Automatic Eye Detection Using Semivariogram Function and Support Vector Machine

João Dallyson Sousa de Almeida[1], Aristófanes Corrêa Silva[2], Anselmo Cardoso Paiva[3]
Federal University of Maranhão - UFMA
Applied Computing Group - NCA/UFMA
São Luís-MA, Brazil
jdallyson@yahoo.com.br[1], ari@dee.ufma.br[2], paiva@deinf.ufma.br[3]

Abstract—Several computational systems which depend on the precise location of the eyes have been developed in the last decades. Aware of this need, we propose a method for automatic detection of eyes in images of human faces using semivariogram functions and support vector machines. Patterns of eyes region and other areas of the face are represented through the semivariogram function to precisely distinguish the region of the eyes. The method was tested in the ORL human face database, which contains 400 images grouped in 40 persons, each having 10 different expressions. The detection obtained results of sensitivity of 84.6%, specificity of 93.4% and accuracy of 88.45%.

Keywords- Automatic eye detection, semivariogram function, support vector machine (SVM).

I. INTRODUCTION

Many applications need to detect eyes, among which we may cite: determination of facial characteristics, biometric systems based on face and iris, analysis of facial expressions, monitoring of drivers' tiredness. This detection is fundamental no just to analyze eyes (open, partially open, closed, sad, scary, etc.) but to supply the position of the mouth and the nose, and constitutes a complex problem due to variation on the illumination, background and face expressions. From the geometrical characteristics, contrast and eye movement, information can be extracted for the iris recognition and detection systems and human machine interface.

The main difficulty in locating the eyes in human face images is the variety of facial expressions and positions that a person can show by the moment of the detection of the eyes.

The analysis of texture in digital facial images is very important in the detection of eyes. This analysis is accomplished usually with classical measures of image processing, as Spatial Gray Level Dependence Method, Gray Level Difference Method, template matching, wavelets, Zernike moment etc.. In geostatistics, there are functions denominated semivariogram that are frequently applied to the analysis of soil texture. The contribution of this work is to apply these functions to verify their accuracy in the discrimination and classification of faces regions as eye and non-eye using Support Vector Machine [17].

In this paper, we propose the analysis of texture by means of the semivariogram function. This function is widely used in Geographical Information Systems (GIS), but has not yet been explored for the eye detection in facial image. This function has the advantage of analyzing simultaneously the variability and the space correlation of pixels, and they work with textures in 2D and 3D. The textural measure is used as input features for a Support Vector Machine.

This paper is organized as follows. In Section 2 we present some works related to eye detection in facial image. In Sections 3 and 4, the techniques for feature extraction and classification are presented. Next, in Section 5, the results are shown and the application of the techniques under study is discussed. Finally, Section 6 presents some concluding remarks.

II. RELATED WORKS

In this section we present some works that have been developed to automatically locate the eyes in digital images. In [1] it is presented a method for detection of eyes in digital facial images using Zernike moment with Support Vector Machine (SVM). There, the eye and non-eyes patterns are represented in terms of the magnitude of the Zernike moment and classified by the SVM. The method achieves matching rates of 94.6% for detection of eyes in face images from the ORL human face database [12].

With similar objective, the authors in [2] made use of the template-based method to find the center of the iris, achieving matching rate of 95.2% for images without glasses from the ORL image base.

In [3] a probabilistic classifier is used to find the region of the eyes. The detection of the eye is done by the extraction of features using Nonparametric Discriminant Analysis (NDA) and the AdaBoost classifier. For such, 500 pairs of eye images from the FERET base were used obtaining a global matching rate of 94.5%.

The authors in [4] proposed a method for real-time monitoring of drivers' tiredness. It uses features such as:

eyelid, look, and head movement and facial expressions. A probabilistic model was developed using two templates, one rectangular and the other inserted in the center of the rectangular template, based on the visual suggestions to predict tiredness. The method obtained matching rate of 96.4% while detecting the glint of the pupil.

In [5] a method for detection of eyes in face using grayscale images, binary template matching and support vector machine (SVM) is developed. First, the image is improved using homomorphic filter and then binarized. All the candidates to be a pair of eyes are rescaled to a pre-determined size and sent to an SVM classifier that verifies the candidates and obtains the pairs of eyes. This method obtained precise eye detection rate of 96.8% for 1,521 images in the BioID faces base.

