

# Zoombasiertes Interaktionskonzept für die Suche in großen, heterogenen Bilddatenbanken

Fredrik Gundelsweiler, Harald Reiterer

Arbeitsgruppe Mensch-Computer Interaktion, Universität Konstanz

## **Zusammenfassung**

In diesem Beitrag stellen wir die Konzeption, Gestaltung und Realisierung eines interaktiven Systems zur Bildersuche vor. Unternehmen verfügen oft über umfangreiche heterogene Bildsammlungen, die sehr unterschiedlichen Zielgruppen innerhalb und außerhalb des Unternehmens zur Verfügung gestellt werden müssen. Die Größe und Heterogenität dieser Bildsammlungen bedingt, dass es oft schwierig ist relevante Bilder nach spezifischen Kriterien zu finden. Wir haben ein webbasiertes System zur Bildersuche am Beispiel einer Bildersammlung von ca. 13.000 Fahrzeugbildern entwickelt. Damit zeigen wir eine innovative Lösung auf, die über ein zoombasiertes Interaktionskonzept mit umfangreichen Such-, Kategorisierung- und Filtermöglichkeiten verfügt und dieses mit einer Komponente zur Ähnlichkeitsuche kombiniert. Erste informale Evaluationen werden beschrieben. Sie zeigen, dass wir dadurch eine neue Qualität bei der Suche in großen heterogenen Bildersammlungen ermöglichen. Diese basiert vor allem auf einer visuellen und kategorialen Darstellung der Inhalte bzw. der Suchergebnisse. Ein fließender Übergang zwischen zielgerichteter Suche und explorativem Browsen wird dadurch unterstützt.

## 1 Einleitung

Bilder und Zeichnungen werden zunehmend digital angefertigt und über das Internet verfügbar gemacht. Die Verwaltung der Bilder wird mit ansteigenden Datenmengen immer schwieriger und stellt eine interessante Forschungsfrage für die Mensch-Computer Interaktion dar. Diese beinhaltet vor allem auch die Suche nach einem geeigneten Interaktionskonzept. Die Hauptproblematik bei der Suche in großen Bilddatenmengen ist die Suche nach Bildern von denen der Benutzer nur eine vage Vorstellung hat. Auch die Präsentation der Suchergebnisse spielt eine wichtige Rolle, da einerseits sehr viele oder aber gar keine Treffer aus den Anfragen resultieren können. Eine weitere Herausforderung stellt die meist unzureichende Verschlagwortung der Bilder dar. Anwender machen Tippfehler oder verwenden unterschiedliche Begriffe, um die Bildinhalte zu beschreiben. Die Bildung von Metadaten Gruppen erleich-

tert zwar die Einschränkung, aber die Ergebnismenge verringert sich oft nicht ausreichend. Wie kann also der Benutzer einerseits bei der Suche nach bekannten sowie unbekanntem Bildern und andererseits bei der visuellen Exploration auf einer sehr großen Menge an Bildern optimal unterstützt werden? In diesem Beitrag stellen wir unseren Lösungsansatz für diese Fragestellungen vor und illustrieren die Konzeption und Gestaltung einer interaktiven Webanwendung. Dazu verwenden wir beispielhaft Bild- und Metadaten von Fahrzeugen. Die konzipierte Bildersuche muss dazu für mehr als 600.000 heterogene Bilder skalieren, darunter z.B. Strichzeichnungen, Photographien, 3D-Bilder, Explosionsdarstellungen und technische Zeichnungen. Die Metadaten enthalten IDs, Verwendungsbereiche (Diagnose, Literatur, Werkstatt, ...), Umsetzungsarten (Strichbild, Fotografie, technische Grafik, 3D Bild, ...), Bildformat, Farbtiefe, Erstellungsdatum und weitere Informationen. In Abschnitt 2 beschreiben wir zunächst die Grundlagen. Dazu stellen wir vorhandene Systeme zur Bildersuche vor, untersuchen die Vorgehensweisen der Benutzer bei der Suche und leiten Anforderungen und Designprinzipien ab. In Abschnitt 3 erläutern wir das von uns entwickelte Bildersuchsystem und stellen auch erste Ergebnisse der Untersuchungen zu Gebrauchstauglichkeit und Aufgabenangemessenheit unseres Prototyps vor. Am Ende fassen wir die gewonnenen Erkenntnisse zusammen und geben einen Ausblick auf zukünftige Arbeiten.

