LOCATING THE HUMAN EYE USING FRACTAL DIMENSIONS

Kwan-Ho Lin, Kin-Man Lam and Wan-Chi Siu

Centre for Multimedia Signal Processing
Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

In this paper, a new method for locating eye pairs based on valley field detection and measurement of fractal dimensions is proposed. Fractal dimension is an efficient representation of the texture of facial features. Possible eye candidates in an image with a complex background are identified by valley field detection. The eye candidates are then grouped to form eye pairs if their local properties for eyes are satisfied. Two eyes are matched if they have similar roughness and orientation as represented by fractal dimensions. We propose a modified approach to estimate the fractal dimensions that are less sensitive to lighting conditions and provide information about the orientation of an image under consideration. Possible eye pairs are further verified by comparing the fractal dimensions of the eye-pair window and the corresponding face region with the respective means of the fractal dimensions of the eye-pair windows and the face regions. The means of the fractal dimensions are obtained based on a number of facial images in a database. Experiments have shown that this approach is fast and reliable. This shows that the texture of the eyes can be represented very well by fractal surfaces.

1. INTRODUCTION

Facial feature extraction is a challenging task in human computer studies, and also an essential step in face recognition. Most of the facial feature extraction techniques that have been reported [1][2] are computationally intensive or require the location of the face contour in advance. Yang *et al.* [3] proposed a hierarchical knowledge-based system for face detection, which consists of three levels of detection. Levels 1 and 2 of the system use mosaic images of different resolutions. In these two levels, two sets of rules are defined for the mosaic images, which are based on the characteristics of the human face region. Possible face candidates from these two levels are then further verified at the third level by comparing them with the edges of the facial features. However, this method is unable to detect or locate a tilted human face reliably. The example-based learning approach [4] is a method for locating the vertical frontal views of human faces in complex scenes. A classification procedure is trained based on a number of "face" and "non-face" examples. Six "face" clusters and six "non-face" clusters are obtained based on 4,150 normalized face patterns. From these clusters, faces are detected by matching the window patterns at different image locations and scales with the distribution-based face and non-face models. Although this approach can achieve a high detection rate, it is computationally intensive. In this paper, we will present a fast approach based on valley field detection and fractal dimensions (*FD*) [5] to extract eye pairs in a complex background, which can then be used to represent a face region.

Fractals are highly detailed, complex geometric shapes. One measure of the complexity of an image is fractal dimension, which can be estimated using the box-counting [6][7] technique. Advanced research and development in fractal analysis have practical applications in video coding, pattern recognition and natural texture analysis [8], etc.

The approach to locating eye pairs proposed in this paper consists of three levels. The first level is to detect in an image the location of the possible eye candidates, which exhibit themselves as valleys in the image space. These valleys are used to define the possible eye candidates. Valley positions that have features similar to the eyes are considered as possible eye candidates, and will be passed to level 2 to form possible eye pairs. Two possible eye candidates with similar fractal dimensions are grouped to form a possible eye pair in level 2. As fractal dimension is sensitive to lighting conditions, a normalization process is performed before estimation of the fractal dimension in order to reduce the effect of lighting. Furthermore, the two eyes of a valid pair should have similar orientations. An oriented fractal dimension is therefore devised to increase the matching accuracy. In stage three, in order to reduce the effect of uneven lighting conditions on an eye-pair window, its edge image is considered in computing the fractal dimension. Based on the input eye pair, the corresponding face region can be formed. The input eye-pair window and the corresponding face region are then binarized, and their respective fractal dimensions are measured. These two fractal dimensions are then compared with the average fractal dimensions of the eye-pair window and face region. The performance of this approach has been evaluated using the ORL database, MIT database, and various complex images. Experimental results show that the eye pairs can be detected reliably under different orientations, lighting conditions, and slight perspective variations.

The rest of the paper is organized as follows. Section 2 describes the box-counting method for estimating the fractal dimension. The details of the eye detection procedure based on valley field detection and measurement of fractal dimension are presented in Section 3. Experimental results will be given in Section 4. Finally, a conclusion will be drawn in Section 5.

2. ESTIMATION OF FRACTAL DIMENSION

A number of methods such as the Fourier power spectrum [8] and

This work is supported by RGC Grant B-Q322 and the Centre for Multimedia Signal Processing, PolyU.

box-counting [6][7] can be used to estimate the fractal dimension. The box-counting method is an efficient technique for estimating fractal dimensions, and it can be applied to gray-level images and binary images, as described below.

