Image Digitization

Hubertus Kohle

https://orcid.org/0000-0003-3162-1304

Abstract A picture is worth a thousand words, as the saying goes. And yet – or precisely because of this – it has been marginalized in European intellectual history or branded as misleading and unreliable. In the Digital Humanities, too, there has been a delay in devoting attention to the image, which is itself reflected in its constitution at German universities. The article deals with the technical characteristics of the digital image, its fundamental peculiarities as an entity of pixel sets with different resolutions and raster graphics as opposed to vector graphics. This is followed by considerations on the structure of image databases and their elementary properties as well as the description of complex search options across astronomically large amounts of image data, which are also becoming increasingly semantically accurate as part of the development of multimodal large language/image/audio models. The presentation concludes with some cursory considerations on AI-controlled image generators.*

Keywords Digital Image, Image Databases, Digital Image Analysis

1. Introduction

The prioritization of the word in European intellectual history and, to a certain extent, the consistently prevailing skepticism towards the image, among others also in the field of reformed theologies, are likewise common in the history of the Digital Humanities. This field of study began in the middle of the 20th century in a philosophical-theological context with the work of the Jesuit Roberto Busa (Father Busa), who used his relationship with Thomas J. Watson, the founder, and head of the *International Business Corporation* (IBM), to use the computer to create an index of the works of Thomas Aquinas. Obviously, the computer was used exclusively for text processing. Not until a quarter of a century later, did the image begin to be used, initially more with textual metadata. In the 1980s, the image as such received research attention (Vaughan 1987, 215–221).

The reasons for this development are both ideological and technical. In the European context and the European history of philosophy, which, according to Alfred

* This chapter, including quotations in foreign languages, was translated from German by Brandon Watson.

Hubertus Kohle: Image Digitization, in: Christopher A. Nunn and Frederike van Oorschot (eds.): Compendium of Computational Theology 1. Introducing Digital Humanities to Theology, Heidelberg: heiBOOKS, 2024, S. 117–126. DOI: https://doi.org/10.11588/heibooks.1521.c21943

North Whitehead, is a series of “footnotes to Plato,” the word and its proximity to the idea take precedence. However, the fact that an image is technically more complex to process digitally is more relevant. Vaughan once said he was unable to test his image-oriented search system (called “Morelli”) because there were hardly any large quantities of digital images available at the time of his invention.[1] There were reasons for the lack of availability: A reasonably high-resolution image – the meaning of which will be shown – could be several hundred KB in size, and were the image close to slide film quality, the image could easily be several MB. High-resolution images, such as those in certain cases by Google’s Arts Project,[2] even reached several hundred MB. Storage space was scarce in the 1980s, and a 20 MB hard drive costed several thousand marks. Yet, a large hard drive could easily hold 20 books with only written text.

Despite technological advances, the positionality of the image itself within the Digital Humanities remains marginal. These are mostly located in linguistics and literature and are therefore purely text oriented. Only in a few cases do academic fields investigate the field of objects, yet do not primarily get into the field of images. Recently, there have been more professorships in the field of visual computing being advertised within the Digital Humanities, despite their popularity in the general field of Computer Science.[3] The focus is on augmented and virtual reality, machine vision, and image processing. Applications can be found primarily in medicine, architecture, design, and gaming. In most cases, AI methods are crucial for successful applications.

2. Technical Basics

Since the late 19th century, photographic reproductions based on originals have determined teaching and research practice, such as in art history, which is the most important auxiliary academic field for theology in the field of images (cf. Dilly 1979). Prior to the 19th century, there were graphic techniques used on a large scale for the reproduction of works of art, transforming the exclusive unique specimen into one of social communication. Copperplate engraving was used in the first place for these types of image reproductions. In the late 20th century, digital reproductions became an additional technique used, which quickly replaced traditional analog forms of production. Currently, chemical-based slide libraries are disposed of, moved to university archives, or stored in the basement. That they are sometimes the subject of art-historiographical and scientific analyses proves rather than disproves their marginality.

1 Verbal communication.
2 See https://artsandculture.google.com (Accessed: 16 June 2024).
3 The DFG has been funding an interdisciplinary priority program with a focus on cultural studies that is dedicated to digital imaging for several years (https://www.digitalesbild.gwi.uni-muenchen.de [Accessed: 16 June 2024]).

