Gaze Relations: Looking as Power
Critical Analysis of Human Gaze and Computer Vision

A thesis paper submitted in partial satisfaction
of the requirements for the degree of

Masters of Fine Arts
in
Digital Arts and New Media

by
Anthony Assi
June 2018

This Thesis of Anthony Assi approved by:

Professor Edward Shanken, Chair

Professor Elliot Anderson

Professor Angus Forbes

Professor Warren Sack

Tyrus Miller
Vice Provost and Dean of Graduate Students
Copyright © by

Anthony Assi

2018
# TABLE OF CONTENTS

Table of Contents........................................................................................................... iii
List of Figures................................................................................................................... iv
Abstract.............................................................................................................................. v
Dedication.......................................................................................................................... vi
Acknowledgement............................................................................................................ vii

1- Introduction.................................................................................................................... 1
2- Project Description......................................................................................................... 3
3- Technical Description.................................................................................................... 7
4- Art Context..................................................................................................................... 15
5- Theory and Concept....................................................................................................... 26
6- Conclusion..................................................................................................................... 30

Works Cited....................................................................................................................... 33
LIST OF FIGURES

FIG. 1: Gaze Relations, 2018………………………………………………………………………………...3

FIG. 2: Fashion design and hairstyle inspired by Haar features………………………………………5

FIG. 3: OpenCV body detection…………………………………………………………………………8

FIG. 4: OpenCV visualisation tool………………………………………………………………………..10

FIG. 5: Frames from CriticalCV…………………………………………………………………………11

FIG. 6: Gaze Sampling……………………………………………………………………………………13

FIG. 7: Gaze trail…………………………………………………………………………………………14

FIG. 8: Adam Harvey, OpenCV Face Detection: Visualized, 2010…………………………………15

FIG. 9: Adam Harvey, CV Dazzle, 2010-2018…………………………………………………………16

FIG. 10: Shannon McMullen and Fabian Winkler, 20/X, 2015-2016…………………………..18

FIG. 11: Shannon McMullen and Fabian Winkler, 20/X, 2015-2016……………………………19

FIG. 12: Behnaz Farahi, Caress of the Gaze, 2015………………………………………………….21

FIG. 13: Happy Valley Band, Organvm Perceptvs, 2017……………………………………………23

FIG. 14: John Bates………………………………………………………………………………………25
Gaze Relations is an art installation that visualizes how human gaze and computer vision see bodies in visual media to compare how they operate perceptually and culturally. Gaze, in visual media, refers to the way the filmmaker and audience look at bodies and how the filmmaker behind the camera influences the way the photographic subject is represented to an audience. Body detection, on the other hand, is a computer vision technique employed by surveillance institutions to automate the process of detecting bodies in surveillance footage. Gaze Relations critically visualizes and materializes these abstract phenomena of gaze and surveillance through visual art that represent the way these modes of vision operate. Gaze-tracking software tracks how a person looks at an image and then custom-built software reconstructs the image through their gaze demonstrating how they visually engage with the photographic body. With respect to surveillance, critical software was developed to visually analyze the processes of how a popular body detection algorithm models a human body. In addition to video, costume designs represent how the algorithm sees the body projecting the machine vision back into the world it is looking at. The piece highlights the differences between the complexities and nuances of human gaze to the algorithmic processes of computer vision that reduce the human form to simple features. Gaze and computer vision will be discussed as powerful social and cultural force capable of influencing the external world, not just the phenomena of taking in light and processing it.
DEDICATION

To all the computer-readable people living in the 21st century.
ACKNOWLEDGEMENT

Thank you to Edward Shanken, Warren Sack, Angus Forbes, and Elliot Anderson for your mentorship and inspiration. Thank you to Kristin Erickson Galvin for your patience and encouragement. Thank you to John Weber for your skillful curation. Thank you to Digital Arts and New Media faculty and students for your feedback and criticism. Thank you to Dude Haley, Harmony Reynolds, Snelly Bushnell, Joy Knighton, and Lucia Flexer-Marshall for willing to be the objects of gaze for this weird experiment.
Introduction

