

Content Management multimedialer Inhalte auf mobilen Endgeräten

Wassim Ben Hamadou
benhamad@ifi.lmu.de

Universität München
Amalienstrasse 17, 80333 München, Germany

Zusammenfassung Im ersten Teil dieser Ausarbeitung geht es um Problemstellungen und um Ansatzpunkte zur Bewältigung dieser in Form einer Softwarelösung namens *Video Semantic Summarization System*, die sich mit der Generierung bzw. Verwaltung von Videoinhalten und deren Metadaten auf mobilen Endgeräten auseinandersetzen. Im zweiten Teil liegt der Fokus auf Fotoinhalte. Hier werden Probleme wie Verwaltung auf mobilen Endgeräten und Bildaustausch mit anderen Benutzern erläutert. Zu diesem Thema werden drei Softwarelösungen vorgestellt die diese Problematik auf verschiedene Arten angehen. Zudem wird anhand einer Studie gezeigt wie in Bezug auf Fotos das Benutzerverhalten von der verwendeten Software abhängt. Es wird ein Einblick in Möglichkeiten und Grenzen von Content Management auf mobilen Endgeräten gegeben, die derzeit bestehen.

1 Einleitung

Mit der Zeit entwickeln sich die technischen Möglichkeiten von mobilen Endgeräten weiter, woraus auch erhöhte Erwartungen der Nutzer an die Funktionalität dieser Geräte resultieren. Daher beschränken sich die Anwendungsgebiete nicht nur auf SMS, Terminverwaltung und Telefonate, sondern fokussieren sich immer mehr auf multimediale Bereiche wie Aufnahme, Verwaltung und Austausch von Bildern und Videos. Die Anzahl der zu verwaltenden Multimedia Inhalte nimmt dadurch drastisch zu. Um dem Nutzer dennoch eine komfortable Nutzung der Dateien und Funktionalitäten zu ermöglichen gewinnt die Effizienz von Content Management Systemen auf mobilen Endgeräten immer mehr an Bedeutung. Ein zentraler Aspekt von Content Management Systemen sind so genannte Metadaten, die thematisch daher auch den Kern dieser Ausarbeitung darstellen.

2 Fokus und Gliederung

Bei der folgenden Untersuchung der Metadaten Thematik definiert sich der analytische Rahmen aus der Betrachtung möglicher Problemstellungen anhand von Anwendungsszenarien und der Vorstellung bereits vorhandener Lösungsansätze. Als konkrete Anwendungsgebiete wurden Video und Foto ausgewählt. Vor Beginn der eigentlichen Analyse werden im folgenden Kapitel grundlegende Begriffe erklärt. Die Kapitel vier und fünf widmen sich dann jeweils einem der thematischen Schwerpunkte Videos bzw. Bilder. Einleitend dazu wird jeweils ein frei erfundenes aber dennoch realistisches Anwendungsszenario vorgestellt.

Danach werden anhand dessen mögliche Problemstellungen aufgezeigt und erläutert. Abschließend werden bereits vorhandene Software Lösungen und Ansätze zur Bewältigung der vorher aufgezeigten Probleme vorgestellt. Kapitel sechs fasst die wichtigsten Ergebnisse in einem Fazit zusammen.

3 Grundlegende Begriffe

3.1 Mobiles Endgerät

Unter einem Endgerät versteht man "... in der Informationstechnik (IT) und der Telekommunikationstechnik (TK) ein Objekt (zum Beispiel PC oder Telefon), welches an einen Netzabschluss eines öffentlichen oder privaten Daten- oder Telekommunikationsnetzes angeschlossen ist. Dabei kann es sich um eine direkte elektrische Verbindung handeln, beispielsweise mit einem Steckverbinder, oder auch um eine Funkverbindung wie bei einem Mobiltelefon." ([1]) Demnach ist ein mobiles Endgerät ein Objekt mit dem man sich unterwegs mittels einer drahtlosen Verbindung in eines der oben genannten Netze verbinden kann. Beispiele hierfür sind Smartphones oder PDA.

3.2 Content Management

"Content Management (abgekürzt CM) ist die Zusammenfassung aller Tätigkeiten, Prozesse und Hilfsmittel, die den Lebenszyklus digitaler Informationen in Form von Unterlagen und Dokumenten unterstützen. Die digitalen Informationen werden oft als Content (Inhalt) bezeichnet, sie können als Dateien vorliegen, die einzeln verarbeitet werden oder auch als zusammenhängende Dateisysteme, wie z.B. Webseiten. Der Lebenszyklus solcher Unterlagen und Dokumente durchläuft typischerweise sechs grundlegende Phasen: Erzeugung, Überarbeitung, Veröffentlichung, Übersetzung, Ablage und Archivierung, Ausscheidung." ([2]) Folgendes Beispiel soll dies verdeutlichen: Bei Aufnahme eines Videos mittels eines Smartphones ermöglicht ein Content Management System dem Nutzer dateibezogene Aktionen wie Speichern, Umbenennen, Überschreiben oder gar Löschen. Auch die Suche der Datei ist ein dem CM zugeordneter Prozess.

3.3 Metadaten

"Als Metadaten oder Metainformationen bezeichnet man allgemein Daten, die Informationen über andere Daten enthalten." ([3]) So könnte eine Metadatei eines aufgenommenen Bildes Informationen enthalten wie Datum, Ort und Zeitpunkt der Aufnahme oder sogar die Beschreibung der Bildinhalte.

