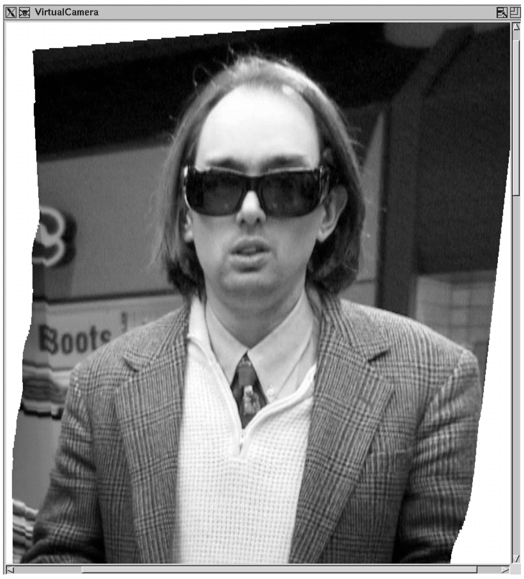


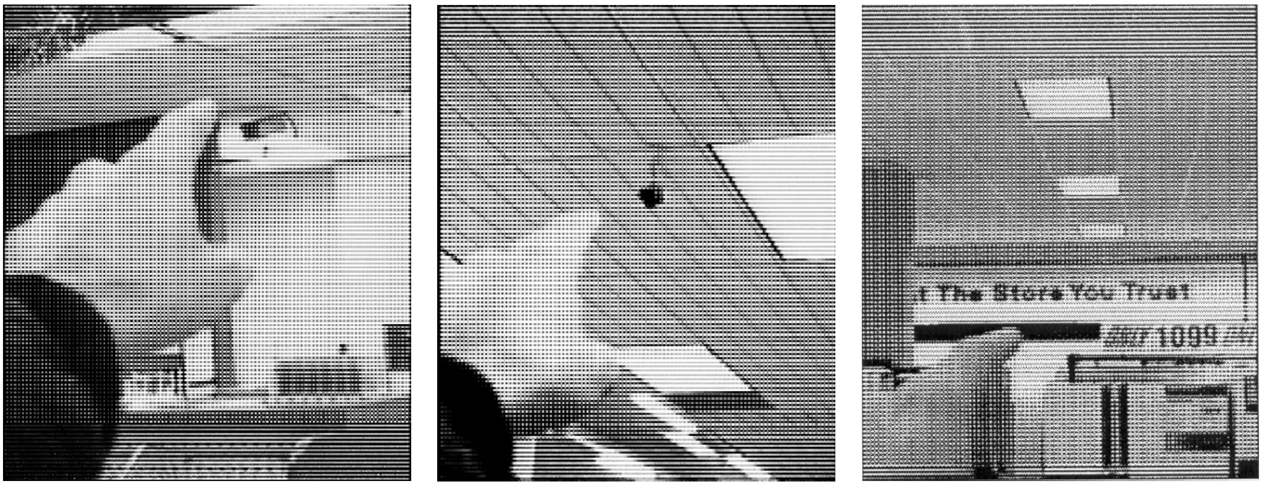
Figure 4. Author wearing covert embodiment of WearComp photo/video memory system suitable as a personal visual memory prosthetic, or for investigative documentary/photojournalism. The system incorporated fully functional UNIX-based computers concealed in the small of the back, with the rest of the peripherals, analog to digital converters, etc., also concealed under ordinary clothing. Sensor arrays were concealed within the eyeglasses, used in the context of Personal Imaging. The full-color version completed in 1996 included special-purpose digital signal processing hardware based on TMS 320 series processors connected to a UNIX-based host processor, concealed in the back of the underwearable. The cross-compiler for the TMS 320 series chips was run remotely on a SUN workstation, accessed wirelessly through radio and antennas concealed in the apparatus. Thus it could be programmed (re-configured) while being worn, without the need to dock to a programming station.

2. Deconfigured eye: On becoming a camera

It should be noted that the methodology of the new cinematographic and photographic genre characterized by personal imaging [4] differs from current investigative journalism (e.g., miniature cameras hidden in the jewel of a tie clip, or in a baseball cap), in the sense that a long-term adaptation process, as described in [5] (e.g., often taking place over a period of many years) makes the camera behave as a true extension of the mind and body, and that the ability to augment, diminish, or otherwise alter the perception of reality is exploited fully, in the capturing of a much richer and more detailed perception of reality that captures the essence of wearer-involvement in his/her interaction with the world.

It should also be noted that the underlying principles of mediated reality (MR) [5] differ from augmented reality (AR) where additional information is *added* onto the real world (e.g., through a display with a beamsplitter). Mediated reality involves, in addition to the capability of augmenting reality, the capability of also diminishing or altering the perception of visual reality. Thus the personal imaging device must be fully immersive, at least over a certain so-called *mediation zone* [5]. A simple example of the utility of *diminished reality* is quite evident in the documentary video ShootingBack [6,7] when, for example, the author is asked to sign a bank withdrawal slip. Because of the deliberately diminished reality, it is necessary that the author bring his head very close to the written page (distance depending on the size of the lettering), in order to see it. A side effect of doing so is that video is produced in which the audience can also see the fine print, whereas

Figure 5. Covert belt-based input device operated by right hand, reaching behind back. Typically this device may be hidden underneath an untucked T-shirt or the like. The units that look like toggle switches are really spring-loaded extremely light-touch lever rockers.



(a) (b) (c)

Figure 6. Living in a 2-D world, through long-term adaptation. Fingerpointing from the perspective of *life through the screen*. Adaptation was in the "rot90" coordinate transformation described in the text. Notice how natural interactions with the world (such as pointing at the video surveillance cameras on the ceiling of the department store, or the signage on the wall) take place in 2-D projection on the image plane rather than 3-D space.

shooting in a traditional investigative documentary style, this would not be so.

Once the reality mediator is worn for some time, and one becomes fully accustomed to experiencing the world through it, there is a certain synergy between human and machine that is not experienced with a traditional camera. More subtle differences between a recording made from the output of an MR and that made from a conventional body worn camera (such as might be hidden in the jewel of a tie clip) include the way that when one is talking to two people the closing of the loop forces one to turn one's head back and forth as one talks to one person, and then to the other. This need arises from the limited peripheral vision the apparatus imposes on the wearer, which is yet another example of a deliberate diminishing of reality in order to heighten the experience of reality.

