IST-Africa 2018 Conference Proceedings
Paul Cunningham and Miriam Cunningham (Eds)
IIMC International Information Management Corporation, 2018
ISBN: 978-1-905824-59-5

Automatic Attendance Capturing Using Histogram of Oriented Gradients on Facial Images

Abejide ADE-IBIJOLA[1], Kehinde ARULEBA[2]
[1]*University of Johannesburg, Johannesburg, South Africa*
Email: abejideai@uj.ac.za
[2]*University of the Witwatersrand, Johannesburg, South Africa*
Email: arulebak@gmail.com

Abstract: Humans mostly use faces to identify/recognise individuals and the recent improvement in the capability of computing now allow recognition and detection automatically. However, there still exist quite a number of problems in the automatic recognition of facial images. Histogram of Oriented Gradients (HOG) has been recently adopted and seen as a standard for efficient face recognition and object detection generally. In this paper, we investigate and discuss a simple but effective approach to capturing student's attendance register in a lecture hall by making use of HOG features for detecting and recognising students face at different moods, orientations, and illuminations. Our experiment detection and recognition output show a good performance on our facial image database obtained from the University of Johannesburg, this performance is due to HOG descriptors attributes which are robust to changes in rotation and illuminations. Our system will help to save instructional staff/lecturer time by eliminating manual calling of students name and also help monitor students.

Keywords: Face recognition, Histogram of oriented gradients, Feature extraction.

1. Introduction

Face recognition is a computer vision research area that has attracted a remarkable interest recently due to its application in pedestrian detection [1], [2], crime detection [3], [4], and biometrics [5], [6], also, the access to cheap computers and digital cameras have helped facial recognition research. Generally, facial features could be grouped as photometric or geometric for recognition. Various face recognition algorithms have been proposed such as principal component analysis [7], local binary pattern [8], independent component analysis [9], a detail discussion of this algorithms is presented in [10].

In this paper, we used a histogram of oriented gradients (HOG) [11] - HOG is a feature descriptor used in computer vision field for object detection, we used HOG to extract the features of our facial images. We applied HOG to a database of images obtained from twenty participants at the University of Johannesburg, for extracting distinctive features of their facial images at different moods, looks and face positioning. The approach counts the number of times the gradient orientation in the localised portions of the facial image occurs. Based on [12] approach, we used a regular grid to extract the HOG descriptors to reduce errors found in the facial feature during recognition that might occur due to illuminations, pose and occlusions. Also, a useful structure for recognition of facial images was achieved by a grouping of HOG descriptors at different levels.

This paper is structured as follows. Section 2 discusses the background of the study and related work. Section 3 presents the histogram of oriented gradients algorithm. Section 4 presents face recognition using HOG. Section 5 discusses the dataset and methodology used in the experiments. Section 6 presents the conclusion and future work.

2. Related Work

Human face detection is a cumbersome task in computer vision due to the varieties of looks, appearances, and facial features. In recognising a human face, a robust set of features that allows the face to be discriminated clean, even under bad illumination or clustered backgrounds will be needed. Shireesha and Raghunadh [13] proposed a system for managing attendance, this system is based on face detection and recognition algorithms. Once a student enters the lecture venue, the system automatically detects his/her face and takes the student attendances after recognising him/her. This system was evaluated and the result showed that it saves time and more efficient compared to the manual attendance system.

Visar and Agni [14] proposed a real-time algorithm for facial image recognition used in automatic attendance taking. This system is integrated into an existing learning management system, it detects and takes students attendance automatically on attending a lecture. The system uses some machine learning algorithms and adaptive learning approach to keep track of changes in facial images over a long period of time. Navneet and Bill [15] studied the problem of feature sets for human detection by using linear support vector machine (SVM) [16] based on human detection as a test case and found out that HOG descriptors normalised locally give better performance compared to other existing feature sets like wavelets [17]. Alberto et al. [12] proposed a new algorithm for face recognition using the common EBGM (Elastic Bunch Graph Matching). The result of their recognition experiments using available public face databases showed a better performance compared to other approaches, the better performance result is attributed to the properties of HOG which are - small displacements and rotation, changes in illumination. HOG descriptors have also be used for pedestrian detection [15], [18], [19]. In this, objects are seen as a fixed scale and are broken down into smaller connected regions at fixed positions. At each region, a descriptor is gotten and the combination of these HOG descriptors is taken as the object representation.

