FACE IMAGE MATCHING USING FRACTAL DIMENSION

A.Z. Kouzani, F. He, and K. Sammut

School of Engineering, Flinders University of SA, GPO Box 2100, Adelaide, SA 5001, Australia
egazk, egfh, egks@flinders.edu.au

ABSTRACT

A new method is presented in this paper for calculating the correspondence between two face images on a pixel by pixel basis. The concept of fractal dimension is used to develop the proposed non-parametric area-based image matching method which achieves a higher proportion of matched pixels for face images than some well-known methods.

1. INTRODUCTION

A basic requirement for any automated face recognition system is that the collection of example faces images must be set into correspondence with one another. Although, the correspondence problem for most face recognition systems is defined specifically as the matching of a subset of predefined points (i.e., facial landmarks), in this paper the face correspondence problem is cast into the more general case where the aim is to match all of the data points.

In image matching no guarantee can be given that a solution exists, is unique, and is stable with respect to small variation in the input data. For a given point in one image, a corresponding point may not exist in the other image due to occlusion, there may be more than one possible match due to repetitive patterns or a semi-transparent object surface, and the solution may be unstable with respect to noise due to poor texture. These issues can be dealt with by introducing additional knowledge about the problem. A range of assumptions usually hold true when dealing with face image matching: (i) the face surface is piecewise smooth; (ii) the gray level values of various face images are acquired using a similar spectral band; (iii) the two face images share approximately similar perspectives with small variations in scale, translation, and orientation. Making these assumptions can result in a reasonable establishment of correspondence between the input and the known face image that significantly enhances the performance of an automatic face recognition system.

There are several approaches used for image matching [1, 2, 3]. It has been demonstrated in the literature [4, 5, 6] that some area-based methods produce higher numbers of matched pixels than other methods where only two images are considered. However, these area-based methods are not perfect and they still produce some mismatches. In this paper, a new area-based method is proposed that calculates the correspondence between two face images for every pixel locations. The proposed method uses the concept of *Fractal Dimension* (FD) and develops a non-parametric transform as a basis for establishing the correspondence between two face images.

This paper is organised as follows. In Section 2, the existing image matching methods are reviewed. In Section 3, the FD is explained. The FD calculation method is described in Section 4. In Section 5, the proposed image matching method is given. In Section 6, experimental results are given. These results are then discussed in Section 7. Finally, concluding remarks are given in Section 8.

2. REVIEW OF EXISTING METHODS

A variety of methods are used for the calculation of the 2D displacement field between pixels across two images, and other new methods continue to appear. A good overview of these methods can be found in [1, 2, 3]. Some of the more important existing methods are discussed in the following.

Feature-Based Methods: In feature-based methods, good matchable features (such as corner points) are sparse, while poor to match features (such as edges) are denser. Even when reasonably unique features are available, establishing the correct correspondence can be problematic. To complicate matters further, occlusion of features may lead to errors in matching. Area-based matching methods are less sensitive to these problems [7].

Intensity-Based Differential Methods: In these methods, accurate numerical differentiation may be impractical. In these cases differential approaches may be inappropriate and it is natural to turn to area-based matching because of noise, because of the existence of a small number of frames, or because of aliasing in the image acquisition process. Several methods for automating the correspondence finding procedure in human face images are reported in the literature. Most of these methods [8, 9] use a gradient-based optical flow algorithm [10, 11] which is classified as belonging to the intensity-based differential group. The application of the gradient-based optical flow algorithm to the face image correspondence finding procedure does not perform well, however, due to the following reasons: (i) only two frames exist; (ii) the two images contain pictures of two different persons; (iii) the face images are not rigid, but rather deformable; (iv) the illumination might slightly change between two face images.

Area-Based Methods: These methods are less sensitive to the above problems and perform better under such circumstances [2, 5, 6]. With the area-based correspondence algorithm, a small window of pixels centred around the pixel to be matched, is compared with equally-sized regions in the other image. Different matching metrics are used to provide a numerical measure of the similarity between a window in one image and a window in another image, and hence are used to determine the optimum match. There are a number of classical matching metrics available, such as the*Normalised Cross Correlation* (NCC) [4].

