Person Detection Using Image Covariance Descriptor

Ms. Vinutha Raj B¹, Dr. M B Anandaraju²,

¹P.G Student, Department of ECE, BGSIT-Mandya, Karnataka
²Professor, Head of Department ECE, BGSIT-Mandya, Karnataka

Abstract -Person re-identification can be defined as finding the correspondences between the images of a probe set representing a single person and the corresponding images in a gallery set. Due to the complexity and the high performance requirement of person detection application, the design of embedded systems is the subject of different types of design constraints such as execution time, time to market, energy consumption, etc. Some methodologies were proposed in order to satisfy the different design constraints.

This paper presents a mixed HW / SW design methodology for the implementation of an image covariance descriptor. This implementation is intended for person detection and re-identification systems. It begins with a pure software implementation of the descriptor and the critical functions require a hardware implementation. Using softcores the process is carried out.

Keyword: Covariance Descriptor, Softcore Processor

I. INTRODUCTION

Identification techniques can be considered from different levels depending on information which are currently available in a video surveillance system. We take into account only the approaches which use camera to acquire information.

These approaches can be categorized into two main families:
1. The first category of methods uses biometrics
2. The second category is by appearance.

Appearance is based in three types of descriptors which are available to structure the different image features:

1. The first type is the global descriptors. It describes an image by using all pixels to calculate image characteristics. It is fast and simple to be employed. Also, the combination of global features can achieve good results.

2. The second type is the local descriptors. It characterizes a local point of interest with a vector of attributes. It is robust to occlusions, rotation, illumination and scale.

3. The third type is Region based descriptors. This descriptor decomposes the image into a set of regions and characterizes each of these regions. This kind of descriptors can be used to describe a region or the overall of the image. It is beneficial in classification. This type of descriptors is widely used in detection and re-identification applications.

The covariance descriptor is in general the best descriptor that is applicable for person detection. As a distance operator of covariance descriptor is computationally intensive, we propose a new FPGA-based implementation which significantly speeds up the computations. FPGA (Field Programmable Gate Array) provides an important computing power, a parallel calculation, a reconfiguration of the functional components and a low manufacturing cost. The software implementation of this descriptor on FPGA showed long execution times that does not meet the real-time constraint for video surveillance applications. Therefore, an appropriate SW/HW methodology is required to achieve the real time constraint.
Human detection concept is one of the outstanding and active research area in many applications like face recognition, facial expression recognition, face tracking, facial feature extraction, gender classification, identification system, document control and access control, clustering, biometric science, human computer interaction (HCI) system and many more. Face detection is often the first step in these applications. Many researchers have proposed different methods for addressing the problem of face detection by combining feature and color-based methods to obtain a high performance and high speed results.

To detect the human, here we use color based segmentation technique or localize a face region from both single and multiple face images. There are many of color space models, which are used to detect a face but for simplicity we will choose only three of them, which are RGB, HSV and YCbCr.

In case of color based segmentation, choosing an effective color space is very important because this may cause different output. Hence in this project we have considered the HSV technique to achieve better and good performance.

II. RELATED WORK

The literature survey is an important part of the project. It enables assimilation of knowledge required for the project right from the problem definition, finding a Solution for the same and its execution.

W. Ayedi, H. Snoussi, and M. Abid, in the year [2012] explained that, Person re-identification in the literature has been considered either as a on the fly or as an offline problem. More formally, person re-identification can be defined as finding the correspondences between the images of a probe set representing a single person and the corresponding images in a gallery set. Depending on the number of available images per individual (i.e., the size of the probe set), different scenarios have been addressed: (a) Single versus Single (S vs. S) if only one exemplar per individual is available both in probe and in gallery sets. (b) Multiple versus Single (M vs. S) if multiple exemplars per individual are available in the gallery set. (c) Multiple (M vs. M) if multiple exemplars per individual are available both in the probe and gallery sets.

S. Bak, in the year [2010] proposed that, in many surveillance systems there is a requirement to determine whether a given person of interest has already been observed over a network of cameras. This is the person re-identification problem. The human appearance obtained in one camera is usually different from the ones obtained in another camera. In order to re-identify people the human signature should handle difference in illumination, pose and camera parameters. We propose a new appearance model based on spatial covariance regions extracted from human body parts. The new spatial pyramid scheme is applied to capture the correlation between human body parts in order to obtain a discriminative human signature. The human body parts are automatically detected using Histograms of Oriented Gradients (HOG). The method is evaluated using benchmark video sequences from i-LIDS Multiple-Camera Tracking Scenario data set. The re-identification performance is presented using the cumulative matching characteristic (CMC) curve.