Déniz et al. [20] extracted HOG descriptors from a regular grid, this reduces the error in facial features recognition that might be caused by pose, occlusions, and changes in illumination. They performed dimensionality reduction to get rid of noise, thereby making the classification stage less prone to overfitting. Mikolajczyk et al. [21] combined binary-threshold gradient magnitudes with orientation position histograms to develop a detector for faces, side and front parts of the body and heads. The developed detector is based on an efficient algorithm i.e. parts to be recognised uses the same features initially.

Yang et al., [22] presented an efficient approach to recognising human from sequences of depth maps, this gives more information about motion and body shape for action recognition. Their method projected depth maps into orthogonal planes and accumulate global activities through the entire video sequences to generate Depth Motion Maps (DMM). DMM are then used to compute the HOG. Geng and Jiang [23] studied Scale Invariant Feature Transform (SIFT) [24] for object detection/recognition and proposed two new approaches based on SIFT namely: Partial-Descriptor-SIFT (PDSIFT) and Volume SIFT(VSIFT) for face recognition. These approaches are based on the original SIFT algorithms. They compared the holistic approaches: Eigenfeature Regularisation and Extraction, Fisherface and the null space approach with the feature based approaches: PDSIFT and SIFT. Their experiments on the ORL[1] and AR[2] databases indicates that PDSIFT performance compared to the original SIFT algorithm is better. Principal Component Analysis [25], [26] is a statistical approach well used in recognising facial images. It reduces the number of variables in facial images. In this, the training set images

[1] www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html
[2] www2.ece.ohio-state.edu/~aleix/ARdatabase.html

are depicted as a linear combination of weighted eigenvectors known as eigenfaces. These eigenvectors are gotten from the covariance matrix of each image in the training set.

3. Histogram of Oriented Gradients

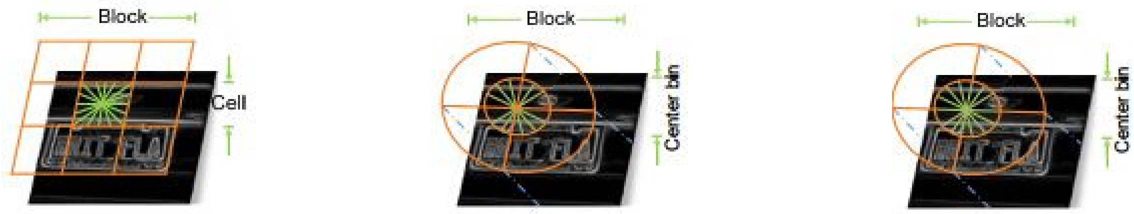
The HOG approach is similar to scale-invariant feature transform [27], shape contexts [28] and edge orientation histogram [29]. HOG is a feature descriptor used in object detection. This approach counts occurrences of gradient orientation in localised portions of an image. The representation of an image patch by extracting useful features and ignoring other details. Basically, the description converts an input image of heightH×widthW×3 (RGB) to a feature vector of length *l*. In HOG descriptor, the standard input image size is 64×128×3 and the feature vector size of the output is 3780 [15]. Mathematically, HOG can be expressed as –

$$G_a = I(a+1,b) - I(a-1,b) \qquad (1)$$
$$G_b = I(a,b+1) - I(a,b-1) \qquad (2)$$

where $I(a,b)$ is the pixel intensity at position (a,b), and G_a and G_b are the vertical and horizontal properties of the gradients.

$$M(a,b) = \sqrt{G_a^2 + G_b^2} \qquad (3)$$
$$\theta_{a,b} = tan^{-1}\frac{G_a}{G_b} \qquad (4)$$

$M_{a,b}$ is the gradients magnitude, $\theta_{a,b}$ is the gradient angle at a given position. HOG descriptors fall into two major categories namely: Rectangular HOG (R-HOG), see Figure 1(a) and Circular HOG (C-HOG), see Figure 1(b) and Figure 1(c) [15].
In this paper, we use R-HOG descriptor to divide our input image into blocks made up of grids.



