Image and Vision Computing 26 (2008) 1269–1276

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Distinctive and compact features

Ayelet Akselrod-Ballin *, Shimon Ullman

Department of Computer Science and Applied Math, The Weizmann Institute of Science, Rehovot 76100, Israel

A R T I C L E I N F O

Article history:
Received 16 March 2006
Received in revised form 3 March 2008
Accepted 4 March 2008
Available online

Keywords:
Feature selection
Object recognition
Distinctive features
Facial features

A B S T R A C T

We consider the problem of extracting features for multi-class recognition problems. The features are required to make fine distinctions between similar classes, combined with tolerance for distortions and missing information. We define and compare two general approaches, both based on maximizing the delivered information for recognition: one divides the problem into multiple binary classification tasks, while the other uses a single multi-class scheme. The two strategies result in markedly different sets of features, which we apply to face identification and detection. We show that the first produces a sparse set of distinctive features that are specific to an individual face, and are highly tolerant to distortions and missing input. The second produces compact features, each shared by about half of the faces, which perform better in general face detection. The results show the advantage of distinctive features for making fine distinctions in a robust manner. They also show that different features are optimal for recognition tasks at different levels of specificity.

1. Introduction

In performing recognition, the visual system, either human or artificial, must cope with the problem of image variability, that is, that an object's appearance is highly variable due to changes in shape, viewing direction, illumination, and occlusion. At the same time, the task often requires making fine distinctions between objects, such as between similar faces. It is particularly surprising given these difficulties that reliable recognition can be obtained on the basis of reduced and distorted representations, such as caricatures and drawings produced by artists, e.g. [1], see examples in Fig. 1. In such images, the faces consist only of a few informative features that are distorted, often represented schematically, and placed in an inaccurate spatial arrangement. This illustrates a fundamental general question: how is it possible to reliably distinguish between multiple similar classes, and yet be tolerant to reduced and distorted information?

To approach this problem, we define and compare two natural strategies for extracting classification features in problems involving multiple similar classes, and apply them to face examples. Both are based on maximizing information for classification, but they produce notable different features. One method divides the problem into multiple binary classification tasks, while the other uses a single multi-class scheme. We show that the first leads to a **sparse** representation based on distinctive features, which is tolerant to large distortions and missing input, and better for robust face identification, requiring only a few distinctive features for reliable identification. The second leads to **compact** coding where each features is shared by about half of the faces, and which performs better in general face detection. The distinctive features are also shown to be similar to the ones selected by an artist specializing in producing reduced face representations [1], and the algorithm is the first to automatically produce such distinctive features. The focus of the study is on feature selection for multi-class recognition, rather than face recognition. Face images are used as a testing domain, for which there are example of distinctive features selected by human experts.

The rest of the paper is organized as follows: Section 2 reviews past relevant approaches to face recognition and detection, with emphasis on the type of features used by these approaches. Section 3 describes the two selection strategies, and automatic extraction of sparse and compact features. Section 4 presents experimental results, comparing sparse and compact features in face recognition and detection. We also compare between the distinctive fragments obtained by the current method and the representations produced by an artist. Section 5 includes a discussion of the results and conclusions.

2. Previous work

The current study considers the problem of extracting features for multi-class recognition problems, and compares two alternative feature selection strategies. Since we evaluate the two schemes in the domain of faces, we briefly review relevant aspects of past approaches for feature extraction and use it in this domain.

A large number of face recognition schemes have been developed in the past, using different families of features and different

* Corresponding author. Tel.: +972 8 9344268; fax: +972 8 934 4122.
E-mail address: ayelet.akselrod-ballin@weizmann.ac.il (A. Akselrod-Ballin).

0262-8856/$ - see front matter
doi:10.1016/j.imavis.2008.03.005

classification methods (for recent reviews, see [2–5]). Often, the same type of classifier, for instance, support vector machines (SVM, [6,7]), can be used with different feature types, leading to different classification performance. We focus below on the main approaches and the type of features they selected and used, since this is the most relevant aspect to the current work.

A wide range of features have been used for both face recognition and face detection. Appearances based methods use image examples of face regions for learning models, and typically apply statistical analysis and machine learning techniques for recognition. The image appearance is used directly for recognition, using either global descriptions (e.g. PCA [8], ICA [9]), or the appearance of local face regions such as [10] for face detection. Decision can then be reached using for instance projection distance, [8] or linear discriminant analysis (LDA/FLD) [11].

Structural matching methods based on geometrical constraints use as features measured distances and angles between key points of the face [12,13]. A recent example within this category is the active shape model (ASM) [14], which is a statistical shape model, representing faces with shape and intensity information.

Deformable templates methods use a geometric model of the face, but allow it to deform in a controlled manner during the matching process. For example, in [15], facial features are described by parameterized templates, which are matched to an image by minimizing an energy function.

