Extended LCCF for Object Detection

Disha Davis[1] Nicy Johnson[2]
[1]Student [2]Assistant Professor
[1,2]Department of Computer Science and Engineering
[1,2]IES College of Engineering, Thrissur, India

Abstract— Object detection is a series of segments containing the features of interest, which are taken as pre-processing and widely applied in various vision tasks. Of these, face recognition is one of its major application. It is a computer vision technology related to the modern method called image processing that deals with identifying instances in the objects of a certain class such as humans that get resided in both images and videos. However, most of existing approaches only utilizes the proposals to compute the location; the aim is to propose a fastest algorithm for making the process as an efficient task. So the correlation procedure has been introduced in order to overcome these existing disadvantages. It is a highly accurate method in order to solve the optimization problems in the correlation filters, more specifically, filters called extended latent constrained correlation filters (LCCF) are proposed. Extended LCCF will maps the correlation filters to a given latent subspace that is created thereby establish a new learning framework. It will relate each pixel in the real time video that is given as the input and will extract the feature points by feature extraction method. The feature points will be stored in the subspace where it recreates using a newer algorithm called autopath algorithm. It will accurately detect the person in the video when it contains noises and occlusions. The extended LCCF method will outperform significantly better than other competing approaches.

Keywords: Object Detection, Correlation Filters, SADMM, Extended LCCF, Autopath, Feature Extraction, Classifiers

I. INTRODUCTION

Today, computer vision has become an integrated part of people's lives. It has, and continues to influence many aspects of daily life and has allowed better working in robotics, real time videos and can also be used for image enhancement, object detection etc. Image processing is one of the rapidly growing technologies and is one of the important parts of computer vision. Of these, image enhancement or the object detection method has become the recently important task due to its many possible applications, such as object recognition, face recognition and video co-segmentation. However, regardless of the ongoing growth in the field of object detection, it is still a challenging job as the process of detecting objectis affected not only by internal factors like accuracy, efficiency, but also by external factors, such as noises, illuminations and occlusions. Several ways has been introduced to detect the object that is associated with the image or the video. One of the important methods is by using correlation. They are usually trained in the frequency domain with the aim of producing a strong correlation peak on the pattern of interest while suppressing the response to the background. To this end, a regression process is usually used to obtain a Gaussian output that is robust to shifting. Of these filters, a newly designed filter called extended LCCF is introduced which can be considered as an image enhancement method. By using this filter, the object that gets resided in the video can be easily detected.

II. MOTIVATION

Object detection is a prominent task related to image processing that deals with identifying occurrences of objects in a particular classes such as humans, buildings or cars in digital images and videos. At present, face recognition has become a relevant task and researchers are finding the faster methods for accurately detecting the persons. Every object class has its own special features that help in classifying the class. Several methods had been introduced in order for detection. But the procedures and the algorithms used in each method are different. The main concern is that, how to efficiently and accurately detects the object without causing any delays. In views of that, handling the noises, illuminations, background clutters and occlusions becomes more important. So that the filters must be designed in order to increase the accuracy, efficiency, speed and performance thereby reducing the human time. There are many popular real world applications that are directly or indirectly related to object detection. With the availability of large amounts of data, faster GPUs, and better algorithms, we can now easily train computers to detect and classify multiple faces of persons within an image with high accuracy. Object detection through automated systems is useful when we need to label an object that gets resided in an image or a video more efficiently.

III. LITERATURE SURVEY

Correlation filters are considered as the special classifiers which are used for shift invariant object recognition and also robust for patters distortion. The idea behind the correlation filter is first proposed by Hester and Casasent namely multivariant technique for multiclass pattern recognition [1]which focuses more on formulating the theory. Correlation with an example template works well if the target does not change notably from image to image and video to video. Unfortunately, the appearance of the target does change across images in most of the domains, due to variations among target instances and changes in imaging condition. It highly focuses on formulating the theory and later on more practical applications. It is a technique for multiclass optical pattern recognition of different standpoint views of an object is described. Each multiclass depiction of an object is recounted as an orthonormal basis function expansion, and a single averaged matched spatial filter is then obtained from a weighted linear combination of these functions. The technique is manifested for a terminal missile guidance application using IR tank imagery. Though it is a simple method and later works had been proposed in order to obtain high efficiency.