(a) A rectangular HOG (R-HOG) descriptor with 3×3 blocks of cells

(b) Circular HOG (C-HOG) descriptor with the central cell divided into angular sectors as in shape contexts

(c) C-HOG descriptor with a single central cell

Figure 1: HOG Descriptors [15]

4. Face Recognition Using HOG Description

As discussed earlier, HOG has emerged as one of the most successful algorithms for facial image description/detection due to its efficiency in handling image transformations like illumination, image rotation, and scale changes. In this paper, input facial images window is first divided into cells or small spatial regions, each cell in the window accumulates a local 1-D histogram of edge orientations or gradient directions over the pixel of the cells. These cells are pixel regions that can either be circular or rectangular in shape, and the histogram bins are distributed evenly from $0^0 - 180^0$ (when there is signed orientation, we have

$0^0 - 360^0$), making every histogram bin have a spread of 200. Combination of all the histogram entries makes up the representation. We performed contrast normalisation on the local output before using them to get better invariance to illumination. This was achieved by getting a measure of the local histogram over a block or bigger spatial regions and the result was used to normalise all the small spatial regions in the block. The normalised descriptor blocks are also known as HOG descriptors.

The main steps involved in this algorithm are:

- scale-space extrema detection
- orientation assignment
- keypoint descriptor

Scale-space extrema detection aims at achieving scale invariance. This is performed by extracting mainly features at the local extrema of the representation (scale-space) of the input image. Orientation searches for the gradient orientation using the gradient information of the image, after this, the gradients of the image are made relative to this dominant direction.

ALGORITHM 1: Overview of HOG

Input: Facial Image

Output: Feature vector containing extracted features

Methodology

Step 1: Preprocess input images
Step 2: Calculate the gradient image
Step 3: Compute HOG in 8×8 cells
Step 4: 16×16 block normalisation
Step 5: Compute the HOG feature vector
Step 6: Visualise the HOG

5. Dataset and Methodology

All experiments carried out in this paper were done on our database of images from the University of Johannesburg, containing 250 facial images from 25 different people. It contains images with different moods, actions, and pose. The HOG descriptors used in this paper is computed on 92×112 patch of a facial image, 8-bit gray levels (all images are converted from RGB color space to grayscale, see Figure 3) organised in 25 directories - one for a single face in different moods and looks. We analysed the image patches at multiple scales in numerous image locations. Patches are computed on a fixed aspect ratio. An aspect ratio of $92{:}\ 112$ and a patch size of 100×200 was used for computing our HOG feature descriptor. For easy computation, images are normalised into a dynamic range of $[0,1]$ and converted to double precision.

