

Ansichten eines Fotoalbums zu Multimedia Information Retrieval

Susanne Boll, Department für Informatik, Universität Oldenburg
Philipp Sandhaus, OFFIS Institut für Informatik, Oldenburg
Ansgar Scherp, OFFIS Institut für Informatik, Oldenburg
Sabine Thieme, CeWe Color AG, Oldenburg

Kurzfassung

Nach einer Reise oder einer Feier möchte man oft aus den aufgenommenen Erlebnissen die schönsten Momente heraussuchen und die Bilder zum Erlebnis in einem Album einkleben. Allerdings sind die richtige Auswahl der Bilder und deren schönes Arrangement in einem Album bisher sehr zeitaufwändig. Auch in Zeiten digitaler Fotografie fehlt immer noch eine Unterstützung, aus der Vielzahl der vorliegenden Bilder die schönsten Erinnerungen auszuwählen und in einem Album festzuhalten.

Was grundlegend fehlt ist eine automatische Annotation der Fotos, die es den Nutzern erlaubt, ihre Bilder anhand des Erlebten und Erinnerten zu finden, zu betrachten und in verschiedenen Anwendungen wie der Erstellung eines Fotoalbums zu nutzen. Die automatische Annotation persönlicher Fotos stellt nach wie vor eine besondere Herausforderung für das Multimedia Information Retrieval (MIR) dar.

In diesem Artikel erarbeiten wir Anforderungen an das MIR, die sich aus der Sicht der Erstellung eines Fotoalbums mit persönlichen Fotos ergeben und präsentieren Lösungsansätze mit denen MIR hier einen Beitrag leisten kann. Für die Medienanalyse und das Retrieval spielen dabei inhaltsbasierte Ansätze ebenso eine Rolle wie die Ausnutzung von Kontextinformationen und Nutzerverhalten. Wir stellen in diesem Artikel unsere Ergebnisse im Bereich kontextgetriebener automatischer Annotation von Fotos aus Sicht der automatischen Erstellung von Fotoalben vor. Basierend auf extrahierten und abgeleiteten Metadaten werden mit geeigneten Auswahlverfahren Bilder für ein Album selektiert und passend in Arrangement und Stil für ein Fotoalbum zusammengestellt. Dabei nehmen wir in diesem Artikel eine bewußt anwendungsorientierte Sicht auf das Gebiet MIR ein. Die Ergebnisse wurden mit unserem Projektpartner CeWe Color, Europas größtem Fotodienstleister in eine Fotobuch-Software überführt, die im Juni 2006 auf den Markt gekommen ist.

1 Hintergrund und Motivation

Die Erstellung eines Fotoalbums wird meist nach einem schönen Urlaub oder einem besonderen Fest in Angriff genommen, wenn das Erlebnis noch in guter Erinnerung ist. Alle Bilder werden dazu betrachtet und sortiert, ausgewählt und dann liebevoll auf die einzelnen Seiten des zuvor erworbenen Fotoalbums verteilt und einzeln eingeklebt. Nicht selten werden dabei in langer Kleinarbeit Fotos ausgewählt oder wieder weggelegt und schließlich passend zum Erlebten gruppiert, solange bis alle Bilder den rechten Platz gefunden haben. Das Ziel ist es, die besondere Erinnerung an das Ereignis in einem Fotoalbum festzuhalten, für sich oder als ein besonderes Geschenk.

Mit dem aktuellen Übergang von der analogen zur digitalen Fotografie, haben der Wunsch und das Interesse an gedruckten Bildern und Alben nicht nachgelassen. Wir befinden uns in einer Marktsituation, in der geschätzt etwa 20 Milliarden Digitalbilder pro Jahr in Europa aufgenommen werden [fCC06]. Gleichzeitig läßt sich beobachten, dass viele der Bilder nicht weiter genutzt werden. Man nimmt an, dass nur etwa 20% der Fotos gedruckt oder in einem physischen Produkt wie einem Album oder einer Tasse verwendet werden [fCC06]. Dies steht im deutlichen Unterschied zum analogen Film, bei dem zumeist alle Bilder des Films auch auf Papier abgezogen wurden.

Der digitale Abzug von Bildern und Fotoalben spielt jedoch nach wie vor eine wichtige Rolle für das persönliche Erlebnis, das man für sich und andere bewahrt und auch zeigen möchte [fCC06]. Neben der klassischen Form, Bilder auszudrucken, bieten Fotodienstleister wie unser Projektpartner CeWe Color Softwareprodukte an, um direkt vom PC aus Bilder für ein Album auswählen zu können und in einem ansprechenden Fotobuch ausdrucken zu lassen. Es bleibt für die Nutzer dennoch das Problem der mühevollen Auswahl aus einer Vielzahl von Fotos und deren Anordnung auf den Seiten des Albums.

Die Frage, warum digitale Bilder nicht weiter genutzt werden, beantworten wir damit, dass die digitalen Abbilder unserer Erlebnisse rar an beschreibenden Informationen sind, die es uns erleichtern könnten, sie wiederzufinden. Die Frage, welche Ereignisse auf den Bildern zu sehen sind, auf welchem Bild Tante Erna so freundlich lächelt oder wie alle Landschaftsaufnahmen in Cornwall aus den letzten Jahren zusammengestellt werden können, kann bisher nur teilweise beantwortet werden. Und gerade hier kann und wird MIR „at the end of the early years“ [SWS⁺00] einen enormen Beitrag leisten. Durch die im sog. EXIF-Header mit den Digitalbildern verfügbaren Metadaten wie Belichtungszeit und Blende aber auch durch situationsbezogene Information wie dem Zeitpunkt der Aufnahme ergeben sich interessante Ansätze, um die automatische Annotation von Bildern deutlich zu verbessern.

