5. Image Search[5]
(Martin Warnke)

The World-Wide Web confronts us with the paradox that the amount of published material grows at exceedingly accelerating speeds and, at the same time, finding relevant items becomes easier than ever. Very often, generic search techniques like the one by Google give more reliable results than hand-crafted navigational structures made by the publishers of the sites themselves. Full text indexes made with brute computational force easily outperform all those refined, semantically oriented retrieval methodologies.

This hare-and-tortoise relationship between abundance and indexing pinpoints the mediality of search itself. A needle will be found in a haystack only as the result of a proper search[6]. So that part of the web that has not been indexed by the search engines is very accurately called the "dark web". Between perception and object sits a medium, and this medium is called a "search engine". Like light and sound for seeing and hearing, search has to be called a medium of perception for the web.

But as breathtakingly efficient as an automatic search for words on the web may be – and this is our second paradox – this does not work at all with images. This shrew has not been tamed by those glorious systems like Google. This is the very point of this essay. The punchline will be that this is not a computational problem solved by faster algorithms or machines but one of media practice.

Search Basics

To learn what the peculiarities of image searches are, we briefly revise the stunning accomplishments of text search on the web.

Every page in the web has its address, the URL, which is similar to the page number that serves as an "address" in a book index. Exactly as with

---

[5] This chapter is a modified translation by the author; an earlier version was previously published in German as: Warnke, M. (2009). Bildersuche. In Gesellschaft für Medienwissenschaft (Ed.), *Zeitschrift für Medienwissenschaft, Heft 01/2009 – Schwerpunkt: Motive* (pp. 28-37). Berlin, Germany: Akademie-Verlag.

[6] Joseph Weizenbaum compared the internet more with a dung hill than with a haystack: Merschmann, H. (2008, January 8). Joseph Weizenbaum – Der zornige alte Mann der Informatik. *Spiegel Online*. Retrieved from http://www.spiegel.de/netzwelt/tech/0,1518,527122,00.html

a book index, every word has to be associated with its address for an index of the web. The process is the same in both cases: for every address – URL or book page number – all of the words to be indexed must be listed, what is much easier with a book than with the web, since the book pages are already there. By contrast, web search engines have to crawl through the sea of words first.

This yields a list that shows all index entries for a certain page or URL. Afterwards this list is inverted, since we don't want to know the words by their places but rather the place to find the words that interest us. This inverted list now is the basis of the search, since it yields the addresses for every word. In books, these addresses are sorted numerically; for the web, complexity has to be reduced in the first place because of the very high number of findings for common words. The algorithms that rate the addresses are the business secrets of the search engines: Google calls it the page rank, and it has a surprisingly high correlation with the significance of a word and vice versa.

Computational power and the smart technique of the inverted lists allow to compute a search for text passages, thus making them Turing computable. This is due to the fact that no semantics are involved. Text search engines operate purely syntactically, ideally fitting into a computational model.

The thesaurus of a language like German comprises some 100,000 words[7]. This is an extremely small number compared to all writeable letter combinations. If you count just all the six-letter combinations of twenty-six letters, you get more than 300 million strings: aaaaaa, aaaaab, ..., zzzzzy, zzzzzz.

The thesaurus of lexically proper words consisting of six letters comprises approximately 5,000 words. That is a reduction by a factor of 60,000: only every 60,000th writable six-letter string is indeed a German word: *abakus*, *abbild*, ..., *zypern*, *zysten*. This is roughly the same with, e.g., the English language.

Niklas Luhman may have had this in mind, too, when he defined a "medium" as follows: "We will call those evolutionary achievements media that work at ruptures of communnication and function precisely to transform improbables into probables".[8]

But what becomes more probable in this case? The event of understanding, and this by at least a factor of 60,000, as far as six-letter

[7] http://www.dict.cc (Retrieved 6 May 2009).
[8] Luhmann, N. (1994, p. 220). *Soziale Systeme – Grundriß einer allgemeinen Theorie.* Frankfurt am Main, Germany: Suhrkamp. [M. Warnke, Trans.]

words are concerned. Words help to distinguish between the noise of sound and the significant parts of speech. Understanding in the medium of text is such an improbable event that trying to do it without the use of words is like betting on a high-score lottery win.

Coding Grammar

Sesame Street, by the Children's Television Workshop, is famous for turning dry stuff into fun. There is that section *many words begin with …,* and then Grover comes to chant the letter of the day, say the popular *C*. Under this letter in his dictionary, the abecedarian will later on find his favorite word, be it *cookie* or *Christianity*. Grover's training will have paid off then.

