Scientific Journal Impact Factor (SJIF) : 3.518
International Journal of Modern Trends in Engineering and Research
ISSN (ONLINE) : 2349 - 9745
ISSN (PRINT) : 2393 - 8161

OBJECTS DETECTION USING GABOR FILTER AND LABELING IN REALTIME ENVIRONMENT

Mrs .P.NALINI [1]& Mrs. M. SUJATHA[2]
[1]Department of ECE, Dadi Institute of Engineering and Technology
[2]Department of ECE, Lendi Institute of Engineering and Technology

Abstract: The greatest challenge on monitoring characters from monocular video scene is to track targets under occlusion conditions. The system tracks people from a video sequence and is robust to varied lighting conditions and complex crowded scenes. The method effectively handles small occlusions and invokes a manual tracking procedure under severe occlusions. The system also computes other parameters like velocity, average number of people crossing the region, and maximum. This paper is based on classifications of the features of an object detected using Gabor filter feature extraction techniques in image processing. The feature vector based on Gabor filters used as the input of the classifier, which is a Feed Forward Neural Network (FFNN) on a reduced feature subspace learned by an approach simpler than Principal Component Analysis (PCA). The effectiveness of the proposed method is demonstrated by the experimental results on testing large number of images and comparisons with state of the art method. Object detection and recognition has many applications in a variety of fields such as security systems, video conferencing and identification.
Keywords: *Object detection, Gabor Filter, Feed Forward Neural Network (FFNN) classifier, Multilayer perceprotn, Image Processing, Occlusion.*

I. INTRODUCTION

Visual-based object detection is critical to automatically monitor object activities in video sequences. To view and detect object actions from a monocular scene, object occlusions usually incur detection errors due to objects in crowded environments. In this work, an object detection scheme is proposed to overcome the occlusion effects and then to improve the accuracy of labeling the characters. Crowd analysis finds its usage in large number of applications like crowd management, public space design, virtual environments, visual surveillance and intelligent environments. Many computer vision algorithms have addressed the research problems related to the scenarios involving large crowds of people. These computer vision methods focus on extracting quantitative features and in object detection and tagging. Object detection and tagging has been the main focus of visual surveillance of human activity. Occlusion caused by the high clutter of people in the crowd scene is the major challenge for crowd detection and labeling. The object detection algorithms have to overcome this problem to effectively track people. The various light intensities in the video and camera motions also make the tracking a challenging problem in the field of computer vision. In this paper we present a simplified object tracking method to effectively track people in a crowd scene. Object detection and counting is an active area of research spanning several disciplines such as image processing, pattern recognition and computer vision. Object detection and recognition are preliminary steps to a wide range of applications such as personal identity verification, video-surveillance, lip tracking, facial expression extraction, gender classification, advanced human and computer interaction. Most methods are based on neural network approaches, feature extraction, Markov chain, skin color, and others are based on template matching. Pattern localization and classification is the step, which is used to classify object and non-object patterns. Many systems dealing with object classification are based on skin color. In this paper we are interested by the design of an ANN algorithm in order to achieve image classification. This paper is organized as follows: In section 2, we give an overview over classification for object detection. Description of our model and training algorithm is discussed in Section 3. Section 4 deals with the Gabor Filter.

II. CLASSIFICATIONS FOR OBJECT DETECTION

While numerous methods have been proposed to detect Object in a single image of intensity or color images. A related and important problem is how to evaluate the performance of the proposed detection methods [1]. Many recent Object detection papers compare the performance of several methods, usually in terms of detection and false alarm rates. It is also worth noticing that many metrics have been adopted to evaluate algorithms, such as learning time, execution time, the number of samples required in training, and the ratio between detection rates and false alarms. In general, detectors can make two types of errors: *false negatives* in which Objects are missed resulting inlow detection rates and *false positives* in which an image is declared to be Object.

$$\text{False negative} = \frac{\text{Number of Missed Falses}}{\text{Total Number of Actual Faces}}$$

$$\text{False Positive} = \frac{\text{Number of Incorrect Detected Faces}}{\text{Total Number of Actual Faces}}$$