## 2 Grundlagen und Anforderungen

Es existieren bereits eine Reihe interessanter Anwendungen zu Bildersuche. Grundlegende Konzepte dieser Anwendungen umfassen die Organisation der Bilder nach Kategorien, die Erstellung von Alben, die Verschlagwortung nach verschiedenen Kriterien und die Darstellung der Bilder. Aus der Menge an Anwendungen haben sich vier durch besonders interessante Konzepte herausgehoben. Dies sind die Desktopanwendungen Picasa (<http://picasa.google.com/>) und Photomesa (Bederson 2001), sowie die beiden Webanwendungen retrievr (<http://labs.systemone.at/retrievr/>) und xcavator (<http://www.xcavator.net/>). Die genauere Untersuchung der Anwendungen ergibt neben einigen Schwachpunkten insbesondere die in Tabelle 1 aufgeführten Besonderheiten.

Anwendung	Besonderheit
Picasa	Organisation und Browsen der Bilder nach Zeit mit interessanten Interaktionsmöglichkeiten (Zeitschieberegler, Verschieben des Zeitstrahls, direkte Selektion von Bildern, Visualisierung der Anzahl)
Photomesa	Zoomkonzept zur Exploration der Bilder von der Übersicht bis zum Einzelbild, Hinzufügen von Metadaten zu den Bildern (Personen, Orte, ...)
Retrievr	Webanwendung mit inhaltsbasierter Skizzen-Ähnlichkeitssuche auf der Flickr Bilddatenbank
Xcavator	Webanwendung mit inhaltsbasierter Ähnlichkeitssuche im Verbund mit textueller Anfragemöglichkeit über mehrere Online-Datenbanken mit Drag'n'Drop und Favoritenbereich

Tabelle 1: Besonderheiten der untersuchten Anwendungen zur Bildersuche und -verwaltung

Die Nachteile der Anwendungen sind, dass die Suche nach einem ganz bestimmten Bild nur dann einfach ist, wenn der Anwender seine Bilder in einer logischen Ordnerstruktur verwaltet, das ungefähre Datum des Bildes kennt oder alle Bilder mit Metadaten versieht. Sind zu

viele Bilder in den Ordnern enthalten, gestaltet sich auch dann die Suche als schwierig. Eine Verbindung der Suche mit Browsingkomponenten könnte die Einschränkung der Treffermenge erleichtern. Die zoombare Ergebnisvisualisierung von Photomesa stellt zunächst alle Bilder wie auf einer Art Pinnwand als kleine Icons dar. Dann hat der Benutzer die Möglichkeit in Bereiche hinein zu zoomen, bis das gewünschte Bild in Vollansicht auf dem Bildschirm zu sehen ist. Es ist zwar sehr intuitiv die Bilder visuell zu explorieren aber durch die langen Ladezeiten bei einer großen Anzahl von Bildern (über 1.000) wird das Zoomkonzept unbenutzbar. Eine Kategorisierung ist daher unabdingbar. Ein weiteres interessantes Feature ist die Anzeige von repräsentativen Bildern für Suchergebnisse oder Ordner. Dabei werden leider oft Bilder angezeigt, die nicht repräsentativ für die ausgewählten Ordner sind.

## 2.1 Theorie zu Suche und Exploration

Zur Suche und zum *Information Retrieval Prozess* gibt es zahlreiche Theorien und Modelle aus der Informationswissenschaft, der Kognitionspsychologie und der Mensch-Computer Interaktion. Beispiele sind das klassische Modell des Information Retrieval (Robertson 1977), das Berrypicking Modell (Bates 2002), der Prozess der Informationssuche (Marchionini 1997), das 5-Phase Framework (Shneiderman & Plaisant 2004) und der Information Search Process (Kuhlthau 2004). In allen Modellen wird festgestellt, dass die Suche in der Regel ein iterativer Prozess ist. Das bedeutet, dass die Suchanfrage nach und nach durch die Bewertung der erhaltenen Ergebnisse verfeinert wird, bis die relevante Information gefunden ist. Eine weitere Gemeinsamkeit ist die Unterscheidung der direkten und der explorativen oder browsenden Suche. Die direkte oder zielgerichtete Suche wird durchgeführt, wenn der Benutzer eine konkrete Vorstellung vom Ergebnis hat. Die explorative Erkundung wird vor allem zum Verschaffen eines Überblicks über die Inhalte, zur Durchsicht einer größeren Treffermenge oder zur Inspiration für eine neue Suchanfrage eingesetzt. Zu jeder Zeit während der Suche kann ein Wechsel zwischen der zielgerichteten und der explorativen Suche stattfinden. Dieser Wechsel hängt zum einen von den Vorkenntnissen des Benutzers, andererseits von Informationen, die er auf dem Weg zum Suchergebnis entdeckt, aber auch von der Auffassungsgabe und Wahrnehmung ab. Deshalb sollte der Benutzer zu jeder Zeit Einfluss auf den Suchprozess nehmen können. Die Ergebnisdarstellung sollte demnach auch einfach erweiterbar und explorierbar sein bzw. reduziert werden können. Insbesondere in Kombination mit Zoomable User Interfaces (ZUIs) unterstützen Visualisierungen den Menschen bei der Suche und Exploration im virtuellen Informationsraum. Der Mensch denkt visuell und versucht sich abstrakte Information räumlich einzuprägen, um sich besser daran erinnern zu können (Miller 1968; Ware 2004). Diese Eigenschaft kann durch Abbildungen, Positionen im Raum, Sortierungen oder andere visuelle Merkmale von Objekten unterstützt werden. Die Navigationsmöglichkeiten Zooming (Skalierung) und Panning (Verschiebung des sichtbaren Bereichs) der ZUIs entsprechen der menschlichen Bewegung im Raum und vereinfachen deshalb auch die Navigation im virtuellen Raum. Die Organisation der Daten im Raum und weitere Vorteile sind in (Bolt 1984) beschrieben. Kontinuierliche bzw. animierte Transitionen erleichtern zudem die Orientierung des Benutzers im virtuellen Informationsraum (Zhang 2005). Zusammengefasst lassen diese Vorteile vermuten, dass ZUIs gegenüber diskreten Ergebnisdarstellungen besser zur Darstellung und Navigation virtueller Informationsräume geeignet sind. Bolt erwähnt bereits skalierbare Benutzungsschnittstellen,