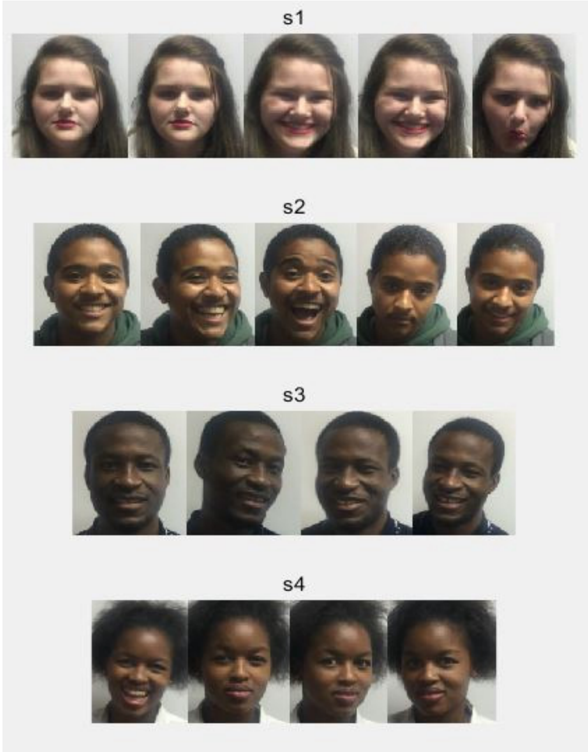


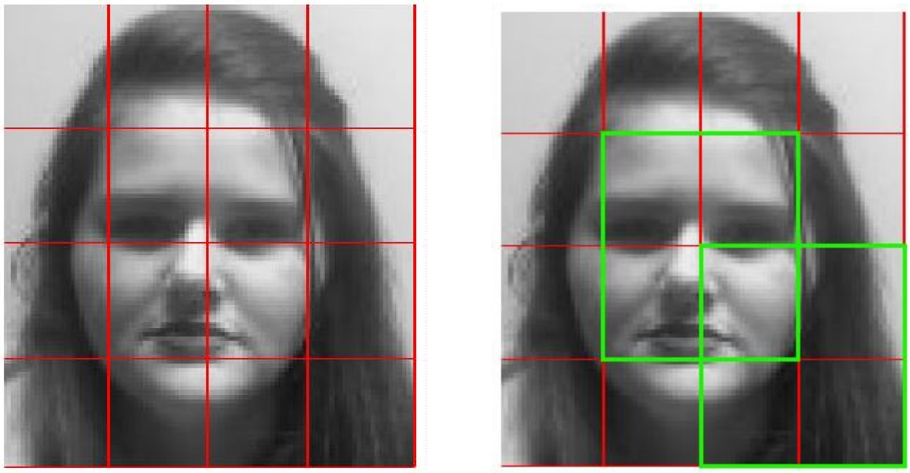
Figure 2. Sample Images from our Database

Then, the image gradients are processed by lapping the image with mask $[+1; 0; -1]$ along a and b axes without smoothing, at each direction G_a and G_b are gotten as the gradient matrices as illustrated in Figure 3. The gradient magnitudes are then computed by the norm of G_a and G_b as shown in Equation 3 which shows the contour and some other texture details of the face image.

(a) Original image (b) Image after Preprocessing

Figure 3. Preprocessing for HOG Feature Extraction

The image window is then split into the same dense sampled grid of points. At every point, the square pixel region in the image is centered, this is known as block and it is divided into cells (see Figure 4), and the block steps in both a and b directions are 8 pixels, this implies that the rate of overlapping is set as 50% to allow each cell contribute at least once to block histograms.



(a) An image divided into cells of size 8×8. (b) Cells grouped into blocks with 50% block overlapping rate.

Figure 4. Preprocessing for HOG Feature Extraction

The vertical and horizontal gradients were first calculated, the histogram gradients were computed using the Sobel operator with kernel size 1. Then, the magnitude and direction of the gradient were computed. The gradient image get rides of much irrelevant information making the image still visible i.e. easy recognition of the face in the image $8 \times 8 \times 3$. For images with color, the gradients of the RGB are evaluated. In this, the facial image is divided into 8×8 cells and a HOG is computed for each cell. An 8×8 image has pixel values. Magnitude and direction per pixel (i.e. $8 \times 8 \times 2$ pixels) are the two values contained in the gradient of the image (patch). The computed histogram over a patch allows this system to be robust to noise. A 8x8 patch makes the system less prone to noise irrespective of individual gradient noise.

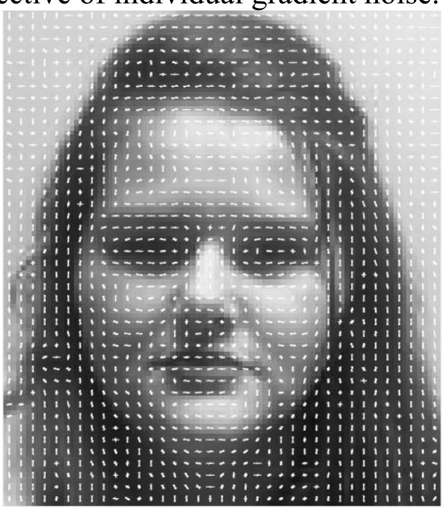


Figure 5. HOG visualisation for Figure 3(b)

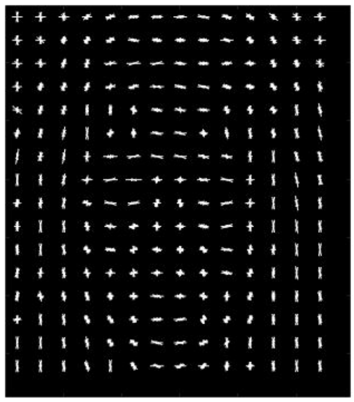


Fig. 6. Extracted Histogram of Oriented Gradient Features for Fig. 3(b)

Figure 6 shows the blocks, each block shows the histogram. At every block, we obtained a vector of length 36, this is computed by multiplying the number of bins in each histogram by the number of cells in a block. The process is iterated for every block, and normalisation is carried out on the resulting vector before been transformed into a large HOG feature vector in sequence.

