

A Symmetry-based Validator and Refinement System for Pedestrian Detection in Far Infrared Images

M. Bertozzi, A. Broggi, M. Del Rose and M. Felisa

Abstract—This paper presents a refinement and validation stage for a tetra-vision based pedestrian detection system. The complete system is based on the use of both visible and far infrared cameras; in an initial phase it produces a list of areas of attention in the images which can contain pedestrians. Starting from the assumption that human shapes are mostly symmetrical and feature a number of vertical edges, the symmetry and edge density content of these areas is computed. The results is used to filter the list, discarding areas of attention that feature a scarce symmetry content. At the same time, the size of each area is refined to match the symmetry content and therefore to better adapt to human shapes. Moreover, this module can be successfully used to partition areas that contains two or more human shapes.

I. INTRODUCTION

A reliable pedestrian detection system is interesting for a number of applications, such as video surveillance or automatic driver-assistance systems for vehicles; moreover, the heavy investments made by almost all car manufacturers on this kind of research proves that particular attention is now focused on improving road safety, especially for reducing the high number of pedestrians being injured every year.

Also military organizations, like the U.S. Army, are actively involved on this research topic for robotic applications like obstacle detection, automatic path following, or anti-tamper surveillance [9], [13]. For military applications, the use of vision as a primary sensor for the detection of human shapes is widely exploited, since cameras do not emit signals that can be detected by third parties.

Unfortunately, vision-based pedestrian detection, especially when performed from a moving vehicle, is a difficult task. Pedestrians are *deformable* obstacles and can therefore assume several different poses and wear different clothes with different colors changing their appearance. Moreover, the automotive environment is often barely unstructured, incredibly variable, and apparently moving, due to the fact that the camera itself is in motion; therefore, really few assumptions can be made on the scene.

This paper describes a tetra-vision based pedestrian detection system to be installed on an autonomous military vehicle.

This system is able to detect all obstacles appearing in the scene and is based on the simultaneous use of two stereo camera systems: two far infrared cameras and two daylight

cameras [5]. The first stages of such system provide a reliable detection of image areas that potentially contain pedestrians. Unfortunately, some of the detected areas do not contain pedestrian or contain more than one pedestrian; therefore a refinement and a preliminary filtering of these areas is mandatory. This refinement and filtering stage is based on the assumption that a pedestrian features a high symmetry and edge density content. The process is only performed on far-infrared images in order to reduce the impact of the background on the whole process. Following filters are used to furtherly analyze the list of refined areas of attention in a multi-vote fashion. A final validation stage inputs all the votes coming from the different filters and evaluates how much the obstacle is likely to be a pedestrian.

This paper details the symmetry-based subsystem for the refinement and filtering of the list of areas of attention and

is organized as follows: section II describes related work in pedestrian detection systems based on artificial vision. The whole pedestrian detection system is briefly discussed in paragraph III while the symmetry based resizing and filtering is detailed in section IV. Finally, section V ends the paper presenting few results and remarks on the system.

II. RELATED WORK

Vision-based systems for pedestrian detection have been developed exploiting different approaches, like the use of monocular [18], [14] or stereo [19], [22] vision. Daylight cameras are widely used for pedestrian detection, but, in the last years, thanks to the decreasing costs and to the improved technology, also infra red devices have been employed for vision based pedestrian detection [8], [3].

Many systems based on the use of a stationary camera employ simple segmentation techniques to obtain foreground region; but this approach is not suitable when the pedestrians has to be detected from moving platforms. Most of the current approaches for pedestrian detection using moving



Fig. 1. Examples of typical scenarios in FIR and visible images.

This work has been supported by the European Research Office of the U. S. Army under contract number N62558-05-P-0380.

