Bayesian Face Annotation in Family Albums

Lei Zhang[1], Longbin Chen[2], Mingjing Li[1], Hongjiang Zhang[1]
Microsoft Research Asia, 49 Zhichun Road, Beijing 100080, China
[1] {leizhang,mjli,hjzhang}@microsoft.com, [2] chenlongbin@msrchina.research.microsoft.com

Abstract

Automatic annotation of photographs is one of the most desirable needs in family photograph management systems. This paper presents a new framework for semi-automated face annotation in family photo album applications. Based on the Bayesian similarity measure, which integrates face recognition and content-based image retrieval, the system is able to provide both face annotation and similar face retrieval, and thus significantly improve the management of family albums.

1. Introduction

With the rapid development of digital cameras and scanners, digital photographs are very polular and automated tools for organizing these photographs become extremely desirable. Unfortunately, though there are many commercial products available, annotating semantic content of photographs, which is a tedious task in organizing photographs, is still left to users.

The most commonly used entries for indexing family photographs are related to *when*, *where*, *who* and *what*. With the advance in digital camera technology, date and time as well as location data is or will be readily available in cameras. In this paper, we focus on how to automatically extract "*who*" in family photographs.

Compared to general image collections, most images in family photo albums are in color. Also, a same individual often appears in a number of photographs taken in the same day or event. In this paper, we utilize these two characteristics to compensate the weakness of current face recognition algorithms and improve the annotation accuracy.

In this system, we propose a new framework for semi-automated face annotation in family photo album applications [3]. The user scenario we target at is in typical digital family photo albums, in which only a handful number of people, e.g. ten to fifty, are of our concern and appear frequently. When a new photograph is imported to the system and a face is detected, the system will allow the user either to provide a name label for that face, or to confirm or select a name from a recommended name list derived from prior face annotation history.

2. Bayesian Face Annotation

2.1 Framework

The framework of the proposed face annotation system is illustrated in Figure 2. First, a face detector derived from the robust AdaBoost-based algorithm [2] is used to detect faces in newly uploaded images or images already in the album. The facial features are extracted from each detected face area, as well as the contextual features. The extracted features include color, eigen-face feature and facial component features in the extended face area. To derive the similarity measure, a large set of training samples are collected offline to train the probability model for each feature. Then these probability models are integrated into a Bayesian framework to measure the similarity between faces. Based on this statistically derived similarity measure, the system generates a list of name candidates for a given query or a new face by statistical learning approaches. Meanwhile, the system also allows users to search for similar faces by specifying either a face or a name and then annotating multiple faces in a batch way. Furthermore, relevance feedback can also be introduced to refine the retrieval result.

Also, our system provides a simple user interface to reduce users' interaction. As illustrated in Figure 1, if a user moves the mouse onto a labeled face, a tool-tip will popup to show the name of this individual. If the face has not been labeled before, a popup menu with a list of candidate names will appear. The user can either select one of the names to annotate the face, or set a new name for that face. To further simplify the face annotation, similar face retrieval and relevance feedback are allowed for labeling multiple faces in a batch way.

