Advances in Intelligent Systems and Computing 1375

Ajith Abraham · Thomas Hanne ·
Oscar Castillo · Niketa Gandhi ·
Tatiane Nogueira Rios ·
Tzung-Pei Hong *Editors*

Hybrid Intelligent Systems

20th International Conference on Hybrid Intelligent Systems (HIS 2020), December 14–16, 2020

Springer

Advances in Intelligent Systems and Computing

Volume 1375

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Indexed by DBLP, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST).

All books published in the series are submitted for consideration in Web of Science.

More information about this series at http://www.springer.com/series/11156

Ajith Abraham · Thomas Hanne ·
Oscar Castillo · Niketa Gandhi ·
Tatiane Nogueira Rios · Tzung-Pei Hong
Editors

Hybrid Intelligent Systems

20th International Conference on Hybrid Intelligent Systems (HIS 2020), December 14–16, 2020

Springer

Editors
Ajith Abraham
Scientific Network for Innovation
and Research Excellence
Machine Intelligence Research Labs
(MIR Labs)
Auburn, WA, USA

Oscar Castillo
Division of Graduate Studies
Tijuana Institute of Technology
Tijuana, Mexico

Tatiane Nogueira Rios
Universidade Federal da Bahia
Salvador, Bahia, Brazil

Thomas Hanne
Institute for Information Systems
University of Applied Sciences and Arts
Northwestern Switzerland
Olten, Solothurn, Switzerland

Niketa Gandhi
Scientific Network for Innovation
and Research Excellence
Machine Intelligence Research Labs
(MIR Labs)
Auburn, AL, USA

Tzung-Pei Hong
Computer Science and Information
Engineering
National University of Kaohsiung
Kaohsiung, Taiwan

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-030-73049-9 ISBN 978-3-030-73050-5 (eBook)
https://doi.org/10.1007/978-3-030-73050-5

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Welcome to the 20th International Conference on Hybrid Intelligent Systems (HIS 2020) and the 12th World Congress on Nature and Biologically Inspired Computing (NaBIC 2020) held in the World Wide Web during December 14–16, 2020. In 2020, HIS and NaBIC were held at Bhopal and Hyderabad, respectively.

Hybridization of intelligent systems is a promising research field of modern artificial/computational intelligence concerned with the development of the next generation of intelligent systems. A fundamental stimulus to the investigations of hybrid intelligent systems (HISs) is the awareness in the academic communities that combined approaches will be necessary if the remaining tough problems in computational intelligence are to be solved. Recently, hybrid intelligent systems are getting popular due to their capabilities in handling several real-world complexities involving imprecision, uncertainty, and vagueness. HIS 2020 received submissions from 22 countries, and each paper was reviewed by at least five reviewers in a standard peer-review process. Based on the recommendation by five independent referees, finally fifty-eight papers were presented during the conference (acceptance rate of 37%).

Nature and Biologically Inspired Computing brings together international researchers, developers, practitioners, and users. NaBIC invited authors to submit their original and unpublished work that demonstrates current research in all areas of Nature and Biologically Inspired Computing, as well as industrial presentations, demonstrations, and tutorials. NaBIC 2020 received submissions from twelve countries, and each paper was reviewed by at least five reviewers in a standard peer-review process. Based on the recommendation by five independent referees, finally twenty papers will be presented during the conference (acceptance rate of 30%).

Many people have collaborated and worked hard to produce this year successful HIS–NaBIC conferences. First and foremost, we would like to thank all the authors for submitting their papers to the conference, for their presentations and discussions during the conference. Our thanks to the program committee members and reviewers, who carried out the most difficult work by carefully evaluating the submitted papers. Our special thanks to the following plenary speakers, for their exciting plenary talks:

- Prof. Dr. Sushmita Mitra, Indian Statistical Institute, Kolkata, India
- Prof. Dr. Bassem Jarboui, Sfax University, Tunisia
- Prof. Dr. Tzung-Pei Hong, National University of Kaohsiung, Taiwan
- Prof. Dr. André Rossi, Université Paris-Dauphine, France
- Prof. Dr. Amine Nait-Ali, University Paris-Est Créteil, France
- Dr. Rudy A. Oude Vrielink, University of Twente & TwiQel, Netherlands
- Prof. Dr. Frederico Gadelha Guimaraes, Federal U. of Minas Gerais, Brazil
- Prof. Dr. Fakhri Karray, University of Waterloo, Canada
- Prof. Dr. Saman K. Halgamuge, University of Melbourne, Australia
- Prof. Dr. Salvador García, University of Granada, Granada, Spain
- Prof. Dr. Vincenzo Piuri, Università degli Studi di Milano, Italy
- Prof. Dr. Dries F. Benoit, Ghent University, Belgium
- Prof. Dr. Dijiang Huang, Arizona State University, USA
- Prof. Dr. Ketan Kotecha, Symbiosis Institute of Technology, India
- Prof. Dr. Sheela Ramanna, University of Winnipeg, Canada

Our special thanks to the Springer Publication team for the wonderful support for the publication of this proceedings. We express our sincere thanks to the session chairs and organizing committee chairs for helping us to formulate a rich technical program. We express our sincere thanks to the organizing committee chairs for helping us to formulate a rich technical program. Enjoy reading the articles!