M. Bertozzi, A. Broggi, M. Felisa are with the Dip. Ing. Informazione, Università di Parma, ITALY. <http://www.vislab.it>

M. Del Rose is with U.S. Army TARDEC, U.S.A. <http://www.tacom.army.mil>

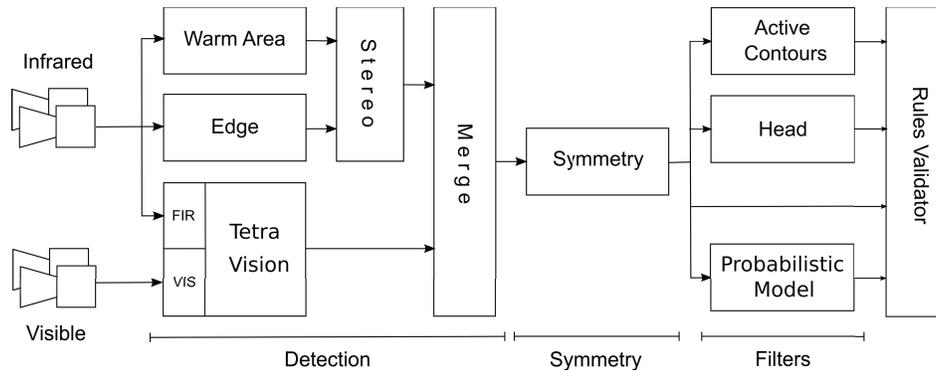


Fig. 2. Overall algorithm flow.

cameras treats the problem as a recognition task: a foreground detection is followed by a recognition step to verify the presence of a pedestrian. Some systems use motion detection [7] or stereo analysis [22], [20] as a means of segmentation.

Other systems, substitute the segmentation step with a focus-of-attention approach, where salient regions in feature maps are considered as candidates for pedestrians; a following recognition approach is used to validate the presence of pedestrians. In the GOLD system [5] a DSI-based approach is used to detect potential pedestrians and the recognition is obtained by a number of independent validators. A similar approach is used in [11] where region of interests are preliminary selected using a stereo vision approach and following shape-based and texture-based stages are used for the validation. In [6] the local image entropy directs the focus-of-attention and is followed by a model-matching module.

For what concerns the recognition phase, recent research are often motion-based, shape-based or multi-cue based. Motion based approaches use the periodicity of human gait or gait patterns for pedestrian detection [7], [23], [1]. These approaches seems to be more reliable than shape based ones, but they require temporal information and are unable to correctly classify pedestrians that are still.

Shape-based approaches are based on pedestrians' appearance, so can detect both moving and stationary people [14], [17]. In these cases, the challenge is to model the several variations in the shapes, pose, size and appearance of humans and their backgrounds. Basic shape analysis methods consist in matching a template onto candidate foreground regions. In [10] a tree-based hierarchy of human silhouettes is constructed and the matching follows a coarse-to-fine approach. In [15], [21] probabilistic templates are used to take in account the possible variations in human shape.

Symmetry assumptions are used as well, in [16], [17] combination of symmetry and edge density features are extracted and fed into an SVM-classifier using a statistical approach. Also the use of temporal tracking of symmetry features seems promising but, analogously to the previous mentioned motion-based recognition systems, it fails to detect still pedestrians [12].

III. SYSTEM DESCRIPTION

The symmetry-base filtering and refinement module described in this work has been developed and is part of a more complex system for the detection of pedestrians in far infrared and daylight images [4], [5]. The whole architecture is based on the simultaneous use of two stereo systems; the first system is composed by two CMOS daylight cameras, while the latter comprises two far infrared cameras in the $7-14\mu$ spectrum. Thanks to this solution, the system is able to detect obstacles and pedestrians when the use of infrared devices is more appropriate (night, low illumination conditions...) or, conversely, in the case visible cameras are more suitable for the detection (hot, sunny environments...).

In fact, in the far infrared domain, the image of an object is quite different from the same image in the visible spectrum and depends on the amount of heat it emits, namely, it is generally related to its temperature (see fig. 1). Conversely, in the visible domain, objects appearance depends on how the surface of the object reflects the incident light as well as on the illumination conditions.

Usually, pedestrians are warmer than other objects like trees, background, or road artifacts. In such cases the thermal shape can be often successfully exploited for human shape detection, since pedestrians are in fact brighter than the background. Unfortunately, other road participants or artifacts emit heat as well (cars, heated buildings...). Moreover, infrared images are blurred, have a poor resolution and the contrast is low compared with rich and colorful visible images and therefore, while being suitable for an initial areas of attention detection, may be too poor for a reliable pedestrian detection.

Conversely, visible images are more detailed than infrared ones. In such a case, areas of the images that contain pedestrians can be investigated to have a reliable detection. At the same time, a cluttered background can make an initial detection of areas of attention difficult. Moreover, daylight cameras can not be used in low illumination conditions.