*Gaze Relations* critically visualizes the relationship between bodies, gaze, and computer vision to compare and contrast how people and computer vision algorithms see images of the body. Gaze, within the context of media studies, refers to the way people see bodies from the perspective of the person behind the camera and the audience viewing the image. It examines how the gaze of the photographer influences the way the photographic subject is represented and how spectator views the subject in the image. Although body detection algorithms perceive the body in a very different way than people do with varying degrees of consent, both project their internal biases onto the external world whether it’s a male cinematographer sexualizing the way a female actor appears on screen or a computer vision algorithm affecting how people dress as to be seen or unseen. In this model vision is not a one way channel, but a two way channel of exchange where the gaze of the observer is sculpting the way the subject interacts with the camera. Historically, gaze and surveillance have been discussed almost exclusively in theoretical and abstract terms without digging into the material and technical details. The micro-details that make up these forms of vision will inform how they function within our culture.

*Gaze Relations* empirically compares and contrasts how these paradigms of human and computer vision operate. These critical visualizations provide tangible results in the form of video and costume that will inform the theoretical discourse around gaze and surveillance in a fresh new way. The relationships between bodies, gaze, and computer vision are shown through digital video that illustrates how the
body is seen and costumes that physically represent the way the body is modeled by body detection software. I designed custom software, CriticalCV (see Technical section), to visualize the operations of the computer vision algorithm. Visualizing the processes of gaze and computer vision demonstrates the complexities and nuances of human gaze in contrast to the algorithmic procedures of body detection that simplify the human form to a set of rigid patterns. This contrast explicitly illustrates the differences between the mechanics of human and computer vision, while implicitly comparing the way in which the photographer and the surveillance algorithm exert power over the body. Fashion will be framed as a tool that mediates the visual and power relationship between bodies, gaze, and machine vision.

This work demonstrates art practice as a method of experimental research into the social and technological forces that shape everyday life. As the labor of surveillance shifts from people to computer vision algorithms it becomes important to critically compare how these two forms of oversight operate on a perceptive and structural level. The technical investigation of computer vision is necessary to understanding how it functions as a tool within the structure of surveillance that bridges the technological and the cultural.
**Project Description**

*Gaze Relations* is a two-screen video installation demonstrating human gaze and computer vision. The black and white costumes in the photographs represent the way the computer vision algorithm sees, which gets explained in more detail in the technical description. Similarly, the installation walls were painted to mimic the black and white rectangular patterns employed by this algorithm.

![Image of Gaze Relations installation](image)

**Figure 1: Gaze Relations, 2018.**

The two screens are oriented vertically and the videos are synced so that both are observing the same source image. The human screen on the left reconstructs the image through the gaze of a person. This animation was created by sampling image chunks from the source image, according to where on the image the person was looking, and plots them on a blank canvas. The visualization softens the edges of the gaze sample to reinforce the human quality of the gaze visualization.
The computer vision animation on the right visualizes a body detection algorithm by plotting each feature in the image that it analyzes. This approach is unique in comparison to previous attempts to visualize this algorithm (Harvey 2010 and OpenCV 2001) because each feature in the body detection model is composited with the corresponding body part from the image and then plotted on a white background emphasizing the machinic quality of dismembering the body. The soft, fuzzy samples of the gaze visualization contrast the rigid Haar features of computer vision to emphasize the differences between them.

Six models in nine photographs add enough variety so that the audience doesn’t get bored of the repeating content. Even though the videos are looped, the meandering performance of the gaze visualization keeps the audience engaged with the piece and reinforces the dialogue between the two modes of vision. The relatively close distance between the two screens creates a tension that forces the viewer to engage with both of them at the same time.

Within the setting of the gallery these screens invoke the tradition of framed portraits with a slight twist. The audience member is not viewing a portrait, but instead viewing how someone — and something — else viewed this portrait. This second order mode of viewership lets the audience experience these two vastly different ways of seeing that sparks a dialogue about the relationship between gaze and surveillance.

