Toposculpting: Computational Lightpainting and Wearable Computational Photography for Abakographic User Interfaces

Steve Mann Ryan Janzen Tao Ai Seyed Nima Yasrebi Jad Kawwa Mir Adnan Ali
hydraulophone@gmail.com rejanzen@gmail.com taoai.ece@gmail.com nima.yasrebi@mail.utoronto.ca jad.kawwa@mail.utoronto.ca maali@socialdynamics.ca

Department of Electrical and Computer Engineering
University of Toronto

Abstract—Toposculpting is the creation of virtual objects by moving real physical objects through space to extrude patterns like beams and pipes. In one example, a method of making pipe sculptures is proposed, using a ring-shaped object moved through space.

In particular, computational lightpainting is proposed as a new form of data entry, 3D object creation, or user-interface.

When combined with wearable computational photography, especially by way of a true 3D time-of-flight camera system, such as the Meta Spaceglasses (extramissive spatial imaging glass manufactured by Meta-View), real physical objects are manipulated during an actual or simulated long-exposure 3D photography process.

I. INTRODUCTION

A. Background: Digital Eye Glass, and Wearable Computing

Whereas AI (Artificial Intelligence) is an attempt at emulating human intelligence using computers [1], HI (Humanistic Intelligence) attempts to merge humans with computers so as to create intelligence from the human being in the feedback loop of a computational process [2].

A childhood experiment of author S. Mann, back in the 1970s involved the creation of DEG (Digital Eye Glass).

For more than 30 years, Mann lived everyday life in a computer-mediated world. The pictures below document various stages of this “Augmediated Reality” work that ran from the 1970s to the present:

![Steve Mann: Evolution of wearable computing in everyday life](image)

These early explorations in wearable computing were directed toward a vision aid, and a mechanism for seeing and understanding the world better, through “Augmediated Reality” (AR), as separate from the use of optical-only glass [3]. By “Augmediated”, we mean that visual reality can be Mediated, Deliberately-Diminished, or Augmented, in order to see and understand the world better.

Having learned to weld at age 4, Mann had envisioned a glass that would diminish/mediate the bright areas of an electric arc while augmenting the dark areas, thus “augmediating” the field-of-view.

Welding glass is well known [4][5][6], and apart from autodarkening helmets like 3M’s “Speedglas™” (everything goes dark when the arc starts), not much has changed in that world. Even with Speedglas (http://3m.com) “auto-darkening welding filter, a combination of liquid crystal elements, polarizers and electronics” (http://3m.com), once the arc is struck, the whole glass goes dark, and you can’t see much other than the tip of the electrode’s arc.

B. Theory of Glass

The Digital Eye Glass is a peripheral for the wearable computer, that reconfigures the eye automatically:

!["DIGITAL EYE GLASS"](image)

Rays of eyeward-bound light strike a “Lightspace Analysis Glass” (which need not necessarily be flat, and is therefore depicted as being curved), are converted to numbers which may then be processed by the wearable computer. The resulting numbers are fed to a “Lightspace Synthesis Glass” to be converted back into light. This allows the wearable computer to become a visual intermediary, to, for example, diminish the bright areas of the Subject Matter, and Augment/Mediate the bright areas, before resynthesizing the rays of light into the Eye, as shown in the above figure.

This “Theory of Glass” was originally motivated by the development of a seeing aid for welding. Various computerized welding helmets were constructed as a commercial enterprise, as a small Mann family business, producing a product known as the MannVis Glass™, MannGlas™, Digital Eye...
Glass™, GlassEyes™, or just Glass™, as an alternative to 3M’s Speedglas™ product introduced in 1981, and still the “gold standard” in the welding industry.

C. Generation-1 Glass

Mann built a rough approximation to the idealized Digital Eye Glass in 1978, using a television camera as the “Lightspace Analysis Glass” a miniature glass cathode-ray tube as the “Lightspace Synthesis Glass” (over the right eye), and various electric circuits as the wearable computer.

Because the camera was located beside the eye the long-term effects after many hours of wearing the apparatus consisted of an adaptation to this strange way of seeing, and the adaptation persisted after removal of the apparatus.

