AUTOMATIC PEDESTRIAN DETECTION SYSTEM – SUPPORT FOR AUTONOMIC COOPERATIVE SYSTEMS

Title Ilona Madalina Costea
Position Lecturer
Institution University "Politehnica" of Bucharest

Title Catalin Dumitrescu
Position Research Associate
Institution University "Politehnica" of Bucharest

Title Claudiu Virgil Banu
Position Ph D Student
Institution University "Politehnica" of Bucharest

Title Florin Codrut Nemtanu
Position Senior Lecturer
Institution University "Politehnica" of Bucharest

Abstract

The present paper purposes and analyses a automatic pedestrian detection system. Pedestrian detection by a vehicle will be based on specific infrastructure cooperative systems, the transmission of information to all vehicles near pedestrian V2V and V2I communications through. Applying concepts and solutions are imminent required in recognizing the pedestrian located in danger areas (street or intersection) and transmitting of informations to other vehicles and equipments in related infrastructure (road side units).

1. STATE-OF-ART

Pedestrian detection based on optical systems is a challenge in multiple ways. Pedestrians appear in highly crowded environments and they have a wide range of appearances, from the different poses, body size, clothing and outdoor lighting conditions. They can be located in quite large distances to the camera and appear quite small in the images. In addition it is not possible to use the background extraction methods to obtain the regions of interest that may include pedestrians due to camera shake and the restrictions imposed by the real-time requirements. So far, various methods have been developed to detect a possible candidate for pedestrian and subsequent recognition as a person. For the detection phase motion are used methods as movement detection, applying thresholds on images, or forms-based detection, while walking analysis and patterns of shaped are part of the recognition stage. The movement is the key to detect regions of interest, which may contain pedestrians and relies on temporal information. Also, we need to analyze a sequence of frames. Forms detection method based on direct analysis performed on the entire image on previous information about the shape of the human body. This approach avoids any problem of motion of the background and rarely integrates temporal information in a sequence of analysis.

The frequency of the human walking is an important clue for the recognition of pedestrians moving, especially when they go sideways to the direction of view. To make the system be able to provide a concrete answer, it must be analyzed a number of frameworks and a follow-up of pedestrian is required. Studying methods of pedestrians walking depend greatly of the side of view and can not detect pedestrians stationary.
The systems that implements shaped patterns for the recognition of pedestrians do not need validation of temporal information and can provide stationary of pedestrians. This approach is more sensitive to false teaching and needs a good detection phase.

COST EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

2. THE INITIAL CONDITIONS AND IMPEDIMENTS

In the pedestrian detection techniques, the main problems are:

The appearance of passersby. A large variability in the local and global appearance of passersby may be caused by different types and styles of clothing. Therefore, a technique that is based on a certain characteristic, such as appearance of the legs may not be applied to all classes of pedestrians, such as women wearing skirts.

Exposure the passer. A passerby is not a rigid body, for this reason the shape will go through a variety of changes due to the variety of possible positions which can be assumed.

Orientation of the passer. A passerby can be viewed from a variety of possible directions regarding video camera image. For example, pedestrians may face the camera directly, ie at 0 degrees or parallel to it, ie ± 90 degrees.

The position of passersby. A passerby can be placed in a scene at different distances from the camera. The emergence of the passer near the camcorder may vary significantly from a greater distance.

Occlusion its own. The silhouette of a passerby can also be disrupted by a variety of accessories such as backpacks, hats, bags and shopping bags.

Occlusion in groups. A passerby may be covered by one or more pedestrians or objects, especially if the passer is situated in a crowd.

The vision of the camera. A passerby can be partially in the field of view of the camera. This may indicate that it is a passerby who enters or leaves the scene. This can occur in several areas within the scene.

Non-pedestrian objects. Not all foreground objects in a scene can be pedestrians, even if the shape and size are similar. It is difficult to predict and to develop models for all other real-world objects that may appear in a scene.

Environmental conditions. Detecting pedestrians in real time scenarios brings additional difficulties due to the movement, changing backgrounds, brightness and weather conditions due to reflections from windows and shadows of pedestrians and other foreground objects.

Variation of intensity. There may be a lack of variation in color intensity between clouds, background and foreground objects, leading to a difficult image segmentation.

The aspects presented so far represents significant challenges for pedestrian detection techniques in single frames. However, the goal of most artificial vision applications, such as those used for security, requires information not only about the location of people at some point in time, but what these people do. More useful information about the passerby trajectory can be obtained when persons are tracked over several scenes.

