ISSN: 2321-7782 (Online)
e-ISJN: A4372-3114
Impact Factor: 7.327

Volume 6, Issue 1, January 2018

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

Face Recognition using Histogram of Oriented Gradients

Sourabh Hanamsheth[1]
Dept. of Electronics Engineering
Vishwakarma Institute of Technology
Pune – India

Milind Rane[2]
Dept. of Electronics Engineering
Vishwakarma Institute of Technology
Pune – India

Abstract: Various methods or experiments can be used for face recognition and detection however two of the main include an experiment that evaluates the impact of facial landmark localization in the face recognition performance and the second experiment evaluates the impact of extracting the HOG features from a regular grid and at multiple scales. We study the question of feature sets for robust visual object recognition. The Histogram of Oriented Gradients significantly outperform other existing methods like edge and gradient based descriptors. We study the influence of each stage of the computation on performance, concluding that fine-scale gradients, fine orientation binning, relatively coarse spatial binning, and high-quality local contrast normalization in overlapping descriptor blocks are all important for good results. Comparative experiments show that though HOG is simple feature descriptor, the proposed HOG feature achieves amazing results with much lower computational time.

Keywords: Histogram of Oriented Gradients, Face Recognition, Histogram

I. INTRODUCTION

Various selection methods and feature extraction methods are widely being used. Other than the holistic methods such as LDA, PCA and Fisher Face the local descriptors have been studied recently.[7] The descriptors having large inter-class variance and small intra-class variance are considered to be ideal descriptors for local facial regions. Among the various descriptors that have been developed for appearance of the image patches, local binary pattern feature yields some of the best results when used for facial images. The idea behind using local binary pattern is that the faces can be seen as a composite of micro-patterns, which are very well defined by this operator.[8] In practice the system has to reduce the number of possible scales or the number of local regions to form a reasonable length feature vector as there are too many micro-patterns.

As one of the most successful applications of image analysis and understanding, face recognition has recently received significant attention, especially during the past several years.

Histogram of Oriented Gradients (HOG) is one of the descriptors among many others. There are a few advantages of HOG over the other descriptors. As HOG operates on local cells, it is invariant to geometric transformations and photometric transformations, except for object orientation. These changes will only appear in larger spatial regions. The discovery by Dalal and Trigg's clearly states that the fine orientation sampling, coarse spatial sampling and strong photometric normalization permits the movement of pedestrians to be ignored as long as they maintain an upright position.[2] The local features and shapes in HOG can be characterized by the distribution of local intensity gradients or edge orientation even without having the exact knowledge of the specific gradient or edge position. As the histogram gives translational variance it is robust to lighting changes. The HOG feature summarizes the distribution of measurements within the image regions and is particularly useful for recognition of textured objects with deformable shapes.

We all know that Intensity of an image is the measure of light in an image, it also signifies the numerical value of a pixel. Intensity isn't an absolute measure, but relative. It is found to be helpful in determining the contour shapes, difference in intensity, texture and shades of the constituent elements. Various methods are in practice for implementation of Face Recognition System, examples include- Eigen Faces, Fisher Faces and Local Binary Pattern Histogram. Each have their own associated pros and cons.[9] However, after experimentation and comparative study, we found that Histogram of Oriented Gradient gave the most competitive result.

II. Histogram of Oriented Gradients

In computer vision and image processing, histogram of oriented gradients is feature descriptor used for the purpose of object detection. The appearance of gradient orientation in localized portions of image are counted. This method is identical to that of scale invariant feature transform descriptors, edge orientation histograms and shape contexts, but it uses overlapping local contrast normalization and is computed on a dense grid of uniformly spaced cells to improve the accuracy

A .Basic Theory

Like the SIFT and EBGM method, the HOG feature is generated for each key-point of an image. The neighboring area around each key-point in the image is divided into uniformly spaced cells. For each cell a local 1-D histogram of edge orientations or gradient directions is accumulated over all the pixels of the cell. The feature of a key-point is formed by the histogram entries off all cells around that key-point. The image is represented by combining the histogram features of all key-points.

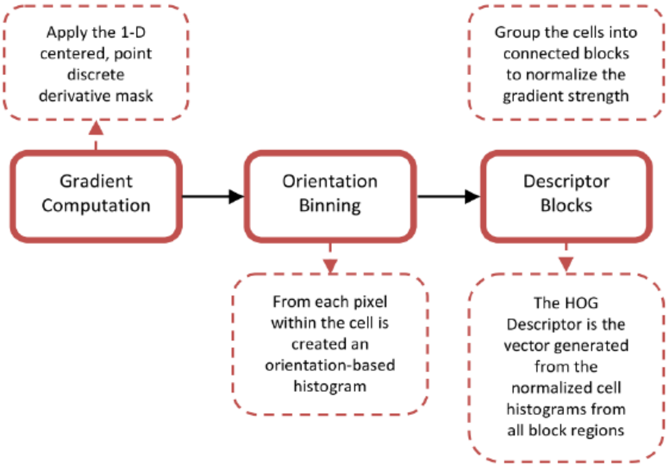

Fig.1. Block diagram of the overall process.