Ajith Abraham
Thomas Hanne
Oscar Castillo
Tatiane Nogueira Rios
Tzung-Pei Hong
Niketa Gandhi
Kun Ma

HIS–NaBIC 2020 Organization

General Chairs

Ajith Abraham	Machine Intelligence Research Labs (MIR Labs), USA
Thomas Hanne	University of Applied Sciences and Arts Northwestern Switzerland, Switzerland
Oscar Castillo	Tijuana Institute Technology, Mexico

Program Chairs

Tatiane Nogueira Rios	Universidade Federal da Bahia, Brazil
Tzung-Pei Hong	National University of Kaohsiung, Taiwan

Publication Chairs

Niketa Gandhi	Machine Intelligence Research Labs, USA
Kun Ma	University of Jinan, China

Publicity Chairs

Jyotshna Dongardive	University of Mumbai, India
Pooja Manghirmalani Mishra	University of Mumbai, India

International Program Committee

Adrien Brilhault	Université de Toulouse, France
Bruno Silva	Federal University of Ceará, Brazil
C. Alberto Ochoa-Zezatti	Universidad Autónoma de Ciudad Juárez, Mexico
Carlos Pereira	ISEC, Portugal

Efrén Mezura-Montes	University of Veracruz, Mexico
Elizabeth Goldbarg	Federal University of Rio Grande do Norte, Brazil
Frantisek Zboril	Brno University of Technology, Czech Republic
Gloria Cerasela Crisan	Vasile Alecsandri University of Bacau, Romania
Isabel S. Jesus	Instituto Superior de Engenharia do Porto, Portugal
János Botzheim	Budapest University of Technology and Economics, Hungary
Jolanta Mizera-Pietraszko	Opole University, Poland
José Everardo Bessa Maia	State University of Ceará, Brazil
José Raúl Romero	University of Córdoba, Spain
José Yaír Guzmán-Gaspar	Universidad Veracruzana, Mexico
Jyotshna Dongardive	University of Mumbai, India
Kaushik Das Sharma	University of Calcutta, India
Lee Chang-Yong	Kongju National University, South Korea
Mansi Sharma	Indian Institute of Technology Delhi, India
Mario Giovanni	University of Pisa, Italy
Matheus Menezes	Universidade Federal Rural do Semi-Árido, Brazil
Niketa Gandhi	Machine Intelligence Research Labs (MIR Labs), USA
Octavio Ramos	Universidad Veracruzana, Mexico
Oscar Castillo	Tijuana Institute Technology, Mexico
Patrick Siarry	Université Paris-Est Créteil, France
Paulo Moura Oliveira	UTAD University, Portugal
Petala Estrela	UFBA, Salvador
Pooja Manghirmalani Mishra	University of Mumbai, India
Radu-Emil Precup	Politehnica University of Timisoara, Romania
Rafael Barbudo Lunar	University of Córdoba, Spain
Ricardo Rios	Universidade Federal da Bahia, Brazil
Romerito Campos	Federal University of Rio Grande do Norte, Brazil
Sidemar Fideles Cezario	Federal University of Rio Grande do Norte, Brazil
Sílvia M. D. M. Maia	Universidade Federal do Rio Grande do Norte, Brazil
Simone Ludwig	North Dakota State University, USA
Subodh Deolekar	REDx Innovation Lab WeSchool, India
Tatiane Nogueira	Universidade Federal da Bahia, Brazil
Thiago Soares Marques	Federal University of Juiz de Fora, Brazil
Thomas Hanne	University of Applied Sciences Northwestern Switzerland, Switzerland

Tzung-Pei Hong	National University of Kaohsiung, Taiwan
Varun Ojha	University of Reading, UK
Xochitl Siordia	Universidad Veracruzana, Mexico

Contents

Vertical Interior Distance Ratio to Minimum Bounding Rectangle of a Shape

Zekun Li, Baolong Guo(✉), Xia Ren, and N. N. Liao

Xidian University, Xi'an 710071, People's Republic of China
blguo@xidian.edu.cn

Abstract. This paper proposed a simple shape descriptor, based on the vertical interior distance ratio of the minimum bounding rectangle (VIDR). Shape descriptor is widely used to describe shape, especially in leaves matching and trademark retrieval. VIDR is the proportional distribution of the vertical interior distance between the shape contour points and its four sides of the minimum bounding rectangle. The minimum bounding rectangle can change according to the change of shape scale and direction, which is extremely suitable for describing the changing shape, and it has global characteristics. Compared with the descriptor of centroid contour distance (CCD) and the shape context descriptor (SC), which are all use contour points to describe shape, our descriptor has a vertical direction and is able to distinguish the same centroid shapes or similar contour shapes, it also has simpler feature extracting description process. More importantly, the experimental results show that our descriptor has a higher precision and faster speed.

