Available online at www.sciencedirect.com
ScienceDirect
Image and Vision Computing 27 (2009) 2–9
image and vision computing
www.elsevier.com/locate/imavis

Object detection using image reconstruction with PCA

Luis Malagón-Borja [a], Olac Fuentes [b,*]

[a] Computer Science Department, I.N.A.O.E., Tonantzintla, Puebla 72840, Mexico
[b] Computer Science Department, University of Texas at El Paso, El Paso, TX 79968, USA

Received 25 January 2006; received in revised form 8 November 2006; accepted 5 March 2007

Abstract

In this paper, we present an object detection system and its application to pedestrian detection in still images, without assuming any a priori knowledge about the image. The system works as follows: in a first stage a classifier examines each location in the image at different scales. Then in a second stage the system tries to eliminate false detections based on heuristics. The classifier is based on the idea that Principal Component Analysis (PCA) can compress optimally only the kind of images that were used to compute the principal components (PCs), and that any other kind of images will not be compressed well using a few components. Thus the classifier performs separately the PCA from the positive examples and from the negative examples; when it needs to classify a new pattern it projects it into both sets of PCs and compares the reconstructions, assigning the example to the class with the smallest reconstruction error. The system is able to detect frontal and rear views of pedestrians, and usually can also detect side views of pedestrians despite not being trained for this task. Comparisons with other pedestrian detection systems show that our system has better performance in positive detection and in false detection rate. Additionally, we show that the performance of the system can be further improved by combining the classifier based on PCA reconstruction with a conventional classifier using a Support Vector Machine.

Keywords: Object detection; Pedestrian detection; Principal Component Analysis; Support Vector Machines

1. Introduction

The object detection problem can be seen as a classification problem, where we need to distinguish between the object of interest and any other object. In this paper, we focus on a single case of the object detection problem, detecting pedestrians in images.

Pedestrian detection is more difficult than detecting many other objects due to the fact that people can show widely varying appearances when the limbs are in different positions. In addition, people can dress in clothes with many different colors and types. For the characteristics of the pedestrian class we need a robust method that can learn the high variability in the class.

Many object detection systems that have been developed focus on face detection. An early very successful system was presented by Rowley et al. [20], which consists of an ensemble of neural networks and a module to reduce false detections. Similar example-based face detection systems have been developed by Sung and Poggio [22], Osuna et al. [17], and Yang et al. [28].

Most pedestrian detection systems use motion information, stereo vision, a static camera or focus on tracking; important works include [5,8,10,29]. Papageorgiou has reported a system to detect pedestrians in images, without restrictions in the image, and without using any additional information [15,16,9]. It uses the wavelet template to represent the image and a Support Vector Machine (SVM) to classify. The system has been improved in [12,4], detecting pedestrians through the detection of four components of the human body: the head, legs, left arm and right arm. Viola, Jones and Snow developed a system to detect pedestrians from image sequences. This system uses a large set of

* Corresponding author.
E-mail addresses: jmb@inaoep.mx (L. Malagón-Borja), ofuentes@utep.edu (O. Fuentes).

0262-8856/$ - see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.imavis.2007.03.004

simple filters as features, and then applies the Adaboost algorithm to generate a cascade of classifiers [26].

We present an object detection system to detect pedestrians in gray level images, without assuming any a priori knowledge about the image. The system works as follows: in a first stage a classifier based on Principal Component Analysis (PCA) examines and classifies each location in the image at different scales. Then, in a second stage, the system tries to eliminate false detections based on two heuristics.

The system uses PCA as a classification tool; the main idea is that PCA can compress optimally only the kind of images that were used to do the PCA, and that any other kind of image will not be compressed well in a few attributes, so we do PCA separately for positive and negative examples; when a new pattern needs to be classified we compare the reconstruction made by both sets of principal components (PCs). In order to improve the performance of the classifier we have used the edge image as additional information for it. Additionally, we show that the performance of the system can be further improved by combining the classifier based on PCA reconstruction with a conventional classifier using a Support Vector Machine.

The organization of the reminder of this paper is as follows: Section 2 presents a detailed description of the system. In Section 3 the performance of our system, and a comparison with similar systems are presented. Section 4 reports conclusions and possible directions for future work.

2. The detection system

2.1. Overview of system architecture

The system works scanning the whole image by means of a detection window of size 105×45 pixels; the window is shifting with two pixel jumps to accelerate the process without losing much information from one window to another. We need a classifier that decides for each window if it contains a pedestrian or not. The construction of the classifier is the most complicated stage, we have created a classifier based on image reconstruction with PCA, this classifier uses the edge image in addition to the gray level image.

The scanning of the whole image is part of an iterative process where the image is resized several times to achieve multi-scale detection. For our experiments, the image has been scaled from 0.26 up to 1.35 times its original size, with increases of approximately 17% in every cycle, thus the image is processed at the following 12 different scales: 0.26, 0.3, 0.35, 0.4, 0.47, 0.55, 0.64, 0.74, 0.86, 1, 1.17 and 1.35, this implies that pedestrians of sizes between 78×33 and 404×173 pixels will be detected by the system.

When the system has finished examining the image in all scales, a second process eliminates some detections that are believed to be false detections. The form in which this process works is eliminating the detections that do not repeat at least twice times and eliminating the detections that overlap.

Fig. 1 shows the complete process to detect pedestrians in an image, starting with the gray level image and finishing with the image with the detected pedestrians.

2.2. Stage 1: a classifier based on image reconstruction with PCA

In this stage, we present a classifier that decides if an image of size 105×45 belongs or does not belong to the pedestrian class. This classifier is based on doing image reconstruction using PCA and comparing the reconstructed with the original images. First, the reasons to work with both the gray level image and the edge image are explained, later we explain how the reconstruction of an image is performed using PCA, and finally, we present the way in which a classifier can use these reconstructions to decide if an image belongs or does not belong to the pedestrian class.

2.2.1. Edge images

Because pedestrians appear in many colors and different textures, it is not advisable to use characteristics based on color or texture to do pedestrian detection. For this reason, we have chosen to use the edge image with the idea of obtaining the typical silhouette of a pedestrian and to eliminate useless information for the classifier.

The edge images were computed using x and y Sobel filters, this edge image serves as complementary information to the gray level image and it allows the classifiers to obtain more data to decide if an image is a pedestrian or not.

In Fig. 2 we can see examples of the corresponding edge images of some pedestrian gray level images. In these images we can observe that although the gray level images are very different in color and background, the edge images present fewer changes from one image to another. This is the reason why the edge images are very important to aid in the classifier's task.

2.2.2. Image reconstruction with PCA

Principal Component Analysis is a popular technique for data compression and has been successfully used as an initial step in many computer vision tasks, including face recognition [2,23] and object recognition [14]. The formulation of standard PCA is as follows. Consider a set of m images, each of size $r \times c$. Each image I_i is represented by a column vector v_i of length rc. The mean object of the set is defined by

$$\mu = \frac{1}{m} \sum_{i=1}^{m} v_i$$

C, the covariance matrix, is given by

$$C = \sum_{i=1}^{m} (v_i - \mu)(v_i - \mu)^{\mathrm{T}}$$