Object detection can be viewed as two-class Recognition problem in which an image region is classified as being a “Object” or “non-Object”. Consequently, Object detection is one of the few attempts to recognize from images a class of objects for which there is a great deal of within-class variability. Object detection also provides interesting challenges to the underlying pattern classification and learning techniques. The class of Object and non Object image are decidedly characterized by multimodal distribution function and effective decision boundaries are likely to be nonlinear in the image space. Pattern localization and classification are CPU time intensive being normally implemented in software, however with lower performance than custom implementations. Custom implementation in hardware allows real-time processing, having higher cost and time to- market than software implementation. Some works [2,3,4] uses ANN for classification, and the system is implemented in software, resulting in a good performance (10 sec for localization and classification). A similar work is presented in [5], aiming to object localization and classification. We are interested in the implementation of an ANN algorithm& design of a Gabor filter in order to provide better image classification. The MLP *(Multi-layer Perceptron)* algorithm is used to classify Object and non-Object patterns before the recognition step.

2.1 NEURAL NETWORK

A neural network is an information-processing system that has been developed as generalizations of mathematical models matching human cognition. They are composed of a large number of highly-interconnected processing units (neurons) that work together to perform a specific task. According to Haykin, a neural network is a massively parallel distributed processor that has a natural prosperity for storing experimental knowledge. It resembles the brain in two respects:

- Knowledge is acquired by the network through a learning process;
- Inter-connected connection strengths known as synaptic weights are used to store the knowledge;

- Each neuron has an internal state called its threshold or activation function (or transfer function) used for classifying vectors. Neural classification generally comprises of four steps:
- Pre-processing, e.g., atmospheric correction, noise suppression, band rationing, Principal Component Analysis etc;
- Training - selection of the particular features which best describe the pattern;
- Decision - choice of suitable method for comparing the image patterns with the target patterns;
- Assessing the accuracy of the classification.

2.2 MULTI-LAYERS PERCEPTRON

The MLP neural network has feed forward architecture within input layer, a hidden layer, and an output layer. The input layer of this network has N units for an N dimensional input vector. The input units are fully connected to the I hidden layer units, which are in turn, connected to the J output layers units, where J is the number of output classes. A Multi-Layers Perceptron (MLP) is a particular of artificial neural network. We will assume that we have access to a training dataset of l pairs (xi, yi) where x i is a vector containing the pattern, while yi is the class of the corresponding pattern. In our case a 2-class task, yi can be coded 1 and -1.

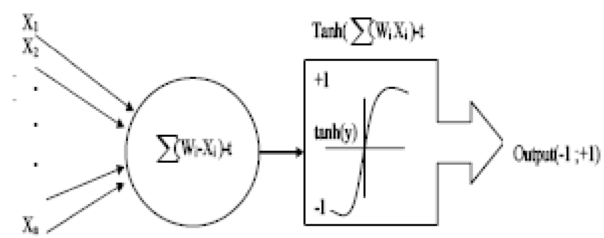

Fig.1 The neuron of supervised training

We are designing a feed forward neural network with one hundred neurons in the hidden layer and one neuron in the output layer which prepares images for training phase. All data form both “Object” and “non-Object” folders will be gathered in a large cell array. Each column will represent the features of an image, which could be a Object, or not. Rows are as follows:

Row 1: File name
Row 2: Desired output of the network corresponded to the feature vector.
Row 3: Prepared vector for the training phase

We will adjust the histogram of the image for better contrast. Then the image will be convolved with Gabor filters by multiplying the image by Gabor filters in frequency domain. To save time they have been saved in frequency domain before Features is a cell array contains the result of the convolution of the image with each of the forty Gabor filters. These matrices have been concated to form a bif 135x144 matrix of complex numbers. We only need the magnitude of the result. That is why “abs “ is used.135x144 has 10,400 pixels. It means that the input vector of the network would have 19,400 values, which means a large amount of computation. So we have reduced the matrix size to one-third of its original size by deleting some rows and columns. Deleting is not the best way but it save more time compare to other methods like PCA.We should optimize this function as possible as we can .First training the neural network and then it will return the trained network. The examples were taken from the Internet database. The MLP will be trained on 200 Object and 300 non-objects.