For these reasons, in the proposed system, the simultaneous use of far infrared and daylight technologies are used in a synergical fashion to improve vision-based pedestrian detection.

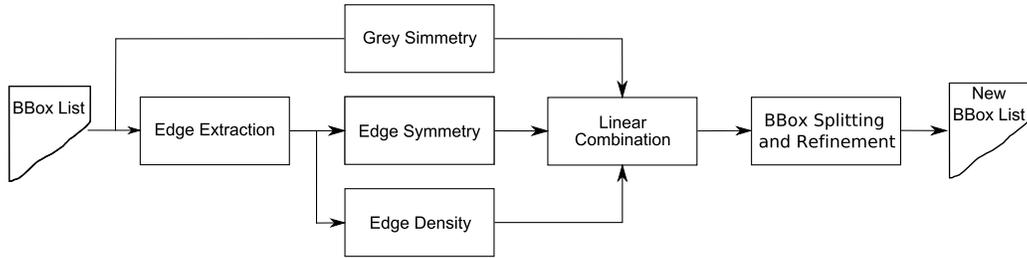


Fig. 3. Flow of symmetry processing.

Figure 2 depicts the overall algorithm flow for the complete pedestrian system. Different approaches have been developed for the initial detection in the two image domains: warm areas detection, vertical edges detection, and an approach based on the simultaneous computation of disparity space images in the two domains [5].

These first stages of detection output a list of areas of attention in which pedestrians can be potentially detected. Each area of attention is labelled using a bounding box.

The symmetry-based approach is used after this first stage. In fact, in order to build this list of areas of attention, few characteristics of the human shape are exploited; therefore, a number of bounding boxes includes objects other than pedestrians and, at the same time, a single bounding box can contain more than one pedestrian. Therefore, a refinement and initial filter step is mandatory as a preliminary stage for a pedestrian presence validation. The rough list is examined and the size of bounding boxes is refined. When a bounding box contains more pedestrians it is splitted in two or more bounding boxes. In addition, each resulting bounding box is labelled with a vote that is used by the final validation stage.

In [2], a symmetry-based approach was used to process monocular images in order to detect a list of regions of interest that potentially contain a pedestrian. In the system described in this paper, the list of areas of attention is produced thanks to FIR and daylight stereo process and the symmetry has proven to be more useful to refine or validate the areas of attention previously detected.

After this step, independent validators are used to evaluate the presence of human shapes inside the bounding boxes. These stages exploits specific pedestrian characteristics to discard false positives from the list of bounding boxes.

A final decision step is used to evaluate the presence of a human shape inside each bounding box balancing the votes of these validators and of the symmetry stage.

IV. SYMMETRY

As previously discussed, the detection phase generates a bounding boxes list in the infrared and daylight images.

Unfortunately, the same bounding box sometimes may enclose different pedestrians especially if they are close and at a similar distance from the cameras (figure 4.a).

In addition, both objects and pedestrians can be detected as well. Therefore a refinement and filtering process has

been developed to get one bounding box for each object, and to get rid of non-pedestrian obstacles. More specifically, a symmetry-based processing is used to split, refine and preliminarily validate each bounding box. In fact, pedestrians are generally symmetrical, therefore a symmetry computation can be used both as a validator and as refinement step.

Experimental results demonstrated that infrared images are more suitable than daylight images for a symmetry computation, since they contain a lower number of small details. Therefore, this phase is performed on the right infrared images only; the whole process is depicted in figure 3. This step produces an updated bounding boxes list for both infrared and visible domains.

Two different vertical symmetry measures are performed: one on the grey-level values and one on the vertical edge values. The density of edges is also considered to remove homogeneous edgeless areas.

For each bounding box all the possible positions of a symmetry axis are considered. For each axis, different symmetry window widths are evaluated considering the size of a human shape at the bounding box distance. Two symmetry maps are computed: one for grey levels (figure 4.b) and one for edges (figure 4.c). In these maps, the horizontal axis encodes the symmetry axis horizontal position within the bounding box; the vertical axis represents the symmetry window width. Figure 4.d shows the density map that is computed in a similar way to the one used for the symmetry map. The triangular shape of the maps is due to the limitation in scanning large windows for peripheral vertical axis.

