ID Card Detection

M. Azhagiri[1], Aditya Vashishtha[2], Rahul Saha[3], Debashree Kalita[4].

[1]Assistant Prof., Computer Science, SRMIST, Tamil Nadu, India.

[2]UG Scholar, Computer Science, SRMIST, Tamil Nadu, India.

[3]UG Scholar, Computer Science, SRMIST, Tamil Nadu, India.

[4]UG Scholar, Computer Science, SRMIST, Tamil Nadu, India.

Abstract

The paper describes detection of an image of 64x 128 pixel. An optimized algorithm is proposed based on conventional Histograms of Oriented Gradient descriptor and Support Vector Machine classifier algorithms for hardware implementation. By implementing the proposed algorithm on a field-programmable gate array (FPGA) of a high-speed vision platform, multi- object in an image can be detected at 12,000 fps under complex background. In hardware implementation, 64 pixels were processed in parallel with 80 MHz camera clock. The source image and detection results can be moved to desktops or laptops in real-time for recording or further processing.

Keyword:- *HOG, SVM, FPGA, RGB, Image Detection.*

I. INTRODUCTION

Many descriptors and algorithms have been proposed for object detection and recognition. Some of them are- Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradient (HOG), etc. The HOG descriptor has some advantages over other descriptors. As the computational complexity and resource consumption is high, HOG descriptor is implemented on CPU or Graphics Processing Unit (GPU). It is hard to achieve object detection and recognition at a speed of 1000 fps by using CPU or GPU, because of limited processing power and image transportation speed from camera to desktop.

To overcome shortage of the system with CPU or GPU, a proposal on object detection algorithm based on HOG descriptor and SVM classifier. This is a low complexity, low memory consumption, and high parallelism algorithm. By implementing our proposed algorithm in the FPGA of a high-speed vision platform, 12,000 fps object detection is realized for a 64x128 pixel image.

II. RELATED WORK

HOG features are extracted of the edges by taking into one file trained and tested using WEKA software. WEKA‘s functionality can be accessed through the graphical user interface. For training process, extracted HOG features from the images in the database are trained with the classifier in the WEKA.

The time taken to build the classification of different users is 0.02ms. For testing, this software can correctly identify the person using the trained data of the features. By using this method the identification will be more accurate.[1]

The histogram equalization has been a widely used image processing technique for speech enhancement, which has the property of increasing the global contrast of an image; while simultaneously compensating for the illumination conditions present at the image acquisition stage.

The main objective of this technique is to enhance the discriminative information contained in the facial images and ensure that environmental factors, such as the ambient illumination present at the image acquisition stage, do not influence the process of facial-feature-extraction. On the ideal case, the contrast of an image would be optimized if all the 256 intensity levels were equally used.[2]

The histogram of oriented gradient (HOG) descriptors could capture information about the gradient orientations in localized areas of an image. The implementation of these descriptors can be attained by dividing the image into small connected regions, that sometimes can be called cells, and for each cell a local 1-D histogram of gradient directions is accumulated over the pixels of the cell.

Each image is divided into a sequence. The number of points in each grid cell is then recorded. The number of points in a cell at one level is simply the sum over those contained in the four cells it is divided into at the next level thus forming a pyramid representation. [3]

An image is divided into two forms: (a) original images with corresponding annotation files, and (b) positive images in normalized 64x128 pixel format with original negative images. Some noteworthy points are:

Only upright persons (with person height > 100) are marked in each image. Annotations may not be right; in particular at times portions of annotated bounding boxes may be outside or inside the object. Normalized positive training or test images centered on the person with their left-right reflections.[4]

To generate negative training windows from normalized images, a fixed set of 12180 windows (10 windows per negative image) are sampled randomly from 1218 negative training photos providing the initial negative training set. For each detector and parameter combination, a preliminary detector is trained and all negative training images are searched repetitively for false positives. All examples with score greater than zero are considered hard examples. The method is then re-trained to produce the final detector. [5]

For extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. The features are highly distinctive, in the sense that a single feature can be correctly matched with high probability against a large database of features from many images. The features are invariant to image scaling and rotation, and partially invariant to change in illumination.

This approach has been named the Scale Invariant Feature Transform (SIFT), as it transforms image data into scale-invariant coordinates relative to local features. The quantity of features is important for image recognition, where the ability to detect small objects in cluttered backgrounds requires that at least 3 features be correctly matched from each object for reliable identification.[6]

III.DESIGN AND IMPLEMENTATION

This process begins with calculating the gradients. Calculate the x and the y gradient images, and from the original image. Now this can be done, by filtering the original images with kernels. Using the gradient images, the magnitude and orientation of the gradient. The calculated gradients are "unsigned" and therefore is in the range 0 to 180 degrees.

Next divide the image into 8×8 cells. At each pixel in an 8×8 cell the magnitude and direction is known, and therefore we have 64 magnitudes and 64 directions — i.e. 128 numbers. Every pixel votes for either one or two bins in the histogram. A pixel where the direction of the gradient is not exactly 0, 20, 40 … 160 degrees splits its vote among the two nearest bins based on the distance from the bin.