Im Folgenden stellen wir Anforderungen der Erstellung eines Fotoalbums in Abschnitt 2 und verwandte Ansätze in Abschnitt 3 vor. Unsere Lösungsansätze im Gebiet kontextsensitives MIR zur automatischen Annotation von persönlichen Fotos beschreiben wir in Abschnitt 4. Die entwickelten Ansätze sind in der Fotobuch-Software unseres Projektpartners CeWe Color zu automatischen Erstellung von Fotoalben implementiert, die wir in Abschnitt 5 vorstellen bevor der Artikel mit einem kurzen Ausblick schließt.

2 Die Erstellung eines Fotoalbums

Um die geforderten Lösungsansätze besser verstehen zu können, betrachten wir zunächst die „Ansichten eines Fotoalbums“: Welche Aufgaben müssen bei der Erstellung eines Fotoalbums

gelöst werden und welche Anforderungen ergeben sich daraus für MIR?

Ein Fotoalbum beinhaltet typischerweise die Bilder eines zeitlich zusammenhängenden Ereignisses in einer angemessenen Darstellung. So werden die Bilder von Urlauben, Familienfeiern und Festen häufig kurz nach dem Erlebten in einem Album arrangiert. Eine weitere Form von Alben sind jene, die die Bilder thematisch zusammenstellen: ein Jahrbuch, das alle wesentlichen Ereignisse des Jahres zeigt; ein Reisealbum, das mehrere Reisen zusammenfasst; eine Biographie, die Bilder aus den verschiedenen Lebensjahren einer Person darstellt. Zentrale Eigenschaft für alle Alben und ihre Betrachter ist, dass das Album ein „schönes“ Abbild des Erlebten darstellen soll. Dabei ist zu bemerken, dass man nicht nur eine schöne Erinnerung schaffen sondern dabei durchaus auch „gut dastehen“ möchte. Daher werden einzelne Bilder weggelassen, die unscharf sind, ungünstige Aufnahmen darstellen oder auch mehrfach die gleiche Situation zeigen. Meist werden mehr Fotos mit der Kamera aufgenommen als später tatsächlich für gut befunden und für das Album ausgewählt werden. Im Zeitalter der digitalen Fotografie ist dies durch die nahezu beliebige Verfügbarkeit von Speicher noch deutlicher geworden. Eine generelle Anforderung an die Albumerstellung ist also, aus einer Menge vorliegender Bilder die am besten geeigneten auszuwählen.

Persönliche Fotos sind typischerweise von deutlich schlechterer Qualität als etwa Testbilder im Bereich der Bildanalyse. Dabei führen zum Beispiel die Spontanität der Aufnahme, eventuell ungeübtes Fotografieren oder Umgang mit der Kamera aber auch ungünstige Beleuchtungsverhältnisse dazu, dass einige Bilder unscharf werden bzw. unter- oder überbelichtet sind. Solche Bilder gilt es entsprechend zu erkennen und für die Fotoauswahl auszublenden, allerdings nicht uneingeschränkt. Die Reduktion der Anzahl der Bilder stellt keine einfache Folge von Ausschlusskriterien wie etwa Unschärfe oder Überbelichtung dar.

Nehmen wir dazu einmal an, ein bestimmtes Bild zeige die versammelte Familie bei einem großen Familienfest. Leider ist das Bild unscharf geworden und würde so nach einfachen Ausschlusskriterien nicht für das Album ausgewählt werden. Zu diesem Zeitpunkt wurde aber kein zweites Bild aufgenommen und so ist es das einzige Foto zu diesem Ereignis. Daher würde das Bild, wenn auch unscharf, vermutlich doch mit zum Album hinzugenommen werden. Andererseits wurde eine bestimmte Szene wie etwa der Sonnenuntergang oder das neue Auto gleich mehrfach fotografiert. An dieser Stelle würde man nicht unbedingt alle, wohl aber die schönsten Aufnahmen in das Album übernehmen. Eine weitere Rolle spielt die Balance der Bilder zum Erlebten. Werden zum Beispiel auf einer Urlaubsreise durch die Toskana verschiedene Städte besucht, so soll das Album doch auch jeden Ort angemessen darstellen und an die Städtereise erinnern, auch wenn die Zahl der Bilder pro Stadt durchaus unterschiedlich ist. Diese Beispiele sollen illustrieren, dass verschiedene Faktoren aus dem Inhalt wie Schärfe, Farbe, Helligkeit aber auch wer und was darauf zu sehen ist in Kombination mit Parametern der Situation wie etwa Ort und Zeit wichtig für die Bildauswahl sind. Die Parameter einer angemessenen Repräsentation des zeitlichen Verlaufs und der besuchten Orte aber auch eine ausgewogene Auswahl der Arten von Bildern, der gezeigten Personen sind wichtig. Das Album wird also bestimmt von einer besonderen Balance verschiedener Parameter, die die Bildauswahl bestimmen.

Für das MIR bedeutet dies, dass das Erreichen einer hohen Zufriedenheit der Nutzer bei einer automatischen Auswahl von Bildern für ein Album sehr stark damit zusammenhängt, wie genau man die Domäne der Albumerstellung für die verschiedenen Ereignisse aber auch die Vorlieben der individuellen Nutzer versteht und in entsprechende Retrievalverfahren einbetten kann.