George Stratton’s work in 1896 with the upside-down eyewear (done optically rather than electrically), reported on adaptation effects to long-term wearing of “reality modifying” eyewear.

Mann observed that mappings that deviate moderately from what the unaided eye would see, were harder to “forget” than mappings that were either closer to or further from what the unaided eye saw (e.g. Stratton’s work). Thus theory and practice suggests one needs to either get the mapping perfect, or make it quite different from normal reality (e.g. present the image upside-down, or backwards, if one can’t get close to reality) [7].

D. Generation-2 Glass

An eyeglass with a camera and display integrated into it, is referred to as a Generation-1 Glass.

Due to many of the long-term adverse effects encountered with Generation-1 Glass, Mann proceeded to invent Generation-2 Glass, which causes the eye itself to, in effect, become both the camera and the display:

\[
\text{GENERATION-2 GLASS}
\]

Rays of eyeward-bound light are diverted into a Camera system that feeds to the Wearable Computer which then drives an Aremac (Lightspace Synthesis Glass). In this way rays of light that would have entered the unaided eye are collinear with rays of light presented by the Glass.

E. Generation-3 Glass

Observing that the focus distance at which objects appear was a problem, Generation-3 Glass was created, which includes a focus control mechanism so that if one looks through the glass the eye will focus to the same point as the unaided eye would have (i.e. in the absence of the glass).

\[
\text{GENERATION-3 GLASS}
\]

F. Generation-4 Glass

Finally, while looking at objects in various focal planes, such as looking at a distant object through a nearby screen or chainlink fence, some problems remained.

For this reason Generation-4 Glass was created using a laser light source with a spatial light modulator, and the like, to attain infinite depth-of-focus:

\[
\text{GENERATION-4 GLASS}
\]

Generations 2 to 4 of the Digital Eye Glass were known as “EyeTap Digital Eye Glass” [8].

The result was a natural experience with zero eyestrain which could be worn continuously for many hours a day for many years. The Generation-4 Glass was completed in 1999, described in detail in 2001 [9], and featured as the lead article (the cover, first page, and pages 31-32) of Abilities Magazine, Issue 69, Winter 2006:

Leftmost: Generation-4 Glass completed in 1999. The eye itself
is the camera exactly! That is why the “Digital Eye Glass” is also known as the “GlassEye” (The eye looks exactly like a camera lens). This eliminates long-term dizziness, vertigo, and flashback effects that can otherwise persist long after the Glass is removed. Rightmost: Google Glass, 2012. The camera being to one side of the eye makes this a Generation-1 Glass. Long-term neurological issues, dizziness, vertigo, and flashback effects, etc., can result from effectively moving the eye out of its eye socket and putting it off to one side.

Commercially made products such as Google’s Glass bear a similarity to this EyeTap Glass (same slender aluminum strip, same two silicone-pads on the nose, similar glass over the right eye), but Google’s Glass is a Generation-1 Glass (camera is to the right of the eye, not the eye itself as in Generations 2-4 of the Glass).

A number of companies are making and selling Generation-1 Glass (glass that does not align the camera in a natural eye position, and therefore does not meet the important EyeTap criteria [9]):

- **space:** spatial (collinearity) and spatial frequency alignment (orthoscopic and orthofocal);
- **time:** orthotemporality [9] and temporal-frequency (non-lagging, etc.);
- **light:** comparametric/superposimetic/orthoquantigraphic [2] [9].

Such glass could have long-term adverse effects [7].

**G. Spaceglass**

The Meta Spaceglass product is a Generation-5 Glass. Generation-5 is defined as being bi-directional in its light sensing and light output. In this way, an active vision system is implemented, which also provides a true 3D time-of-flight sensing capacity.

True 3D images are captured. This allows rendering of images at the Point-of-Eye (i.e. as the camera images can be re-rendered for point of left eye and point of right eye).

**H. Self-Gesture-Sensing**

Spaceglasses may be used to make objects using “in-the-air” gestures as illustrated in Fig. 1. Moreover, Spaceglasses can be used collaboratively, using wireless communications, antenna arrays, and the like [10][11][12] which signals can themselves be visualized. For example, the radio waves themselves can be made visible, adding new dimensions to wireless communications.