Keywords: Minimum bounding rectangle · Vertical · Interior distance · Contour

1 Introduction

Many shape descriptors have been proposed [1–5]. It can be roughly divided into four categories: global descriptors, local descriptors, multi-scale descriptors and multi-faceted descriptors. Shape descriptors are used to extract shape features of objects. FD (Fourier descriptor) [6], CCD (centroid contour distance), FD-CCD (CCD based FD) [6], FPD (farthest point distance) [7], WD (wavelet descriptor) [8], SC (shape context) are some classic descriptors. There are some improved descriptions of FMSCCD (Fourier descriptor based on multi-scale centroid distance) [9], IDSC (shape context based on inner distance) and ASD&CCD (angular scale descriptor based on distance from centroid) [10].

For example, ASD is a descriptor that contains directional characteristics, which can calculate the eigenvectors of Angle sequences of different proportions, including local and global information. The ASD&CCD adds Angle information according to the CCD. The ASD&CCD is more accurate than the CCD, but a lot of calculations are inevitable and run slowly. In [10], ASD&CCD has the highest bull's eye test score, but the matching time of mpeg-7 CE1 part B is longer than that of CCD. IDSC is improved on the basis of

A. Abraham et al. (Eds.): HIS 2020, AISC 1375, pp. 1–10, 2021.
https://doi.org/10.1007/978-3-030-73050-5_1

SC. Although the experimental results are better than SC, the matching stage still needs a lot of time, and the shape feature is not directional, and there is a direction defect.

The ability to evaluate a shape descriptor is mainly through translation, rotation and scaling of the shape to see whether it will be unaffected and still be able to correctly identify the shape. The minimum bounding rectangle of an image is one-to-one related to the image. If the image has different positions, angles and sizes, the minimum bounding rectangle will change accordingly. Therefore, the position, Angle and size of the image can be reflected in the parameters of its minimum bounding rectangle. Using the minimum bounding rectangle to describe the image has strong descriptive ability and robustness.

Shape histogram has a relatively intuitive representation of quantity, and it has a strong ability to describe the distribution of quantity, so it is used by most scholars to describe the features of shape. For example, in the classical shape context method, shape histogram is used to describe the region where the contour point belongs to relative to the current point. Since shape histogram is the regional distribution of all quantities, it has a natural advantage for describing shape features, that is, in the process of describing shape, the uncertainty of initial description position can be ignored, and the computational cost of describing process can be greatly reduced, with strong robustness.

The general principle of simple shape descriptor is simple, simple calculation, but there is no direction, can only on behalf of the general characteristics of shape, and the lack of detailed description such as directional characteristics, such as CCD and SC, only for contour point, is to distinguish the similar centroid contour shape context of distance and similar problems, as shown in Fig. 1. The method proposed in this paper uses a simple feature - the vertical interior distance ratio of the minimum bounding rectangle to give shape descriptors a vertical orientation feature. Therefore, this descriptor complements the directivity defect of the simple descriptor and has higher resolution capability without increasing the time complexity of the general descriptor algorithm.

2 Related Work

2.1 CCD (Centroid Contour Distance)

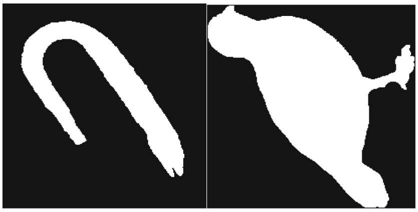
Centroid contour distance (CCD) [11] is a simple and effective shape descriptor. The descriptor consists of the shapes of each contour point of the shape to the center of mass. The sequence $P = \{p_1, p_2, p_3, \ldots, p_n\}$ is the contour sample points after uniform sampling of the shape contour.

The centroid contour point p_c of these contour points is first calculated, using Equation:

$$p_c = \frac{1}{n}\sum_{i=1}^{n} p_i = \left(\frac{1}{n}\sum_{i=1}^{n} x_i, \frac{1}{n}\sum_{i=1}^{n} y_i\right) \tag{1}$$

where n means how many sampling points there are on the contour, (x_i, y_i) is the *ith* contour point p_i of a shape. Then, the Euclidean distance between each contour point $p_i(x_i, y_i)$ and the centroid contour point $p_c(x_c, y_c)$ is calculated, using Equation:

$$d_{ic} = ((x_i - x_c)^2 + (y_i - y_c)^2)^{\frac{1}{2}} \tag{2}$$



(1) similar centroid contour distance

(2) similar shape context

Fig. 1. Two shape pairs, upper pair having similar centroid contour distance and the following pair having similar shape context.

Finally, The Euclidean distance d_{ic} is normalized to obtain the centroid contour distance d^i_{ccd}, calculated with Equation:

$$d^i_{ccd} = \left(\frac{d_{ic}}{(1/n)\sum_{j=1}^{n} d_{jc}} \right) / \sqrt{n} = d_{ic}\sqrt{n} / \sum_{j=1}^{n} d_{jc}, i = 1, 2, 3, \ldots, n \quad (3)$$

The sequence $\{d^1_{ccd}, d^2_{ccd}, d^3_{ccd}, \ldots, d^n_{ccd}\}$ is the CCD feature of a shape.