4 Videos auf mobilen Endgeräten

4.1 Anwendungsszenario

Ein Anwender möchte über sein Smartphone (bzw. PDA) Videoinhalte suchen und wiedergeben, die sich auf einem zentralen Video - Server befinden und auf Abruf zum mobilen Gerät überspielt werden. Der Benutzer gibt hierfür

auf seinem Gerät frei gewählte Suchbegriffe ein anhand derer ihm eine Liste verfügbarer Videoinhalte angezeigt wird die seinen Präferenzen möglichst genau entsprechen soll. Die letzten Endes gewählte Datei wird dann auf das mobile Endgerät transferiert und von diesem abgespielt.

4.2 Problemstellungen

Suche und Verwaltung von Videoinhalten Im Allgemeinen geht einer Wiedergabe von Videos auf mobilen Endgeräten eine Suche danach voraus. Dabei ist es unerheblich ob sich das Video zum Zeitpunkt der Suche bereits auf dem Endgerät befindet oder noch auf einem entfernten Server. Die Suche nach Videoinhalten auf einem mobilen Endgerät ist mit einigen Problemen verbunden die meist auf technische Restriktionen des Endgeräts zurückzuführen sind. Die Geräte weisen oft Einschränkungen bezüglich Ausführungsgeschwindigkeit, Energieverbrauch sowie Bilddarstellung auf [4]. Das Problem ist demnach, trotz der genannten Restriktionen eine für den Benutzer komfortable, intuitive und schnelle Möglichkeit zu schaffen gezielt nach Videoinhalten zu suchen. Zudem besteht die Problematik dass ein Benutzer einfach Suchbegriffe eingibt die seiner Meinung nach für das Video kennzeichnend sind [5]. Demnach ist die Suche subjektiv und es ist möglich dass die Suche erfolglos ist, obwohl der Film existiert. Hier entsteht also eine Diskrepanz zwischen den systematisch generierten Daten von Videoinhalten und der subjektiven Beschreibung des Benutzers. Dieses Delta zu minimieren ist ein wesentliches Ziel, möchte man eine Volltextsuche in mobilen Endgeräten zur Verfügung stellen. Eine andere Möglichkeit wäre, den Benutzer anhand von Genren wie "Action", "Dokumentation", "News", etc. sein Suchergebnis minimieren zu lassen [4]. Hierbei könnte es allerdings sein, dass das Suchergebnis dennoch zu viele Treffer enthält und eine weitere Eingrenzung der Ergebnisliste mühselig wird. Wenn das gesuchte Video gefunden wurde stellt sich die Frage was dann mit dem Video passieren soll. Abhängig vom mobilen Endgerät des Benutzers hätte er zum einen die Möglichkeit das Video zu streamen [4], oder zunächst auf dem Endgerät zu speichern um das Video bei Bedarf anzusehen. Im zweiten Fall wäre zu entscheiden ob das Endgerät das Video automatisch in einem vorgeschriebenen Ordner seiner Dateistruktur abspeichert [5], oder ob der Benutzer den Ort für das Video selbst auswählen darf. Hier stellt sich erneut die Frage welche Option es dem Benutzer am leichtesten erlaubt nach seinen Videoinhalten zu suchen sobald sie abgespeichert wurden. Um den Suchprozess hinsichtlich der knapp bemessenen Ressourcen des mobilen Endgeräts zu optimieren, ist intelligente Zusammenfassung von Videoinhalten in Form von Metadaten erforderlich [4].

Segmentierung von Videoinhalten Die Notwendigkeit der Segmentierung eines Videos ergibt sich aus der Anforderung Videoinhalte effizient und sinnvoll mit Metadaten zu versehen, und aufgrund derer Suchprozesse zu präzisieren. Die Metadaten eines kurzen Videos können Informationen über den Videoinhalt als Ganzes enthalten. Handelt es sich jedoch um ein längeres Video, kann dessen Segmentierung in kleinere Einheiten den Prozess der Metadateneingabe vereinfachen [4]. Hierzu werden verschiedene Clustering - Algorithmen benutzt

um die Auswahl von repräsentativen Schlüsselbildern aufgrund ihrer visuellen Eigenschaften zu optimieren [4]. Es wurden jedoch auch andere Methoden entwickelt um semantisch wichtige Ereignisse innerhalb eines Videos zu isolieren. Aizawa et. al. beispielsweise benutzen Gehirnwellen um spannende Momente im betrachteten Videomaterial zu identifizieren [4], und das Video anhand dessen zu segmentieren. Die Summe der Metadaten zu allen Schlüsselbildern ergibt dann eine wünschenswerte Zusammenfassung des Videos [4].

Darstellung von Videoinhalten Mit der Anzahl der verschiedenen mobilen Endgeräte wächst auch die Anzahl verschiedener technischer Möglichkeiten der einzelnen Geräte. Sie unterscheiden sich in Akkulaufzeit, Ausführungsgeschwindigkeit, Bildschirmgröße und der Fähigkeit Videos zu dekodieren [4]. Diese Problematik wirft die Frage auf, ob die Wiedergabe eines Videos auf einem mobilen Endgerät überhaupt beginnen soll, wenn es aufgrund fehlender Funktionalität nicht in der Lage ist dieses korrekt darzustellen. Mobile Endgeräte haben in der Regel eine kleine Bildauflösung, sowohl was die Größe des Bildschirms als auch was die Farbtiefe angeht. Zudem haben einige Geräte wie Palm-OS PDAs keine Audio - Wiedergabe Funktionalität. Wenn beispielsweise im Decoder keine Audiodaten vorhanden sind könnte man stattdessen das Video mit mehr Textinformationen versehen, oder Videos mit Untertitel zur Verfügung stellen [4].