die entweder (1) aus einer Informationslandschaft in einer Ebene, (2) aus einer Informationslandschaft mit geschachtelten Bereichen auf einer Ebene oder (3) aus einer Informationslandschaft mit weiteren geschachtelten Informationslandschaften über mehrere Ebenen bestehen. Bei der Realisierung eines ZUIs sollten diese drei Arten auf Vor- und Nachteile im Zusammenhang mit dem Nutzungskontext geprüft werden. Eine weitere relevante Komponente zur Strukturierung von Informationen ist die Verwendung von Kategorien. Sind diese dem Benutzer geläufig, so können sie die Suche, Exploration und Orientierung wesentlich erleichtern. Dadurch beschleunigen Kategorien den Suchprozess, wobei sie zur Übersicht und Sicherheit des Benutzers beitragen (Yee et al. 2003). Andererseits können sie aber auch zu Unsicherheit beitragen, wenn sie dem Benutzer unbekannt sind und deshalb keine Informationen zu den enthaltenen Daten liefern (Lambert & Shanks 1997).

## 2.2 Anforderungen und Nutzungskontext

In der Anforderungsermittlung wurden das bisherige System zur Bildersuche, die Benutzertypen sowie deren Aufgaben untersucht, um problemadäquate Konzepte für den Nutzungskontext der Anwendung erstellen zu können. Alter, Geschlecht, Computerkenntnisse und Vorkenntnisse bzgl. der Fahrzeugdomäne unterscheiden sich sehr stark. Als grundlegende Benutzertypen unterscheiden wir hier den Beginner (nur grundlegende Computerkenntnisse, wenig Domänenwissen) und den Experten (fortgeschrittene Computerkenntnisse, viel Domänenwissen), die beide adäquat vom System unterstützt werden müssen. Das bisherige System funktioniert mittels einer rein formularbasierten Suche. Alle Metadaten sind in der Suchmaske verfügbar und können vom Benutzer angegeben werden. Nach der Eingabe der Metadaten als Filterkriterien, muss der Benutzer zunächst die Anzahl an Bildern ermitteln, da das System momentan nur 600 Bilder auf einmal anzeigen kann. Werden zu viele Treffer gefunden, muss der Benutzer weitere Filterkriterien angeben. Die Ergebnisanzeige ist eine Bildmatrix ohne Sortiermöglichkeiten, die zunächst nur das Bild und dessen ID anzeigt und je nach Treffermenge mehrere Seiten enthält. Ein Doppelklick auf ein Bild führt zur erweiterten Ergebnisanzeige, die das Bild in Vollansicht sowie alle Metadaten anzeigt. Die rein formularbasierte Suche versteckt die Daten vor den Benutzern, die neben der Hauptaufgabe der Suche noch weitere Aufgaben haben können. Bei einem Formular müssen Benutzer wissen nach was sie suchen, da sie genaue Angaben zu den Ergebnissen machen müssen. Was aber wenn es bessere Bilder gibt, die nur nicht gefunden werden, weil die Benutzer nicht wissen wie man nach diesen sucht? Die Untersuchung des bisherigen Suchsystems offenbart viele weitere Probleme wie das Umschalten zwischen Anfrageformular (Angabe der Filterkriterien und ermitteln der Trefferanzahl) und Ergebnisvisualisierung bei jeder neuen Anfrage, die hier nicht im Detail aufgeführt werden. Die typischen Benutzeraufgaben sind die Suche nach einem oder mehreren Bildern mit bestimmten IDs, die Suche nach Bildern mit bestimmten Metadatenkombinationen, die Suche ähnlicher Bilder zu einem vorgegebenen Bild, das Abspeichern von Bildern und die visuelle Exploration der Treffermenge(n).