The lexic order of words – not a semantic one – relies on several media technologies from times before the computer that are of utmost sophistication. Firstly, it relies on writing that conveys the ephemeral sound of speech durability and opens it up for post-processing. For ideographical script an arrangement of words is right at hand, though that is hard to learn, like in Chinese. It is much easier with short phonetic alphabets, since the order of words results directly from the order of the letters, even without the use of computers. But first, speech must be separated into its lexic atoms: words must be invented first.

Ivan Illich writes: "We sometimes forget that words are creatures of the alphabet. […] Our 'words' became significant, as other syntactical elements of speech, only after centuries of alphabet use, with which they have been 'bred'".[9] Writing later reacted to this breeding by inventing the blank space, this very significant nothingness, worthy sister of the zero, which in arithmetics plays a comparable role as the explicit presence of absence. Just as the zero was necessary to complete the system of positional notation, the blank space was necessary for the invention of the word. Early writing in phonetic alphabets don't show spaces that separate words. Only later did spaces help to subdivide the stream of text by breaking it into words.

Scholasticism, relying on the word, has invented all these well-known technical aids: paragraphs, tables of contents, chapters, and the preconditions of the address system of pagination in books.

To find something in text without words, we had to do pattern matching of substrings within strings. Probably under similar conditions, mankind

---

[9] Illich, I. (1991, p. 42). *Im Weinberg des Textes – Als das Schriftbild der Moderne entstand*. Frankfurt am Main, Germany: Luchterhand. [M. Warnke, Trans.]

would have invented the concept of the word. In the Abchasian language – a language with no highly developed text technology, which has been written since 1932 in Latin, from 1938 in Georgian and since 1954 in Russian – "The awkward question 'how could she, this poor thing, not give it to him?' is expressed by a single word: *Jeschpaleseymtagweschasaj?"*[10]

To separate "this poor thing" from "him" requires an explicitly coded grammar, which could help describe how to construct complex expressions from lexemes that could be separately listed and found in dictionaries. And grammar, following pertinent authors[11], has been brought about by the printed book alone.

Pecularities of Image Search

But with images there is neither syntax nor lexis, which is of the utmost importance for word searches. The number of isolated image signifiers is infinite, likely not even countable, and thus the cardinality of its set bigger than that of all computable numbers[12]. This is very different from the finite set of words, with some synonyms, that can be searched linearly in a dictionary.

What kind of strategies could we think of to search for images, since that simple running along the number line does not work as it does with words?

There is, in the first place, image search as text search. You may look for images that are somewhere placed with certain words, as with Google's image search. That this is nothing more than word search could be simply proven by searching for "invisible" images, getting lots of results. Genuine image search engines that don't do word search, on the other hand, have to look for image characteristics such as color or form, because an image has to be classified for similarity, not for exact identity with the search term as with text. Image and similarity or likeness are very close together: "And God said, Let us make man in our image, after our likeness."[13] Mitchell writes: "'image' is to be understood not as

---

[10] Zand, B. (2001, May 28). Klang der Kieselsteine. *Der Spiegel*, p. 221. [M. Warnke, Trans.]

[11] Giesecke, M. (1992). *Sinnenwandel Sprachwandel Kulturwandel*. Frankfurt am Main, Germany: Suhrkamp.

[12] Turing, A. M. (1937). On computable numbers, with an application to the *Entscheidungsproblem*. *Proceedings of the London Mathematical Society, s2-42*(1), 230-265. doi:10.1112/plms/s2-42.1.230

[13] Genesis 1:26 King James Version.

'picture', but as spiritual similarity."[14] This spiritual similarity that clings to every image, and also to the picture, has to be transformed into a computable measurement. We would expect an image to be found even if it is a bit lighter or darker than the search image, actually being different in every pixel. A hit would be allowed even if the number if pixels would differ because of a different image resolution.[15]

Images can't be classified by identity or difference. Images could be the same if they are very much alike, but they become very different ones by small alterations alone. Since modern times images are, following Foucault, "no longer the form of knowledge but rather the occasion of error, the danger to which one exposes oneself when one does not examine the obscure region of confusions."[16] That is the reason why images have such a bad scientific reputation: they unambiguously lack unambiguousness.

So there is only the classification for similarity, which is a pre-rationalistic category. Similarity then has to be expressed by a number, measuring the distance between the characerics of the searched image and the resulting hits. Subjects to these measurements are colour distribution, forms, and patterns.

This is called "Query by Image Content", and the St. Petersburg Hermitage Museum has implemented such a system.[17]

The results are actually similar in respect to the chosen characteristics, but their purely syntactical nature becomes quite obvious. Contrary to text search, there is no strong relationship between the syntactic dimensions and the semantics of an image, due to the lack of a word concept. Even facing some progress in computer graphics there is no hope that this will ever change.

The image not only demands its two dimensions of the search process, in contrast to the one of linear search, it also possesses such variety that computed similarity does not yield satisfactory results. Video adds yet another dimension to the two that are already present in images, yielding a third for the elapsing time, thus not facilitating the situation at all.