However tracking passers introduces new challenges:

The movement of pedestrians. A pedestrian can go in a nonlinear and unpredictable way, and this is leading to a difficult tracking of it. For example, in real world scenarios a pedestrian can walk, run, stop or turn unexpectedly.

Occlusion. One or more paths can be merged into one. Depending on the tracking technique, where two or more paths overlap, it is possible that one of them be lost temporarily. Tracking algorithm must be able to consider this, so that after overlapping each trajectory to maintain the proper object (the one pursued before occlusion).

The division. A path can be divided into two or more parts.This can occur due to poor segmentation within the current framework , poor segmentation in previous frames, or possibly the introduction of any object in the scene.

The new trajectories. Passersby who enters the scene must be recognized and not confused with pedestrians previous track.

Lost trajectories. A path may be lost completely due to noise, poor segmentation of the object of interest or exit the visibility of the camcorder. The path loss would be marked as inactive and not confused with other trajectories.

3. PEDESTRIAN DETECTION BASED ON FORMS

Forms-based approach attempts to solve difficult issue of pedestrian recognition with unique images, treating pedestrians moving and also those stationary. The biggest challenge that launches this problem is the huge amount of variation modeling of shapes, positions, size and stature of men. Knowing the extrinsic parameters of the system it can be use fixed scenes for reducing the candidates searching in a certain limited part of the image. After this low level processing, the next step is to make an analysis of vertical symmetry maps derived from the gray levels of the image horizontal gradient. Then there is the identification of regions that can be characterized as human shapes. This stage produces frames with high probability to characterize a pedestrian. Finally, with filteres is removed a part of the candidates which actually do not represent the human shapes, as shown in Figure 1.

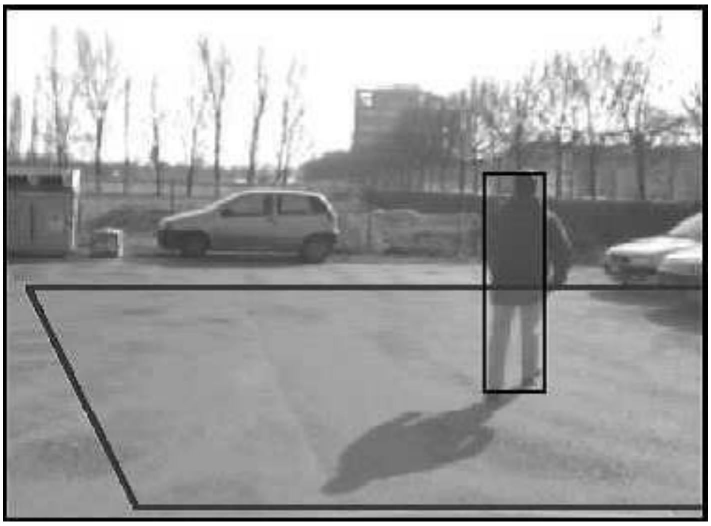


Figure 1. The result of the algorithm for an example captured from a vehicle in motion

In the case of pedestrian detection systems based on infrared images there are used, first, images represented by pixels. For the classification of target objects as a pedestrian or not, is used intensity of the value of each pixel.Target objects are extracted from scenes by simply filtering threshold. Pixels belonging to the background are assigned a value of 0 and pixel corresponding to heat emitting objects as a pedestrian are assigned a value of 1. It uses a training set to develop a probabilistic model that consists of a set of 1000 rectangular images 128x48 , all containing people. Each template is translated so that the center of gravity of the non-zero pixels to coincide with the geometrical center of the image. For each pixel in the template is calculated the probability p (x, y), that characterize a pedestrian, based on how often appears as 1 in the training set. For each pixel the probability that it has the correct classification, so that the window contains a pedestrian or not, is p(x, y), where the pixel has the value 1 and 1-p (x, y), where the the pixel has the value 0. The sum of these probabilities is calculated for all pixels and this gives the combined probability to contain in the given window a pedestrian or not, see Figure 2.

Figure 2 Probabilistic model and output frames (heads of pedestrians are marked by blue contours)

Presence Detection Using Full Image

The proposed method for presence detection based on Full Image (DPII) uses local properties taken from the image instead of working directly with the pixels in the image. The background estimate is an image of properties. For each frame, the image of properties must be calculated and compared with the background. The advantage of working directly with images is the speed, the only deficiency being the fact that it only accounts for a single pixel in differentiating, with no regard to its neighborhood. Starting from the fact that, in so many cases, we are not interested in only one pixel or very small objects, we develop a method which considers the pixel and its neighborhood, removing the shadows.