De-/Kompression von Videos Echtzeit-Kompression oder Dekompression einfacher Multimedia Standards kann auf mobilen Endgeräten mit technisch eingeschränkten Möglichkeiten sehr schwierig bis unmöglich werden [4]. Auch wenn die Prozessorgeschwindigkeit mobiler Endgeräte in Zukunft erhöht werden kann bleibt die Energieversorgung der Geräte ein Faktor der Multimedia Anwendungen auf den Endgeräten beeinträchtigt. Laut [4] werden PDA der nächsten Generation Prozessoren mit geringerem Energieverbrauch enthalten wie beispielsweise Intels StrongARM der dem Betriebssystem ermöglicht Akkulaufzeit zu sparen indem sowohl seine Taktfrequenz als auch die seine Betriebsspannung gesenkt wird.

4.3 Software Lösungen

Video Semantic Summarization System Belle L. Tseng et al. stellen in [4] ein semantisches System zur semantischen Zusammenfassung von Videoinhalten (engl.: Video Semantic Summarization System, kurz: V3S) vor, das sie entworfen und implementiert haben. V3S ist ein Beispiel für ein System mit dessen Hilfe man das Anwendungsszenario (siehe 4.1) umsetzen könnte. V3S besteht aus drei Modulen die im folgenden näher beschrieben werden.

Client Der Client ist auf dem mobilen Endgerät installiert. Belle L. Tseng et. al benutzen in [4] einen Palm OS Emulator um einen PDA zu simulieren. Mit Hilfe des Client ist der Benutzer in der Lage nach Videoinhalten zu suchen. Er kann Suchkriterien in Form von bevorzugten Themengebieten, Schlüsselwörter oder auch zeitliche Vorgaben bzgl. der Länge des Videos in sein Endgerät eingeben.

Zeitgleich werden automatisch verschiedene Clientprofile an das Kontrollmodul der Video Middleware gesendet, wie Benutzerprofil (z.B. englisch sprechender Amerikaner, Benutzer wohnt in New York City), Geräteprofil (z.B. Palm IIIc mit eingeschränktem Farbbildschirm, keine Audiowiedergabe, Speicherkapazität verfügbar) und Übertragungsprofil (Internetzugang über 56K Modem) [4]. Aufgrund dieser Daten wählt die Video Middleware dynamisch die für das mobile Endgerät qualitativ beste Videoversion aus, und sendet dem Client diese. Das empfangene Video wird dann über den Video Client auf dem mobilen Endgerät wiedergegeben [4]. Abbildung 1 zeigt die Benutzeroberfläche des Client auf einem Palm OS PDA in drei verschiedenen Szenarios. In Abbildung 1(a) kann der Benutzer aus verschiedenen Kanälen von Videoinhalten auswählen. Das nächste Szenario (Abbildung 3(b)) zeigt wie der Benutzer mittels des Client bevorzugte Themengebiete auswählen kann. Ähnlich dazu kann in Abbildung 3(c) Schlagworte über die Benutzeroberfläche des Client eingeben nach denen dann die Videoinhalte gesucht werden. Im Folgenden wird ein weiteres Modul von V3S, der Datenbank Server genauer betrachtet.



Abbildung 1. Benutzeroberfläche des Client aus [4].

Datenbank Server Der Datenbank Server versorgt die Video Middleware sowohl mit Beschreibungen zu Videos als auch mit deren Pfade damit die Middleware die zugehörigen Videoinhalte findet. Im Datenbank Server werden Videoinhalte gespeichert, analysiert und mit MPEG-7 Metadaten versehen. Da die Video Middleware verschiedene Videoformate, Bitraten und Auflösungen unterstützt, kann die Datenbank Videoinhalte in den gängigen Formaten (z.B. MPEG-1/2, AVI und QuickTime) speichern [4]. Um die Metadaten zu den Videos zu erzeugen existieren Applikationen mit deren Hilfe man entweder manuell Beschreibungen zu Videos eingeben kann, oder diese (halb-) automatisch generiert werden [4]. Diese Beschreibungen werden als MPEG-7 XML Dateien gespeichert. Diese Da-

teilen identifizieren den Inhalt der zugrunde liegenden Videos [4]. Diese können bei einer Suchanfrage an die Video Middleware gesendet werden.

Video Middleware Dieses Modul stellt die Schnittstelle dar zwischen dem Client und dem Datenbank Server [4]. Es besteht aus dem Semantik Transcoder und dem Übertragungsserver. Der Transcoder leitet die vom Benutzer in den Client eingegebene Suchanfrage an den Datenbank Server weiter und erhält von diesem als Antwort das Resultat in Form von MPEG-7 Metadaten. Diese werden dann zunächst zum MPEG-7 Parser gesendet der Teil des Semantik Transcoders ist. Hier werden dann die gewünschten Suchergebnisse aus den Metadaten extrahiert [4]. Diese Ergebnisse identifizieren die dazugehörigen semantischen Beschreibungen und deren Video Segmente. Nun werden die gefundenen Videos zusammengefasst und an den Client übertragen. Der Übertragungsserver wurde für das drahtlose Streamen von Videoinhalten auf unterschiedliche mobile Endgeräte konzipiert [4]. Er erhält zum Zeitpunkt der Suchanfrage gleichzeitig die Clientprofile (siehe Client). Mit Hilfe dieser Informationen wird festgelegt in welchem Format bzw. in welcher Kompressionsart die gewünschten Videoinhalte an den Client gesendet werden, sodass dieser sie in möglichst guter Qualität wiedergeben kann. Abschließend werden die Videoinhalte vom Übertragungsserver an den Client gesendet [4]. Abbildung 2 bietet eine Übersicht des Video Semantic Summarization System und zeigt dessen modularen Aufbau und die Zusammenhänge zwischen den Modulen.