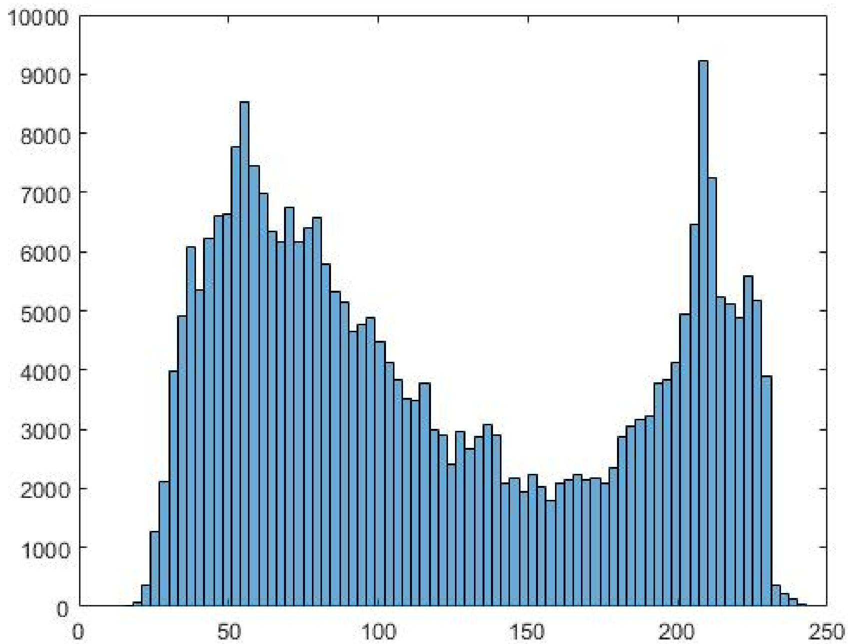


Figure 7. Histogram of Figure 3(b)

6. Conclusion and Future Work

This paper describes and investigates the application of HOG features for facial image recognition. Our experiment recognition outputs show a good performance on our facial image database obtained from the University of Johannesburg, this performance is due to HOG descriptors attributes which are robust to changes in rotation and illuminations. Central to the good performance of our technique is that the features are taken from a big set of blocks at multiple positions aspect ratios and sizes. For the future, an efficient classifier using Support Vector Machines (SVM) will be developed.

Also, this paper is based on snapped images, developing a recognition system that will work for a video camera with real-time recognition will be a thing to look at.

References

[1] S. Paisitkriangkrai, C. Shen, and A. Van Den Hengel, “Pedestrian detection with spatially pooled features and structured ensemble learning”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 6, pp. 1243–1257, 2016.

[2] C.S. Fahn, C.P. Lee, and Y.S. Yeh, “A real-time pedestrian legs detection and tracking system used for autonomous mobile robots”, in International Conference on Applied System Innovation (ICASI), 2017, pp. 1122–1125, IEEE, 2017.

[3] M. Sharif, S. Bhagavatula, L. Bauer, and M. K. Reiter, “Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition”, in Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, pp. 1528–1540, ACM, 2016.

[4] Y. Gao, J. Ma, and A. L. Yuille, “Semi-supervised sparse representation based classification for face recognition with insufficient labeled samples”, IEEE Transactions on Image Processing, vol. 26, no. 5, pp. 2545–2560, 2017.

[5] H. C. I. Jonathan, S. U. Pankanti, and N. K. Ratha, “Facial analysis by synthesis and biometric matching”, Dec. 6 2016. US Patent 9,514,354.

[6] D. J. Berini, G. A. Van Beek, I. Arnon, B. J. Shimek, R. B. Fevens, and R. L. Bell, “Multi-biometric enrolment kiosk including biometric enrolment and verification, face recognition and fingerprint matching systems”, Feb. 9, 2016. US Patent 9,256,719.