D. S. Bolme et al. [2] proposed a newer type of correlation filter called average of synthetic exact filters that differs from the traditional filter is that it will specify the entire correlation output on each training image by tuning filters for particular tasks. This is in marked contrast to prior methods which only specify a single output value per training image. It has two important properties. First, an entire correlation response surface is specified for each training instance during filter construction. Second, the resulting filters are simply averaged. The resulting filters are less susceptible to over-fitting the training data than other methods, and can therefore be trained over larger and more inclusive training sets. As a result, they outperform previous methods. One consequence of completely specifying the correlation output is a perfect balance between constraints and degrees of freedom for each training image, and therefore a complete "exact filter" is determined for every training image. Moreover, the proposed approach implicitly learns correlation filters over an embedded dense sampling strategy which is inherited from the shift circular property of the convolution operation in the Frequency domain. This approach demonstrated superior empirical results for both object detection and real-time tracking compared to current state of thearts.ASEF is shown to out sail all these methods, locating the eye within the radius of the iris approximately 98.5%. The work had been extended to locate more points on faces, such as the nose, eyebrows, and mouth with excellent results.

To work with more variety problems including face detection, face verification, automatic target recognition, and medical image registration, VitomirStruc et al. [3] proposed principal directions of synthetic exact filters. It exhibits some desirable properties such as relatively short training times, computational simplicity, high localization rates and real time capabilities. Instead of only relying on the average of a set of synthetic exact filters, as it is the case with the ASEF filters, eigenvectors of the correlation matrix of the SEFs as correlation templates (or filters) are employed hence, the name PSEFs. After determining the location of the facial regions in all images, 640 images are selected from the LFW database and manually label the locations of the left and right eye. Through these transformations, the initial set of 640 images are augmented to a set of 25600 images (of size 128×128 pixels) and employed for training of the ASEF and PSEF filters. For testing purposes, the same random transforms are applied to 3815 images from the FERET database. Here, only 12 modifications of each facial regions are produced, which results in 45780 facial images being available for the assessment.

Hamed Kiani Galoogahi et al. [4] proposed multi channel correlation filters that can be efficiently posed as a correlation/convolution between a multi-channel image and a multi-channel detector/filter which results in a single channel response map indicating where thepattern (e.g. object) has occurred. To demonstrate the effectiveness of this strategy, it is evaluated across a number of visual detection/localization tasks where it superior computational and memory efficiencies compared exhibit superior performance to current state of the art correlation filters, and to state of the art spatial detectors. This approach is applied across a myriad of detection and localization tasks including: eye localization, car detection and pedestrian detection and demonstrated superior performance to current state of the art single-channel correlation filters, and superior computational and memory efficiency in comparison to spatial detectors (e.g. linear SVM) with comparable detection performance.

Object recognition algorithm using maximum margin correlation filter & SVM is proposed by SamyaBagchi et al. [5]. It considers the problem of detecting objects in two dimensional images and proposed a new technique which uses support vector machine (SVM) along withmaximum margin correlation filter. This algorithm detects objects well and is robust with respect to scale changes. Introduction to support vector machine (SVM) helps maximum margin correlation filter to compacts with non-linearly separable data. It also detects same object, if it is found several times at various scales, thus it helps avoiding detection of same object and finally selects the best version of it. Application of SVM on top of this filter also helps us to detect and select the best peak corresponding to the object. It also helps in detecting object by choosing the best among several scaled versions. This algorithm can be useful for a good number of applications like automatic target recognition, robot vision etc. which requires some more attention.

IV. PROPOSED SYSTEM

Increase in the images has urged for the development of robust and efficient face recognition techniques. The proposed system is based on the idea of constrained correlation filter. More specifically an extension to the latent constrained correlation filters are proposed called extended latent constrained correlation filter (LCCF) is designed. It will maps the correlation filters to a given latent subspace. Once the video or image is obtained, it will does the correlation part along with the preprocessing step by using histogram algorithm in order to avoid noises and occlusions. In the subspace, the system uses a LCCF and SADMM algorithms inorder to recreates the data values based on the relation specified and to avoid the repetition of processing A new algorithm called autopath algorithm is performed on the resulted values. Sections of Knn and ball tree algorithms are used along with the autopath algorithm from dynamic library. Different functions and classes are obtained from dlib(). Ball tree algorithm is used for partitioning of data points in to a nested set of hyper spheres and knn algorithm is used in order to train the system with single image. After comparing the resulted values with the training image, it will detect the face thereby displays the details for the known persons and displays unknown for unknown individuals. The system architecture of proposed system is shown in figure 1.

