Precise 3D Measurements for Tracked Objects from Synchronized Stereo-Video Sequences

Panagiotis Agrafiotis, Andreas Georgopoulos, Anastasios D. Doulamis, and Nikolaos D. Doulamis

National Technical University of Athens, School of Rural and Surveying Engineering
Lab. of Photogrammetry, Zografou Campus, Heroon Polytechniou 9, 15780 Zografou, Greece
{pagraf,drag}@central.ntua.gr, {adoulam,ndoulam}@cs.ntua.gr

Abstract. This paper presents a system suitable to perform precise and fast 3D measurements from synchronized stereo-video sequences and provide target's georeference in a known reference system. To this direction we combine a robust tracker with photogrammetric techniques into a fast and reliable system. For tracking objects and people, we adopt and modify a stable human tracker able to cope efficiently with the trade-off between model stability and adaptability. For achieving accurate and precise 3D measurements, camera calibration was implemented in order to recover the intrinsic parameters of the cameras of the configuration. Finally, for precise and reliable calculation of the 3D trajectory of the moving person, we apply bundle adjustment for all frames. Bundle adjustment is a very accurate algorithm and has the advantages of being tolerant of missing data while providing a true Maximum Likelihood estimate. The results have been tested and evaluated in real life conditions for proving the robustness and the accuracy of the system.

1 Introduction

Object detection and tracking is one of the most critical research problems within the computer vision community and thus many related algorithms and tools have been proposed in the literature. The main problem of these tools is the fact that they process and analyze single optical sources without providing 3D information. However, outdoor surveillance and security applications present new research challenges as far as vision technologies and analysis tools are concerned. Firstly, the data are being captured from such distances that the imagery resolution sometimes becomes far lower than the appropriate definition range. Secondly, capturing is performed in outdoor conditions and thus the analysis is strongly affected by weather variations (rain, night, dawn-light, reflections due to clouds, fog, mist, wind and temperature conditions, etc.). Thirdly, the cameras used to monitor the outdoor areas are continuously displaced for obvious reasons, a fact which poses new challenges to the detection and identification process. Fourthly, it is possible that a target be split into several pieces (e.g. people falling into the sea and swimming towards different directions) thus making tracking from one data source an inefficient process. All the

G. Bebis et al. (Eds.): ISVC 2014, Part II, LNCS 8888, pp. 757–769, 2014.

above verify the fact that the existing research tools are not adequate for complete outdoor wide areas surveillance and new algorithms should be developed.

For this reason, in this paper we propose a research tool that recovers the position of a tracked object in space while providing at the same time high precision results. It is worth mentioning that precise estimation of the position in space of detected targets from an optical system is a very important attribute for surveillance systems. Even more important is to maintain the precision of 3D positioning as the camera-to-object distance is dramatically increasing. Thus, the need is emphasized for precise and reliable recovery of the 3D coordinates of a detected target similar to that required in many terrestrial photogrammetric applications. This provides the possibility to project the coordinates of the target to a reference system, and make the identification procedure by the relevant services or by aerial surveillance methods (e.g. UAV) easier and faster while enabling them to record and study the routes of moving people and generally monitor human activities.

All necessary tools for the implementation of the above are provided by computer systems and computer vision. More specifically, computer systems, implement the above by finding correspondences between points, visible and imaged by two or more cameras. It is possible to calculate the three-dimensional (3D) position of these points using these correspondences and knowing the distance between the two cameras. The detection of points – targets in every image is demanding. Therefore the detection area in each picture is reduced exploiting the geometry of the stereopair in order to reduce computational cost.

This paper proposes a system for calculating three-dimensional coordinates of detected targets which is running concurrently with the tracking module and is designed to operate in real-life applications, achieving high positioning accuracy and minimum computational cost. This level of precision is necessary in long distance tracking applications since the increasing object's distance from the camera results in consecutive loss in precision.

2 Related Work

In the literature, there are not many reports on the issue of precise 3D measurements for tracked objects from synchronized video stereo-pairs. Key element for accurate and reliable results is the development of a robust, fast and accurate tracking system.

