FACE DETECTION ALGORITHM USING EMD

Application to Biometric Recognition System

Jordi Solé-Casals
Digital Technologies Group, University of Vic, Vic, Spain

Carlos M. Travieso-González, Marcos del Pozo and Jesús B. Alonso
Signals and Communications Department, Institute for Technological Development and Innovation on Communications (IDETIC), University of Las Palmas de Gran Canaria, Las Palmas de GC., Spain

Keywords: Face detection, Empirical Mode Decomposition (EMD), Biometrics.

Abstract: In this work we present a new face detection algorithm for biometric recognition systems, based on Empirical Mode Decomposition (EMD) that helps us to detect the central skin part of the face. After a binarization and dilatation process, we get an approximation of the face location that is finally established by adjusting an ellipse. Finally, a mask is created using the perimeter of this ellipse and applied to the original image, giving as a result the detected face. Experiments show the performance of the method, with a recognition rate of 97.60%.

1 INTRODUCTION

Face detector is an important step towards biometric face recognition. If we have a person in an image and we want to recognise it, the first step is to detect the face in order to apply then a biometric system.
Of course, the most interesting possibility is to do it automatically, without any manual step. This can be a requirement in real time systems, or when we have an important amount of pictures to analyze.

Different work has been done in this area. Some references are showed in the next paragraphs.

In (Yongqiu et al., 2010) a combination of three classifiers is used: skin color detector, AdaBoost detector based on haar-Iike features, and eye-mouth detector. A semi-serial architecture is designed to combine the three detectors. They selected 103 pieces of color images including faces, and 100 pieces of color images of background from their own database. Success rate was 85.5%.

The (Qiang-rong and Hua-lan, 2010) paper proposes a novel face detection algorithms based on combining skin color model, edge information and features of human eyes in color image The size of the database was 544 samples, reaching a recognition rate up to 94.9%.

In (Venkatesh and Marcel, 2010) authors use an additional classifier that predicts the bounding box of a face within a local search area. Then a face/non-face classifier is used to verify the presence or absence of a face. The database is mixed from CMU+MIT and Fleuret face databases, with a total amount of 375 images. Authors have improved between 15-30% in detection rate or speed when compared to the standard scanning technique.

On the other hand in (You-jia and Jian-wei, 2010) authors propose a rotation invariant multi-view color face detection method combining skin color segmentation and multi-view AdaBoost algorithm. The database is composed by 153 internet images. Authors reached, with a simple background, 98.04% of recognition rate; and a smaller value of 82.35% with complex background.

In (Zhao et al., 2010) authors present a combination of Contrast-limited Adaptive Histogram Equalization (CLAHE) and multi-step integral projection. The JAFFE database was used and they obtained a recognition rate of 95.318%.

(Yihu et al., 2010) have designed a novel scheme integrating skin color segmentation and facial component localization to detect face. They have used 214 PC camera images, and reached a 95.8% of detection rate, and for 218 Web images, a 93.5% of

Solé-Casals J., Travieso-González C., del Pozo M. and Alonso J..
FACE DETECTION ALGORITHM USING EMD - Application to Biometric Recognition System.
DOI: 10.5220/0003297305650569
In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* (MPBS-2011), pages 565-569
ISBN: 978-989-8425-35-5

detection rate.

In (Zhang and Shi, 2008) authors introduce a new nonlinear transformation method of the *YCbCr* color skin space, which is the color segmentation of the human face for regional analysis and extraction. They implemented image rotation and template matching. The database is composed by 400 images. The success rate for Face Average Brightness was 92%.

Face detection method for the color image with complex background is presented in (Xue-wu et al., 2009), which is a mixed skin-color segmentation model in both *YCbCr* and *HIS* color space constructed. Authors used 95 pictures with different brightness for testing. Result shows a success detecting rate over 91.7%.

In our work, the proposed algorithm can be summarized by the following block diagram (see Figure 1):

On each image $x(t)$ we apply Empirical Mode Decomposition (EMD) algorithm in order to decompose the image to its Intrinsic Mode Functions (IMFs). First 50% of the modes are combined together in order to generate a new image $y(t)$ that contains basically the different regions present into the original image.