The digital image differs from the analog image in that the digital image is a discrete collection of pixels (picture elements) arranged in vertical and horizontal rows in a grid, which are created in an electronic exposure process (see Besser et al. 2003). To say a digital image is a "discrete" collection means that the transition from one pixel to another is made in jumps, whereas it is continuous in an analog image. With digital cameras and scanners, devices have been available since the 1980s that carry out this exposure process either synchronously or diachronically. Scanners are available for a range of different analog source products, such as slide scanners used to digitize analog slides.

Related to the fact that the computer only deals with zeros and ones (this only in parenthesis), the discrete organization of the pixels is often understood from the humanities as an indication that digitally supported analyses only research unambiguity and leave no room for ambiguous results. However, if one considers that the discrete approaches the analog asymptotically as the resolution increases, then there is only a theoretical difference. The related problem of blurring has become a topic of discussion in the Digital Humanities (see Borek et al. 2022).

Depending on the resolution, the raster consists of a different number of pixels, whereby the greater the number of pixels, the greater the detail of an image. One unit of measurement for resolution is dpi (dots per inch), which denotes the number of pixels per inch = 2.54 cm. The size of the original is a decisive factor. If a 2.54 × 1.27 meter oil painting is digitized at 100 dpi, then there are 50 million dots (100 × 100 for the height × 50 × 100 for the width); if I do the same with an analog reproduction of the painting measuring only 25.4 × 12.7 cm, one gets 500,000 dots (10 × 100 for the height × 5 × 100 for the width). In addition to black-and-white and grey-scale images, in which the former is reduced to either black or white, and different brightness values are generated in the latter, color images are currently prevalent, which are intended to produce the most realistic representation of the world. The greater the color depth, the more varied the colors are displayed. The predominant color model in which colors are combined is the RGB color model, in which the primary colors red, blue, and green can be mixed additively in all color tones, i. e., by superimposing them. Realistic color effects are achieved when each of the three colors is coded with 8 bits = 256 colors, which results in a total of 16.7 million colors. High color depth (256 × 256 × 256) combined with high resolution produces large files. While storing and distributing these files is done using modern hard disks with large storage capacities and the large network bandwidth, the calculation of these images is still a task requiring the most powerful machines. The task is simplified though, given that the images can be saved in compressed form, which can reduce their size by a two-digit factor, but the common jp(e)g format *(Joint Photographic Experts Group)* also leads to a loss of quality initially barely visible to the human eye yet can be noticeable with greater compression. The compression level can be preselected when processing the files.

In furtherance of the above-mentioned jpg format, several other formats have been developed, starting with the unchanged original formats. These include bmp *(bitmap)* and tiff *(tagged image file format);* a tiff file can also be compressed. In contrast to jpg, png *(portable network graphics)* does not lose quality upon compression, although the file size is also compressed – albeit to a lesser extent. The gif *(graphics interchange format)* is suitable for moving animations. The manufacturer specific *raw* format contains the most information, which is preferred by professionals as an output format for image processing, but like bmp requires a lot of storage space.

In addition to raster graphics, there are vector graphics, which do not encode each individual point, but mathematically define the geometric and color nature of a pictorial object. The latter has advantages and disadvantages: if a uniform red line does not have to be represented by a whole number of equally colored (red) pixels, but as an instruction, making all points between coordinate x and coordinate y equally red, then it requires less memory space and is more scalable, i.e., the image can be enlarged without loss. With a raster graphic, the appearance in this case becomes less clear because the individual pixels emerge more and more clearly. The disadvantage of vector graphics is a slower processing time, since a coordinate is not visible as a number and must first be converted (rendered) back into a pixel. Incidentally, the principle of vector graphics works best with schematic visualizations with few local color variations.

3. Image Databases

Within the Visual Arts, there is a series of databases in which digital reproductions of two- and three-dimensional works of art are stored. Museum directories are a primary storage location, with the advantage that the digital reproductions are most likely produced according to the originals and not according to printed reproductions, where moiré effects repeatedly occur, one of the negative effects of digitizing a paper print. Comprehensive databases are available primarily in the American *artstor*[4] and in the *Foto Marburg* database[5], which also has historical relevance, as it ultimately draws on photo campaigns dating back to the early 20th century and therefore partly includes objects that have not yet been affected by the potentially extensive losses of the two World Wars. There is also *Prometheus*[6], which represents the synthesis of over a hundred individual directories searchable by a common interface. All large collective databases, not referring to individual museums but to entire clusters of museums or architecture, contain several million images and are continually expanding their holdings.

