

INTERACTIVE ABSTRACT MIRRORS ON MOBILE DEVICES

Javier Villegas and Angus Forbes

Creative Coding Lab
School of Information: Science, Technology, and Arts
University of Arizona

ABSTRACT

This paper presents analysis/synthesis strategies for generating abstract, creative representations via the camera input on a mobile device. Mobile devices are well suited for interactive video processing since they are simultaneously capable of image capture, display, and manipulation. Analysis/synthesis methods are particularly powerful in interactive arts projects as they enable even drastic manipulations of the input image while still maintaining fundamental aspects of its original identity. Moreover, by using abstract synthesis elements (i.e., coherent elements larger than single pixels), we are able to directly interact with the image and to manipulate its final output. We describe some of the exciting capabilities of video processing and interaction on mobile devices and introduce a series of mobile applications that use analysis/synthesis techniques.

1. INTRODUCTION

Nearly all modern smart phones and tablets are capable of handling multimedia data, including video input and output. Moreover, most mobile devices include dedicated GPUs that are powerful enough to allow real-time manipulation of video. These devices are also equipped with various sensors that enable many interaction possibilities. This paper presents an investigation of analysis/synthesis (hereafter, A/S) techniques that take advantage of these these capabilities.

1.1. Analysis/synthesis techniques

A/S techniques provide a high level of manipulability in the input signal without altering its fundamental identity. This is the main reason why A/S techniques are extensively used in creative audio processing (e.g. Vocoder, Autotune). But few examples exist in the image or the video domain. Some examples of previous work using A/S approaches for creative image manipulation include the computer-assisted mosaics of Kenneth Knowlton [1, 2]. And similar approaches from other artists in the 60s, including Waldemar Cordeiro [3], Charles Csuri, and James Shaffer [4], and the Japanese Computer Technique Group (CGT), also make use of A/S

techniques [5]. More recent A/S related works include Robert Silvers' Photomosaics [6] and works of Golan Levin, such as "Segmentation and Symptom," [7] and "Floccular Portraits" [8]. Real-time installations using A/S approaches have been created by Daniel Rozin in his collection of mechanical mirrors [9]. Jim Campbell's "Low Resolution" series [10], and two of Rafael Lozano-Hemmer pieces in his "Shadow Box" series, "Eye Contact" [11] and "Third Person," [12], are other examples of this relatively unexplored technique.

1.2. Non-photorealistic rendering

An area of research closely related to our work in A/S techniques is called "non-photorealistic rendering" (or NPR), which aims to creatively interpret the raw data from realistic images and videos. NPR is primarily concerned with automatically recreating the look of different styles of hand-made paintings. In fact, most of the image-based NPR algorithms are A/S processes, in which an input image is used to calculate the position, color, orientation or texture of the synthesis elements [13]. In a seminal paper, Haeberli demonstrates different alternatives for abstract representation of natural and synthetic images [14]. He explores the use of different primitives, such as brush strokes, and successfully mixes together automatic and semi-automatic techniques. Stylization effects, including "mosaic," "pixelate," and "cubism," are now part of standard digital photo-manipulation tools. These effects can be interpreted as A/S processes (although they are usually not referred to as such).

One of the reasons why A/S methods for video have failed to be as ubiquitous as their audio equivalents involves the problem of *temporal coherence* [15]. Temporal coherence is lost when video frames are analyzed independently, producing a set of drawing primitives that jump arbitrarily from place to place as every new frame is processed. Different alternatives to guarantee temporal coherence on non-photorealistically rendered animations have been explored. For instance, Litwinowicz describes a technique to create image animations that seem "hand-painted" from live action video [16]. He uses the edge map of an input

image to constrain the length of the strokes, then applies an optical flow algorithm to ensure the coherence of the strokes between successive frames. Bénard, et al., has explored the use of dynamic textures and Gabor noise primitives to create time-coherent stylizations [17, 18]. In “Animosaics,” Smith, Liu and Klein explore rules that govern the smooth motion of mosaic tiles in animated mosaics [19]. They examine not only the coherence between individual elements, but also the cohesion in the movement of groupings of tiles. A detailed review of techniques and content of non-photorealistic animation can be found on the papers from Bénard, Bousseau and Thollot [15], and Agrawal [20].

