International Journal of Innovative Research in Computer Science & Technology (IJIRCST)
ISSN: 2347-5552, Volume-3, Issue-3, May-2015

Intelligent Transport System for Human Detection with an Efficient HOG Extraction Method

Elavarasi.K, Dr. R. Kalpana

***Abstract*— A robust human detection system in an intelligent transportation system is desired by people and becomes essential to industries such as surveillance, automotive systems, and robotics. However, there are still many encounters to attain ideal human detection, such as the diversity of object appearance and the interference of an image due to light changing. These challenges make human detection a more challenging and unreliable task. Histograms of Oriented Gradients (HOG) are proven to be able to knowingly outpace existing feature sets for human detection. In this work, motivation only on the feature extraction method using HOG for real-time applications. For simplicity, a linear support vector machine (SVM) is used as a baseline classifier throughout the study. It is obvious that the calculation of HOG feature extraction is computationally complicated and unsuitable for hardware implementation. Hence, to adopt some approximate techniques, it will reduce implementation complexity and to improve extraction speed.**

***Index Terms*— Human Detection; Feature extraction; Histograms of Oriented Gradients (HOG); Complexity; baseline classifier; object appearance.**

I. INTRODUCTION

Human detection has been an important issue and has been widely used in many applications, such as surveillance, automotive systems, and robotics. A robust human detection system in intelligent transportation systems is desired by people and becomes essential to industries. However, there still are many challenges to attain ideal human detection, such as the diversity of object appearance and the interference of an image due to light changing.

These challenges make human detection become more challenging and unpredictable. Many scholars have proposed diverse techniques to deal with object detection. Generally object detection includes two significant processes, i.e., feature extraction and classification.

There are various feature extractions, such as Harr wavelets, scale - invariant feature transform descriptors; Gabor filters shape contexts, histograms of oriented gradients (HOG), etc. [1]. Human detection under occlusion is a challenging problem in computer vision. To address this problem through a Framework this integrates face detection and person detection. First investigate how the response of a face detector is correlated with the response of a person detector. From these observations, to formulate hypotheses that captures the intuitive feedback between the responses of face and person detectors and uses it to verify if the individual detectors outputs are true or false. To illustrate the performance of our integration framework on challenging images that have considerable amount of occlusion, and demonstrate its advantages over individual face and person detectors.

II. DATASET AND METHODOLOGY

Datasets are tested our detector based on two different data sets. That is the well-established MIT pedestrian database, containing 509 training and 200 test images of pedestrians in city scenes (plus left-right rejections of these). It contains only front or back views with a relatively limited range of poses. Our best detectors give essentially perfect results on this data set, so we produced a new and significantly more challenging data set, `INRIA', containing 1805 64_128 images of humans cropped from a varied set of personal photos. The people are usually standing, but appear in any orientation and against a wide variety of background image including crowds. [1]

A. HOG

Histograms of Oriented Gradients (HOG) are feature descriptors used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

Manuscript received May 15, 2015
ELAVARASI. K PG Scholar, Department Of Computer Science & Engineering, IFET College of Engineering. Tamilnadu, India.
Dr. R. KALPANA, Professor, Department of Computer Science and Engineering,, IFET College of Engineering, Villupuram, India

B. ORIGINAL HOG ALGORITHM

Scanning on the input image is based on detection window. The window is divided into cells, for each cell accumulating a histogram of gradient orientations over the pixels of the cell. For better invariance to illumination, histogram normalization can be done by accumulating a measure of the local histogram energy over blocks and using the results to normalize all cells in the block. The normalized histograms (HOG features) are collected over the detection window. The collected features are fed to a linear SVM for object/non object classification.

III. PROPOSED SYSTEM

Sketch based image retrieval system to retrieve the better image from the database. But above studied algorithm have some disadvantages; to overcome these disadvantages, here we present some work which is useful for retrieve the better image from the database than the previous system. In our proposed system, combine the HOG (Histogram of Oriented Gradient) Descriptor and the K-mean algorithm together.

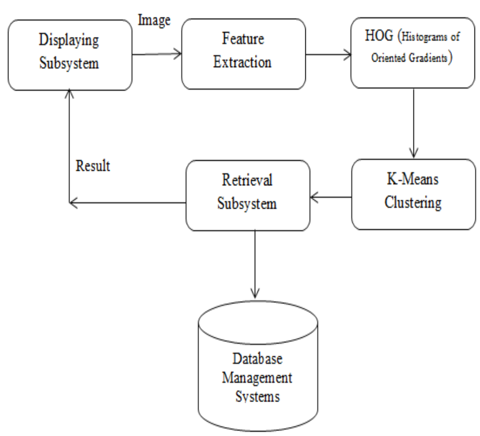

Figure1. HOG Descriptor and the K-mean algorithm

Our architecture adopts a cell-based pipeline flow, shows a relation between cells, blocks, windows, and a frame. One cell contains 8 × 8 pixels. One block is composed of 2 × 2 cells. One window is made up of 7 × 15 blocks. Each block overlaps with neighboring blocks. Cell-based pipeline processing is conducted as follows:

1. A cell histogram is generated with cell-based scanning.
2. When the process described above reaches the block level, a block-level cell histogram is normalized; then the block-level HOG feature is extracted.
3. Block-level HOG features and SVM coefficients corresponding to each window are multiplied and accumulated.
4. An accumulation result of window level is compared with the SVM threshold. Then the detection result is obtained.