4 See https://www.artstor.org (Accessed: 16 June 2024).
5 See https://www.bildindex.de (Accessed: 16 June 2024).
6 See https://www.prometheus-bildarchiv.de (Accessed: 16 June 2024).

The *Prometheus* database is predominant in the German-speaking world, partly because it is inexpensive compared to the American databases. The principle of this meta-database, which does not provide any data of its own apart from the search functionality, is suitable for the peculiarities of the internet. It is based on a notion of "shared authorship" and relies on the influence of the crowd, which in total is unbeatable. However, this principle is not without disadvantages, as the quality of the data provided by the individual suppliers is rarely checked to reach a critical mass as quickly as possible. The makers of *Prometheus* recognized this limitation at the outset, just like all the internet entrepreneurs who invest heavily in the rapid generation of large numbers of users: if one wishes to penetrate the market – which in the case of *Prometheus* means, above all, being used by university institutions – one must make a number of artwork reproductions that can be used primarily for teaching purposes available very quickly. Only then will teachers feel compelled and justified to convert the projection system from slide to digital – which has now occurred on a large scale. As mentioned, doing so comes at a cost, namely, the at times poor quality of the images and the doubling or tripling four to fivefold increase in reproductions after better-known works of art. The creators of *Prometheus* have responded to the issue by asking users to mark the individual images with asterisks so that the poorer ones can be eliminated, a procedure ingeniously relying again on "collective wisdom" (style-defining: Surowiecki 2005). Reducing the number of poor-quality images could also be accomplished using the procedure based on SIFT proposed by Schneider (2019). That *Prometheus* increasingly integrates the databases of individual large museums likewise serves to increase the quality.

The resolvable disadvantages are offset by a significant advantage resulting from the distribution of the work over several parties, which in turn can be used everywhere on the network. In the past, any institution involved with images had its own archive. The limited workforce led to a patchy stock of images, noticeable even in extensive collections. If hundreds (if not thousands) of digitizers are active across the network, then there are much larger amounts of data/images able to be used by any (registered) internet user, despite any duplications. The joint use of the various, often technically divergent, systems has been successfully realized for some years with software solutions such as IIIF *(International Image Interoperability Framework).*[7]

The data is entered into a relational database where the information is organized in tables that are related to each other and queried using sql *(search query language).* Recently, graph databases are coming into the limelight provided their flexibility. In them, data entities called *knots* (e.g., objects, but also names, etc.) are defined, which are related to each other via *edges* and stored in rdf *(resource description framework)* triples. The system can reach any level of complexity.

7 See https://iiif.io (Accessed: 16 June 2024).

4. Image Search

One of the advantages of digitization in the field of image databases is that individual objects can not only be located in a specific place but can also be accessed in a differentiated way depending on the level of detail of the description data. In the traditional slide library, Michelangelo's "Moses" is found under the artist's name, which is of little help to anyone interested in the iconography of Moses, if one does not know that Michelangelo also devoted himself to the subject. Of course, it is possible to set up an alternative category "Iconography" and classify "Moses," but someone interested in Michelangelo who does not know that he created a "Moses" might not find what they are looking for. The solution is to place two reproductions of the same work under both the artist and the iconography. However, doing so would inflate the picture collection, and if a category of date of creation were added (in this case approx. 1514) under which all works from this period were categorized, or one of the preserving institutions (here the church of San Pietro in Vincoli in Rome), then the number could be increased three or fourfold. From a cost and space perspective, this solution would not be desirable.

The digital realm is different. Each work can be digitally indexed in a variety of different ways without needing to be stored multiple times. If the work is registered in a database in which, artist, title, date, technique, and storage location are stored, then each of these categories can be searched, even in combination. A search for the date "1514," for example, would return all the works stored in this database from the year 1514. Then, when combined with the artist "Michelangelo," everything that the Renaissance artist created in that year will be shown. In this way, statistically oriented investigations are possible such as: which museum/church (repository) owns the most works by Michelangelo? In which period did Michelangelo create the most works? What are Michelangelo's preferred themes? The latter, however, is somewhat less trivial: Work titles are ambiguous and thus cannot be determined by simply naming the title; it can be "Maria," "Mother of God," or even "Mary." A classification system such as *iconclass*[8] can eliminate this ambiguity and the different names can be *mapped* onto each other. A relatively easy-to-use tool for carrying out demanding tasks is available in *Museum Analytics*[9], which was programmed expressly for interested parties who do not come from IT. It remains to be seen whether LLMs *(Large Language Models),* which are currently in vogue in the form of ChatGPT, e.g., will be able to simplify the research possibilities.