In [6] it is presented a method for detection of eyes in faces using Wavelets and Back Propagations Neural Networks. The inputs of the Neural Network are the maximum coefficients of the Wavelet vicinity. The output of the network is the classification as eye or non-eye. The results obtained for the ORL image base was 88% of matching.

In [7] it is proposed a method for identification of persons through the analysis of iris texture using the geostatistical functions of semivariogram and correlogram. The method obtained a success rate of 98.14% using an iris database called CASIA.

III. Semivariogram Function

Geostatistics is the analysis of spatially continuous data. It treats geographic attributes as random variables which depend on joint distributions on their locations. The semivariogram functions summarize the strength of associations between responses as a function of distance, and possibly direction [8] [9] [10].

Semivariance is a measure of the degree of spatial dependence between samples. The magnitude of the semivariance between points depends on the distance between the points. A smaller distance yields a smaller semivariance and a larger distance results in a larger semivariance. The plot of the semivariances as a function of distance from a point is referred to as a semivariogram.

A semivariogram (and also the functions studied in the sections below) has three main features: its sill, range, and nugget. The sill is the ordinate value at which the semivariogram levels off, that is, its asymptotic value; the range is the distance at which this leveling off occurs, that is, the spatial extent of the structure in the data; and the nugget is the semivariance at a distance 0.0, that is, the intercept. A nonzero nugget can imply either intrinsic variability in the data (the component typically ascribed to sampling error), or it might indicate that the sampling was conducted at an inappropriate spatial scale, that is, there is considerable variability at scales smaller than the smallest between-point distance.

The semivariogram is defined by

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (x_i - y_i)^2 \tag{1}$$

where h is the lag (vector) distance between the head value (target pixel), y_i, and the tail value (source pixel), x_i, and $N(h)$ is the number of pairs at lag h.

IV. Materials and Methods

Fig. 1 illustrates the proposed methodology consisting of two stages: training and test. The steps of pre-processing with homomorphic filter, extraction of features through the semivariogram and the classification with SVM are executed in both stages. Besides those, in the training we have the manual segmentation of the region of interest and the selection of the most significant features through Fisher's stepwise discriminant analysis [11]. In the test stage we still have: the automatic extraction of the region of the eyes through the projection gradient and segmentation of the eye candidates by applying region growing [13].

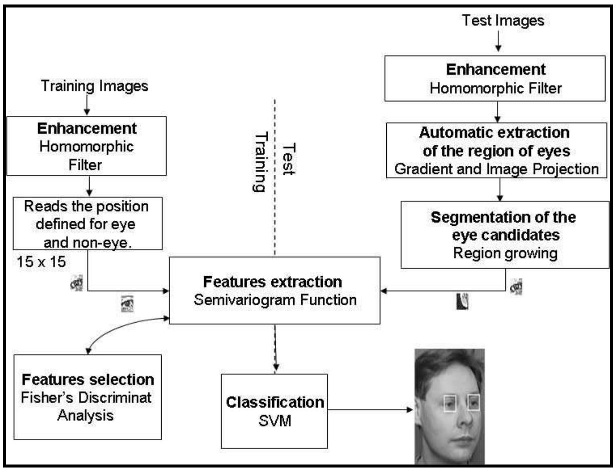

Figure 1: Proposed methodology.

A. Database

The ORL [12] image base was used for both stages. It is formed by 40 people with different facial expressions, hair styles, illumination conditions, with and without glasses, adding up to 400 grayscale images with 92x112 pixels. In the present work, we have used 164 images for training and 327 for testing.

B. Training Methodology

Initially, the images pass through a pre-processing using homomorphic filter [5] in order to solve luminosity divergences. From the 164 training images, 24 regions were manually selected, which were formed by 15x15-pixel windows, being 18 associated to the eye class and 6 associated to the non-eye class (eyebrows, glasses, ears, etc.) in each of them. Next, from each window texture features are extracted using the semivariogram function explained in Section 3.

The parameters used by the semivariogram functions for extraction of features were the directions 0º, 45º, 90º and 135º with angular tolerance of 22.5º and lag (distance) increment of 1, 2 and 3, corresponding to 14, 7 and 4 lags and tolerance of each lag distance of 0.5, 1.0 and 1.5, respectively. The adopted directions are those mostly used in the literature for image analysis. In order to choose the lag tolerance according to Isaaks and Srivastava [14] the commonest choice is to adopt half the lag increment.

