Image Knowledge | Image Analysis

September 13, 2024

George Legrady
University of California, Santa Barbara

Image Knowledge

"Image knowledge" refers to the information and insights that can be derived from analyzing and interpreting images. This includes understanding visual content, context, objects, patterns, spatial relationships, and sometimes even inferred meanings or cultural references.

In practice, image knowledge can come from trained models that analyze images and provide insights based on learned patterns, often enabling applications in various fields such as art, science, security, retail, and more.

Examples

Scene Understanding: Understanding the general context of a scene, like recognizing a beach, forest, or urban setting.

Image Complexity Appreciation: Creative explorations in constructing an image based on the configuration of forms, texture, screen subdivision, visual elements suggesting direction, movement, balance between chaos and order, challenges to formal rules.

Sentiment and Expression Analysis: Determining emotions or moods in images, especially those with human faces or social contexts.

Pattern Recognition: Identifying patterns, such as color schemes or recurring shapes, which can be important in fields like medical imaging or fashion.

Object Recognition: Identifying objects in an image, like cars, animals, people, etc.

Facial Recognition: Recognizing or matching faces, often used in security and tagging applications.

Text Recognition (OCR): Detecting and interpreting any text within an image.

Abstract

Images play an important role in visualization. As users are more willing to adopt a product if it evokes pleasurable feelings the aesthetic appeal of interfaces becomes more important. Thus, there is a growing need to generate also images which appear aesthetically to the user. Starting with the modularities of the human visual system, we derive six dimensions of visual aesthetics. For each dimension we explore, inspired by principles of the visual arts and insights of cognitive neuroscience, which pecularities of the dimensions are particularly adequate for an aesthetic impression. Accompanied by a fair number of image examples, these considerations result in an easy to understand guideline for computer scientists and interface designers how to deal with images in terms of aesthetics.

1. Introduction

The roles of images[1] in visualization are manifold. Stone et al. [1] name four main benefits. Images motivate and attract the attention of the user and have the function to persuade her. They communicate information, which is often exploited in computer-based learning. Furthermore, they have the great power to overcome language barriers, and they support interaction. Images are especially powerful whenever it is difficult to describe the depicted information by words or numbers. This is the paradigm for most human-computer interaction applications. For example, in web design images are utilized mainly for two different purposes. They can have the function to attract the user and may be used as anticipation of the overall topic of the website. On the other hand, small thumbnails promote interaction. In e-learning images usually are the support of the informa-

Figure 1. Images as key components of visualization: navigation and maintenance.

the visualization, are shown in figures 1 and 2. The left picture of figure 1 shows an image of the environment which is augmented by data indicating a possible path for a vehicle. One could be of the opinion that such a real-time navigation system has to show ”just the image the camera captures”. But the interface designer has to decide for the specification of numerous variables that determine how the captured image is presented in the user interface. To name but a few, she has to choose color space, contrast, dynamic range, spatial arrangement of the image components (e.g., the position of the horizon), depth of field, and focal length. The right part of figure 1 shows an example for maintenance instructions for an engine. The previous statements hold true for this example, as well. The last example is given in figure 2. A virtual guide for the Guggenheim Museum Bilbao refers the visitor to architectural features of the building or gives explanations of exhibits. Two screenshots from the user interface are shown. In the context of information visualization and human-computer interaction topics such as the importance of aesthetic qualities of graphical elements of user interfaces [2] or the aesthetics of interaction [3] have frequently been addressed. Also the relation between visual and verbal information in presentations has been dealt

Effective Complexity & Formal Organizational Control

Something almost entirely random, with practically no regularities, would have effective complexity near zero. So would something completely regular, such as a bit string consisting entirely of zeroes. ***Effective complexity can be high only in a region intermediate between total order and complete disorder***

Spatial Organization
The sub-division of the image
Visual tension between forms
Simple to complex structures
Balanced / imbalanced spatial grouping of forms