ShootingBack was shot through what amounted to "rot90" (90 degree rotating) eyeglasses, similar to George Stratton's upside-down glasses [8] and Kohler's left-right reversing glasses [9], but where the altering of visual reality was achieved through computational means rather than optical means. Furthermore, in addition to being another long-term psychophysical adaptation experiment, ShootingBack enabled new heights to be reached in concentration and seeing everything in a much more intensified way. Thus, just as copy editors often read their typed manuscripts in a mirror, so that they will spot all the typographical errors to which they had previously been *error blind*, in ShootingBack, the author learned how to see all over again. Similarly, just as the artist's sketch pad, because of its crudeness, forces us to concentrate and to really "see", mediated reality becomes an intensifying artifact in which the act of recording forces one to look and experience with more intensity and enjoyment than might otherwise be the case.

2.1. From "fly on the wall" documentary to 'fly in the eye' personal documentary

The reality mediator goes beyond merely functioning as a viewfinder to set the camera correctly because the human operator is completely in the loop (e.g., one will trip and fall if it is not set correctly, since one will not be able to see properly), so that it transcends being a mere compositional tool, toward allowing the camera to "become" the eye of the wearer.

The nature of Personal Imaging is such that the combination of camera, processor, and display results in a visual adaptation process, causing it to function as a true extension of the mind and body: after an extended period of time, it does not seem to the wearer to be an external entity. Images captured on WearCam have a certain expressive quality that arises from having the wearer be "in the loop" – part of the complicated feedback process that synergizes both human and machine into a single unit.

2.2. Living in a 2-D world

In addition to adapting to transformations of the perceptual world, the author noticed a loss of the perception of 3-D depth. (A typical video camera lacks depth from stereo, depth from focus, etc.) In this sense, the author developed a "photographic mindset" in which an enhanced sense of awareness of light and shade, and of simple renaissance perspective were attained. It was found [4] that this effect persisted, even when the apparatus was removed, and that the effect would revisit in the form of 2-D "flashbacks", so that the world was seen in two ways, much like we see the Necker cube illusion in two possible ways. This discovery gave rise to the fingerpointing process (figure 6) where it was found that pointing at objects was as though a 2-D plane projection. The notion of attaching a light to

(a)

(b)

Figure 7. Examples of tracing out a locus of points in 3-D space that are mapped onto a 2-D image. Here a small light source, attached to the author's finger, takes the form of a pointing device, which is used to outline objects in 3-D space, but falling upon their 2-D projection. (a) One of the early lightpaintings using this technique. (b) Image which won best color entry, in the National Fuju Film competition, 1986. Here the method is perfected somewhat, where we begin to see how the Personal Imaging system functions more like the artist's sketch pad than like a traditional camera. (C) Steve Mann, 1985.

the finger arose out of various expressive lightpainting efforts, where the world was viewed as 2-D video, while a light source, attached to the finger, was moved around in 3-D space (figure 7), while recording the process simultaneously on film (a beam splitter being used to combine video and film cameras).

In mediated reality the drawing takes place right on top of the video stream, so that registration is, for all practical purposes, exact to within the pixel resolution of the devices [5], in contrast to the registration problem of augmented reality [10]. This characteristic of mediated reality (perfect registration) has been suggested as a means of using the finger as a mouse to outline actual objects in the scene [2]. This form of interaction with the real world, through the apparatus, is yet another example of the human–machine symbiosis that is at the core of Personal Imaging.

3. *Lookpainting*: Towards developing a new camera user-interface

3.1. Introduction: What is lookpainting

One characteristic of the reality mediator is that, after wearing it for an extended period of time, one forgets that one is wearing it[1], while at the same time, it provides an output signal of whatever passes from the real world through to the human visual system of the wearer [4]. In addition to the personal documentary capabilities of this "video tap", into the visual system, there is the possibility of creating high resolution environment maps by standing in one place and looking around. Assuming that one can look around faster than objects in the scene can change, images are typically found to be in approximately the same orbit of the projective group of coordinate transformations [11], modulo any changes in overall exposure level brought on by the automatic exposure mechanism of the camera [12].

3.2. Building environment maps by looking around

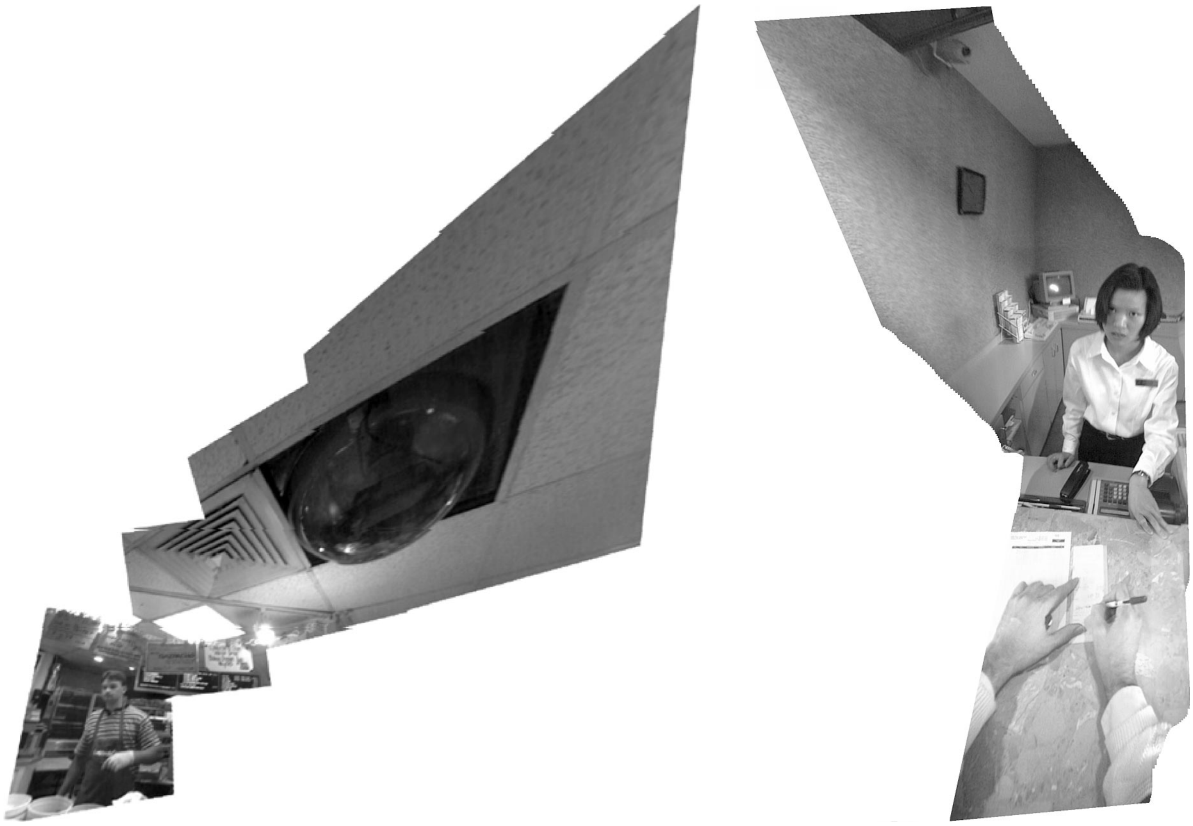
An environment map is a collection of images, seamlessly "stitched" together, into some unified representation of the quantity of light that has arrived from each angle in space, over the range of angles for which there exist measurement data. Examples of environment maps appear in figure 8.