The proposed technique considers a pixel as belonging to the objects of interest, if most of the neighboring pixels belong to the object of interest. Rapid determination of the pixels' nexus is achieved by using rectangular properties. The simplest property fulfilling all the conditions is the amount of pixels in a region, which is similar to an average of these pixels. This amount can be calculated very quickly using the Full Image.

The Algorithm Based on the Method of Background Composition

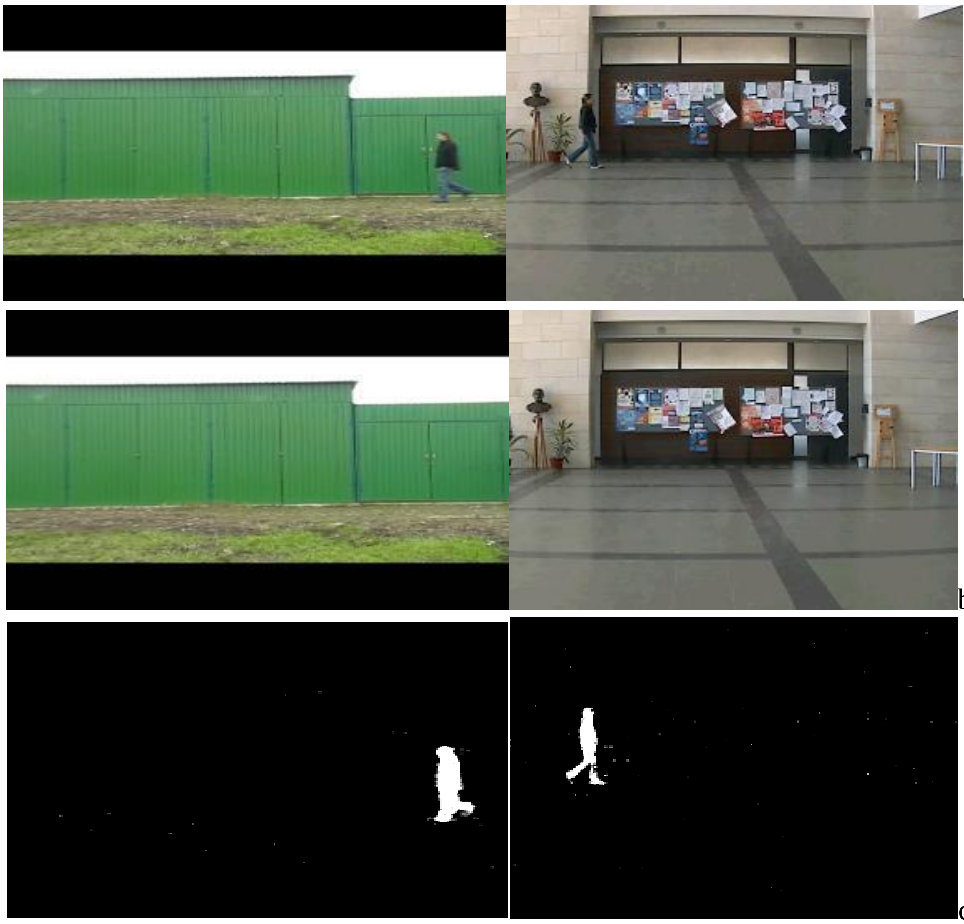
Most of the algorithms in this category are very complicated and very costly in terms of deployment. The present work introduces a simpler algorithm obtained by starting from a classic solution of deployment, but with very good results.

The algorithm starts from a frame with a RGB format at a 24 bpp color depth, hereinafter called \`current image\`. The grayscale version, obtained just like the previous algorithm – by a RGB-YUV conversion and dropping the U and V components, it will be called \`current frame\`. Moreover, there will also be an image called \`background\` (background frame), which will also be a greyscale version; this will be the frame to which the new frames are compared to.

At first, the first frame content will be copied directly in the background frame. Afterwards, the new frames will be compared to the first one. However, a problem arises in frame modification (adding or deleting objects) which will determine the alarm alert until the reboot of the algorithm.

The problem can be solved by updating the background (frame), not suddenly, but progressively, so that this frame will tend towards the current frame (in terms of content). This inclination towards the current frame can be achieved by changing the light intensity with one or more units in order to reduce the difference between the amounts of pixels at the same coordinates, at every frame received.

The computing power for this approach is quite high. In order to reduce it, some filtering may be made, to reduce the number of colors in the acquired frames, such as color expansion.



a.

b.

c.