Formal Relationships
Foreground / background
Depth perspective
Regular to irregular forms (repetition / variation)
Blur value between visual elements

Texture
Light / dark tones for subsections
Color range, color hue(s), color saturation
Darkness / brightness
Complementary color contrast
Complex to simple texture

Direction / Movement
Angle of forms and lines
Rhythmic variation
Clustered groupings to suggest movement

Gell-Mann, M. and Lloyd, S. (1996), Information measures, effective complexity, and total information. Complexity, 2: 44-52

Ricking the Reed(1886) Peter Henry Emerson, UK, https://en.wikipedia.org/wiki/Peter_Henry_Emerson

Ricking the Reed(1886) Peter Henry Emerson, UK, https://en.wikipedia.org/wiki/Peter_Henry_Emerson

Ricking the Reed(1886) Peter Henry Emerson, UK, https://en.wikipedia.org/wiki/Peter_Henry_Emerson

Tones (1976) Michael Bishop, Chicago Center for Contemporary Photography

Girl with Leica (1934)Alexander Rodchenko, Moscow

Laszlo Moholy-Nagy (1924), Berlin

Self-Portrait (1924) Laszlo Moholy-Nagy

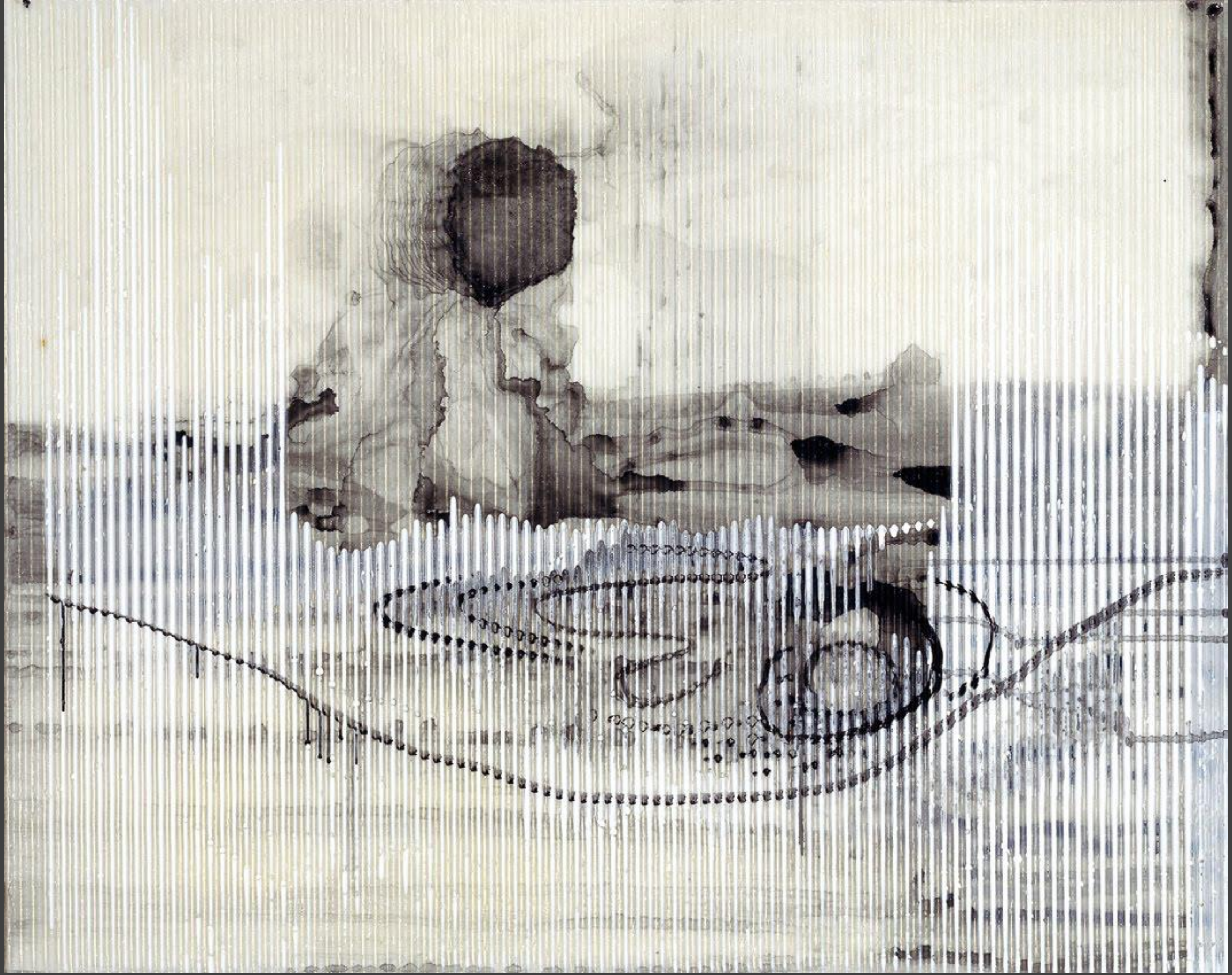
Self-Portrait with Leica (1931), Ilse Bing www.eastman.org/node/6733