## 2.3 Gestaltungs- und Designprinzipien

Anhand des Nutzungskontextes, der Untersuchung des bisherigen Systems, der Analyse anderer Bildsuchsysteme und der Theorie (Abschnitt 2.1) lassen sich grundlegende Anforder-

rungen an eine interaktive Suchanwendung ermitteln. Das System muss eine schnelle und einfache Suchkomponente enthalten, mit deren Hilfe Benutzer in der Lage sind sofort Bilder anzuzeigen deren IDs bekannt sind oder die nach anderen Metadaten eindeutig beschrieben werden können. Des Weiteren sollte eine Art persönlicher Bereich eingeführt werden, der es ermöglicht Bilder zu speichern, damit die Benutzer nicht gezwungen werden sich Nummern zu merken oder extern eine Liste der wichtigsten Bilder zu führen. Zur Einschränkung des Suchraums ist es notwendig verschiedene Filter für Metadatengruppen (z.B. Umsetzungsart) und zu speziellen Metadaten (z.B. Bildformat, Farbtiefe, Erstellungsdatum) in die Anwendung zu integrieren. Einerseits muss eine Direktsuche und andererseits eine visuelle Suche sowie Exploration in der Bildkollektion möglich sein. Eine weitere Möglichkeit die Benutzer bei der Suche nach unbekanntem Bildern zu unterstützen ist die Suche nach Bildern über inhaltsbasierte Ähnlichkeit (Content-based Image Retrieval - CbIR). Dazu sollte es möglich sein ein Bild von der Festplatte oder aus der Bildkollektion selbst als Anfrage abzusenden und ähnliche Bilder als Ergebnis zu erhalten. Aus unseren Untersuchungen haben wir die folgenden Designprinzipien für die interaktive Bildersuche abgeleitet. Der Schwerpunkt liegt auf den Such- und Filtertechniken, sowie auf der Ergebnispräsentation.

Designgegenstand	Beschreibung des Designprinzips
Direkte Suche	Suche nach ID und Schlagworten mit Autovervollständigung.
Explorative Suche	Verschiedene Bildkategorien als Einstieg anbieten, in Kombination mit dazu passenden Visualisierungen.
User Interface, Visualisierung und Ergebnisdarstellung	Wechsel zwischen direkter Suche und Exploration soll unterstützt werden. Sofortige Anzeige erster Ergebnisse mit Angabe der Anzahl an Bildern. Dadurch soll dem Benutzer das Gefühl von Kontrolle und Übersicht vermittelt werden.
Filter	Logische Gruppen für die Filter aus den Bildmetadaten bilden. Hierarchische Gliederung des Informationsraums über verschiedene hierarchische der Metadaten (z.B. Verwendungs-, Umsetzungs- und Darstellungsart). Mittels Dynamic Queries werden allgemeinere Metadaten wie Farbtiefe und Erstellungsdatum gefiltert.
Kombination von Suche, Filter und Ergebnisanzeige	Textsuche, Filter, Hierarchien und Ergebnispräsentation müssen so kombiniert werden, dass ein Wechsel sehr einfach innerhalb der Benutzungsschnittstelle möglich ist.
Ein- und Ausblenden von Filtern	Die Filter werden nach Möglichkeit ein- und ausgeblendet, damit der Benutzer mehr Platz zur Ergebnisanzeige hat. Ausblenden von Interaktionselementen, die für die Hauptaufgaben irrelevant sind. Einblenden nur nach Bedarf.

Tabelle 2: Übersicht über die aufgestellten Designprinzipien für interaktive Suchanwendungen

Es existieren eine Reihe von Richtlinien für Webanwendungen und Applikationen, die weitere Empfehlungen zur Gestaltung der Suche machen. Dazu gehören z.B. die Richtlinien zur Gestaltung von Suchanwendungen nach (Rose 2005) und die Richtlinien der Nielsen Norman Group (Molich et al. 2000). Darin wird empfohlen, dass verschiedene User Interfaces (oder Interaktionsformen), die zu den verschiedenen Suchzielen passen, zur Verfügung gestellt werden sollten. Das Interface soll außerdem die Auswahl des passenden Suchkontexts vereinfachen und den iterativen Prozess der Suche unterstützen. Zur detaillierteren Gestaltung der Benutzerschnittstelle gibt es allgemeine Richtlinien wie die DIN EN ISO-Norm 9241 (Teil 110 und 11), sowie ISO 14915 (Teil 1, 2 und 3) und ISO/IEC 18035, die bei der Gestaltung unseres Prototyps bisher nur teilweise beachtet wurden.

