Content-based Photo Album Management using Faces' Arrangement

Mohamed Abdel-Mottaleb and Longbin Chen
Department of Electrical & Computer Engineering, University of Miami
1251 Memorial Drive, Coral Gables, FL 33146
Email: mottaleb@miami.edu,lchen@umsis.miami.edu

Abstract

Photo album management, an application of content-based image browsing/retrieval, has attracted much attention in recent years. Identities of individuals appearing in the photos are the most important aspect for photo browsing. However, face recognition generally doesn't work effectively in such situation due to the large variations in pose, illumination and sometimes poor quality of images. In this paper, we present a system for browsing photo albums. The system automatically detects and stores information about the locations of faces in the photos. A similarity function based on face arrangement is defined. Photos are then clustered based on the similarity function using a clustering algorithm proposed in this paper. The system also represents photos of an event using a composite image. The composite image is built from representative faces and an image that represents the event. Experiments indicate that the face arrangement features are effective in representing the semantic content of the photos and are appropriate for photo albums.

1. Introduction

Content-based image browsing/retrieval using low-level features such as color, texture, and shape was used in many image retrieval systems. However, low-level features generally can not fully represent image content. In this paper, we present a system for Photo album management.

Users usually remember photos by date, scene type and the individuals in the photos [3]. Date and other metadata could be obtained from the photos. Therefore to automate photo album management, the "who is in the photo" needs to be automatically obtained. However, automatic face recognition usually doesn't work well in this case because of the large variance in pose, illumination and partly poor photo qualities [2]. Therefore, other individual-related features, instead of identity, should be used for representation. In [3], a semi-automatic face annotation mechanism is proposed. "Cloth features" are used for photos of the same scene or event as a replacement for facial features to identify individuals appearing in multiple photos. In [9]the idea is extended by using a Bayesian framework to identify faces from image to image in photo albums, where features such as image date and other face features are used.

In this paper, we present a system for image browsing. The system builds composite images, which represent photos captured for the same event. The system clusters the images within a single event based on face arrangement, which includes the locations of the faces, their sizes and identities.

A composite image is built from an image that represents the event as well as faces of people who appear in the event. In fact, the individuals involved in an event could be a very discriminant feature to recognize the event. If combined with the representative image that shows the background, the user can easily recognize the event.

Our system has the following features:

1) Automatic Face Detection with probabilistic output;

2) An effective similarity measurement function based on face arrangement;

3) Event representation using composite images that include faces which appear in the event;

4) Photos are clustered based on the similarity metric to facilitate browsing; each cluster contains photos that could be displayed in one screen.

2. The digital photo album

In this section we describe our system for image browsing based on faces and their arrangement. During archiving, a fast and robust face detector, based on the algorithms in [7][8], is applied to the images. This face detector could a) detect faces with large yaw-rotation and in-plane rotation b) produce face detection results with probabilistic measures that indicate the confidence in the detection accuracy. The system allows the user to modify the results of the face detection if there are any false positives or false negatives, which are usually rare. The face detection results are then archived in the database. It is also used to create composite images, which are used for browsing, that represent the folders of images in the database as explained in Section 2.1. In this paper each folder contains images from the same event.

Figure 1 shows the browsing user interface of our system. The interface is composed of three panels. The left one is the folder browsing section, where folders are listed using the folder representative images. The middle one is the image browsing section where all images in one folder are listed by cluster. The photos in the same folder are clustered, based on face arrangement similarity as

0-7803-8603-5/04/$20.00 ©2004 IEEE

explained in Section 2.2, so that each cluster contains exactly the number of photos that could be shown in one screen. Using the scroll bar, the user could browse photos of different clusters in the folder. These clusters are sorted by average number of faces. The right panel is where images could be viewed in more detail.