The walls of the installation (Figure 1) were painted black and white to implicate the audience by surrounding them with computer readable patterns that are
blown up to size of an actual human body. Bringing the patterns outside of the digital realm gives the machinic way of seeing a physical and material agency that makes the audience more aware of the surveillance systems that are always around them, but not usually seen.

Figure 2: Fashion design and hairstyle inspired by Haar features.

Similarly, the fashion designs in the photographs that I wore during the exhibition and my hairstyle project the algorithmic representations of the body back onto the body it is examining. The visual design also borrows from the tradition of Mod fashion, which shares a similar black and white geometric aesthetic to the
algorithmic features. The intention of assigning the computer vision systems this material agency isn’t about how the world should or shouldn’t be, but how the world is.
Technical Description

This section separately discusses the technical details of the gaze and computer vision visualizations. I wrote custom software, CriticalCV, that critically visualizes the algorithmic processes of a popular computer vision tool called OpenCV. In addition I developed a tool, GazeViz, that visualizes the gaze data collected by an open source software tool called xLabs Gaze Tracking.

Computer Vision

Visualizing how computers see the world is a metaphor for how computer vision algorithms represent the physical world in code and data files. Gaze Relations specifically visualizes the OpenCV implementation of the Haar cascade object detection algorithm introduced by Paul Viola and Michael Jones in 2001. This section discusses the technical process of visualizing how OpenCV’s Haar cascade object detection method models the human body.

OpenCV is a popular open source computer vision and machine learning library whose main object detection method uses a algorithm proposed by Paul Viola and Michael Jones in their "Rapid Object Detection using a Boosted Cascade of Simple Features" paper from 2001.

The high level structure of OpenCV’s implementation of the Haar cascade method is relatively simple. It takes in a large dataset (images), trains a model (XML file) based off the input data, and applies this trained model to some real world problem (object detection) (Viola and Jones, 2011).
Haar features are computed by subtracting the sum of the pixels in the white area from the sum of the pixels in the black area and then comparing that value to a threshold value. In other words, the algorithm is comparing the brightness/darkness or the contrast of adjacent rectangles within an image. The arithmetic comparison then yields a value that indicates whether that particular Haar feature represents that section of the image. In figure 3, it is clear that the Haar feature matches the pixels in the original image because the costumed body in the black rectangle is darker than the white background to the left and right. During the training process the algorithm tries to find the Haar features that best represent the set of training images. This algorithm builds a cascade of stages that increase in feature number and complexity. The beginning stages have a small amount of features that are tested against sub windows.
of an image. If the sub window passes then it is tested against stages with increasingly more features. This cascade method makes the algorithm run significantly faster because sub windows that do not have the object we are looking for are rejected at the beginning saving thousands of feature comparisons (Viola and Jones, 2001).

OpenCV object detection method stores the trained model in an .XML file, which is made up of metadata, threshold/leaf values, and Haar features. Here are a few lines from the “haarcascade_fullbody.xml” model:

```
<internalNodes>
  0 -1 0 -5.5820569396018982e-02
</internalNodes>
<leafValues>
  5.8697921037673950e-22 -6.2811422348022461e-01
</leafValues>
...
<rects>
  1 5 12 21 -1.
  5 5 4 21 3.
</rects>
```

The shapes of the Haar features are expressed by the integers in the <rects> tag, which represent the X-coordinate, Y-coordinate, width, and height in that order. The computed value of the Haar feature discussed above is compared to the threshold value inside the <internalNodes> tag and then one of two values in the <leafValues> tag is added to a sum. This sum represents the algorithm’s confidence that the object is present in the image (Viola and Jones, 2001).

Although it is possible begin to understand this trained model by reading it in
small doses, trying to gain high-level insight of 33,314 lines of .XML code is practically impossible. How can we translate or transform this model, which is really just a long text file, into another type of data that is more palpable to a human?

