A Cognitive Approach to Face Detection System for Real-Life Images

S.Fathima Shuruma,
Student of Master of Engineering,
Kamaraj College of Engineering and Technology,
Virudhunagar.

A.Geetha,
Assistant Professor,
Kamaraj College of Engineering and Technology,
Virudhunagar.

Abstract:
This paper presents a detailed idea about the face detection technique which reduces false positive rate and improves the rate of detection. Face detection remains as one of the challenging problems in the field of image analysis and computer vision. The color image is converted into a grayscale image. Based on the spatial localities and orientation characteristics of the Gabor filter, we extract the facial feature from the facial image by appropriate filter design. The input of feed-forward neural network classifier is the feature vector obtained by the Gabor filter. The convolution is done by multiplying the image with the Gabor filter in the frequency domain and is saved in a cell array in order to save time. Thus the input of the network will have larger values and is optimized by reducing the matrix size to one-third by deleting appropriate number of rows and columns. The results obtained in the output layer in testing with real-life images shows that the false-positive rate is reduced by 97%.

Index Terms: Gabor filter, Feed-Forward Neural Network, computer vision, color image processing.

1.INTRODUCTION
Face detection is the technology in computer vision that helps in detecting faces and thereby its features and recognize them optimally irrespective of several variations in faces and its associated background. The main motivation behind this approach is to find best methodology to solve problems that emerge when there are multiple variation in the input image like facial sizes, occlusion and illumination. Many novel methods have been proposed like template-matching methods and appearance based methods. The template-matching methods are used for face localization and detection by computing the correlation of an input image to a standard face pattern. Whereas the appearance based methods are used for face detection with Eigen faces and neural network.

Face detection can be done by fine tuning the Gabor filter parameter, and recognition is done by using the Gabor wavelets [1][5]. In [6], recent advances in face detection techniques are discussed which clearly shows that the available face detection techniques are time-consuming, and found to have more number of false positives. So, in our paper, Gabor filter is used for feature extraction and Neural Network is chosen as Classifier. The existing Viola-Jones face detector [7], also found to be more effective and better in performance but time-consuming. The false positive rate is found to be the ratio of number of incorrect detected faces to the total number of actually detected faces. The distributed class of face or non-face is characterized by multimodal distribution function and effective decision boundaries. In

[26], object localization is focused and this paper uses the technique of object localization and classification.

The image pre-processing is done based on the well-suited technique based on the input image pixel intensities and occlusion. Histogram equalization is done inorder to equally distribute the brightness levels over the whole brightness scale whereas median filtering is used for removing salt and pepper noise.

2.FEATURE EXTRACTION
The simpler technique to perform filtering operation is by convolution in spatial domain and can be improved by utilizing the separability of Gabor filter, their symmetry, and their wavelet characteristics for reducing the number of arithmetic operations needed. And for general improvement we go for filtering in the frequency domain. Thus the Gabor features are based on the Gabor filter responses of the given input image. The convolution of the Fourier Image F(I) and a Fourier Gabor kernel F(ψ,μ,ν(x,y)) is called as Fourier Gabor Feature and is defined by the equation (1) as follows,

\[ O_{μν}(x,y) = F(I(x,y)) \ast F(\Psi_{μν}(x,y)) \] (1)

As the response \( O_{μν}(x,y) \) to each Fourier Gabor kernel is a complex function with a real part: \( \text{Real}(O_{μν}(x,y)) \) and an imaginary part: \( \text{Imag}(O_{μν}(x,y)) \). We use its real \( \text{Real}(O_{μν}(x,y)) \) to represent Fourier Gabor features. The complete set of Gabor filter representation of an Image I(x,y) is of size 16x16 by 8 orientations and is defined by equation (2).

\[ G(I) = \{O_{μν}(x,y); μ \in \{0,1,...,7\}, ν \in \{0,1,...,4\}\} \] (2)

To extract useful features of an image a set of Gabor filters with different orientations and frequencies are used. The real and the imaginary components of a filter may be formed into a complex number representing orthogonal directions and is given by equation (3).

\[ g(x,y;λ,θ;ψ,σ,γ) = \exp \left[ -\frac{x'^2 + y'^2}{2σ^2} \right] \exp \left[ i\left(\frac{2πx'}{λ} + ψ\right) \right] \] (3)

Where

\[ x' = x\cosθ + y\sinθ \] (4)

\[ y' = -x\sinθ + y\cosθ \] (5)
3. FACE DETECTION

With the detected facial features via Gabor filter, the face or non-face is classified based on the Neural Network Classifier.

3.1 Neural Network Classifier

The objective here is to design a feed-forward neural network with an input layer for acquiring the features of an input image and a hidden layer consisting of hundreds of neurons, and an output layer neurons. The input neuron is connected to a set of neurons in the hidden layer and the result obtained by the hidden layer is then sent to output layer neurons. The input of feed-forward neural network is a feature vector obtained from the Gabor filter. The input image is of size 27x18. The convolution is done by multiplying the input image with the Gabor filter in the frequency domain. Thus the images are prepared for the training phase. All the pictorial data obtained from face and non-face folders are gathered and stored in a large cell array. Optimization of larger input is done by deleting some of the unused rows and columns.

If the feature vectors are obtained for the color image, then numerical value of color at each pixel is calculated. And for each pixel one node is allotted for computation. Now, this information in the input layer is sent as input to the hidden layers of multiple neurons. The resulting output obtained in the output layer comprises of binary values. The value of “1” at the output node means the face of the person is detected. An algorithm is used in order to improve the performance of the network called Scaled Conjugate Gradient Algorithm for fast supervised learning.