[7] I. T. Jolliffe, “Principal component analysis and factor analysis”, in Principal Component Analysis, pp. 115–128, Springer, 1986.

[8] T. Ahonen, A. Hadid, and M. Pietik¨ainen, “Face recognition with local binary patterns”, Computer Vision-eccv 2004, pp. 469–481, 2004.
[9] A. Hyvarinen, J. Karhunen, and E. Oja, Independent component analysis, vol. 46. John Wiley & Sons, 2004.
[10] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, “Face recognition: A literature survey”, ACM Computing Surveys (CSUR), vol. 35, no. 4, pp. 399–458, 2003.
[11] G. Tsai, “Histogram of oriented gradients,” University of Michigan, 2010.
[12] A. Albiol, D. Monzo, A. Martin, J. Sastre, and A. Albiol, “Face recognition using HOG-EBGM”, Pattern Recognition Letters, vol. 29, no. 10, pp. 1537–1543, 2008.
[13] S. Chintalapati and M. Raghunadh, “Automated attendance management system based on face recognition algorithms”, in Computational Intelligence and Computing Research (ICCIC), 2013 IEEE International Conference on, pp. 1–5, IEEE, 2013.
[14] V. Shehu and A. Dika, “Using real-time computer vision algorithms in automatic attendance management systems”, in Information Technology Interfaces (ITI), 2010 32nd International Conference on, pp. 397–402, IEEE, 2010.
[15] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection”, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005. vol. 1, pp. 886–893, IEEE, 2005.
[16] S. Suthaharan, “Support vector machine”, in Machine Learning Models and Algorithms for Big Data Classification, pp. 207–235, Springer, 2016.
[17] N. Kingsbury, “Image processing with complex wavelets”, Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 357, no. 1760, pp. 2543–2560, 1999.
[18] T. Watanabe, S. Ito, and K. Yokoi, “Co-occurrence histograms of oriented gradients for pedestrian detection”, Advances in Image and Video Technology, pp. 37–47, 2009.
[19] F. Suard, A. Rakotomamonjy, A. Bensrhair, and A. Broggi, “Pedestrian detection using infrared images and histograms of oriented gradients”, in Intelligent Vehicles Symposium, 2006 IEEE, pp. 206–212, IEEE, 2006.
[20] O. Deniz, G. Bueno, J. Salido, and F. De la Torre, “Face recognition using histograms of oriented gradients”, Pattern Recognition Letters, vol. 32, no. 12, pp. 1598–1603, 2011.
[21] K. Mikolajczyk, C. Schmid, and A. Zisserman, “Human detection based on a probabilistic assembly of robust part detectors”, Computer Vision-ECCV 2004, pp. 69–82, 2004.
[22] X. Yang, C. Zhang, and Y. Tian, “Recognizing actions using depth motion maps-based histograms of oriented gradients”, in Proceedings of the 20th ACM international conference on Multimedia, pp. 1057–1060, ACM, 2012.
[23] C. Geng and X. Jiang, “Face recognition using sift features”, in Image Processing (ICIP), 2009 16th IEEE International Conference on, pp. 3313–3316, IEEE, 2009.
[24] D. G. Lowe, “Distinctive image features from scale-invariant keypoints”, International Journal of computer vision, vol. 60, no. 2, pp. 91–110, 2004.
[25] X. Luan, B. Fang, L. Liu, W. Yang, and J. Qian, “Extracting sparse error of robust PCA for face recognition in the presence of varying illumination and occlusion”, Pattern Recognition, vol. 47, no. 2, pp. 495–508, 2014.
[26] P. Liton and A. S. Abdulla, “Face recognition using principal component analysis method”, International Journal of Advanced Research in Computer Engineering and Technology, vol. 1, no. 9, pp. 135–139, 2012.
[27] H. Lad and M. A. Mehta, “Feature-based object mining and tagging algorithm for digital images”, in Proceedings of International Conference on Communication and Networks, pp. 345–352, Springer, 2017.
[28] S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 509–522, 2002.
[29] X.M. Shen, J.X. Zhou, and T.W. Xu, “Minority costume image retrieval by fusion of color histogram and edge orientation histogram”, in 15th International Conference on Computer and Information Science (ICIS), 2016 IEEE/ACIS, pp. 1–7, IEEE, 2016.