Finally, the selection of the most significant variables is done by using Fisher's stepwise discriminant analysis [11] through the software Statistical Package for the Social Sciences (SPSS) [15]. We used cross validation leave-one-out [16] in the elaboration of the training model to evaluate its discrimination power.

We selected 26 features from 100 – 4 directions x 25 lags (14+7+4), corresponding to the increments 1, 2 and 3, respectively. After the selection of variables, the samples are trained in the SVM.

C. Test Methodology

In the test stage we have the automatic detection of eyes. It is composed by following steps: pre-processing using homomorphic filter, automatic extraction of the region of the eyes by adapting the method proposed in [2] and segmentation of eye candidates through the adaptation of the method proposed in [1].

The automatic extraction of the region of the eyes aims to reduce the search space by generating a sub-image with the possible region of the eyes and excluding regions with no interest (mouth, nose, hair and background). In the detection of the region of the eyes, a smoothing is initially executed using the Gaussian filter of 3 x 3 mask, and next, we calculate the gradient of the input image by using the Sobel filter [13]. In Fig. 2c we have an example of the calculation of the gradient magnitude applied to Fig. 2b.

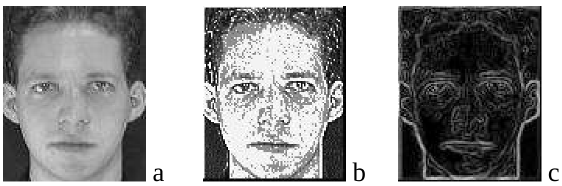

Figure 2: (a) Original image. (b) Image 2a after application of homomorphic filter. (c) Image after application of Sobel filter on 2b.

We apply a horizontal projection of this gradient, obtaining as result the mean of the three higher peaks of this projection. Knowing that the eyes are found in the superior part of the face and that joined to the eyebrows they correspond to the 2 peaks closer to each other, this physiologic information, a priori known, can be used to identify the region of interest. The peak of the horizontal projection will provide the horizontal position of the eyes. We apply a vertical projection for all the pixels in the horizontal region, and the valley of this projection will provide the center of the face. At the same time, we apply a vertical projection in the gradient image. There are two peaks in the left and in the right that correspond to the limits of the face. From these limits, the length of the face is estimated. Combining the results of the projections, we achieve an image that corresponds to the region of the eyes.

In the segmentation stage we used the minimal filter [13] sizing 7x7 in the detected region of the eyes image. Next, we performed the thresholding using the mean of the intensity of the pixels as threshold. In the binarized image resultant from the thresholding we apply the region growing technique [13] to locate the center of the objects. This very same center will be the center of the 15x15 windows that is formed to extract the region to which the semivariogram function is applied.

The samples generated by the application of the semivariogram function are submitted to classification by the already trained SVM. After executing this stage, we evaluated the results using the measurements sensitivity, specificity and accuracy.

Sensitivity is defined by TP/(TP + FN), specificity is defined by TN/(TN + FP), and accuracy is defined by (TP + TN)/(TP + TN + FP + FN), where TN is true-negative, FN is false-negative, FP is false-positive, and TP is true-positive.

V. Results and Discussion

The proposed methodology was evaluated with use of the ORL image base. First, we tested 400 images in the stage of automatic extraction of the region of the eyes. In this test, it was possible to detect the region of 81.5% of the images, that is, 327 images. In Fig. 3 we show some examples of successfully detected regions. On the other hand, in Fig. 4 we have examples of images for which the detection failed. We observed that most of the errors occurred due to the position of the face, especially when the face was turning aside, and due to the various types of glasses and hair styles.

Figure 3: Region of the eyes correctly detected.

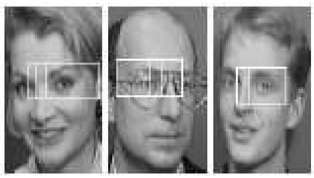

Figure 4: Failure in the detection of the region of eyes.