When the starting point position of the contour changes, the distance sequence will also change accordingly. Therefore, the distance between two shapes, $s1$, $s2$, is computed with Equation:

$$dis_{ccd}(s1, s2) = \min_{0 \le m \le n} (\sum_{i=1}^{n} (d^i_{ccd} - dd^{m+i}_{ccd})^2)^{\frac{1}{2}}, m \in Z \quad (4)$$

where, $d^{m+i}_{ccd} = d^{m+i-n}_{ccd}$. This distance is used for shape matching. When the distance is less than a given threshold, it is regarded as in the same class, when the distance is greater than this threshold, it is regarded as in the different class.

The CCD descriptor has translational invariance because its European distance is normalized. Under the above matching method, it has rotation invariance, but the efficiency of matching method is low, the speed of matching is slow, and the time complexity is high. Although the idea of this descriptor is simple, it is a scalar distance without direction and lacks the description of shape direction information.

2.2 SC (Shape Context)

The shape context (SC) [12] is a classic descriptor based on shape contour features introduced by Serge Belongie et al. in 2002, it is a global shape descriptor. The shape contour is generally represented by a uniformly sampled set of points, the shape context of each point in the point set is a histogram of the point distribution, the target shape context is represented by a set of histograms.

The contour of a shape is represented by a sequence of continuous point sets, after sampling uniformly, a contour sample point is obtained $P = \{p_1, p_2, p_3, \ldots, p_n\}$. Take any one of these points p_i as the reference point, and take p_i as the center of the circle, N concentric circles are constructed with R as radius and logarithmic distance intervals. Then, the concentric circle region is divided into M equal parts along the circumference to form the template as shown in the figure. The number of points in each small area of the template is counted to constitute the histogram $h_i(k)$, $h_i(k)$ is the shape context of p_i, calculated with Equation:

$$h_i(k) = \#\{p_j : j \neq i, p_i - p_j \in bin(k)\} \tag{5}$$

Where, $k = \{1, 2, 3, \ldots, M \times N\}$, $h_i(k)$ is a $M \times N$ matrix square. When p_i is the reference point, take as the origin of coordinates and construct the polar coordinate system on the template, the abscissa of the histogram is θ, the ordinate is $\log r$. Each square represents a small area in the template, and the darker the square, the more points fall within this area.

For the whole point set P, take the point $p_1, p_2, p_3, \ldots, p_n$ as the reference point respectively, the shape context of each point is calculated in turn, and finally the shape context of the target is obtained.

The SC is invariant to translation, scale and rotation, but in terms of discernibility, the descriptor lacks the orientation feature description of the target.

3 VIDR

An in-depth study of CCD and SC reveals that these two feature-description algorithms based on contour have the same weaknesses. SC has better anti-interference performance in describing shape contour features, so the experimental results are better than CCD. However, in the two pairs of shapes as shown in Fig. 1, the naked eye can distinguish well, and since both algorithms only process and calculate contour points, neither algorithm can distinguish well. Figure 2 show the CCD feature vector curve of the pair (1) in Fig. 1.

It is clear from Fig. 2 that the centroid contour eigenvectors of the two shapes are very similar.

Figures 3 and 4 show SC shape histogram of the pair (2) in Fig. 1, respectively.

Although SC has stronger robustness and higher recognition accuracy than CCD, it is difficult to distinguish a pair of shapes with very similar shapes because SC and CCD only consider shape contour points, but not the direction of the shapes.

We proposed VIDR method is through from a contour point perpendicular to the shape of the minimum circumscribed rectangle do respectively, so as to calculate the four vertical after shape inside of Euclidean distance and four vertical general Euclidean distance, to calculate the contour points within the vertical distance, finally calculate the shape within all the contour points of vertical distance ratio distribution, to say the shape characteristics of the shape.

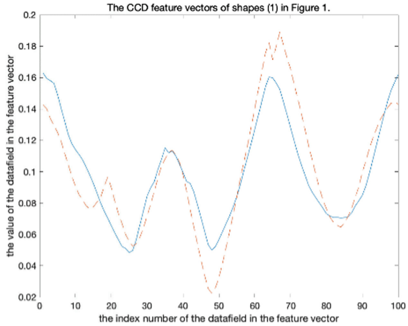


Fig. 2. The dashed line represents the CCD feature of the left shape (1), and the solid line represents the CCD feature of the right shape.