2.1 Tracking Objects

Recently a great effort has been dedicated to handle object tracking as a classification problem [1], [2]. However, these approaches exploit no adaptable mechanisms to update the performance of the classifier and thus the structure of the model remains fixed. One of the first approaches towards an adaptable classification for object tracking has been presented in [3]. These works, however, do not face efficiently the general trade-off between model stability and adaptability. A highly specific model significantly increases the reliability of the tracker to capture the target but loses adaptability. On the other hand, a highly general model copes with the legitimate changes in appearance but with a cost in reliability.

Other methods exploit the semi-supervised paradigm [4], [5], a co-training strategy [6] a combination of generative and discriminative trackers [7] or finally coupled layered visual models [8].The realization that local information sometimes is not enough to make a correct decision led to the development of global optimization trackers. These trackers operate on larger time scales and make use of high-level reasoning to re-solve ambiguities [9], [10]. Other works handle the problem as a minimum flow cost problem [11], [12].

2.2 3D Position Estimation

Recently, a novel method for dealing with this problem was presented from [13]. This method presented a new stereo rectification method for dual-PTZ camera system in order to increase the efficiency of stereo matching. In addition, an interesting approach to solve stereo vision problems by means of rotating cameras has been proposed recently with its analytic formulation [14]. There, an off-line initialization process is performed in order to initialize essential matrix using calibration parameters. During on-line operations the rotation angles of the cameras are retrieved and exploited to compute the essential matrix. When the zoom is considered, it would require the calibration for any zoom level of both cameras. In [15], the authors propose a stereo system based on two PTZ cameras from a network of cooperative sensors. The proposed solution is able to accurately localize a moving object in outdoor areas. Finally, and most commonly the problem is addressed by a triangulation procedure, where already knowing the camera positions and after a stereo-calibration process, we can compute the 3D position of a point on a stereo-pair.

2.3 Contribution

Main contribution of the work presented in this paper is that through the use of bundle adjustment for frames orientationthe georeference of the detections in a real reference system such as the World Geodetic System (WGS) 84is obtained . The reason that led us to use bundle adjustment is that most of the tracking cameras are used in outdoor environments and thus noise is inevitably introduced into the images. This leads to noisy image measurements and the equations of projective transformation are not exactly satisfied. Bundle adjustment, which is used in our case for the orientation of the resulting frames, is the optimal procedure in order to overcome the aforementioned problems, because it appears to be tolerant of missing data while providing the true Maximum Likelihood estimation. Through this, it is now possible to calculate with great accuracy the position of a target necessary for detection and tracking at long distances as this system remains robust and accurate. Moreover bundle adjustment is a well-tested algorithm not only for the case of aerial and satellite imagery, but also for large terrestrial image sequences. Today algorithms have been developed that allow bundle adjustments to run fast and with reliability and at almost real time, providing robust and reliable results with the best possible accuracy that may be extracted from the data available.

3 System's Overview

For achieving fast and precise 3D coordinate estimation of the target, the ground tracking system is modified with additional photogrammetric routines (Figure 1).Initially a camera calibration procedure was applied in order to compensate for all distortions caused by optical anomalies, like housing of the cameras, radial distortion, etc. Camera calibration is performed using an easy to use 2D calibration pattern in order to automate the procedure and achieve very accurate results. Next, a tracker is applied in order to fast and efficiently extract the moving object from the background. To do this, we introduce a stable human tracker able to efficiently cope with the trade-off between model stability and adaptability. More specifically, we adopt probabilistic mixture models like the Gaussian Mixture Models (GMMs) which exploit geometric properties for background modelling. Then, we integrate iterative motion information methods, concerned by shape and time properties, to estimate image regions of high confidence for updating the background model. Finally and after the final stage of the tracking scheme which is the extraction of target's image coordinates, the system performs bundle adjustment in order to estimate the 3D coordinates of the aforementioned target. In contrast with most of relevant works, the proposed system is developed aiming on the achievement of measurements of high accuracy and precision at very long distances, up to 100 meters.