The high number of reproductions in these image databases in conjunction with their content indexing can lead to a methodological decanonization. While under analog conditions the access possibilities were limited and depended largely on the knowledge of the searcher, so that when visiting the card catalog, one needed to know

8 See https://iconclass.org (Accessed: 16 June 2024).

9 See https://dhvlab.gwi.uni-muenchen.de/max (Accessed: 16 June 2024).

in advance the intended outcome, the digital age has changed the process entirely. If the media logic of digitization is followed and it is not seen as a simple extension of analogization, there are more diverse search options, implying that the search results are also diverse. Even if the old rule that new media imitate the old applies on both the supply and, above all, on the demand side, so that the methodical transformation remains manageable, then the tendency towards diversification will increase as the range of offerings continues to differentiate.

5. Image Calculations

From a purely practical perspective, digitization offers considerable advantages in the management of and search for images. However, image digitization also enables new analytical possibilities, which undoubtedly represent a paradigm shift in the field. Until now, it has been possible to describe images, identify their forms and content, and place them in a historical or stylistic context. However, it was not possible to go into them directly, or even indirectly. A digital image is not just a grid of pixels; each individual pixel can be principally addressed, identified, and changed. Whereas the search options described above were limited to the information referred to as metadata, i.e. information added by a human or technical intelligence, here there is the completely new possibility of focusing on the object itself, not the secondary information about it, e.g. in a histogram, with which the distribution of color and brightness values in an image can be displayed diagrammatically. The conclusions are intriguing.

The computer is a formalistic machine; it recognizes color dots at a particular location, and nothing more, but it does so with tremendous speed and accuracy yet without recognizing semantics. For the computer, the image of a dog is simply a cloud of pixels. In Art History, formalism has once again received attention, which would progress further by the image-addressing procedures in the digital age. Heinrich Wölfflin, who achieved worldwide success more than a hundred years ago with his *Kunstgeschichtliche Grundbegriffe [Basic Concepts of Art History]* and distinguished works of art from the Renaissance and Baroque periods purely on the basis of their formal characteristics, is experiencing a renaissance in this context (Wasielewski 2023). Media scholars assume a freeing of the image from the clutches of human categorizations, which ultimately leads to a de-historicization, as historical thinking is ultimately founded in human language (Pias 2003). Along with this intimate connection, the problem takes on an almost eschatological meaning, often associated with the discourses of Posthumanism.

Just as in *Optical Character Recognition* (OCR) of texts, where the available letters as pixel clouds are given meaning in a training process, these pixel clouds can also be trained for dogs and all other objects, which is currently being done using neural networks trained for semantics in a machine learning process. These neural networks are modeled on neuronal connections in the brain and, in the eyes of techno-utopians, are still less powerful than the brain, mainly because the latter still enables far more nerve cell connections than even the most powerful computers. By way of example, one of the consequences is that a small child only needs to be shown a few images of a dog to gain an abstract idea of the dog and to recognize dogs that do not correspond exactly to the examples shown, whereas the computer must be given several hundred, if not thousands, of these examples to derive a general model.

For image searches, this means that going forward the searches will increasingly be based on the images themselves rather than on metadata. From a purely technical point of view, this means that larger quantities of images can also be searched, i.e., those that go beyond art databases containing millions of images. The problem sounds purely theoretical, but when one considers that 350 million photos are uploaded to Facebook alone every day, the practical side quickly becomes apparent. However, much research needs to be done in Art History in order to present a functional and reputable search system based on direct image addressing. It is already possible to recognize a person in a picture, even a person pouring water over another person's head is no longer an insurmountable problem for advanced AI programs. But recognizing John the Baptist baptizing Jesus Christ in the Jordan River is hardly feasible with the existing approaches. To do so, one needs to increase massively the amount of training data and supplement the 400 million image-text combinations used to train models such as CLIP from the Open AI Initiative, especially with more specific combinations. In the case of individual sciences, this result could primarily be achieved by drawing on a large amount of expert knowledge available in the form of machine-readable research texts. Such approaches also show how important it is to publish scientific texts as comprehensively as possible in open access regarding potential solutions.

