# Automatic Text Searching For Personal Photos

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

***Abstract*— This demonstration presents the MediAssist prototype system for organisation of personal digital photo collections based on contextual information, such as time and location of image capture, and content-based analysis, such as face detection and recognition. This metadata is used directly for identification of photos which match specified attributes, and also to create text surrogates for photos, allowing for text-based queries of photo collections without relying on manual annotation. MediAssist illustrates our research into digital photo management, showing how a combination of automatically extracted context and content-based information, together with user annotation and traditional text indexing techniques, facilitates efficient searching of personal photo collections.**

***Index Terms*— Personal Photo Management, Text Search, Context**

## I. Introduction

In recent years digital photography has become increasingly popular, resulting in the accumulation of large numbers of personal digital photos. The MediAssist project [6] at the Centre for Digital Video Processing (CDVP) addresses this situation by developing tools for the efficient searching of photo archives. The system uses both automatically generated contextual metadata (eg. time, location) and content-based analysis tools (eg. face detection and recognition). Semi-automatic annotation allows the user to interactively improve the automatically generated annotations. Retrieval tools allow for complex query formulation, in addition to the facility to create simple text queries, based on these features. In previous work using context for photo management, Davis et al [1] utilised context to recommend recipients for sharing photos taken with a context-aware phone, although their system does not support retrieval. Naaman et al [4] use context-based features for photo management, but they do not use content-based analysis tools, or facilitate semi-automatic annotation or text-based searches. There is also a huge body of work on content-based image retrieval [10], but it has been shown that users to not find this facility useful for personal photo management [9].

## II. Content and Context-Aware Photo Organisation

The MediAssist photo archive contains over 17,000 location-stamped photos taken with a number of different camera models, including camera phones. Over 11,000 of these have been manually annotated for a number of concepts, including buildings, indoor/outdoor and the presence and identity of faces. This manually annotated dataset serves as a ground truth for the evaluation of content-based analysis tools, and also can be used to bootstrap semi-automatic tools (which depend on a certain level of user annotation). All photos are indexed using both context and content-based analysis. Time and location of photo capture are used to derive additional contextual information such as daylight status, weather and indoor/outdoor classification [7]. A face detection system is used to detect the presence of frontal-view faces [2]. Other content-based tools used include body patch (the area under the face, modelling the clothes worn by the individual) feature extraction [2], face recognition using ICA (Independent Component Analysis) and building detection based on the distribution of edges in the image [7]. All of this information can prove very useful for searching photo collections.

All authors are members of the Centre for Digital Video Processing, Dublin City University. Alan F. Smeaton and Noel E. O'Connor are members of the Adaptive Information Cluster, Dublin City University.
email: nohare@computing.dcu.ie

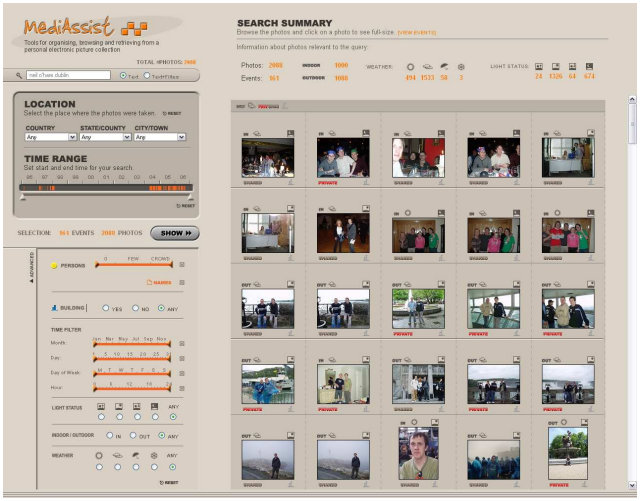


Fig. 1. The MediAssist Photo Management System

## III. The MediAssist Web Demonstrator System

The MediAssist Web-based desktop interface allows users to search through their personal photo collections using the contextual and content-based features described above. The MediAssist system interface is shown in Fig. 1. Our earlier version of the MediAssist prototype supported filter-based searching using the photo metadata features [6]. The new version presented here has been extended to include free-text ranked information retrieval functionality.