In constructing these environment maps, the computer performs basic calculations which it is good at, while the human operator makes higher level decisions about artistic content, what is of greatest interest or importance, etc.

As described earlier, the human becomes at one with the machine (in this case the camera) through a long-term adaptation process, so that, as one experiences one's life through the apparatus (living in a computer-mediated world), the subject matter of interest is automatically captured by the human operator. Note that in this simple case, there is no Artificial Intelligence, but instead, there is a synergy between human and machine, where the "intelligence" arises through having the human operator in the feedback loop of the overall image acquisition process.

Personal imaging, which is facilitated through WearComp, image processing, machine vision, and computer-

[1] Furthermore, owing to its lightweight and unobtrusive nature, others do not regard it is unusual either.



(a) (b)

Figure 8. Environment maps captured from a first-person perspective, through the process of looking around, are called “lookpaintings”. Lookpaintings are characterized by irregularly-shaped boundaries which capture the gaze pattern of the wearer. (a) Lookpainting made from 4 input images. Individual image boundaries are clearly visible. Note the unified perspective and the lack of distortion (e.g., lines on the ceiling tiles are almost perfectly straight, despite the extreme perspective). (b) Lookpainting made from 226 input images. This composite image illustrates the nature of first-person perspective. Note that both the author’s hands are visible in the picture. Because the apparatus is wearable, it captures a new point of view, while at the same time, capturing what is important in a scene. (Here, for example, the lookpainting has included the video surveillance camera on the ceiling because the author has looked there.)

mediated reality, enables the user to effortlessly capture high-quality images of a new genre characterized by not only enhanced tonal range and spatial resolution, but also by the ability to include and exclude areas of interest (figure 9).

3.3. Mathematical framework for lookpainting: Video Orbits and the Wyckoff principle

The mathematical framework for determining the unknown nonlinear respose of a camera from nothing more than a collection of differently exposed images, as well as a means of extending dymanic range by combining differently exposed images was first published in 1993 [14]. This framework is now described in detail.

A set of functions,

$$I_i(\mathbf{x}) = f\left(k_i q\left(\frac{\mathbf{A}_i\mathbf{x}+\mathbf{b}_i}{\mathbf{c}_i\mathbf{x}+d_i}\right)\right),$$

$$\mathbf{x} \in \mathbb{R}^{N\times 1}, \quad \mathbf{A}_i \in \mathbb{R}^{N\times N}, \quad \mathbf{b}_i \in \mathbb{R}^{N\times 1},$$

$$\mathbf{c}_i \in \mathbb{R}^{1\times N}, \quad d_i \in \mathbb{R}, \quad k_i \in \mathbb{R}, \qquad (1)$$

is known as a projective-Wyckoff set [4,12,13]. When $N = 2$, the projective-Wyckoff set describes, within a common region of support, a set of images, I_i, where $\mathbf{x} = (x, y)$ is the continuous spatial coordinate of the focal plane of an electronic imaging array or piece of film, q is the quantity of light falling on the sensor array, and f is the unknown nonlinearity of the camera’s response function (assumed to be invariant to $\mathbf{x}$). In this case, the constants $\mathbf{A}$, $\mathbf{b}$, $\mathbf{c}$, and d are the parameters of a projective spatial coordinate transformation, and the scalar parameter k is the parameter of a tonal transformation (arising from a change in exposure from one image to the next). In particular, the constant $\mathbf{A} \in \mathbb{R}^{2\times 2}$ effects a linear coordinate transformation of the image. (This constant represents the linear scale of the image, as might compensate for various zoom settings of a camera lens, rotation of the image, and image shear). The constant $\mathbf{b} \in \mathbb{R}^2$ effects translation of the image, and the constant $\mathbf{c} \in \mathbb{R}^2$ effects “chirping” of the image. The constants $\mathbf{b}$ and $\mathbf{c}$ may each be regarded as a vector with direction and magnitude (e.g., translation magnitude and direction, as well as chirp rate and chirp direction).

Methods of simultaneously estimating the parameters of this relationship (1) between images having a common region of support, allowing the parameters, $\mathbf{A}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$, d_i,

Figure 9. This image depicts a group of people to whom the author is lecturing. Here the author is able to quickly sweep out the important details of this scene (namely all of the participants), while leaving out areas of the room where nobody is seated. Note how the image is close-cropped, leaving out the two empty chairs in the center, while at the same time, extending out to include the feet of those sitting to the left and right of the empty chairs. This natural selection of subject matter happens often without conscious thought or effort, through the process of simply "looking around" at everyday scenes or objects, and is characteristic of the symbiotic relationship between human and machine that arises when the two become inextricably intertwined through a constancy of user-interface extending over a time period of many years.

and k_i, as well as the function f, to be determined from the images themselves, have been proposed [4,12–14]. An outline of these methods follows, together with some new results. The method presented in this paper differs from that of [14] in that the method presented here emphasizes the operation near the neighbourhood of the identity (e.g., the "video orbit").

For simplicity of illustration, let us consider the case for which $N = 1$ (e.g., so that pictures are functions of a single real variable). In actual practice, $N = 2$, but the derivation will be more illustrative with $N = 1$, in which case, $\mathbf{A}_i$, $\mathbf{b}_i$, $\mathbf{c}_i$, d_i, and k_i, are all scalar quantities, and will thus be denoted a_i, b_i, c_i, d_i, and k_i, respectively.

For simplicity of illustration (without loss of generality), also suppose that the projective-Wyckoff set contains two pictures,

$$I_1 = f(q) \quad \text{and} \quad I_2 = f\left(kq\left(\frac{ax+b}{cx+d}\right)\right),$$

where I_2, called the *comparison frame*, is expressed in the coordinates of I_1, called the *reference frame*. Implicit in this change of coordinates is the notion of an underlying group representation for the *projective–Wyckoff operator*, p_{12}, which maps I_2 as close as possible to I_1, modulo cropping at the edges of the frame, saturation or cutoff at the limits of exposure, and noise (sensor noise, quantization noise, etc.):

$$\hat{I}_1 = p_{12} I_2, \tag{2}$$

where $\hat{I}_1$ is the best approximation to I_1 that can be generated from I_2.