**II. TOPOSCLuptING AND ABAKOGRAPHY (“פָּאַבָא”)**

Making hand gestures in mid-air has certain drawbacks regarding a lack of tactile feedback, and a lack of engagement with real physical objects.

We now introduce two new concepts:

1) a form of sculpting that uses real physical objects such as arrays of light sources. We name this sculpting “topological sculpting” or “toposculpting” from the Greek word “topos” (“place”, e.g. spatial relations and geometric constancy under continuous scale and shape variation); and

2) a form of 3D computational lightpainting in which time exposures can be captured, and manipulated as graspable 3D objects in a computer-mediated reality. We name this process “abakography” from the Greek word “αβαξ” (“abax”) meaning a surface covered in sand or dust for doing calculations or drawing mathematical figures, itself derived from the Hebrew word פָּאַבָא (“dust”), i.e. the root word from which the modern word “abacus” derives. In this sense, abakography is often informally referred to as “dusting”.

A camera may be regarded as a time-integrating device. It integrates light, over time, at each pixel, or each point on a film plane, to form an exposure. In this sense, a long exposure shows the time-integral of the energy of the light received, at each point or pixel, as illustrated in Fig. 2 (leftmost). Here a wand or pen with a green LED (light emitting diode) on its tip has been waved through a space, to “write” or “draw” out a simple pattern.

When captured with a 3D camera, we can use such a light source to draw out a sculptural form that can then be manipulated in space, where other objects, can be rendered with it. For example, a bead can be rendered in 3D space on the path that the light took, and the bead can be slid along the path, as if it were the wire of an abacus. This “bead on a wire” metaphor leads us to consider the long exposure photograph an abakagram, and its mathematical form in space as an abakograph, the latter shown in Fig. 2 (rightmost).

A plurality of light sources moved together can be made to operate like a dot-matrix printer, i.e. to create text, graphics, and other patterns in three-dimensional space. The experimen-
III. 3D/Visible Image Alignment

In this section, we define the calculation used to overlay each image $I$ obtained from the visible-light 3-channel 2D camera onto the single-channel depth image $D$ obtained from the infrared time-of-flight camera. The basic problem is that the spatial coordinates of the raw image data returned in $I$ and $D$ are not aligned in correspondence for our objects of interest, namely, light-emitting object within the user’s field-of-view.

The reasons for this mis-registration are firstly that the cameras are mounted apart from one another, giving each a different point-of-observation, and secondly the lens systems of the two kinds of cameras have differing fields-of-view. Since the region of interest is far enough away from the cameras to neglect parallax error in the object, we aligned the two images (RGB and depth) only by shifting and scaling, proportional to the range (distance) of the object from the depth camera.

To find this transform, we construct binary images based on the RGB and depth data, using threshold values for separating the lights from the background. Then we compare the spatially weighted means and standard deviations of these binary images, to find the shifting and scaling factors, respectively. After this, by overlaying the depth information on top of the 2D image, we can construct the 3D abakogram and abakograph.

Since we use a binary threshold function to isolate light sources and foreground depth, let us define a general threshold function

$$T_C(v) = \begin{cases} 0, & \text{if } v < c \\ 1, & \text{if } v \geq c \end{cases}$$

Now let $I^{(b)} = T_{200}(I)$ and $D^{(b)} = 1 - T_{800}(D(B))$ represent thresholded binary images for the RGB and depth image, respectively. The threshold is reversed for the depth case (subtracting from 1) since we are interested in the foreground of the scene. The threshold constants (200 and 800) were simply chosen to suit the signal ranges observed in our current setup. Each pixel of any image $X$ may be indexed along rows by $m$, and columns by $n$. Let us define a normalized weighting $w$ by $X$ as

$$w_{m,n}(X) = \frac{X(m,n)}{\sum_{m,n} X(m,n)}$$ (1)