The differential box-counting approach [6] has been used to estimate fractal dimensions. This method counts the number of boxes that cover the image intensity surface. The surface can be considered as a 3D space in which the two coordinates (x,y) represent the 2D position and the third coordinate (z) represents the image gray-level intensity. For a given image of size $I \times I$, the image is partitioned into grids of size $S \times S$. The grids are numbered as (i,j), where $0 \le i,j < r$ and $r = \lfloor I/S \rfloor$. Each grid is stacked with a column of boxes of size $S \times S \times S'$. Suppose that the maximum gray-level intensity is G, then $\lfloor G/S' \rfloor = \lfloor I/S \rfloor$. The boxes on a grid are assigned a number with the box at the bottom as box one and the one on the top as $\lfloor G/S' \rfloor$. If the minimum and maximum gray-levels of the image in the $(i,j)^{th}$ grid fall in the boxes numbered k and l, respectively, then $n_r(i,j) = l - k + 1$, where $r = \lfloor I/S \rfloor$ and $0 \le i,j < \lfloor I/S \rfloor$. $n_r(i,j)$ represent the number of boxes covering the image intensity surface over the $(i,j)^{th}$ grid, as shown in Fig. 1. The total number of boxes required to cover the surface is $N_r = \sum_{i,j} n_r(i,j)$. With different grid size S, different values of r and N_r can be obtained. The fractal dimension can then be estimated from the least-square linear fit of $\log(N_r)$ against $\log(1/r)$.

For binary images, only two levels exist: black and white. A black pixel represents an image point or an edge point, while a white pixel is regarded as a point in the background. An image of size $I \times I$ is divided into grids of size $S \times S$. At each grid, the number of boxes, $n_r(i,j)$, that contain any black pixels is counted. The total number of boxes over all the grids is denoted as $N_r = \Sigma n_r(i,j)$, where $i,j, = 1,\ldots\ldots, I/S$ and $r = \lfloor I/S \rfloor$. The grid size $S \times S$ is changed such that different values of r and the corresponding total box count N_r are obtained. Again, the fractal dimension can be estimated by using the least-square linear fit of $\log(N_r)$ against $\log(1/r)$.

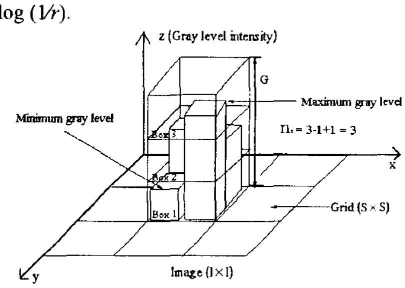


Figure 1: Estimation of fractal dimension for a gray-scale image using the box-counting technique

3. DETECTING THE EYE PAIRS

In this section, we propose a new method for locating the eye pairs in an image with a complex background. The detection is divided into three stages. First, all the valley regions in an image are detected and tested for possible eye candidates, which are selected based on a number of criteria. In the second stage, possible eye pairs are formed from the eye candidates based on eye features and oriented fractal dimension. The possible eye pairs are then further verified based on the fractal dimensions of the eye-pair windows and face regions in the third stage. The detailed procedures for locating the eye pairs and the face regions are described below.

3.1. Valley Detection of the Eyes

As the iris has a relatively low gray-level intensity in a human face, a valley exists at an eye region. The valley field, Φ_v, is extracted by means of morphological operators. A possible eye candidate is identified at position (x,y) if the following two criteria are satisfied:

$$f(x,y) < t_i \text{ and } \Phi_v(x,y) > t_v \quad (1)$$

where $f(x,y)$ is a facial image, and t_i and t_v are thresholds. A number of regions of possible eye candidates are detected, and are then reduced to a point by choosing the best candidate in each of the regions. Two functions, $v_1(x,y)$ and $v_2(x,y)$, are used to locate the best eye candidate in each region. The two functions are defined as follows:

$$\begin{aligned} v_1(x,y) &= C_{1,1}\Phi_{1,1}(x,y) + \\ &\quad C_{1,2}\left(\frac{f(x-2,y)+f(x+2,y)}{2} - S_{1,1}(x,y)\right) \\ v_2(x,y) &= C_{2,1}\Phi_{2,1}(x,y) + \\ &\quad C_{2,2}\left(\frac{f(x-3,y)+f(x+3,y)}{2} - S_{2,1}(x,y)\right) \end{aligned} \quad (2)$$

where C's are weighting factors, $\Phi_{1,1}(x,y)$ and $S_{1,1}(x,y)$ are the average valley intensity and the average gray-level intensity inside a 3×3 window, respectively, while $\Phi_{2,1}(x,y)$ and $S_{2,1}(x,y)$ are the corresponding values inside a 5×5 window.