A suitable group representation is given by

$$p_{a,b,c,d,k} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & k \end{bmatrix}. \tag{3}$$

Thus, using the group representation (3), we may rewrite the coordinate transformation from any frame to any other, as a composition of pairwise coordinate transformations. For example, to obtain an estimate of image frame I_1 from any image, say, I_n, we observe

$$\hat{I}_1 = p_{12} p_{23} p_{34} \ldots p_{n-1,n} I_n, \tag{4}$$

where $p_{i-1,i}$ is the coordinate transformation from image frame I_i to image I_{i-1}. The group representation (3) provides a law of composition for these coordinate transformation operators, so that it is never necessary to resample an image more than once.

Photographic film was traditionally characterized by the so-called "Density versus log Exposure" *characteristic curve* [15,16]. Similarly, for the CCD sensor arrays typically concealed in the sunglass-based reality mediators, logarithmic exposure units, $Q = \log(q)$, may also be used, so that one image will be $K = \log(k)$ units darker than the other:

$$\log\left(f^{-1}(I_1(\mathbf{x}))\right) = Q = \log\left(f^{-1}\left(I_2\left(\frac{ax+b}{cx+d}\right)\right)\right) - K, \tag{5}$$

where the difference in exposure, K, arises from the fact that the camera will have an automatic exposure control of sorts, so that, while looking around, darker or lighter objects will be included in the region of the image which causes a global change in exposure.

The existence of an inverse for f follows from the semi-monotonicity assumption. Semi-monotonicity follows from the fact that we expect[2] pixel values to either in-

[2] Except in rare instances where the illumination is so intense as to damage the imaging apparatus, as, for example, when the sun burns through photographic negative film and appears black in the final print or scan.

crease or stay the same with increasing quantity of illumination, q. This is not to suggest that the image content is semi-monotonic, but, rather, that the response of the camera is semi-monotonic. The semi-monotonicity assumption thus applies to the images after they have been registered (aligned) by a spatial coordinate transformation.

Since the logarithm function is also monotonic, the problem comes down to estimating the semi-monotonic function $F() = \log(f^{-1}())$ and the parameters a, b, c, d, and $K = \log(k)$, given two pictures I_1 and I_2:

$$F(I_1(\mathbf{x})) = F\left(I_2\left(\frac{ax+b}{cx+d}\right)\right) - K. \tag{6}$$

Rather than solve for F, it has been found that registering images to one reference image is more numerically robust [4]. In particular, this is accomplished through an operation of the form

$$\hat{I}_1(\mathbf{x}) = F^{-1}\left(F\left(I_2\left(\frac{ax+b}{cx+d}\right)\right) - K\right), \tag{7}$$

which provides a recipe for spatiotonally registering the second image with respect to the first (e.g., appropriately lightening or darkening the second image to make it have, apart from the effects of noise – quantization noise, additive noise, etc. – the same tonal values as the first image, while at the same time "dechirping" it with respect to the first). This process of "registering" the second image with the first differs from the image registration procedure commonly used in much of image-based rendering [17–19], machine vision [20–23] and image resolution enhancement [11,24–26] because it operates on both the *domain* and the *range*, $f(q(\mathbf{x}))$, (tonal range) of the image $I(\mathbf{x})$ as opposed to just its *domain* (spatial coordinates) $\mathbf{x} = (x, y)$. Image processing done on range-registered images is also related to the notion of nonlinear mean filters [27], and range-registration, as well as other forms of range-based processing are also of special interest. Whether processing in a range-registered function space, or processing quantigraphically [4], it is often useful to consider the relationship between two differently exposed images [14].

Proposition 3.1. When a function $f(q)$ is monotonic, the parametric plot $(f(q), f(kq))$ can be expressed as a function $g(f)$ not involving q.

Definition 3.1. The resulting plot $(f, g(f)) = (f(q), f(kq))$ is called the *comparametric* plot [12].

The function g is called the *comparametric function*, and expresses the range of the function $f(kq)$ as a function of the range of the function $f(q)$, and is independent of the domain, q, of the function $f(q)$.

Separating the estimation process into two stages also allows us a more direct route to "registering" the image ranges, if, for example, we do not need to know f, but only require a recipe for expressing the range of $f(kq)$ in the units of $f(q)$.

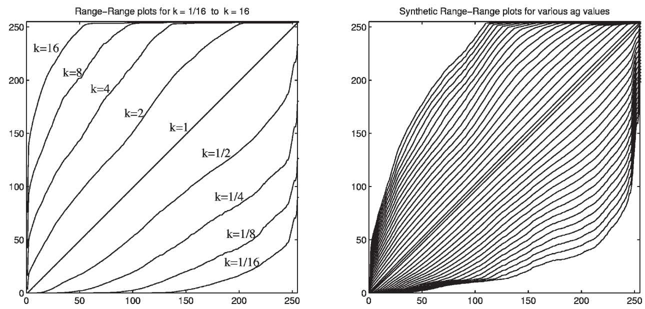


Figure 10. The family of comparametric curves, g_i, for various values of k_i capture the unique nonlinear response of the particular camera under test (a 1/3 inch CCD array concealed inside eyeglasses as part of a reality mediator, together with control unit). The form of this curve may be determined from two or more pictures that differ only in exposure. (a) Curves as determined experimentally from pairs of differently exposed pictures. (b) Curves as provided by a one-parameter model, for various values of the parameter.

The determination of the function g (or f) can be done separately from the determination of the parameters **A**, **b**, **c**, d, and k. The determination of g is typically done by comparing a variety of images differing only in exposure. The function g is thus parameterized by the exposure k, and for a given camera, the camera's response function is characterized by a family of curves g_i, one for each possible k_i value, as illustrated in figure 10. These curves are found by slenderizing the joint histogram between images that differ only in exposure (each such image pair determines a curve g for the particular k value corresponding to the ratio of the exposures of the two images). The slenderizing of the joint histogram amounts to a non-parametric curve fit, and is what gives rise to the determination of g.

The above method allows us to estimate, to within a constant scale factor, the continuous (unquantized) response function of the camera without making any assumptions on the form of f, other than semi-monotonicity.