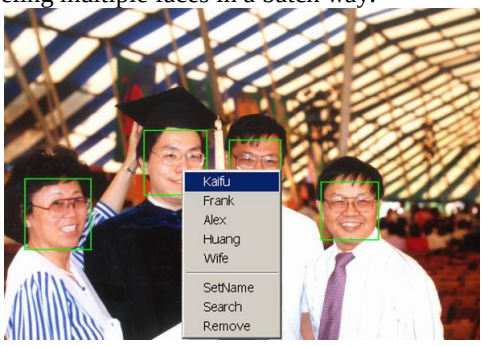

Figure 1. A popup menu in the face annotation system

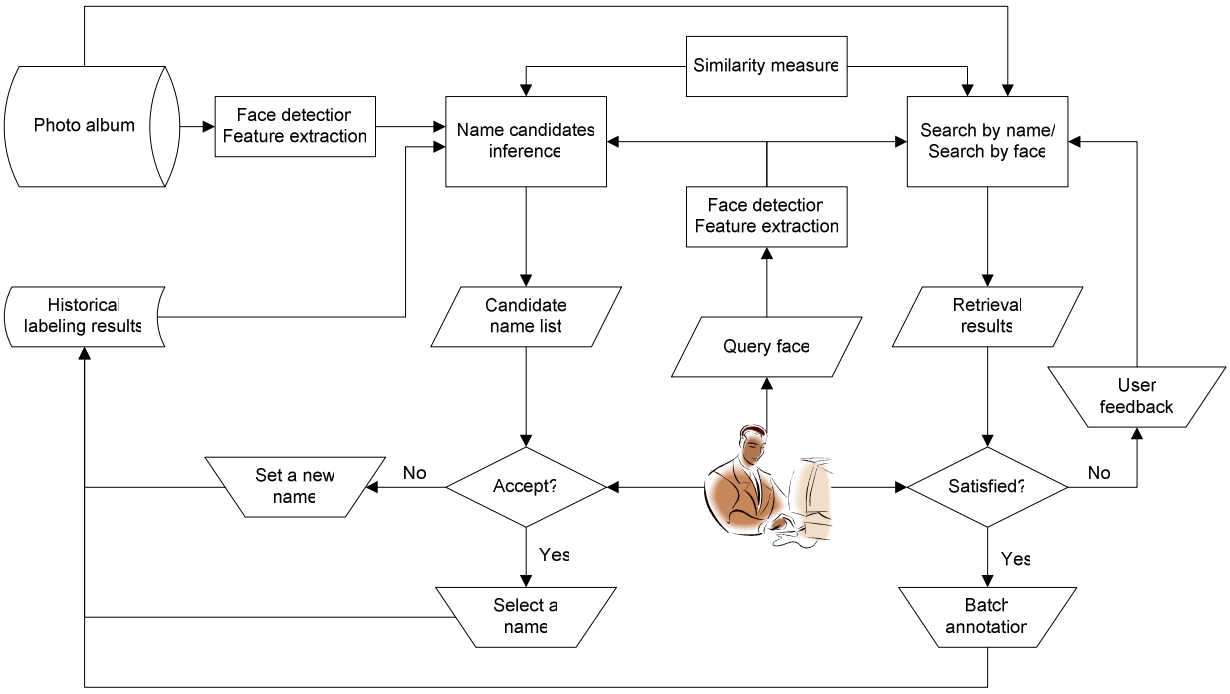

Figure 2. Overview of the proposed framework

2.2 Face Annotation

In the proposed framework, face similarity is defined as *maximum a posteriori* (MAP) estimation, which measures similarity between two faces by:

$$S(\boldsymbol{F}_1, \boldsymbol{F}_2) = p(\Delta \boldsymbol{F} \in \Omega_I) = p(\Omega_I \mid \Delta \boldsymbol{F})$$

where $\boldsymbol{F}_1$, $\boldsymbol{F}_2$ are the features corresponding two faces, Ω_I corresponds to difference appearances of the same individual. By introducing the *posteriori*, multiple features, such as facial appearance feature and contextual feature are well integrated [3].

Given an unknown face, based on the nearest neighbor or *K*-nearest neighbor algorithms, among its *K* nearest labeled faces, the name candidates can be generated by sorting the names according to the sum of similarities to the unknown face.

2.3 Name Propagation and Batch Annotation

Because the similarity function is actually a posterior probability, it can be regarded as a confidence score of how confident to annotate the unlabeled face with the top first name among the recommended name candidates.

By setting an appropriate confidence threshold in accordance with high annotation accuracy, substantial faces may be annotated automatically once similar faces were annotated before. For example, if an individual appears in several photos which were taken in the same event and one of the faces was annotated before, the other faces may be annotated by propagation with high accuracy.

Also, given the face similarity measure function, the system provides the function of similar face retrieval. In this way, users are allowed to search for similar faces by specifying either a face or a name and then annotate multiple faces in a batch way. On the other hand, similar face retrieval itself is of great use in family photo album management as faces and names are typically the most important index of photographs. Furthermore, relevance feedback techniques can also be introduced to refine the retrieval result.

3. References

[1] Zhao W., Chellappa R., Rosenfeld A. and Phillips P., "Face recognition: A literature survey", Technical Report, Maryland University, CfAR CAR-TR-948, 2000

[2] Xiao R., Li M.J., Zhang H.J., "Robust Multi-Pose Face Detection in Images", to be appeared in IEEE Trans. on CSVT Special Issue on Biometrics, 2003

[3] Zhang L., Chen L., Li, M.J., Zhang H.J., "Automated Annotation of Human Faces in Family Albums", to be appeared in ACM Multimedia, 2003.