Street Photography (2019) Camilo Jose Vergara

Gaza (2023) News photo

Prisoners Sleeping in a Phillipines jail (2023) News photo

Fissures and Ribbons (2004) John Baldessari, https://eastofborneo.org/articles/john-baldessari-cut-to-the-chase/



Sigmar Polke

Sigmar Polke, https://www.artistdatabase.com/sites/default/files/top-artist-works/Sigmar%20POLKE.jpg

Sigmar polke

John Baldessari Non-Rectangular Framing

https://www.mat.ucsb.edu/g.legrady/rsc/bald/index.html

Newly Released

SWIMMERS

Available from MACK

PICTURES FROM HOME



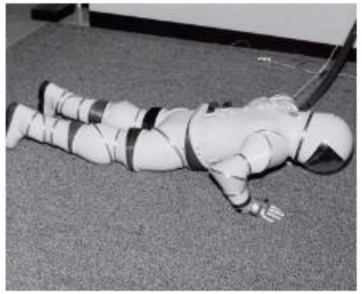
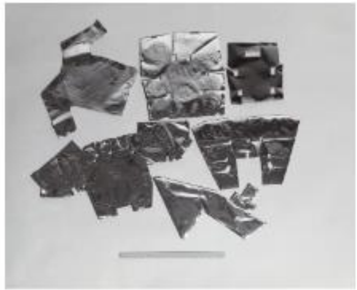