Dabei enden die Anforderungen der Erstellung eines Fotoalbums nicht bei der Auswahl der relevanten Bilder für das Album. Die Seitenorientierung der Alben erfordert auch eine passende Verteilung der Fotos auf die einzelnen Seiten. Dabei spielen erneut inhaltsbasierte Faktoren aber auch Kontextparameter eine Rolle. So würde man in aller Regel Fotos eines Ereignisses auch nah beieinander etwa links und rechts auf einer aufgeschlagenen Fotobuchseite platzieren und verhindern wollen, dass eines der Bilder auf die nächste Seite „rutscht“. Je nach Art des Albums können auch Farbaspekte eine Rolle spielen, um ähnlich eingefärbte Bilder oder Bilder mit ähnlicher Bildkomposition, gerade zusammenzustellen, oder aber als Auflockerung gerade eine Vermischung von Stil und Farbe zu erreichen. Kontext und Inhalt der Bilder bleiben also auch für die Anordnung der Bilder im Album und die Gestaltung des Albums relevant.

3 Verwandte Arbeiten

In den vergangenen Jahren finden digitale Fotos eine immer stärkere Aufmerksamkeit im Forschungsgebiet MIR. Im Jahr 2003 untersuchten Rodden und Wood [RW03], ob Techniken der Spracherkennung und Bildanalyse im Gebiet persönlicher Fotosammlungen hilfreich sein können. Die Autoren kommen zum Schluss, dass Nutzer die eigenen Bildersammlungen meist intuitiv nach dem zeitlichen Verlauf sowie den damit zusammenhängenden Ereignissen ordnen und nicht nach detaillierten Multimedia-Eigenschaften. Weiterhin stellen die Autoren zu recht fest, dass manuelle Annotationen nicht von allen Nutzern erwartet werden können.

Im etablierten Feld der inhaltsbasierten Analyse und des Retrieval von Bildern wurden in den letzten Jahren einige Arbeiten im Anwendungsgebiet persönlicher Fotosammlungen und digitaler Fotoalben veröffentlicht. In diesen Ansätzen wird inhaltsbasierte Analyse, teilweise auch in Kombination mit „user relevance feedback“ eingesetzt, um persönliche Fotos zu annotieren und zu organisieren. Prominente erste Beispiele sind etwa MiAlbum [WSZ00], AutoAlbum [Pla00] und SmartAlbum [TCMK02]. In den nachfolgenden Jahren wurden beispielsweise im DIVA Projekt Lernansätze verwendet, um inhaltsbasiert Fotos zu annotieren [PMK03]; in [ZCLZ03] werden durch inhaltsbasierte Gesichtserkennung die Namen von Familienmitgliedern annotiert. Da die Fotos im Retrieval auch den Erwartungen der Nutzer gerecht werden müssen, stellen [MCM05] Hypothesen bezüglich von Wahrnehmungswerten wie Größe und Fragmentierung auf und schlagen dafür ein Gewichtungsmodell für das Retrieval von Bildern vor. Obgleich diese Arbeiten interessante Ergebnisse erzielen zeichnet sich im Bereich des inhaltsbasierten Retrieval ab, dass inhaltsbasierte Annotation alleine nicht ausreicht. Am Ende der ersten Jahren des Gebietes MIR läßt sich sagen, dass Quellen außerhalb des Bildes [SWS⁺00] dazu beitragen werden, die sog. semantische Lücke zu schließen.

Mit der Verfügbarkeit von Zeit und Ort in Digitalkameras entstanden Arbeiten, die diese Kontextinformationen, teilweise in Kombination mit inhaltsbasierten Eigenschaften, für die Organisation und Annotation von Bildern verwenden. In PhotoTOC [PCF02], werden der Zeitstempel des Fotos und Farbhistogramme genutzt, um die Bildersammlung in einer visuellen Nutzungsschnittstelle zu organisieren. Die Zeit der Aufnahme wird in verschiedenen Arbeiten dazu verwendet, um Ereignisse zu erkennen, zu gruppieren und die zugehörigen Bilder gemeinsam darzustellen [ML03, GGMPW02, CFGW05]. Ebenso wird der Ort in Form einer Geokoordinate als Kontextelement verwendet, um die Bilder räumlich zu organisieren und zugreifbar zu machen, wie etwa im ATLAS Projekt [PG04, PG05]. Die Bedeutung der Kombination von Zeit und Raum für die Bild-

organisation zeigen die Arbeiten im Stanford Digital Library Project in [NSPGM04, NHWP04] wie auch im MediAssist Projekt [OGL⁺05, NGCA05]. Nicht nur die Kombination von Inhalt und Kontext, auch Kontextinformation alleine wird verwendet, um Bilder zu annotieren wie etwa in [NYGMP05], wo Vorschläge für "identity labels" auf temporalem, räumlichem und sozialem Kontext basieren. Der EXIF-Header wird zur Szeneklassifikation [BL05] oder aber als Basis für kontextabhängige Vorschläge zur Verschlagwortung [LWP⁺06] für mobile Videoclips verwendet.

Bestehende kommerzielle Anwendungen wie Picasa [Goo06], ACDSsee[ACD06] oder iPhoto[App06] erlauben heute zur Verwaltung von Fotosammlungen Fotos in vielfältiger Weise zu annotieren und zu kategorisieren, allerdings noch manuell. Im AKTive Media Annotation Tool [THD⁺06] werden verschiedene Eigenschaften von Foto und EXIF Header genutzt um Annotationsvorschläge für Fotos zu generieren, die dann vom Nutzer nachbearbeitet werden können. Die seit 2006 im Beta-test verfügbare Anwendung Riya [Ojo06] verwendet Bildanalyse zur (semi-)automatischen Annotation von Namen basierend auf Gesichtserkennung. In ihrer Ausrichtung beschäftigen sich diese Arbeiten mit der semi-automatischen Annotation einzelner Bilder zur Organisation und Suche. Eine umfassende automatische inhalts- und kontextbasierte Annotation sowie die besonderen Anforderungen der balancierten Auswahl der Bilder und das Arrangement in einem Fotoalbum werden nicht gesondert berücksichtigt.