1.3. Art in mobile devices

Many artists have taken advantage of the various possibilities of mobile devices and have recognized their potential to engage audiences in novel ways. This potential includes both the increasing computational power of current hardware and also the distribution model of the applications that can be easily downloaded to run on them. Well-known artists, such as Scott Snibbe [21] and Lisa Jevbratt [22], have chosen mobile devices as the target platforms or their latest creations, and digital art festivals are nowadays likely to include a specialized category for mobile devices. However, most of the art-oriented applications that involve video processing use simple pixel-based techniques, and moreover include only a limited amount of interaction. Much of the processing methods in these art projects involve basic color space conversions, linear filters, or geometrical transformations. Recent work by the Creative Coding Lab at the University of Arizona explores various methods to creatively manipulate video, mostly based on A/S approaches, and implements these approaches on mobile devices. In our work, each video frame is analyzed and then recreated after incorporating the creative manipulations generated by user input. For programming of the mobile devices we utilize the *Aluminum* framework, a powerful multimedia codebase developed at the Creative Coding Lab that gives developers access to current, low-level features of OpenGL and OpenGL ES [23]. Specifically, the framework provides access to the GPU capabilities of modern iOS mobile devices while simplifying most of the objective-C configuration code that iOS devices require.

The power of A/S strategies is based on their ability to distort the input image parameters in various ways before reconstructing an output. In the next section, we illustrate some of the parameters that can potentially be manipulated before this reconstruction.

2. ABSTRACT MIRRORS

Humans give meaning to images by grouping the existing elements and associating them with previous knowledge. There are many different ways in which picture elements can be grouped [24], and these different grouping principles can be exploited to create *abstract mirrors*, or non-photorealistic representations of a raw video capture.

2.1. Regions

A *region* can be defined as a group of connected pixels that share a similar color, luminance, or texture. Figure 1 shows different ways of re-drawing regions that were detected in an analysis stage. Manipulations can then occur before the reconstruction of the image.

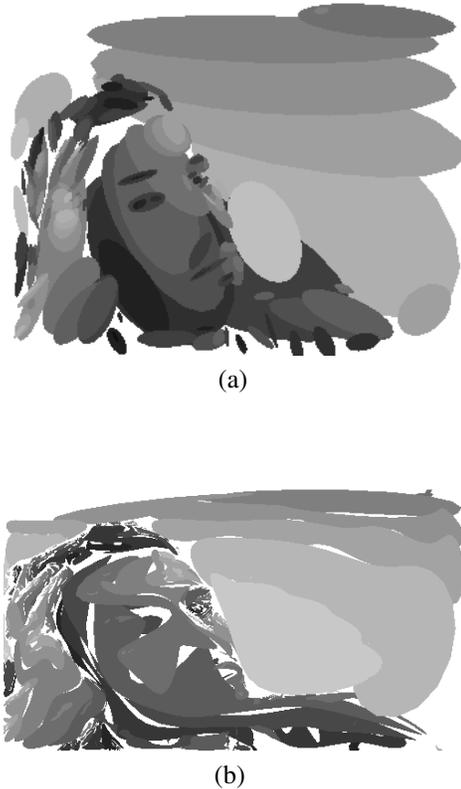


Fig. 1. Different ways to present identified regions: (a) replacing regions with ellipses; (b) re-drawing the regions after modifying their Fourier descriptor.

2.2. Edges and lines

The identification of contours is fundamental to object perception [24]. A complete new family of examples can be

created if, in the analysis stage, parametric curves are fitted to the edge of the map on the input image (Figure 2).

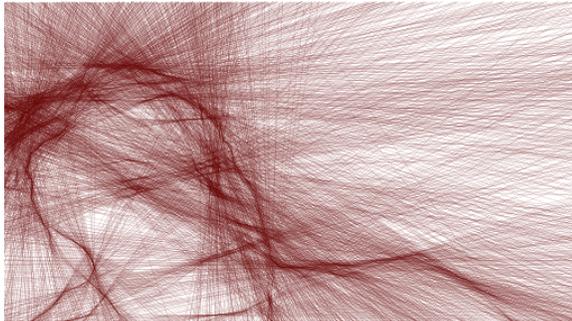


Fig. 2. Edges of the image recreated with straight lines.

2.3. Orientation

The human visual system can also separate regions based on orientation. A direct mapping from gray value to line orientation produces the results shown in Figure 3.