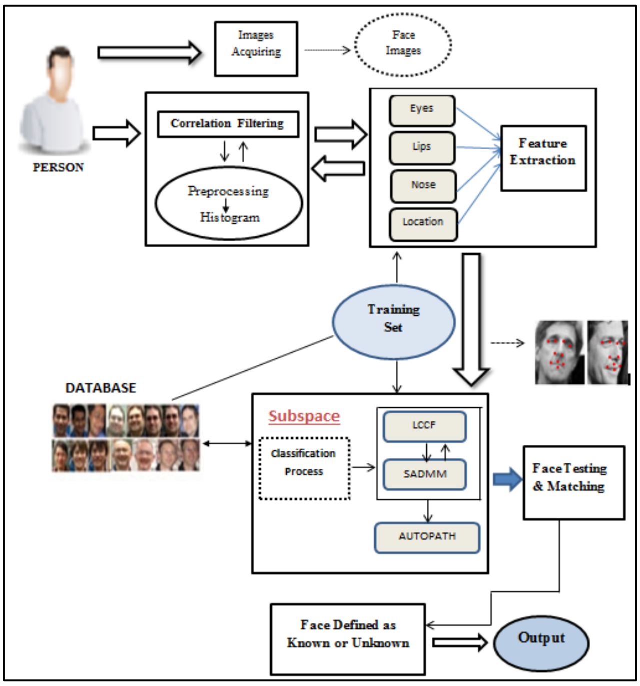

Fig. 1: Architecture Diagram

V. METHODOLOGY

A. Technologies

1) Correlation

Correlation filters have attracted increasing attention due to its simplicity and high efficiency. They have successfully been applied to various problems ranging from pattern recognition tasks such as face to basic computer vision problems related to object detection and tracking. In this paper the correlation method will relate the adjacent cells in the video and will obtain data values that are to be stored in the subspace. So that, the values in the subspace are only need to process on further steps instead of processing the entire values.

B. Algorithms

1) Histogram Algorithm

In an image processing context, the histogram of an image mainly refers to a histogram of the pixel intensity values. The horizontal axis of the graph represents the tonal variations, while the vertical axis represents the number of pixels in that particular tone. The proposed system uses histogram method in the preprocessing stage in order to make the given input video clearer. If the video comes up with noises and occlusions, it will be highly difficult in order to identify the person in the video. So by using the histogram algorithm, a clear peak values will obtained by removing the noises thereby it is easy to extract the feature points by correlation.

2) LCCF Algorithm

The LCCF algorithm is used while creating the subspace and for recreating the data values in one of the layer of the subspace with respect to the relation specified. The results will be passed to the next layer where it is performed again. This process continues until correct data values are obtained. This algorithm will works in a nonlinear fashion. That is, it will choose the values in a random manner so that it doesn't follow any steps. As a result, the output obtained will be highly accurate and it will also reduce the human time.

3) SADMM Algorithm

Subspace based alternating direction method of multipliers are the extension of classical ADMM methods which can be used in the case of multipliers. By using this algorithm, repetition of processing can be avoided. It takes inherent visual data structure for solving optimization problem. It is used to solve both linear & non-linear correlation filtering based on latent subspace constraint.

4) Autopath Algorithm

Autopath algorithm is designed in order to make the recognition process faster without causing any delays. As the name implies it takes data by automatically roots to each node & choose the path. It is an extension to ordinary LCCF algorithms. The existing approaches require a lot of trained images in the dataset but while using this autopath algorithm, it requires only a single image for training purpose. This is because it uses sections of knn and ball tree algorithms along with the autopath. These sections are called from the dynamic library called dlib. Ball tree as the name implies is used for partitioning of data points into a nested set of hyperspheres. The data values will form the nodes of tree so that autopath algorithm will root to the node which it likes to take the values. Knn is used to search through the training dataset for the most nearest instances.

VI. SYSTEM IMPLEMENTATION

Implementation begins with the selection of preferred OS. Operating system selected is Linux. Configuration settings are done according to the users needs. The language used is python. Python program is created using the IDE Pycharm and the library used is Opencv.

Package manager is ANACONDA. It simplifies the package management and deployment. Various packages are installed from the ANACONDA repository. Packages installed are NumPy, SciPy, Analysis and Matplotlib.