Then let $\mathbf{x} = (\mathbf{x}_m, \mathbf{x}_n)$ be the spatial weighted sample mean of binary image $X$, for example in $m$ we have

$$\mathbf{x}_m = \sum_{m} \sum_{n} w_{m,n}(X) \cdot m,$$

and similarly, we let $\mathbf{x}_n = (\mathbf{x}_m, \mathbf{x}_n)$ be the weighted sample standard deviation of image $X$. For example, again in $m$, we have

$$\sigma_m = \sqrt{\sum_{m} \sum_{n} (w_{m,n}(X) \cdot (m - \mathbf{x}_m))^2},$$

where we use $I^{(b)}$ and $D^{(b)}$ as $X$ to calculate the means $\mathbf{T}^{(b)} = (\mathbf{T}_m^{(b)}, \mathbf{T}_n^{(b)})$ and $\mathbf{D}^{(b)} = (\mathbf{D}_m^{(b)}, \mathbf{D}_n^{(b)})$, and the standard deviations $(\sigma_{I_m^{(b)}}, \sigma_{I_n^{(b)}})$ and $(\sigma_{D_m^{(b)}}, \sigma_{D_n^{(b)}})$. Now, the amount of shifting is

$$\mathbf{T}^{(b)} - \mathbf{D}^{(b)}.$$
and the ratios
\[ \frac{\sigma_{I_{mn}}}{\sigma_{D_{mn}}} \quad \text{and} \quad \frac{\sigma_{I_{mn}}}{\sigma_{D_{mn}}} \]
are used to register the scale of the RGB and depth images.

IV. VIRTUAL LIGHT

A. Computationally generating a virtual grid of light

The information provided by the previous step gives depth information for each pixel in each frame of input video. Knowing the real-world location of the light sources allows us to then in a 2 step process of integration in time, and integration in space, construct the abakograph and abakogram, as shown in Figure 4.

Each motion of the toposculptor (user of the toposculpting system) may create an element that is in itself used in later creations, as in Figure 5. Organic variations on a theme that can be assembled at will to create a superstructure. For example, architecture often contains organic forms that are repetitive in nature, mimicking the structure of living plants and creatures, as in Figure 8.

B. 3D rendering of virtual light traces

The algorithm generates the virtual integrated light in 3D space, along a trajectory interpolating between the detected real-life points in light detected by the 3D camera.

The interpolation process creates a trajectory of discrete artificial light sources, to replicate the discreteness of the points sampled in time. That is, while the “time” line segments are discretized by being in different frames of the video, correspondingly we discretize the virtual light sources in “space”.

These virtual light sources were created with maximum intensity at the centre, and decaying intensity radially outward, to replicate the properties of the real light sources dynamically (as will be described in the following section).

The segments of virtual light connect between the real-life light sources, in sequence. The light-point connection sequence is computationally determined, rather than connecting every point to every other point in the source image. (i.e. making a closed loop of virtual light rather than a star shape).

C. Dynamic virtual integrated light sources to match the computationally-determined dimensions of the real-world lights

In order to create the virtual connecting grid lines, we dynamically compensate for changes in the apparent diameter of the moving lights, as they move closer and farther from the camera, in order to cause the virtual connecting grid lines to match the integrated lines.

We compensate for ellipse eccentricity in light motion: Since a moving light produces a stretched-out smear when integrated over time, we compute the eccentricity of each detected region of light in order to dynamically estimate the true diameter of the light source (in a stationary reference frame).

The algorithm then draws the virtual connecting grid lines by integrating a sequence of blurred virtual light sources that replicate the frame-by-frame captures of the estimated real-life light sources.
Figure 6. (a) Pipe sculpture generated from a ring light moved through space. (b,c) Meshed pipe sculptures (note space-rings along time-lines, which together form a spacetime mesh), showing manipulation of 3D pipe sculpture length, following the hands (e.g. pulling the pipe out longer or shorter).

Figure 7. Pipe sculptures by author S. Mann take the form of interactive public art installations constructed from TIG-welded Type 316 stainless steel pipes, sometimes combined with pipes made from other materials (brass, bronze, etc.). Elements of these sculptures are designed and prototyped as computer models using toposculpting (top row), and then installed as interactive sound sculptures that make music from water flowing through variously curved pipes. The 2nd to last picture is courtesy of CNIB (Canadian National Institute for the Blind) Compass 2008, and depicts Mann’s rehabilitation sculpture used for rehabilitation of blind and deafblind children (finger tactility, touch, and haptics).
Figure 9. Gantries are typically made of straight I-beams. Here the outdoor part of the gantry is not well-connected to the indoor part. We will now show how we can use Tangible Lightpainting to design and architecturally previsualize a better gantry.