In addition, a number of non-parametric methods have been used in area-based matching. These methods are based on the relative ordering of pixel intensities within a window, rather than the intensity values themselves. Consequently, these methods are robust with respect to distortion, since differences in gain and bias between two images will not affect the ordering of pixels within a window. In addition, these transforms can tolerate a small amount of random noise. Two non-parametric transforms which have demonstrated robustness in image matching are: the *rank transform* and the *census transform*. The rank transform is defined as the number of pixels in the window whose value is less than the centre pixel [5]. The census transform maps the window surrounding the centre pixel to a bit string [6].

It has been demonstrated that non-parametric methods produce higher numbers of matched pixels than other area-based matching methods [4, 5, 6]. However, they are not perfect and they still produce some mismatches. In the following subsection, the concept of the FD is presented and the method for calculating the FD of an image is explained. The concept of the FD is used in a new non-parametric image matching method which achieves a higher proportion of matched pixels for face images than the rank and census methods.

3. FRACTAL DIMENSION

Fractal dimension is an important measure of roughness and self-similarity in pictures [12]. It separates important classes of images and characterises information which is not characterised by other texture analysis methods. The FD has been used in several image processing and pattern recognition applications. Pentland [13] developed a 3D fractal model which has been used in texture segmentation and classification, estimation of 3D shape information, and classification of perpetually smooth and perpetually textured surfaces. Peleg et al. [14] derived a set of 48 features using the FD for texture analysis. Rigaut [15] and Keller et al. [16] employed the FD for image segmentation.

In the realm of classical geometry, dimensions of objects are defined by whole numbers. Fractals, however, have dimensions which fall between whole numbers. This is because fractals will have a finite area or volume but an infinite perimeter or surface size. There exist several approaches for estimating the FD of an image. Peleg et al. [14] views an image as a hilly terrain surface whose height from the normal ground is proportional to the image gray value. Pentland [13] suggests a method of estimating the FD by using the Fourier power spectrum of image intensity surface. Keller et al. [16] propose a modification of a method due to Voss [17]. They introduced new features based on the concept of lacunarity which capture the second-order statistics of fractal surfaces. They used these measurements together with an improved estimate of fractal dimension as both global signatures of texture and as local measurements of texture. However, their method compresses the results of the estimated dimensions toward the middle of the true range, a phenomenon that becomes more apparent as the FD moves towards 3.0. The main reason for this is the quantisation effect. As the dimension approaches 3.0, the surface becomes highly irregular and after quantisation the gray level values are widely spaced. This problem can be relieved by interpolating the neighbouring gray values, however, the interpolation would be linear and hence not accurate. Sarkar and Chauduri [12] propose a method for calculation of the FD of an image that does not suffer from this problem since their counting method gives a better approximation to the boxes intersecting the image intensity surface. Conversely, the box counting method developed by Keller et al. does not cover the image surface so well and hence struggles to accurately capture the FD for a textured surface. This technique also requires far more computations which makes it considerably slower.

The method which is used in this paper for the calculation of the FD of an image is a variation of the Sarkar-Chauduri's box-counting procedure [12] which is efficient and sufficiently accurate for this application.

4. FRACTAL DIMENSION CALCULATION

The FD of a 2D image can be calculated using the following algorithm.