Initially feasibility study has to be done before starting the project. It is about checking whether the system can work with the packages and tools. For setup the system, several functions are used. The main functions used are Click() function to trigger the click event, dlib() which is the dynamic library in order to use the code which is created by someone that can be loaded anywhere in the memory so that section of knn and ball tree can be called. Tox() and flake() functions are used for functional test in order to check the interaction between the modules.

Final stage is the system interfacing step. The system is interfaced with the program to work the extended LCCF properly. The main concern is that, as it is a real time project, webcam reference must be available. Many inbuilt functions are used for the interfacing purpose. The webcam reference is done using the function cv2.Videocapture(0). Then the images are trained using the function load_image_file in order to detect the face. Facial locations & distances in the input faces are obtained using lccf and sadmm algorithms. The obtained faces are made as a batch file using batch_face_locations. Subspace is created using arrays and the face encodings are done for the known faces. If face_encoding = 0, no faces is encoded and if face_encoding has a value then, a face is given. Opencv is the main library used as the front end of the system. It uses BGR color format. So the system will convert this format to RGB for the face recognition purpose. Based on these features, comparisons are done. For the training process, a single image is only given. If a match is occurred then the details are displayed on the bottom of the bounding box. It is displayed using the function called cv2.imshow().

VII. Results

A lot of methods had been introduced in order for better detection of the person but the aim is to make this detection more and more fasterin fewer steps. Each method is different in its on way of usage and algorithms that are used. So that we have to propose a newer algorithm that obtains better performance than the others. In order to overcome the disadvantages of the existing systems, the proposed system called extended LCCF for object detection with the application of face recognition is introduced. It is important in its way of detection where it produces the result as efficient as possible.

Real time videos and images are used for the experiment. Initially a real time video is passed as the input. Here we are considering two cases. The first case is correlation part in which it will relate the adjacent pixel values. The relation may be similarity or differences in the image or the video. Then it will extract the feature points or the data values. The second case is that creating the subspace. The extracted feature points are stored in the subspace and are compared with the training images given. The images have to be trained based on the image or video given as the input. For comparing, it uses autopath algorithm for better result. By using this algorithm, it will identifies the face based on the training image as faster as possible. That is, it will detect the person in the video and displays the details as output for a known individual while displays as unknown for unknown individual. This system can be placed in public places and security places so that it will be very helpful in investigation purposes.

Fig. 2: Output of Extended LCCF

VIII. Conclusion

In this paper latest works in the field of object detection were discussed. A new filter called extended LCCF was introduced and applied to the application called face recognition. Many researchers had contributed and are still working in this field. Though there are a number of problems in existing systems that need to be addressed such as occlusions, noises and background clutters that are associated with the object. However some of the problems like redundant images on various scales, detecting different objects on same image have already been solved but still have room for improvement. It is concluded, by using extended LCCF all the existing disadvantages can be avoided by relating the pixels like difference between foreground and background image, illumination variations etc., thereby it can accurately detect the person by working in a non linear fashion. The experimental results have shown consistent advantages over the state-of-the-arts when applying extended LCCF to several computer vision applications. In future work, it can be incorporated in CUDA frameworks.

Acknowledgment

I would like to convey my heartfelt gratitude towards my guide Mrs. Nicy Johnson for her constant guidance, encouraging help and inspiring words. I am thankful to the department of computer engineering for their support.

References

[1] C. F. Hester and D. Casasent"Multivariant technique for multiclass pattern recognition", Applied Optics, vol. 19, no. 11, pp. 1758–1761, 1980.
[2] D. S.Bolme, B. A. Draper and J. R. Beveridge, "Average of synthetic exact filters", IEEE Conference on Computer Vision Pattern Recognition (CVPR), pp. 2105–2112, 2009.
[3] VitomirStruc, JernejaZganecGros and Nikola Pavesi"Principal directions of synthetic exact filters for robust real-time eye localization", Springer, pp.180-192, 2011.
[4] HamedKianiGaloogahi, Terence Sim and Simon Lucey, "Multi-channel correlation filters" IEEE International Conference on Computer Vision, pp. 3072–3079, 2013.
[5] SamyaBagchi and Poonacha P. G, "An object recognition algorithm using maximum margin", IEEE Twentieth National Conference on Communications (NCC), pp.1 – 6, 2014.