ISSN No. 0976-5697

Volume 6, No. 1, Jan-Feb 2015

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Object Detection for Markerless Augmentation using Haar Training

Shubham Paul, Sourabh Dixit
CSE/ IT department,
Jaypee Institute of information Technology
Noida, India

Dr.Chetna Dabas
CSE / IT Department
Jaypee Institute of Information Technology
Noida, India

Abstract: Computer Vision is a field that includes methods for acquiring, processing, analyzing, and understanding images and, in general, high-dimensional data from the real world in order to produce numerical or symbolic information etc[1]. As a highly scientific field, computer vision is concerned with the theory behind artificial systems that extract information from images involving image processing and object recognition. In our research we aim to explore this extensive domain of object recognition and image processing in terms of real time environment and deploy its application on various open source environments like Android. For achieving this we have incorporated a recently introduced Viola Jones, rapid object detection scheme based on a boosted cascade of simple features. For the object detection and recognition, we performed Haar Training on a set of positive and negative samples, which result in creation of Haar Cascade Classifier files in xml formats. These files are then used for detection. Post successful completion, the proposed project could also find its application in Markerless Augmented Reality as recent convergence of imaging sensors and general purpose processors on mobile phones creates an opportunity for a new class of augmented reality applications.

Keywords: Computer Vision, Object Recognition, Viola Jones, Android, Haar Training, Haar Cascade Classifier, Markerless Augmented Reality

I. INTRODUCTION

The project deals with providing object recognition algorithms which will help and guide the users of their respective devices by helping them with a better understanding of an unknown device in order to set up its working.

Real-Time object detection and recognition of the object using the device's camera. Haar cascade classifier files were created by performing haartraining on the object and its ports' images, for detection purposes.

This paper extends Viola Jones Algorithm's rapid object detection framework in two important ways: [2]Firstly, their basic and over-complete set of haar-like feature is extended by an efficient set of 45° rotated features, which add additional domain-knowledge to the learning framework and which is otherwise hard to learn. These novel features can be computed rapidly at all scales in constant time. Secondly, we derive a new post optimization procedure for a given boosted classifier that improves its performance significantly.

Our prototype implementation has been applied to several practical applications such as image search, rapid object recognition and augmented reality applications. The project will be an Android Application to ensure better portability and thus keeping in mind the vide scope of Mobile Technologies in coming future.

II. RELATED WORK

Many researchers have earlier tried to improve object recognition algorithms performance and accuracy. The common approaches used by them are Scale Invariant Feature Transform (SIFT) ,Speeded Up Robust Features (SURF) and Edge Detection which we have discussed and put forward a comparative analysis for each in comparison to Haar Training Approach which we have preferred to the latter ones.

Scale Invariant Feature Transform (SIFT) uses the technique in which key points of objects are first extracted from a set of reference images and stored in a database repository. An object is recognized in an image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. SIFT falls under Feature based Detection and was discontinued in our next phase as it did not prove to be efficient and failed to recognize the target object.

Speeded Up Robust Features (SURF) is a robust image detector and descriptor. The standard version is several times faster than SIFT and claimed by its authors to be more robust against different image transformations than SIFT based on sums of approximated 2D Haar wavelet responses and made efficient use of integral images. Also a feature based detection method, but is more robust than Sift and makes use of interest points generated using viola Jones algorithm.

Edge Detection Algorithm Uses edge detection techniques such as the Canny edge detection, to find edges .Changes in lighting and colour usually don't have much effect on image edges. The module was compiled on android emulator and tested with various images. The edges detected were not sufficient for the recognition of target in the scene.

Also, comparison of an already launched android application 'Aurasma' which is similar to the app to be developed in the project. Aurasma is an augmented reality platform created by Autonomy Corporation. Mainly designed for 3G and 4G mobile devices it uses the device's video camera to recognize pre-trained images and overlay an image or video so that the video tracks as the camera is moved.

The concept behind Aurasma and that of our android app is similar, in a way that Aurasma too deals with markerless augmented technology and object recognition methods, and our app which is supposed to detect a target(router in this particular case) and not a marker, is also markerless. Although the concept is same, but the procedure implemented for augmentation in Aurasma is entirely different to our approach.

III. EMPIRICAL ANALYSIS

Many researchers have been done in the field of object detection. We have tweaked the methods so far proposed in research papers by applying them after with markerless augmentation on mobile.