K. Loukil, N. Ben Amor, M. Abid in the year [2010] came across with the problem related to the hardware/software partitioning of reconfigurable multimedia system on chip. We proposed a new method of HW/SW partitioning. Indeed, our method is based on a static/dynamic mixed approach. For the static part, we used the design trotter tool and for the dynamic part we used the dynamic profiling of the application.

Walid AYEDI, Hichem SNOUSSI, Fethi SMACH and Mohamed ABID in the year [2006] explained as, Object matching is the process of determining the presence and the location of a reference object inside a scene image. Matching accuracy requires robust image description and efficient similarity measures. In this paper, we present a tree based object matching approach using a descriptor proposed in a previous work. Visual objects are described by a collection of multi-scale covariance matrices structured in a tree form. Tree matching is then performed to match visual objects. With this approach, matching accuracy considerably increases compared to traditional image matching techniques. The proposed matching approach is evaluated on CAVIAR.
dataset. Overall, our approach is an important contribution to a complete system for object reidentification and tracking over different camera views.

III. METHODOLOGY

Visual feature techniques is used to detect a persons in accurate and efficiently. The HSV is one of the most common methods used for detection of person in any environment.

The differences of colour distribution in image frames provide important values that could be used to group similar frames together to detect oneshot. The extracted features are given to gabor filter, then applied to covariance descriptors. Finally, concatenation of the result from descriptor gives the desired output.

---

**Figure 1.** Flowchart of the proposed approach: (1) color images are split into three color channels (HSV), (2) for each channel, Gabor filters are computed at different scales, (3) pairs of neighboring scales are grouped to form one band, (4) magnitude images are produced by applying the MAX operator within the same band, (5) magnitude images are divided into small bins and each bin is represented by a covariance descriptor, and (6) the difference of covariance descriptors between two consecutive bands is computed for each bin and concatenated to form the image representation.

---

**Figure 2.** A pair of images and their BIF Magnitude Images. From left to right the original image, its three HSV channels, six BIF Magnitude Images for different bands.

---

The extracted bicov features are then implemented in hardware to accelerate system execution.

IV. IMPLEMENTATION

The implementation of the proposed approach undergo following steps:

1. Color images are split into three color channels RGB, HSV and finally converted as gray scale images.
2. Images features are extracted using covariance descriptor.
3. The features extracted are gradient function, magnitude, normalization value of X and Y, integral value of P and Q
4. Image is resized and then simulated on modelsim.
5. Once simulation is done the obtained result is synthesized using xilinx.
6. The implementation is done on vertex-4 board using softcore processors.
Softcore belongs to FPGA synthesis that is implemented using logic synthesis. In designing softcore FPGA has to be split up into three main structures

- Defining the co-efficients
- Profiling
- Designing of hardware

Mixed approach allow avoiding complex programming using HDL, through the instantiation of a processor core in the reconfigurable device. These methodologies provide a gain of time development.

The hardware implementation process is shown in figure 3

![Figure 3. Hardware architecture](image)

**V. CONCLUSION**

In this paper we implemented an image covariance descriptor using an FPGA mono processor platform. This implementation is intended for person detection and reidentification systems. Here we use a mixed HW/SW design methodology. The processing time of the covariance descriptor is considerably reduced while using the hardware master accelerator. Large rotation and illumination changes are also absorbed by using covariance matrix. This method also reduces the loss of information during processing of an image by using covariance descriptors and softcore processors.

For future work we aim to design and implement the complete covariance descriptor based person detection system providing unique value for every detected object, while meeting real-time, cost and energy consumption constraints.

**REFERENCES**


[5]. Walid AYEDI, Hichem SNOUSSI, Fethi SMACH and Mohamed ABID “Tree based object matching using multi-scale covariance descriptor”Charles Delaunay Institute (FRE CNRS 2848), University of Technology of Troyes, 10010 Troyes, France Sfax University, National Engineering School of Sfax, 3052 Sfax, Tunisia,October 2011.


European conference on Computer Vision-Volume Part II
Pages 589-600, 2006.