With the reduction of the search space through the detection of the region of eyes, we start the stage of eyes detection. For this stage we used 164 random images from the ORL base, what is equivalent to 50% of the images that had the region of eyes correctly detected, to train the SVM. From each image submitted to the training, 18 eye samples and 6 non-eye samples were taken, totalizing 3984 samples. Such samples were manually selected. The SVM library LIBSVM [16] was used for training and testing. A basic radial function was used as kernel and the used parameters were $C = 8$ and $\gamma = 2$.

To evaluate the efficiency of our proposal we applied the methodology to the 327 images on which the detection of the region of eyes succeeded. From the stage of detection of the possible eyes, we obtained 864 samples to be classified by the trained SVM. The result of the test was a sensitivity rate of 84.6%, specificity of 93.4% and accuracy of 88.45%.

Fig. 5 shows some examples of images for which the methodology succeeded in detecting the eyes, presenting *TP* of 413. On the other hand, in Fig. 6 we have examples where the methodology failed, obtaining *FP* of 25. Analyzing the results we notices that most of the errors occurred in the regions of eyebrows and structures of glasses.

Figure 5: Correct location of the eyes

Figure 6: Failure in the location of the eyes.

Analyzing the classification of the non-eye regions, we observed that the amount of TN and FN was of 351 and 75. In Fig. 7 we have examples of eye regions which were classified as non-eye. We can notice that the error occurred in the images whose eyes regions were darkened, becoming similar to eyebrow, hair or background regions.

Figure 7: Eye region classified as non-eye.

VI. Conclusion

In this paper we presented a methodology for detection of eyes in human faces, which can be applied to systems that need to locate the region of the eyes. This methodology is subdivided into training and detection of eyes. For both stages the pre-processing using homomorphic filter is performed, the semivariogram function is used as feature descriptor and the support vector machine is used in training and classification.

The results indicate that the technique used in the methodology is promising because it achieved matching rate superior to 88%. However, it is necessary to perform tests with other facial image bases, in order to confirm the efficiency of the proposed methodology.

References

[1] H. Kim and W. Kim. Eye Detection in Facial Images Using Zernike Moments with SVM. ETRI Journal, (2008). 30(2):335–337.

[2] K. Peng, L. Chen, S. Ruan and G. Kukharev. A robust algorithm for eye detection on gray intensity face without spectacles. Journal of Computer Science and Technology, (2005). 5(3):127–132.

[3] P. Wang, M. Green, Q. Ji, and J. Wayman. Automatic eye detection and its validation. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, (2005). pages 164–164.

[4] F. Jiao and G. He. Real-Time Eye Detection and Tracking under Various Light Conditions. Data Science Journal, (2007). 6(0):636–640.

[5] Q. Wang and J. Yang. Eye detection in facial images with unconstrained background. Journal of Pattern Recognition Research, (2008). 1(1):55.

[6] M. Motwani, R. Motwani and Jr, F. H. Eye detection using wavelets and ann. Proceedings of Global Signal Processing Conferences & Expos for the Industry (GSPx), (2004).

[7] O. S. Junior, A. C. Silva and Z. Abdelouah. Personal Identification Based on Iris Texture Analysis Using Semivariogram and Correlogram Functions. International Journal for Computacional Vision and Biomechanics, (2009). 2(1).

[8] I. Clark. Practical Geostatistics. Applied Sience Publishers, London, (1979).

[9] N. A. C. Cressie. Statistical for Spatial Data. John Wiley & Sons, New York., (1993).

[10] A. G. Journel and C. J. Huijbregts. Mining Geostatistics. Academic Press, London, (1978).

[11] P. Lachenbruch and M. Goldstein. Discriminant analysis. Biometrics, (1979). pages 69–85.

[12] A. L. Cambridge. ORL Face Database:. Database available at http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase. html, (2009).

[13] R. Gonzalez and R. Woods. Digital image processing, (2002). ISBN: 0-201-18075-8.

[14] E. Isaaks and R. Srivastava. An introduction to applied geostatistics, (1989).

[15] L. Technologies. SPSS forWindows vs 12.0. LEAD Technologies, (2003).

[16] C. Chang and C. Lin. LIBSVM: a library for support vector machines. Software available at http://www. csie. ntu. edu. tw/cjlin/libsvm, (2001). 80:604–611.

[17] C. J. C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. Kluwer Academic Publishers, 1998.