6. Image Forgeries

AI cannot only be used directly to address images. It can also be used to produce images that no longer have a depictive reference to reality. This was known before as well, but only as image modifications, not as completely new inventions. These depictions are thus referred to as fake images or deep fakes. AI supported systems such as *DALL-E, Midjourney,* or *stable diffusion,* in addition to serious applications in which the non-realistic focus cannot be doubted, allow such uses. The sociopolitical effects remain to be seen. If a picture of the Pope running away from the police recently went viral on the internet, then a well-informed individual should recognize the falsity of the image, just because of its extreme improbability. However, a picture of the same Pope amused by a caricature of Mohammed is likely to cause displeasure in the Islamic world, even if the Pope has never actually seen such caricatures.

Literature Cited

Besser, H. (2003). *Introduction to Imaging. Revised Edition.* Ed. by S. Hubbard & D. Lenert. Los Angeles: Getty Research Institute. URL: https://www.getty.edu/publications/virtuallibrary/0892367334.html (Accessed: 16 June 2024).

Borek, L., Zeppezauer-Wachauer, K., & Ketschik, N. (2022). Eindeutig Uneindeutig. Zur Modellierung von Unschärfe in der Mediävistik. In *Mittelalter. Interdisziplinäre Forschung und Rezeptionsgeschichte* [Blog]. DOI: https://doi.org/10.58079/rha5 (Accessed: 16 June 2024).

Dilly, H. (1979). *Kunstgeschichte als Institution. Studien zur Geschichte einer Disziplin.* Frankfurt a. M.: Suhrkamp.

Pias, C. (2003). Das digitale Bild gibt es nicht. Über das (Nicht-)Wissen der Bilder und die informatische Illusion, *zeitenblicke,* 2(1), no pag. URL: https://www.zeitenblicke.de/2003/01/pias/pias.pdf (Accessed: 16 June 2024).

Schneider, S. (2019). Über die Ungleichheit im Gleichen. Erkennung unterschiedlicher Reproduktionen desselben Objekts in kunsthistorischen Bildbeständen. In *DHd 2019. Digital Humanities multimedial und multimodal.* Frankfurt a. M./Mainz: Zenodo. DOI: https://doi.org/10.5281/zenodo.4622051 (Accessed: 16 June 2024).

Surowiecki, J. (2004). *The Wisdom of Crowds. Why the Many are Smarter than the Few and How Collective Wisdom Shapes Business, Economics, Society and Nations.* New York: Doubleday.

Vaughan, W. (1987). The Automated Connoisseur. Image Analysis and Art History. In P. Denley, & D. Hopkin (Eds.), *History and Computing* (pp. 215–221). Manchester: Manchester University Press.

Wasielewski, A. (2023). *Computational Formalism. Art History and Machine Learning.* Cambridge, Mass.: MIT Press.

Wölfflin, H. (1915). *Kunstgeschichtliche Grundbegriffe. Das Problem der Stilentwickelung in der neueren Kunst.* Munich: Bruckmann. DOI: https://doi.org/10.11588/diglit.27250 (Accessed: 16 June 2024).

Literature Related

Altenhöner, R., Berger, A., Bracht, Ch., Klimpel, P., Meyer, S., Neuburger, A., Stäcker, Th., & Stein, R. (2023). DFG-Praxisregeln "Digitalisierung". Aktualisierte Fassung 2022. Online: *Zenodo.* DOI: https://doi.org/10.5281/zenodo.7435724 (Accessed: 16 June 2024).

Brown, K. (2020). *The Routledge Companion to Digital Humanities and Art History,* New York/London: Routledge.

Cohen, D., & Rosenzweig, R. (2006). *Digital History. A Guide to Gathering, Preserving, and Presenting the Past on the Web.* Philadelphia: University of Pennsylvania Press.

Jannidis, F., Kohle, H., & Rehbein, M. (Eds.). (2017). *Digital Humanities. Eine Einführung.* Stuttgart: J.B. Metzler Verlag.

Kuroczynski, P., Bell, P., Dieckmann, L., Hoppe, S., & Münster, S. (Eds.). (2018), *Computing Art Reader. Einführung in die digitale Kunstgeschichte.* Heidelberg: arthistoricum.net [= *Computing in art and architecture,* 1].