3.2. Grouping the best eye candidates to form possible eye pairs

Two eye candidates are paired to form a possible eye pair. We assume that the rotation of the human face is less than 45°. The two eyes of a face should have similar orientations and be of similar size. Moreover, the inter-distance between the two eye candidates should not be less than a threshold, which also represents the size of the human face. Furthermore, the roughness of the two eye regions should be close to each other.

Based on the inter-distance between the two eye candidates, we can set up the corresponding windows to cover each of the two eye regions. The height, h, and width, w, of the windows are equal to a quarter of the inter-distance, L. As fractal dimension is sensitive to lighting conditions, the average gray-level intensities of the two eye windows are therefore adjusted to 180. This arrangement alleviates the effect of uneven lighting conditions on both halves of the face. When two eye candidates are paired, they should have similar fractal dimensions, as well as orientation. However, due to the use of square boxes in counting, fractal dimension lacks this information about orientation. In our approach, we use two rectangular boxes, namely a vertical rectangular box and a horizontal rectangular box of different orientations instead of a square box in box-counting. The two measured fractal dimensions are called horizontal fractal dimension, FD_h, and vertical fractal dimension, FD_v, respectively. Table 1 tabulates the horizontal and vertical fractal dimensions of the left and right eyes under different orientations for a number of facial images. The results demonstrate that the FD_h of the left and right eyes of a human face are very similar at a given orientation. The same applies to the FD_v. The two fractal dimensions, FD_h and FD_v, for a human face are different from each other, and change according to the orientation of the eyes. For any given image region, the texture inside will remain more or less the same under different rotations, but both the FD_h and FD_v will change. Nevertheless, when the FD_h decreases, the FD_v will increase, and vice versa. This is due to the fact that the total number of boxes required to cover the image area or space should change only slightly even when the image is rotated. As shown in Table 1, the sums of FD_h and FD_v for the eye samples under different rotations remain fairly constant. Thus, a valid eye pair can be selected if the following criteria are satisfied.

$$|FD_h(x_0,y_0)-FD_h(x_1,y_1)|<t_1 \text{ and}$$
$$|FD_v(x_0,y_0)-FD_v(x_1,y_1)|<t_2 \text{ and}$$
$$|(FD_h(x_0,y_0)+FD_v(x_0,y_0))-M_{eye}|<t_3 \text{ and}$$
$$|(FD_h(x_1,y_1)+FD_v(x_1,y_1))-M_{eye}|<t_4 \quad (3)$$

where (x_0, y_0) and (x_1, y_1) are the locations of the left- and right-eye candidates, M_{eye} is the average fractal dimension of the eye windows, and t_1, t_2, t_3, and t_4 are thresholds. For a valid pair, the respective differences should be less than certain thresholds. The possible eye candidates are then validated by the oriented fractal dimension to form possible valid eye pairs. Our experiments show that by using the oriented fractal dimension we can form eye pairs more reliably and can reduce the number of eye pairs by 50% when compared to cases where the conventional fractal dimension is used. The use of oriented fractal dimension provides a better pairing result due to the addition of information on orientation. As the oriented fractal dimension matches eye pairs more accurately, the total amount of computation required in stage three can be reduced.

Sample	FD_h of left eye	FD_v of left eye	FD_h of right eye	FD_v of right eye	FD_h + FD_v of left eye	FD_h + FD_v of right eye
C1	1.1696	1.1450	1.1701	1.1464	2.3146	2.3165
C2	1.1744	1.1516	1.1716	1.1441	2.3260	2.2957
C3	1.1786	1.1524	1.1858	1.1608	2.3310	2.3466
C4	1.1701	1.1406	1.1613	1.1505	2.3107	2.3118
C5	1.1769	1.1572	1.1776	1.1464	2.3341	2.3240

(a) Upright

Sample	FD_h of left eye	FD_v of left eye	FD_h of right eye	FD_v of right eye	FD_h+ FD_v of left eye	FD_h+ FD_v of right eye
C1	1.1715	1.1550	1.1635	1.1494	2.3265	2.3129
C2	1.1598	1.1428	1.1533	1.1494	2.3026	2.3027
C3	1.1757	1.1359	1.1693	1.1414	2.3116	2.3107
C4	1.1786	1.1432	1.1748	1.1393	2.3218	2.3141
C5	1.1808	1.1381	1.1759	1.1324	2.3189	2.3083

(b) 15° rotation

Table 1: The oriented fractal dimensions, FD_h and FD_v, and their corresponding sums, FD_h+FD_v, for a number of eye samples under different orientations: (a) upright, (b) 15° rotation.