In unserem Context-aware Smart Multimedia Authoring Tool (xSMART) [SB05] haben wir die Aspekte von kontextsensitivem Authoring in der Domäne der Erstellung von Fotoalben erprobt und die besonderen Anforderungen der automatischen Annotation von Fotos und an eine Albumsoftware erarbeitet, die wir in diesem Artikel vorstellen.

4 Lösungsansatz

Aus der Zielsetzung der automatischen Erstellung von digitalen Fotobüchern ergibt sich die Notwendigkeit, digitale Fotos mit Hilfe von Kontextinformation mit Semantiken anzureichern, um so durch geeignetes Retrieval Fotos auswählen und auf den Seiten platzieren zu können. Unser Lösungsansatz für ein effektives MIR basiert auf der Idee, Bild- und zugehörige Kontextdaten sukzessiv zu extrahieren und abzuleiten. So ist es möglich, auch aus bereits erzeugten Metadaten weitere Metadaten abzuleiten. Dieser Ansatz wird durch unsere komponentenbasierte, workflowgetriebene Softwarearchitektur realisiert.

Abbildung 1 verdeutlicht MetaXa, unsere „Content-based and context-driven metadata enhancement architecture“ näher. Dargestellt sind mehrere Extraktions- und Anreicherungskomponenten, die von einer Managerkomponente mit Hilfe einer Workflowkomponente angesprochen werden. Als Eingabe dienen Bilder, die von einer Digitalkamera stammen, zusammen mit ihren Kontextdaten, die z.B. durch die Kamera im EXIF-Header des Bildes gespeichert sind. Die Fotos durchlaufen eine Reihe von Extraktions- und Anreicherungskomponenten, wobei in jedem Schritt die bereits existierenden oder abgeleiteten Metadaten als Eingabe für den folgenden Anreicherungsschritt dienen. Nach und nach wird so das Foto mit immer mehr Metadaten angereichert, so dass auch eine immer bessere semantische Beschreibung entsteht. Durch diese modulare Struktur ist es leicht möglich, das System um zusätzliche Komponenten zu erweitern bzw. Komponenten auszutauschen.

Jede der Anreicherungskomponenten ist dabei durch entsprechende Vor- und Nachbedingungen näher spezifiziert. Durch diese Bedingungen wird angezeigt, welche Metadaten die Komponente als Eingabe benötigt und welche sie liefert. Mit Hilfe

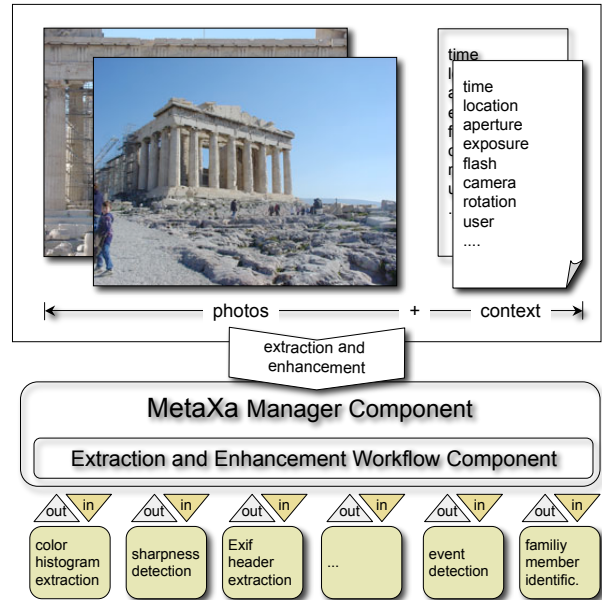


Abbildung 1: Gesamtarchitektur

dieser Angabe erstellt die Workflowkomponente dynamisch eine geeignete Reihenfolge, in der die Komponenten aufzurufen sind. Abbildung 2 zeigt einige der entwickelten Anreicherungskomponenten und ihre Abhängigkeiten untereinander.

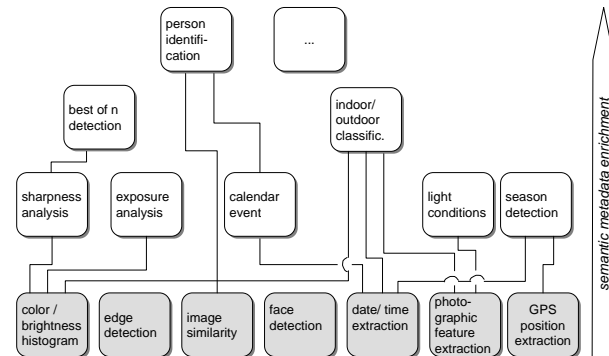


Abbildung 2: Verschiedene Extraktions- und Anreicherungskomponenten

Extraktions- und Anreicherungskomponenten Um einen genaueren Einblick in die Architektur und die Extraktion und Anreicherung von Fotos mit Metadaten zu geben, sollen im folgenden ein kleiner Überblick der entwickelten Komponenten gegeben werden. Wir unterscheiden zwischen Extraktions- und Ableitungskomponenten. Extraktionskomponenten extrahieren Inhalts- und Kontextinformationen aus den Bildern und benötigen dafür keine bereits von anderen Komponenten generierten Metadaten. Wir haben Extraktionskomponenten entwickelt, die Farbhistogramme, Kanten und weitere Features aus dem Bild extrahieren. Darüberhinaus werden Kontextinformationen wie Zeit, Datum und Ort aber auch fotografische Informationen wie Belichtungszeit und Blende aus einem vorhandenen EXIF-Header extrahiert. In darauf aufbauenden Anreicherungskomponenten werden aus den so extrahierten Metadaten weitere Metadaten abgeleitet. So lässt sich beispielsweise aus einem Helligkeitshistogramm und den im EXIF-Header gespeicherten Informationen über Blende und Belichtung schliessen, ob das Bild eine Innen- oder Aussenaufnahme darstellt. Hierzu wird aus der Blendens- und Verschlusszeit abgeleitet, welche