In order to peek into the black box, the model needs to be translated or decompressed into another form of data. With respect to the Haar cascade method, the most natural way to peek inside the algorithm would be to visualize all the Haar features. The OpenCV installation comes with a utility called “opencv_visualisation” that visualizes all the Haar features of each stage of the object detection algorithm.

![OpenCV visualisation tool](image)

Figure 4: OpenCV visualisation tool.

This tool outputs low resolution, grayscale image sequences for each stage of the cascade. Although this tool is helpful, it doesn’t offer any other output format besides the one in the figure above. The major contribution of this visualization technique is mapping the abstract software processes of this machine learning algorithm back into visual space.
To address the limitations of the standard OpenCV visualization tool, I built a custom software tool, CriticalCV, to further investigate this algorithm. CriticalCV takes in an image and the object it’s looking for, and then outputs an animation of each feature in its dataset.

![Figure 5: Frames from CriticalCV.](image)

Translating the algorithm into a visualization introduces a few aesthetic decisions such as the choice to represent the features as black and white rectangles. Although the object detection code above describes the shape of the features as rectangles, it doesn’t explicitly describe them as black and white. But the aesthetic choice of black and white demonstrates how the algorithm is comparing the darkness/brightness or contrast of neighboring rectangles.

CriticalCV diverges from OpenCV’s and Adam Harvey’s approach by only visualizing the particular feature and corresponding body part. Omitting the rest of the photograph emphasizes the machinic quality of this form of gaze illustrating the
way it dismembers the body and jumps around in unnatural ways. This approach is also unique in how it composites the source image with the Haar feature to make the connection between bodies and algorithmic representation.

CriticalCV made it possible see how the features represent the body serving as a visual guide for the costume construction. The techniques used to make the costumes range from fabric paint and fabric collage, to sewing them from scratch. Collaging different black and white garments together worked well because it emphasized the way the software dismembers the body and doesn’t require constructing it from scratch.

**Gaze-Tracking**

Visualizing human gaze is a two-part process: recording and tracking gaze as data and then visualizing that data to make sense of it. Capturing and computing the gaze-tracking data was done with the help of an open source tool called xLabs Gaze Tracker. xLabs’ calibration system tells the user to look at particular points on the screen and then click them so that the tool can build a model that maps an input image from a webcam to a gaze point on the screen. I modified one of xLabs’ examples to display an image on the screen for 15 seconds, track the viewers gaze, and then output the gaze-tracking data. The output of this process is a set of coordinates that correspond to the points on the image where the viewer was looking over time.
In order to visualize this data, I wrote a Python tool called GazeViz that takes in gaze data and the source image and then outputs an animation. GazeViz goes through each xy data point, grabs a sample from the source image at that point, and then pastes the sample on a blank canvas.

Figure 6: a) Source image b) Mask c) Result

GazeViz grabs a square chunk from the source image (Figure 6a) and masks its with a radial filter (Figure 6b) so that the sample becomes transparent towards the edges (Figure 6c). This gets rid of the hard edges producing a softer gaze sample more representative of the way people, as opposed to computers, see.
Figure 7: Gaze trail.

GazeViz adds a trail of gaze samples with decreasing opacity to illustrate the trajectory of the gaze. The trajectory demonstrates the meandering gaze path of the human that contrasts it from the computer vision visualization. This trail can also be thought of as representing the short-term memory of the spectator.
Art Context

Adam Harvey, *CV Dazzle*, 2010-2018

*CV Dazzle* uses experimental makeup and hair design to hide from OpenCV face detection. In order to trick the face detection system, Harvey needed to understand the technical processes of how the software was detecting faces so he developed a visualization of how the software works.

![Image of CV Dazzle](image.jpg)

Figure 8: Adam Harvey, *OpenCV Face Detection: Visualized*, 2010.

www.vimeo.com/12774628.