3.3. Verification of eye pairs

The eye pairs selected in stage two are passed to the next stage for further verification. At this stage, the measurement of each possible eye-pair region and its corresponding face region will be computed by means of fractal dimension. In order to reduce the effect of lighting conditions, the edge images are used in the computation of the fractal dimensions.

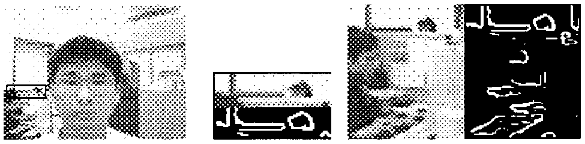
Based on the inter-distance L between two eye candidates, the eye-pair region and its corresponding face region can be extracted. The height, EH, and width, EW, of an eye-pair window are defined to be 0.5L and 1.67L, respectively. Furthermore, the corresponding width of the face region, WF, is also set to 1.67L. The height of the face region is set in reference to the position of the eye pairs: the top and bottom lines of a face window are 0.3L and 1.5L, respectively, from the eye pair.

In order to verify whether a selected eye-pair candidate is valid, the average fractal dimensions of the eye-pair regions, M'_{eye}, and the face regions, M'_{face}, are computed. In our study, we computed the fractal dimensions of a number of eye-pair windows and face windows. The corresponding means and variances of the fractal dimensions are 1.7298 and 0.0015, respectively, for the eye-pair region and 1.9469 and 0.0027, respectively, for the face region.

The computed fractal dimensions of the eye-pair regions and face regions show a small variance under different scales. A big difference is found between the fractal dimensions for facial images and those for non-facial images. Fig. 2(a) illustrates a possible eye-pair candidate which is actually the background of an image. Fig. 2(b) and 2(c) show an invalidly selected eye-pair window and its corresponding face window. The corresponding fractal dimensions for these two binarized windows are 1.564 and 1.679, respectively, which is a big difference from the corresponding average fractal dimensions. Thus, a valid eye pair can be selected if the following criteria are satisfied.

$$|F_{face}(x,y) - M'_{face}(x,y)| < t_5 \text{ and}$$
$$|F_{eye}(x,y) - M'_{eye}(x,y)| < t_6 \quad (4)$$

where (x,y) represents the position of the eye-pair window and the face windows, F_{eye} and F_{face} are the fractal dimensions of the eye-pair and face regions, M'_{eye} and M'_{face} are the average fractal dimensions of the two windows, and t_5 and t_6 are the thresholds. However, it is possible to detect more than one valid eye pair, which clusters around the valid eye-pair region. In making a selection among the overlapping valid eye pairs, the one with the lowest value of $|F_{eye}(x,y) - M'_{eye}(x,y)| + |F_{face}(x,y) - M'_{face}(x,y)|$ should be selected as the best eye pair in the overlapping region.



(a) (b) (c)

Figure 2: (a) A possible eye pair is selected in the background, (b) the invalid eye pair and its binarized image: FD of the binarized image is about 1.5638, and (c) the invalid face region and its binarized image: FD of the binarized image is about 1.6789.

3. EXPERIMENT RESULTS

The new approach for extracting the human eye pairs is tested using the MIT face database, ORL face database, and other complex images with a number of faces. Some of the results are shown in Fig. 4. Two eyes are located at the position of their corresponding irises in the face region, which is used to define the criterion for a hit of the eye pair. All the upright faces in the MIT face database and in other complex images with a number of faces under head-on lighting are detected without any error. Table 2(a) shows the hit rates and miss rates when the conventional fractal dimension is used under different conditions. Table 2(b) shows the hit rates and miss rates when our new approach is used under different conditions. Table 2(c) shows the results of locating the human eye and face [10] using the Gabor Wavelet. Our new approach achieves an overall hit rate of 96.9% without head tilt and under head-on lighting. When the heads are tilted to the left or to the right, the hit rate is 90.6%. When the lighting source to the faces is from 45°, the hit rates for the upright and rotated faces are both 84.4%. When the lighting is 90°, the hit rate for the upright face is 84.4% and the hit rate for the tilted face is 75%. Table 3 shows the hit rates and miss rates of our method based on the ORL database for upright faces and faces with perspective variations. Our approach achieves an overall hit rate of 92.9% for upright faces. When a face is subject to a slight perspective variation, the average hit rate is 82.4%. The processing time for locating the eye pairs in an image of size 128×120 or 92×112 is less than 1.5s. The experiment was performed on a Pentium II 450MHz PC. In conclusion, our new approach can greatly improve the hit rate under a lighting source of 90°. The approach also works well even if the eye pairs and faces are under uneven lighting conditions, or subjected to slight perspective variations.