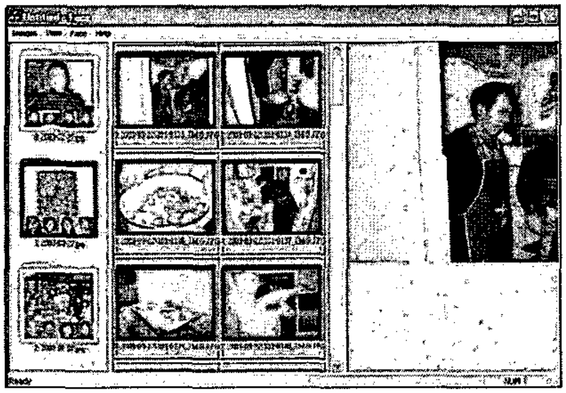

Figure 1. The user interface for photo album browsing

2.1. Building Composite images to represent the folders

Users tend to remember an event by WHEN and WHERE it happened and WHO participated in that event. The Date and Time of an event could be extracted from digital photos' metadata and represented as text. The place of the event could be reasonably represented by a photo of the scene. However, it is not easy to automatically select such a representative photo for the event.

In this paper, we propose to build a composite image to represent the photos in a folder. The composite image is built from an image from the folder that has the largest number of faces as well as a set of representative faces from the folder. This criterion is natural and usually effective, especially when the user is searching for photos of someone.

Each individual usually wears the same clothes in a single event, therefore the color and texture features of the cloth can be an effective feature to roughly identify individuals [3]. The *correlogram* feature[4], which is an effective feature for representing color and texture, is adopted in this paper to find representative individuals in the folder. The faces of these individuals are then used in creating a composite image that represents the event. Figure 3 is the algorithm for generating representative images for folders.

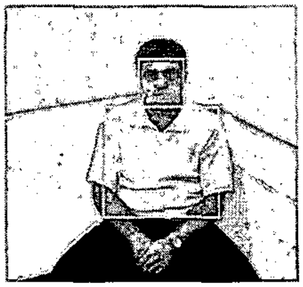

Figure 2. facial region and body region

1.Select the photo with maximum number of faces as the main representative photo;
2. Select the body regions (e.g., Figure 2) based on the results of face detection, extract color-texture features from the body regions;
3. Cluster the body features using 4-center clustering algorithm; then select the faces that are nearest to the clustering centers as sub-photos. These faces are, with great possibility, for different persons.
4. Combine the main representative photo and sub-photos of the faces to build square images as the representative images of the folder, Fig. 4.

Figure 3. Algorithm for Generating Folder Representative Image

The Folder Representative Photo, i.e., built in step 4 above, is described in Figure 4. If the photo orientation is landscape, the left frame is used; otherwise the right one is used.

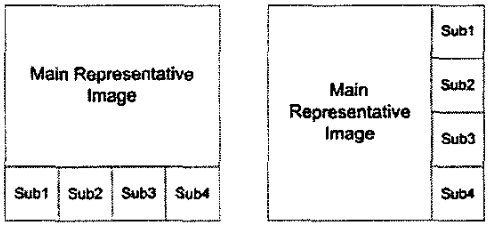

Figure 4. Illustration of Folder Representative Images

2.2. Browsing folders by Face Arrangement

In Section 2.2.1, we describe our measure of distance/similarity between images based on face arrangement. It is used to cluster the images in a folder for browsing as explained in Section 2.2.2.

2.2.1. Similarity Function Based on Face Arrangements

There are several approaches for representing spatial relations between objects in images. In this paper our goal is to establish similarity between images based on the number of faces, their sizes and locations. Each face is represented by its bounding rectangle. To calculate the distance between two images, correspondence between the faces in the images has to be established as shown in Fig. 5. The distance between two images is defined as the weighted sum of four measures, T_N, T_D, T_A, T_{ov}.

$$dis(i,q)=\alpha*T_D+\beta*T_{OV}+\gamma*T_A+(1-\alpha-\beta-\gamma)*T_N$$

T_D measures the relative spatial location.

$$T_D=\frac{1}{M}\left(\frac{\sum_{j=1}^{M}|L_i^j-L_q^j|}{W}+\frac{\sum_{j=1}^{M}|W_i^j-W_q^j|}{W}+\frac{\sum_{j=1}^{M}|T_i^j-T_q^j|}{H}+\frac{\sum_{j=1}^{M}|H_i^j-H_q^j|}{H}\right)$$

T_A measures the average ratio of the area of corresponding faces.

$$T_A=\frac{1}{M}\left[\sum_{j=1}^{M}\frac{\min(A_i^j,A_q^j)}{\max(A_i^j,A_q^j)}\right]$$