The video above visualizes each feature that the face detection software uses to detect a face while simultaneously displaying the feature number and confidence score. This visualization is crucial to hiding from the software because it allowed Harvey to understand what the algorithm was looking for in face. Then it’s a matter
of altering the face with makeup and hair to break the expectation of how OpenCV models the structure of a face.

Figure 9: Adam Harvey, *CV Dazzle*, 2010-2018.

Adam Harvey’s creative team came up with a series of makeup and hair designs that used this vulnerability to their practical and creative advantage. They use face paint, jewelry, and hair extensions to obscure certain facial structures that disrupt the face detection in visually compelling way that draws influence from contemporary fashion. The crisp, sleek presentation of the photographs is reminiscent of a cosmetics advertisement or the cover of a fashion magazine. Harvey’s choice to draw influence from contemporary fashion not only makes *CV Dazzle* more aesthetically interesting, but also recontextualizes the function of fashion styling in the 21st century.
CV Dazzle is a significant precedent for the theoretical motivations and visual manifestations of Gaze Relations. While CV Dazzle focuses on OpenCV’s face detection, Gaze Relations instead focuses on OpenCV’s body detection, which uses the same algorithmic structure but on a different object. CV Dazzle uses makeup and hair to style the face, where Gaze Relations uses clothing to style the body. Similar to CV Dazzle, Gaze Relations draws influence from contemporary fashion to inform the aesthetic presentation of the costume design and photographs while simultaneously questioning the social and technological function of fashion in the 21st century with respect to surveillance.

Although CV Dazzle and Gaze Relations are both using fashion to mediate the relationship between people and OpenCV, their end motivations are very different. On one hand, the main motivation of CV Dazzle is to hide from computer vision algorithms by styling the face to break the expectation of the algorithm. Gaze Relations, on the other hand, highlights the expectation of the body detection algorithm to visualize how computer software sees the body in contrast to how people see the body. CV Dazzle takes a more adversarial approach to using fashion to mediate faces and computer vision. Gaze Relations doesn’t use fashion to hide from OpenCV, but instead as a tool to visually communicate how the algorithm is representing the human body, which consequently influences the fashion design.
Figure 10: Shannon McMullen and Fabian Winkler, 20/X, 2015-2016.

20/X invites the audience to immerse themselves in an interactive installation that allows them to see the processes of how a machine vision algorithm classifies an object. Audience members step inside of an enclosed box where they experience the process of machine vision analyzing a plant. A camera fixed on the plant feeds a real-time signal into the speculative software that gets broadcasted in the visualization box. The speculative software, similar to CriticalCV, is critically visualizing how the machine vision algorithm itself is learning. The plant being analyzed is located in front of the visualization box creating a weird separation of visual perceptions between the person seeing the plant as a machine and the people outside the installation seeing the plant normally. The person in the visualization box can interact
with a set of knobs that change the overlay, network depth, and feature maps of the visualization.

McMullen and Winkler were interested in showing the internal process of machine vision that led to the final classification, not just the end result that people are used of seeing. McMullen and Winkler argue “in order to understand, critique and shape the impact of machines that can see in ways that exceed human vision capability, humans will need to learn to see like machines, to understand their abstractions…” (2015-16). This critical analysis frames their art as a method of investigation into how the social and technological forces of machine vision shape everyday life.

![Figure 11: Shannon McMullen and Fabian Winkler, 20/X, 2015-2016.](image)

The title of the piece, 20/X, is a clever play on the notion of 20/20 being the ideal score of human vision encouraging the audience to engage with the dichotomy of human and computer vision. Even though 20/X is implicating machine vision within the framework of human vision in a very subtle way, it introduces an
interesting comparison. *Gaze Relations*, on the other hand, frames the comparison of human and computer vision in a more explicit way by presenting human gaze and computer vision side by side. Both of these projects use visualizations to understand how machine vision algorithms represent the world in their own language by translating it into something that people can understand.