Bright points encode the presence of a high symmetric content: the value of each symmetry map pixel depends on the number of symmetric points in the related symmetry window. In the grey level symmetry computation, two points are regarded as symmetric if they are equidistant from symmetry axis and have about the same grey value. Conversely, in the edge symmetry computation, they must also feature an opposite edge sign in order to match edges belonging to the same object.

Let $A = (x_a, y_a)$ and $B = (x_b, y_b)$ be two symmetric edge points with $x_a < x_b$ and let $\rho_a, \rho_b \in \pm 1$ be respectively their edge sign. A and B are regarded as warm symmetric edge points if $\rho_a = 1$ and $\rho_b = -1$, namely if A is a transition from dark to bright and B is a bright to dark ones. The edge symmetry computation mainly exploits the warm

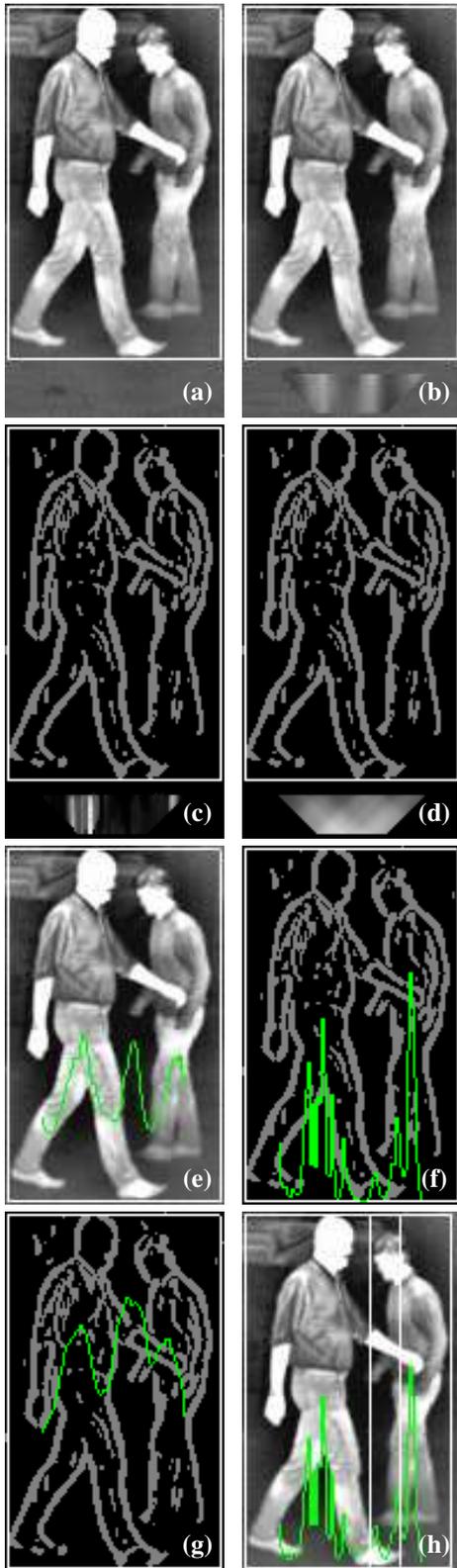


Fig. 4. Symmetry computation steps: (a) original bounding box, (b) grey and (c) edge symmetry maps, (d) edge density map, (e) grey and (f) edge symmetry histograms, (g) edge density histograms, (h) overall histogram and new splitted bounding boxes.

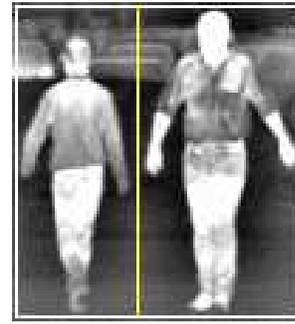


Fig. 5. A wrong axis is not detected thanks to the use of edge sign.

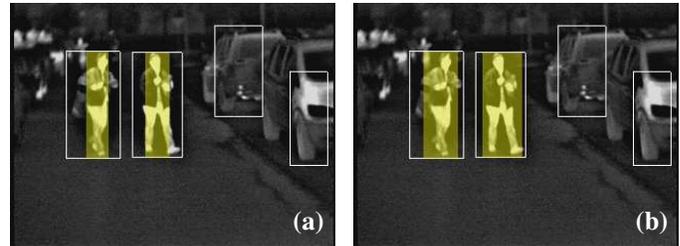


Fig. 6. The bounding boxes before and after the expansion.

symmetric points. This particular approach is aimed to regard as symmetric the objects that are brighter than background (figure 5). Also the match between bright to dark and dark to bright ones is considered for computing the symmetry maps; in fact, when, given an axis position and a symmetry window, the edge symmetry is higher than a given threshold, also this contribution is computed and added to the symmetry map. The underlying ratio of this approach is to further enhance the symmetry weight related to warm objects.