Afterwards, we apply a binarization process followed by a dilatation, in order to emphasize the central part of the skin of the face against all the other parts of the image. The result is image $z(t)$.

Finally, we detect the region of interest (ROI) and we fit to it an ellipse in order to obtain face's borders (image $w(t)$) and generate a mask that indicates which pixels belongs to the face (with value equals to 1) and which not (with value equals to 0).

The output image $s(t)$ is then generated by applying this mask to the original image.

This paper is organized as follows: After this introduction, EMD technique is presented in Section 2 with details on how to use it in our case; Section 3 is devoted to the image processing step, whereas ROI procedure and boundary calculations are explained in Section 4. Experiments and results are shown in Section 5; and finally conclusions are presented in Section 6.

2 EMPIRICAL MODE DECOMPOSITION

The first block of the proposed method (see Figure 1) consists in applying the Empirical Mode Decomposition (EMD) technique to the input image $x(t)$.

EMD is a relatively new data analysis method that decomposes the signal into waveforms modulated in both amplitude and frequency by extracting all of the oscillatory modes embedded in the signal.

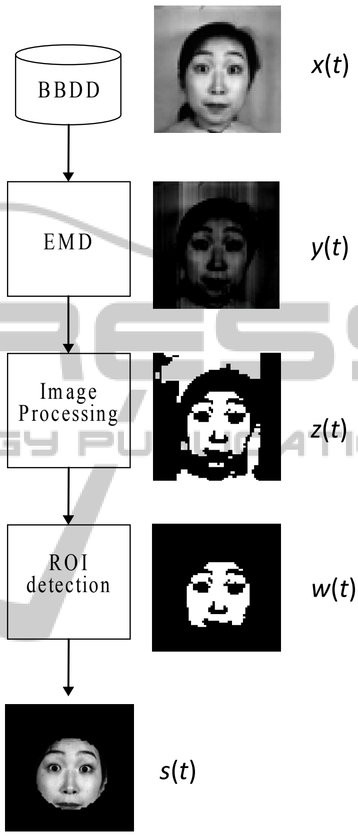

Figure 1: Block diagram of the proposed method.

2.1 Obtaining Intrinsic Mode Functions

As explained in (N. Huang, et al., 1998), the decomposition can be viewed as an expansion of the data in terms of the so called Intrinsic Mode Functions (IMF's) that are the waveforms extracted by EMD. Each IMF is symmetric and it is assumed to yield a meaningful local frequency traces. Different IMF's do not exhibit the same frequency at the same time. Then, these IMF's, based on and derived from the data, can serve as the basis of that expansion which can be linear or nonlinear as dictated by the data, and it is complete and almost orthogonal.

The decomposition is an intuitive and adaptive signal-dependent decomposition and does not require any conditions about the stationarity and linearity of the signal.

IMF in each cycle is defined by the zero crossings. Every IMF involves only one mode of oscillation, no complex riding waves are thus allowed. Notice that the IMF is not limited to be a narrow band signal, as it would be in traditional Fourier or wavelets decomposition. In fact, it can be both amplitude and frequency modulated at once, and also non-stationary or non-linear.

The process of IMF extraction, from a signal $x(t)$, known as *sifting process* (N. Huang, et al., 1998), is based on the following steps:

1. Determine the local maxima and minima of the analyzed signal $x(t)$.
2. Generate the upper and lower signal envelopes by connecting those local maxima and minima respectively by the chosen interpolation method (e.g., linear, spline, cubic spline, piece-wise spline).
3. Determine the local mean $m(t)$ by averaging the upper and lower signal envelopes.
4. Subtract the local mean from the data: $h_1(t) = x(t) - m_1(t)$.
5. Test if $h_1(t)$ is an IMF.
    - If yes, stop the procedure. The first IMF labelled as $c_1(t)$ is already obtained.
    - If not, replace $x(t)$ with $h_1(t)$ and repeat the procedure from step 1.

In order to obtain the second IMF, one applies the sifting process to the residue $r_1(t) = x(t) - c_1(t)$, obtained by subtracting the first IMF from $x(t)$; the third IMF is in turn extracted from the residue $r_2(t) = r_1(t) - c_2(t)$ and so on.