	Lighting	Head On		45°		90°	
	Head tilt	No	Tilt	No	Tilt	No	Tilt
Full Scale	Hit	16	28	13	26	13	17
	Miss	0	4	3	6	3	15
Medium Scale	Hit	14	26	11	24	9	14
	Miss	2	6	5	8	7	18

Table 2(a): Experiment results based on the MIT face database using the conventional fractal dimension.

	Lighting	Head On		45°		90°	
	Head tilt	No	Tilt	No	Tilt	No	Tilt
Full Scale	Hit	16	28	14	28	14	24
	Miss	0	4	2	4	2	8
Medium Scale	Hit	15	30	13	26	13	24
	Miss	1	2	3	6	3	8

Table 2(b): Experiment results based on the MIT face database using our new approach.

	Lighting	Head On		45°		90°	
	Head tilt	No	Tilt	No	Tilt	No	Tilt
Full Scale	Hit	16	28	14	24	14	19
	Miss	0	4	2	8	2	23
Medium Scale	Hit	14	26	12	26	10	15
	Miss	2	6	4	6	6	17

Table 2(c): Experiment results based on the MIT face database using the Gabor Wavelet.

Perspective Variation	No	Slight
Hit	70	51
Miss	5	9

Table 3: Experiment results based on the ORL face database using our new approach.

4. CONCLUSION

In this paper, we have proposed an oriented fractal dimension to accurately extract eye pairs in a complex image. Instead of searching the whole image space to determine the different scales of the eye pairs, only possible eye pairs are investigated. These possible eye pairs are identified by means of valley detection and by measuring their fractal dimensions. In order to reduce the effect of uneven lighting conditions on the left-eye and the right-eye, eye windows are normalized to a specific gray-level intensity. In addition, when grouping a left eye candidate with a right eye candidate to form an eye pair, it is proposed that the oriented fractal dimension be used because it can provide a higher level of matching accuracy. The corresponding average fractal dimensions of the binarized eye-pair regions and the face regions are then used to verify whether the eye pairs selected are valid. The proposed method can successfully extract valid eye pairs under different scales, orientations, lighting conditions, and slight perspective variations.

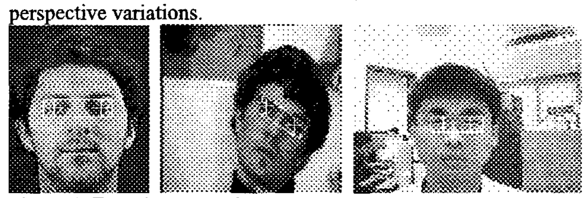


Figure 4: Experiment results.

5. REFERENCES

[1] K. Sobottka and I. Pitas, "A Novel Method for Automatic Face Segmentation Facial Feature Extraction and Tracking", *Signal Processing: Image Comm., 12,* pp. 263-280, 1998.

[2] K.M. Lam and H. Yan, "Locating and Extracting the Eye in Human Face Images", *Pattern Recognition,* Vol. 29, No 5, pp. 771-779, 1996.

[3] G. Yang and T.S. Huang, "Human Face Detection in a complex background," *Pattern Recognition,* Vol. 27, No 1, pp.53-63, 1994.

[4] K.K Sung and T. Poggio, "Example-Based Learning for View-Based Human Face Detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. 20, No 1, pp.39-51, 1998.

[5] P.J. Saupe and M. P. Yunker, "Fractals for the classrooms", *Strategic Activities Vol. One,* Springer-Verlag.

[6] N. Sarkar and B.B. Chaudhuri, "An Efficient Differential Box-Counting approach to Compute Fractal Dimension of image", *IEEE Trans. on Systems, Man, and Cybernetics,* Vol. 24, No 1, pp.115-120, 1994.

[7] J. Feng, W. C. Lin and C.T. Chen, "Fractional Box-Counting Approach to Fractal Dimension Estimation", *Proc. of the 13th Int. Conf. on Pattern Recognition,* Vol.2, pp. 854-858, 1996.

[8] K.L. Chan. "Fractal Based Texture Analysis", *Singapore ICCS/ISITA '92. Comm. on the Move,* Vol.1, pp. 102-106, 1990.

[9] Y. Y. Tang and Y. Tao, "Feature Extraction by Fractal Dimension", *ICDAR '99. Proc. of the 5th Int. Conf. on Document Analysis and Recognition,* pp. 217-220, 1999.

[10] K. M. Lam, "A fast approach for detecting human faces in a complex background ", *ISCAS '98, Proc. of the IEEE Int. Sym. on Circuits & Systems,* Vol 4, pp. 85–88, 1998.