3.4. Parametric methods

In situations where the image data is extremely noisy, and/or where a closed-form solution for $f(q)$ is desired, a parametric method is used, in which a function $g(f)$ that is the comparametric function of a suitable response function $f(q)$ is selected. The method amounts to a *curve fitting* problem in which the parameters of g are selected so that g best fits one or more joint histograms constructed from two or more differently exposed spatially registered images under analysis. Various parametric models have been considered for this purpose [12]. For example, the one parameter model, $f(q) = \exp(q^\Gamma)/e$, is equivalent to applying gamma correction $\gamma = k^\Gamma$ to f in order to obtain $g = f(kq) = f^\gamma$, where Γ is typically between 0 and $1/2$. Note that when tonally registering images using this gamma correction, one makes an implicit assumption that $f(q)$ does not go all the way to zero when $q = 0$. Thus a better model is $f = \exp((1 + q^b)^c)$, which also nicely captures the toe and shoulder of the response function.

However, one of the simplest models from a parameter estimation point of view is the classic response curve,

$$f(q) = \alpha + \beta q^{\gamma}, \tag{8}$$

used by photographers to characterize the response of a variety of photographic emulsions (and suitable also to characterize the response of electronic imaging devices which attempt to emulate the appearance of film), including so-called *extended response* film [15].

Proposition 3.2. The comparametric plot corresponding to the standard photographic response function (8) is a straight line. The slope is k^{γ}, and the intercept is $\alpha(1 - k^{\gamma})$.

Proof. $g(f(kq)) = f(kq) = \alpha + \beta(kq)^{\gamma}$. Re-arranging to eliminate q gives $g = k^{\gamma}(\alpha + \beta q^{\gamma}) + \alpha(1 - k^{\gamma})$ so that

$$g = k^{\gamma} f + \alpha(1 - k^{\gamma}). \tag{9}$$

This result suggests that $f(q)$ can be determined from two or more differently exposed images by applying linear regression to the joint histogram, J_{12}, of the images, I_1 and I_2.

When two successive frames of a video sequence are related through a group-action that is near the identity of the group, one may think of the Lie algebra of the group as providing the structure locally, so that (7) may be solved as follows: From (8) we have that the (generalized) brightness change constraint equation [28] is: $g(f(q(x,t))) = f(kq(x,t)) = f(q(x + \Delta x, t + \Delta t)) = k^{\gamma} f(q(x,t)) + \alpha - \alpha k^{\gamma} = k^{\gamma} I(x,t) + \alpha(1 - k^{\gamma})$, where $I(x,t) = f(q(x,t))$. Combining this equation with the Taylor series representation,

$$\begin{aligned} I(x + \Delta x, t + \Delta t) = {} & I(x,t) + \Delta x I_x(x,t) \\ & + \Delta t I_t(x,t) + \cdots \end{aligned} \tag{10}$$

where $I_x(x,t) = (\mathrm{d}f/\mathrm{d}q)(\mathrm{d}q(x)/\mathrm{d}x)$, at time t, and $I_t(x,t)$ is the frame difference of adjacent frames, we have

$$k^{\gamma} I + \alpha(1 - k^{\gamma}) = I + \Delta x I_x + \Delta t I_t. \tag{11}$$

Thus, the brightness change constraint equation becomes

$$I + uI_x + I_t - k^{\gamma} I - \alpha(1 - k^{\gamma}) = \varepsilon, \tag{12}$$

where, normalizing, $\Delta t = 1$. □

3.5. *Estimating in the the neighbourhood of the identity of a group of coordinate transformations*

The procedure for determining the coordinate transformation is a repetitive procedure in which, for each image pair, first an estimate is made of the parameters of an approximate model of the transformation going from one image to the other, and then the corresponding parameters of the actual ("exact") transformation are determined from those of the approximate model. The parameters of the "exact" model are then used to transform one of the images in a first attempt to register it with the other image. The registered images should, ideally, be identical within

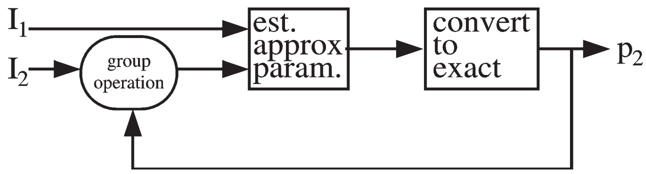


Figure 11. Diagram depicting the "lookpainting" algorithm: parameters are estimated from pairwise successive image frames, such as I_1 and I_2 in the image sequence I_i. An approximate model that operates in the neighbourhood of the identity may be used, and converted to the actual ("exact") model. As a result of the feedback loop, the parameters of the exact model p_{21} relating images I_1 and I_2 are estimated. Although this method involves repeated estimation of the parameters, it should be noted that it has advantages over the method presented in [25,26] in the sense that the estimation of the approximate parameters is non-iterative. It should thus be emphasized that the estimation process each time around the loop is direct (non-iterative).

their region of common support (overlap), but, there will in practice be some residual error. Thus the process is repeated each time, in order to reduce the error, as depicted in figure 11. In this way, the approximate model operates in the feedforward path, and the exact model is implemented by way of the feedback path. It should be noted that cumulative applications of coordinate transformations (e.g., coordinate transformations of coordinate transformations) are not done, but, rather, the group structure and its law of composition are used so that there is only one composite coordinate transformation done each time.

The approximate model used for the spatial coordinate transformation is that of a quadratic Taylor series of the projective coordinate transformation. Substitution of the approximate model into (12) gives:

$$\begin{aligned} & I + \left(q_2 x^2 + q_1 x + q_0\right) I_x + I_t - k^{\gamma} I \\ & \quad - \alpha + \alpha k^{\gamma} = \varepsilon, \end{aligned} \tag{13}$$

where $q_2 = (bc/d^2 - a/d)c/d$, $q_1 = a/d - bc/d^2$, and $q_0 = b/d$.

Minimizing $\sum \varepsilon^2$ yields a linear system of equations from which the values of the parameters can be easily solved:

$$\begin{bmatrix} \sum x^4 I_x^2 & \sum x^3 I_x^2 & \sum x^2 I_x^2 & -\sum x^2 I I_x & -\sum x^2 I_x \\ \sum x^3 I_x^2 & \sum x^2 I_x^2 & \sum x I_x^2 & -\sum x I I_x & -\sum x I_x \\ \sum x^2 I_x^2 & \sum x I_x^2 & \sum I_x^2 & -\sum I I_x & -\sum I_x \\ \sum x^2 I I_x & \sum x I I_x & \sum I I_x & -\sum I^2 & -\sum I \\ \sum x^2 I_x & \sum x I_x & \sum I_x & -\sum I & -\sum 1 \end{bmatrix}$$

$$\times \begin{bmatrix} (bc/d^2 - a/d)c/d \\ a/d - bc/d^2 \\ b/d \\ k^{\gamma} \\ 1 - \alpha k^{\gamma} \end{bmatrix} = -\begin{bmatrix} \sum x^2 I_x (I + I_t) \\ \sum x I_x (I + I_t) \\ \sum I_x (I + I_t) \\ \sum I (I + I_t) \\ \sum (I + I_t) \end{bmatrix}, \tag{14}$$

where $I(x,t) = f(q(x))$ at time t, $I_x(x,t) = (\mathrm{d}f/\mathrm{d}q)(\mathrm{d}q(x)/\mathrm{d}x)$, at time t, and $I_t(x,t)$ is the frame difference of adjacent frames. The physical interpretation of k is the gain, and that of α is the bias. These two constants amount to approximating g with an affine relation (e.g., the plot of g