Grey level symmetry (figure 4.e), edge symmetry (figure 4.f) and edge density (figure 4.g) histograms are computed scanning the related maps and detecting the highest values for each column. Symmetry histograms and edge density are then combined to compute the overall histogram (figure 4.h), as shown in the following formula:

$$TotSym = (EdgeSym + k \times GreySym) \times EdgeDens$$

where k is a constant value used to weight grey level symmetry contribution.

A new bounding boxes list is then generated by thresholding the resulting histogram. Therefore, for each bounding box, the over-threshold peaks generate new bounding boxes contain the shape of a potential pedestrian. This step allows both to refine bounding boxes width and to split bounding boxes that contain more than one object.

The new bounding boxes width is, in fact, computed maximizing the edge density. Unfortunately, the so-computed bounding boxes tend to be narrower than the object they contain (figure 6.a). Therefore, the bounding boxes are furtherly expanded using a vertical histogram of edges, in order to better match the objects width (figure 6.b).

Also the symmetry content of each bounding box is added to the bounding box list. In fact, the list of refined and filtered bounding boxes is used by a following validator stages, and

by a final decision step; these stages are depicted in figure 2 and not detailed in this article [5].

V. DISCUSSION

The developed system has been tested in different situations using an experimental vehicle equipped with the tetra-vision system.

Figure 7 shows some results of the symmetry computation step. The validated pedestrians are shown using a superimposed yellow box. Conversely, the white rectangles represent the bounding boxes generated by previous steps. As examples, figures 7.e, 7.g, and 7.h show the bounding box width refinement; figures 7.b, 7.c, 7.d, 7.e, and 7.f depict the splitting of bounding box that contain more than one object and figure 7.e shows the bounding box validation.

As it can be noticed, some symmetrical objects, like cars are trees, are validated as well (figures 7.p and 7.q). This is also true for other vulnerable road users like cyclists and motorbikers. Moreover, some validation problems are encountered when the far infrared images are not optimal, like those acquired in summer under heavy direct sunlight; in these conditions, many object in the background become warm, and the assumption that a pedestrian features a higher temperature than the background is not satisfied. This causes some problems in the edge and symmetry computation, and therefore in the validation step (figure 7.o).

A quantitative analysis of results of the symmetry step has not yet been performed since the symmetry computation is an intermediate stage of the pedestrian detector system and its output is meant to be used by following steps.

Two different enhancements are currently under development: the use of horizontal edges in the symmetry computation and the bounding boxes height refinement. The horizontal edges are useful to the validation: in fact the bounding boxes, that contain more horizontal edges than vertical ones, are regarded as obstacles and not as pedestrian and therefore discarded. A single bounding box can contain some obstacle with different height. Therefore, the bounding boxes height refinement is needed in order to match the exact height of each obstacle.

ACKNOWLEDGMENT

This work has been supported by the European Research Office of the U. S. Army under contract number N62558-07-P-0029.