Algorithm 1 *(Fractal Dimension of a 2D Image) Using the basic FD equation*

$$D = \frac{log\ (N_r)}{log\ (1/r)}, \tag{1}$$

the FD of an image is calculated as follows. Consider that the image of size $M \times M$ is scaled down to a size

$s \times s$ where $M/2 \geq s > 1$ and s is an integer. Then, an estimate of r is defined as $r = s/M$. Also, consider the 2D image as a 3D space with (x, y) denoting the 2D position and the coordinate z denoting gray level. The (x, y) space is partitioned into grids of size $s \times s$. On each grid there is a column of boxes of size $s \times s \times s'$. If the total number of gray levels is G then $\lfloor G/s' \rfloor = \lfloor M/s \rfloor$. Let the minimum and maximum gray level of the image in (i, j)th grid fall in box number k and l, respectively. Then $n_r(i, j) = l - k + 1$ is the contribution of N_r in (i,j)th grid. Taking contributions from all grids, N_r can be obtained from

$$N_r = \sum_{i,j} n_r(i, j), \tag{2}$$

in which N_r is counted for different values of r (i.e. different values of s). Then using Equation 1, the fractal dimension D can be estimated from the least square fit of $\log(N_r)$ against $\log(1/r)$. A typical plot of $\log(N_r)$ vs $\log(1/r)$ for a 312×312 face image of 256 gray levels is shown in Figure 1. Let $y = mx + c$ be the fitted straight line, where y denotes $\log(N_r)$ and x denotes $\log(1/r)$. Then error of fit E can be expressed as

$$E = \frac{\sqrt{\sum_{i=1}^{n} \frac{(mx_i + c - y_i)^2}{(1+m^2)}}}{n}, \tag{3}$$

where n is the number of data points in the plot. This error provides a measure of fit so that the lower the value of E, the better is the fit.

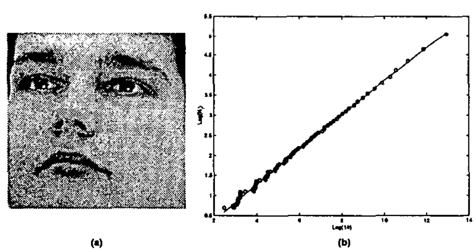


Figure 1: Plot of $\log(N_r)$ vs $\log(1/r)$ of a face image. (a) The face image. (b) The plot.

The procedure described above is the method that calculates the FD of an entire face image. However, in the proposed image matching method the correspondence between two images is calculated by computing the FD of local areas in an image. Each image pixel is represented by the FD of a particular region of the image that surrounds the pixel. The calculation of the FD can be done by taking the image region as a full image and applying the above algorithm to it. But this approach has two shortcomings: (i) it is possible that two local image regions with different textures and optical differences have the same FD; (ii) the regions surrounding the pixels located near the image borders become smaller than the regions representing other pixels. Therefore, different number of pixels in different regions participate in the calculation of the FD.

The size of the region is also important. The region size determines the amount of pixels of the image that will be used for extraction of the FD. When the region size is very small (e.g. 3×3) the true FD cannot be obtained. If the region size is very large, similar FD might be obtained for neighbouring overlapping regions. Choosing an appropriate window size is an important issue for which a solution is offered in the following algorithm.

5. PROPOSED METHOD

The proposed image matching method makes use of the concept of the FD, and develops a non-parametric local transform as a basis for establishing correspondence between two face images. The overall algorithm proposed for matching two face images is shown in the block diagram of Figure 2. The source and target face images are directed to a fractal dimension transform stage. The transformed images are presented to a normalised cross-correlation stage in which the best matches are found and the corrected image is produced based on the matching results.

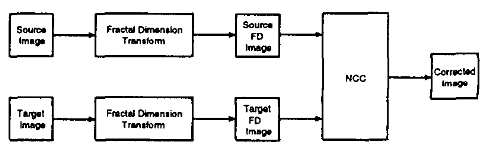


Figure 2: Block diagram of the proposed image matching method.

To handle the shortcomings of the FD calculation method described in the previous section, the FD of image regions of different sizes around a pixel are calculated and averaged. This solves the problem of having similar FDs for local image regions with different textures and of considering the effect of regions of different sizes. This will also remove the problem of selecting an effective region size. To handle the problem which the pixels near the image borders face, the image extension process is proposed. This is done by circularly shifting the image. The proposed algorithm is explained in detail in the following.

