

Wolfgang Ernst  
 Kunsthochschule für Medien  
 Peter Welter Platz 2  
 50674 Köln  
 Tel.: 0221/20189-245  
 Fax: 0221/20189-17  
 E-mail: ernst@khm.de

Stefan Heidenreich  
 Alte Schönhauser Str. 29  
 10119 Berlin  
 Tel.: 030/282 77 54  
 E-mail: stefan.heidenreich@rz.hu-berlin.de

### **Bildarchive**

In jüngster Zeit haben Bildarchive einen Schritt gemacht, dessen Tragweite noch nicht ganz überschaut werden kann. Durch inhalts- und ähnlichkeitsbezogene Bildsuche (Content- oder Similarity-based Image Retrieval) verweisen Bilder zum ersten Mal wieder auf Bilder. Der Umweg über Schrift und Sprache entfällt.

Mit dem Kurzschluß der Zeiger innerhalb von Bildermengen könnte ein Grundstein gelegt werden für eine Art von bildlichem Wissen, die zuvor schlicht aus technischen Gründen nicht zu denken war. Bisher konnten Bilder nur über eine verschriftete Suche und damit eine begrenzte Anzahl von Entscheidungen in einem auf den Kanal der Buchstaben beschränkten Feld ausgewählt werden. Die Auswahl hat auf die Art des abbildbaren Wissens einen enormen Einfluß. Man kann soweit gehen zu sagen, daß die Adressierung eines Archivs dessen Wissen bestimmt. Es ist immer nur das verfügbar, was tatsächlich abgefragt oder ausgewählt werden kann. Bildarchive, die über Sprache abgefragt werden, bleiben sprachliche Archive, so viele Bilder sie auch enthalten. Das läßt sich an den Archiven der Kunstgeschichte (Foto Marburg, Iconclass) genausogut nachvollziehen wie an den meisten der neueren, auf sprachlich fixierbaren Inhalten aufgebauten Datenbanken. Mit der Programmierung von Datenbanken, in denen Bilder wieder auf Bilder verweisen, wird eine neue Art von Auswahl möglich. Erst unter diesen Bedingungen könnte tatsächlich visuelles Wissen abfragbar werden. Wenn dieses Wissen über buchstäbliches Wissen hinausgeht, dann hätten bislang schriftbezogene Abfrageformulare diesen Überhang aus den Archiven hinausgefiltert. Erst unter der Bedingung visueller Adressierbarkeit könnte das Wissen eines Bildarchivs wirklich gewußt werden.

### **Bildinhalt, Ähnlichkeit**

Dem stehen noch immer einige Hindernisse im Weg. Die Begriffe Content oder Similarity bergen Reste von Schrift, selbst dort, wo nicht mehr geschrieben wird. Content, der Inhalt eines Bildes, tut so als sei der Inhalt eines Bildes etwas anderes als das Bild selbst (Picard 1996). Damit fällt der Inhalt in das Modell der Semiotik zurück, das zwischen einer Anwesenheit (dem Zeichen) und dem notwendig Abwesenden (seiner Referenz) differenziert. Similarity, die Ähnlichkeit zwischen Bildern, vermeidet diesen Fehler, um an einer anderen Stelle in die Falle zu gehen. Das Problem ist hier subtiler. Ähnlichkeit tendiert dazu, etwas zu setzen, worin sich Bilder ähnlich sind. Sobald versucht wird, dieses Etwas der Ähnlichkeit der menschlichen Wahrnehmung anzupassen - und das ist die Regel -, produziert auch ein Konzept der Ähnlichkeit eine "bedeutsame" Aussage, das heißt ein Verhältnis, in dem das Bild etwas bedeutet und darüber erreichbar wird. Üblicherweise wird Ähnlichkeit über einen Kriterienkatalog, oft nach statistischen oder psychologischen Untersuchungen, definiert (Jain/Santini 1996). Dadurch werden Bezüge zwischen Bildern immer wieder von vorneherein auf einen bestimmten, nämlich einen "sinnvollen" Anschluß festgezurt. Dagegen stellt sich die Frage, ob eine Ordnungsrelation innerhalb der Bildermenge nicht desto eher erfolgreich sein wird, je mehr sie auf einen bestimmten Anschluß und einen bestimmten Sinn verzichtet. Man kann zur Begründung dieser Hypothese auf einen vergleichbaren Fall im Medium der Schrift verweisen. Gegenüber beiden Strategien sollte ein Zugriff auf Bilder diskutiert werden, der von jeglicher Sinnbildung, sei es durch Ähnlichkeit oder Inhalt, absieht und "Sinn" nicht als Bedeutung, sondern im Sinne der etymologischen Ableitung durch die Gebrüder Grimm als Deuten auf etwas, als vektorielle Bestimmung, mithin also als Zeiger auf Daten liest. Das bedingt ein Sich-Einlassen auf den genuin "archäologischen", d. h. kalten, von Einbildungskräften freien Blick des Computers, der Signifikanten rein signifikant, also (im Sinne der Hermeneutik) deutungsfrei liest. Der Unterschied ist dem zwischen Scanning, also dem einfachen Einlesen von Daten, und OCR als Deutung der Daten, also tatsächlichem Lesen, vergleichbar.

