People Detection, its Applications in learning-based computer vision

Nom: Nagi OULD TALEB ,Adil CHERGUI, Mohamed Larbi BEN MAATI, Mohamedade Farouk NANNE
Corresponding author Email: nagismi@gmail.com

Abstract: Detecting humans in video/images is a challenging task, with many applications that has attracted lot of attention in recent years. The area of computer vision is concerned with the recognition of activity from video sequences. It's application techniques are very numerous. There may be mentioned the video surveillance, power management and evaluation of advertising policies. Several algorithms have been developed to achieve very good results.

Some approaches are based on motion information and other are focused on appearance information. Considering the huge demand existing in the area, it would be very important to talk about the learning-based computer vision.

In this article, we will discuss the main techniques of person detection on video sequence and present an overview oncomputer vision routines based on learning. After that we will describe in a clear manner the steps for create a cascade of classifiers based on Haar-like features.

I- Introduction

Computer vision is an exciting topic of research to develop the computer systems analysis and interpretation capabilities of the visual contents of a scene close to those of human vision. Recognition techniques and activity analysis of human behavior from an observed scene remain a major challenge. This task remains problematic given the wide variations in how to perform the activities, the appearance of the person and changes in vesting conditions.

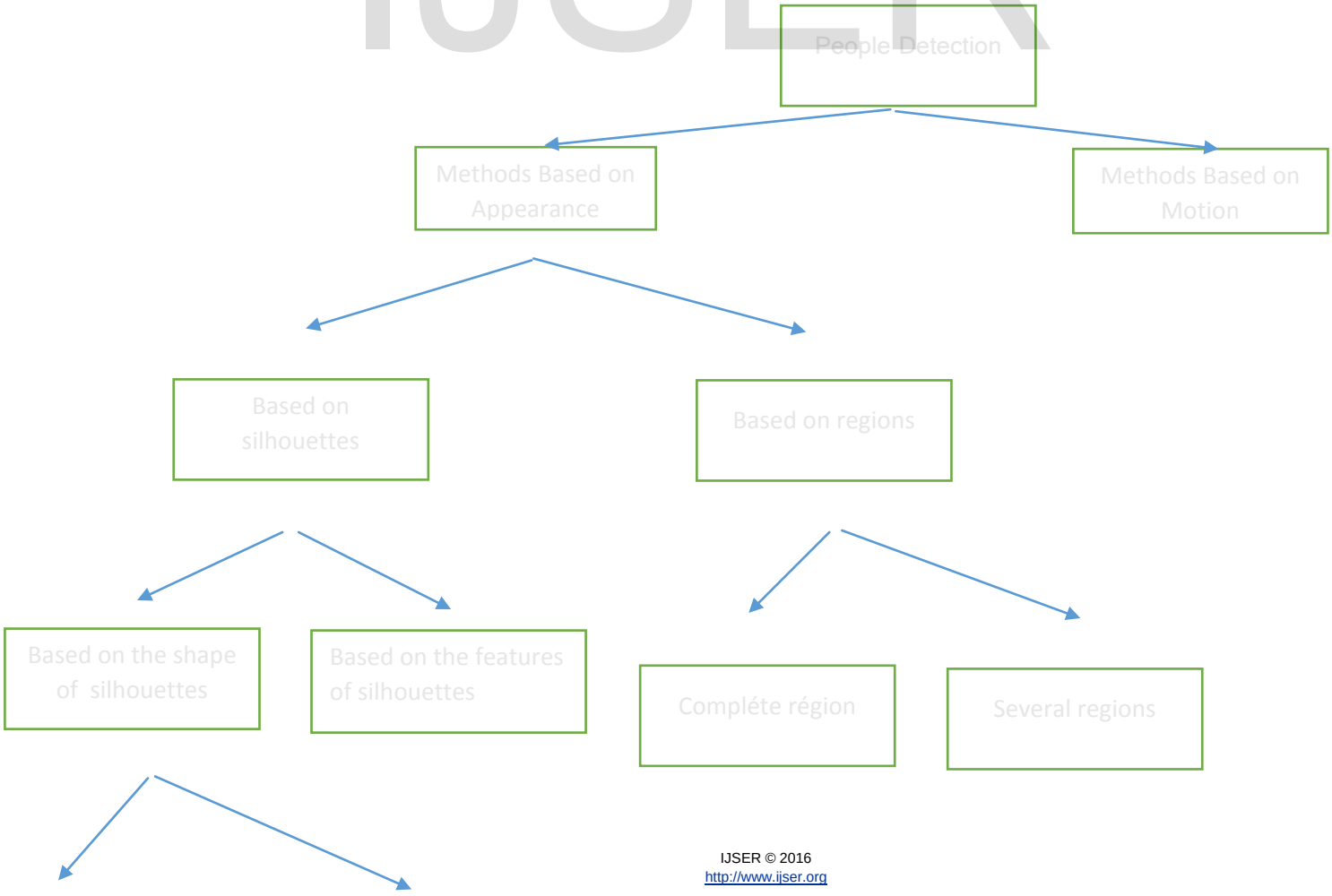
We will present an overview of the effective recognition methods that are robust to different variability factors. Before we begin the analysis stage, it is essential to go through the detection and identification steps of the human shape. We can see that the people present a big variety of physical appearances, poses, movements and interactions between different people and objects. In this sense, the automatic people detection is a complex problem because of the huge variability in people appearance and motion. Several methods have been developed to be able to detect a human in different situations. Many constraints