T_{ov} measures the average overlapped area between corresponding faces.

$$T_{OV}=\frac{1}{M}\left[\sum_{j=1}^{M} Overlapped\ Size(A_i^j,A_q^j)\right]$$

T_N is a measure that captures the difference in the number of faces between the query and a test image.

$$T_N=\frac{|I_N-Q_N|}{M}$$

Where I_N and Q_N are the numbers of faces in images I and Q. In case of retrieval, image I can be an image in the database and image Q can be the query image. M is $\min(I_N,Q_N)$, (L_i^j,T_i^j) is the coordinate of the left-top corner of the rectangle for face j in database image i. When q is used it means for the query image. W_i^j and H_i^j are the width and height of the rectangle for face j in image i. A_i^j is the area of the rectangle for face j in image i. W and H are the width and height of the image (all images are normalized). In the experiments we used $\alpha=0.2,\beta=0.1$ and $\gamma=0.1$ and the similarity is obtained by the following equation:

$$Similarity(i,q)=\lceil(1-dist(i,q))\times 100\rceil$$

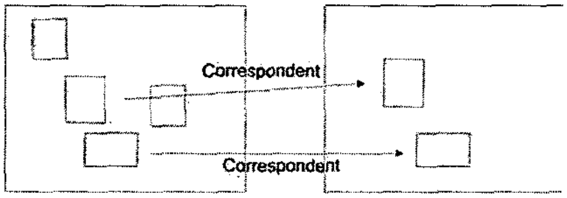

Figure 5. Correspondence between objects in two images

Correspondence between objects is accomplished by minimizing T_D for every object in one image with all other objects in the other image.

2.2.2. Clustering the Photo Images According to Face Arrangement

When the number of photos in a single folder is larger than maximum number of images that can be displayed in one screen, the photos should be arranged in some order to facilitate browsing. In this paper, we cluster the photos based on their faces arrangement. Each cluster contains photos that could be displayed in one screen. Clustering images in a database was used before [6], however, in this paper, the clustering is based on the face arrangement similarity. Each cluster contains a set of images that can be displayed in one screen.

Our distance/similarity measure of face arrangement, defined in Section 2.2.1, is a non-metric measure. Therefore, general metric-based clustering algorithms (e.g. k-means), could not be applied in our system. In order to cluster the photos in a folder based on the this measure we propose the clustering algorithm in Figure 6

1. Given a set of m photos, and the number of photos, n, which each cluster should contain
2. Calculate the number of cluster centers k

$$k=\left\lfloor\frac{m-1}{n}\right\rfloor+1$$

3. Sort the photos based on the number of faces and divide them into k clusters; each of the first k-1 clusters contains n photos and the last one contains m – n * (k-1) photos
4. For each pair of photos (x_i, x_j) from different clusters, check whether their exchange could reduce the sum of with-in distance of clusters, if so, exchange them. The with-in distance of clusters is defined as:

$$S_w=\frac{1}{n}\sum_i^n\sum_j^n Similarity(x_i,x_j)$$

5. Repeat step 4 until there are no such pairs.

Figure 6. Cluster algorithm for photo album browsing

After the clusters of photos are built, they are sorted by the average number of faces in each cluster. And the clusters can be browsed based on this order.

3. Experimental results

3.1. Representative Images of Folder

Using the method described in Section 2, we can generate the representative photos of a folder. Figure 7 shows two representative folder images generated by our system. This result is compared with the "Thumbnails-view" used in *Windows XP explorer.* The first folder contains 78 photos from a party. Our algorithm selected a representative photo that has the largest number of faces and selected four faces as sub-photos (see algorithm in Figure 3) to combine with the representative photo, while the Windows explorer selected some photos that have furniture and are not good representatives for the event. The second folder contains photos from a trip that contains an outdoor event with gathering in a cabin. Our algorithm selected a representative photo that contains faces of most individuals in that event and attached to it four different faces from the event.