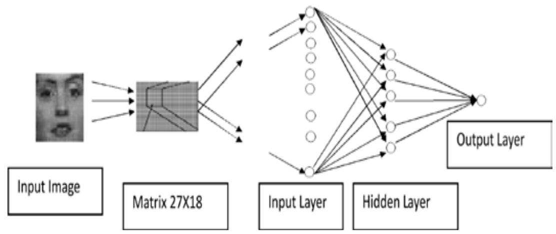

Fig.2 (a) Architecture of proposed System

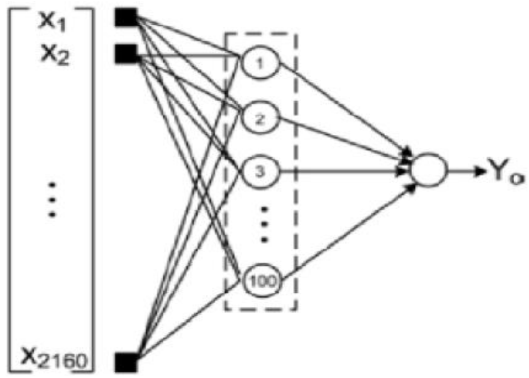

Fig.2 (a) Architecture of proposed System

2.2TRAINING ALGORITHM

The MLP with the training algorithm of feed propagation is universal mappers, which can in theory, approximate any continuous decision region arbitrarily well. Yet the convergence of feed forward algorithms is still an open problem. It is well known that the time cost of feed forward training often exhibits a remarkable variability. It has been demonstrated that, in most cases, rapid restart method can prominently suppress the heavy-tailed nature of training instances and improve efficiency of computation. Multi-Layer Perceptron (MLP) with a feed forward learning algorithms was chosen for the proposed system because of its simplicity and its capability in supervised pattern matching. It has been successfully applied to many pattern classification problems [9]. Our problem has been considered to be suitable with the supervised rule since the pairs of input-output are available. For training the network, we used the classical feed forward algorithm. An example is picked from the training set, the output is computed.

III. 3.GABOR FILTER

Features based on Gabor filters have been used in image processing due to their powerful properties. Gabor kernels are characterized as localized, orientation selective, and frequency selective. A family of Gabor kernel is the product of a Gaussian envelope and a plane wave. A 2D Gabor filter is expressed as a Gaussian modulated sinusoid in the spatial domain and as shifted Gaussian in the frequency domain. The Gabor wavelet representation of images allows description of spatial frequency structure in the image while preserving information about spatial relations. The Equation of Gabor is customized as

$$G(x, y; \varphi, \sigma, \psi) = exp\ (-(x2 + y2) / 2\ \sigma) * cost\ (w * y * cos\ (\varphi) + w * y * sin\ (\varphi) + \psi)$$

Filter of Gabor can be drawn as follows. We can vary φ (i.e. angle); ψ (i.e. psi) of the filter to derive at these different filters as below.

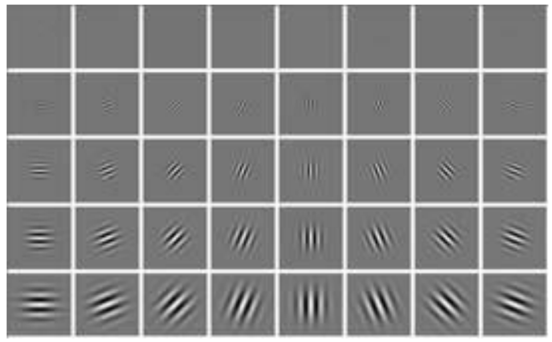

Fig 3 Forty bank filter of Gabor

Using these filters images we can convolve it with the Object image to produce the Gabor filtered images. From these images we can extract the features of the object like eye brows, eyes, mouth, etc., and a feature vector is generated and then given to the classifier. Different people use different ways to form the feature vector for training the classifier. Some of them even use whole image as a feature vector and perform classification which needs high computation. So here feature vector is taken from important values of the image from each filter Energy, mean and standard deviation forming a 32 value feature vector for every image.