The histogram of oriented gradients (HOG) is a dense feature extraction method for images. Dense means extracting features for all locations in the image (or a region of interest in the image) as opposed to only the local neighborhood of key-points like SIFT (Scale Invarient Feature Transformation) Intuitively it tries to capture the shape of structures in the region by capturing information about gradients. It does so by dividing the image into small (usually 8x8 pixels) cells and blocks of 4x4 cells. Each cell has a fixed number of gradient orientation bins. Every pixel in a cell votes for the bin for orientation of gradient with a vote commensurate to the gradient magnitude at that pixel. To reduce aliasing, the pixels votes are bi-linearly interpolated. This interpolation happens in both the orientation and position. This statement is important, it means that a pixel will not only vote for its orientation bin, but also to neighboring orientation bins (e.g. the gradient orientation at a pixel is 45 degrees, it will vote with a weight of 0.5 for the 35 to 45 degree bin and a weight of 0.5 for the 45 to 55 degree bin). Likewise it will also vote for the other two orientation bins not only in its cell, but also in the neighboring four cells of its cell. The distance of the pixel from the cell center is used to determine the weight. Histograms are also normalized based on their energy (regularized L2 norm) across blocks. As the block has a step size of one cell, a cell will be a part of four blocks. The four differently normalized versions of cell's histogram are thus defined. These 4 histograms are concatenated to get the descriptor for the cell.

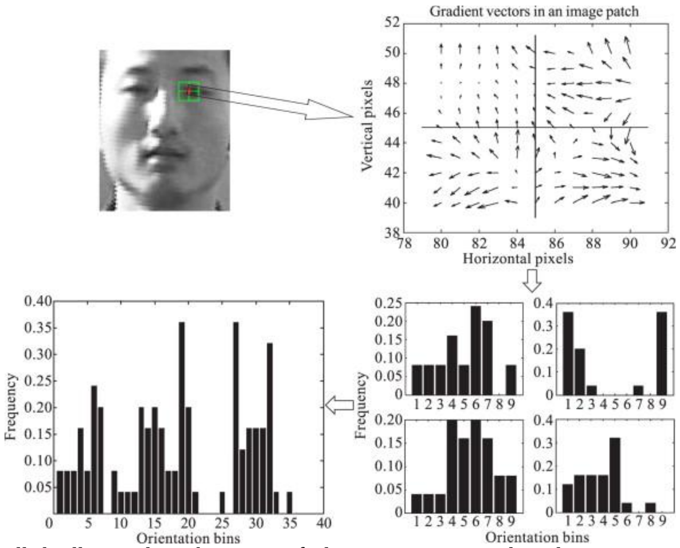

Fig.2. Image divided into small regions called cells. Local 1-D histogram of edge orientation or gradient direction are accumulated and concatenated to form the final histogram feature.[1]

B. Orientation

Orientation can either be described as a single angle orientation or a double angle orientation. A single angle treats a given edge as having opposite orientation whereas the double angle maps it into the same orientation. More patterns may be distinguished by using single angle representation. Since, we are using images from an already existing Database we will only need two variables in order to describe the gradient vector. In this experiment we have used single angle orientation to allow more differentiation between patterns. Euler Angles can be used to describe the orientation of a three dimensional body.

C. Normalization

Normalization is a process that changes pixel intensity values. It is often also referred to as 'Histogram Stretching' or 'Contrast Stretching' The process of normalization is usually carried out after calculating the vectors. It helps us bring the signal or image into a range that is more familiar or normal to the senses, it improves contrast. The purpose of normalization is usually to bring the image, or the type of signal under study, into a range that is more relevant to us, the goal being achieving uniformity for better interpretation. Normalization has three approaches: first is Normalization to standard interval - (0,a) or (0,255). Second approach is Normalization to zero mean and unit variance where mean when changed to zero- wipes all the information in the image. Hence, it is advisable that we keep the mean value the midpoint value of intensity. Lastly, Histogram Equalization spreads the values to full range, i.e. 0-255; set the mean to midpoint and standard deviation or variance to 0.288 of the range. We have used L2 norm in our experiment. This scheme was used owing to its better comparative performance.