3.2 Optimization Strategy

The method is to minimize functions inorder to reduce the computational complexity. The minimization is a local iterative process in which an approximation to the function in the neighbourhood of the current point in the weight space is minimized. And this strategy is illustrated in the pseudocode given below.

**Optimization Strategy Pseudocode**

Choose initial weight vector $w_1$ and set $k = 1$.
Determine a search direction $p_k$ and a step size $\alpha_k$ so that

$$E(w_k + \alpha_k p_k) < E(w_k)$$

Update vector:

$$w_{k+1} = w_k + \alpha_k p_k$$

If $E'(w_k) \neq 0$ then set $k = k+1$ and go to 2
else return $w_{k+1}$ as the desired minimum.

**Step 1:** Choose the weight vector $w_i$ and scalars $> 0$.

$$\lambda_2 > 0, \lambda_1 > 0$$

Set $p_1 = r_1 = -E'(w_1), k=1$ and success = true.

**Step 2:** If success = true then calculate the second order information:

$$\sigma_k = \frac{1}{|p_k|^2}$$

$$s_k = \frac{E'(w_k + \sigma_k p_k) + E'(w_k)}{\sigma_k}$$

$$\delta_k = p_k^T s_k$$

**Step 3:** Scale $s_k$:

$$s_k = s_k + (\lambda_k - \bar{\lambda}_k)p_k$$

**Step 4:** If $\delta_k \leq 0$, then make the Hessian matrix positive definite:

$$s_k = s_k + (\lambda_k - \frac{2}{|p_k|^2})\delta_k$$

$$\bar{\lambda}_k = 2(\lambda_k - \frac{\delta_k}{|p_k|^2})$$

$$\delta_k = -\delta_k + \frac{1}{\lambda_k} |p_k|^2$$

where $\lambda_k = \bar{\lambda}_k$

**Step 5:** Calculate the step size:

$$\mu_k = p_k^T r_k$$

$$\alpha_k = \frac{\mu_k}{s_k}$$

**Step 6:** Calculate the comparison parameter:

$$\Delta_k = \frac{2\delta_k[E(w_k) - E(w_k + \alpha_k p_k)]}{\mu_k^2}$$

**Step 7:** If $\Delta_k \geq 0$, then a successful reduction in error can be made:

$$w_{k+1} = w_k + \alpha_k p_k$$

$$r_{k+1} = -E'(w_{k+1})$$

If $\lambda_k = 0$, then success = true.

**Step 8:** If $k \mod N = 0$, then restart algorithm with $p_{k+1} = r_{k+1}$

Else create new conjugate direction:

$$\beta_k = \frac{|r_{k+1}|^2 - r_{k+1}^T r_k}{\mu_k}$$

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

**Step 9:** If $\Delta_k \geq 0.75$, then reduce the scale parameter

$$\lambda_k = \frac{1}{2} \lambda_k$$

Else a reduction in error is not possible. $\bar{\lambda}_k = \lambda_k$, success = false.
Step 10: If $\Delta_k < 0.25$, then increase the scale parameter: $\lambda_k = 4\lambda_k$

Step 11: If the steepest descent direction $r_k \neq 0$, then set $k = k+1$ and go to 2.
Else terminate and return $w_{k+1}$ as the desired minimum.

For each iteration there is one call of $E(w)$ and two calls of $E'(w)$, which gives a calculation complexity per iteration of $O(3N^2)$. When the algorithm is implemented this complexity can be reduced to $O(2N^2)$, because the calculation of $O(N^2)$ can be built into one of the calculations of $E'(w)$. In comparison with backpropagation, scaled conjugate gradient algorithm involves twice as much computation per iteration, since backpropagation algorithm has a computational complexity of $O(N^2)$ per iteration.

**4. EXPERIMENTAL RESULTS**

By varying the arbitration heuristics or thresholds, the system can be made more conservative. In order to focus the detector’s attention some standard tracking methods can also be used. Also by applying more suitable normalization techniques, and separate network to recognize the hurdles that occur in detection like various illumination, facial orientations, faces of different races and so on.

**Figure 1 Block diagram of Face Detection Technique**

**Figure 2. An illustration of Facial Feature Extraction using Gabor filter**

**Table 1 Performance measure of face detection rate of images of different characteristics**

<table>
<thead>
<tr>
<th>Type Of Input Image</th>
<th>Face Detection Rate(%)</th>
<th>Minimization of false positive rate(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image containing a face with spectacles</td>
<td>100</td>
<td>99.8</td>
</tr>
<tr>
<td>Image containing faces of multiple illuminations</td>
<td>98</td>
<td>96.8</td>
</tr>
<tr>
<td>Image with multiple faces in different orientations</td>
<td>94</td>
<td>96.4</td>
</tr>
<tr>
<td>Image with faces of multiple races</td>
<td>98</td>
<td>97.5</td>
</tr>
</tbody>
</table>

**Figure 3. Result of Face Detection Technique**
5. CONCLUSION
In this paper, a cognitive approach to face detection is proposed in order to reduce false-positive rates in processing real-life images. Processing various real-life images clearly prove that Gabor feature extraction and Neural Networks approach is faster but have low accuracy when comes for images in uncontrolled environment. There arise a problem in finding the factor of luminance and also it is found to be less efficient when finding faces of multiple races. By minimizing false positive rate and by improving face detection rate, we can apply this technique for Gender Classification and for improving the robustness of image in uncontrolled Environment.

6. REFERENCES
[8] Stan Z. Li, Anil K. Jain (Eds.), ”Handbook of Face Recognition”Springer