## **Das Alphabet als Beispiel**

Der angestrebte Effekt könnte dem ähneln, den die Übernahme des phönizischen Alphabets in Griechenland hatte. Erst als (analog zum Schicksal von Archiven in Umbruchsituationen) der Inhalt der Ordnung, nämlich die ursprünglich buchhalterische Funktionalität der einzelnen, auf Bildersprachen zurückweisenden Buchstaben, nicht mehr gewußt wurde und der bloße Konsonantenzichensatz übriggeblieben war (Goody 1986; Hafemann 1996: 13f.), konnte das Alphabet phonetisch aufgeladen zu dem Kanal des Wissens werden, als den wir es heute kennen: eine Schrift, welche diskrete Lautstrukturen wiedergibt (Phonographie) und die Zeichen nicht mehr ideographisch an Wortbedeutungen koppelt (Logographie, in welcher das Zeichen immer noch strukturanalog zur Welt des Bezeichneten steht). Wo ein Zeichensatz inhaltsleer, also als rein externes Zeichenmaterial ansehbar und damit zur Aufzeichnung rein differentieller Phoneme einsetzbar wird, gerät er in die Abhängigkeit konkreter Sprachen und verliert als universaler Kode (womit der griechische Begriff "barbarisch" für Fremdsprachen ganz zutreffend ist, die damit vom Klang her benannt sind) (Havelock 1990:95).

## **Bildordnungen**

Wenn aber (im schriftarchäologischen Sinne) Zahlen und Buchstaben eine gemeinsame Matrix als Informationsspeicher haben, ist der Anschluß an einen Bildbegriff, der als Kalkül gerechnet wird, denkbar, und damit die Befreiung des Bildarchivs von seiner logozentristischen Unterwerfung unter die schlagwortartige Adressierung. Es kann nicht darum gehen, analog zur Sprache ein Alphabet, also eine Reihe von Bildern anzustreben. Dem steht die Mehrdimensionalität des Bildkanals entgegen. Was aber an der Errungenschaft des Alphabets abzulesen und auf den Kanal der Bilder übertragbar sein könnte, ist folgendes:

1. Erst durch eine Ordnung, die in dem Sinn allgemein ist, daß sie nicht eine bestimmte Bedeutung impliziert, kann Information sortiert und gefunden, also zu Wissen werden. Die Ordnung der Bilder ist in einer im Vergleich zum Alphabet komplexen Relation von Bildern zu suchen, mit der ein visuelles Interface korrespondiert.
2. Erst unter den Bedingungen dieser Ordnung kann visuelle Information tatsächlich über ihre sprachliche Eingrenzung hinaus wieder auf weitere Bilder verweisen. Damit kann die noch immer bestehende Bindung des Bildes an Sprache und deren Informationsfluß entfallen. Es stellt sich die Frage, welche Anschlüsse, welches Wissen und welche Kommunikationsformen ein derartiger Bildkanal hervorbringen könnte.

## **Das Programm "Suchbild"**

Das Programm beläßt Bilder zur Archivierung in dem Format, das sie ohnehin haben, also als zweidimensionale Flächen von Farbwerten. Die Suchfunktion ist so einfach wie möglich. Jeder Versuch, über Datenkompression oder Extraktion von Merkmalen bereits Entscheidungen über Bilder zu treffen würde die Suche mit einer Intelligenz ausstatten, die es, wie oben ausgeführt, gerade zu vermeiden gilt. Eine Suchfunktion, die in einer dem Alphabet vergleichbaren Einfachheit läuft, kann allgemein eingesetzt werden. Sie läuft über Suchbilder, die die Originalbilder auf eine Größe von 32x32 Pixel verkleinern. Als Maß der Ähnlichkeit gilt die durchschnittliche Abweichung pro Pixel. Mit diesem "rohen" Begriff der Ähnlichkeit wird keine bildliche Übereinstimmung in irgendeinem sinnvollen Bezug garantiert. Das ist von vorneherein auch nicht die Absicht des Programms, dem es nur darum geht, eine bildlich vorstellbare Verbindung zwischen Bildern wiederzugeben. Im Sinne einer quasi "gedankenlosen" Nutzung der Suche wird sowohl dem Nutzer als auch dem Programm eine Analyse des Bildes in Hinsicht auf etwas, das es zeigt, erspart. Eine erfolgreiche Suche setzt immer voraus, daß der Nutzer gemäß derselben Struktur denkt, die das Archiv vorgibt. Der Entschluß, auf eine möglichst simple bildliche Relation zurückzugreifen, geht davon aus, daß Sehgewohnheiten sich weit eher einer formalen Ordnung anpassen als umgekehrt. Neben der einfachen Verkleinerung der Bilder auf Suchicons werden weitere Möglichkeiten der Vorverarbeitung erprobt. Eine komplexe Vorsortierung der Bilder auf der Basis von Bildpyramiden und Octrees hat sich als sehr speicherintensiv und nur unwesentlich schneller erwiesen (vgl. Samet 1990:23). Auch eine Analyse der Bilder auf Texturenergie oder Konturen führt nicht zu einer höheren Anschaulichkeit der Suchergebnisse. Die Verarbeitung der Suchbilder durch eine Fourier-Transformation befindet sich in der Erprobung. Eine Suche auf der Grundlage von ortsinvariant transformierten Bildern ermöglicht es, von der Position visueller Einheiten im Bildrahmen zu abstrahieren.

