

Content Based Retrieval in Image Databases Workshop / Einführung

Workshop / Introduction

Gerd Stanke

Gesellschaft zur Förderung angewandter Informatik e. V. (GFal)

Rudower Chaussee 30, 12489 Berlin

Tel.: +49-30-6392-1610, Fax: +49-30-6392-1661

E-mail: stanke@gfai.de, Internet: www.gfai.de

Zusammenfassung:

Gegeben wird eine Einführung, in welchem Umfeld sich seitens der Anwendungen und der Forschung die Arbeiten zum Content Based Image Retrieval (CBIR) bewegen. Relationen zu Präsentationen auf vergangenen EVA-Konferenzen und aktuellen Workshopbeiträgen werden gezogen.

Abstract:

An Introduction is given respecting the user requirements and the state of the art of Content Based Image Retrieval research. Relations to papers of former EVA conferences as well as to the workshop papers are communicated.

„Ein Bild sagt mehr als tausend Worte“ - dem Menschen ja, nicht aber dem Suchmechanismus in einer Bilddatenbank. Ein passendes Anfragewort filtert deutlich sicherer relevante Informationen aus großen Datenbeständen heraus als „tausend Bilder“, eine Erfahrung, die viele von uns haben und der in den Beiträgen zu dem Workshop widersprochen werden kann.

In einem Bild sind auf Bildpunktniveau „unendlich“ viele Informationen enthalten. Wir haben in der Evolution Mechanismen entwickelt, diese Informationsmenge in effektiver Weise parallel zu verarbeiten, zu filtern und in eine Handlungsbedeutung zu überführen. Wie das abläuft, wissen wir bisher nur partiell (vgl. *Workshopbeitrag Hagendorf*) und sind so auch nicht in der Lage, dieses Vorgehen dem Rechner zu vermitteln. Die Informationen in einem Bild sind nicht linear strukturiert. Das Bild repräsentiert sich selbst. Ein Wort, ein Suchbegriff auf der Basis eines Alphabets ist kompakt als sequentieller Kode dargestellt und so herausragend für eine strukturierte Suche in Datenbanken geeignet. Das Wort kategorisiert. Die Suche nach Begriffen geht auch deshalb so gut, weil Begriffe normieren. Eine Sonne ist eine Sonne in vielen Bildern, in denen sie in unterschiedlichster Erscheinungsform auftritt. Auf eine „Kultur des Suchens“ geht *Bienert* ein

Die gegenwärtig entstehende Vielfalt der Bilder ist „unendlich“, die der Worte „begrenzt“. Mit jedem neuen Bild entsteht 1 MB oder mehr, mit einem neuen Buch(inhalt) auch 1 MB. Die Verbreitung der digitalen Bildtechnik tut ein Übriges, um neue Bildinhalte in Mengen zu produzieren. In die gleiche Richtung laufen Entwicklungen, die für ein Bild nicht ein Bild sondern viele erstellen. Anwendungsbereiche, die einerseits Bildmengen produzieren und andererseits bildinhaltsorientierte Anfragen fordern sind: Verlage und Agenturen mit Bilddatenbanken, Privatnutzer, die Kriminal- und Überwachungstechnik, die Werbebranche, Versandhäuser, Bauteilkataloge im Ingenieurwesen und elektronische Produktkataloge, Designer und Architekten, Wissenschaft und Medizin, kunsthistorische Anwendungen, z. B. [8]. Die Motivation hinter einer Suche ist entsprechend vielgestaltig: „Was war schon da“ (Logos etc.), „was gibt es noch“ (Variationen von Produktbildern).

Schlagwortssysteme sind etabliert (z. B. ICONCLASS) und in ihrer Leistungsfähigkeit durch die Indexierung determiniert. Sie sind schnell und leistungsfähig. Die Verschlagwortung unterliegt subjektiven Einflüssen. Eine Sprachabhängigkeit und ein hoher Arbeitsaufwand sind dem Vorgehen innewohnend. CBIR Systeme arbeiten objektiv und automatisch, sie sind aber auf die Entwicklung passender automatischer Indexierungsverfahren angewiesen, genau hierin liegt das Problem.

Eine Vielzahl von Ansätzen zu CBIR-Systemen ist im Netz präsent, die ersten sind gut 10 Jahre alt. Der interessierte Sucher stößt auf: Cires, Qbic, ImageMinderr, Simba, retrievr, Blobworld, Viper, Simplicity, Webseek, Ikona, Fire, ImageFinder, Photospy, PictureFinder, Excalibur, Image Sorter und weitere. Sie arbeiten mit Bildvorlagen, interaktiven Skizzen oder auch Descriptoren. Einige Eigenschaften und Systeme werden bei *Lambeck* in einer Übersicht charakterisiert, in [9,10] sind untersetzende Informationen und Hinweise auf die Systeme zu finden.