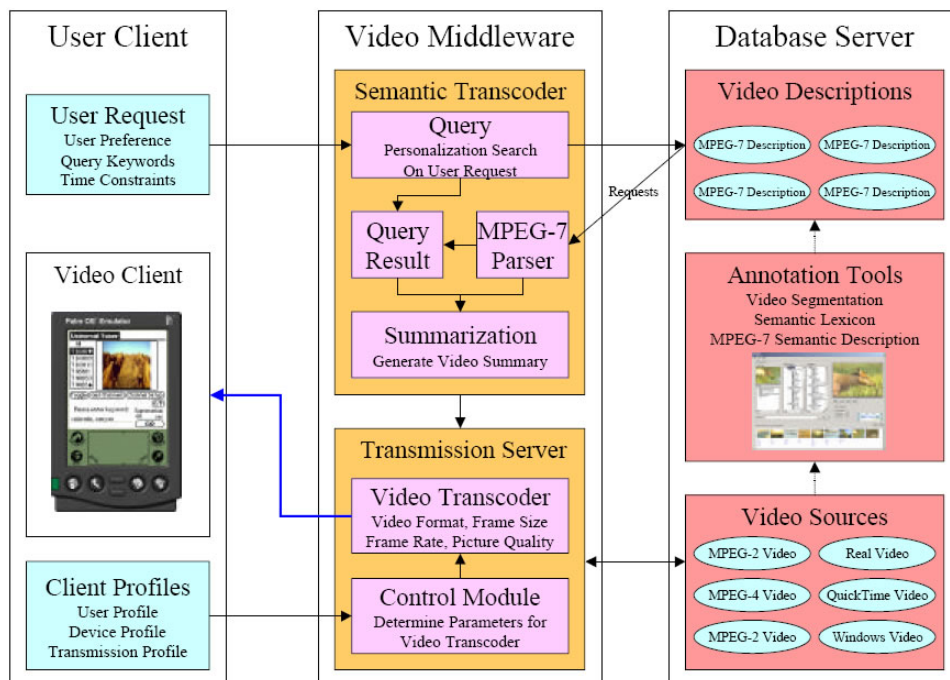


Abbildung 2. Video Semantic Summarization System (Blockdiagramm) aus [4].

5 Bilder auf mobilen Endgeräten

5.1 Anwendungsszenario

Eine Benutzerin ist auf einer Hochzeit eingeladen, und im Laufe des Abends nimmt sie mit ihrem mobilen Endgerät Fotos auf [6]. Ihr Endgerät ist in der Lage automatisch sowohl Datum, Uhrzeit und Ort der Fotoaufnahme zu ermitteln, als auch andere mobile Endgerät die über Bluetooth erreichbar sind zu erkennen. Diese Informationen werden dann vom Endgerät dem aufgenommenen Foto als Metadaten zugewiesen (siehe Abbildung 3) [6]. Die Benutzerin macht Aufnahmen von Gästen und einige Fotos extra für ihre Freundin die derzeit ihre eigene Hochzeit für nächstes Jahr plant. Die Hochzeitstorte fällt der Benutzerin besonders auf, und möchte deshalb von der Mutter der Braut wissen wo diese bezogen wurde. Diese Information notiert die Benutzerin in ihrem Endgerät und weist diese dem Foto und Datum zu [6]. Zudem möchte die Benutzerin ihre schönsten Fotos auch auf einen Webserver übertragen, um ihr persönliches Online - Fotoalbum zu aktualisieren [7]. Im Folgenden werden Problemstellungen und Lösungsansätze für dieses Anwendungsszenario vorgestellt.

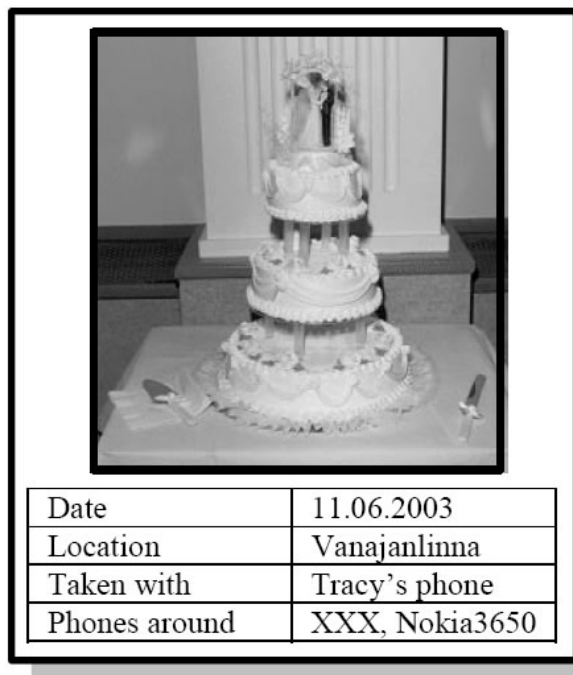


Abbildung 3. Beispiel für Kontext Metadaten aus [6].

5.2 Problemstellungen

Suche und Verwaltung von Bildern Ein effizientes Content Management System ist nötig um Benutzern die Verwaltung ihrer multimedialen Inhalte zu ermöglichen. Dies beinhaltet die Teilprozesse Erstellung, Anzeige/Wiedergabe,

Archivierung, Suche und Finden der Inhalte [5]. Das heißt dass mobile Endgeräte in der Lage sein müssen Metadaten der multimedialen Inhalte dynamisch zu verwalten und zu ändern [5]. Derzeit ist die Verwaltung von multimedialen Daten auf mobilen Endgeräten Dateibasiert und nicht eingängig [5], demnach nicht für die Verwaltung großer Mengen multimedialer Inhalte geeignet [5]. Der Benutzer hat hier das Problem statische Dateistrukturen zu schaffen, und sich diese dann, im Falle der Suche nach seinen Inhalten, wieder ins Gedächtnis zu rufen [5]. Dies gilt nicht nur für die Suche von Inhalten, sondern auch für die Verwaltung derer, da einem Verwaltungsprozess wie beispielsweise Löschen oder Kopieren eines Inhalts eine vorherige Suche danach zugrunde liegt. Im Gegensatz zum Dateibasierten Ansatz können Kontext Metadaten multimedialer Inhalte semantisch informativere Daten enthalten die von Content Management Systemen verwendet werden können [5] um mehr Effizienz zu gewährleisten.