Numbers, text, and images are three basic media that have been unified by the binary code of their representation in the computer, but remain

---

[14] Nelson Goodman, after Mitchell, W. J. T. (1986, p. 31). *Iconology: Image, text, ideology*. Chicago, London: University of Chicago Press.

[15] How fragile similarities are becomes obvious in the duckomenta collection: http://www.duckomenta.de.

[16] Foucault, M. (1994, p. 50). *The order of things*. New York, NY: Random House.

[17] http://www.hermitagemuseum.org

separate in terms of their indexability for cultural use. So it is not enough to have better recognition algorithms for color, form, or pattern search: what is actually lacking is the *reduction of the image variety to a limited number of image atoms*, as text has done with speech, breeding the word. Nelson Goodman's position on the difference between images and text convinces me very much: "The boundary line between text and images, pictures and paragraphs, is drawn by a history of practical differences in the use of different sorts of symbolic marks, not by a metaphysical divide."[18]

But this means we will have to wait to see if image atoms actually will emerge, so to say as side effects of image search in the internet. The work that is done on facial recognition and searching for certain types of similar images that the big search engines do at the moment will be very interesting to observe.

Toward Systemization

There were attempts to systemize image elements. A catalogue that carries the notion of textual image description in its name is the one of iconography. In a German version,[19] it starts with "Alpha und Omega" and ends with "Zypress", as any other dictionary could start and end. With *IconClass*[20] a controlled vocabulary has been proposed that allows for precise verbal representation of the image content of Western European art history, but in which an image no longer is an image. Art history itself talks about images more than it shows them.[21]

For those lacking words, obviously not for the art historians, there is the classic image dictionary "Bildwörterbuch" by the German publisher Duden. In Webster's Dictionary, images go in the margins. In the course of the Duden's many editions since 1937, the image standards have changed so much that we can observe another fundamental difference between text and image: the time scale on which forms evolve. With images, the time periods at which characteristics change significantly is much shorter than they are for text. Especially for art: "Language has to be old, artworks have to be new."[22]

---

[18] Ibid.

[19] Mohr, G. H. (1984). *Lexikon der Symbole – Bilder und Zeichen der christlichen Kunst*. Köln, Germany: Diederichs-Verlag.

[20] http://www.iconclass.nl

[21] Claus Pias once declared the absence of the image to be the precondition of the dicipline of art history. Personal communication.

[22] Luhmann, N. (1999, p. 40). *Die Kunst der Gesellschaft*. Frankfurt am Main, Germany:

Recently biometric procedures have served as cataloguing means even for private image collections, as in Google's Picasa or Apple's iPhoto, by training the systems with names given to facial regions. Here we have a case for computability, since the metrics of a human face are robust and precise. The mere syntactical nature could, as always, be observed at the failures: the author, e.g., has been mistaken by iPhoto for a wheel cap pictured from the side, which is not very flattering but thought-provoking nonetheless.

To generate a powerful order of images for search purposes, similar to the one for text, seems not to be a matter of technology alone, but mostly one of the cultural use of images that first of all must establish image dictionaries with isolated image atoms. It is a very interesting question to ask, to what extent and in which way the image search algorithms themselves will influence those future dictionaries, as the media technology of print has done for text, triggering word dictionaries – especially given the way Google has left its traces in our culture of script, e.g. in the way presentations and scholarly papers are written. That the imagination, images, and text would be the same in principle does not seem maintainable to me, considering the many differences between these basic media. Let Mitchell paraphrase Goodman's position, explaining the significant opposition between text and image for the phenomenon we are discussing here: "The image is syntactically and semantically dense in that no mark may be isolated as a unique, distinctive character (like a letter in an alphabet), nor can it be assigned a unique reference or 'compliant'. Its meaning depends rather on its relation with all the other marks in a dense, continuous field."[23]

In a Turing Galaxy where we could only expect answers to questions that are processable on digital computers, we can expect image catalogues and orders of images to co-evolve with image search engines. It would not surprise me very much if only those images survive effectively that are usable results of an image search process. All the others may disappear into the dark web, where search engines may no longer reach.

An icono-Grover could then approach our children and chant: *There are many images that look like* … for which we may effectively search with our engines in later life. Without having to use that old-fashioned *C*, that then is reserved exclusively for *cookie* or *Christianity*.

Suhrkamp. [M. Warnke, Trans.]

[23] Mitchell, W. J. T. (1986, p. 67). *Iconology: Image, text, ideology*. Chicago, London: University of Chicago Press.

The power of images would not have vanished but drifted to marginal realms, taking along with it their moving variety, the video: to art, e.g., as subversive form of communication, as a projection screen for wild thought.