



An on-line learning method for face association in personal photo collection[☆]

Liliana Lo Presti^{*}, Marco La Cascia

University of Palermo, Italy

ARTICLE INFO

Article history:

Received 10 May 2011

Received in revised form 6 January 2012

Accepted 26 February 2012

Keywords:

Face descriptor

Data association

On-line learning

Semi-supervised learning

Digital libraries

ABSTRACT

Due to the widespread use of cameras, it is very common to collect thousands of personal photos. A proper organization is needed to make the collection usable and to enable an easy photo retrieval. In this paper, we present a method to organize personal photo collections based on “who” is in the picture. Our method consists in detecting the faces in the photo sequence and arranging them in groups corresponding to the probable identities. This problem can be conveniently modeled as a multi-target visual tracking where a set of on-line trained classifiers is used to represent the identity models. In contrast to other works where clustering methods are used, our method relies on a probabilistic framework; it does not require any prior information about the number of different identities in the photo album. To enable future comparison, we present experimental results on a public dataset and on a photo collection generated from a public face dataset.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Personal photo albums show characteristics that make them quite different from a generic image database. In general, such photo collections are characterized by the presence of few persons, and photos are taken in several places. Based on common sense, users would like to browse their own collections considering properties such as when and where a photo was taken and on who is in the photo [1]. While the first two properties are based on contextual information and they can be extracted – if available – from the EXIF data embedded within each photo [2], organizing the sequence based on who is in the photo is still a very challenging problem that requires each face to be associated to a tag. In this paper, we present a method aimed to minimize the user effort in tagging photos. Our method detects the faces in the photo sequence to process, and arranges them in groups corresponding to identities. Users can tag faces associating a label to the whole face group. A similar approach is followed also in [3–6]. However, these methods apply clustering techniques to get a coarse face partition later refined by applying post-processing steps.

In case of personal photo album organization, clustering-based methods do not consider the mutual exclusivity constraint: “two faces in the same photo cannot be assigned to the same partition/identity”. Our method, instead, has been developed to exploit this important property and adopts a semi-supervised approach for grouping similar faces. The detected faces are processed in bunches (each bunch corresponding to the set of faces detected in the same

photo) and each face is associated to an identity. Identities are then iteratively estimated by means of the discovered associations. From this point of view, our method reduces to a multi-target visual tracking where the detected faces are used to on-line estimate and update an appearance model for each identity. The main difference between standard tracking applications and our domain is the lack of temporal smoothness for the observations.

Motivated by the success of recent works such as [7–10], where visual tracking is performed by modeling the appearance information via on-line trained classifiers, in this paper we adopt a similar approach to capture the face information available across the sequence while new associations between the identities and the depictions are discovered in each processed photo. Our main contribution is the use of a classifier ensemble to represent the set of identities, and a new probabilistic framework to enforce the mutual exclusivity constraint; such framework ensures the collaboration among the classifiers and permits to label the faces to be used for retraining the whole ensemble (see Fig. 1).

In our method, no prior information about the number of identities in the photo album is needed, in contrast with clustering methods (for instance K-means) where this information is sometimes required. If prior information about the collection is available, such as already tagged photos, it could be possible to train offline the classifiers and use them to classify faces detected in the new incoming photo sequence. However, an on-line strategy better accounts for the discovery of new identities when matches with the already found ones become too uncertain. In fact, in photo collections it is very common that new identities, never seen in previous pictures, appear in the collection.

In the following, we will present some of the most relevant related works, then we will give an overview of our method and we will explain deeply its main components. Finally we will present

[☆] This paper has been recommended for acceptance by Matti Pietikainen.

^{*} Corresponding author. Tel.: +39 091 238 42 524.

E-mail address: lopresti@dinfo.unipa.it (L. Lo Presti).