Versand von Bildern in das Internet Eine Studie von Marc Davis et. al. ergibt in [7] dass die Schwierigkeit darin besteht Fotos vom mobilen Endgerät in das Internet zu versenden um diese effizient zu verwalten bzw. mit anderen Benutzern zu teilen. Das Kernproblem ist dass es zwar zahlreiche, aber bei Weitem keine einfache Möglichkeiten gibt diese Hürde zu überwinden. Sowohl die Vielfalt an Übertragungsmöglichkeiten - Email, Bluetooth, Infrarot, USB Kabel, MMS - als auch der Mangel einer einheitlichen Empfangsschnittstelle für diese Übertragungstechniken führen zur Verwirrung der Benutzer [7]. Diese Tatsache führt bei vielen Benutzern dazu, dass sie ihre multimedialen Inhalte nicht vom mobilen Endgerät ins Internet verschicken können. [7] Sie können ihre Inhalte nur mit anderen Benutzern austauschen.

Bildaustausch mit anderen Benutzern Genau wie beim Versand von Bildern in das Internet besteht auch beim Austauschen multimedialer Inhalte mit anderen Benutzern das Problem der Vielfalt an Übertragungsmöglichkeiten und dem Mangel an einheitlichen Verfahren Bilder zwischen Kamerahandys auszutauschen [7]. Bei den meisten mobilen Endgeräten mit integrierter Kamera kommen weitere Problematiken hinzu: Sowohl die Hardware als auch die Software zweier kommunizierender Endgeräte können inkompatibel sein. Es könnte beispielsweise sein, dass der Empfänger keine MMS empfangen kann weil sein Endgerät diese Funktionalität nicht unterstützt [7]. Zudem kann ein Austauschvorgang zeitintensiv sein, da oft eine Email an ein persönliches Web Postfach geschickt werden muss bevor man von dort aus dann den Versand des Fotos an den Empfänger initiieren kann. Ein weiterer Aspekt der den Bildaustausch zwischen mobilen Endgeräten erschwert ist die Unwissenheit der Benutzer bezüglich des Kostenmodells bei einem Bildtransfer (Wer zahlt für was wie viel?) [7]. Im Folgenden werden verschiedene Software Lösungen vorgestellt die sich mit diesen Problemstellungen auseinandersetzen.

5.3 Software Lösungen

Media Assistant Der Media Assistant ist eine Applikation die von Motorola Labs [5] konzipiert und implementiert wurde. Sie ist ein Beispiel dafür wie man

effizient nach Bildern auf mobilen Endgeräten suchen kann. Media Assistant ist eine Java Applikation die auf einer Plattform mit einem 400MHz Intel XScale Prozessor und 64 MB Speicher läuft. Die Benutzeroberfläche wird von einem Touchscreen mit einer Auflösung von 320x240 Pixel angesprochen [5]. Wird ein Bild aufgenommen, ordnet Media Assistant dem Bild die Uhrzeit bei Aufnahme des Bildes sowie den aktuellen Ort hinzu, indem diese Informationen in eine Metadaten Datenbank geschrieben werden. Der Benutzer kann dann noch optional andere Informationen zum Bild angeben wie Titel, oder die Namen der Personen auf dem Foto, die dann ebenfalls in der Datenbank gespeichert werden [5]. So muss der Benutzer das Bild zum einen nicht manuell in einer hierarchischen Dateistruktur in einem Ordner abspeichern, und zum anderen muss er für das Bild keinen Dateinamen erfinden um es später in einer Liste von Dateinamen wieder zu finden [5]. Folgendes Szenario aus [5] wäre mit Hilfe dieses Tools möglich: Angenommen es wurden einige Bilder während eines Familienurlaubs in Boston im Jahr 2002 zur Weihnachtszeit aufgenommen. Einem Benutzer könnten bei der Suche nach einem bestimmten Bild davon Stichworte wie *Reise nach Boston, Dezember 2002, Reise nach Boston irgendwann nach 2000*, oder viele andere einfallen. Der Media Assistant erlaubt dem Benutzer beliebige Kriterien an die er sich erinnert auszuwählen, und gibt ihm dann eine Liste von Werten zu diesen Kriterien aus. Durch Auswahl aus dieser Liste wird dem Benutzer die manuelle Eingabe von Werten mittels der meist spärlichen Tastatur des Endgeräts erspart [5]. Sollte der Benutzer also als Kriterium Jahr auswählen, würde das Tool ihm unter anderem als auswählbaren Wert in einer Liste 2002 anzeigen. Wählt er das Jahr dann aus, sieht er in Mineaturansicht alle Bilder die in dem Jahr aufgenommen wurden. Sollte die Ergebnisliste der Bilder immer noch zu groß sein, kann der Benutzer ein anderes Kriterium wie Ort angeben um seine Ergebnisliste weiter einzugrenzen. Sobald der Benutzer sein gesuchtes Foto oder seine gesuchten Fotos gefunden hat, ist er in der Lage das Bild bzw. die Bilder mit dem Tool an ein anderes Endgerät zu versenden. Werden Daten versendet, dann werden sie mitsamt ihrer MPEG-7 Metadaten versendet, sodass das der Benutzer des empfangenden Geräts bei der Suche nach einem Foto ebenfalls von den übermittelten Metadaten profitieren kann [5]. Der Media Assistant ist kompatibel zu Metadaten Formaten wie MPEG-7 MDS, EXIF und ID3. Basierend auf dieser Software erhoffen sich die Entwickler bei Motorola Labs (Bhavan Gandhi et. al.) komplexere Systeme zu entwickeln die einen reibungslosen Austausch multimedialer Inhalte und deren Metadaten in einem heterogenen Netzwerk mobiler Endgeräte gewährleisten [5].

