Study of Efficient Algorithms for Face Detection: Cascaded and Parallel GSLDA

Jay Prakash Maurya¹, Abhilasha R Wakde²
¹Assistant Professor Department of Computer Science and Engineering, IES Institute Of Technology And Management, Bhopal MP (India)
²Department of computer science and Engineering, IES Institute Of Technology And Management, Bhopal MP (India)

Abstract—Face detection is a computer technology that determines the locations and sizes of human faces in arbitrary (digital) images. It detects facial features and ignores anything else, such as buildings, trees and bodies. Face detection can be regarded as a more general case of face localization. In face localization, the task is to find the locations and sizes of a known number of faces (usually one). The work proposed in this paper is based on advance technique for face detection from image using (GSLDA) i.e. Greedy Sparse Linier Discriminate Analysis. With the help of this technique images can be processed extremely rapid and high detection rates can be achieved as compare to AdaBoost[1]. We have managed to organize classifiers in parallel and in cascaded manner, with critical order of grid that detects faces rapidly. The paper has research work for face detection has following steps: first is the method of feature selection the method of introducing “Integral image” by which new image is represented which allows the features used by our detector to be computed very quickly for each classifier. The second step is learning algorithm, which is based on GSLDA which selects weak classifiers based upon the maximum class separation criterion [2, 3].

Keywords—Integral image, greedy sparse linear discriminate analysis (GSLDA), feature selection, cascade classifier.

I. INTRODUCTION

Number of vision based applications has face detection as an important task. This task is made a first step in applications such as intelligent video surveillance, content based image retrieval, face and activity recognition. Nowadays face detection is a challenging scenario because there are large variations in visual appearances, object poses, illumination, camera motion, etc. All these issues have made the problem very challenging from a machine vision perspective. Object detection is a extremely imbalanced task for classification. Face detection can be explained in three steps as preprocessing, feature extraction and designing of classifier as given in Figure 1. Feature extraction methods are Eigen face, PCA, ICA, LDA[11]. Principal component analysis (Karhunen-Loeve or Hotelling transform). PCA belongs to linear transforms based on the statistical techniques.

The method describes a strong tool for analysis of data and pattern recognition, as a technique for data dimension reduction or their decor relation and data compression as well, but it consists of a problem of small sample size. ICA is known as a statistical method that linearly transforms a multidimensional vector data into components that are statistically as independent from each other as possible[10]. A standard natural image consists of many more negative background patterns than object patterns. That means, if one wants to achieve a high detection rate, together with a low false detection rate, one need to design a specific and sensitive classifier that takes the imbalanced data distribution into consideration [4].

Linear Discriminate Analysis (LDA) is a dimension reduction method which finds transformation that maximizes the class separation. Cascading of classifier is faster but it requires time for training for classifier [4]. The simple, rich image representative features used are Haar basis functions which have been used by Papa Georgiou et al[5].

![Figure 1: Principle of face detection system](image-url)
For the rapid computation of the features at many levels we introduce the concept of integral image representation. The integral image can be computed from an image only with a few operations per pixel. Once computed, any one of these Haar-like features can be computed at any scale or location in constant time. The another contribution of the research work is a process for developing a classifier with selection of a small number of features using LDA. The weak classifier is constrained so that each weak classifier returned can depend on only a single feature. The third contribution is of the method that combines successively simple to complex classifiers in a parallel and cascaded structure which increases the speed of the detector by paying attention on promising regions of the image. In the domain of face detection it is possible to achieve less than 1 per cent false negatives and 40 per cent false positives with a classifier developed from five Haar-like features.

This paper is organized as follows: Section II, introduce feature types and evaluation. Section III explains the proposed cascading algorithm for training weak classifier. Section IV explains the parallel algorithm and Section V gives experiments and discussions.

II. FEATURE TYPES AND EVALUATION

The feature employed by the detection framework universally involves the sums of image pixels within rectangular areas. As such, they bear some resemblance to Harr basic function, which have been used previously in the realm of image-based object detection. However, since the features used by Viola and Jones all rely on more than one rectangular area, they are generally more complex. The figure at right illustrates the four different types of features used in the framework. The value of any given feature is always simply the sum of the pixels within clear rectangles subtracted from the sum of the pixels within shaded rectangles. As to be expected, rectangular features of this sort are rather primitive when compared to alternatives such as steerable filters. Although they are sensitive to vertical and horizontal features, their feedback is considerably coarser. However, with the use of an image representation called the integral image, rectangular features can be evaluated in constant time, which gives them a considerable speed advantage over their more sophisticated relatives. Because each rectangular area in a feature is always adjacent to at least one other rectangle, it follows that any two-rectangle feature can be computed in six array references, any three-rectangle feature in eight, and any four-rectangle feature in just nine.

A. Learning algorithm

There are two algorithms that are used to study and simplify the basic face detection systems and they can be classified as: Cascading and parallel LDA. The Figure III shows flow diagram of face detection using of Haar like feature calculation.