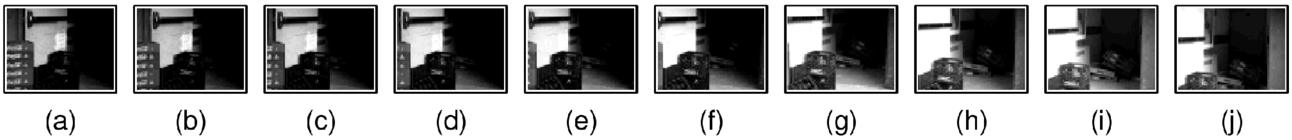


Figure 12. Ten images from the ‘cluttered and unsafe fire-exit’ investigative journalism sequence (taken using a covert eyeglass-based system in which the camera had automatic exposure). As the camera pans across to take in more of the open doorway, the image brightens up showing more of the interior, while, at the same time, clipping highlight detail.

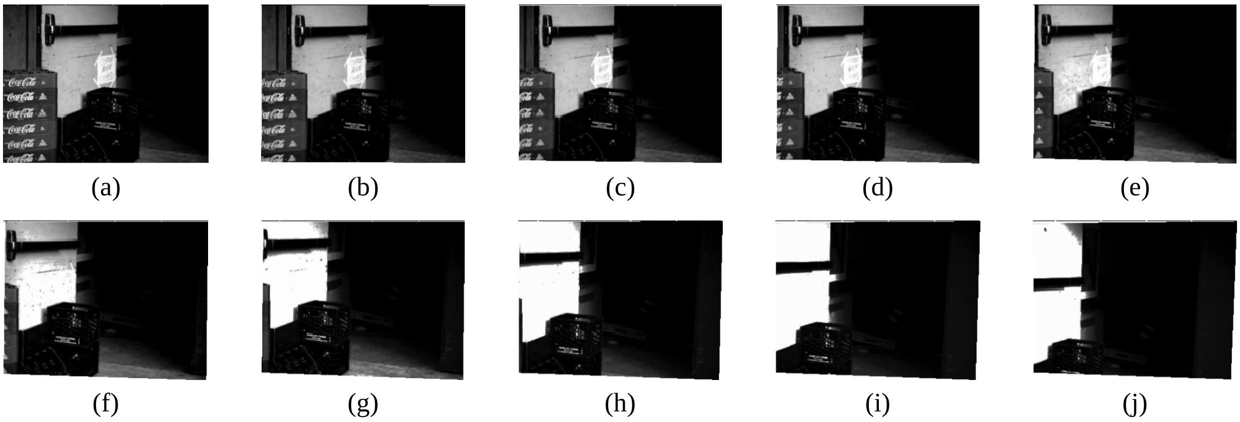


Figure 13. All images in the “cluttered and unsafe fire-exit” sequence expressed in the spatiotonal coordinates of the first image in the sequence. Note both the “keystoning”, or “chirping” of the images toward the end of the sequence, indicating the spatial coordinate transformation, as well as the darkening of the lighter images, indicating the tone scale adjustment, both of which make the images match (a). Prior to quantization for printing in this figure, the darker images (e.g., (i) and (j)) contained a tremendous deal of shadow detail, owing to the fact that the quantization step sizes are much smaller when compressed into the domain of image (a).

with a best-fit straight line), which is equivalent to approximating f with a power law (proposition 3.2).

It should be noted that the parameters of the projective coordinate transformation are determined indirectly, assuming that $d \neq 0$. The condition for which $d = 0$ corresponds to two successive pictures where the optical axis of the camera had turned through a right angle. That is, if one picture is taken, and then the camera is rotated so it is pointing in a direction 90 degrees from the direction in which it was pointing during the taking of the first picture.

Since it is very difficult to turn the head a full 90 degrees in the time between capture of two successive frames (1/60 second), especially given the tendency of the apparatus to make one feel dizzy and nauseated with such rapid changes in motion[3], it is sufficient to rule out the $d = 0$ possibility.

Another way in which the algorithm might fail is if the images are not close enough to being in the same orbit of the projective group of coordinate transformations. Thus an underlying assumption of the method is that the wearer can generate most of the image motion by turning his/her head much more quickly than the lesser motion produced by either scene motion or change of center of projection (e.g., that head turning is faster than appreciably moving the body from one location to another).

[3] Not to mention the fact that such images will likely not have any overlap unless an extremely wide-angle lens were being used, and even so, such images would not have enough overlap to provide a meaningful estimate of p.

Once the parameters of the projective coordinate transformation, as well as the parameters of the response curves relating the images are found, (e.g., once all the unknowns, $(bc-a)c$, $a-bc$, b, k^γ, and $1-\alpha k^\gamma$ are found) then the view into the environment map is rendered from the entire set of images which have overlapping scene content, weighted by their certainty functions, as follows:

$$\hat{q}(\mathbf{x}) = \frac{\sum_i c_i(\frac{ax+b}{cx+d})\frac{1}{k_i}f^{-1}(I_i(\frac{ax+b}{cx+d}))}{\sum_i c_i(\frac{ax+b}{cx+d})}. \qquad (15)$$

Then the desired rendered view, constructed from the estimate q is given by

$$\hat{I}_i(x,y) = f\left(k_i q\left(\frac{a^{-1}(dx-b)}{1-ca^{-1}x}\right)\right). \qquad (16)$$

An example of rendering $\hat{I}_0$ from each of the frames I_0 through I_9 of the “cluttered and unsafe fire-exit” sequence (original input data shown in figure 12) is illustrated by way of figure 13. In particular, the process of rendering $\hat{I}$ for any value of A, b, c, and k, may be explored interactively on a computer system, as illustrated in (figure 14). This process turns the personal imaging apparatus into a telematic camera in which viewers on the World Wide Web experience something similar to a QuickTime VR environment map [29], except with some new additional controls allowing viewers to move around in the environment map both spatially and tonally.