**Behnaz Farahi, *Caress of the Gaze*, 2015**

*Caress of the Gaze* by Behnaz Farahi is a kinetic garment that changes form in response to a person’s gaze. The 3D printed garment uses a camera and computer vision software embedded in a microcontroller to detect gaze and respond by triggering an actuation system that controls the movement of nodes on the external skin. The garment doesn’t conform to the body like a shirt, functioning more as body sculpture than clothing. The project shows how wearable art has struggled to create electronic and mechanical garments that fit the body in a natural way. Although the garment is mechanical and futuristic, its complex texture is organic reminiscent of hair or tentacles. The dichotomy between the mechanic and organic visual qualities mirror the conceptual tension between this sophisticated technology and the rawness of human gaze.
The novelty of this project is its ability to materially communicate something as intangible and private as gaze in a physically and visually compelling way that surfaces social phenomena of intimacy, gender, and personal identity. Traditionally, the beholder of the gaze is thought of as exerting power over the subject of the gaze (Mulvey, 1975). *Caress of the Gaze* flips this power dynamic by using computer vision to expose the beholder of the gaze making them the vulnerable one.

This relationship between human gaze and computer vision staring back informs *Gaze Relations* in a few different ways. On the surface both *Gaze Relations* and *Caress of the Gaze* seek to visualize or otherwise materialize the phenomenon of gaze that typically operates in more subtle ways. While *Caress of the Gaze* uses kinetic apparel to represent gaze, *Gaze Relations* uses animation to reconstruct an image of the body through the gaze of the beholder. An important difference between the two is *Caress of the Gaze* deals with gaze and the body as opposed to *Gaze Relations* that deals with gaze and photographic images of the body. There are many
psychological and social factors that differentiate how people gaze upon bodies in real life and bodies in visual media. These differences stem from the fact that bodies in real life can gaze back where photographic bodies are passive and lifeless. This is significant because photographic images of bodies are subject to a more rigorous and forensic gaze by people and surveillance software for extensive periods of time with or without their knowledge or consent.

Another connection between these two pieces is how they use clothing as a medium to communicate an invisible everyday phenomenon. The physical form of *Caress of the Gaze* responds according to how people are gazing upon the body while the fashion design in *Gaze Relations* represents how a body detection algorithm models an image of the human body. Both of these approaches use the subject’s body as a vehicle to spread a social awareness.

**Happy Valley Band, *Organvm Perceptvs*, 2017**

Happy Valley Band’s *Organvm Perceptvs* is an album that tries to sonify how machines hear popular music. David Kant, the bandleader, built custom software that uses machine learning to deconstruct pop songs into its individual instruments and then transcribe them back into musical notation. Then the Happy Valley Band performs these reconstructed songs that represent how the machine learning software heard them. In practice the performed songs sound like jittery and dissonant versions of familiar pop songs emphasizing the machinic perception.
On one hand, this project functions as an advanced research in acoustics and machine listening, while simultaneously exposing the imperfections of this technology in a clever, humorous way. It is these imperfections that highlight the absurdity of training a machine to hear like a person establishing a disjunction between computation and perception. Kant’s choice of training the machine learning software on popular music introduces another layer of disjunction between computation and popular culture.

Framing the original pop songs from *Organvm Perceptvs* as cultural products, instead of music, broadens the implications on this project offering an interesting dialogue to other forms of machine perception such as machine vision. *Gaze Relations* for example demonstrates how machines see the body in the context of fashion photography. Similar to *Organvm Perceptvs*, the content that *Gaze Relations* presents to the machine learning software is encoded with cultural references that the algorithm knows nothing about. Exposing this disjunction between computation and
culture adds subtle sense of humor and playfulness to broader philosophical questions of human and machine.

Another connection between these two projects is how they both give a voice and agency to the machine by bringing it back into the real world. By physically performing the machinic interpretations of these pop songs, the Happy Valley Band is bringing the algorithmic processes of machine listening back into the space of a live musical performance. The costume design in Gaze Relations represents the way the computer sees the body by projecting the algorithmic processes of computer vision back onto the body. Both of these projects bring this machine representation of the world back into the world it’s looking at and listening to.