Several recognition systems use constellations of simple local features, including wavelets, Gabor patches, edges, lines and curves, for representing and recognizing faces. In such approaches the face is described by the constellation, sometimes modeled as joint distribution, of the features. The face detection algorithm developed in [16] uses a multi layer network to directly learn input image intensities. The algorithm presented in [17] classifies objects based on a set of rectangular features, where each feature computes the sum and difference of pixel intensities within a number of sub-rectangles. In the Elastic Bunch Graph Matching system of [18], faces are represented as graphs, with nodes positioned at key points on the face (eyes, tip of nose, mouth, etc.), and the features used are based on wavelet responses. Wavelet transforms were used also by Schneiderman and Kanade [19] and applied to the detection of faces and cars. In general, previous methods used the same set of features, often extracted in an ad hoc manner, for all recognition tasks, and did not compare features optimized for a single individual, multi-class recognition, and general face detection.

Psychological studies support the claim that in human vision some type of distinctive features are used for face recognition [20,21]. A recent study [22] showed that in performing recognition, humans focus on restricted regions in the face, and that the selected regions are task-dependent. The study supports the notion that the visual system does not rely on a fixed set of features, but learns for each task to use a small subset of critical features that are the most informative for the task.

The methods described above rely on an accurate geometrical agreement between the face model and the input image. They therefore have severe limitations in their ability to deal with reduced and distorted images. These limitations can be illustrated by comparing real images with artists drawings (as in Fig. 1), which are recognizable by human observers despite the large distortions and features omission in the input images.

In the present work, we compare two alternative strategies to the selection of useful features in multi-class problems in general, and face recognition in particular. We show that one of these strategies produces a representation that relies on the presence of a small number of distinctive features, and can use them for recognition without relying on exact geometric agreement between the model and the input image. These features and their extraction are described in the following section.

3. Feature extraction

3.1. Sparse and compact features

We contrast below two alternative approaches to extracting useful visual features for classifying a novel image, into one of n known classes. For example, the training may consist of face images taken from n different individuals under different viewing conditions (see Fig. 3), and the task is to then classify a novel image of one of the known individuals. One strategy results in sparse, the other in compact representation. Compact coding uses features

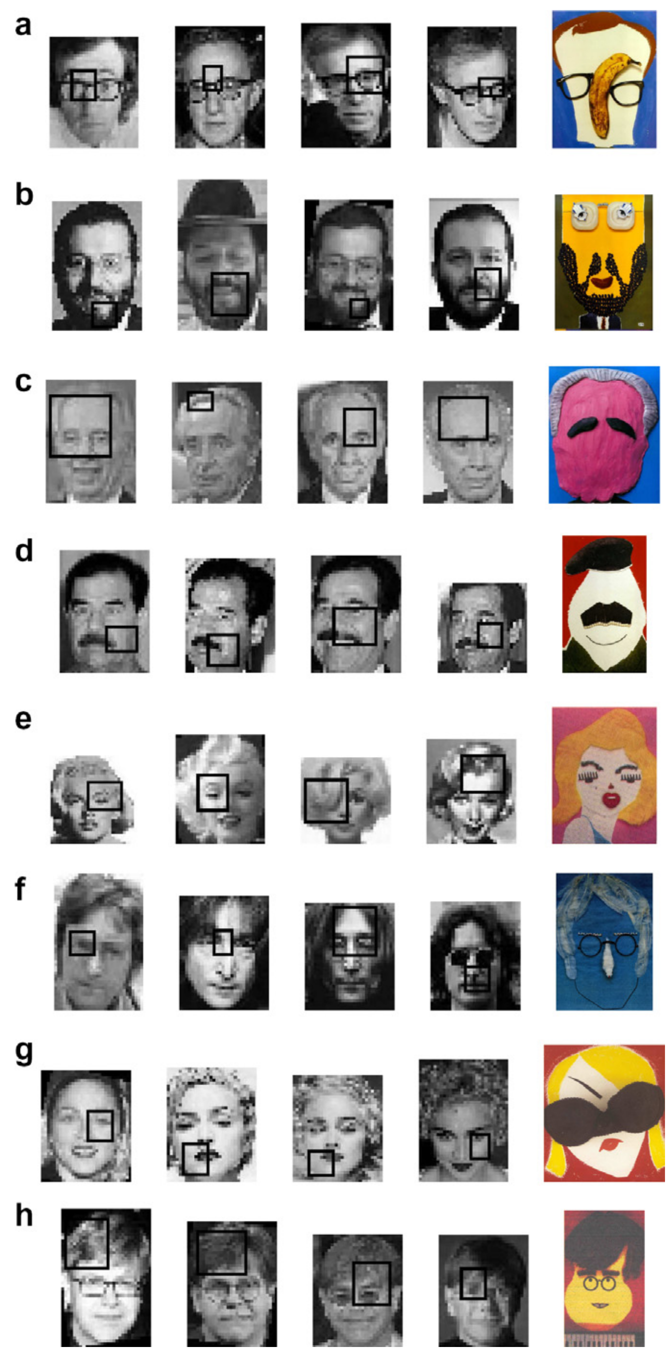

Fig. 1. Sparse fragments extracted for several individuals. The black rectangles displayed on the images illustrate the set of informative extracted fragments (in decreasing order). The corresponding artist's images (by H. Piven) for these individuals is shown on the right column in each panel. (a) Allen. (b) Deri. (c) Peres. (d) Sadam. (e) Monroe. (f) Lennon. (g) Madonna. (h) Elton.