Algorithm 2 *(Fractal-Dimension Image Matching) The pixel-based correspondence between a source image I_S and a target image I_T is calculated as follows:*

1. *The $n \times n$ face images I_S and I_T are extended by circularly shifting their contents towards different directions (see Figure 3). This will create two new $m \times m$ images $\hat{I}_S$ and $\hat{I}_T$ whose sizes are*

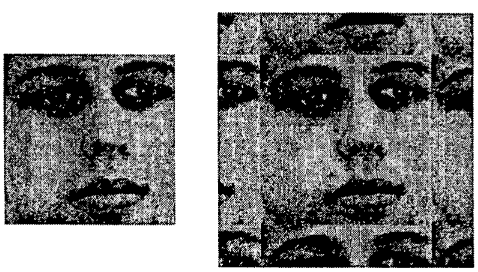
$$m = \begin{cases} n + \lceil n/2 \rceil - 1 & \textit{if } \lceil n/2 \rceil \textit{ is odd} \\ n + \lceil n/2 \rceil & \textit{if } \lceil n/2 \rceil \textit{ is even.} \end{cases} \quad (4)$$

2. *For each pixel P_{ij} in I_S and I_T, the following operations are separately performed:*
   - (a) *The window size W is initialised to 5.*
   - (b) *The associated image region of size W, with pixel P_{ij} being its centres, is extracted from $\hat{I}_S$ or $\hat{I}_T$.*
   - (c) *The fractal dimension D_W^{ij} of the extracted image is calculated.*
   - (d) *The window size W is increased by 2.*
   - (e) *If W is less than or equal to $n/2$ then a jump to Step 2b is executed.*
   - (f) *The fractal dimension associated with P_{ij} is calculated from*

$$D^{ij} = \frac{1}{N} \sum_{W} D_W^{ij}, \quad (5)$$

   where N is the number of windows.

3. *P_{ij}s is replaced with D^{ij} (obtained from I_S) in I_S and with D^{ij} (obtained from I_T) in I_T. This creates two new matrices $\bar{I}_S$ and $\bar{I}_T$.*
4. *The NCC is performed on $\bar{I}_S$ and $\bar{I}_T$.*
5. *The correspondence field is stored.*



(a) (b)

Figure 3: Generating a new 384 × 384 face image from a 256 × 256 face image by circularly shifting it.

A validation technique adopted from [18] is used to identify incorrect matches. In this technique, the roles of the two images are reversed and the matching is performed a second time. This validity test detects and removes invalid matches.

6. EXPERIMENTAL RESULTS

The proposed method together with a number of other matching metrics have been implemented and their behaviour have been explored on a collection of face images.

The methods which have been implemented and used for comparison against the proposed image matching method, are the *Bergen-Hingorani gradient-based optical flow* (BH) [11], the NCC, the rank, and the census. For the NCC and rank metrics, a correlation window size of 9 × 9 is used. For the census metric, a correlation window size of 9 × 9 and a transform window size of 7 × 7 are used. Other parameters, such as the number of disparities considered, are held constant for all methods.

(a) (b)

Figure 4: A test pair in which the source image (a) is 30% brighter than the target image (b).

Ten pairs of face images of resolution 314 × 229 have been used in the experiments. Each pair of face images contain two face images of different people with approximately similar poses. Each face image may have a maximum pose variation of ±5°. The ten pairs of face images are separated into two groups. Five pair of face images, taken under similar lighting conditions, are put in the first group. The remaining five pairs, in which the source face image is 30% brighter than the target image (see Figure 4 for an example), are put in the second group. Both pose and lighting conditions vary across the ten different pairs. For each group of the test images, the portion of matches remaining after each validity check are added and then divided by the number of pairs in that group, and are displayed in Table 1.

Table 1: Portion of matched pixels for each correspondence method, for each group of test images.