### 3 Zoombasierte Bildersuche

Zur praktischen Umsetzung wird Adobe Flex 3 mit amfphp (<http://www.amf-php.org>) zur Kommunikation zwischen Webanwendung und einer DB2-Datenbank verwendet. Zudem ist die Installation eines Frameworks zur inhaltsbasierten Ähnlichkeitssuche notwendig. Die Suche über eine einfache Texteingabe ist kombiniert mit verschiedenen Metadatenfiltern (Kachelkategorisierung), einer Ähnlichkeitssuche (mittels Datei laden oder Drag'n'Drop) und der visuellen Präsentation der Ergebnisse in skalierbaren Informationskacheln. Der Informationsraum kann über die Konfiguration der Kacheln in Hierarchieebenen zerlegt werden, die dann mittels Zooming und Panning exploriert werden können.

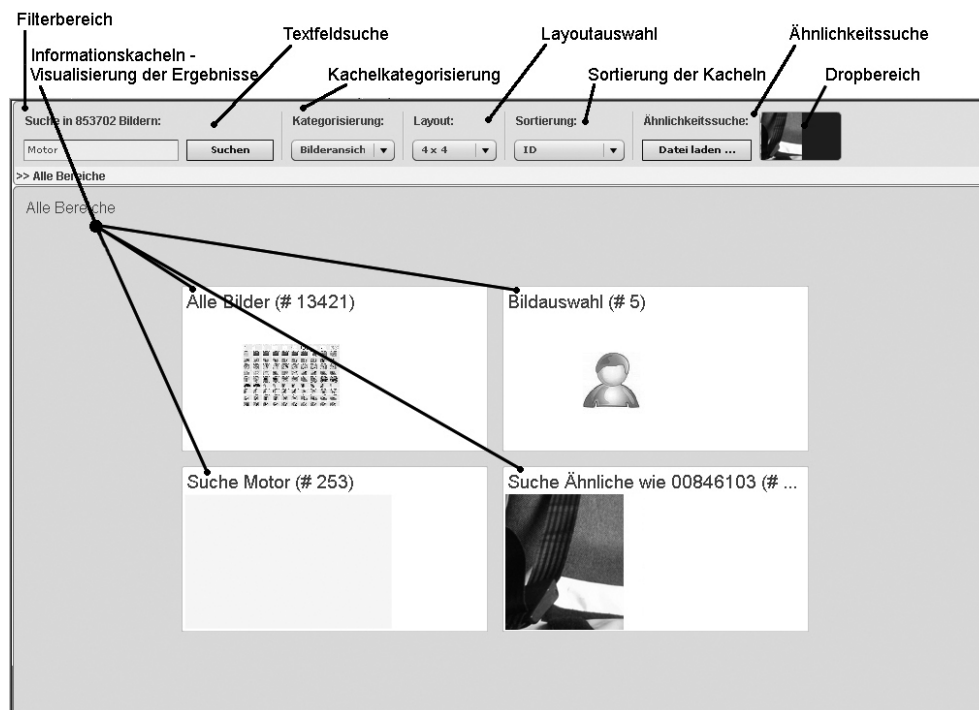


Abbildung 1: Benutzungsschnittstelle des High-Fidelity Prototyps zur Bildersuche

Neben diesem stehen weitere Konzepte zur Diskussion, wie eine Suche direkt am abstrakten 3D-Fahrzeugmodell, eine Suche anhand einer Metadaten-Baumdarstellung und eine Pixelansicht der kompletten Bilddatenbank. Das zoombasierte Konzept mit Informationskacheln wurde von uns in einer Expertenuntersuchung favorisiert, da die anderen Konzepte später leicht als zusätzliche Kacheln integriert werden können. Ein weiterer Grund, der auch für die zoombare Ergebnisvisualisierung spricht ist die Annahme, dass sich zoombare Browser besser für die Exploration einer großen Anzahl Bilder eignen (Combs und Bederson 1999).

Die Benutzungsschnittstelle (siehe Abb. 1) besteht aus zwei Hauptteilen. Der obere schmale Bereich wird zur Suche, Filterung und Konfiguration der Informationskacheln verwendet. Der untere Bereich zeigt die Ergebnisse in Form von Kacheln kategorisiert an. Ein Klick auf eine der Informationskacheln führt zum Zoom und damit zur Skalierung und Verschiebung des sichtbaren Bereichs auf diese. Dies ermöglicht dem Benutzer bestimmte Kacheln zu fokussieren und semantisch mit weiteren Informationen anzureichern. Zunächst werden in den Kacheln Vorschaubilder angezeigt und erst bei einer Skalierung auf Vollansicht werden die enthaltenen Bilder geladen. Die Bilder können wiederum per Klick skaliert werden, was zur Anzeige des vergrößerten Bildes und dessen Metadaten führt (Abb. 2).

