Visual Semiotics, Media Theory, and Cultural Analytics: a Personal Journey

Lev Manovich

When I was 17, I have read a book which shaped my intellectual preoccupations for the next few decades. This book was Yuri Lotman's Structure of an Artistic Text (*Struktura khudozhestvennogo teksta*, published in Russia in 1971).[1] While Lotman focused on the semiotic analysis of literary texts, for me this was a revelation that we can study in similar ways visual art and media. What are the elements and structures in an artistic image, and how do they determine the meanings, aesthetic impact and value of this image? Soon thereafter I have read Rudolf Arnehim's *Art and Visual Perception* (1954)[2] and Sergei Eisenstein's analysis of audio-visual montage in the opening sequence from his film *Alexander Nevsky* (1938). The texts of these theorists added to my fascination with the idea that we can analyze multiple visual dimensions in artistic works in great detail and predict the aesthetic response of the viewer.

During my 20s, I gradually realized the extreme difficulty – and eventually, the theoretical impossibility – of this project. In contrast to literary texts that use natural languages, images in most cases are not constructed from a pre-defined vocabularies of possible elements (a system of traffic lights is one of the rare exceptions). Nor do artis-

---

[1] English translation: Yuri Lotamn, *The Structure of the Artistic Text*. Translated from the Russian by Gail Lenhoff and Ronald Vroon (Michigan Slavic Contributions 7). Ann Arbor: University of Michigan, 1977. http://monoskop.org/File:Lotman_Jurij_The_Structure_of_the_Artistic_Text_1977.pdf

[2] See Rudolf Arnheim, *Art and Visual Perception*, University of California Press, 1974. Expanded and revised edition of the 1954 original book. http://monoskop.org/images/e/e7/Arnheim_Rudolf_Art_and_Visual_Perception_1974.pdf

tic images share some general grammar. This makes it impossible to develop a general visual semiotics which would enumerate possible elements of all images and ways in which they can be combined. Instead, we have to examine every image (or a series of similar images) individually to understand what its "elements" are. For example, if we take expressionist paintings by Jackson Pollock and slightly change shapes or colors of a few selected paint lines, the viewers will not notice this because these paintings have hundreds of such lines with different shapes and colors. But in Joseph Albers works from *Homage to the Square* series (1949-) that consists of only a few rectangular shapes of different colors, modifying even very slightly the size, brightness, saturation or a hue of a single shape would change the work.[3] What was invisible and insignificant in one image became visible and influential in another.

However, I have not given up on semiotic project to understand visual media. Instead, I have shifted my attention to what to me looked like an easier task – and also more urgent one. Rather of trying to understand signifying elements, their interactions and effect on the viewer in single artworks in "old media" (such as paintings), I started to think about the new artistic dimensions of "new media." This term emerged around 1990 to refer to computer-based cultural artifacts (I started to work in computer graphics and animation professionally in 1984). Like with "old media", you could talk about visual dimensions of new media such as colors, composition, rhythms. In fact, the design language theorized and developed into teaching systems at Vkhutemas and Bauhaus in the 1920s was perfectly suitable of describing visual aspects of new media. But it also had new dimensions such as interactivity, interface, database organization, or spatial navigation and new authoring paradigms such as writing code, using filters, digital compositing, and 3D modeling. My research and publications throughout 1990s have focused on analyzing such dimensions and paradigms, and this work was brought together in my book *The Language of New Media* (completed in 1999 and published in 2001).[4]

---

[3] For an example of works from this series, see http://www.metmuseum.org/toah/works-of-art/59.160/

[4] Lev Manovich, *The Language of New Media*, The MIT Press, 2001.

While the interrogation of "new media" was my main research focus in the 1990s, gradually I started to wonder what happens to the concept of "media" itself in digital era. Digital media authoring and editing tools and workflows were gradually replacing all kinds of cultural instruments used previously. While artists continued to rely on them, in culture industries painting, drawing, photo editing, creation of 3D objects and environments, graphic design, media design and sound editing were now carried out with software tools such as Photoshop and After Effects from Adobe, Apple's Final Cut, Autodesk Maya, Microsoft Office, and Avid's Pro Tools. How do these tools shape the aesthetics of contemporary media and design? What happens to the idea of a "medium" after previously media-specific tools have been simulated and extended into software? Is it still meaningful to talk about different mediums at all? What was the thinking and motivations of people who in the 1960 and 1970s created concepts and practical techniques that underlie contemporary media software?