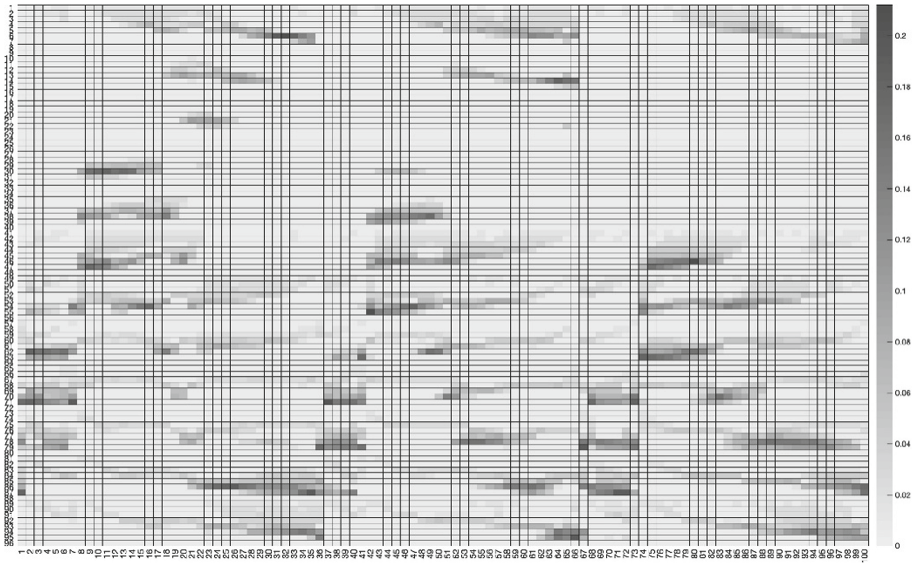


Fig. 3. The shape context histogram of the first picture in Group (1) in Fig. 1

Finding the shape of the smallest bounding rectangle is the most critical step in calculating the vertical interior distance. The method of finding the smallest bounding rectangle is introduced first, and then the specific calculation method of vertical interior distance ratio is introduced.

3.1 Minimum Bounding Rectangle

There are two types of restricted minimum rectangles for general graphics, minimum area limited minimum rectangle and minimum perimeter limited minimum rectangle. In this article, we use the smallest bounding rectangle with the smallest area. The following steps are key to calculating the minimum bounding rectangle.

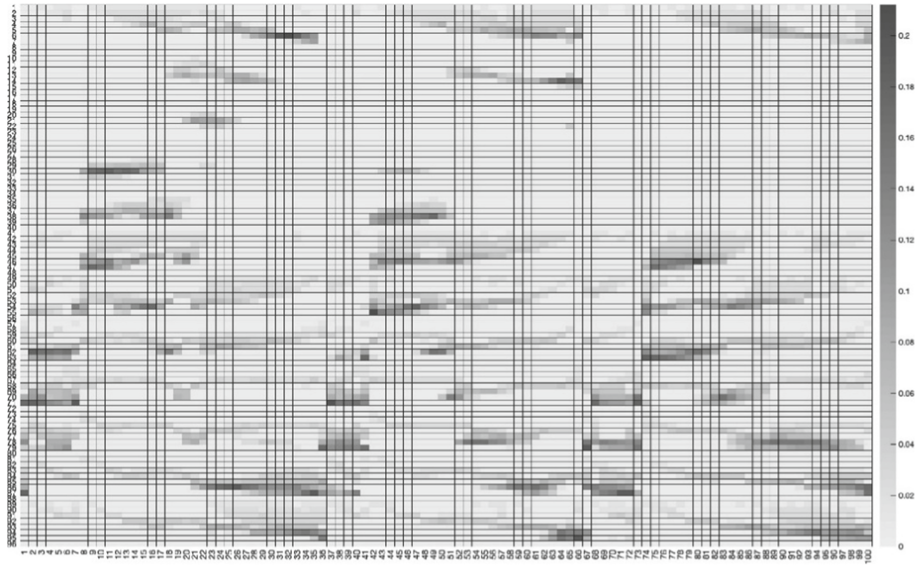


Fig. 4. The shape context histogram of the first picture in Group (2) in Fig. 1

Step 1: Use the function to find four points with maximum and minimum horizontal and vertical coordinates.
Step 2: Use these four points to construct the four tangents to the shape.
Step 3: If one (or two) lines overlap one side, the rectangular region identified by the four lines is calculated and saved as the current minimum. Otherwise, the current minimum is defined as infinity.
Step 4: Rotate the lines clockwise until one of them coincides with the edge of the polygon.
Step 5: Calculate the area of the new rectangle and compare it to the current minimum. If it is less than the current minimum, update and save the rectangle information that determines the minimum.
Step 6: Repeat steps 4 and 5 until the line is rotated more than 90°.
Step 7: Output the minimum region of the boundary rectangle.

Figure 5 is a schematic diagram of the smallest bounding rectangle for two pairs of shapes in Fig. 1.

3.2 Vertical Interior Distance Proportional Distribution

Within the vertical distance calculation is based on external rectangular shape, when calculating the shape of the external rectangular, starting from the shape of a contour point p_i separately to the four of the minimum circumscribed rectangle do vertical edge, $d_{i1}, d_{i2}, d_{i3}, d_{i4}$ are four vertical shape after the internal length of Euclidean distance, respectively, and seeking the sum of the Euclidean distance of four vertical d_0, by formula $R_i = (d_{i1} + d_{i2} + d_{i3} + d_{i4})/d_0$, pray that point within the vertical distance ratio, finally

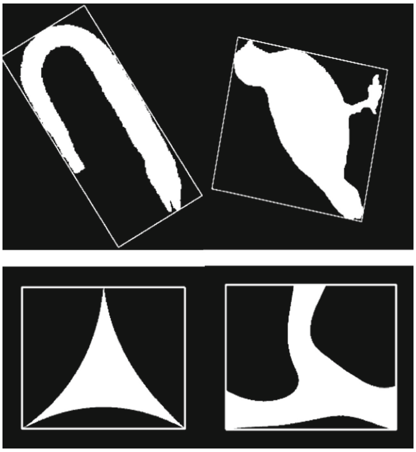


Fig. 5. The minimum bounding rectangles of two sets of shapes in Fig. 1

calculate the shape within all the contour points of vertical distance ratio R, forming the shape of the shape of the histogram, The histogram is the shape feature of the shape.