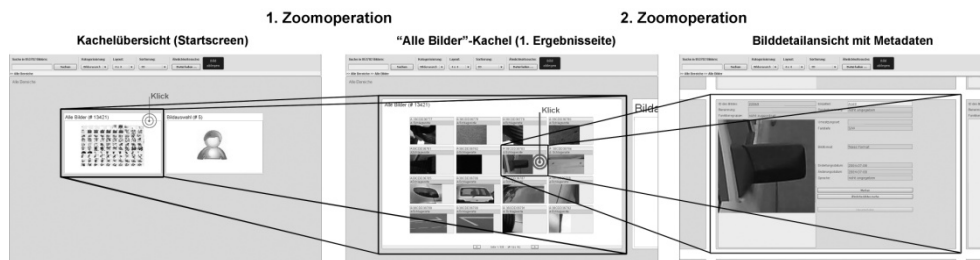


Abbildung 2: Prototyp zur Bildersuche, Zoomoperationen durch Klick auf Informationskacheln

Durch die semantische Skalierung von Kategorien und Bildern kann einerseits ein Überblick über eine große Anzahl Bilder gegeben werden und andererseits leicht zu den Details eines bestimmten Bild gezoomt werden. Durch eine Verschachtelung mehrerer Informationskacheln ist es möglich den Informationsraum nach den Benutzerbedürfnissen einzuschränken. Zunächst kann die Kachel „Alle Bilder“ herangezogen werden. Jetzt kann der Benutzer eine weitere Kategorisierung (z.B. Verwendungsart) wählen (siehe Abb.1 oben). Daraufhin erstellt das System in der momentan fokussierten Kachel die entsprechenden Kacheln und zeigt diese inklusive der Anzahl enthaltener Bilder an. Anstatt der gesamten Bildmenge sieht der Benutzer nun also z.B. nur noch 6 neue Kacheln zu den verschiedenen Verwendungsarten. Er kann sich für eine entscheiden und bekommt nach einem erneuten Zoom wieder die Bilder, allerdings eingeschränkt auf die selektierte Verwendungsart, angezeigt. Diese können erneut durch Auswahl einer Kategorisierung in Kacheln aufgeteilt werden. Dem Benutzer ist es also möglich die filternden Ebenen selbst zu wählen und in jeder Hierarchieebene neue Kacheln durch Suche oder Filterung zu erstellen. Die Ergebnisse einer neuen Suche werden auf der obersten Ebene, neben der „Alle Bilder“-Kachel und der Kachel mit gespeicherten Bildern (Bildauswahl), als zusätzliche Kachel angezeigt. Dazu berechnet das System ein neues Layout für die Anzahl an Kacheln und teilt diesen ihre Koordinaten und die Größe zu. Zur einfacheren Erledigung der Benutzeraufgaben bietet der Suchfilter am oberen Bildschirmrand eine Direktsuche an, mit deren Hilfe nach IDs oder Schlagworten gesucht wird. Das Browsen des Informationsraums anhand von Kategorien ermöglicht den Benutzern unterschiedliche Einstiege und unterstützt damit verschiedene Vorgehensweisen bei der Suche. Relevante Suchergebnisse können über einen „Merken“-Knopf in der Bilddetailansicht der persönlichen Bildauswahl hinzugefügt werden. Über die zu jeder Zeit sichtbare Breadcrumb-Navigation erkennt der Benutzer, an welcher Stelle und wie tief er in die Kachelhierarchie

navigiert hat. Die Ähnlichkeitssuche wird direkt per Knopfdruck oder per Drag'n'Drop von Bildern in den Bereich der Ähnlichkeitssuche gestartet. Dies führt ebenfalls zur Erstellung einer neuen Kachel mit den ähnlichen Ergebnisbildern. Es können zudem Dateien von der Festplatte als Anfragebilder in die Anwendung geladen werden. Zum besseren Verständnis des Konzepts ist eine detailliertere Beschreibung und ein kurzes Video auf unserer Webseite unter <http://hci.uni-konstanz.de/bildersuche> einsehbar.

### 3.3 Ähnlichkeitssuche

Sowohl in der Forschung als auch in der Praxis existieren sehr viele unterschiedliche Systeme zur inhaltsbasierten Ähnlichkeitssuche von Bildern und multimedialen Daten (Veltkamp & Tanase 2000). Ein weiteres Framework zur inhaltsbasierten Ähnlichkeitssuche ist GIFT - **G**nu **I**mage **F**inding **T**ool (Müller et al. 1999), das in seiner ersten Version als VIPER (Squire et al. 1998) herausgegeben wurde. Wir verwenden GIFT, da es im Gegensatz zu anderen Frameworks kostenfrei und relativ einfach zu installieren ist. Zudem hat es den Vorteil, dass bereits ein Webservice integriert ist, der von unserer Webanwendung per XML angesteuert wird. Aus diesen Gründen indexierten wir unsere Bildkollektion von 13.000 Bildern mit GIFT, obwohl in (Rummukainen et al. 2003) Performancenachteile im Vergleich zu anderen System aufgeführt sind. Dabei berechnet das System bis zu 150 Featurevektoren aus jedem Bild und speichert diese in einer Ähnlichkeitsmatrix. Die Ergebnisse der Ähnlichkeitsanfragen sind bei Farbbildern recht gut, bei schwarz-weißen Bildern werden aber oft Bilder für ähnlich befunden, die nicht mehr viel mit dem Anfragebild gemeinsam haben. Die optimale Vorgehensweise für unser sehr heterogenes Bildmaterial zu finden ist schwierig, da sowohl verschiedene Algorithmen für die Indexierung und die Ähnlichkeitsmaße zur Verfügung stehen, als auch die Bilder in unterschiedliche Kollektionen aufgeteilt werden können. Für eine erste Einbindung belassen wir das GIFT Framework in seinem Ausgangszustand.