D. Overlapping in HOG

For matching of the two facial images accuracy is very important. Though 100% accuracy is not possible the accuracy further decreases drastically due to motion blurring or bad light conditions. The histogram provides some equilibrium to it, but it is not enough. Thus to overcome this, HOG feature is introduced. This was inspired by Dalal and Trigg's conclusion that the redundant information introduced by HOG significantly improves the performance.[2]

Overlapping in HOG significantly improves the performance of detection and identification which would have been otherwise quite difficult in presence of poor lighting conditions and motion and thus the information available, though redundant is highly effective in reducing the rate of false-positives. The method we used here is to divide the image into small connected regions. These small regions are called cells. Then for each individual cell compute a histogram of edge orientation or gradient direction for the pixels within the cell. Separate each cell according to its orientation into angular bins. The pixel of each cell provides weighted gradients to its respective angular bin. The neighboring cells form a group and are termed as spatial regions called blocks. This grouping of cells into blocks forms the basis for grouping and normalization of histograms. The normalized group of histograms represents the block histogram and the set of these block histograms represent the descriptor. The descriptor is a concatenation of histogram of gradient directions which was obtained for every pixel within each cell of the image.

To generate an overlapped HOG descriptor, several HOG descriptors are first generated independently based on unique HOG grid. These grids may have cells of different sizes, but for simplicity we have used cells of the same size in our experiments. Thus the cells in different HOG grid may overlap with one another. Due to overlapping the HOG feature is more robust to small variations. Regardless of other parameters, overlapping significantly increases the performance for the HOG feature.

III. Excremental Results

Various tests were used to evaluate the HOG feature on the Yale face database. In the yale face database there are 165 grayscale images in GIF format of 15 individuals. For every individual, there are eleven images, one per different face expression or configuration. As shown in the image below the face expression and configuration are as follows: center light, with glasses, happy, left light, without glasses, normal, right light, sad, sleepy, surprised and wink.

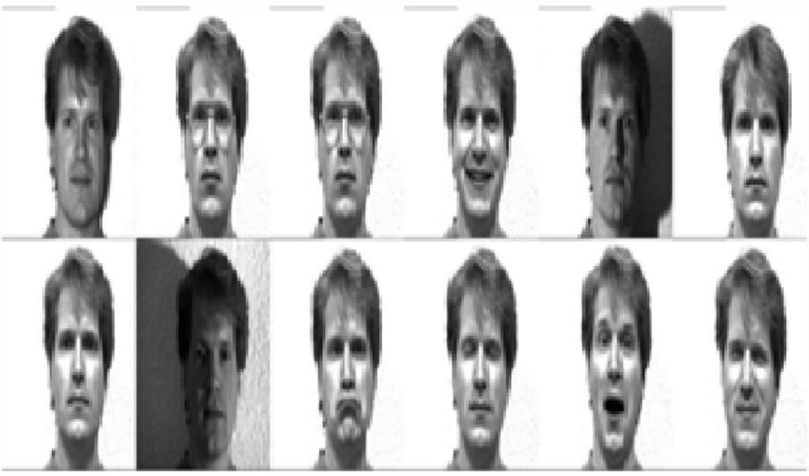

Fig.3. Images in yale face database.

Eight images of each subject were randomly selected for training. The HOG feature was compared with other existing methods like PCA, LDA, etc. The comparative results of each method are summarized in the table below.

Method	2 train	3 train	4 train
PCA	46.0	50.0	55.7
Fisherface	45.7	62.3	73.0
2DLDA	43.4	56.3	63.5
S-LDA	57.6	72.3	77.8
HOG	77.44	84.00	85.00

There are various factors that have different effects on the performance of the HOG feature for face recognition. All these factors do not affect the HOG performance the same way as in pedestrian detection.[4] Several experiments were carried out to evaluate the factors. These considered variations of scales, cell size, orientation bins, overlapping, angle representation, etc. Face recognition and also its accuracy is shown in the Fig.4.