expand detection approaches: a camera movable or fixed, static or dynamic background, multi-scale and multi-detection.

Classical approaches pedestrian detection use the pattern recognition and statistical learning. They therefore suffer performance degradation when theappearance of pedestrians or elements of the scene is too different from the one studied during learning.

Some approaches have built detectors based on motion information or detectors based on appearance information but we can see a method that combines both sources of information in a single detector.

There are various techniques designed to detect people/pedestrian in a frame sequence. This allows us to draw the following classification of these methods (Fig1).



The remainder of the paper is organized as follows. In section 2 we review the current state of the people detection approaches. In Section 3 we give a display of some well-knowing learn-ing-based techniques. We conclude in Section 4.

II- Related Approachs

Object detection in images and videos has received a lot of attention in the computer vision andpattern recognition communities in recent years. It is actually a complexproblem with multiple applications, not only in video surveillance, but also different areas like intelligent systems (robotic), video games, etc. Currently many different systems exist which try to solve this problem. The first people detection method classification consists of methods based on people appearance or methods based on people movement analysis. Sidenbladh [1] has proposed a method that detects humans based on motion patterns and used Support Vector Machine. Another approach [2] based (founded) on motion information proposes an object classification system based on periodic motion analysis. We can quote Phantom [3] which is a people-tracking system. It determines their positions in 3D using monochromatic stereo imaging. Another method was conducted [4,5,6,7] for person's

detection with a measure of their activity in video sequences.However, most of the existing approaches use people appearance information to decide if the detected objects are people or not. In this case, we can see two different types of algorithms:

- The algorithms based on people silhouettes, which are based on their contour.we can classify them as follows:

- Methods Based on Features of Silhouette

For the detectors based on features of silhouette F. Xu and al proposed a method [8] which not use background subtraction, but uses depth information to segment objects. Fernández-Carbajales [9] proposes a people detection system based on the fusion of three simple independent people detectors. Each detector is separately applied to each detected blob and the final decision is a combination or fusion of those three evidences.

- Methods Based on Silhouette's Shape

Methods based on the shape of the silhouette are divided in two branches. On the one hand those which are based on a model of person as a complete silhouette [10,11,12], and on the other hand those which are based on a model of person as the union of parts of the same silhouette [13,18].

- Complete Silhouette

Defines a people detection system [10] using a trained codebook of people shapes in order to classify between humans and other objects.

- Parts of Silhouette

Haritaoglu and al [13] describes an approach to locate body parts of people using silhouettes. The people model consists of six primary body parts (head, two hands, two feet and torso) and ten secondary parts (elbows, knees, shoulders, armpits, hip and upperback).