### 3.4 Qualitative Evaluation

Zur Diskussion, Untersuchung bzw. einer ersten formalen Evaluation und zur Ermittlung weiterer Anforderungen ist ein Workshop abgehalten worden. Sechs Benutzer, vier Domänenexperten und drei weitere Personen (zur Dokumentation und Moderation des Workshops) haben daran teilgenommen. Zunächst ist eine kurze Vorstellungsrunde abgehalten worden, um in das Thema einzuführen. Dann wurde der Prototyp vorgeführt und die einzelnen Funktionen diskutiert. Die Diskussion ist dokumentiert worden, um insbesondere die Kommentare der Benutzer festzuhalten. Nach der Präsentation des Prototyps hatten die Benutzer die Gelegenheit diesen selbst auszuprobieren. Dabei wurden sie gebeten laut zu denken und sich zu jeder ihrer Aktionen zu äußern. Als Resultat wurden einige Fehler in der Programmierung entdeckt und sehr viele Verbesserungsvorschläge von den Benutzern und Domänenexperten gemacht. Mit diesen Ergebnissen ist es nun möglich das Konzept und Design im Hinblick auf die Suchstrategien, Aufgaben und Vorgehensweisen der Benutzer in einer weiteren Iteration zu verbessern.

Eine erste verbesserte Version des Prototyps wurde bereits erstellt. Dabei wurde das Feedback für die Benutzer erweitert. Das Überfahren einer Kachel mit der Maus hebt diese nun

durch Farbänderung der Kachelränder hervor und zeigt mittels einer Lupe am Mauszeiger an, ob die Kachel bei einem Klick vergrößert, verkleinert oder lediglich verschoben wird. Eine weitere grundlegende Änderung ist die Integration des Filterbereichs, der sich zuerst am oberen Bildschirmrand befand, in die Kacheln. Ein globaler Filterbereich existiert nun nicht mehr, sondern die Filterung wird für jede Kachel vorgenommen. Ein Auszug weiterer wichtiger Erkenntnisse aus dem Workshop ist im Folgenden aufgeführt. Bei der Diskussion mit den Benutzern stellte sich im Bezug auf den Nutzungskontext heraus, dass das Sucheingabefeld über alle Metadaten eingesetzt werden können sollte. So wäre es möglich direkt bestimmte Metadatenwerte in Kombination mit IDs und Schlagworten einzugeben. Das Sucheingabefeld kann weiter verbessert werden, indem ein Thesaurus integriert wird der eine automatische Vervollständigung von Suchbegriffen anbietet. Die Benutzer verlangen außerdem eine Mehrfacheingabe von IDs, um mehrere Bilder nach ID auf einmal anzuzeigen. Weiterhin wurde erkannt, dass die unterschiedlichen, domänenspezifischen und aufgabenbezogenen Suchstrategien der Benutzer von großer Bedeutung sind. Das System muss also alle unterschiedlichen Suchstrategien unterstützen und auch eine Kombination dieser ermöglichen. Die drei generellen Suchstrategien sind die Suche per Texteingabe (IDs oder Schlagworte), die Suche durch Einschränkung der Bilder mittels Kategorisierung des Informationsraums und das iterative „Herantasten“ an ein Bild nach Ähnlichkeit. Ein weiterer wichtiger Punkt ist die Personalisierung, die es ermöglicht die Benutzungsschnittstelle anzupassen und die Bildmenge schon zu Beginn auf eine relevante Untermenge einzuschränken. Die Ähnlichkeitssuche ist als besonders hilfreich empfunden worden und kann in Zukunft mittels Themen- oder Schlagwortlandkarten eine weitere kombinierbare, alternative Übersicht über den Datenraum zur Verfügung stellen.