In the image, we convert the ratio of inner distance to the ratio of the number of pixels in the image, so that the computation is smaller and the algorithm is faster. The dot product operation of the image was used to calculate the intersection lines of the image interior and four perpendicular lines of a certain contour point p_i. The number of pixel points on the intersection line was first calculated, marked as n_i, and then the number of pixel points on the four perpendicular lines was calculated, marked as n_0. Finally, the formula $R_i = n_i/n_0$ was used to calculate the vertical interior distance ratio of this point. Within the specific range forms as shown in Fig. 6, respectively from the P point to do vertical rectangular four side, hand in the shape of other contour points P_1, P_2, P_3, P_4, and through the shape internal segment of PP_1, PP_2, PP_3, PP_4, which is the sum of the contour points of the distance, in the four lines contains the number of pixels to n_i, said again to find the external rectangle contained a long and a wide number of pixels n_0, the n_i/n_0 said that point within the vertical distance ratio.

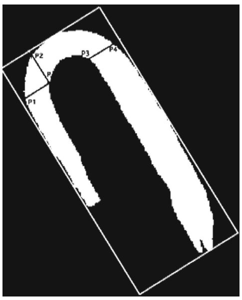


Fig. 6. Inner distance for shape contour point P

The VIDR method is applied to the two pairs of shapes in Fig. 1, and the shape histogram is shown in Fig. 7 respectively. The first picture is VIDR feature histogram in Group (1) in Fig. 1, and the second picture is VIDR feature histogram in Group (2) in Fig. 1. Obviously, in the shape features obtained by this method, the two pairs of shapes in Fig. 1 are significantly different, which can be distinguished by this method. And in the feature matching stage, this method only needs to do the Euclidean distance difference of two shape histograms, which has a faster matching speed.

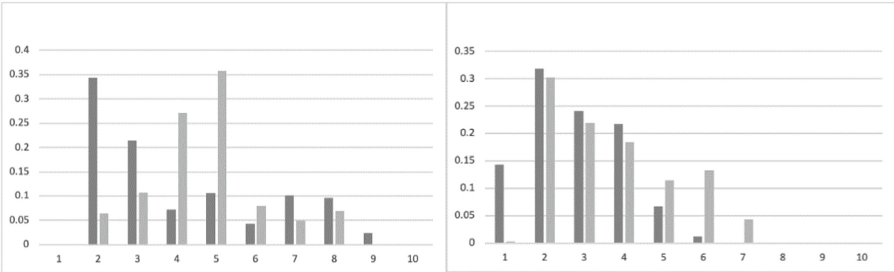


Fig. 7. VIDR feature histogram of two groups of shapes in Fig. 1

3.3 Shape Matching

The interior distance scale histogram of each shape can be expressed as sequence $\{x_j, j = 1, 2, 3\ldots 10\}$, we use the following formula to calculate the distance between the two shapes, the calculation formula is as follows:

$$dis(s1, s2) = \left(\sum_{i=1}^{10} (x_i - y_i)^2\right)^{\frac{1}{2}} \tag{6}$$

Where, x_i, y_i represents the histogram of the interior distance ratio of the shape $s1, s2$ respectively.

4 Experimental Results and Analysis

Mpeg-7 CE1B is an image template library used for image shape template by a large number of scholars, and many articles [1–9] use this database for experiments. The library contains 1400 shaped template images (70 categories, each containing 20 shapes), which are ideal for studying the shape characteristics of images. The proposed method is also tested on this data set. The experimental results and the test results of some other shape descriptors are shown in Table 1. Figure 8 shows a part of this datasets.

Table 1. Bull's eye test scores and matching time of some descriptors in MPEG-7 CE1 B

Descriptor	Bulls-eye-test score (%)	Matching time (ms)
VIDR (our)	76.88	80.0
CCD	68.60	112.3
ASD&CCD [10]	76.20	230.5
SC	64.59	85.0
IDSC [13]	68.83	85.2

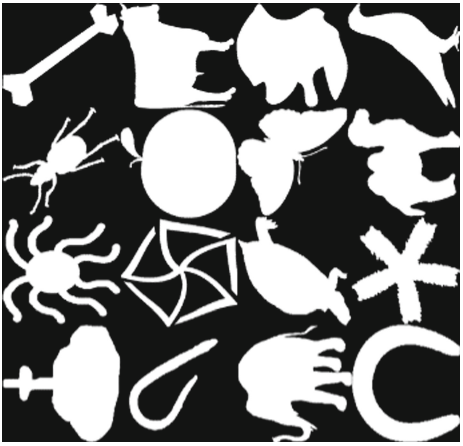


Fig. 8. Some shapes in MPEG-7 Part B

5 Conclusion

VIDR is a simple, effective shape descriptor with directional features. Because the descriptor uses shape histogram to represent features, it has global features, making it easier to match features. Because it can distinguish the shapes of similar centroid and similar shape context in Mpeg7 Part B database, such as Fig. 1, it can be seen from the experimental results that compared with other classical descriptors using contour, such as CCD and SC, which supplement the defects of directional features, it has more accurate test accuracy, stronger description ability and robustness. Compared with the improved CCD and SC, such as ASD&CCD and IDSC, it also has more accurate test accuracy and faster feature matching speed.