Mod Fashion

Artistically, Gaze Relations is inspired by the mod fashion movement from 1960’s England. The mod fashion aesthetic is well suited to represent Haar features because of its emphasis on black and white, minimal design, and geometric patterns.
Figure 14: Diana Riggs as Mrs. Emma Peel for *The Avengers*, Costume by John Bates for Jean Varon, 1965.

The costumes in *Gaze Relations* use a monochromatic color palette because OpenCV’s object detection only operates on grayscale images. The algorithm’s disregard for RGB color shows that it is looking for patterns of form, not color. This makes sense because while there is a lot of variation in the color of our clothes and skin, the form of our bodies can be generalized to fit a pattern. *Gaze Relations* uses geometric patterns to represent the geometric nature of the Haar features exposing the absurdity of trying to detect a human body using rectangular features.

The mod fashion movement was culturally inspired by modernism’s principles of embracing science, technology, and minimal design. Similarly, the visual design of *Gaze Relations* reflects the cultural and technological developments of the 21st century with regard to machine vision technology.
Theory and Concept

Male Gaze

Gaze, in the context of media studies, refers to a very specific practice of looking at visual media such as film and photography from the perspective of the camera, the audience viewing final product, and the characters within the frame. In her essay, *Visual Pleasure and Narrative Cinema*, Laura Mulvey writes about the male gaze that describes the phenomenon of depicting women in visual media as objects from a masculine, heterosexual point of view (1975). Although her essay specifically theorizes the male gaze and female object in cinema, the implications are much more broad. The phenomenon of male gaze is not “intrinsic to film, but it is only in the film form that they can reach a perfect and beautiful contradiction, thanks to the possibility in the cinema of shifting the emphasis of the look [gaze]. It is the place of the look that defines cinema, the possibility of varying it and exposing it” (Mulvey 1975). Mulvey suggests gaze is not unique to film exclusively, but that the nature of cinematography serves as a record of looking that is encoded with social relationships and gender politics.

Establishing the camera as a tool to record gaze, frames cinematography as a primitive form of gaze-tracking. *Gaze Relations* uses gaze-tracking software to record and track the precise location on the image that the person is looking. In general, cinematography emphasizes a particular frame in a 3-dimensional space, whereas gaze-tracking emphasizes a small sub-frame within that cinematic frame. *Gaze*
Relations takes a more literal approach to studying the phenomenon of gaze than Mulvey’s theory that operates on at the level of camera position and bodies in space. By presenting human gaze next to computer vision Gaze Relations implicates computer vision, or more broadly the institution of surveillance, as a social force comparable to human gaze. Mulvey says the “...male gaze projects its phantasy on to the female form which is styled accordingly” (1975). In practice, this means the male filmmaker uses visual tools such as cameras, costumes, makeup, lighting, and editing to style the female form according to his fantasy. Gaze Relations imagines a world where the photographic subject is styled according to the fantasy of the body detection algorithm. The fantasy of the body detection algorithm is a metaphor for the algorithm’s idealized representation of the body. The costumes in Gaze Relations visually represent the way that this body detection algorithm models the body.

New Aesthetic

In 2012 at SXSW, James Bridle gave a presentation about his project called the New Aesthetic about “...this way of seeing, that is itself about ways of seeing”. Bridle spent most of the presentation showing images of real world objects—furniture, buildings, clothes, tanks, and sculptures—that were in some way visually influenced by digital aesthetics. He says, “So we have pixels and a digital way of seeing appearing in the world, as an aesthetic effect. We then realise it comes from our interactions with technologies like satellites, like these things looking down on us, like digital photography, that’s producing all these new interactions in the world”
Bridle argues that the digital patterns that are becoming more ubiquitous in our visual culture are a symptom from our interaction with technologies that visually interact with the physical world. The conceptual motivations and visual manifestations of *Gaze Relations* rigorously investigate this interaction between the physical and digital by visualizing how algorithms see people.