B. Cascaded Architecture

The evaluation of the strong classifiers generated by the learning process can be done quickly, but it isn’t fast enough to run in real-time. For this reason, the strong classifiers are arranged in a cascade in order of complexity, where each successive classifier is trained only on those selected samples which pass through the preceding classifiers. If at any stage in the cascade a classifier rejects the sub-window under inspection, no further processing is performed and continue on searching the next sub-window (see figure at right). The cascade therefore has the form of a degenerate tree. In the case of faces, the first classifier in the cascade – called the attentional operator – uses only two features to achieve a false negative rate of approximately 0% and a false positive rate of 40%.
To reduce to roughly half the number of times the entire cascade is evaluated is the work of this singly set classifier. The cascade architecture has interesting implications for the performance of the individual classifiers. Since each classifier’s activation is depended fully on the behavior of its predecessor, the false positive rate for entire cascade can be summarized as:

\[ F = \prod_{i=1}^{K} f_i. \]

Same way, the detection rate can be given as:

\[ D = \prod_{i=1}^{K} d_i. \]

Therefore, in order to match the false positive rates which are typically achieved by other detectors, each classifier has to away with having surprisingly poor performance. For example, to achieve a false positive rate of each classifier for a32-stage cascade needs only to achieve a false positive rate of about 65%. However, at the same time each classifier has to be exceptionally capable if it has to achieve detection rates adequately. For example, in order to achieve a detection rate of about 90%, it is necessary for each classifier in the aforementioned cascade to achieve a detection rate of approximately 99.7%.

The cascaded scenario permits most of the non-face patches to be rejected instantly before reaching to the final node, hence results in fast performance. When a test image patch passes tests in all nodes then and only then it is reported as a face. In this way, most non-face patches are rejected at the early nodes.
IV. THE PARALLEL GSLDA ALGORITHM

Algorithm 1: The simple training process for building a parallel of GSLDA face detector.

Input:
A positive training set and a negative training set;
A set of Haar-like rectangle features h1, h2, h3;

1 For each feature does
2 Train a weak learner (e.g. decision stump parameterized by a threshold) that yields the maximum class separation to the set of selected weak classifiers using LDA;

Output: A set of weak learners [1, T] that best separates the training set.

Simple types of classifiers are used to reject the majority of sub-window before more complex classifiers are called upon to achieve false positive rates. The overall problem of the detection process is that of a degenerate decision tree, what we call “parallel cascade”.

Figure VI: Schematic depiction of a parallel detection cascade.

A series of classifiers are applied to every sub-window. The preceding classifiers eliminate many number of negative sub-windows with very less processing. The first layer contains four classifier of the same Haar features. The sub-windows are forwarded in a sequential manner i.e., sub-window1, 5, 9, etc. to layer first node 1. Additional negatives are eliminated by subsequent layers but require additional computation. The number of sub-windows has been reduced radically after few stages of processing. Further processing can take any form such as additional stages of the cascade (as in our detection system) or an alternative detection system. A positive result from the first classifier stage triggers the initiation of a second classifier which also has been adjusted in such a way to achieve high detection rates. A positive result from the second classifier stage triggers a third classifier, and so on.
A negative outcome at any point leads to the immediate rejection of the sub-window. Stages in the cascade are developed by making the classifiers trained using LDA and then adjusting the threshold value to minimize false negatives rates.

In the further evaluation, detections are made together in a very simple manner. The set of detections are first partitioned into the disjoint subsets. Two detections are put into the same subset if bounding regions of them overlap. Each partition yields a single final detection. The average of the corners of all detections in the set is the corners of the final bounding region. Detections are considered to be true or false positives based on the area of overlap with ground truth bounding boxes. Considered to be a accurate detection, there must be at least a 50 per cent overlap between the predicted bounding box and ground truth bounding box. If multiple detections occur in the same face in an image then they are considered false positive detections.

V. Observations and Discussions

This section of paper consists of the data sets and the increased performance of using sets as well as ROC of face detection given in Figure 5. We have made a test on our face detectors on the low resolution faces dataset, MIT+CMU frontal face test set. The overall set contains 130 images with 507 frontal faces. In this experiment, we have made the scaling factor to 1.2 and window shifting step to 1 pixel. Also its negative error rate is less as compare with GSLDA (shown in figure 5) [1]. Table I describes the time required to detect face from the given image. From observation we can observe that image containing less number of faces required less time to process as compare to the image containing more number of faces.

![Figure VII: LDA and GSLDA curve for face detector on MIT+CMU test set](image)

![Figure VIII: Comparison between PGSLDA and GSLDA for train and test error sets.](image)

![Figure IX: Output of face detector on number of test images from the MIT+CMU test set.](image)

<table>
<thead>
<tr>
<th># of Faces</th>
<th>Single classifier</th>
<th>Parallel classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>276.24ms (3.62fps)</td>
<td>71.174ms (14.05fps)</td>
</tr>
<tr>
<td>2</td>
<td>282.49ms (3.54fps)</td>
<td>72.57ms (13.78fps)</td>
</tr>
<tr>
<td>6</td>
<td>316.45ms (3.62)</td>
<td>99.30ms (10.07fps)</td>
</tr>
<tr>
<td>7</td>
<td>315.46ms (3.16fps)</td>
<td>106.26ms (9.41fps)</td>
</tr>
</tbody>
</table>
VI. CONCLUSION

An alternative approach for the detection of faces is provided in this research paper. The improved scenario is GSLDA. Both the approaches of parallel and cascaded algorithms are systematically studied. With the development of the parallel classifier and then experimenting gives result shows that double cascade classifier required less time as compare required is to single cascade classifier (GSLDA). The accuracy of face detection is very high. Future work will be focus on the search for more efficient weak classifier and apply this technique for character detection.

REFERENCES