One stops extracting IMF's when two consecutive sifting results are close to identical. Then, the empirical mode decomposition of the signal $x(t)$ may be written as:

$$x(t) = \sum_{i=1}^{n} c_i + r_n \qquad (1)$$

Thus, we obtain a decomposition of the data into n-empirical modes, and a residue, r_n, which can be either the mean trend or a constant.

2.2 Applying EMD Procedure

Decomposing the input image $x(t)$ whit EMD procedure explained before will result in a decomposition of the signal in a set of IMF's plus a residue, as shown in equation 1.

As the first modes capture high frequencies of the signal, we will generate a new signal $y(t)$ (pre-processed image) by taking into account only the first 50% of the modes. This pre-processed image is obtained by adding the considered mode:

$$y(t) = \sum_{i=1}^{n/2} c_i \qquad (2)$$

3 IMAGE PROCESSING

The image processing block uses classical image processing techniques that are necessary in order to facilitate the detection of the Region Of Interest (ROI) that will be detailed in the next section.

Using the image $y(t)$ obtained from EMD step, we apply a binarization procedure in order to have only two different values for the pixels. Homogenous region of the face will basically be in white color, while most of the rest of the image will be in black color. Edges are easily detected after binarization, as for example those due to eyes, nose, mouth, hair, etc.

In order to group the possible unconnected regions of the face, and taking into account that our goal is to detect the position of the face into the image, we apply a dilatation process with a rectangle of size 1x2 as structural element. This step is very useful as very often eyes, nose and mouth breaks the face into different unconnected regions. The output signal of this step will be black and white image $z(t)$ (see figure 1).

4 ROI DETECTION

The goal of this last block is to determine which part of the image is the face and which part is not.

In order to implement the ROI detector, we label all the connected regions of the image $z(t)$ and we analyze the different obtained regions:

The biggest region is due to the background part of the image, as in all the images the face is always smaller than the rest of the pixels.

The second biggest region is habitually the region that we are looking for, as the face (especially due to the dilatation process) is an important part of the image. But sometimes it may happen that the frontal region (upper part), the neck (lower part), or other parts of the image are larger than the region of the face itself (central region).

In order to avoid a mistake selecting the ROI, we propose the following procedure:

1. Sort of all the regions by the total number of pixels that contains;
2. Select the second bigger region (the first one is omitted as it corresponds to background);
3. Calculate the mass centre of the region;
4. If the mass centre is located near the centre of the image, take this region as ROI and stop the procedure;
5. If not, select the next important region and go to point 3 until a ROI is detected.

Once we have a ROI, the next step is to fit an ellipse that will define the portion of the image considered as face and the rest of the image. In order to calculate the axes and centre of the ellipse, we obtain the coordinates of the minimum and maximum of rows and columns of the ROI, which will be used in order to calculate the centre and the axes of the ellipse. Using them we get the borders of the ellipse that fits the ROI (see $w(t)$ in Figure 1), and we create a mask with the aim of extracting the face of the image by an AND operation (pixel by pixel) between the original image $x(t)$ and the mask. The resultant image is the output of the signal, labelled as $s(t)$ in Figure 1.

5 EXPERIMENTS

In order to test our algorithm we apply the proposed method to the entire JAFFE database that contains 213 different images.

As our method is not based on any learning paradigm, we just take (randomly) 5 images for adjusting some few parameters of the method (number of considered modes and threshold for binarization process), keeping the rest of the images (208) for testing it.

Figure 2: Up: original faces from JAFFE database. Down: Correct face detection obtained using the proposed method.

Figure 3: Up: original faces from JAFFE database. Down: Bad face detection obtained using the proposed method.

Considering that the system works correctly if eyes, nose and mouth falls into the detected face, and works improperly if some (or all) of these elements fall out of the ROI, we get 203 good detections and 5 bad detections. Calculating the quotient between good detections and total number of available images (excluding the images used for training process), we obtain a recognition rate of 97.60%.

In order to show different obtained results, we present some of the correctly detected faces in figure 2, and incorrectly detected faces in figure 3.