- The algorithms based on the regions that represent the person. we can classify them as follows:

- Complete Region

Viola and al [19] presents a pedestrian detection system that integrates image intensity information with motion information. CUI and al [20] proposes an extension of the previous algorithm. In this case the algorithmdefines seven types of volume filters in the 3D space, instead of using rectangle filter in the 2D space.

- Several Regions

Sprague [21] defines a people detection system based on color segmentationand cloth people detection. Harasse and al[22] proposes a people detection method based on human model of three body parts and color information to front and side views.

III- Learning-Based Computer Vision

A- OpenCv

OpenCV is a free Open Source Computer Vision Library.Th e library is written in Cand C++ and runs under Linux, Windows and Mac OS X. Th ere is active development on interfaces for Python, Ruby, Matlab, and other languages. It has strong and growing support for learning-based vision. OpenCV has played a role in the growth ofcomputer vision by enabling thousands of people to do more productive work in vision. With its focus on real-time vision, OpenCV helps students and professionals efficientlyimplement projects and jump-start research by providing them with a computer vision and machine learning infrastructure that was previously available only in a few mature research labs.

B- Creating a Cascade of Haar-Like Classifiers

The technique has used in computer-visionfor faceand eye detection.OpenCV comes now already with a trained classifier for frontal face and eyedetection. We will show that we can create a classifier for whatever object. These training steps to create a Haar-like Classifier are organized as follows:

a- Collection of positive and negative training images.

You should have a lot of positive and negative sample images for training.

All the required tools and positive/negative image dataset are provided here:

https://www.cs.auckland.ac.nz/~m.rezaei/Tutorials/Haar-Training.zip

The positive images are those images that contain the object (Human body, face or eye, etc.), and negatives are those ones which do not contain the object.Having more number of positive and negative (back ground) images will normally cause a more accurate classifier.

To create the positive images, we can useobjectmarker.exe or ImageClipper tools.

b- Creating a vector (.vec) file based on positive marked images using createsamples.exe

We need to create a data file (vector file) that contains the names of positive images as well as the location of the objects in each image. We can use the command:

opencv_createsamples.exe -info positive/info.txt -vec vector/ahumanvector.vec -num 8 -w 24 -h 24.

The parameters num, w and h mean successively Number of positive files to be packed in a vector file, width of objectsand heightof objects.

c- Training the classifier using haartraining.exe

In this step we can use this command:

haartraining.exe -data cascades -vec vector/human.vec -bg negative/image.txt -npos 150 -nneg 150 -nstages 15 -mem 1024 -mode ALL -w 24 -h 24 –nonsym

-data cascades	Path and for storing the cascade of classifiers
-vec data/vector.vec	Path which points the location of vector file
-bg negative/bg.txt	Path which points to background file
-npos 150	Number of positive samples ≤ no. positive bmp files
-nneg 150	Number of negative samples (patches) ≥ npos
-nstages 15	Number of intended stages for training
-mem 1024	Quantity of memory assigned in MB
-mode ALL	Look literatures for more info about this parameter
-w 24 -h 24	Sample size
-nonsym	Use this if your subject is not horizontally symmetrical

To run haartaining.exe you also needs the files cv097.dll, cxcore097.dll, and highgui097.dll

d- Creating the XML File

After finishing Haar-training step, in folder ../training/cascades/ you should have. catalogues named from “0” upto “N-1” in which N is the number of stages you already defined in haartraining.bat.

In each of those catalogues there should be AdaBoostCARTHaarClassifier.txt file.

Now you have your own XML file, run your detection program.

IV- Conclusion

In this paper, we presented an overview on human detection methods in an image/video. We give a particular interest to learning-based techniques. Through this article, we have commented some advantages and disadvantages of different approaches to solve the people detection problem in video sequences. After that, we talked about the importance of computer vision. Finally we have technically shown a way to create a Haar-like Classifier.