Helligkeitsverhältnisse zum Aufnahmezeitpunkt herrschten. In Kombination dieses Helligkeitswertes mit dem Zeitpunkt der Aufnahme aus dem EXIF-Header und dem Helligkeitshistogramm des Bildes kann eine sehr zuverlässige Aussage darüber getroffen werden, ob es sich um eine Innen- oder Aussenaufnahme handelt. Ein Test mit 437 Fotos eines 4-Wöchigen Urlaubs, von denen 68 Innen- und 369 Aussenaufnahmen waren, zeigte eine Fehlerrate von 0,6 % bei den Aussen- und 10,1% bei den Innenaufnahmen. 10,1% des gesamten Testsets wurden als nicht eindeutig klassifizierbar eingestuft. Wie dieses Beispiel zeigt, ist gerade die Kombination von inhaltsbasierten Eigenschaften mit dem Kontext der Aufnahme besonders hilfreich für die Ableitung höherwertiger Metadaten. Wir haben in unseren Arbeiten verschiedene dieser Komponenten entwickelt und in die in Abbildung 1 vorgestellte Architektur integriert.

Generierung von Fotoalben Die generierten Metadaten können in vielerlei Hinsicht für die Fotoalbumerstellung genutzt werden. Eine Herausforderung ist die Möglichkeit der automatischen Auswahl von Fotos aus einem Datensatz von beispielsweise 1000 Fotos, die während eines 4-wöchigen Urlaubs gemacht wurden. Üblicherweise werden hier die *besten* herausgesucht und in einem Fotoalbum genutzt. Was aber sind nun die besten Fotos? Die Auswahl der Fotos gliedert sich in verschiedene Schritte: Zunächst werden alle Fotos nach bestimmten Kriterien (entweder Ort oder Zeit) zu Clustern oder sog. Mini-Alben zusammengefasst. Ausgehend hiervon können nun ganz pragmatisch zunächst Bilder aussortiert werden, die üblicherweise als nicht gelungen gelten, da etwa die Belichtung nicht stimmt oder das Bild unscharf ist. Dabei muss darauf geachtet werden, dass zumindest die Mini-Alben nicht komplett gelöscht werden: üblicherweise möchte man Bilder, wenn sie die einzigen sind, die zu einem Ereignis gehören, auch dann in einem Fotoalbum erhalten, wenn sie „schlecht“ sind. In einem weiteren Schritt können Fotos entfernt werden, die sehr ähnlich zu zeitlich benachbarten sind. Häufig kommt es vor, dass mit einer Digitalkamera mehrere Aufnahmen von einem Motiv gemacht werden, um sicherzugehen, dass mindestens eine gute dabei ist. Dies deutet darauf hin, dass das Motiv für den Benutzer wichtig war, also wenigstens ein Foto aus der Reihe auch in dem Fotoalbum enthalten sein sollte. Aus einer solchen Reihe kann jeweils das *beste* Foto (in Bezug auf Belichtung und Schärfe) ausgewählt werden, um die Auswahl an Fotos weiter zu reduzieren. In unseren Arbeiten werden bereits verschiedene Kriterien eingesetzt, um eine automatische Auswahl als Vorschlag für ein Fotoalbum vorzunehmen. Basierend auf den Rückmeldungen der Markteinführung der kommerziellen Fotobuchsoftware unseres Projektpartners (vgl. Abschnitt 5) werden wir das Retrieval der Fotos verfeinern und um weitere inhalts- und kontextbasierte Kriterien erweitern, die in die Relevanzfunktion für die Auswahl einfließen.

Gestaltung von Fotoalben Um ein Fotobuch durch einen Computer generieren zu lassen, muss zunächst genauer spezifiziert werden, durch welche Parameter es bestimmt ist. Unser Partner CeWe Color liefert gedruckte Fotoalben in mehreren Ausführungen. Unter anderem unterscheiden sich diese Ausführungen durch ihre Seitengröße, ihre Seitenzahl und ob der Umschlag bedruckbar ist, oder nicht. Die Seiten eines Fotobuches können verschiedene Hintergründe besitzen, auf denen die Bilder aber auch textuelle Annotationen platziert werden können. Innerhalb einer Seite können die Bilder beliebig angeordnet, gekippt und skaliert werden. Dabei können auch nur Ausschnitte eines Bildes gewählt oder Bilder überlappend angeordnet werden. All diese Möglichkeiten, bis auf die Skalierung eines Bildes und Gestaltung des Hintergrundes, bestehen auch

bei einem herkömmlichen „analogen“ Fotoalbum und lehnen sich so stark an die herkömmliche Fotobuchgestaltung an.