As we can see in these figures, the method fails sometimes and takes as face a wrong region. The principal problem is due to the disconnected regions that can appear after EMD. This is why this point must be developed in more detail as is the key factor in order to obtain good results.

On the other hand, different illumination or background conditions may affect results. More research must be done for these cases in order to get a general face detector system.

We can compare our method versus the method proposed in (Zhao et al., 2010) because in both cases the same database (JAFFE database) is used. Our proposal improves over 2% the detection face rate. Therefore our method shows better proprieties versus (Zhao et al., 2010) proposal.

6 CONCLUSIONS

The proposed method for face detection presented in this work is based in EMD technique, and uses very popular and simple image tools. Using EMD, the central part of the skin on the face is easily identified, and then a binarization and dilatation process are applied in order to localize the face on the image. Finally, an ellipse is fitted and a mask is created to be used on the original image for extracting the face.

Some of the aspects of the algorithm can be explored in more detail. For example, the number of considered modes (IMFs) is an important and critical parameter. More experiments, with other databases, must be done in order to determine the best number of modes, in a way that this number is independent of the image (database).

Also in this preliminary work we have not tested other possible structural elements for the dilatation step, but it is an interesting point to investigate in future works, as the success of the results depends on the capability of this step. If we can merge the different parts of the face in a sole region, the procedure works properly; but if the face is split in many different parts, the systems fails and face is not correctly detected.

On the other hand, we have not explored yet the possibility of having more than one face in the image, but this is another interesting point to investigate in detail.

Finally, future work must be done in order to test the procedure in other situations, as for example in bad illumination conditions, dark or shadow places, brilliant places, etc. Of course, these scenarios, where illumination can be very different, are not easy, but they are realistic and must be taken in consideration for real world applications.

We are already conducting investigations following these points, and preliminary results for some cases are very promising.

ACKNOWLEDGEMENTS

This work has been partially supported by the University of Vic under a mobility grant for Dr. Jordi Solé-Casals, by "Cátedra Telefónica ULPGC 2009-10", and by the Spanish Government under funds from MCINN TEC2009-14123-C04-01.

REFERENCES

Huang N., Shen Z., Long S., Wu M., Shih H., Zheng Q., Yen N. C., Tung C., and Liu H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. of the Royal Society A*, vol. 454, no. 1971, pp. 903–995,

Qiang-rong J., Hua-lan, L., 2010. Robust human face detection in complicated color images. *The 2nd IEEE International Conference on Information Management and Engineering,* pp.218-221

Venkatesh, B. S., Marcel, S., 2010. An alternative scanning strategy to detect faces. *IEEE International Conference on Acoustics, Speech and Signal Processing.* pp. 2122-2125

Xue-wu, Z., Ling-yan, L., Dun-qin, D., Wei-liang, X., 2009. A novel method of face detection based on fusing YCbCr and HIS color space. *IEEE International Conference on Automation and Logistics,* pp.831-835

Yihu, Y., Daokui, Q., Fang, X., 2010. Face detection method based on skin color segmentation and facial component localization. *2nd International Asia Conference on Informatics in Control, Automation and Robotics,* vol.1, pp.64-67

Yongqiu, T., Faling Y. Guohua, C., Shizhong, J., Zhanpeng H, 2010. Fast rotation invariant face detection in color image using multi-classifier combination method. *International Conference on E-Health Networking, Digital Ecosystems and Technologies,* pp.211-218

You-jia,F, Jian-wei, L. 2010. Rotation Invariant Multi-View Color Face Detection Based on Skin Color and Adaboost Algorithm. *International Conference on Biomedical Engineering and Computer Science*, pp.1-5

Zhang, Z., Shi, Y., 2008. Face Detection Method Based on a New Nonlinear Transformation of Color Spaces. *Fifth International Conference on Fuzzy Systems and Knowledge Discovery, 2008.* vol.4, pp.34-38

Zhao, Y., Georganas, N. D., Petriu, E. M., 2010. Applying Contrast-limited Adaptive Histogram Equalization and integral projection for facial feature enhancement and detection. *Instrumentation and Measurement Technology Conference*, pp.861-866.