	BH	NCC	rank	census	Proposed
Group 1	0.59	0.68	0.66	0.81	0.91
Group 2	0.32	0.42	0.54	0.77	0.86

7. DISCUSSIONS

It can be seen from Table 1 that the BH method achieves the poorest matching rate amongst the tested methods. This is due to the problems highlighted in Section 2. Amongst the NCC, the rank, and the census methods, the later achieves the highest proportion of matched pixels for the two test groups. However, the use of the proposed image matching method allows higher matching rates than the census for both test groups.

The results prove that the proposed image matching method is an efficient method for computing correspondence between images and that it also accounts for variations between the source and target images.

8. CONCLUSIONS

In this paper, a new face image matching method is presented. The proposed method calculates the correspondence between two face images for every pixel location. The concept of fractal dimension is used to develop a non-parametric local transform as a basis for establishing correspondence between two face images.

The proposed image matching method achieves a higher proportion of matched pixels between two face images than the existing counterparts. This improvement is due to the establishment of correspondence based on a multiresolution method which explores the partial self-similarity of the images as provided by the fractal dimension measurement.

9. ACKNOWLEDGEMENTS

The first author would like to thank DEETYA, Australia, and FUSA, for providing him with scholarships to undertake this research.

10. REFERENCES

[1] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *Int. J. of Computer Vision*, 12(1):43–77, 1994.

[2] S.S. Beauchemin and J.L. Barron. The computation of optical flow. *ACM Computing Surveys*, 27(3):33–467, 1996.

[3] C. Heipke. Overview of image matching techniques. In *Proc. OEEPE Workshop on the Application of Digital Photogrammetric Workstations*, Lausanne, March 1996. dgrwww.epfl.ch.

[4] J. Banks, M. Bennamoun, and P. Croke. Fast and robust streo matching algorithm for mining automation. In *Proc. International Workshop on Image Analysis and Information Fusion*, pages 139–149, Adelaide, Australia, November 1997.

[5] D. Bhat and S. Nayar. Ordinal measures for visual correspondence. In *Proc. Computer Vision and Pattern Recognition*, pages 351–357, San-Fransisco, 1996.

[6] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proc. 3rd European Conference on Computer Vision*, Stockholm, 1994.

[7] J.J. Little. Accurate early detection of discontinuties. *Vision Interface*, pages 2–7, 1992.

[8] D. Beymer and T. Poggio. Face recognition from one example view. Technical Report 1536, MIT AI Lab., 1995.

[9] T. Vetter and N.F. Troje. A seperated linear shape and texture space for modeling two-dimensional images of human faces. Technical Report 15, Max Planck Inst., April 1995.

[10] E.H. Adelson and J.R. Bergen. the extraction of spatiotemporal energy in human and machine vision. In *Proc. IEEE Workshop on Visual Motion*, pages 151–156, Carlston, 1986.

[11] J.R. Bergen and R. Hingorani. Hierarchical motion-based frame rate conversion. Technical report, David Sarnoff Research Center, Princeton NJ, 1990.

[12] N. Sarkar and B.B. Chaudhuri. An efficient approach to estimate fractal dimension of textural images. *Pattern Recognition*, 25(9):1035–1041, 1992.

[13] A. Pentland. Shading into texture. *Art. Intell.*, 29:147–170, 1986.

[14] S. Peleg, N.R. Hartley, and D. Anvir. Multiple resolution texture analysis and classification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-6(4):518–523, 1984.

[15] J.P. Rigaut. Automated image segmentation by mathematical morphology and fractal geometry. *J. Microscopy*, 150(1):21–30, 1988.

[16] J.M. Keller and S. Chen. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing*, 45:150–166, 1989.

[17] R. Voss. Random fractals: Characterisation and measurment. In R. Pynn and A. Skjeltorp, editors, *Scaling Phenomena in Disordered Systems*. Plenum, New York, 1986.

[18] P. Fua. A parallel stereo algorithm that produces dense depth maps and preseves image features. *Machine Vision and Applications*, 6:35–49, 1993.