Neural Network Regression of Eyes Location in Face Images

Krzysztof Rusek and Piotr Guzik

AGH University of Science and Technology, Department of Telecommunications, Krakow, Poland
krusek@agh.edu.pl, guzik@kt.agh.edu.pl
http://kt.agh.edu.pl

Abstract. Automatic eye localisation is a crucial part of many computer vision algorithms for processing face images. Some of the existing algorithms can be very accurate unfortunately at the cost of computational complexity. In this paper the new solution to the problem of automatic eye localisation is proposed. Eye localisation is posed as a nonlinear regression problem solved by standard feed-forward multilayer perceptron (MLP) with two hidden layers. Additionally the procedure for artificial training samples generation is proposed.

The input feature vector is constructed from coefficients of two dimensional discrete cosine transform (DCT) of a face image. Both, the feature extraction and neural network prediction have known efficient implementations, thus the entire procedure can be very fast.

Obtained results indicate that the accuracy of the proposed approach is comparable or better than existing ones.

Keywords: eye localisation, neural network, DCT, computer vision.

1 Introduction

The problem of automatic eye localisation had arisen several years ago, along with the invention of automatic facial recognition algorithms. In fact it was the facial recognition research that motivated the development of eye localisation algorithms. It was quickly realised that the accuracy of most of facial recognition algorithms strongly depends on appropriate alignment of the detected face. This problem was broadly investigated in [1]. Sometimes satisfactory results can be obtained without face alignment [12]. Having said that it is reasonable to assume, that additional alignment can only improve the accuracy. The face alignment with use of the position of eyes seems to be a natural choice here.

Although the automatic eye localisation problem has been investigated for a few decades now, its performance still is not as good as one would need. The main problem with this task is a variability of eye appearance. The eye appearance on an image not only depends on a personal differences but also on the lighting conditions, pose, expression, etc.
In this article we present another approach to the problem of eye localisation with the problem stated as a regression one. DCT coefficients are proposed as a feature vector and artificial neural network is used to find regression parameters.

The rest of the paper is organised as follows. The section 2 describes the related work on eye localisation, in 3 our solution is presented. Results are discussed in section 4. Finally, section 5 concludes the paper.

2 Related Work

The problem of finding eyes (or any part of the face) location in face image can be posed as a classification or a regression problem. In classification approach a classifier is trained to recognize whether a given part of an image contains an eye. Such classifier is then applied to sliding window at different scales. The combined results from neighbouring windows and scales define the final rectangle containing the eye. Any classifier can be used but since the algorithm is repeated many times, its speed is an important factor.

The most common approach is to use AdaBoost algorithm with Haar or LBP features. Exactly this technique is used in popular computer vision library OpenCV in an implementation of eye detection algorithm. All algorithms using template matching also fall into this category.

While the classification approach is much more common, the problem was also stated as a regression one [5]. In this formulation the eyes locations are estimated from the feature vector $x$ extracted from the face image.

Everingham and Zisserman limited their analysis to kernel regression model and subset of pixels being the features. In this paper we use the similar approach but instead of using a subset of image we reduced the dimensionality by taking most important DCT coefficients as a feature vector. Having performance, accuracy and model simplicity under consideration we decided to use an artificial neural network for regression. Similar approach was presented in [8] where authors suggested using MLP with DCT from YCbCr planes for classification.

3 Proposed Approach

Following Everingham and Zisserman [5] we formulate the problem of finding the eye location as a multinomial regression problem. Having a feature vector $x$ we want to predict the output vector $y$ representing coordinates of both eyes. This is the standard regression problem and variety of solutions were proposed to deal with that. Among them, artificial neural networks and support vector regression (SVR) are state-of-the-art nonlinear regression algorithms. Since SVRs are commonly defined only for one dimensional predicted variable they would require four independent models, one for each component. Thus, we used neural networks. In such a case the single model gives multidimensional prediction. Furthermore it is significantly more compact, and hence faster to evaluate, than a support vector machine having the same generalization performance [2]. Usefulness of a neural network for such task is presented in [9], where authors used neural network to refine eye location indicated by different algorithm.