The cell-based pipeline architecture greatly reduces the memory bandwidth because it prevents reloading of input pixels in different detection windows.

For HDTV resolution requires memory bandwidth of 55Gbps. In general, the mobile system under limited battery conditions adopts a lower operating frequency. Therefore, the memory bandwidth must be reduced as low as possible for low-power and real-time operation. Our approach adopts a cell-based algorithm and architecture to reduce the memory bandwidth to 0.499Gbps.

A. Gray Image

A gray scale or grey scale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black-and-white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. Gray scale images are distinct from one-bit bi-tonal black-and-white images, which in the context of computer imaging are images with only the two colors, black, and white (also called bi-level or binary images). Gray scale images have many shades of gray in between.

Gray scale images are often the result of measuring the intensity of light at each pixel in a single band of the electromagnetic spectrum (e.g. infrared, visible light, ultraviolet, etc.), and in such cases they are monochromatic proper when only a given frequency is captured. But also they can be synthesized from a full color image; see the section about converting to gray scale.

Figure2. Gray image

B. Binary Image

A binary image is a digital image that has only two possible values for each pixel. Typically the two colors used for a binary image are black and white though any two colors can be used. The color used for the object(s) in the image is the foreground color while the rest of the image is the background color. In the document-scanning industry this is often referred to as "bi-tonal".

Binary images are also called bi-level or two-level. This means that each pixel is stored as a single bit—i.e., a 0 or 1. The names black-and-white, B&W, monochrome or monochromatic are often used for this concept, but may also designate any images that have only one sample per pixel, such as gray scale images. In Photoshop parlance, a binary image is the same as an image in "Bitmap" mode.

Binary images often arise in digital image processing as masks or as the result of certain operations such as segmentation, thresholding, and dithering. Some input/output devices, such as laser printers, fax machines, and bi-level computer displays, can only handle bi-level images.

A binary image can be stored in memory as a bitmap, a packed array of bits. A 640×480 image requires 37.5 KiB of storage. Because of the small size of the image files, fax machine and document management solutions usually use this format. Most binary images also compress well with simple run-length compression schemes.

Binary images can be interpreted as subsets of the two-dimensional integer lattice Z2; the field of morphological image processing was largely inspired by this view.

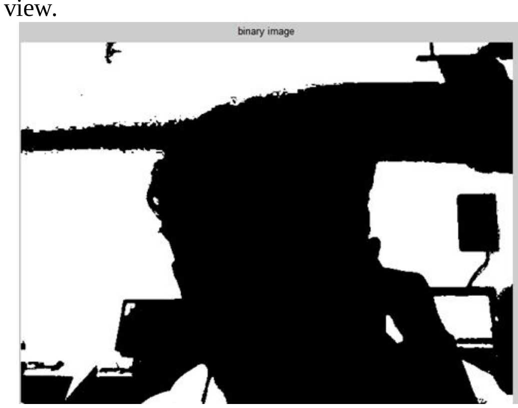

Figure 3: Binary Image

C. Smoothing and Filtering

In an image processing, to smooth a data set is to create an approximating function that attempts to capture important patterns in the data, while leaving out noise or other fine-scale structures/rapid phenomena. In smoothing, the data points of a signal are modified so individual points (presumably because of noise) are reduced, and points that are lower than the adjacent points are increased leading to a smoother signal.

Smoothing may be used in two important ways that can aid in data analysis by being able to extract more information from the data as long as the assumption of smoothing is reasonable and by being able to provide analyses that are both flexible and robust. Many different algorithms are used in smoothing. Data smoothing is typically done through the simplest of all density estimators, the histogram.

Figure 4: Smoothing

Filtering is a technique for modifying or enhancing an image. For example, you can filter an image to emphasize certain features or remove other features. Image processing operations implemented with filtering include smoothing, sharpening, and edge enhancement.

Filtering is a neighborhood operation, in which the value of any given pixel in the output image is determined by applying some algorithm to the values of the pixels in the neighborhood of the corresponding input pixel. A pixel's neighborhood is some set of pixels, defined by their locations relative to that pixel. Linear filtering is filtering in which the value of an output pixel is a linear combination of the values of the pixels in the input pixel's neighborhood.