### 3.5 Zusammenfassung und Ausblick

In diesem Beitrag wird aufgezeigt wie Theorie und Praxis bei der Konzeption, Gestaltung und Realisierung einer Webanwendung zur Suche und Exploration von heterogenen Bildern ineinander greifen können. Es wird eine innovative Lösung vorgestellt, die über ein zoombasiertes Interaktionskonzept mit umfangreichen Such-, Kategorisierung- und Filtermöglichkeiten verfügt und dieses mit einer Komponente zur Ähnlichkeitssuche kombiniert. Erste informale Evaluationen werden beschrieben. Sie zeigen, dass wir durch das neue Konzept und die bessere Unterstützung der Aufgaben und Suchstrategien im Vergleich zum bisherigen System eine neue Qualität bei der Suche in großen, heterogenen Bildersammlungen erreichen können. Die Verbesserung basiert vor allem auf einer visuellen und kategorialen Darstellung der Suchergebnisse. Ein fließender Übergang zwischen zielgerichteter Suche und explorativem Browsen wird dadurch unterstützt. Zukünftig muss zur besseren Kopplung der Metadaten mit der inhaltsbasierten Ähnlichkeitssuche eine eigene Komponente zur Ähnlichkeitssuche entwickelt werden, die eine Verbesserung der Algorithmen und eine Optimierung der Feature-Berechnung für unsere Bildkollektion zur Verfügung stellt.

#### **Literaturverzeichnis**

- Ahlberg C., Shneiderman B. (1994). Visual information seeking: tight coupling of dynamic query filters with starfield displays. Proc. of the SIGCHI conference on Human factors in computing systems, ACM Press, 1994, 313–317.

- Bates, Marcia J. (2002). Toward an Integrated Model of Information Seeking and Searching. The Fourth International Conference on Information Needs, Seeking and Use in Different Contexts, Lisbon, Portugal, September 11-13, 2002.
- Bederson, B. B. (2001). PhotoMesa: A Zoomable Image Browser Using Quantum Treemaps and Bubblemaps. *UIST 2001, ACM Symp. on UI Software and Technology, CHI Letters*, 3(2), pp. 71-80.
- Bolt, Richard A. (1984). *The Human Interface: Where People and Computers Meet*. Belmont, California: Lifetime Learning Publications, 1984. Paperback, ISBN 0534033873.
- Combs, T. T. A. und Bederson, B. B. (1999). Does zooming improve image browsing? In *Proceedings 10th International Conference on Digital Libraries*.
- Kuhlthau, C. C. (2004). *Seeking Meaning: A Process Approach To Library and Information Services*. 2<sup>nd</sup>. Ed. Westport, CT: Libraries Unlimited, 2004.
- Lambert K. & Shanks, D. 1997. *Knowledge, Concepts and Categories (Studies in Cognition)*. Cambridge, Massachusetts: Psychology Press Ltd; Auflage: MIT Press Ed (Oktober 1997).
- Marchionini, Gary (1997). *Information Seeking in Electronic Environments*. Cambridge Press 1997.
- Molich, Rolf; Snyder, Carolyn; Farrell, Susan; Nielsen, Jakob (2000): *E-Commerce User Experience*. Nielsen Norman Group, 2000.
- Müller, H., David McG. Squire, W.M., Pun., T. (1999). Efficient access methods for content-based image retrieval with inverted files. In: *Proceedings of Multimedia Storage and Archiving Systems IV (VV02)*, Boston, MA, USA, 1999.
- Robertson, S. E. (1977). 'Theories and Models in IR', *Journal of Documentation*, Ausgabe 33, Nr. 2, 1977, Seiten 126-148.
- Rose, D. E. (2005). Reconciling Information-Seeking Behavior with Search User Interfaces for the Web. *Journal of the American Society of Information Science and Technology*, 2005.
- Rummukainen, M.; Laaksonen, J. and M. Koskela (2003). An efficiency comparison of two content-based image retrieval systems, GIFT and PicSOM. In *International Conference on Image and Video Retrieval*, pages 478-487. *Lecture Notes in Computer Science*, vol. 2728, Springer, 2003.
- Shneiderman, Ben; Plaisant, Catherine (2004). *Designing the User Interface: Strategies for Effective Human-Computer Interaction (4th Edition)*. Addison Wesley, 4th edition, March 31, 2004.
- Squire, D.M., Müller, W., Müller, H., Pun, T. (1998): Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback. Technical Report 98.04, Computer Vision Group, Computing Centre, University of Geneva, rue Gnral Dufour, 24, CH-1211 Genève, Switzerland, 1998.
- Veltkamp, R.C.; Tanase, M. (2000). *Content-Based Image Retrieval Systems: A Survey*, Department of Computing Science, Utrecht University, working material, 2000.
- Yee, K.-P., K. Swearingen, K. Li, and M. Hearst (2003). Faceted metadata for image search and browsing. In *CHI '03: Proceedings of the conference on Human factors in computing systems*, pp. 401-408. ACM Press.
- Zhang, X. (2005). Space-scale animation: enhancing cross-scale understanding of multiscale structures in multiple views. In *Coordinated and Multiple Views in Exploratory Visualization, 2005. (CMV 2005). Proceedings. Third International Conference on*, pages 109-120.