Die Gestaltung von Seiten, das Arrangement der Bilder und die Hintergründe hängen ebenfalls von den zuvor extrahierten Metadaten ab. Typischerweise tendiert man dazu, das Buch in verschiedene Abschnitte oder Gruppen von Fotos aufzuteilen. Bei Urlaubsreisen sind diese Gruppen beispielsweise verschiedene Orte oder Tage; bei einer Geburtstagsfeier können es verschiedene Personengruppen sein, die auf den Fotos dargestellt sind. Wir kommen diesem Konzept nach indem wir ein Fotoalbum aus verschiedenen Mini-Alben zusammensetzen, die jeweils Bilder einer dieser Gruppen enthalten. Jedes Mini-Album repräsentiert einen Cluster, der Fotos entsprechend ihrer Eigenschaften gruppiert. Abhängig von der Art des gewünschten Albums können dies zeitlich zusammengehörige, nach Personen oder auch nach Ort gruppierte Fotos sein. Durch entsprechende Clustering-Verfahren werden die Mini-Alben gefüllt und dienen für das Arrangement auf den Albumseiten. Ziel ist es nicht nur die Fotos in Mini-Alben zu gruppieren sondern die Mini-Alben auch visuell voneinander abzugrenzen. So können sich z.B. die Hintergründe aufeinanderfolgender Mini-Alben unterscheiden oder jedes Mini-Album immer auf einer neuen Doppelseite beginnen. Wenn einem Mini-Album eine Annotation wie etwa ein Ortsname zugeordnet wurde, kann diese als Überschrift für den Abschnitt des Mini-Albums oder auch ein automatisch generiertes Inhaltsverzeichnis genutzt werden. Für die Seitengestaltung verwenden wir inhaltsbasierte Eigenschaften der Bildgruppe, um aus den Farbhistogrammen einen Vorschlag für den Hintergrund zu generieren. Weiterhin arbeiten wir daran, die Semantiken, die sich aus der Albumerstellung und einer Nachbearbeitung durch die Nutzer ergeben, für die Ableitung weiterer Metadaten der Bilder zu nutzen [SBC06].

5 Die Fotobuch-Software

Die vorgestellten Arbeiten stehen im Zusammenhang mit einer Kooperation mit CeWe Color, Europas größtem Fotodienstleister. Mit dem Übergang von der analogen zur digitalen Fotografie ergab sich für das Unternehmen nicht nur ein Umbruch der Dienstleistungen sondern auch die Möglichkeit, die digitalen Fotos und ihren Kontext für neue Dienstleistungen zu nutzen. Konkret fließen die Ergebnisse der Arbeit in die Entwicklung der Fotobuch Software¹ von CeWe Color ein, die im Juni 2006 auf den Markt gekommen ist. Ziel war die Entwicklung von Konzepten für eine Fotobuch-Software, die den Nutzern und Nutzern eine einfache Erstellung von gedruckten Fotoalben mit nur wenigen Klicks erlaubt. Sie zeigt auf, wie die Forschungsergebnisse aus dem Gebiet MIR in Context einen Beitrag für den Markt der Verwaltung und Nutzung persönlicher Medien leisten können. Als Produkt bietet CeWe Color dazu Fotobücher an, wie sie in Abbildung 3 illustriert sind. Diese Bücher sind so gestaltet, dass die Bilder gemeinsam mit verschiedenen Hintergründen gedruckt werden. Zu den Fotos können zusätzliche Beschriftungen eingegeben werden. Mit unterschiedlichen Umschlägen als Hardcover oder Softcover mit Titelbild und Beschriftung werden die Fotoalben gebunden.

Für die Generierung eines Albums aus einer zuvor ausgewählten Bildersammlung können verschiedene Einstellungen vorgenommen werden, die auf den zuvor vorgestellten Konzepten zur automatischen Extraktion und Annotation für die Auswahl und Gestaltung von Fotoalben basieren. Vom Fotoalbumnutzer kann über einfache Auswahlfelder festgelegt werden, ob ein automatischer Vorschlag überhaupt gemacht werden soll, ob unscharfe und unter- oder überbelichtete Bilder sowie ähnliche Bilder aus der Auswahl entfernt werden dürfen. Weiterhin können

¹<http://www.cewe-fotobuch.de/>



Abbildung 3: Eine aufgeschlagene Fotobuchseite

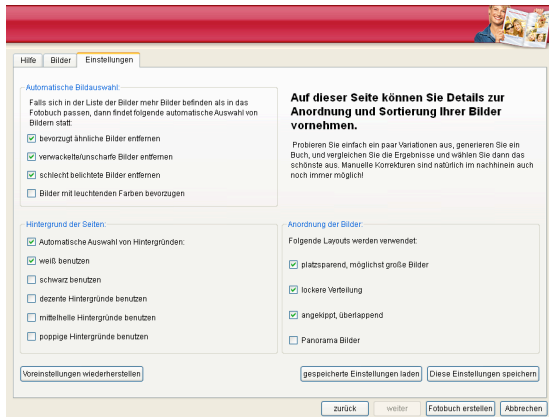


Abbildung 4: Einstellungen für die automatische Generierung und Gestaltung des Fotobuches

Angaben zu Gestaltung und Layout vorgenommen werden wie etwa ob eine aufgelockerte Platzierung der Bilder erwünscht ist oder ob farbige Hintergründe vorgeschlagen werden dürfen.

Mit diesen Einstellungen kann nun nach der Auswahl der Fotosammlung ein Vorschlag für das Fotobuch erstellt werden. In diesem Schritt erfolgt die Auswertung der Metadaten, Gewichtung und Gruppierung der Bilder und die Verteilung auf die einzelnen Seiten und Layouts. Abbildung 5 zeigt einen Screenshot der Albumsoftware nachdem der Assistent einen Vorschlag gemacht hat. In einer Leiste oben werden alle Seiten des Albums klein angezeigt und die jeweils aktuell angesehenen oder bearbeiteten Seite rechts in der Mitte. Die gewählte Seite zeigt eine automatische Positionierung von Bildern, die in einer gewissen zeitlichen Nähe entstanden sind und gruppiert sie auf den Seiten. Dafür werden verschiedene Layouts verwendet und die Verteilung auf die Seite und in der Größe erfolgt entsprechend der Bildeigenschaften. Die unterschiedlichen Farben der Hintergründe basieren auf den Farbeigenschaften der Bildgruppe auf der Seite. Der Rahmen um das Bild in der Mitte links im Album deutet an, dass nicht nur die Auswahl sondern auch Größe, Lage und Ausrichtung aller Bilder noch nachträglich verändert werden können.