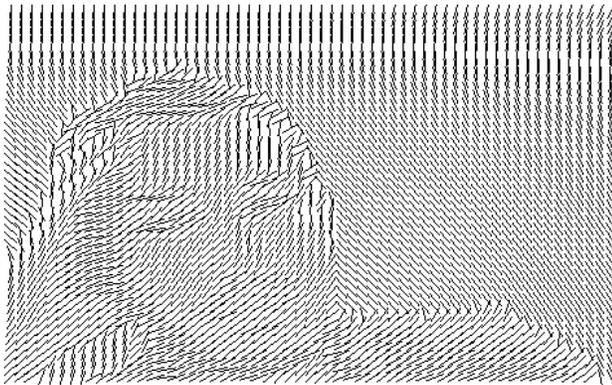


Fig. 3. Gray-levels are mapped to line orientation.

2.4. Size

Objects can also be grouped into different regions by size. Areas represented with larger objects will be perceived as darker than areas represented with smaller objects (assuming dark objects on white background). Figure 4 shows a visual experiment using size as a mapping parameter. A gray-scaled image is used as input and also as the synthesis object.

2.5. Density

We perceptually separate regions that have similar element density [24]. Also, a region can be perceived as darker than others if there is more object density in that region. This

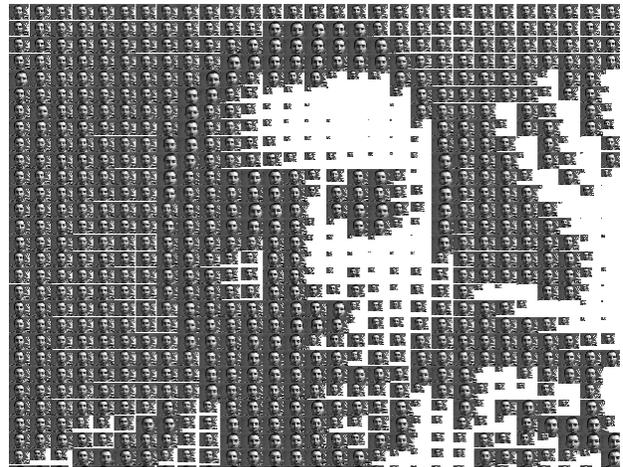


Fig. 4. The gray-scale levels are mapped to the size of the synthesis element.

is the basic concept behind many dithering techniques [13]. In Figure 5 the dot density is changed to recreate the gray levels of the input.

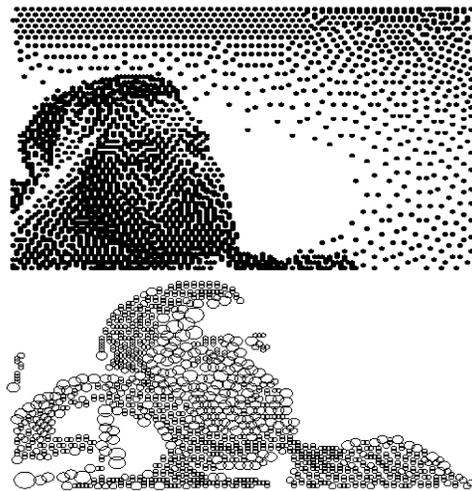


Fig. 5. Image represented by manipulating the density of objects.

2.6. Connectedness

Figure 6 shows different ways of using the output of a dithering algorithm to join the resultant black points with lines following different paths. As stated by Ware [24], connectedness is a strong grouping principle. Small differences in the way that points are connected can drastically change the look of the result.

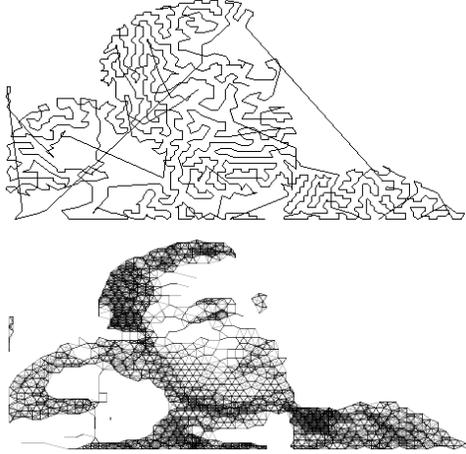


Fig. 6. Two images showing different ways to connect dots.

3. MOBILE IMPLEMENTATIONS

In this section we show how we use some of the grouping principles described above to implement interactive abstract mirrors that work in real-time on current mobile devices. Specifically, we present three different applications that we have implemented using the *Aluminum* framework [23].