IV. EXPRIMENTAL RESULTS

In this section, the implementation results of the detection algorithms based on neural algorithm are studied. Our algorithm can detect approximately 90% of Objects in a set of 200 test images, with an acceptable number of false detections. Depending on the application, the system can be made more or less conservative by varying the arbitration heuristics or thresholds used. The system has been tested on a wide variety of images, with many objects and unconstrained backgrounds. A fast version of the system can process a 320x240 pixel image in 1 to 2 seconds on an I3 Core with 2 Gigabyte Ram. There are a number of directions for future work. The main limitation of the current system is that it only detects upright objects looking at the camera. Separate versions of the system could be trained for each head orientation, and the results could be combined using arbitration methods similar to those presented here. Preliminary work in this area indicates that detecting profiles views of objects is more difficult than detecting frontal views, because they have fewer stable features and because the input window will contain more background pixels. We have also applied the same algorithm for the detection of car tires and human eyes, although more work is needed. Even within the domain of detecting frontal views of objects, more work remains. When an image sequence is available, temporal coherence can focus attention on particular portions of the images. As a object moves about, its location in one frame is a strong predictor of its location in next frame. Standard tracking methods, as well as expectation-based methods, can be applied to focus the detector’s attention. Other methods of improving system performance include obtaining more positive examples for training, or applying more sophisticated image preprocessing and normalization techniques. One application of this work is in the area of media technology. Every year, improved technology provides cheaper and more efficient ways of storing and retrieving visual information. However, automatic high-level classification of the information content is very limited; this is a bottleneck that prevents media technology from reaching its full potential. Systems utilizing the detector described above allow a user to make requests of the form “Show me the people who appear in this video" or "Which images on the World Wide Web contain objects?" and to have their queries answered automatically. Some sample images are shown in Fig. 4.1 and Fig. 4.2.

Fig.4.1. some sample images to test the proposed structure

Fig. 4.2 some sample images failed to work with the proposed algorithm

V. CONCLUSION AND FUTURE WORK

In comparison with other methods, all Gabor wavelet Object based method is better than the base-line method. Down sampled Gabor wavelets transform of object images as features for Object detection in subspace approach is superior to pixel value.Humans are able to identify reliably a large number of objects and neuroscientists are interested in understanding the perceptual and cognitive mechanisms at the base of the object detection process. Those researches illuminate computer vision scientists' studies. Although designers of object detection algorithms and systems are aware of relevant psychophysics and neuropsychological studies, they also should be prudent in using only those that are applicable or relevant from a practical/implementation point of view.Although using 2- D Gabor wavelet transform seems to be well suited to the problem, graph matching makes algorithm bulky. Moreover, as the local information is extracted from the nodes of a predefined graph, some details on a object, which are the special characteristics of that object and could be very useful in detection task, might be lost. In this paper, a new approach to object detection with Gabor wavelets& feed forward neural network is presented. The method uses Gabor wavelet transform & feed forward neural network for both finding feature points and extracting feature vectors. From the experimental results, it is seen that proposed method achieves better results compared to the graph matching and Eigen face methods, which are known to be the most successive algorithms. Although the proposed method shows some resemblance to graph matching algorithm, in our approach, the location of feature points also contains information about the object. Feature points are obtained from the special characteristics of each individual object automatically, instead of fitting a graph that is constructed from the general object idea. In the proposed algorithm, since the facial features are compared locally, instead of using a general structure, it allows us to make a decision from the parts of the object. For example, when there are sunglasses, the algorithm compares objects in terms of mouth, nose and any other features rather than eyes. Moreover, having a simple matching procedure and low computational cost proposed method is faster than elastic graph matching methods. Proposed method is also robust to illumination changes as a property of Gabor wavelets, which is the main problem with the Eigen face approaches. Feature points, found from Gabor responses of the object image, can give small deviations between different conditions (expression, illumination, having glasses or not, rotation, etc.), for the same individual. Therefore, an exact measurement of corresponding distances is not possible unlike the geometrical feature based methods. Moreover, due to automatic feature detection, features represented by those points are not explicitly known, whether they belong to an eye or a mouth, etc. Giving information about the match of the overall facial structure, the locations of feature points are very important. However using such a topology cost amplifies the small deviations of the locations of feature points that are not a measure of match. Although detection performance of the proposed method is satisfactory by any means, it can further be improved with some small modifications and/or additional preprocessing of object images. Such improvements can be summarized as;

1) Since feature points are found from the responses of image to Gabor filters separately, a set of weights can be assigned to these feature points by counting the total times of a feature point occurs at those responses.
2) A motion estimation stage using feature points followed by an affined transformation could be applied to minimize rotation effects. This process will not create much computational complexity since we already have feature vectors for recognition. By the help of this step object images would be aligned.
3) When there is a video sequence as the input to the system, a frame giving the "most frontal" pose of a person should be selected to increase the performance of object detection algorithm. This could be realized by examining the distances between the main facial features which can be determined as the locations that the feature points become dense. While trying to maximize those distances, for example distance between two eyes, existing frame that has the closest pose to the frontal will be found. Although there is still only one frontal object per each individual in the gallery, information provided by a video sequence that includes the object to be detected would be efficiently used by this step.
4) As it is mentioned in problem definition, a object detection algorithm is supposed to be done beforehand. A robust and successive object detection step will increase the detection performance.
5) In order to further speed up the algorithm, number of Gabor filters could be decreased with an acceptable level of decrease in detection performance. It must be noted that performance of detection systems is highly application dependent and suggestions for improvements on the proposed algorithm must be directed to a specific purpose of the object detection.