The histogram calculated in the previous step is weak to lighting changes. Multiplying image intensities by a factor multiplies the histogram bin values as well. To counter these effects the histogram can be normalized — i.e. think of the histogram as a vector of 9 elements and divide each element by the magnitude of this vector. In the original HOG paper, this normalization is not done over the 8×8 cell that produced the histogram, but over 16×16 blocks. The idea is the same, but now instead of a 9 element vector you have a 36 element vector.

The preprocessing process begins with cropping the outline of the scanned image. For calculating the gradient images OpenCV is used. It calculates the derivatives for each color channel separately. Combination of the color channels to come up with one gradient magnitude and one gradient angle is not defined clearly. Reducing the color image to a gray value image. And use this scalar image from here on and visualize the gradient components fx and fy and the gradient magnitude where,

x - no. of cells in horizontal direction which is usually set to 8.

y - no. of cells in vertical direction which is usually set to 8.

For calculating a HOG in a 8x8 cell, write a function HOG*8x8(gx, gy)* that takes two scalar 8x8 images as input where gx is the derivative in x direction and gy the derivative in y direction.

A function *HOG block (w1, w2, w3, w4)* is taken for calculating the HOG feature vector that takes 4 cell histograms from step 3, concatenates them in one large (36) vector and normalizes the vector. In the middle of each 8x8 block a "rose of directions" is drawn.

Every bin in the histogram for that block shows the strength of the gradient in that particular orientation. We draw a line in the orientation (center of the line in the middle of the block) with a length that is proportional to the bin count in the histogram.

After the process gives 3780 feature cells, which determines the structural identity, with specific feature processing it can be detect whether the image is real or fake.

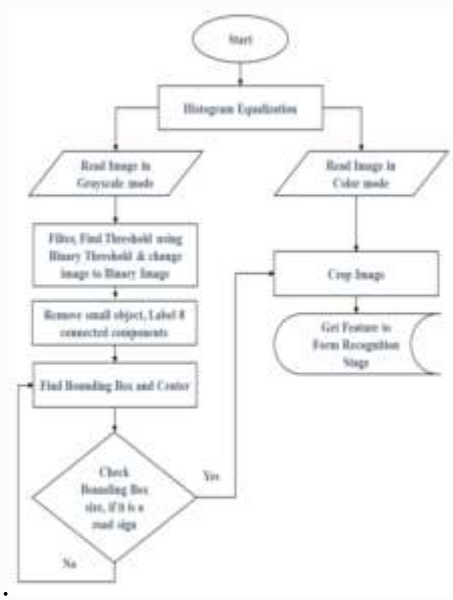

Fig -1 System architecture diagram for ID Card Detection

IV.CONCLUSION AND FUTURE SCOPE

The paper proposes a multi-object detection algorithm based on HOG descriptor and SVM classifier. The algorithm is designed for parallel processing, making it suitable for FPGA with limited Block RAM and Multiplier resources.

Implementation of the proposed algorithm in the FPGA of a high speed vision platform, multi-object detection is achieved at 12,000 fps with 64x128 pixels images. The results show that the design of our proposed algorithm is feasible. The proposed system is efficient for objects with descriptive texture information. Multi-classes objects in a single frame can be recognized by training SVM coefficients of different classes and implementing multi-classes classification circuits in FPGA. For future prospects, improvement on recognition accuracy for complex objects by combining the system with other features, such as moments and brightness histograms.

V. REFERENCES

[1] NavneetDalal and Bill Triggs. Histograms of oriented gradients for human detection. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, International Conference on Computer Vision

[2] K. Ramirez-Gutierrez, D. Sanchez-Perez, H. Perez-Meana, “Image recognition and verification using histogram equalization,” Selected Topics in Applied Computer Science, H. Fujita and J. Sasaki, Eds. WSEAS Oct. 2010, pp. 85-89.

[3] Image Recognition Using Pyramid Histogram of Oriented Gradients and SVM Hui-Ming Huang, He-Sheng Liu, Guo-Ping Liu 7

[4] Kanade Takeo, Picture Processing by Computer Complex and Recognition of Human Faces, Kyoto University, Ph.D. thesis edition,1973

[5] Zhi Zhou, Eliza Yingzi Du, N. Luke Thomas, Edward J. Delp, ―A New Human Identification Method: Sclera Recognition,‖ IEEE Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans, Vol. 42, No. 3, May 2012.

[6] J. Daugman, `` New Methods in Iris Recognition,‘‘ IEEE Transactions on Systems, Man, and Cybernetics, Part B, vol. 37, pp. 1167-1175, 2007. P.FascaGilgy Mary, P.SunithaKency Paul, J.Dheeba, ―Human Identification based on Periocular Biometrics,‖ International Journal of Science, Engineering and Technology Research (IJSETR) Volume 2, Issue 5, May 2013.