3.1. The *Fluid Automata* application

Fluid Automata is an interactive fluid simulation and vector field visualization application based on cellular automata systems [25]. A flow of energy is distributed through the image using local rules that are inspired by the physics of fluids (although the fluids are inherently non-realistic due to the fact that there is no mass conservation condition and because the fluid is compressible). The system is illustrated on Figure 7. Different alternatives for interaction with mobile devices have been explored. Energy can be added to the system by tapping the screen or moving a finger across the screen to change the direction of vectors in the fluid system. Figure 8 shows one of the possible outputs of the system. *Fluid Automata* has been presented as an independent artistic piece [26] and also as a visual accompaniment for dynamic music compositions [27].

3.2. The *Angle Shift* application

Figure 3 illustrates how orientation can be used as a discriminator for regions. *Angle Shift* is an interactive application that combines orientation, color and motion. The synthesis elements are thin lines that change position, direction, and color according to the current camera input. The new state of each line is updated from the previous one in the CPU using the camera image and the interaction information (touchscreen) as inputs. The updated information is

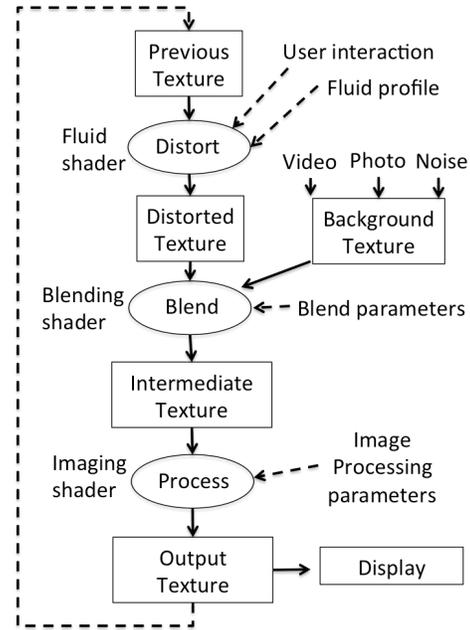


Fig. 7. The Fluid Automata system.



Fig. 8. Running the application.

then sent to the GPU for drawing via modern OpenGL demands. Figure 9 shows a block diagram of the system and Figure 10 show an example of a typical output of the application.

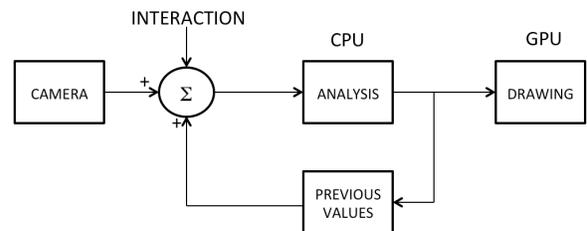


Fig. 9. The Angle Shift System.



Fig. 10. The *Angle Shift* application running on an iPad.

3.3. The *Meshflow* application

In the *Meshflow* application a set of points evolves slowly before finally morphing into the image of the camera input. The motion is continuous and the position of each point directly evolves from its position in the previous frame [28]. *Meshflow* exploits some common expectations and characteristics of the human visual system. On one hand, the impression of having different grayscale values can be created using only black dots on a white background simply by altering the density of dots in a particular area. This halftoning technique is common in print publications [29]. When we are observing a grid, our expectation of regularity creates an impression of dimensionality, as shown for instance in Figure 11. This characteristic has been used for centuries

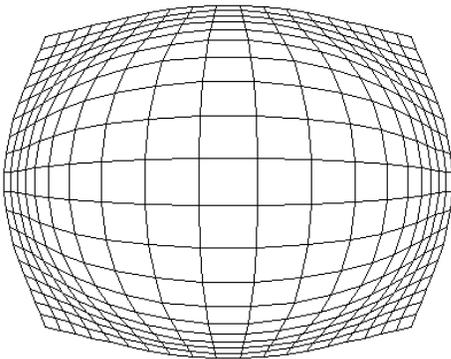


Fig. 11. A distorted grid gives the impression of dimensionality.

to represent three-dimensional objects in the plane. The artist Marius Watz has created a series of grids that when perturbed in their regularity creates a strong feeling of three dimensionality [30]. In *Meshflow*, the nodes on the grid are attracted to the darkest areas of the image, but the grid structure is kept. The motion of the nodes is constrained by physics laws that describe an attraction between neighbor nodes and a drag force. Figure 12 shows the block diagram

of the application.