REFERENCES

- [1] C. Bauckhage, J. K. Tsotsos, and F. E. Bunn. Detecting Abnormal Gait. In *Procs. 2nd Conf. on Computer and Robot Vision*, pages 282–288, Toronto, Canada, May 2005.
- [2] M. Bertozzi, A. Broggi, R. Chapuis, F. Chausse, A. Fascioli, and A. Tibaldi. Shape-based pedestrian detection and localization. In *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems 2003*, pages 328–333, Shanghai, China, Oct. 2003. doi:10.1109/ITSC.2003.1251972.
- [3] M. Bertozzi, A. Broggi, M. Del Rose, and A. Lasagni. Infrared Stereo Vision-based Human Shape Detection. In *Procs. IEEE Intelligent Vehicles Symposium 2005*, pages 23–28, Las Vegas, USA, June 2005. doi:10.1109/IVS.2005.1505072.
- [4] M. Bertozzi, A. Broggi, M. Felisa, G. Vezzoni, and M. Del Rose. Low-level Pedestrian Detection by means of Visible and Far Infra-red Tetra-vision. In *Procs. IEEE Intelligent Vehicles Symposium 2006*, pages 231–236, Tokyo, Japan, June 2006. doi:10.1109/IVS.2006.1689633.
- [5] A. Broggi, M. Bertozzi, M. Felisa, P. Grisleri, S. Ghidoni, G. Vezzoni, C. Hilario Gómez, and M. Del Rose. Pedestrian Detection by means of Far-infrared Stereo Vision. *Computer Vision and Image Understanding*, 106(2):194–204, 2007, doi:10.1016/j.cviu.2006.07.016.
- [6] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas, and W. von Seelen. Walking Pedestrian Recognition. *IEEE Trans. on Intelligent Transportation Systems*, 1(3):155–163, Sept. 2000.
- [7] R. Cutler and L. S. Davis. Robust Real-time Periodic Motion Detection, Analysis and Applications. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(8):781–796, Aug. 2000.
- [8] J. W. Davis and V. Sharma. Robust Background-Subtraction for Person Detection in Thermal Imagery. In *Procs. Intl. IEEE Wks. on Object Tracking and Classification Beyond the Visible Spectrum*, Washington D. C., USA, 2004.
- [9] M. Del Rose and P. Frederick. Pedestrian Detection. In *Procs. Intelligent Vehicle Systems Symposium*, Traverse City, USA, 2005.
- [10] D. M. Gavrila. Pedestrian Detection from a Moving Vehicle. In *Procs. of European Conference on Computer Vision*, volume 2, pages 37–49, June–July 2000.
- [11] D. M. Gavrila and S. Munder. Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle. *Image and Vision Computing Journal*, 2007. In press.
- [12] L. R. Havasi, Z. Szlávik, and T. Szirányi. Pedestrian Detection using derived Third-Order Symmetry of Legs. In *Procs. Intl. Conf. on Computer Vision and Graphics*, Warsaw, Poland, Sept. 2004.
- [13] R. Kania, M. Del Rose, and P. Frederick. Autonomous Robotic Following Using Vision Based Techniques. In *Procs. Ground Vehicle Survivability Symposium*, Monterey, USA, 2005.
- [14] M. Mählich, M. Oberländer, O. Löhlein, D. Gavrila, and W. Ritter. A Multiple Detector Approach to Low-resolution FIR Pedestrian Recognition. In *Procs. IEEE Intelligent Vehicles Symposium 2005*, pages 23–28, Las Vegas, USA, June 2005.
- [15] H. Nanda and L. Davis. Probabilistic Template Based Pedestrian Detection in Infrared Videos. In *Procs. IEEE Intelligent Vehicles Symposium 2002*, Paris, France, June 2002.
- [16] S. Schauland and A. Kummert. Implementation and Optimization of Wavelet and Symmetry Features for Vision-based Pedestrian Detection. In *Procs. Intl. Conf. on Graphics and Visualization in Engineering*, pages 282–288, Clearwater, FL, USA, Jan. 2007.
- [17] S. Schauland, A. Kummert, S.-B. Park, U. Iurgel, and Y. Zhang. Vision-based pedestrian detection – improvement and verification of feature extraction methods and svm-based classification. In *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems 2006*, pages 97–102, Toronto, Sept. 2006.
- [18] A. Shashua, Y. Gdalyahu, and G. Hayun. Pedestrian Detection for Driving Assistance Systems: Single-frame Classification and System level Performance. In *Procs. IEEE Intelligent Vehicles Symposium 2004*, Parma, Italy, June 2004.
- [19] H. Shimizu and T. Poggio. Direction Estimation of Pedestrian from Multiple Still Images. In *Procs. IEEE Intelligent Vehicles Symposium 2004*, Parma, Italy, June 2004.
- [20] M. A. Sotelo, I. Parra, D. Fernandez, and E. Naranjo. Pedestrian Detection Using SVM and Multi-Feature Combination. In *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems 2006*, pages 103–108, Toronto, Sept. 2006.
- [21] C. Stauffer and W. E. L. Grimson. Similarity Templates for Detection and Recognition. In *Procs. IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 221–228, 2001.
- [22] F. Suard, V. Guigue, A. Rakotomamonjy, and A. Bensrhair. Pedestrian detection using stereo-vision and graph kernels. In *Intelligent Vehicles Symposium, Las Vegas, Nevada*, pages 267–272, June 2005.
- [23] P. Viola, M. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. In *IEEE International Conference on Computer Vision*, pages 734–741, Oct. 2003.