Answering such questions leads to what we can call a "semiotics of cultural software". This analysis was developed in my next book *Software Takes Command* (first open source version – 2007; revised version – 2013).[5]

In this way, my semiotic journey proceeded from trying to understand how visual art "works" (1980s) to looking at new aesthetic dimensions of new media (1990s) and then theorizing software applications and platforms that were now used to author, distribute and interact with any media (2000s). But what about my original desire – to describe possible elements and dimensions of visual artifacts and understand how they are processed by our senses and brain, giving us meanings and emotions? In 2005 I realized that I can came back to this idea but approach it in different ways. Instead of focusing on single visual artifacts I could now analyze millions of artifacts together using computers. They would range from digitized historical artifacts to user-generated content such as photos shared Flickr, Instagram, VK and other social media platforms. Such large-scale analysis

[5] Lev Manovich, *Software Takes Command*, Bloomsbury Academic, 2013.

unthinkable in the middle of 1980s became possible twenty years later – both because of the speed of computers and also because "visual culture" was now available in digital form on a massive scale.

Its analysis would now be carried out not manually but by computers using techniques developed in the fields of Computer Vision and Artificial Intelligence. One big advantage of computational analysis is that computers can measure multitude of visual characteristics with arbitrary precision not available in our natural languages. These can range from colors of an image to a degree of smile in photo self-portaits (for example, for our lab's 2013 *Selfiecity*[6] projected we used software that measured the amount of smile on 1-100 scale). This gives us a new language for describing cultural images in ways natural languages can't – and brings us closer to the ideal of visual semiotics.

Another advantage of computers is that they can also qualitatively describe characteristics of images or their parts that are not manifested in distinct visual "elements" such as Pollock's color lines or Albers's squares. The examples of these characteristics are gradients and textures, or degree of sharpness and blur in photo, or speed of movement in a video.

My methodological shift from studying single visual artifacts to analyzing massive collections of such artifacts parallels the shift in how we experience visual culture. Single-artifact research and "close reading" was logical for 20th century when as cultural consumers we also were focusing on single works. We went to the cinema to see a particular movie, or to a museum to see particular artworks, or listened to a single music recording at home over and over. The media available to us was limited in numbers and we would spend significant time with individual artifacts. I remember, for example, that as a teenager looking hundreds of times though the same books with art reproductions in our home library. A few

[6] *Selfiecity*, 2013, http://selfiecity.net/

images of modern art from these books that particularly touched me would be imprinted in my memory.

And now? Visual search and recommendations in Google, Yandex, YouTube, Instagram or Pinterest expose us to endless images and video, while websites of major museums invite us to browse hundreds of thousands of digitized artworks and historical artifacts. A visual "message" or a "sign" (to use semiotic terms) is now never isolated but instead is a part of the large series which we experience as infinite (do you have a feeling for how two billion images people share daily look like? If it was four billion, would you notice?).

In 2007 I established the Cultural Analytics Lab to both think theoretically about the use of computers for the study of visual culture, and to work on concrete projects with a variety of datasets.[7] After completing over 40 projects I put together everything I learned about using computers to analyze visual culture at scale into a new book also titled *Cultural Analytics* (2020).[8]

One of the ideas in this book particularly relevant to semiotic imagination of the third quarter of the 20th century is a possibility of a "science of culture". The agglomeration of people in the growing megacities during the 19th century made the "society" directly visible in one way, and the growth in government statistics made it in another way, and the idea of "sociology" (the science of the social) was born. The parallel explosion in numbers of cultural artifacts and interactions with them agglomerated together in the web and social media platforms in the early 21st century, together with the relatively easy ways to collect billions of them and analyze via computers similarly suggests the idea of a "science of culture". Such science does not need to try to discover hard "cultural laws", but can instead be uncovering many cultural patterns, and in fact in the last ten years we see hundreds of thousands of numbers of publications that do this within computer science, information science and computational social science.

---

7 *Cultural Analytics Lab* site, http://lab.culturalanalytics.info/
8 Lev Manovich, *Cultural Analytics*, The MIT Press, 2020.