MMM Risto Sarvas und sein Team haben hat am HIIT (Helsinki Institute for Information Technology) eine Software namens MMM (Media Metadata Creation with Mobile Phones) entwickelt, die der automatischen Generierung von Metadaten zu Multimediadaten wie Bildern dient [8]. MMM ist zudem auch in der Lage größere Datenmengen automatisch zu verwalten. Dieses System kombiniert sowohl die automatische Generierung von Metadaten zu multimedialen Inhalten mit der Möglichkeit den Benutzer manuell Metadaten hinzuzufügen, um Vorteile aus beiden Methoden zu vereinen [8]. Unmittelbar nach Aufnahme

eines Fotos werden Kontext Metadaten extrahiert. Das ist technisch möglich da das mobile Endgerät Zugriff auf Informationen wie Ort (durch die Zellen ID), Uhrzeit, Datum und Benutzername des Benutzers (z.B. Telefonnummer des Benutzers) hat [8]. Diese automatisch gesammelten Daten werden dem aufgenommenen Foto zugewiesen. Zusätzlich wird der Benutzer aufgefordert das Thema des Fotos einzugeben (z.B. Person, Ort, Objekt, Aktivität oder anderes) [8]. Das ausgewählte Thema ist relevant für den serverseitigen Prozess der Metadatenerstellung. Alle gesammelten Daten zum aufgenommenen Foto werden dann inklusive Foto an den Server gesendet. Auf dem Server werden dann die empfangenen Daten benutzt um weitere Metadaten automatisch zu generieren. Dort werden bereits abgespeicherte Fotos und deren Metadaten mit dem neuen Foto und dessen Beschreibung verglichen um Gemeinsamkeiten zu finden [8]. Handelt es sich bei dem aktuellen Foto beispielsweise um eine Person, und hat der Benutzer zuvor schon Fotos von Personen aufgenommen, dann ist die Wahrscheinlichkeit höher dass es sich beim neuen Foto um eine vom System bereits "bekannte" Person handelt [8]. Ein anderes Beispiel ist die Aufnahme einer Sehenswürdigkeit. Die Idee besteht darin dass potentiell mehrere Personen ein und dieselbe Sehenswürdigkeit an einem bestimmten Ort fotografieren. Ein großer Vorteil von MMM ist die Tatsache dass alle Fotos und deren Metadaten auf dem Server allen Benutzern zur Verfügung stehen. Das Beispiel der Sehenswürdigkeit zeigt dass obwohl verschiedene Benutzer in keiner sozialen Beziehung stehen (sich beispielsweise nicht kennen) sie dennoch Gemeinsamkeiten haben bezüglich Ort und Motiv des aufgenommenen Fotos. Auf diese Art werden Metadaten von einem Benutzer eingegeben von denen jeder andere Benutzer profitieren könnte [8]. Der serverseitige Prozess der Metadatenerstellung durch Vergleich mit bereits gespeicherten Inhalten basiert trotz allem auf "Vermutungen" [8]. Manche von diesen Vermutungen treffen hundertprozentig zu (z.B. Tag der Woche oder Ort einer Zellen ID), während andere vom System als unsicher klassifiziert werden. In dem Fall werden diese Vermutungen zum Benutzer auf sein mobiles Endgerät zurückgeschickt damit der Benutzer diese entweder bestätigen oder aber ablehnen kann. So wird der automatisierte Prozess auf Serverseite mit der Interaktion des Benutzers kombiniert [8]. Um dem Benutzer die Auswahl der fragwürdigen Vermutungen komfortabel zu machen werden ihm diese in einer drop-down Liste angezeigt, wobei die wahrscheinlichste Vermutung an oberster Stelle steht [8]. Die vom Benutzer überprüften und bewerteten Vermutungen werden dann an den Server zurückgeschickt der mit Hilfe dieser Daten den Metadaten Erstellungsprozess fortführen kann. Diesmal kann der Server andere Algorithmen anwenden oder wahrscheinlichere Vermutungen generieren da der Benutzer ihm genauere semantische Informationen geliefert hat. Dieses Verfahren wird solange fortgeführt bis der Server keine weiteren Vermutungen generiert oder der Benutzer entscheidet den Prozess zu beenden [8]. Abbildung 4 zeigt den Ablauf des Prozesses anhand eines Beispiels. MMM wurde primär für die automatische Erstellung von Metadaten zum Zeitpunkt der Aufnahme von Fotos konzipiert. Das System unterstützt keinen Bildaustausch mit anderen Benutzern. Es sollte eine Grundlage für andere Systeme sein die von den Ergebnissen von MMM profitieren.

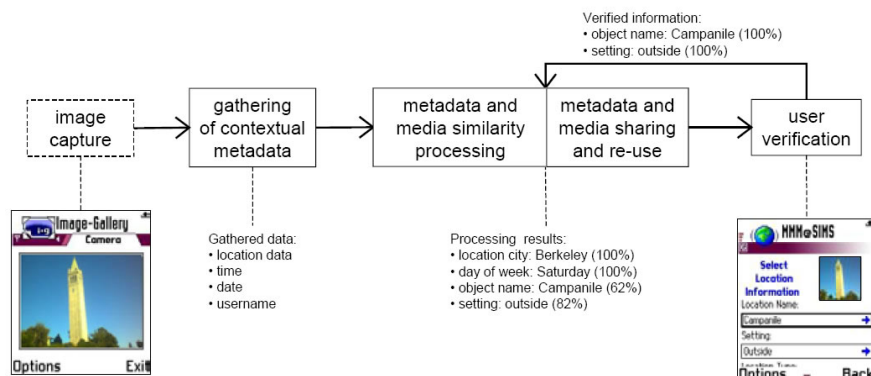


Abbildung 4. Übersicht des Metadaten Erstellungsprozesses [8].