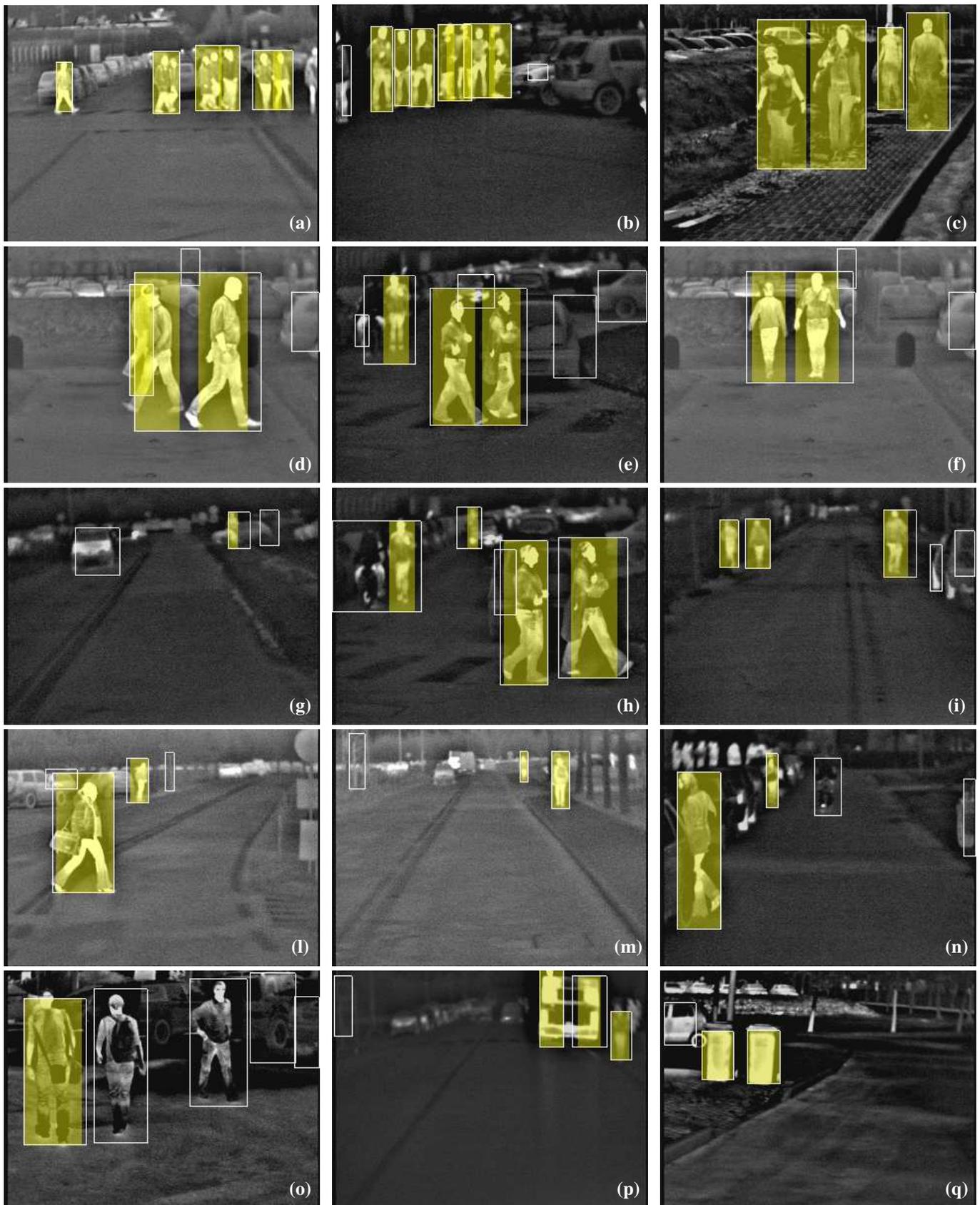


Fig. 7. Symmetry analysis results: validated pedestrians are shown using a superimposed yellow box; the white rectangles represent the bounding boxes generated by previous steps and fed to the symmetry based validation and refinement step.