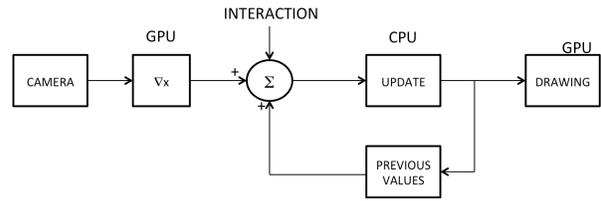


Fig. 12. Block diagram of the *Meshflow* application

The gradient of the image is calculated using the Sobel operator in the GPU. This information is then encoded into two 8-bit channels and then read back into the CPU. In the CPU, the previous values of the nodes of the grid are updated using the gradient information and using the coordinates of the interaction in the touchscreen. The geometry information is then sent back again to the GPU for final rendering to the screen. Figure 13 shows the application running in an iPad.

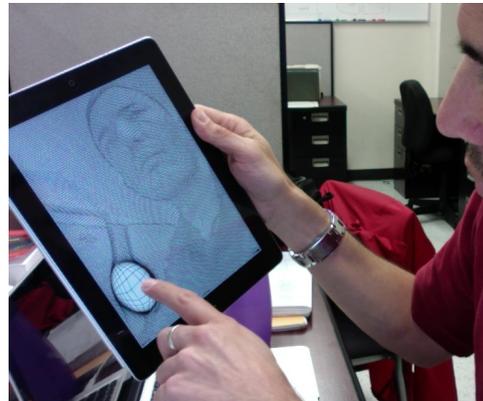


Fig. 13. The *Meshflow* application running on an iPad.

4. SUMMARY AND FUTURE WORK

We believe that A/S strategies are an effective means to generate creative mobile applications that function as abstract mirrors. Using A/S techniques, an input video frame is reduced to a set of parameters, then those parameters are manipulated, and finally an abstract representation is created with the modified parameters. In each of the examples presented here, the input information is not used to recreate the image from scratch, but instead as a way to update the previous state of the synthesis elements. This guarantees temporal coherence on the output image. Temporal coherence is a factor that has perhaps the strongest impact in the perceived quality of an abstract animation. We are exploring ways to measure the quality of temporal coherence, a

surprisingly difficult task. Some potential evaluations we are exploring include: the displacement distribution of the synthesis objects; the minimum jerk rule [31] for smoothness evaluation; and the spatial dispersion of the direction of motion. Part of our present research involves finding correlations between these quantitative measurements and the qualitative evaluations gathered via user studies.

In each of the three applications we presented, users interact via the touchscreen. Future versions may also include other sensors common to mobile devices (e.g., accelerometers, gyroscopes, audio inputs) to manipulate various parameters before the output synthesis.

Lastly, in future versions we would like to incorporate the concept of “cross-synthesis,” that is, mixing the parameters of two images mixed *before* the synthesis stage. One application of cross-synthesis is the creation of what we call “double meaning” images, where a single image has both a local and a global narrative occurring simultaneously. Future work will extend our previous research on double meaning images [32] and explore its application to mobile devices.