Figure 3. Examples for presence detection in outdoor and indoor environments with the integrated image-based method; a) current image, b) estimated background, c) generated mask.

4. DETECTING PEDESTRIANS USING THE PERIODICITY OF HUMAN MOVEMENT

This method treats scenes involving people who walk sideways from the direction of view. The frequency of human walking is an important clue for recognizing pedestrians that are walking. The recognition is based on the movement characteristic of the legs of a pedestrian that moves parallel to the image plane. Each image is segmented into regions by grouping pixels into a combined color-position. The grouping method associates groups of consecutive frames, allowing them to be tracked along a sequence of frames. Based on the observation of groups over time, a classifier extract those groups of pixels that represents most likely legs of a pedestrian. A polynomial classifier performs a rigorous selection of groups by assessing temporal changes of group characteristics dependent on shape. The final classification is performed using a neural network trained to detect human walking.

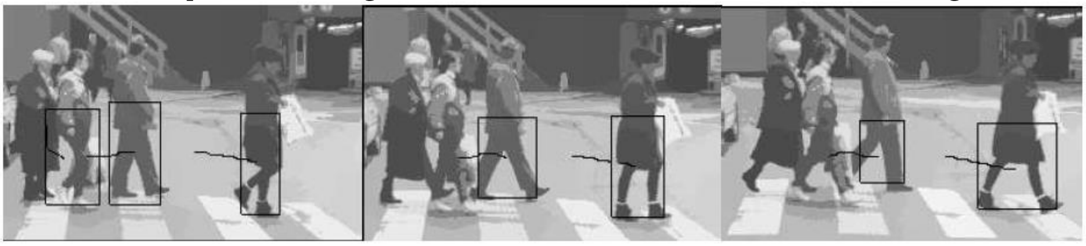
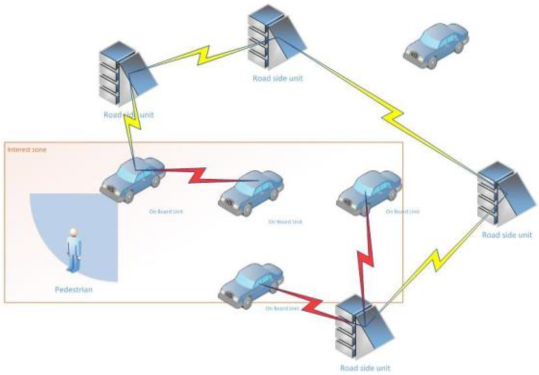


Figure 4. Sequence of images of pedestrians. Groups recognized as pedestrians are marked by borders

4. CONCLUSION

Pedestrian detection based video systems remains an open research. People merges with the background, are walking or stationary, walking direction changes unexpectedly. The background is complex and sudden changes are unavoidable in video systems mounted on moving vehicles. Different approaches have been developed to treat these complex things. Some systems have a proven track record in testing on the road, which encourages further research. Future researches involves the segmentation of regions background / foreground which may contain pedestrians, accurate color detection and analysis of video sequences taken by a camera mounted on a vehicle.



Pedestrian detection by a vehicle will be based on specific infrastructure cooperative systems, the transmission of information to all vehicles near pedestrian V2V and V2I communications through. Applying concepts and solutions are imminent required in recognizing the pedestrian located in danger areas (street or intersection) and transmitting of informations to other vehicles and equipments in related infrastructure (road side units). Based on cooperative technologies, blind vehicles (in terms of technologies described in this paper) may have information about the objects or pedestrians in certain areas of interest. Autonomous properties characterize both equipment installed on the vehicle (on-board units) and also those that are in the road infrastructure.

AKNOWLEDGEMENTS

The work has been funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/134398.

REFERENCES

[1]. A.J. Hugo, "Process Controller Performance Monitoring and Assessment", http://www.controlartsinc.com/Support/Articles/PerformanceAssessment.PDF, 2002

[2]. GARG V., Wireless communications and networking, ed.1, Morgan Kaufmann Series in Networking, USA, 2006

[3]. *KEITHLEY Instruments*, SUA, ADwin DSP System, Technical Data, 2004

[4]. BOVIK A.C., The Essential Guide to Video Processing, Academic Press, USA,
ISBN: 978-0-12-374456-2, 2009

[5]. *STERN R., ACERO A., LIU F.H., and OHSHIMA Y.*, Signal processing for robust speech recognition

[6]. *CHIBELUSHI C. C.,DERAVI F. , and MASON J. S. D.*, A review of speech-based bimodal recognition. IEEE Trans. Multimedia, 4(1):23–37, 2002