A. Logical Database Requirement:

The database requirement for our project is based on the images of Router (target object) which need to be trained for detection of the Router. Haartraining for image detection is the approach applied for training the sample images and detecting the Router. For HaarTraining, we require two sets of images in our database-

a. ***Positive Images*** - The set of images containing the target object (router) in them to be used as positive samples. To obtain set of positive images , we use 'ffmpeg' which is ffmpeg is a command-line tool to convert one video file format to anoth-er. It can also grab and encode in real-time from videos as done by us.

b. ***Negative Images*** - The set of images that do not contain the object to be detected, in our case Router. Using the above images, on Linux OS, Haartraining was performed, in which the database was used to create samples first in form of vector files (.vec format). These vector files were then used to create Haarcascades.

B. Interface Requirement:

a. The application will be started on the android device and it is required that user focuses the device's inbuilt camera in the direction of the router.
b. Router Detection- The router will be detected from the moving video frame captured from the device camera.
c. Ports Detection and Recognition- After the router has been detected by our application, the ports shall be detected and then recognized for further functions.

C. Android Integration:

Our project is based on developing an Android Application which will be able to guide the users about installation of Router to set up a working internet connection. Eclipse Indigo is integrated with the libraries of OpenCV (for image detection and recognition) and OpenGL (for rendering 3D graphics).

D. Sample Target Object (Router):

Figure: 1 Sample Router Image (Positive Image)

IV. TECHNIQUE APPLIED FOR OBJECT RECOGNITION

A. HAARTRAINING - Creating Haar Cascade Classifiers:

In this technique we first collect positive and negative images of the object (i.e. Router in this case). We create samples using these images which are converted to a vector file (.vec file). Now after integrating all the vectors in a single file comes HaarTraining. After several hours of training we get haar cascades of the samples and which are further converted into a cascade classifier XML file.

Haar-like features are digital image features used in object recognition.this approach basically deals with creating .xml file also known as boosted haar cascade classifier file, generated upon performing Haar training over a set of positive and negative samples of the target to be tracked.

A cascade of classifiers is degenerated decision tree where at each stage a classifier is trained to detect almost all objects of interest while rejecting a certain fraction of the non-object patterns[4].

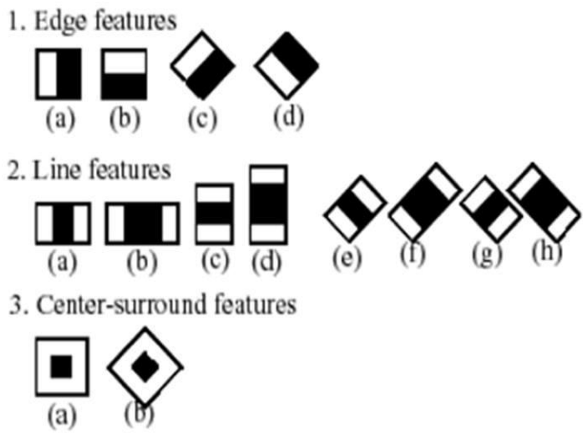

Figure: 2 Haar-Like Features

a. Mathematical Representation:

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[5]. 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[6]. 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%.The effect of this single classifier is to reduce by roughly half the number of times the entire cascade is evaluated.

The cascade architecture has interesting implications for the performance of the individual classifiers. Because the activation of each classifier depends entirely on the behavior of its predecessor[5], the false positive rate for an entire cascade is:

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

Similarly, the detection rate is:

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

Thus, to match the false positive rates typically achieved by other detectors, each classifier can get away with having surprisingly poor performance. For example, for a 32-stage cascade to achieve a false positive rate of 10^{-6}, each classifier need only achieve a false positive rate of about 65%. At the same time, however, each classifier needs to be exceptionally capable if it is to achieve adequate detection rates. For example, to achieve a detection rate of about 90%, each classifier in the aforementioned cascade needs to achieve a detection rate of approximately 99.7%.

B. Viola-Jones Algorithm:

The Viola–Jones object detection framework is the first object detection framework to provide competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones. Although it can be trained to detect a variety of object classes, it was motivated primarily by the problem of face detection[7]. This algorithm is implemented in OpenCV as cvHaarDetectObjects()[8].

a. Training a Cascade of Classifiers:

The cascade training process involves two types of tradeoffs[9]. In most cases classifiers with more features will achieve higher detection rates and lower false positive rates. At the same time classifiers with more features require more time to compute[9]. In principle one could define an optimization framework in which:

a) The number of classifier stages,
b) The number of features in each stage, and
c) The threshold of each stage, are traded off in order to minimize the expected number of evaluated features.

Unfortunately finding this optimum is a tremendously difficult problem. In practice a very simple framework is used to produce an effective classifier which is highly efficient. Each stage in the cascade reduces the false positive rate and decreases the detection rate.