5. REFERENCES

- [1] Frank Dietrich, “Visual intelligence: The first decade of computer art (1965-1975),” *Leonardo*, vol. 19, no. 2, pp. pp. 159–169, 1986.
- [2] Knowlton, “Knowlton mosaics - portraits by computer assisted art pioneer ken knowlton,” <http://www.knowltonmosaics.com/>.
- [3] Annateresa Fabris, “Waldemar cordeiro: Computer art pioneer,” *Leonardo*, vol. 30, no. 1, pp. pp. 27–31, 1997.
- [4] Charles Csurí and James Shaffer, “Art, computers and mathematics,” in *Proceedings of the December 9-11, 1968, fall joint computer conference, part II*, New York, NY, USA, 1968, AFIPS ’68 (Fall, part II), pp. 1293–1298, ACM.
- [5] Jasia Reichardt, *Cybernetic serendipity: The computer and the arts*, Praeger, 1969.
- [6] Robert Silvers, “Robert silvers,” <http://www.photomosaic.com/>, 2003.
- [7] Golan Levin, “Segmentation and symptom - interactive art by golan levin and collaborators,” <http://www.flong.com/projects/zoo/>.
- [8] Golan Levin, “Floccular portraits - interactive art by golan levin and collaborators,” <http://www.flong.com/projects/floccugraph/>.
- [9] Jay David Bolter and Diane Gromala, *Windows and Mirrors: Interaction Design, Digital Art, and the Myth of Transparency*, The MIT Press, Oct. 2005.
- [10] Jim Campbell, “Jim campbell, low resolutions works,” http://www.jimcampbell.tv/portfolio/low_resolution_works/, 1999-2011.
- [11] Rafael Lozano-Hemmer, “Rafael Lozano-Hemmer - project ”Eye contact”,” http://www.lozano-hemmer.com/eye_contact.php.
- [12] Rafael Lozano-Hemmer, “Rafael Lozano-Hemmer - project ”Third person”,” http://www.lozano-hemmer.com/third_person.php.
- [13] Thomas Strothotte and Stefan Schlechtweg, *Non-Photorealistic Computer Graphics: Modeling, Rendering and Animation*, Morgan Kaufmann, 1 edition, June 15 2002, ISBN: 1558607870, pages 472.
- [14] Paul Haeberli, “Paint by numbers: abstract image representations,” *SIGGRAPH Comput. Graph.*, vol. 24, pp. 207–214, September 1990.
- [15] Pierre Bénard, Adrien Bousseau, and Joëlle Thollot, “State-of-the-art report on temporal coherence for stylized animations,” *Computer Graphics Forum*, vol. 30, no. 8, pp. 23672386, December 2011, DOI: 10.1111/j.1467-8659.2011.02075.x.
- [16] Peter Litwinowicz, “Processing images and video for an impressionist effect,” in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1997, SIGGRAPH ’97, pp. 407–414, ACM Press/Addison-Wesley Publishing Co.
- [17] Pierre Bénard, Adrien Bousseau, and Joëlle Thollot, “Dynamic solid textures for real-time coherent stylization,” in *ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games (I3D)*, Boston, MA, Etats-Unis, february 2009, ACM, pp. 121–127, ACM.
- [18] Pierre Bénard, Ares Lagae, Peter Vangorp, Sylvain Lefebvre, George Drettakis, and Joëlle Thollot, “A dynamic noise primitive for coherent stylization,” *Computer Graphics Forum (Proceedings of the Eurographics Symposium on Rendering 2010)*, vol. 29, no. 4, pp. 1497–1506, june 2010.
- [19] Kaleigh Smith, Yunjun Liu, and Allison Klein, “Animosaics,” in *Proceedings of the 2005 ACM SIGGRAPH/Eurographics symposium on Computer animation*, New York, NY, USA, 2005, SCA ’05, pp. 201–208, ACM.

- [20] Amit Agrawal, “Non-photorealistic rendering: Unleashing the artist’s imagination [graphically speaking],” *Computer Graphics and Applications, IEEE*, vol. 29, no. 4, pp. 81–85, july-aug. 2009.
- [21] Scott Snibbe, “Snibbe studio apps,” <http://www.snibbestudio.com/apps/>.
- [22] Lisa Jevbratt, “Zoomorph,” 2013, <http://zoomorph.net/>.
- [23] Angus Forbes, “Aluminum,” <https://github.com/CreativeCodingLab/aluminum>.
- [24] C. Ware, *Information visualization: perception for design*, vol. 22, Morgan Kaufmann, 2004.
- [25] Martin Gardner, “Mathematical games: The fantastic combinations of john conways new solitaire game life,” *Scientific American*, vol. 223, no. 4, pp. 120–123, 1970.
- [26] Angus Graeme Forbes, Tobias Höllerer, and George Legrady, “Generative fluid profiles for interactive media arts projects,” in *Proceedings of the Symposium on Computational Aesthetics*. ACM, 2013, pp. 37–43.
- [27] Angus Graeme Forbes and Kiyomitsu Odai, “Iterative synaesthetic composing with multimedia signals,” in *Proceedings of the International Computer Music Conference (ICMC)*, Ljubljana, Slovenia, September 2012, pp. 573–578.
- [28] Javier Villegas, “Meshflow: A grid warping mirror,” 2010, <https://vimeo.com/21653145>.
- [29] D.L. Lau and G.R. Arce, *Modern digital halftoning*, vol. 8, CRC, 2001.
- [30] Marius Watz, “Marius Watz - grid distortion,” <http://www.unlekker.net/proj/griddistortions/>, 2008.
- [31] P. Viviani and T. Flash, “Minimum-jerk, two-thirds power law, and isochrony: converging approaches to movement planning.,” *Journal of Experimental Psychology: Human Perception and Performance*, vol. 21, no. 1, pp. 32, 1995.
- [32] Javier Villegas and Angus Graeme Forbes, “Double-meaning: Interactive animations with simultaneous global and local narrative,” in *Proceedings of of re-new 2013 media art conference and festival*. Re new, 2013, pp. 300–304.