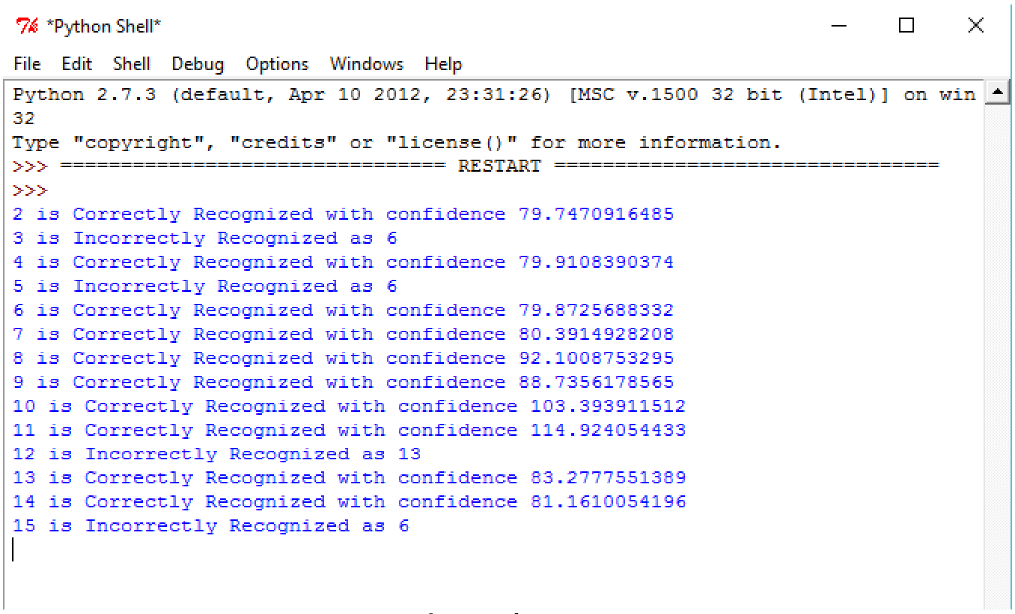

Fig.4. Final output of Face Recognition.

IV. CONCLUSION

The HOG feature is widely used for pedestrian detection but has been rarely used for face recognition. A fast computational method was developed and different factors were evaluated. We explore the use of HOG features for face recognition. The contributions are threefold:

To provide robustness to facial feature detection, we propose uniform sampling of the HOG features.

To remove redundancy in the data, improve computational efficiency and avoid over fitting. We propose to use, dimensionality reduction in the HOG representation.

We show that a decision-level combination of results using HOG features extracted from different image patch sizes significantly improves in choosing a single best patch size.

V. FUTURE SCOPE

The face recognition system used today are quite efficient and work very well for frontal mug-shot images and constant light conditions, but the current face recognition algorithm fails when there are varying light conditions under which the humans need to and are able to identify other people. The next generation system will need to recognize people in much less constrained situations.

The thought of developing a robust system which can perform efficiently under varying light conditions and in presence of noise cannot only rely on single modality. For better systems there needs to be a fusion with other modalities. The technology used has to be unobtrusive and should allow the users to act freely. Considering all the requirements and developments that need to be done, the face recognition systems seems to have the most potential for wide-spread applications.

ACKNOWLEDGEMENT

The authors would like to thank Prof. (Dr.) Vijay Gaikwad, Head of Department of Electronics Engineering-VIT Pune, for his immense support and encouragement.

References

1. Shu Chang, Ding Xiaoquing, Fang Chi, Histogram of the Oriented Gradients for Face Recognition, 2011. [2] ei Jia, Rong-Xiang Hu, Ying-Ke Lei, Yang Zhao, Jie Gui, Histogram of Oriented Lines for Palm Print Recognition, 2014.
2. Dalal N, Triggs B. Histograms of oriented gradients for human detection. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Diego, CA, USA, 2005, 1: 886-893

3. Gabor D. Theory of communication. Journal of Institute for Electrical Engineering, 1946, 93(Ⅲ): 429-457.
4. Zhao W, Chellappa R, Rosenfeld A, et al. Face recognition: A literature survey. ACM Comput. Surv., 2003, 35: 399-458.
5. Wiskott L, Fellous J M, Kruger N, et al. Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Analysis and Machine Intelligence, 1997, 19(7): 775-779.
6. Freeman W T, Roth M. Orientation histograms for hand gesture recognition. In: Intl. Workshop on Automatic Face and Gesture Recognition. IEEE Computer Society, Zurich, Switzerland, 1995: 296-301.
7. Aba M. Patil, Satish R. Kolhe, Pradeep M. Patil and Milind E Rane, "Modified Fisher Face Approach for Robust Biometric Recognition System", ICGST International Journal on Graphics, Vision and Image Processing, GVIP, Vol. 10, Issue IV. 2010, pp. 9-15. Satyajit Kautkar, Rahulkumar Koche, Tushar Keskar, Aniket Pande, Milind Rane, Gary A. Atkinson,
8. "Face Recognition Based on Ridgelet Transforms," ICEBT 2010, Procedia Computer Science 2, pp.35–43, 2010.
9. Shrinivas Kadam, Ashwini Barbadekar and Milind Rane, "Face Detection Methods: A Survey," IJRIT International Journal of Research in Information Technology, Volume 1, Issue 11, Pg. 282-289, November, 2013.

AUTHOR(S) PROFILE

Sourabh Hanamsheth, is a student in the Electronics Engineering Department, Vishwakarma Institute of Technology, Pune-411037, Affiliated to Savitribai Phule Pune University, Pune, Maharashtra. He is currently pursuing B. Tech (Electronics) Degree. His research interests are Computer Vision and Artificial Intelligence.

Milind Rane, is assistant professor of the Electronics Engineering Department, Vishwakarma Institute of Technology, Pune-411037, Affiliated to Savitribai Phule Pune University, Pune, Maharashtra. He has received his M.E in Electronics. His area of interests is Computer Vision and Image Processing.