## Anwendungen

Daß Verfahren zur Archivierung von Bildern und zum Image Retrieval auf ein Vielzahl von Anwendungen hoffen lassen, wird allgemein vorausgesetzt. (vgl. etwa für einen Überblick die Ausgabe der Zeitschrift COMPUTER 9/95) nahliegend sind sowohl Anwendungen im Internet ( Demos laufen z.Zt. unter <http://www.qbic.almaden.ibm.com/~qbic/qbic.html>; [www-white.media.mit.edu/~tpminka/photobook/](http://www-white.media.mit.edu/~tpminka/photobook/); [www.ctr.columbia.edu/~jrsmith/visualSEEK/visualSEEK.html](http://www.ctr.columbia.edu/~jrsmith/visualSEEK/visualSEEK.html); [www.virage.com/](http://www.virage.com/) ), wie auch der Einsatz in bildverarbeitenden Berufen, etwa beim Videoschnitt oder bei Grafikern. Voraussetzung ist selbstverständlich die vollständige Digitalisierung der Bildermengen. Das Programm "Suchbild" wird zur Zeit in Zusammenarbeit mit der Kunsthochschule für Medien, Köln, für den Einsatz auf einer CD-ROM der Plakatsammlung des Museums für Gestaltung, Zürich eingerichtet.

Es wäre aber kurzsichtig, die Perspektiven des "Image retrieval" in gerade naheliegenden Anwendungen zu suchen. Da die gegenwärtige wissenschaftliche wie kulturelle Praxis schriftfixiert ist, werden Verfahren der Bildsuche allzuoft als Anhängsel einer wie auch immer funktionierenden schriftlichen Datenverarbeitung gedacht. Erst wenn sich zu Bildermengen ein schriftunabhängiger Zugriff etabliert hat, können Bilderarchive tatsächlich visuelles Wissen bilden. Dann spielen die Suchmöglichkeiten des "Image retrieval" eine zentrale Rolle, da sie bestimmen, was tatsächlich in einem Bildarchiv gewußt, und das heißt hier: eingegeben und wiedergefunden werden kann.

Auf dem Weg zu diesem Ziel wird vielleicht weit weniger die komplexe Modellierung der menschlichen Bildverarbeitung im Vordergrund stehen, auf die augenblicklich das Hauptaugenmerk der Forschung gerichtet ist, sondern eher einfache und schnelle Suchverfahren sowie das Design einer anschaulichen Schnittstelle. So vorteilhaft sich die Maus zur Erhöhung von Kontrolle und Redundanz erwiesen hat, so sehr könnte sie sich bei der Bildsuche als Hindernis erweisen. Durch ihre Bindung des Blicks verstellt sie den Bildschirm als optische Aktionsfläche. Bildsuche der Zukunft könnte aussehen wie Computerspiele von heute. Die Durchquerung der Bilderflut (man müßte die Metapher ins Positive wenden und von einem "Bilderbad" sprechen) wird sich nicht den ans Auge rückgekoppelten und deshalb vergleichsweise langsamen Zeigeoperationen bemessen, sondern an der Geschwindigkeit des Blickwechsels.

---

**Jack Goody 1986:** The Logic of Writing and the Organization of Society, Cambridge

**Ingelore Hafemann 1996:** Ägyptens Beziehungen zum Sinai in pharaonischer Zeit, in: ARCUS - Berichte aus Archäologie, Baugeschichte und Nachbargebieten, Heft 3 / 1996: Antiker Orient, 1-14, Potsdam

**Eric A. Havelock 1990:** Schriftlichkeit: das griechische Alphabet als kulturelle Revolution, Weinheim

**Ramesh Jain / Simone Santini 1996:** Similarity Queries in Image Databases. Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, S.646-650.

**Rosalind Picard 1996:** A Society of Models for Video an Image Libraries. MIT Media Laboratory Perceptual Computing Section Technical Report No.360. <http://www.media.mit.edu/~picard/>

**Hanan Samet 1990:** Applications of Spatial Data Structures, Reading (Mass)