Figure 11. Architectural visualization and structural engineering design using abakographic user-interfaces. Here a labyrinth of gantries is visualized on the ceiling of a large research lab facility.

V. FURTHER APPLICATIONS OF 3D TANGIBLE LIGHTPAINTING

As another application, we wanted to put a gantry in our lab. Most gantries are made of straight I-beams as shown in Figure 9, but we wanted to visualize a curved gantry running around the perimeter of our lab.

To design the curved gantry’s I-beam, we used the VoxBox planar array of LEDs to display the letter “I”, and then we “painted” out an I-beam along the floor of our lab. We took several tries at this to get the approximate shape we wanted. Some examples of this “extrusion” of the letter “I” are shown in the leftmost 4 images of Figure 10. Our system with the IMU attached to the VoxBox™ LED array allows to capture these lightpaintings into the Unity 3D environment, where they can be manipulated in the Spaceglass wearable computer, using hand gestures. A Unity 3D interactive and graspable lightpainting is shown in the rightmost 2 images of Figure 10. Subsequently we visualized and explored more intricate labyrinths of ceiling-mounted gantries 11.

A. 3D Abakographs in Unity 3D environment

While most of our computational abakograms were mathematically rendered in MATLAB, we also generated 3D renderings in the Unity 3D environment: See Fig 13. A key challenge was to spatially transform the depth image from our SoftKinetic DepthSense DS325 camera, to match the same video orbit [13] as the visible light image. We accomplished this by measuring and aligning features of the light array, while remaining robust against interfering reflections from the bright visible light sources.

An Inertial Measurement Unit (IMU) was added, whereby roll, pitch and yaw acceleration were used to drive the directionality of light pipe extension. The performance of this user-interface was able to reach 30fps running on an Intel i7 processor-based Mac Pro, in the Unity3D environment.

B. Abakographic user-interfaces

Wayfinding, by physically walking through an abakograph is one type of augmented-reality interface modality. By seeing a 3D abakograph in augmented-reality glasses [14], [15], [16], [17], [18], [19], [20] it is possible to either “draw” an abakograph tunnel or corridor, or simply experience a predefined tunnel by walking through it. We used comparometric image processing [21], [22], [9], [23], [24], to be able to quantimetrically combine the abakograph with features of the human subject and external scene.

We built a handheld user-interface, with a digitally addressable light array. It could be set to modulate its colour in both time and space, which translated to light changes in two spatial dimensions in the abakograph. See Figs. 14 and 15. We also applied this to a hand-eye coordination exercise (Fig. 12).

C. Visualizing Electromagnetic Waves using Abakography

We built a system to visualize radio waves as they propagate through space, using abakography to generate a spatial map of wave intensity. A microwave RADAR device was built with two outputs at 90-degree phase-shifted demodulations to baseband (“direct” and “quadrature” outputs, or “real” and “imaginary” if the signal is represented by complex numbers).
Figure 10. **Generating a Unity 3D model using lightpainting:** We used an LED array to display the letter “I” while dragged it along the floor to trace out several curved I-beams. The resulting abakographs are automatically meshed and imported into Unity 3D as manipulable and editable objects. Selecting the one we wanted, we then visualized it on the ceiling using a Meta-View Spaceglass, so we could understand this architectural design by walking through the space to see how it would look in the space, as well as run realtime load-bearing simulations as we moved it through the space.

Figure 16. **Leftmost:** 3D computational lightpainting using a light ring to make a Tangible Media Conduit. **Center:** Unity 3D model automatically generated from the lightpainting appears in the Digital Eye Glass of the user. **Rightmost:** Unity 3D model can be manipulated using hand gestures or Ring gestures. Here the user can rotate the Tangible Media Conduit and move it around in 3D space. The object can be shared among multiple users, each one seeing it situated in space where the others have left it. A smoothing function has been applied to remove some of the wand shake in the original lightpainting. The result is a smooth “pipe” object that can be used for visualization, simulation, or manufacturing, such as by traditional sculptural means or by 3D printing.