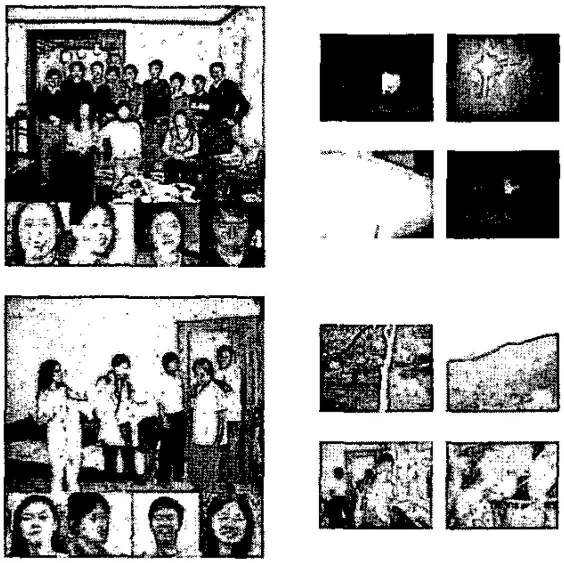

Figure 7. Two Examples of Folder Representative Images, Left are auto-generated folder representative images, right are randomly selected photos from that folder, (similar to Thumbnails-view of folders in Windows XP explorer)

3.2. Face Arrangement Similarity and Clustering

We conducted a subjective evaluation of the face arrangement similarity function described in Section 2.2.1. The evaluation was done by using the similarity function for retrieval, where images were archived based on the face arrangement and queries were either images or sketches. The retrieved images were judged as either relevant or irrelevant by a subject who is not familiar with the system. Then, the performance was calculated. First we explain the performance measure and then present the results.

We used the performance measure used in [5], which is defined to reflect the rank position (*Rp*) of the relevant images:

$$R_p = \frac{1}{NN_R}\left[\sum_{i=1}^{NLR} R_i - \frac{N_R(N_R+1)}{2}\right],$$

where N is the number of images retrieved for browsing (in our experiment we looked at the top 50 images), N_R is the number of the relevant images for the query, R_i is the rank at which the i^{th} relevant image is retrieved. This measure is 0 for perfect performance and approaches 1 as performance decreases.

The query images contained up to four faces in different positions. The average R_p values are shown in Table 1 for queries with different number of faces.

Table 1. Average rank position for different face number

# faces in query image	Average R_p
1	0.0590
2	0.1275
3	0.1299
4	0.0092

4. Conclusions

We have presented a system for archiving and browsing of photos in a photo album. The system allows for browsing based on faces and their arrangements and represents events with a set of composite images. As was demonstrated in the experiments, the results are intuitive and make sense from user's point of view. The archiving process is fully automated.

References

[1] S. K. Chang and S. H. Liu, "Picture indexing and abstraction techniques for pictorial databases", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 6, pp. 475-484, 1984.

[2] R. Chellappa, C.L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," in *Proceeding of IEEE*, vol. 83, pp. 705-740, May 1995.

[3] L. Chen, B. Hu, L. Zhang, M.J. Li and H.J. Zhang, Face annotation for family photo album management, *International Journal of Image and Graphics*,Vol.3, No.1, pp.81-94, 2003

[4] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, Image indexing using color correlograms," In IEEE Conf. Computer Vision and Pattern Recognition 762 (1997).

[5] H. Muller, W. Muller, D. Squire, S. Marchand-Maillet and T. Pun, "Automated Benchmarking in Content-based Image Retrieval", In *Proceedings of the 2001 IEEE International Conference on Multimedia and Expo (ICME2001)*, Tokyo, Japan, August 2001.

[6] S.Krishnamachari, and M. Abdel-Mottaleb, "Hierarchical Clustering Algorithm for Fast Image Retrieval", Storage and Retrieval for Image and Video Databases VII, SPIE Symposium on Electronic Imaging: Science & Technology, San Jose CA, January 1999.

[7] H. Rein-Lien, M. Abdel-Mottaleb, A.K.Jain, Face detection in color images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.24 no.5,pp:696 -706, May 2002

[8] R. Xiao, L. Zhu, H.J. Zhang, Boosting Chain Learning for Object Detection, in *Proceeding of ICCV 2003*, pp:709-715.

[9] L.Zhang, L.B Chen, M.J.Li, H.J.Zhang, Automatic face annotation for family album, ACM multimedia, 2003