REFERENCES

[1] Yang M.H., Kriegman D.J., and Ahuja N., *"Detecting Faces in Images: A Survey"*, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.24, No.1, January 2002.
[2] Rowley H. A., Baluja S. , Kanade T., *"Neural Network- Based Face Detection"*, IEEE Trans. On Pattern Analysis and Machine Intelligence, vol.20, No. 1, Page(s). 39-51, 1998.
[3] ZhenQiu Z., Long Z., Li S.Z., Jiang Z.H., *"Real-time multi-view face detection"*, Proceeding of the Fifth IEEE International Conference on automatic Face and Gesture Recognition, pp: 142-147, 20-21 May 2002.
[4] Gavrila, D.M; Philomin, V., *"Real-Time Object detection for Smart Vehicles"*. International Conference on Computer Vision (ICCV99). Vol. 1. Corfu, Greece, 20-25 September 1999.
[5] Molz R.F., Engel P.M., Moraes F.G., Torres L., Robert M., *"System Prototyping dedicated to Neural Network Real- Time Image Processing"*, ACM/SIGDA ninth international Symposium On Field Programmable Gate Arrays (FPGA 2001).
[6] Theocharides T., Link G., Narayanan V., Irwin M.J., *"Embedded Hardware Face Detection"*, 17th International Conference on VLSI Design, Mumbai, India, January 5-9, 2004.

[7] Yang F. and Paindavoine M., *"Prefiltering for pattern Recognition Using Wavelet Transform and Face Recognition"*, The 16th International Conference on Microelectronics, Tunisia 2004.
[8] Yang F. and Paindavoine M., *"Prefiltering for pattern Recognition Using Wavelet Transform and Neural Networks"*, Advances in imaging and Electron physics, Vol. 127, 2003.
[9] Li X. and Areibi S., *"A Hardware/Software co-design approach for Face Recognition"*, The 16th International Conference on Microelectronics, Tunisia 2004.
[10] Gajski D., Dutt v, Wu A., *"High-Level Synthesis: Introduction to Chip and System Design"*, Kluwer Academic Publishers, Boston, 1992.
[11] Kanade T., *"Picture processing by computer complex and recognition of human faces"*, Technical report, Kyoto University, Dept. of Information Science, 1973.
[12] Lin S., Kung S. , and Lin L. , *"Face Recognition / Detection by Probabilistic Decision-Based Neural Network"*, IEEE Transactions on Neural Networks, Vol.8, pp.114-132, 1997.
[13] Brunelli R., Poggio T., *"Face Recognition: Features vs. Templates"*, IEEE Trans. on PAMI, Vol. 12, No. 1, Jan.
[14] R. Ramamoorthi and P. Hanrahan "On the relationship between radiance and irradiance: Determining the illumination from images of a convex Lambertian object", J. Opt. Soc. Amer., vol. 18, no. 10, pp.2448 -2459 2001
[15] A. Shashua and T. Riklin-Raviv "The quotient image: Class-based re-rendering and recognition with varying illuminations", IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 2, pp.129 -139 2001 .
[16] H. Wang , S. Li and Y. Wang "Generalized quotient image", Proc. IEEE. Comput. Vis. Pattern Recognit., vol. 2, pp.498 -505 2004 .
[17] Q. Li , W. Yin and Z. Deng "Image-based face illumination transferring using logarithmic total variation models", Int. J. Comput. Graph., vol. 26, no. 1, pp.41 -49 2009 .
[18] E. H. Land "The Retinex theory of color vision", Sci. Amer., vol. 237, no. 6, pp.108 -128 1977.
[19] D. J. Jobson , Z. Rahman and G. A. Woodell "Properties and performance of a center/surround Retinex", IEEE Trans. Image Process., vol. 6, no. 3, pp.451 -462 1997 .
[20] R. Gross and V. Brajovie "An image preprocessing algorithm for illumination invariant face recognition", Proc.