### A. Filter-Based Search

The system presents search options enabling a user to enter details of desired locations, times, and advanced options such as people present, weather, light status, indoor/outdoor and

building/non-building. Semi-automatic person identification relies on a combination of automatic methods and manual annotation as described below. Time filters enable powerful time-based queries, for example all photos taken in the evening, at the weekend, during the summer or within certain date ranges.

### B. Text Search Interface

For text-based search the automatic context and content-based features are mined to construct text surrogates for all photos, creating a textual equivalent of each feature (e.g. if the date is October 21th 2006, the text 'october autumn saturday weekend 21 twenty-first 2006' would form a surrogate textual description). So an image might have the text 'dublin ireland september weekend afternoon person alan' associated with it, representing the features location, time, face detection and person annotation. We index the text document associated with an image using a conventional text search engine based on the standard BM25 information retrieval model [8]. We also create text surrogates for 'events' (see below) to allow for text-based searching of events in the 'Event List' view described below. The system presents a text search box to allow for the quick and easy formulation of text queries based on the content and context features described above. We will conduct an evaluation of this search interface in future work.

### C. Collection Browsing

Four different views are available to present the results of searches. The default view, *Event List*, organises the filtered photos into 'events' in which the photos are grouped together based on time proximity, by detecting large temporal gaps between consecutively captured photos, similar to [3]. Each event is summarized by a label (location and date/time) and five representative thumbnail photos selected based on the query. *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. In all of the above presentation options, each photo is presented with its associated automatic annotation information.

### D. Semi-Automatic Annotation

MediAssist allows users to manually change or update any of the automatically tagged information for a single photo or for a group of photos. In *Photo Detail* view, the user can highlight all detected faces in the photo and tidy up the results of the automatic detection by removing false detections or adding missed faces. The system uses a body patch feature (i.e. a feature modeling the clothes worn by a person) combined with a face recognition feature to suggest names for detected faces: the suggested name for an unknown face is the known face with the most similar body patch and face [2]. The user can confirm the system choice or choose from a shortlist of suggested names, again based on similarity. Other work has shown effective methods of suggesting identities within photos using context-based data [5]: in our ongoing research we are exploring the combination of this type of approach with both face recognition and body-patch matching.

## IV. Conclusions

We have presented the MediAssist demonstrator system for context-aware management of personal digital photo collections. Automatically extracted features are supplemented with semi-automatic annotation which allows the user to correct or add to the automatically generated annotations. The system allows the user to formulate precise queries using content and context based features, or alternatively the user can formulate simple text queries, which are enabled without the need for manual annotation. We plan to leverage context metadata to improve on the performance of content analysis tools [7], and we will use combined context and content-based approaches to identity annotation, based on face recognition, body-patch matching and contextual information. We will also extend the integration of person recognition to enable the user to query for a given individual, and the system will (in addition to returning the photos with confirmed annotations) return a ranked list of candidate photos which should contain this person.

## Acknowledgments

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 No 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, N. O'Connor, C. Gurrin, G. Jones, N. O'Hare, and A. F. Smeaton. Identifying person re-occurrences for personal photo management applications. In *VIE 2006*, pages 144–149, Bangalore, India, September 2006.

[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] 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.

[5] 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.

[6] N. O'Hare, C. Gurrin, H. Lee, N. Murphy, A. F. Smeaton, and G. Jones. Digital photos: Where and when? In *ACM Multimedia 2005*, pages 261–262, Singapore, November 2005.

[7] N. O'Hare, H. Lee, S. Cooray, C. Gurrin, G. Jones, J. Malobabic, N. O'Connor, A. F. Smeaton, and B. Uscilowski. Mediassist: Using content-based analysis and context to manage personal photo collections. In *CIVR 2006*, pages 529–532, Tempe, AZ, July 2006.

[8] S. E. Robertson, S. Walker, , S. Jones, M. M. Hancock-Beaulieu, and M. Gatford. Okapi at TREC-3. In *Proceedings of the Third Text REtrieval Conference (TREC-3)*, pages 109–126, NIST, November 1995.

[9] K. Rodden and K. R. Wood. How do people manage their digital photographs? In *CHI 2003*, pages 409–416, Florida, USA, April 2003.

[10] A. Smeulders, M. Worring, S. Santini, A. Gupta, and A. 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.