MMM2 MMM2 (Mobile Media Metadata for Media Sharing) wurde von Marc Davis et al. [7] entwickelt und ist ein Tool das den Fokus nicht nur auf die Metadaten Generierung legt, sondern auch auf den anschließenden Austausch von Bildern mit anderen Benutzern (oder eher ihren mobilen Endgeräten) [7]. Schon im Jahr 2003 wurde die Vorgängerversion MMM1 (nicht zu verwechseln mit MMM) entwickelt und getestet. Die Applikation lief auf Nokia 3650 Kamerahandys und es wurden 60 Testbenutzer über einen Zeitraum von vier Monaten zur Kooperation gebeten. Damals lag das Hauptaugenmerk darin Metadaten zu generieren. Die Ergebnisse dieses Tests ergaben dass der Austausch von Bildern ein wesentlicher Bestandteil eines mobilen Medien Lebenszyklus ist. Beim Übergang von MMM1 zu MMM2 ging man mehr und mehr weg von der Frage: Was habe ich gerade fotografiert? Vielmehr lautete die Frage nun: Mit wem möchte ich das fotografierte Bild austauschen/teilen? [7] MMM2 wurde entworfen um sich mit drei Problemen auseinanderzusetzen:

- Fotos von dem mobilen Endgerät ins Web zu versenden
- Fotos zu finden und zu verwalten
- Fotos mit anderen auszutauschen

Das erste Problem der Vielfalt an Transfermöglichkeiten wurde bereits erläutert (siehe *Versand von Bildern in das Internet*). Das zweite Problem lässt sich wie folgt erklären: Benutzer nehmen viele Fotos auf, während die technischen Rahmenbedingungen wie kleiner Bildschirm, langsames scrolling und umständliche Texteingabe es dem Benutzer erschweren ein bestimmtes Foto zu finden wenn die Anzahl an Fotos ein überschaubares Limit überschreitet. Eng mit diesem Problem verbunden ist auch die Tatsache dass ein Benutzer der Fotos an ein anderes seiner Endgeräte sendet, seine Fotosammlung auf verschiedene Endgeräte verteilt und damit auf keinem die Sammlung vollständig ist. Beim dritten Problem, dem Austausch von Fotos kommt wieder die Inkompatibilität der unterschiedlichen Hardware/Software Kombinationen zum tragen (siehe *Bildaustausch mit anderen Benutzern*). MMM2 bezieht zum Zeitpunkt der Bildaufnahme automatisch Daten wie: Räumlichen Kontext (z.B. durch Zellen ID oder GPS Modul), zeitlichen Kontext (z.B. durch Aktualisierung vom Netzwerk Zeitserver) oder sozialen Kontext (z.B. Erkennung der mobilen Endgeräte mit

aktiviertem Bluetooth und anschließender Vergleich mit der Kontaktliste). Danach kann der Benutzer einen Titel zur Ergänzung der Metadaten eingeben und (oder) das Bild mit einem anderen Benutzer austauschen (siehe Abbildung 5). Währenddessen wird das Bild im Hintergrund auf einen Webserver hochgeladen der die hochgeladenen Bilder des Benutzers in einer Art persönlicher Bibliothek oder persönlichem Album verwaltet [7]. Entscheidet sich der Benutzer zum Austausch des Bildes mit einem anderen Benutzer, öffnet sich auf dem mobilen Endgerät ein Webbrowser der den Benutzer auf eine personalisierte Share - a - photo Seite verweist. Von dieser Seite aus hat der Benutzer die Möglichkeit mit einer Vielzahl vom System vorgeschlagener Empfänger das Bild auszutauschen. Die möglichen Empfänger werden aus einem eigens angelegten Adressbuch extrahiert. Hier ist es ebenfalls möglich mehrere Empfänger in einer Art Verteilerliste zusammenzufassen. Zudem werden, nach Einholung der Erlaubnis vom Benutzer, der Liste dynamisch Empfänger hinzugefügt die sich zum Zeitpunkt des Hochladens in Bluetooth - Reichweite befinden [7].

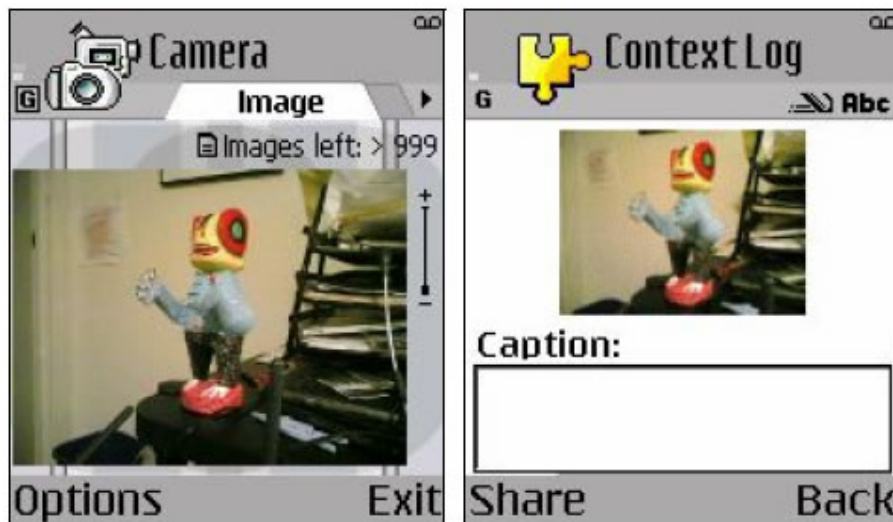


Abbildung 5. Modifizierte Nokia Oberfläche aus [7].