It should be emphasized that the environment map was generated by images obtained from a covert wearable appa-

(a) (b)

Figure 14. **VirtualCamera for Covert Telepresence:** Differently exposed pictures were generated by the natural process of looking around. (See figure 12.) These input images were spatiotonally registered and combined into an environment map of extremely high dynamic range which could then be rendered, and interactively explored in both its spatial domain and its tonal range. (b) Here we are able to see a white note left on a door upon which direct sunlight was incident, while at the same time we can see all the way down into a cluttered and unlit corridor to observe an unsafe fire exit. Needless to say, the management of this department store would not have responded favorably to a traditional photographic camera, yet the image captured here matches or exceeds the quality that would have been attainable by professional camera and lighting crews.

ratus, simply by looking around, and that no special tripod or the like was needed, nor was there significant conscious thought or effort required. In contrast to this proposed method of building environment maps, consider what must be done to build an environment map using QuickTime VR [29]:

> Despite more than twenty years photographic experience, Charbonneau needed to learn new approaches for this type of photography. First, a special tripod rig is required, as the camera must be completely level for all shots. A 35 mm camera. . . with a lens wider than 28 mm is best, and the camera should be set vertically instead of horizontally on the tripod. . . Exposure is another key element. Blending together later will be difficult unless identical exposure is used for all views.

The constraint of the QuickTime VR method and many other methods reported in the literature [30–32], that all pictures be taken with identical exposure, is undesirable for the following reasons:

- Imposing that all pictures be taken with the same exposure means that those images shot in bright areas of the scene will be grossly overexposed, while those shot in dark areas will be grossly underexposed. Normally the AGC would solve this problem and adjust the exposure as the camera pans around the scene, but since it must be shut off when using previously existing methods, shooting all the pictures at the same exposure will mean that most scenes will not record well. Thus special studio lighting is often required to carefully ensure that everything in the scene is equally illuminated.
- It does not benefit from the advantages of the Wyckoff principle of combining differently exposed images.

In contrast to the prior art, the proposed method allows natural scenes of extremely high dynamic range to be captured from a covert eyeglass-based reality mediator, by simply looking around. The natural AGC of the camera ensures that (1) the camera will adjust itself to correctly expose various areas of the scene, so that no matter how bright or dark (within a very large range) objects in the scene are, they will be properly represented in at least some of the input images, and (2) the natural ebb and flow of the gain, as it tends to fluctuate, will ensure that there is a great deal of overlapping scene content that is differently exposed, and thus the same quantities of light from each direction in space will be measured with a large variety of different exposures. In this way, it will not be necessary to deliberately shoot at different apertures in order to obtain the Wyckoff effect.

3.6. Automatic generation of photo albums: Collective photographic/videographic memory over the family–area–network

It is not necessary to press any button on the "lookpainting camera" because it can automatically determine when images are in the same orbit of the projective group of coordinate transformations. This is done by analysis of the error terms in (12), or, by comparing, for each pair of images in the sequence, $\sum_{\mathbf{x}}(\hat{I}_i - p_{ij}I_j)^2$ with a threshold to determine whether or not they are in the same orbit. The choice of thereshold is adaptive, as suggested by a Likelihood Ratio, Neyman–Pearson detection strategy, or CFAR (constant false alarm rate) detector in which the user can, for example, select a rate at which new candidate images are presented.

It is then assumed by the algorithm that, if more than some number (say, a hundred or so, or a number based on orbital distance) of images are captured within the same orbit, the subject matter is of sufficient interest to begin building an environment map. The building process stops as soon as the incoming images are no longer in the same orbit, and if enough new images arrive to form a second orbit, a second lookpainting is generated, and so on. All of the lookpaintings may be posted to a World Wide Web page, or disseminated over a smaller "family–area–network". In this way, for example, when one or more family members go on vacation, wearing the special sunglasses, the family's photo album is generated automatically without the need for any conscious thought or effort on the part of family members.

4. Collective connected *Humanistic Intelligence*

Lookpainting provides more than just personal photographic/videographic memory and real-time "Wearable Wireless Webcam" style telepresence. Because the system is a fully functional networked computer, it also facilitates a new form of interaction. For example, while shopping at the grocery store, a spouse may not only visit (through the affordances and capabilities of Wearable Wireless Webcam as illustrated in figure 14) but may also interact with the wearer, such as by pointing at objects with a shared cursor. For example, a remote spouse may point to the milk to indicate a preference, or select fruits and vegetables over the wireless link. In this way, the capabilities of these shared environment maps are quite similar to the shared transparent whiteboard spaces of computer-supported cooperative work [33,34], except that the proposed methodology overlays the shared "work" space on the real world and the apparatus is wearable, wireless, and covert. Furthermore, typical uses of the apparatus are characterized by the domain of ordinary living (e.g., shopping, banking, or walking home late at night), rather than just "working". Thus this new form of interaction is best described as *computer supported collaborative living* (CSCL), or simply *computer supported collaboration*.

5. Conclusions

In summary, *lookpainting* facilitates the use of a miniature covert personal imaging system to capture environment maps from which extremely high-resolution/high-definition pictures can be rendered. 'Lookpainting' also provides a "look around" user interface which is even more natural than the "point and click" user interface of modern cameras. Furthermore, 'lookpainting' affords the user total control of the process, makes the process of capturing an image more engaging and fulfilling, and results in environment maps that can be shared remotely with others who have access to the World Wide Web or a similar visual communications medium. Furthermore, lookpainting provides a new metaphor for *computer-supported collaborative living*.

Acknowledgements

The author wishes to thank Simon Haykin, Rosalind Picard, Steve Feiner, Don Norman, Jonathan Rose, Charles Wyckoff, Hiroshi Ishii, Chuck Carter, Kent Nickerson, and Chris Barnhart for much in the way of useful feedback, constructive criticism, and other matters related to this research and the general outlook on life it involves.

Additional thanks to Kodak, Digital Equipment Corporation, Compaq, ViA, Xybernaut, VirtualVision, HP labs, WaveRider, Chris Barnhart, Kopin, Colorlink, Ed Gritz, Miyota, Chuck Carter, Bob Kinney, and Thought Technologies Limited for lending or donating additional equipment that made these experiments possible.

References

[1] D. Norman, *Turn Signals Are the Facial Expressions of Automobiles* (Addison–Wesley, 1992).

[2] S. Mann, Wearable computing: A first step toward personal imaging, IEEE Computer, 30(2) (February 1997). http://hi.eecg.toronto.edu/ieeecomputer/index.html.

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

[4] S. Mann, Personal imaging, Ph.D. thesis, Massachusetts Institute of Technology (MIT), 1997.

[5] S. Mann, 'Mediated reality,' TR 260, M.I.T. Media Lab Perceptual Computing Section, Cambridge, Massachusetts, http://wearcam.org/mediated-reality/index.html (1994).