The New Aesthetic has faced criticism, most notably from Bruce Sterling, in regard to its surface level engagement with this movement as an aesthetic. Sterling says “The New Aesthetic is a genuine aesthetic movement with a weak aesthetic metaphysics” (2012). Sterling’s criticism addresses how Bridle’s presentation functioned as an informal collection of images that participate in the New Aesthetic without spending enough time developing a theory to how these ideas fit together to form an “aesthetic metaphysics”. Although Sterling’s criticism is important to developing the ideas that Bridle alludes to in a more meaningful way, the following two art works contradict Sterling’s skepticism.

Bridle discusses Adam Harvey’s *CV Dazzle*, which as described above uses experimental makeup and hair design to hide from face detection. *Gaze Relations* visually demonstrates the processes of body detection with costumes that mimic the way the algorithm models the body. These two pieces inform the theory of the New Aesthetic beyond the surface level pixel-design couch cushions from Bridle’s presentation (2012). *CV Dazzle* and *Gaze Relations* both participate in the visual vocabulary of Bridle’s archive of images while communicating something deeper than just a glitched image. These New Aesthetic artworks offer a solution or a critical
analysis of how the visual and material culture of fashion mediates the relationship between people and software.
**Conclusion**

Overall, the piece developed into a very complex and rich project with a lot of potential for further exploration. Including human gaze and computer vision made the project a lot more conceptually dense than if I had just focused on one of them. That being said, comparing these two modes of vision resulted in a more original artwork than if I had focused on gaze or computer vision exclusively. Similarly, framing fashion as a tool of mediation between gaze, surveillance, and the body added another layer of conceptual depth that made the piece more unique at the expense of making it harder on the audience.

If I had more time, access to space, and actors I would have re-recorded the source content as video instead of still photographs. Videos of the actors posing would have made the piece more dynamic and reminiscent of surveillance.

The two biggest challenges during the last couple weeks before the show were painting the walls and playing two 4k videos synchronously. My lack of experience painting walls and the amount of time I had to do it added a lot of pressure and stress in final weeks before the exhibition. Another challenge that seemed easy in the beginning was trying to play two synchronous 4k videos. After spending multiple sessions with Kristin Erickson Galvin, we were able to get it working by setting up a server on a small computer that sent synchronous messages to two Black Magic media players that would start the videos at the same time and loop them.

*Gaze Relations*, as with most conceptual artworks, tried to balance communicating an idea without being extremely didactic about it. Besides the wall
text, the piece didn’t have any written or verbal explanation of the conceptual motivations. Considering most people didn’t pay too much attention to the wall text, the audience was largely responsible for interpreting the piece on their own. Although most people seemed to understand that the piece was comparing human and computer vision, some needed further explanation. My matching costume made it easy for people to find me and ask questions about the piece if they were confused. Most people could infer that fashion was involved in some regard, but only some understood that the clothes were mimicking the way software sees the body. The people who were already familiar with my work or put enough effort into understanding the piece really appreciated how the painted walls and costumes were informing the piece.

This project allowed me to expand on my existing skills in software while developing new creative skills in video art, fashion design and photography. Although I’ve used computer vision techniques in previous projects, Gaze Relations was less about applying this technology and more about critically examining it. In comparison to previous digital artwork that I’ve done, this work uses software in a more socially engaged way.

Ultimately, Gaze Relations is the first step in a potentially career long creative and intellectual inquiry into the nature of seeing. Over the 10 months that I’ve spent on this project I’ve developed a set of methods and tools that can be applied to future projects in gaze-tracking and critical computing. I plan on focusing on the gaze-tracking component of my project and am going to buy more sophisticated gaze-
tracking hardware after receiving a grant from the Florence French Fund for the Arts. I am currently in the middle of submitting this work to the Los Angeles Fashion Film Festival, 72U creative residency, and a number of grants. This project has already received considerable attention in the UCSC Grad Slam and L.A.S.T. art exhibition and I hope to keep pushing it further.
Works Cited