Abbildung 6 zeigt die Erstellung eines Albums einer Australienreise. Insbesondere vom Ayers Rock wurde eine Vielzahl von Bildern aufgenommen. Die automatische Vorauswahl hat basierend auf Bildschärfe und Ähnlichkeit für diese aufgeschlagene Seite sechs Bilder ausgewählt. Die Bildleiste links zeigt dabei mit einem grünen Häkchen an, welche Bilder ausgewählt wurden und läßt den Kunden jederzeit die Möglichkeit, diese Auswahl zu verändern. Bereits jetzt kann ein Album mit nur vier Klicks erstellt und online oder über eine CD bestellt werden.



Abbildung 5: Automatische Gestaltung des Albums mit der Ce-We Fotobuch Software



Abbildung 6: Automatische Vorauswahl von Bildern in der Ce-We Fotobuch Software

Für die Markteinführung geht das Konzept für die Software davon aus, dass Benutzer tendenziell alle gemachten Fotos auch in dem generierten Album sehen wollen. Die entwickelte automatische Aussortierung wird daher eher von den geübteren Nutzern eingestellt werden. Hier wendet die Software die zuvor beschriebenen Algorithmen an, die Unschärfe, falsche Belichtung und Ähnlichkeit zeitlich nah beieinander liegender Bilder erkennen sollen. Besondere Herausforderung in diesem Rahmen war es, ein Gleichgewicht zwischen diesen Ergebnissen durch eine geeignete Bewertung zu finden. Für die nun marktreife Version der Fotobuchsoftware wurde das Ausfiltern auf wenige Bilder optimiert, d.h. höchstens 20% der Bilder werden nicht mit in das Buch einsortiert. Das bedeutet nicht, dass nicht auch gute Ergebnisse für höhere Raten erzielt werden können. Um das Vertrauen der Kunden zu gewinnen, wird aber zunächst eher vorsichtig in die Fotoauswahl eingegriffen. Die statistische Auswertung der Nutzung der Albumsoftware und Benutzerstudien werden einen wesentlichen Beitrag zur Weiterentwicklung der Konzepte leisten. Ziel ist es, die Auswahl der Bilder und die Gestaltung des Albums immer weiter zu verfeinern und eine manuelle Nachbearbeitung immer seltener werden zu lassen. Dies wird unter anderem durch von CeWe Color durchgeführten Usability Tests und Rückmeldungen von Kunden an die Support Abteilung erreicht.

6 Ausblick

Intuitive Verwaltung und Nutzung von digitalen Fotosammlungen ist nicht nur ein dringendes sondern auch ein kommerziell sehr relevantes Thema. Durch die Kombination von inhaltsbasierten Techniken mit der Ausnutzung von Kontextinformationen ergeben sich ganz neue Potentiale für die automatische Annotation von Bildern. Damit entsteht die Grundlage für eine innovative Nutzung und Verwendung der Fotos in neuen physischen Produkten und digitalen Dienstleistungen.

In unseren weiteren Arbeiten werden wir die verschiedenen Extraktions- und Anreicherungskomponenten verfeinern und erweitern. Basierend auf den abgeleiteten Metadaten entwickeln wir Heuristiken und probabilistische Ansätze für die Vorschläge zur Bildauswahl und erproben diese an repräsentativen Testsammlungen. Weiterhin arbeiten wir an der Erweiterung der inhaltsbasierten und kontextabhängigen Layoutgestaltung der Fotoalben. Die ersten Rückmeldungen von Nutzern des Fotoalbums und der darin eingesetzten Techniken des MIR zeigen, dass Forschungsergebnisse nicht nur wissenschaftlich relevant sein können sondern auch einen Beitrag zur Freude von Menschen über ein schöne Erinnerung in Form eines Fotoalbums leisten können.