Figure 14. **Experimental setup:** A discrete array of physical light sources, before their temporal/spatial extents are processed computationally. To generate a virtual corridor or meshed pipe, the light array is moved along its normal vector direction, plus continuous deviations in direction. An IMU (Inertial Measurement Unit) is located at center of crosshairs, and the frame (including the crosshairs) has a unique red color signature that is visible to the 3D infrared camera as well as the RGB camera. These combined sensory modalities track the exact position and orientation of the ring of light at all times.

Figure 15. **Abakographs provide guidance and warning for wayfinding and obstacle avoidance, as well as gaming and fitness training:** For example players attempt to run through a curved abakographic tunnel, without touching any of the 3D abakographic contours. When a player fails (e.g. by touching the contours), an audible/tactile proximity warning is given, and points are deducted from that player’s total score.

We mapped these baseband outputs to an addressable LED strip, with a bidirectional bar graph display in real-time. We also attached a radio-reflective surface to the bar graph, so the graph formed an abakograph with each point representing the wave reflection at that point in space.

Since frequency \( f \) of the RADAR output is proportional to velocity \( v \) of the moving target (reflector attached to the bar...
(2)

That is, at any point in time, the observed phase $\phi$ of the waveform (displayed on the bar graph) changes in proportion to displacement $x$ of the bar graph. As a result, we can physically move the bar graph scanning through a waveform, viewed as a virtual standing wave, no matter if we move the bar graph quickly or slowly. Moving the bar graph quickly merely “plays” the radio wave faster as well, such that the radio wave always matches the displayed data.

We thus created abakographs to let us “see” radio waves as if they are standing still, thus “stopping time” so as to make radio waves tangible, visible, and “graspable”. See Fig 17.

We chose a green colour to represent the direct signal (i.e. the “real part” of the radar return), and magenta (red+blue) to represent the quadrature signal (i.e. the “imaginary part”). These being complimentary colours allowed us to visualize the two out-of-phase waveforms in space, resulting in white where overlapping at equal strength.

D. Veillance in the Vironment

Abakography bridges the gap between our environment (that which is around us but not including our clothes, skin, and body) and our exvironment (ourselves), thus creating a computing “vironment”.

 Whereas surveillance (watching from above) is typically done by cameras in the environment around the individual under surveillance, sousveillance is done by cameras in the exvironment. As a vironmental interplay between these two vironments, the individual may attempt to create an alibi by way of a NEWStamp$^\text{TM}$ in which a news item is displayed (e.g. abakographically), in a video sequence (to be pieced together later for defense in court trials, for example) or a still image. Subsequently the result is timestamped as well, so that the abakograph forms an alibi that can be used to refute allegations that veillance (be it surveillance or sousveillance) has been temporally falsified. See Fig 18.

VI. Conclusions

We have proposed novel methods for generating volumetric 3D sculptures, that we have coined as “Toposculpting” (topological sculpting). In some embodiments, depth cameras are used to provide a means to capture any movable object to sweep out shapes and forms that can be used as building blocks, pipes, or other patterns that are manipulated directly by a human operator, using an augmented reality system.

Moreover, 3D computational lightpainting has been introduced as both an artistic sculpting medium, as well as for user interface design, data visualization, and the like, for example, allowing surgical visualization, architectural visualization, and allowing people to see and touch and manipulate radio waves and light itself.
This “Abakography” bridges the gap between our environment (that which is around us but not including our clothes, skin, and body) and our exvironment (ourselves), thus creating a computing “vironment”.

ACKNOWLEDGMENT

Authors, including author S. Mann, as Chief Scientist of Meta-View, would like to acknowledge the assistance of Meta-View (http://spaceglasses.com).

REFERENCES


[8] Peter Nowak. The world’s first 'cyborg,' Steve Mann, says always being connected to others can make people see the world in a different — and better — light. CBC News, Monday Dec. 22, 2003.


