DETECTION OF NOSES AND FACES IN CONSUMER IMAGES

Michael S. Richman

Center for Applied Mathematics
Cornell University
Ithaca, NY 14853

Thomas W. Parks

Department of Electrical Engineering
Cornell University
Ithaca, NY 14853

ABSTRACT

The problem of nose detection is studied using a feature-based approach. Preprocessing is used to find image regions which might contain noses. Features extracted from these regions are processed using a support vector classifier. We apply our algorithm to a database of actual consumer images, provided by the Eastman Kodak Company. A discussion is provided on incorporating the proposed nose detection algorithm into an overall face detection scheme.

1. INTRODUCTION

One of the major interests in object detection is the human face, although other objects such as human body parts (for medical imaging), trees, sky, grass, and furniture (for consumer imaging), are also of considerable interest. Face detection has been approached by a method of full-face template matching and many of its variations [1, 2]. These approaches are not flexible nor efficient in dealing with faces of arbitrary poses.

Another approach for object detection is the feature-based method [3, 2]. The idea is to break a complex object into simpler parts or features. By detecting these simpler features, one can then group or verify the spatial relations between detected features to come to a conclusion about the existence or the absence of the target object [3]. This feature-based approach to object detection is more flexible in dealing with 3-D objects of different views.

The basic building block of a general object detection and localization system is a module that can learn from training samples to derive a good representation of the input image patterns and develop a good classifier that can discriminate between the target and the non-target. For example, if the goal is to detect and locate faces in an image using the feature-based approach, one can choose to detect a much simpler object, say the nose. Once a candidate nose is located in an image, one can compute the expected size and orientation of the face that contains the candidate nose, and therefore, the face detection problem is considerably simplified. Detecting a candidate nose is a simpler task than detecting a face because a nose is not as complex and variable as a face. In particular, the cross-section of a nose yields a feature that is uniquely distinguishable. An example of this feature is shown in figure 2. If a large set of nose images are collected, one can work out a good representation of this feature and then develop a good classifier to discriminate between noses and non-noses.

2. NOSE DETECTION USING SUPPORT VECTOR MACHINES

A support vector detector is a support vector classifier constructed for the purposes of detection. Support vector classifiers have been studied by [4, 5], but not in the context of simultaneous detection and estimation. A support vector detector is designed using a collection of images that have already been classified, known as training data. Because of this, they are considered universal detectors, that is, for any posed detection problem, a detector can be designed as long as training data is available.

Support vector machines were recently used for a face detection problem by Osuna, Freund, and Girosi, [6]. While their results are promising, their method has some limitations. They required faces to be oriented forward, which does not address the problem of detecting faces at different poses. Also, they conducted their search for faces in an image using a 19x19 pixel window. To analyze a test image, the position of the window was successively shifted about the image, and the resulting 19x19 cropped image was processed by a support vector classifier. To account for varying face size, they rescaled test images to several different sizes.

SUPPORTED BY ONR N00014-94-1-0102, NSF MIP 9705349, AND THE EASTMAN KODAK COMPANY

Supported by ONR N00014-94-1-0102, NSF MIP 9705349, and the Eastman Kodak Company

To process one test image, then, this method requires the same image be reprocessed several times at different levels of scale, which translates into extended computational times. Lastly, their detector was trained on entire faces. The reason this is a limitation is that one would need many more training samples to account for the large amount of variability within faces i.e. glasses, mustaches, beards, light distribution.

To extend their method, we propose a new support vector machine-based detection method. The key to our method is that our support vector machine will not be trained to detect whole faces, but rather noses. By focusing on noses, we eliminate complications due to varying facial features from our detection scheme. It will also be simpler to include different facial orientations into our detection scheme, as we will be focusing on one key feature that has been rotated rather than many.

The overall algorithm will use a support vector machine over a window much larger than a typical nose cross-section. With this larger window, we will be able to train our detector on a variety of face sizes. This will enable us to avoid having to reprocess the same image at several different levels of scale as we will be allowing for greater variability in the size of the nose cross-section.

3. NOSE DETECTION ALGORITHM

The algorithm we have developed consists of three components: a preprocessor, the support vector detector, and a postprocessor. A block diagram of the overall scheme is shown in figure 1.

3.1. Preprocessor

A preprocessor is used to eliminate regions of the image which obviously do not contain noses. We only include regions which are skin colored by using a mapping from R (red), G (green), and B (blue) to $log(\frac{R}{G})$ and $log(\frac{B}{G})$. This mapping reduces the effect of scene illumination. Since the cross-section in figure 2 resembles the second derivative of a Gaussian pulse, we use matched filtering on the skin colored regions. Slices centered at the remaining pixels are used as feature vectors for the support vector machine.

3.2. Support Vector Machine

The support vector machines employs a radial basis kernel function and classifies each feature vector generated by the preprocessor. To facilitate further processing, our support vector machine does not employ binary classification. Rather, the raw value generated by the detector (without thresholding) is assigned to each feature vector. A binary classifier is not useful in this instance because there are typically several pixels within any given nose that would result in postively-detected feature vectors. Classification of each feature vector as nose or non-nose would not provide information as to the quantity of noses present in an image. Instead, the raw values are supplied to a postprocessor.

3.3. Postprocessor

A postprocessor is used to convert support vector machine values at each candidate pixel into a description of the number and location of noses in an image. First, the pixels passed by the preprocessor are sorted by support vector machine value, and the largest-valued quintile is kept for subsequent steps. At each pixel in the largest quintile, a simple pattern recognition scheme based on the values of neighboring pixels is used to determine which pixels correspond to that of a nose. A small polygon is drawn around each of these pixels, and overlapping polygons, if any are present, are combined to form one polygon. Each detected nose is labeled by such a polygon.