Literatur

- [ACD06] ACDS Systems International Inc. acdsee Pro, 2006. <http://www.acdsystems.com>.
- [App06] Apple Inc. iPhoto, 2006. <http://www.apple.com/de/ilife/iphoto/>.
- [BL05] Matthew R. Boutell and Jiebo Luo. Beyond pixels: Exploiting camera metadata for photo classification. *Pattern Recognition*, 38(6):935–946, 2005.
- [CFGW05] Matthew Cooper, Jonathan Foote, Andreas Girgensohn, and Lynn Wilcox. Temporal event clustering for digital photo collections. *ACM Trans. Multimedia Comput. Commun. Appl.*, 1(3):269–288, 2005.
- [fCC06] GfK-Panel Services für CeWe Color. Nutzungsverhalten Digitalfotografie 2006. Eine repräsentative Befragung von Digitalkamerabesitzern in deutschen Haushalten durch die GfK, 2006.
- [GGMPW02] Adrian Graham, Hector Garcia-Molina, Andreas Paepcke, and Terry Winograd. Time as essence for photo browsing through personal digital libraries. In *Proc. of the 2nd ACM/IEEE-CS Joint Conf. on Digital Libraries (JCDL)*, pages 326–335, Portland, Oregon, USA, 2002. ACM Press.
- [Goo06] Google, Inc. Picasa, 2006. <http://picasa.google.com/>.
- [LWP+06] Lahti, Westermann, Palola, Peltola, and Vildjiounaite. Context-aware mobile capture and sharing of video clips. In *Handbook of Research on Mobile Multimedia*. Idea Publishing, 2006.
- [MCM05] Jean Martinet, Yves Chiaramella, and Philippe Mulhem. A model for weighting image objects in home photographs. In *Proc. of the 14th ACM international conference on Information and knowledge management (CIKM)*, pages 760–767, New York, NY, USA, 2005. ACM Press.
- [ML03] Philippe Mulhem and Joo-Hwee Lim. Home photo retrieval: Time matters. In *Proc. of the Second International Conference on Image and Video Retrieval (CIVR)*, pages 321–330, Urbana-Champaign, IL, USA, July 24–25, 2003. Springer, LNCS.
- [NGCA05] O’Hare N, Jones G Gurrin C, and Smeaton A.F. Combination of content analysis and context features for digital photograph retrieval. In *2nd IEE European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies*, London, U.K., 30 November–1 December 2005.
- [NHWP04] Mor Naaman, Susumu Harada, QianYing Wang, and Andreas Paepcke. Adventures in space and time: Browsing personal collections of geo-referenced digital photographs. Technical report, Stanford University, InfoLab, 2004.
- [NSPGM04] Mor Naaman, Yee Jiun Song, Andreas Paepcke, and Hector Garcia-Molina. Automatic organization for digital photographs with geographic coordinates. In *Proceedings of the 4th ACM/IEEE-CS Joint Conf. on Digital Libraries (JCDL)*, pages 53–62, Tuscon, AZ, USA, 2004. ACM Press.
- [NYGMP05] Mor Naaman, Ron B. Yeh, Hector Garcia-Molina, and Andreas Paepcke. Leveraging context to resolve identity in photo albums. In *Proc. of the 5th ACM/IEEE-CS joint Conference on Digital Libraries (JCDL)*, pages 178–187, New York, NY, USA, 2005. ACM Press.
- [OGL+05] Neil O’Hare, Cathal Gurrin, Hyowon Lee, Noel Murphy, Alan F. Smeaton, and Gareth J.F. Jones. My digital photos: where and when? In *Proceedings of the 13th annual ACM International Conference on Multimedia (MM)*, pages 261–262, New York, NY, USA, 2005. ACM Press.
- [Ojo06] Inc. Ojos. Riya Photo Search, 2006. URL: <http://www.riya.com/>.
- [PCF02] John C. Platt, Mary Czerwinski, and Brent A. Field. PhotoTOC: Automatic Clustering for Browsing Personal Photographs. Technical report, Microsoft Research, 2002.

- [PG04] A. Pigeau and Marc Gelgon. Incremental statistical geo-temporal structuring of a personal camera phone image collection. In *ICPR (3)*, pages 878–881, 2004.
- [PG05] A. Pigeau and M. Gelgon. Building and tracking hierarchical geographical & temporal partitions for image collection management on mobile devices. In *Proc. of the 13th ACM international conference on Multimedia (MM)*, pages 141–150, New York, NY, USA, 2005. ACM Press.
- [Pla00] John C. Platt. Autoalbum: Clustering digital photographs using probabilistic model merging. In *Proc. of the IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL)*, page 96, Washington, DC, USA, 2000. IEEE Computer Society.
- [PMK03] W. K. L. P. Mulhem, J. H. Lim and M. Kankanhalli. *Advances in Digital Home Image Albums*, pages 201–226. Idea Publishing, 2003.
- [RW03] Kerry Rodden and Kenneth R. Wood. How do people manage their digital photographs? In *Proc. of the SIGCHI conference on Human factors in computing systems*, pages 409–416, New York, NY, USA, 2003. ACM Press.
- [SB05] Angar Scherp and Susanne Boll. Context driven smart authoring of multimedia content. In *Proc. of the 13th ACM international conference on Multimedia (MM)*, New York, NY, USA, 2005. ACM Press.
- [SBC06] Angar Scherp, Susanne Boll, and Holger Cremer. Emergent semantics in personalized multimedia content. In *Fourth special workshop on Multimedia Semantics (WMS)*, Chania, Greece, 2006.
- [SWS⁺00] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, 2000.
- [TCMK02] Tele Tan, Jiayi Chen, Philippe Mulhem, and Mohan Kankanhalli. Smartalbum: a multi-modal photo annotation system. In *Proc. of the tenth ACM international conference on Multimedia (MM)*, pages 87–88, Juan-les-Pins, France, 2002. ACM Press.
- [THD⁺06] Mischa M. Tuffield, Stephen Harris, David P. Duppaw, Ajay Chakravarthy, Christopher Brewster, Nicholas Gibbins, Kieron O’Hara, Fabio Ciravegna, Derek Sleeman, Nigel R. Shadbolt, and Yorick Wilks. Image annotation with photocopain. In *First Intl. Workshop on Semantic Web Annotations for Multimedia (SWAMM)*, 2006.
- [WSZ00] Liu Wenyin, Yanfeng Sun, and Hongjiang Zhang. Mi-album - a system for home photo management using the semi-automatic image annotation approach. In *Proc. of the eighth ACM Intl. Conference on Multimedia (MM)*, pages 479–480, New York, NY, USA, 2000. ACM Press.
- [ZCLZ03] Lei Zhang, Longbin Chen, Mingjing Li, and Hongjiang Zhang. Automated annotation of human faces in family albums. In *Proc. of the eleventh ACM international conference on Multimedia (MM)*, pages 355–358, New York, NY, USA, 2003. ACM Press.