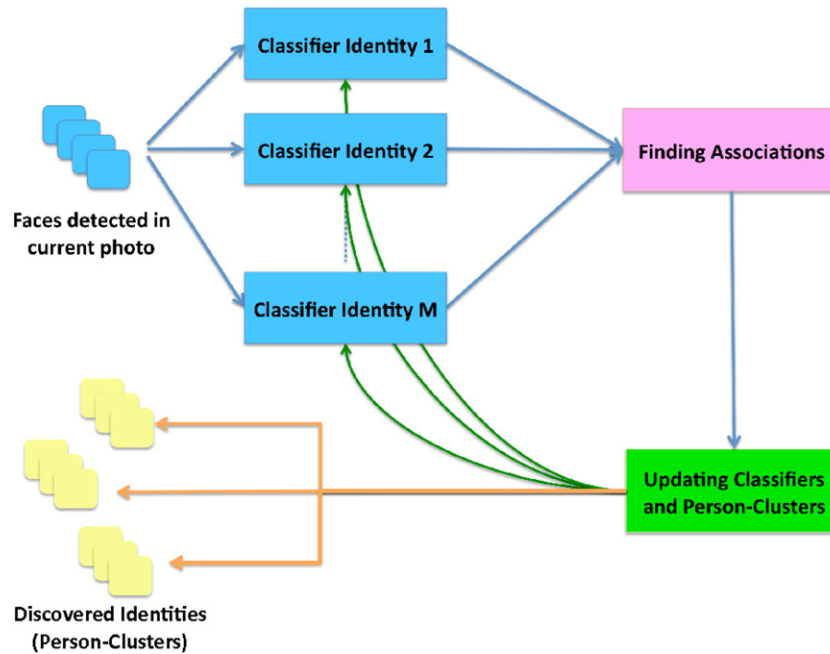


Fig. 1. Schema representing our framework: faces detected in the current photo are classified by the classifiers associated to each already discovered identity. The probabilities corresponding to the classification outcomes are used in a probabilistic framework to fuse the decisions from the classifier ensemble. Once the set of associations with maximal joint probability is found, the faces are used to update the classifiers' parameters and the person-clusters.

experimental results on a public dataset and on a photo sequence generated from a public dataset.

2. Related works

Personal photo albums are photo collections depicting individuals belonging mainly to the same family or the same social group. In general, the number of depicted individuals/identities in the whole collection is small but unknown.

The most natural way to browse a photo collection is, perhaps, considering who is in the photo based on tags attached to each face. Tags can be assigned by the user considering a face at a time but such task is very boring and a user would soon give up when too many photos have to be processed. New methods enabling a user to easily and rapidly tag photos were presented in [11,12]. In such methods, the already tagged photos are used to suggest likely tags for identifying faces in untagged photos. Therefore, face recognition/matching techniques may be applied to perform tag association, considering that such task must be performed in the “wild”, that is the detected faces are affected by large pose variations and abrupt illumination changes.

Recently, new applications such as iPhoto [1] and Picasa [13] provide tools for face detection and recognition that are used to suggest likely tags for each face detected within the collection. In particular, Picasa presents a tool closely related to our work where faces are organized in groups each one containing faces of the same person. The user is assisted in the tagging task asking for a confirmation for every suggested label before propagating it to all the faces within the same group. However, it can be observed that results can be affected by over-clustering, that is several clusters can correspond to the same person (see [14]).

Nevertheless the success of these commercial semi-automatic applications, still the problem of photo organization, remains very challenging. On the one hand, faces in the “wild” processing are not a solved problem and new face descriptors and/or learning algorithms are needed to enhance identity recognition; on the other hand, new methods aiming to minimize user's interactions are required, moving from semi-automatic to fully automatic photo organization.

Many previously proposed papers [3–5] use clustering methods to group faces, each cluster representing an identity. All these methods do not explicitly consider the mutual exclusivity constraint. Only some works, as for example [4,15], use this property as post-processing step in order to remove incorrect associations or to suggest probable tag to each face. In [4], K-means algorithm is applied in several, subsequent steps for identifying persons using both face and clothing information. Faces are rearranged to enforce the mutual exclusivity constraint during a refinement step. This approach has the limitation that the number of face clusters has to be specified by the user.

In [16], active learning has been extended with the mutual exclusivity constraint to determine which faces must be tagged first in order to limit the number of inputs from the user. They propose a probabilistic discriminative model to induce a probability distribution over class labels, and a Markov Random Field to enforce the constraints. However, a considerable user effort is still required to tag photos. In [17], the mutual exclusivity constraint is used within an agglomerative hierarchical clustering to fuse clusters whose faces did not belong to the same photo. However, the constraint is not used for evaluating the best association and, in general, the result depends on the order of cluster processing.

An interesting approach is presented in [6], where users are allowed to multi-select a group of photos and assign a name/tag to the person appearing in them. The method attempts to propagate the assigned name from photo level to face level, i.e. to infer the correspondence between name and face. However, while the user effort for tagging is minimized, still the user has to manually identify the group of photos where that person appears. Moreover, the method is not able to disambiguate between persons that always appear together in the set of photos.

In previous works [18,14], associations across a photo sequence are found by considering face and clothing features locally within a temporal window and by means of a joint probabilistic data association (JPDA) [19]. Basically, associations between identities and depictions in a photo are computed as maximum matching in a bipartite graph [20] by means of the Hungarian algorithm [21]. While the method is able to cluster depictions in an unsupervised way, it is dependent on the probability that a new identity is discovered and/or a person is not

Download English Version:

<https://daneshyari.com/en/article/527106>

Download Persian Version:

<https://daneshyari.com/article/527106>

[Daneshyari.com](https://daneshyari.com)