I see a connection here to 1960s semiotics and structuralism because these paradigms were also to a large extent aimed at making analysis culture less impressionistic and subjective and instead more methodological, organized, and science-like. Thus, Roland Barthes' classical essay *The Photographic Message* reads at first like a science article: «The press photograph is a message. Considered overall this message is formed by a source of emission, a channel of transmission and a point of reception» (Barthes 1961, 194).[9]

Does this mean that the computational analysis and modeling of cultural data with their similarly systematizing and formalizing motivations will also eventually lose their energy and attraction (because we will realize the inability of these approaches to fully account for richness and individually of cultural artifacts and interactions)? Or will they allow us to go beyond limitations of semiotics in 20th century? Time will tell.

For me, the key motivation in working with "big cultural visual data" is about creation of methods and tools that allow us to see it in the first place. It is not about replacing with statistical models or neural networks everything I learned about culture and the social world during my life, or all the ways of thinking about visual media I learned from other theorists, or my intuitions. But before I can think about media today, I need to see it, and this basic act became very problematic when people share billions of images every day. Thus, my trajectory took me from a semiotic perspective on singular works of art I can easily see with our bare eyes to the design of interfaces and techniques of computational "seeing" contemporary media that are necessary because of its scale.

Does this mean that original semiotics project to understand how structures of cultural texts and objects are organized and how they create aesthetic effects is no longer relevant? Or that single images no longer matter? Not at all. Consider only large numbers of young

[9] You can download Barthes' "The Photographic Message" from *A Barthes Reader*, edited by Susan Sontag (New York: Hill and Wang, 1982), 194-210, from here: https://monoskop.org/images/5/59/Barthes_Photographic_Message.pdf

Instagrammers worldwide who spend days editing single photos they would pose, and agonizing over literally single pixels (Manovich 2017).[10] The single, unique and well-crafted did not go away – and understanding why exactly this rather than that image, or this rather than that filter moves us more than another may wait a long time until neuroscience progresses sufficiently. So while we can now describe quantitatively the structures in art and media objects and situations with great detail, understanding how art "means" and "affects" remains unsolved. But do we really want to understand it any more than we want to understand love, desire, memory and other dimensions of human experience?

[10] Lev Manovich, *Instagram and Contemporary Image*. The four parts of the book were written between 12/20/2015 and 12/26/2016 and posted online as they were completed. The final version available as a single PDF combines revised versions of these chapters, an Introduction (written in August 2017 – essentially, it is a new chapter), and an Appendix. This version was released in September 2017. http://manovich.net/index.php/projects/instagram-and-contemporary-image

References

Bibliography

Arnheim, R., *Art and Visual Perception*, Berkeley and Los Angeles, University of California Press, 1974 [expanded and revised edition of the 1954 original book]. http://monoskop.org/images/e/e7/Arnheim_Rudolf_Art_and_Visual_Perception_1974.pdf (Accessed 15/10/2024).

Barthes, R., "The Photographic Message", in Sontag, S., (Ed.), *A Barthes Reader*, New York, Hill and Wang, 1982: 194-210. https://monoskop.org/images/5/59/Barthes_Photographic_Message.pdf (Accessed 15/10/2024).

Lotamn, Y., *The Structure of the Artistic Text*, translated from the Russian by R. Vroon (Michigan Slavic Contributions 7), Ann Arbor, University of Michigan, 1977. http://monoskop.org/File:Lotman_Jurij_The_Structure_of_the_Artistic_Text_1977.pdf (Accessed 15/10/2024).

Manovich, L., *The Language of New Media*, Cambridge, The MIT Press, 2001.

Manovich, L., *Software Takes Command*, New York, Bloomsbury Academic, 2013.

Manovich, L., Cultural Analytics, Cambridge, The MIT Press, 2020.

Sitography

Albers, J., "With Rays" (1959), *Homage to the Square* series, in http://www.metmuseum.org/toah/works-of-art/59.160/ (Accessed 15/10/2024).

Manovich, L., *Cultural Analytics Lab*, established in 2007, in http://lab.culturalanalytics.info/ (Accessed 15/10/2024).

Manovich, L., et alii, *Selfiecity* project, 2013, in http://selfiecity.net/ (Accessed 15/10/2024).

Manovich, L., *Instagram and Contemporary Image*, 2017, in http://manovich.net/index.php/projects/instagram-and-contemporary-image (Accessed 15/10/2024).