Mit MPEG7 steht eine Struktur für die Beschreibung und das Wiederfinden medialer Informationen, darunter auch von Bildern, in Datenbeständen bereit. Datenbanksysteme geben Strukturen vor, der Aufwand für eine Nutzung liegt in der Füllung mit Inhalten. *Meiers* gibt eine Einführung und bringt Beispiele, wie eine Indexierung automatisch laufen kann.

Für das Retrieval im direkten Sinne des Wortes liegen, wenn die Beschreibungs- und Ablagestrukturen definiert sind, Lösungen vor. Die eigentliche (Teil-)Frage, die es für CBIR-Systeme zu lösen gilt, ist, wie kommt ein Objekt, ein Bild, besser seine Beschreibung, in diese Struktur. Vielleicht sollte der Workshop lauten Image Description Generation?

In einer einfachen Weise sind für die Ableitung von Bildinhaltsbeschreibungen drei Stufen unterscheidbar. Die erste umfasst primitive Grundmerkmale, die sich über Bildern oder deren Regionen ableiten lassen. Beispiele sind Kantenverläufe, Formen, Farben und Texturen. Algorithmische Mittel für die Ableitung derartiger Merkmale aus Bildern stehen bereit. Eine zweite Stufe erfordert die strukturelle und logische Verknüpfung von Grundmerkmalen, um das Vorhandensein von Objekten (Tisch, Haus, Mensch, Tier u. ä.) abzuleiten und entsprechend zu beschreiben. Diese Stufe enthält bereits eine semantische Interpretation. Gelingt es in einer dritten Stufe, die Objekte im Sinne einer Szeneninterpretation weiter zu verknüpfen, erhalten wir eine Szenen- oder Ereignisbeschreibung. Heute existierende CBIR-Systeme und -Ansätze sind vor allem der ersten Stufe zuzuordnen. Zu den höheren Stufen klafft eine semantische Lücke (semantic gap), sie ist intensiver Forschungsgegenstand der Bildverarbeitung (Szenenanalyse) und der Künstlichen Intelligenz.

Auf den EVA-Konferenzen hat es immer wieder Beiträge zu CBIR gegeben [1 ... 8]. Die Erstellung von Grundmerkmalen war jeweils Basis, durch die jeweilige Einschränkung auf Bilder einer in sich definierten Klasse mit gemeinsamen Charakteristika (Steinmetzzeichen, Wappen, ...) konnten die Ansätze Erfolge präsentieren und in Richtung höherer Interpretationsstufen erweitert werden. Dies ist kompatibel mit der Aussage, dass leistungsfähige und treffsichere CBIR-Systeme für freiere Bildmengen noch auf sich warten lassen werden. Schränkt man den Grundcharakter der zu betrachtenden Bildmengen geeignet ein, liegen bereits praktikable Ergebnisse vor. Die drei Beiträge *Hermes*, *Köthe* und *Schmidt* folgen dieser Aussage. Sie setzen bei technischen Fragen der Bildverarbeitung zum sicheren Finden von Grundelemente (Striche, geometrische Figuren, Punktumgebungen) an, z. T. sehr exakt im Detail, und versuchen, Relationen zwischen diesen aufzubauen, um eine „höhere“ den Kontext berücksichtigende Beschreibung zu erhalten.

Literatur

- [1] R. Sablatnig, E. Zolda: Rechnergestützte Präklassifizierung von Portraitminiaturen. In: Konferenzband „EVA 1996 Berlin“.
- [2] A. Dammeyer, W. Jürgensen: Digitale intelligente Videoanalyse. In: Konferenzband „EVA 1998 Berlin“.
- [3] S. Kröner, A. Lattner: Automatic recognition of free hand drawings by Eugène Delacroix. In: Konferenzband „EVA 1998 Berlin“.
- [4] V. M. Kiiko, V. Matsello, H. Masuch, G. Stanke: A Fast Algorithm for Retrieval of Images in a Library of Masonmarks. In: Konferenzband „EVA 1998 Berlin“.
- [5] Universität Bremen, Technologie-Zentrum Informatik: Ausstellungspräsentation PictureFinder. In: Konferenzband „EVA 1998 Berlin“.
- [6] St. Müller, F. Wallhoff, St. Eickeler, G. Rigoll: Content-Based Retrieval of Digital Archives using Statistical Object Modeling Techniques. In: Konferenzband „EVA 1999 Berlin“.
- [7] Ch. Breiteneder, H. Eidenberger, M. Wasinger: JEmblazoner: Konstruktion von Wappenbildern aus textuellen Beschreibungen. In: Konferenzband „EVA 2002 Berlin“.
- [8] E. Schallehn, I. Schmitt, N. Schulz: Visual Retrieval for Searching in a LostArt Metasearch Engine System. In: Konferenzband „EVA 2002 Berlin“.
- [9] http://lmb.informatik.uni-freiburg.de/lectures/bild_Seminar/index.de.html
- [10] <http://en.wikipedia.org/wiki/CBIR>