MediAssist: Using Content-Based Analysis and Context to Manage Personal Photo Collections

Neil O'Hare[1], Hyowon Lee[1], Saman Cooray[1], Cathal Gurrin[1], Gareth J.F. Jones[1], Jovanka Malobabic[1], Noel E. O'Connor[1,2], Alan F. Smeaton[1,2], and Bartlomiej Uscilowski[1]

[1] Centre For Digital Video Processing, Dublin City University, Ireland.
[2] Adaptive Information Cluster, Dublin City University, Ireland
nohare@computing.dcu.ie
http://www.cdvp.dcu.ie

Abstract. We present work which organises personal digital photo collections based on contextual information, such as time and location, combined with content-based analysis such as face detection and other feature detectors. The MediAssist demonstration system illustrates the results of our research into digital photo management, showing how a combination of automatically extracted context and content-based information, together with user annotation, facilitates efficient searching of personal photo collections.

1 Introduction

Recent years have seen a revolution in photography with a move away from analog film towards digital technologies, resulting in the accumulation of large numbers of personal digital photos. The MediAssist [4] project at the Centre for Digital Video Processing (CDVP) is developing tools to enable users to efficiently search their photo archives. Automatically generated contextual metadata and content-based analysis tools (face and building detection) are used, and semi-automatic annotation techniques allow the user to interactively improve the automatically generated annotations. Our retrieval tools allow for complex query formulation for personal digital photo collection management. Previous work has reported other systems which use context to aid photo management. Davis et al [1] utilise context to recommend recipients for sharing photos taken with a context-aware phone, although their system does not support retrieval. Naaman et al [6] use context-based features for photo management, but they do not use content-based analysis tools or allow for semi-automatic annotation.

2 Content and Context-Aware Photo Organisation

MediAssist organises photo collections using a combination of context and content-based analysis. Time and location of photo capture are recorded and used to derive additional contextual information such as daylight status, weather and

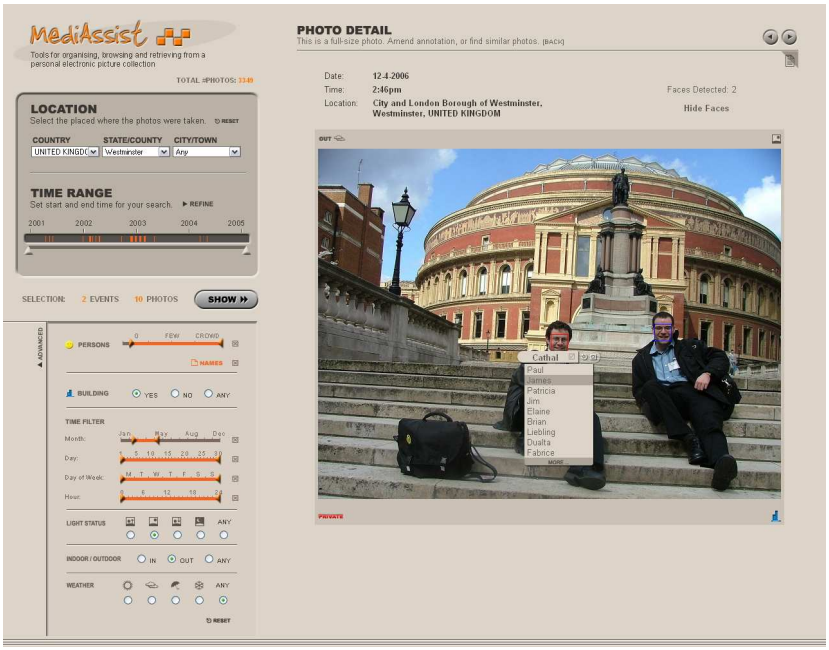

Fig. 1. The MediAssist Photo Management System

indoor/outdoor classification [4]. By using this information the browsing space when seeking a particular photo or photos can be drastically reduced. We have previously shown the benefits of using location information in personal photo management [4]. The MediAssist photo archive currently contains over 14,000 location-stamped photos taken with a number of different camera models, including camera phones. Over 75% of these images have been manually annotated for a number of concepts including buildings, indoor/outdoor, vehicles, animals, babies and the presence and identity of faces, serving as a ground truth for evaluation of our content-based analysis tools.

Our face detection system is built on both appearance-based face/non-face classification and skin detection models [2]. The algorithm detects frontal-view faces at multiple scales, with features supported to detect in-plane rotated faces. We also detect the presence of large buildings in outdoor digital photos, approaching the problem as a building/non-building classification of the whole image using low-dimensional low-level feature representation based on multi-scale analysis and explicit edge detection [5].

3 The MediAssist Web Demonstrator System

The MediAssist Web-based desktop interface allows users to efficiently and easily search through their personal photo collections using the contextual and content-

based information described above. The MediAssist system interface is shown in Fig. 1.