References

1. Premachandran, V., Kakarala, R.: Perceptually motivated shape context which uses shape interiors. Pattern Recognit. **46**, 2092–2102 (2013)
2. Alajlan, N., El Rube, I., Kamel, M.S., Freeman, G.: Shape retrieval using triangle-area representation and dynamic space warping. Pattern Recognit. **40**, 1911–1920 (2007)

3. Yang, X., Bai, X., Latecki, L.J., Tu, Z.: Improving shape retrieval by learning graph transduction. In: Proceedings of the European Conference on Computer Vision, Marseille, France, 12–18 October 2008, pp. 788–801 (2008)
4. Zahn, C.T., Roskies, R.Z.: Fourier descriptors for plane closed curves. IEEE Trans. Comput. **100**, 269–281 (1972)
5. Felzenszwalb, P.F., Schwartz, J.D.: Hierarchical matching of deformable shapes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007, pp. 1–8 (2007)
6. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 509–522 (2002)
7. El-ghazal, A., Basir, O., Belkasim, S.: Farthest point distance: a new shape signature for Fourier descriptors. Sig. Process. Image Commun. **24**, 572–586 (2009)
8. Yang, H.S., Lee, S.U., Lee, K.M.: Recognition of 2D object contours using starting-point-independent wavelet coefficient matching. J. Vis. Commun. Image Represent. **9**, 171–181 (1998)
9. Zheng, Y., Guo, B., Chen, Z., et al.: A Fourier descriptor of 2D shapes based on multiscale centroid contour distances used in object recognition in remote sensing images. Sensors **19**(3), 486 (2019)
10. Fotopoulou, F., Economou, G.: Multivariate angle scale descriptor of shape retrieval. In: Proceedings of the SPAMEC, Cluj-Napoca, Romania, 26–28 August 2011, pp. 105–108
11. Chang, C.C., Hwang, S.M., Buehrer, D.J.: A shape recognition scheme based on relative distances of feature points from the centroid. Pattern Recogn. **24**(11), 1053–1063 (1991)
12. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002). https://doi.org/10.1109/34.993558
13. Ling, H., Jacobs, D.W.: Shape classification using inner-distance. IEEE Trans. Pattern Anal. Mach. Intell. **29**(2), 286–299 (2007)
14. Shu, X., Wu, X.J.: A novel contour descriptor for 2D shape matching and its application to image retrieval. Image Vis. Comput. **29**, 286–294 (2011)
15. Zhang, J., Wenyin, L.: A pixel-level statistical structural descriptor for shape measure and recognition. In: 10th International Conference on Document Analysis and Recognition, pp. 386–390 (2009)

Auto-Encoder Based Wavelet and Extreme Learning Machine for Face Recognition

Mariem Kadri(✉), Salwa Said, and Ridha Ejbali

Research Team in Intelligent Machines, National Engineering School of Gabes, 6072 Gabes, Tunisia
salwa.said@isimg.tn

Abstract. Face recognition is a popular topic in biometrics. It has received increasing interests. Face recognition is widely used in many applications. Such as surveillance, identification in login system and personalized technology. Recognizing multiple face in real-time on the embedded system is very challenging due to the complexity computation of the processing. Extreme learning machine (ELM) has garnered tremendous success and attention as an efficient learning algorithm due to its good generalization performance in various problems such as face recognition. In this paper we present a new method for face recognition using an auto encoder based wavelet analysis and extreme learning machine. For more details, we have used a composite wavelet activation function at the hidden nodes of ELM and the learning step is accomplished by the Deep ELM. For the data collection we have used the Face Recognition Grand Challenge (FRGC) v.2 database.

Keywords: Face recognition · Extreme learning machine (ELM) · Wavelet neural network · Auto-encoder · FRGC v.2

1 Introduction

Face recognition (FR) [1, 2] has been a long-standing research topic and still a challenging question in computer vision due to the rise in security demands and the development of imaging technologies. It is applied broadly in crime detection, identification of genetic diseases, security information, and image understanding. Facial recognition is the process of identifying and classifying person using their face. This process involves three main steps face detection, face capture, face match. The face detection detects faces within the images and video. Face captures convert a set of analogue information into a set of data based on the person's facial features. Whereas face match compares two faces to verifies if these faces belong to the same person (Fig. 1).