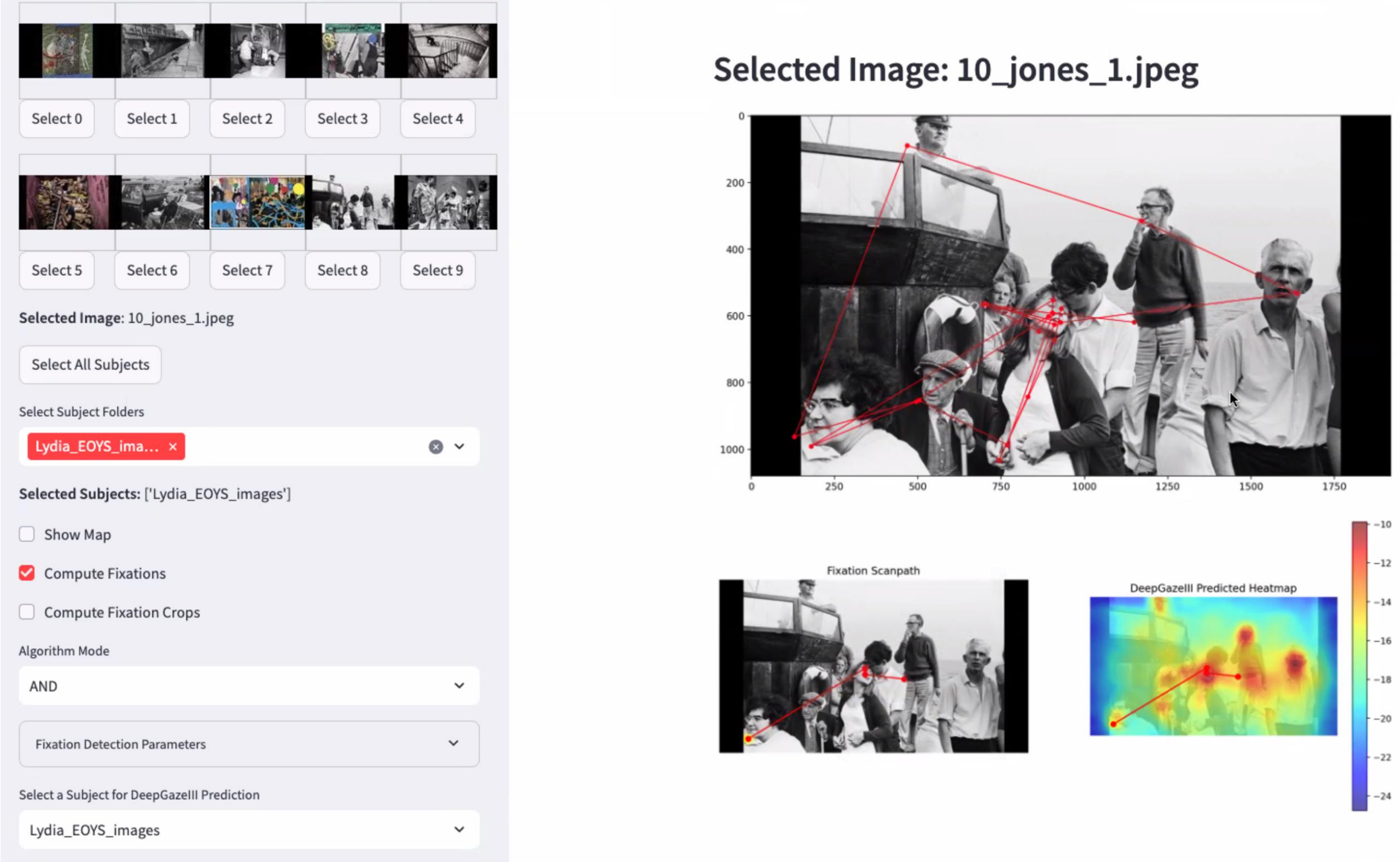


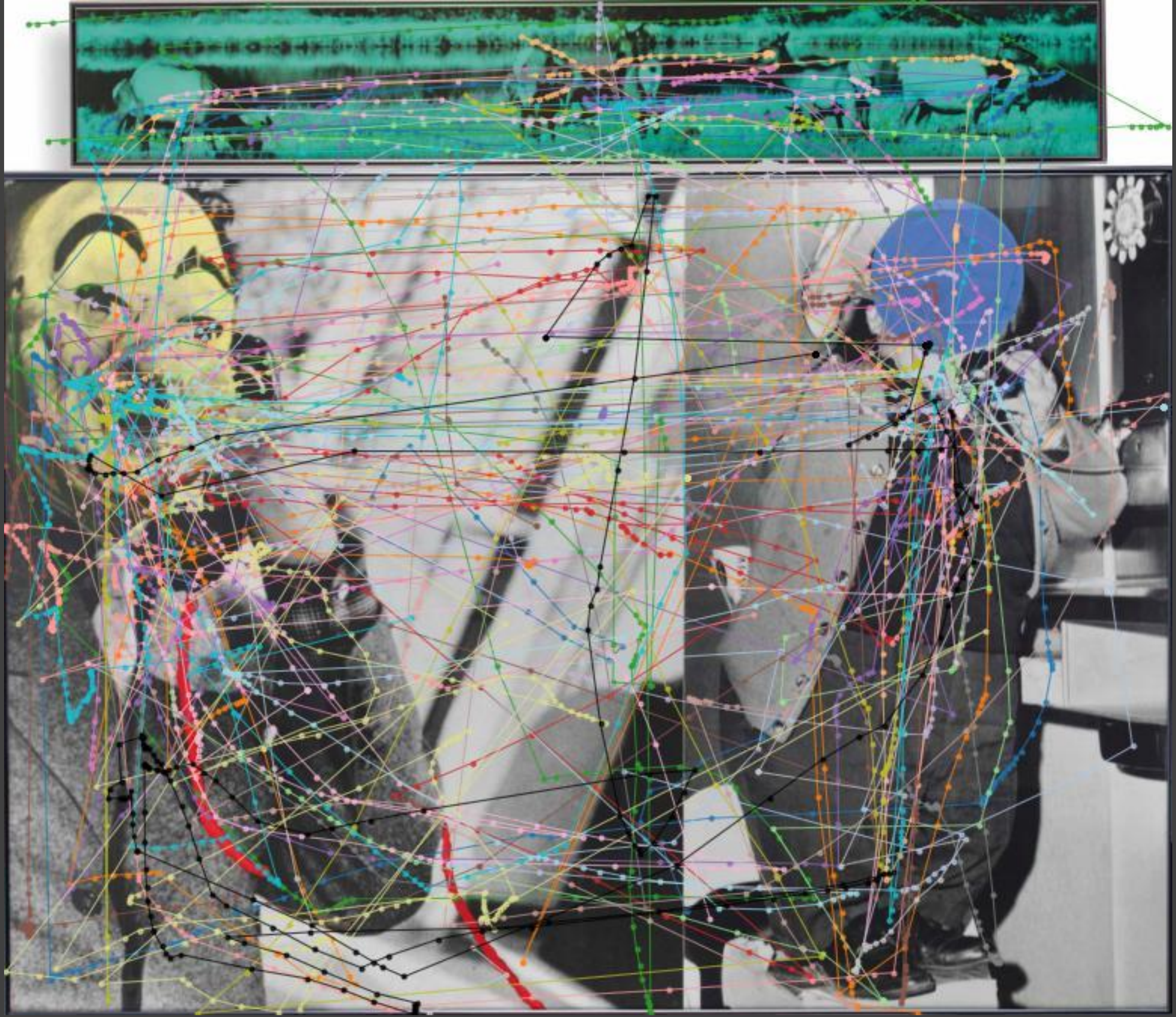


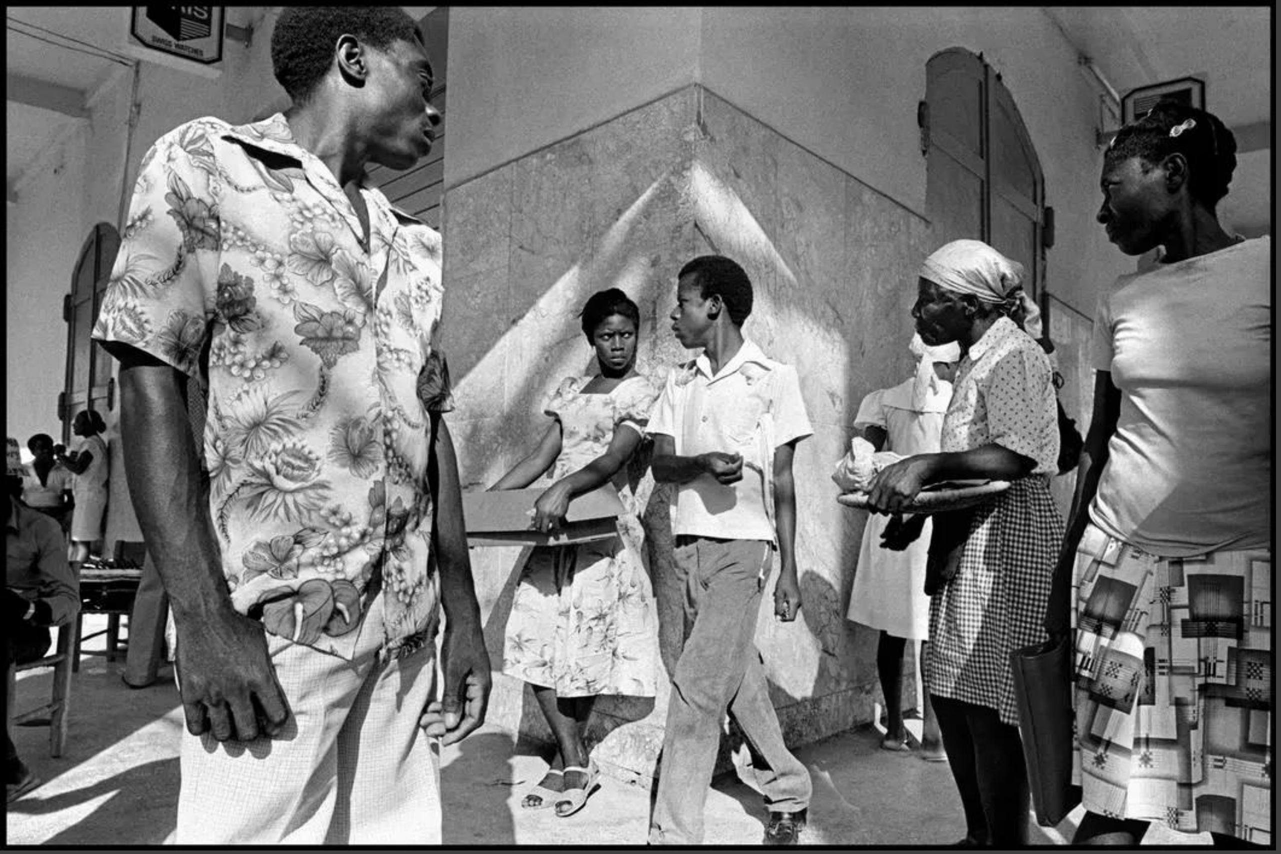

Tones (1968) Tony Ray-Jones, https://americansuburbx.com/2012/06/tony-ray-jones-photographs-of-america-and-england-1968.html

Research: *Visual Study thru Eye-Tracking to Guide AI Image Generation*, Weihao Qiu, Shaw Xiao





Two Men and Telephone (with Animals) (1988) John Baldessari, https://www.christies.com/en/lot/lot-5944715



Bruce Gilden


Hyeres (1932) Cartier-Bresson

Problem Statement

The great challenge of current generative AI image synthesis is to understand *what degree the input* (image or text prompts) can result in outcomes that *align with the image maker's intentions*

ExpVisLab Research

Implement a feedback model into the Stable Diffusion pipeline to evaluate interestingness level to adjust image generation to increase evaluation results

Guiding Perspective

Technologies leave an imprint on the data content they process and therefore the *meaning of an image to a great degree is a consequence of its construction and the technologies used*

https://artillerymag.com/abstraction-studies/

“Generative art refers to any art practice where the artist uses a system, such as a set of natural language rules, a computer program, a machine, or other procedural invention, which is set into motion with some degree of autonomy contributing to or resulting in a completed work of art.”

Generative AI can be defined as an AI system that uses existing media to create new, plausible media.

Process of Generative AI Image Synthesis:

Test prompt -> image generation

image & text prompt -> image generation

This is a transition from current computational practices which are rule-based either through software creation or the use of existing software