3.1 Content and Context-Based Search

The user is first presented with basic search options enabling them to enter details of desired location (placenames are extracted from a gazetteer) and time, and also advanced options that allow them to further specify features such as people. Slider bars can be used to filter the collection based on the (approximate) number of people required. In addition, the user can specify the names of individuals in the photo based on a combination of automatic methods and manual annotation, as described below. Time filters allow the formulation of powerful time-based queries corresponding to the user's partial recall of the temporal context of a photo-capturing event, for example all photos taken in the evening, at the weekend, during the summer. Other advanced features the user can search by include weather, light status, Indoor/Outdoor and Builing/Non-Building.

3.2 Collection Browsing

In presenting the result photos, four different views are used. The default view is *Event List* which organizes the filtered photos into events in which the photos are grouped together based on time proximity [3]. Each event is summarized by a label (location and date/time) and five representative thumbnail photos automatically extracted from the event. *Event Detail* is composed of the full set of photos in an event automatically organized into sub-events. *Individual Photo List* is an optional view where the thumbnail size photos are presented without any particular event grouping, but sorted by date/time. *Photo Detail* is an enlarged single photo view presented when the user selects one of the thumbnail size photos in any of the above views. Arrow buttons allow jumping to previous/next photos in this view. In all of the above presentation options, each photo (thumbnail size or enlarged) is presented with accompanying tag information in the form of icons, giving the user feedback about the features automatically associated with each photo.

3.3 Semi-Automatic Annotation

MediAssist allows users to change or update any of the automatically tagged information for a single photo or for a group of photos. In *Photo Detail* view, the system can highlight all detected faces in the photo, allowing the user to tidy up the results of the automatic detection by removing false detections or adding missed faces. The current version of the system uses a body patch feature (i.e. a feature modeling the clothes worn by a person) to suggest names for detected faces: the suggested name for an unknown face is the known face with the most similar body patch. The user can confirm that the system choice is the correct one, or choose from a shortlist of suggested names, again based on the body patch

feature. Other work has shown effective methods of suggesting identities within photos using context-based data [7]: in our ongoing research we are exploring the combination of this type of approach with both face recognition and body-patch matching to provide identity suggestions based on both content and context.

4 Conclusions

We have presented the MediAssist demonstrator system for context-aware management of personal digital photo collections. The automatically extracted features are supplemented with semi-automatic annotation which allows the user to add to and/or correct the automatically generated annotations. Ongoing extensions to our demonstration system include the integration of mapping tools. Other important challenges are to leverage context metadata to improve on the performance of content analysis tools [8], particularly face detection, and to use combined context and content-based approaches to identity annotation, based on face recognition, body-patch matching and contextual information.

5 Acknowledgements

The MediAssist project is supported by Enterprise Ireland under Grant No CFTD-03-216. This work is partly supported by Science Foundation Ireland under grant number 03/IN.3/I361

References

1. S. Ahern, S. King, and M. Davis. MMM2: mobile media metadata for photo sharing. In *ACM Multimedia*, pages 267–268, Singapore, November 2005.
2. S. Cooray and N. E. O'Connor. A hybrid technique for face detection in color images. In *AVSS*, pages 253–258, Como, Italy, September 2005.
3. A. Graham, H. Garcia-Molina, A. Paepcke, and T. Winograd. Time as essence for photo browsing through personal digital libraries. In *ACM Joint Conference on Digital Libraries*, pages 326–335, Portland, USA, July 2002.
4. C. Gurrin, G. J. F. Jones, H. Lee, N. O'Hare, A. F. Smeaton, and N. Murphy. Mobile access to personal digital photograph archives. In *MobileHCI 2005*, pages 311–314, Salzburg, Austria, September 2005.
5. J. Malobabic, H. LeBorgne, N. Murphy, and N. E. O'Connor. Detecting the presence of large buildings in natural images. In *Proceedings of CBMI 2005 - 4th International Workshop on Content-Based Multimedia Indexing*, Riga, Latvia, June 2005.
6. M. Naaman, S. Harada, Q. Wang, H. Garcia-Molina, and A. Paepcke. Context data in geo-referenced digital photo collections. In *ACM Multimedia*, pages 196–203, New York, USA, October 2004.
7. M. Naaman, R. B. Yeh, H. Garcia-Molina, and A. Paepcke. Leveraging context to resolve identity in photo albums. In *ACM Joint Conference on Digital Libraries*, pages 178–187, Denver, CO, USA, June 2005.
8. N. O'Hare, C. Gurrin, G. J. F. Jones, and A. F. Smeaton. Combination of content analysis and context features for digital photograph retrieval. In *EWIMT*, pages 323–328, London, UK, December 2005.