5.4 Studie über das Nutzerverhalten mit MMM2

Verglichen mit dem MMM1 Projekt aus dem Jahr 2003 bei dem das Nokia 3650 zum Einsatz kam, haben Testbenutzer mit MMM2 auf Nokia 7610 Endgeräten im Durchschnitt viel mehr Fotos hochgeladen wie man Abbildung 6 entnehmen kann. Dieser erhebliche Nutzungsanstieg wurde auf einige Faktoren zurückgeführt: Die Bildqualität wurde erhöht (von VGA auf 1 Megapixel Bildauflösung inklusive Nachtmodus für schlechte Lichtverhältnisse und digitalem Zoom). Zudem kommt hinzu dass beim Test mit MMM2 mehr Testbenutzer bereits vor der Studie ein Kamerahandy besaßen als bei der ersten Studie 2003 (Verhältnis 12:1). Ein weiterer Faktor ist das automatische Hochladen in die Foto -

Verwaltungs - Applikation im Web. Nicht zuletzt war auch die automatische Unterstützung zum Austausch von Bildern im Web oder vor Ort durch MMM2 für den regen Gebrauch verantwortlich [7]. Abbildung 7 zeigt eine allgemeine Statistik wieviele Fotos in einem bestimmten Zeitraum hochgeladen, empfangen und ausgetauscht wurden

STATS	MMM1	MMM2	DIFF
Users	38	40	5%
Days	63	39	-38%
Raw totals			
Personal photos uploaded	155	1478	854%
Total photos uploaded	535	1678	214%
Photos not uploaded	108	52	-52%
Average per user per day			
Personal photos uploaded	0.06	0.95	1363%
Total photos uploaded	0.22	1.08	381%
Photos not uploaded	0.05	0.03	-26%
Upload failure rate	16.8%	3.0%	-82%

Abbildung 6. MMM1 im Vergleich zu MMM2 Upload Statistiken aus [7].

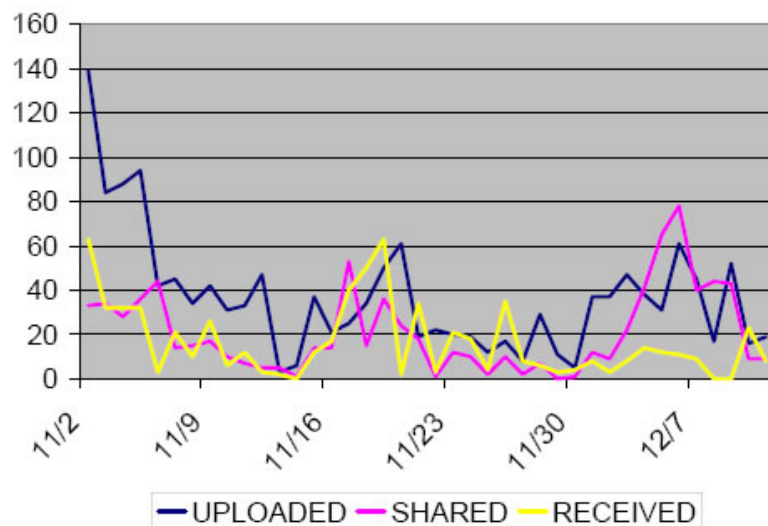


Abbildung 7. Aufgenommene, hochgeladene, empfangene und ausgetauschte Fotoanzahl aus [7].

6 Fazit

Es ist ersichtlich dass mit der zunehmenden Ausstattung mobiler Endgeräte mit größeren Kapazitäten (Akku und Speicher) und Sensoren (Temperatur, GPS, Biometrische Daten des Benutzers) viele der benötigten Metadaten zu aufgenommenen Multimediadaten automatisch generiert werden können. Das bringt dem Benutzer aber in den meisten Fällen kein Mehrwert. Heutzutage erzielt man mit einem Content Management System die besten Ergebnisse, wenn die Metadaten auch für den Benutzer repräsentativ strukturiert sind. Das macht das Eingreifen des Benutzers in die Generierung nötig. Wie man anhand der vorgestellten Lösungen sieht, kann man nur versuchen dem Benutzer die Eingabe und allgemein die Interaktion zwischen mobilem Endgerät und ihm zu vereinfachen und (oder) zu beschleunigen, ganz abnehmen kann man ihm den Aufwand jedoch nicht, sodass die Verwaltung seiner Inhalte ähnlich komfortabel bleibt.

Literatur

1. WIKIPEDIA: Endgerät. <http://de.wikipedia.org/wiki/Endger%C3%A4t> (2005)
2. WIKIPEDIA: Contentmanagement. http://de.wikipedia.org/wiki/Content_Management (2005)
3. WIKIPEDIA: Metadaten. <http://de.wikipedia.org/wiki/Metadaten> (2006)
4. B. L. Tseng, C.-Y. Lin, J.R.S.: Video summarization and personalization for pervasive mobile devices. SPIE Electronic Imaging 2002 - Storage and Retrieval for Media Databases (2002)
5. B. Gandhi, A. Martínez, F.B.: Intelligent multimedia content management on mobile devices. ICME '04. 2004 IEEE International Conference on **3** (2004) 1703–1706
6. A. Sorvari, J.J.e.a.: Usability issues in utilizing context metadata in content management of mobile devices. Proceedings of NordiCHI '04 (2004) 357–363
7. et al, M.D.: Mmm2: Mobile media metadata for media sharing. Proceedings of 13th Annual ACM International Conference on Multimedia (2005)
8. Sarvas, R.: Media content metadata and mobile picture sharing. Proceedings of the 11th Finnish Artificial Intelligence Conference STeP 2004 (2004)