4. EVALUATION

To demonstrate the viability of this face detection algorithm, we trained our support vector machine and tested our algorithm on a database of images developed by Kodak from typical snapshots taken by amateur photographers. The wide variation in content, lighting, scale, and color provides a challenging and realistic test of our algorithm. Of the approximately 2000 color images, 1784 were used in the course of training our support vector machine, leaving the remaining 216 to be used in testing.

Figure 3 contains 4 images from the set of 216 used for testing. For the top-left image, the algorithm has correctly identifies 4 noses and misses the remaining nose present. For the top-right image, the algorithm correctly identifies both noses present and also falsely detects a hand as being a nose. In the bottom-left image, the algorithm correctly identifies both noses present and falsely detects the lower right-hand portion of the image as containing a nose. For the bottom-right image, the algorithm correctly identifies all 4 noses present.

In the test set, there were 266 noses present overall. Of these, 138, or 51.9%, were correctly detected. There were 406 false alarms among all of the images in the test set.

5. CONCLUSION

The results for this algorithm should be considered in the context of the challenging database that was used for testing. The algorithm is reasonably successful despite the immense variation of the database without resorting to such computationally costly measures like image rescaling. While the false alarm rate per detection is high, it should be noted that the proposed algorithm is intended to be a part of an overall face detection method that looks for the presence of several facial-features, such as eye, nose, mouth to determine the presence or absence of a face. Many of the nose false alarms would most likely be eliminated after combining this algorithm with other facial-feature detectors.

6. ACKNOWLEDGMENT

The authors would like to thank Dr. Hsien-Che Lee of the Eastman Kodak Company for his help in developing the proposed algorithm.

7. REFERENCES

[1] R. Brunelli and T. Poggio, "Face recognition: feature versus templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042–1052, 1993.

[2] R. Challappa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–740, 1995.

[3] K. C. Yow and R. Cipolla, "Feature-based human face detection," *Image and Vision Computing*, vol. 15, pp. 713–735, 1997.

[4] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer-Verlag, 1995.

[5] B. Schölkopf, K. Sung, C. J. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik, "Comparing support vector machines with gaussian kernels to radial basis function classifiers," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2758–2765, November 1997.

[6] E. E. Osuna, R. Freund, and F. Girosi, "Support vector machines: Training and applications," tech. rep., (available via http) Massachusetts Institute of Technology, March 1997.

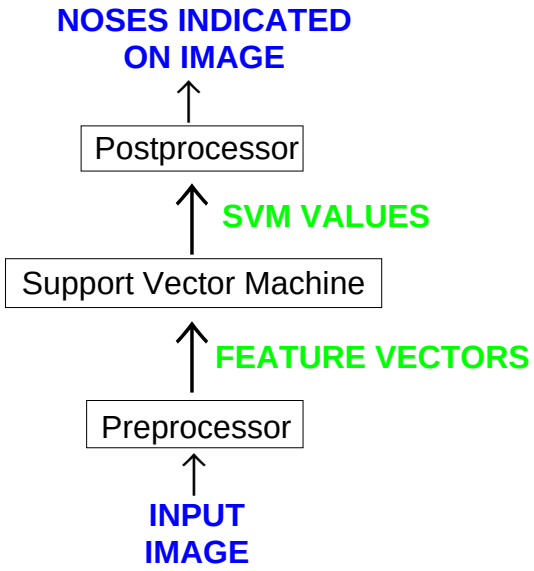

Figure 1: A block diagram of the proposed nose detection algorithm.

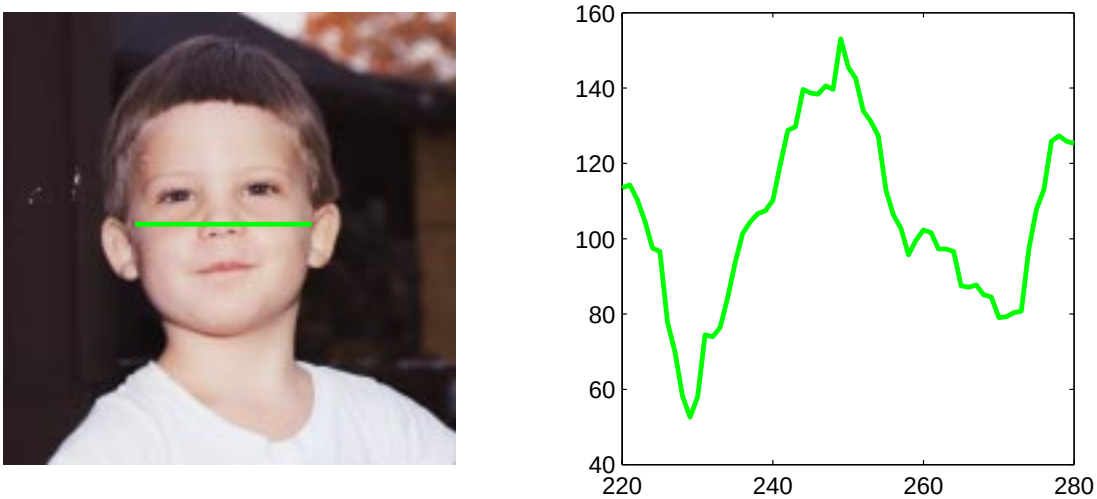

Figure 2: A typical image containing a face is shown on the left. A close–up of the cross-section indicated in the image is shown on the right.

Figure 3: 4 images from the set of test images processed by the proposed nose detection algorithm. The noses detected by the algorithm are indicated by the polygons superimposed on the images.