Face recognition have progressed enormously and becomes more and more ubiquitous due the significant breakthrough made by such machine learning techniques such as DeepIDs, DeepFace, Face++, FaceNet, and Baidu. The most machine learning used is the artificial neural network which is new data representation. Different from the traditional data representation models, this information processing paradigm is based on the

A. Abraham et al. (Eds.): HIS 2020, AISC 1375, pp. 11–21, 2021.
https://doi.org/10.1007/978-3-030-73050-5_2

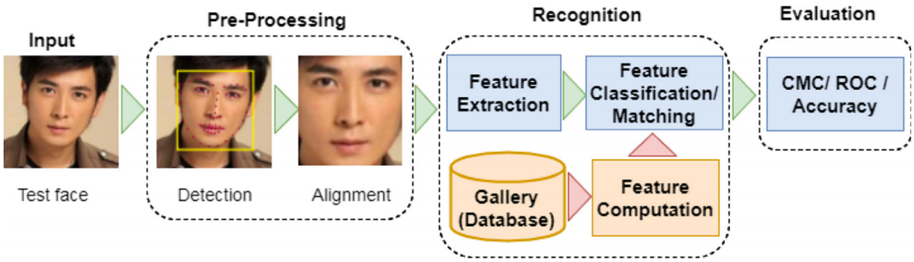


Fig. 1. Face recognition pipeline

model of human biological nervous systems. Its capability to process data has motivated many researchers to introduce new model for data processing basing on the combination of such theories. The combination of the wavelet theory and the artificial neural network make the introduction of new neural network called wavelet neural network. The wavelet neural network is proposed by Zhang and Benvesite [3]. Studies shows that this network yields accurate results on the classification problems and the nonlinear system problems. The key discriminate feature of this network is the properties of wavelet decomposition for localization in both time and frequency domain.

Another prominent technique is the Extreme Learning Machine [4]. It is a specific type of feed-forward neural networks. This network is characterized by a single layer where the parameters of hidden nodes (not just the weights connecting inputs to hidden nodes) need not be tuned. These hidden nodes can be randomly assigned and never updated. The main goal of the ELM is to tackle the barriers between the conventional artificial learning techniques and biological learning mechanism to represent a reliable machine learning technique in which hidden neurons need not to be tuned. Hence this make this technique faster than the traditional neural networks in the training phase. Jian Chen [19] adopt a conjugate gradient extreme learning machine for facial expression recognition. Here authors outline the positive definiteness of the extreme learning machine. Lee [20] introduce a new 2.5D face descriptor that based on the Regional Covariance Matrix (RCM). They present a new variant of ELM named Random Maxout Extreme Learning Machine (RMELM) for FR recognition.

The key feature of these techniques is the feature engineering process. Feature engineering is the craft of extracting reliable feature from dataset and make it visible to learning algorithms. The main drawback of the traditional machine learning algorithms is that the feature extraction is made by experts. Furthermore, to achieve an effective result we need a massive amount of training data to solve complex problems. The training data preparation consist of the creation of labelled data which is tedious and time-consuming task.

Machine learning undergone important developments. A significant contributor to that surge is the deep learning. The main contribution of this network is that learn the feature extraction part.

Recently, deep learning [5, 6] networks giant strides on this domain, typically these networks are classified into 2 types which are supervised and unsupervised networks. The supervised networks involve the convolution and the recurrent neural network. The training of these networks is based on a huge number of labeled data. However today the lack of annotate dataset became a big challenge today. The unsupervised neural networks meet this issue because there is no need for annotated date. Data representation and compression is vital step to ensure a good classification. The auto-encoder provide a good data representation therefore it is widely used for the classification and the recognition problem. Regarding the context of the face detection and classification, this network shows a pertinent result which motivate many researches to adopt it. In their work, Jemel et al. [7] present a new algorithm of classification based on deep learning and sparse auto-encoder. Basing on the auto-encoder, they introduce a new classifier. Researchers combine the auto-encoder architecture and some specific type of neural network such as the wavelet neural network for example Salima et al. [8] propose a cascade model that involves the sparse auto-encoder and the fast wavelet transform algorithms. Ramzi et al. [9] adopt this combination for medical classification.

Ridha and all present a new method based on a denoising auto-encoder with Dropout-based network anomaly detection to enhance the intrusion detection. The key feature of this network is the good features extraction that yields a good detection accuracy. [10] introduce a cascade model which is based on the combination of the denoising and the sparse auto encoders. For the calcification ask they adopt 2 classifiers multi-class SVM and Softmax classifier. They collect data from different databases such as The Yale Face Database (Yale), The ORL face database, by the Olivetti Research Laboratory in Cambridge (ORL). However, this model is slow to train. [11] also adopt a fusion model based on stacked convolutional auto-encoder and sparse task. Here they exploit the ability the auto-encoder to extract deep and high-level feature and the capability of the sparse representation for the classification. For the data collection, they collect data from the LFW face database. The fast training process is a key feature for a good model selection. Researchers adopt the Extreme Learning Machine to ensure a faster training learning and extent it by the auto-encoder for example [12] introduce a new deep model based on the sparse auto-encoder and the Extreme Learning Machine for the object recognition.

2 Proposed Approach

The main goal of this paper is to apply new method based on a deep model using an auto-encoder based wavelet and extreme learning machine for Face recognition.

The building blocks of the proposed architecture are the Extreme Learning Machine, the autoencoder and the wavelet neural network.

2.1 Deep Learning

Deep learning can be considered as a machine learning technique that learns features and tasks directly from the data. While traditional machine learning algorithms are linear, deep learning algorithms are stacked in a hierarchy of increasing complexity and abstraction. This structure based on the artificial neural network whose purpose to imitate the human behavior. This makes a huge impact in the area of machine learning that overcomes the limitations of the modern machine vision algorithms for example