A target is selected for the minimum reduction in false positives and the maximum decrease in detection. Each stage is trained by adding features until the target detection and false positives rates are met (these rates are determined by testing the detector on a validation set). Stages are added until the overall target for false positive and detection rate is met.

C. OpenCV (Open Source Computer Vision Library):

It is a library of programming functions mainly aimed at real-time computer vision, developed by Intel, and now also supported by Willow Garage and Itseez. It is free for use under the open source BSD license. The library iscross-platform. It focuses mainly on real-time image processing. OpenCV is written in C++ and its primary interface is in C++.

It was integrated with the Android SDK in the later stages of our project to develop an android based application.

V. RESULT

Accuracy of the Haar Classifier against the other existing classifiers has been shown in table1. We are validating results with the help of following research questions:

RQ1: Can haar training help in improving the result as compared to other classifiers?

As shown in Table1, we are able to achieve almost 50% better accuracy using haar training as compared to accuracy achieved using the SVM and Random forest classifiers. Haar cascade classifiers getting 83% accuracy in object detection, in this case a router and its ports, as compared to 54% which was with SVM classifier. Similarly, using Random classifier, we achieved an accuracy of only 45% which is the least among the three classifiers.

Hence, we are able to increase accuracy of result amazingly by applying defined haar training steps.

RQ2: Which classifier is able to achieve better results?

We have achieved our best accuracy from Haar cascade classifier that is close to 83%.

Table: 1 Accuracy For Different Algorithms

Classifiers	Precision	Sensitivity	Specificity	Accuracy
SVM Classifier	0.1056	0.3191	0.4783	0.4526
Random Forest Classifiers	0.8738	0.7713	0.9429	0.5492
Haar Cascade Classifier	0.9739	0.7253	06275	0.8323

VI. CONCLUSION

In this paper we applied three different classifiers to detect objects, one being the most accurate for object detection on mobile. We trained the haar classifier with positive and negative images and detected the object on the basis of that classifier using the OpenCV library.

Although haar training used by us have many advantages, it certainly does come with some disadvantages. The disadvantage of haar training is that they are time consuming. The original system took several weeks to train and even on modern computers, training time may range from hours to days. Therefore, accelerating this training phase in a power efficient manner is important to facilitate further exploration of this detector technique and allow for object detectors to be quickly trained in field environments.

VII. REFERENCES

[1] "Computer Vision and Smart Embedded Systems,R&D&I in Applied Research",Carmen Alonso Montes , Basque Center for Applied Mathematics , 2013

[2] "An Extended Set of Haar-like Features for Rapid Object Detection" Rainer Lienhart and Jochen Maydt Intel Labs, Intel Corporation, Santa Clara, CA 95052, USA Rainer.Lienhart@intel.com

[3] "Efficient Extraction of Robust Image Features on Mobile Devices" Wei-Chao Chen , Nokia Res. Center, Palo Alto, CA , Yingen Xiong , Jiang Gao , Gelfand. Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium

[4] "Face tracking based on Haar-Like features and eigenfaces" Paulu Menezes , Jos´e Carlos Barreto , Jorge Dias

[5] "A method for hand detection based on Internal Haar-like features and Cascaded AdaBoost Classifier" - Van-Toi Nguyen,Thi-Lan Le ,International Research Institute Mica Hust - Grenoble Inp Hanoi University of Science and Technology Vietnam

[6] "Ultra Low Power Signal Processing in Wearable Monitoring Systems:A Tiered Screening Architecture with Optimal Bit Resolution",Hassan Ghasemzadeh, University of California, Los Angeles,Roozbeh Jafari, University of Texas at Dallas.

[7] Log-Gabor Filters and Lda Based Facial Expression RecognitionT.Gourham Reddy, Ms.T.Sravanthi - IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) Volume 3, Issue 6 (Nov. – Dec. 2013), PP 37-43e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197

[8] RFID Based Efficient Smartcard Authentication System Using Facial Recognition And One Time Password(OTP)- Ms. Anjali Singh,Asst. Prof. Mrs. Neelam Dhurpate. Prof. Mr. Veera Manickam - International Journal of Advance Foundation and Research in Computer (IJAFRC) Volume 1, Issue 5, May 2014. ISSN 2348 - 4853

[9] "Robust Real-time Object Detection" - Paul Viola, Michael Jones,Second International Workshop On Statistical And Computational Theories Of Vision – Modeling, Learning, Computing, And Sampling Vancouver, Canada, 2001.