References

[1] H. Sidenbladh, “Detecting human motion with support vector machines,” in International Conference on Pattern Recognition (ICPR), 2004.

[2] R. Cutler and L. S. Davis, _Robust real-time periodic motion detection, analysis, and applications," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22(8), pp. 781_796, 2000.

[3] J.Leskovec, Sentjost, and Slovenia "Detection of Human Bodies using Computer Analysis of a Sequence of Stereo Images". 11th European Union Contest for Young Scientists, 1999.

[4] P.Viola, M.J. Jones and D. Snow "Detecting Pedestrians Using Patterns of Motion and Appearance", in International Journal of Computer Vision, pp. 153–161, 2005.

[5] Benezeth, Yannick, and al. "Humain Presence Detection and Charactezation of Activity", In XXIIe conference GRETSI (signal and image processing), Dijon (FRA), 2009.

[6] L.Biancardini,S.Beucher,L.Letelllier "Extraction des Objets en Mouvement : une approche mixte contours-régions", ORASIS 2005, 9ème Congrès Jeunes Chercheurs en Vision par Ordinateurs, Clermont-Ferrand, France, mai 2005.

[7] S. Zhang, et al. "*How Far are We from Solving Pedestrian Detection", Cornell University Library,* (2016).

[8] F. Xu and K. Fujimura,"Human detection using depth and gray images,_"in Conference on Advanced Video and Signal Base Surveillance (AVSS), pp. 115_121, 2003.

[9] V. Fernández-Carbajales, M. A. García, and J. M. Martínez, _Robust people detection by fusion of evidence from multiple methods,_ in WIAMIS '08: Proceedings of the 2008

Ninth International Workshop on Image Analysis for Multimedia Interactive Services , pp. 55_58, 2008

[10] J. Zhou and J. Hoang, _Real time robust human detection and tracking system,_ in Computer Vision and Patter Recognition (CVPR), p. 149, 2005.

[11] M. Hussein, W. Abd-Almageed, Y. Ran, and L. Davis, _Real-time human detection, tracking, and veri_cation in uncontrolled camera motion environments,_ in International Conference on Vision Systems (ICVS), p. 41, 2006.

[12] N. Koenig, _Toward real-time human detection and tracking in diverse environments,_ in International Conference on Development and Learning (ICDL), pp. 94_98, 2007.

[13] I. Haritaoglu, D. Harwood, and L. S. Davis, "Ghost: A human body part labelling system using silhouettes," in International Conference on Pattern Recognition (ICPR), pp. 77_82, 1998.

[14] P. Viola, M. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," in International Conference on Computer Vision (ICCV), 2003.

[15] X. Cui, Y. Liu, S. Shan, X. Chen, and W. Gao, "3d haar-like features for pedestrian detection," in International Conference on Multimedia & Expo (ICME), pp. 1263_1266, 2007.

[16] N. Sprague and J. Luo, "Clothed people detection in still images," in International Conference on Pattern Recognition (ICPR), 2002.

[17] S. Harasse, L. Bonnaud, and M. Desvignes, “Human model for people detection in dynamic
scenes, “ in Computer Vision on Pattern Recognition (CVPR), pp. 335_354, 2006.

[18] B. Wu and R. Nevatia, “Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detector,” in International Conference on Computer Vision (ICCV), pp. 90_97., 2005.

[19] P. Viola, M. Jones, and D. Snow, “Detecting pedestrians using patterns of motion and appearance,” in International Conference on Computer Vision (ICCV), 2003.

[20] X. Cui, Y. Liu, S. Shan, X. Chen, and W. Gao,” 3d haar-like features for pedestrian detection,” in International Conference on Multimedia & Expo (ICME), pp. 1263_1266, 2007.

[21] N. Sprague and J. Luo,” Clothed people detection in still images,” in International Conference on Pattern Recognition (ICPR), 2002.

[22] S. Harasse, L. Bonnaud, and M. Desvignes, “Human model for people detection in dynamic
scenes, “ in Computer Vision on Pattern Recognition (CVPR), pp. 335_354, 2006.