Figure 5: Filtering

D. Segmentation

Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as super pixels). The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyze. Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images.

The result of image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. Each of the pixels in a region is similar with respect to some characteristic or computed property, such as color, intensity, or texture. Adjacent regions are significantly different with respect to the same characteristic(s). When applied to a stack of images, typical in medical imaging, the resulting contours after image segmentation can be used to create

3D reconstructions with the help of interpolation algorithms like marching cubes.

Compression based methods postulate that the optimal segmentation is the one that minimizes, over all possible segmentations, the coding length of the data. The connection between these two concepts is that segmentation tries to find patterns in an image and any regularity in the image can be used to compress it. The method describes each segment by its texture and boundary shape.

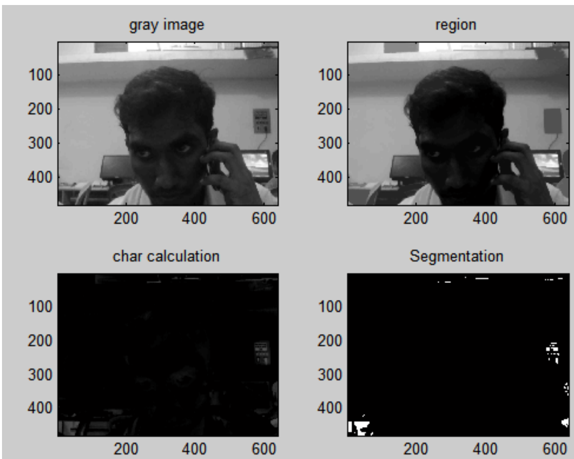

Figure 5. Segmentation

IV. Conclusion

In this paper, we have proposed K means algorithm for HOG feature extraction. By using this algorithm, the proposed HOG scheme can be implemented with smoothing and filtering concepts. It proves to be a good candidate for high-performance, and high-accuracy-rate for human detection. The proposed algorithm can be incorporated with other image processing algorithms, such as noise removal and image enhancement, on a single data set for more applications in an intelligent transportation system.

V. REFERENCES

[1] Pei-Yin Chen, Chien-Chuan Huang, Chih-Yuan Lien, and Yu-Hsien Tsai, "An Efficient Hardware implementation of HOG Feature Extraction for Human Detection", IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 15, NO. 2, APRIL 2014.

[2] S. J. Krotosky and M. M. Trivedi, "Person surveillance using visual and infrared imagery," IEEE Trans. Circuits Syst. Video Technol., vol. 18, no. 8, pp. 1096–1105, Oct. 2008.

[3] C. Papageorgiou and T. Poggio, "A trainable system for object detection,"Int. J. Comput. Vision, vol. 38, no. 1, pp. 15–33, Jun. 2000.

[4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. IEEE Conf. Comput. Vision Pattern Recog., Jun. 2005,vol. 1, pp. 886–893.

[5] K. Mizuno, Y. Terachi, and K. Takagi, "Architectural study of HOG feature extraction processor for real-time object detection," in Proc. IEEE Workshop Signal Process. Syst., Oct. 2012, pp. 197–202.

[6] Y. Pang, Y. Yuan, X. Li, and J. Pan, "Efficient HOG human detection," Signal Process., vol. 91, no. 4, pp. 773–781, Apr. 2011.

[7] T. P. Cao and G. Deng, "Real-time vision-based stop sign detection system on FPGA," in Proc. Digital Image Comput., Tech. Appl. Los Alamitos, Canberra, ACT, Australia, 2008, pp. 465–471, IEEE Computer Society.

[8] S. Bauer, U. Brunsmann, and S. S. -Macht, "FPGA implementation of a HOG-based pedestrian recognition system," in MPC-Workshop, Jul. 2009, pp. 49–58.

[9] M. Hiromoto and R.Miyamoto, "Hardware architecture for high-accuracy real-time pedestrian detection with CoHOG features," in Proc. IEEEICCVW, 2009, pp. 894–899.

[10] R. Kadota, H. Sugano, M. Hiromoto, H. Ochi, R. Miyamoto, and Y. Nakamura, "Hardware architecture for HOG feature extraction," in Proc. IEEE Conf. Intell. Inf. Hiding Multimedia Signal Process., Nov. 2009, pp. 1330–1333.

ELAVARASI. K PG Scholar, Department Of Computer Science & Engineering, IFET College of Engineering. Her research interests include data structures & algorithms, Image processing and big data.

Dr. R. KALPANA received B.E and M.E degrees from Madras University and Sathyabama University in 1991 and 2006 respectively and Ph.D. degree from Anna University Chennai, in 2013, all in Computer Science and Engineering. She is in professor in Computer Science and Engineering, IFET College of Engineering, Villupuram. Her research interests include Network Security, Wireless Networks, Adhoc Networks and Databases.