[6] S. Mann, Humanistic intelligence, in: *Proc. Ars Electronica* (September 8–13, 1997). http://wearcam.org/ars/; http//www.aec.at/fleshfactor.

[7] S. Mann, Shootingback, The winners of the prix ars electronica 1997 (September 8–13, 1997). http://www.rito.com/prix/winners.htm; See also: http://wearcomp.org/shootingback.html.

[8] G.M. Stratton, Vision without inversion of the retinal image, Psychological Review (1987).

[9] I. Kohler, *The Formation and Transformation of the Perceptual World*, Psychological issues, Vol. 3(4) Int. University press (1964).

[10] R. Azuma, Registration errors in augmented reality: NSF/ARPA Science and Technology Center for Computer Graphics and Scientific Visualization (1994). http://www.cs.unc.edu/~azuma/azuma_AR.html.

[11] S. Mann and R.W. Picard, Video orbits of the projective group; a simple approach to featureless estimation of parameters, TR 338, Massachusetts Institute of Technology, Cambridge, Massachusetts, http://n1nlf-1.eecg.toronto.edu/tip.ps.gz (1995). Also IEEE Trans. Image Proc. 6(9) (September 1997).

[12] S. Mann, 'Pencigraphy' with AGC: Joint parameter estimation in both domain and range of functions in same orbit of the projective-Wyckoff group, Technical Report 384, MIT Media Lab, Cambridge, Massachusetts (December 1994); http://hi.eecg.toronto.edu/icip96/index.html. Also in: *Proc. IEEE Int. Conf. on Image Process-*

ing (ICIP '96), Lausanne, Switzerland (September 16–19, 1996) pp. 193–196.
[13] S. Mann and R.W. Picard, Being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures, Technical Report 323, M.I.T. Media Lab Perceptual Computing Section, Boston, Massachusetts (1994). Also *IS&T's 48th Annual Conf.* (May 1995) pp. 422–428.
[14] S. Mann, Compositing multiple pictures of the same scene, in: *Proc. 46th Annual IS&T Conf.*, Cambridge, Massachusetts, The Society of Imaging Science and Technology (May 9–14, 1993).
[15] C.W. Wyckoff, An experimental extended response film, *S.P.I.E. Newsletter* (June–July 1962).
[16] C.W. Wyckoff, An experimental extended response film, Technical Report No. B-321, Edgerton, Germeshausen & Grier, Boston, Massachusetts (March 1961).
[17] E. Chen, QuickTime VR – an image-based approach to virtual environment navigation, *SIGGRAPH* (1995).
[18] M. Levoy and R. Hanrahan, Light field rendering, *SIGGRAPH* (1996) pp. 31–42.
[19] L. McMillan and G. Bishop, Plenoptic modeling, *SIGGRAPH* (1995).
[20] B.K.P. Horn, *Robot Vision*, MIT Electrical Engineering and Computer Science Series (McGraw–Hill, 1986).
[21] B.K.P. Horn and E.J. Weldon JR, Direct methods for recovering motion, Int. J. Computer Vision (1988).
[22] O.D. Faugeras and F. Lustman, Motion and structure from motion in a piecewise planar environment, Int. J. Pattern Recognition and Artificial Intelligence 2(3) (1988) 485–508.
[23] S. Laveau and O. Faugeras, 3–D scene representation as a collection of images, in: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Seattle, Washington (June 1994).
[24] M. Irani and S. Peleg, Improving resolution by image registration, CVGIP 53 (May 1991) 231–239.
[25] R. Szeliski, Video mosaics for virtual environments, Computer Graphics and Applications (March 1996) pp. 22–30.
[26] R. Szeliski and H.-Y. Shum, Creating full view panoramic image mosaics and environment maps, in: *Computer Graphics Proceedings*, Annual Conference Series (August 1997) pp. 251–258.
[27] A.N. Venetsanopoulos, Nonlinear mean filters in image processing, IEEE Trans. ASSP 34(3) (June 1986) 573–584.
[28] B. Horn and B. Schunk, Determining optical flow, Artificial Intelligence (1981).
[29] J. Campbell, QuickTime VR (October 18, 1996). http://www.webreference.com/content/qtvr/.
[30] R. Szeliski and J. Coughlan, Hierarchical spline-based image registration, CVPR (1994) pp. 194–201.
[31] H.S. Sawhney, Simplifying motion and structure analysis using planar parallax and image warping, ICPR 1 12th IAPR (October 1994).
[32] R. Kumar, P. Anandan and K. Hanna, Shape recovery from multiple views: a parallax based approach, in: *ARPA image understanding workshop* (November 1994).
[33] H. Ishii and M. Kobayashi, Clearboard: A seamless media for shared drawing and conversation with eye-contact, in: *Proc. Conf. on Human Factors in Computing Systems (CHI '92)*, ACM SIGCHI (May 3–7, 1992) pp. 525–532.
[34] H. Ishii, M. Kobayashi and J. Grudin, Integration of interpersonal space and shared workspace: Clearboard design and experiments, ACM Trans. Information Systems (October 1993) pp. 349–375.

Steve Mann, inventor of WearComp (wearable computer and personal imaging system), pictured here wearing an eyeglass-based prototype of his invention, is currently a faculty member at the University of Toronto, Department of Electrical and Computer Engineering. Mann has been experimenting with wearable computing and personal imaging as a hobby, since his high school days in the 1970s and early 1980s. More recently, in 1991, he brought his inventions and vision to the Massachusetts Institute of Technology, and established a new direction of research there. He received his Ph.D. degree from MIT in 1997. His previous degree is Master of Electrical Engineering (M.Eng.) from McMaster University. Mann also holds undergraduate degrees in physics (B.Sc.) and electrical engineering (B.Eng.). He was also guest-editor of a special issue on wearable computing and personal imaging in Personal Technologies journal, and one of four organizers of the ACM's first international workshop on wearable computing. He also proposed, to the IEEE, the first International Symposium on Wearable Computing (IEEE-ISWC-97), and was publications chair for that symposium. Mann has also given numerous keynotes and invited plenary lecturs, including, for example, the keynote address at the first International Conference on Wearable Computing in Fairfax VA, May 1998. His present research interests include quantigraphic imaging, formulation of the response of objects and scenes to arbitrary lighting, creating a self-linearizing camera calibration procedure, and wearable, tetherless computer-mediated reality – inventing the camera of the future. Mann has also exhibited his photoquantigraphic image composites and lightspace renderings in numerous art galleries from 1985 to present. See http://www.wearcomp.org and http://www.wearcam.org.
E-mail: mann@eecg.toronto.edu