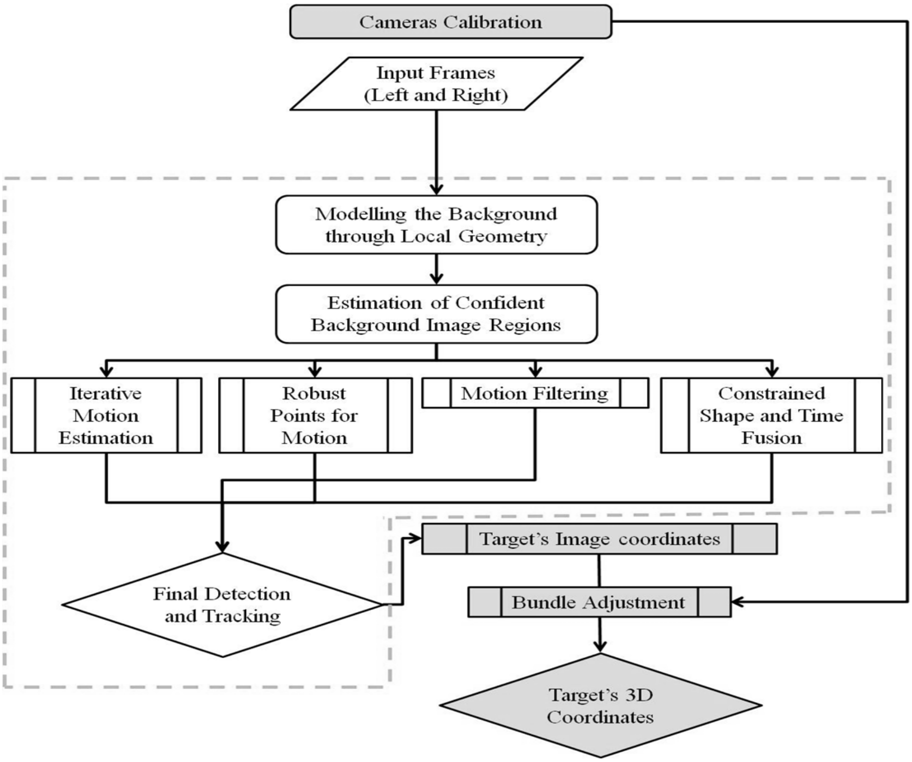

Fig. 1. Overview of the System's Architecture

3.1 Camera Calibration

Accurate camera calibration and orientation procedures are necessary prerequisites for the extraction of precise and reliable 3D metric information from images and videos. Through this procedure, the intrinsic parameters of the camera are calculated consisting of (a) the principal distance, (b) the principal point offset and (c) the lens distortion values. However, in many applications, especially in computer vision, only the focal length is recovered while for precise photogrammetric measurements all the calibration parameters are generally employed. In the system described in this paper, high accuracy and precision are required in order to extract the 3D position of the target in very long distances and for this reason it is necessary to recover all the calibration parameters described above. Various algorithms for camera calibration have been reported over the years in the photogrammetric and computer vision literature. The algorithms are generally based on perspective or projective camera models, with the most popular approach being the well-known self-calibrating bundle adjustment, which was first introduced to close-range photogrammetry in the early 1970s.

For the system described in this paper, camera calibration is performed offline and by using a two-dimensional (2D) control field, having dimensions of 90 cm x 90 cm and containing exactly 100 points, including four coded control points. The high-contrast targets (dots) allow for their automatic marking and having the knowledge of their relative position and the distances between them and the 4 control points, the calibration can be performed automatically by using template matching techniques. It is worth mentioning that the calibration models used in computer vision have traditionally employed reference grids. The intrinsic parameters being determined using images of a known object point array (e.g. a checkerboard pattern). Commonly adopted methods are those described in [16], 17] and [18]. These are all based on the pinhole camera model and include terms for modelling radial distortion.

For these methods and in order to achieve better and more stable results, it is necessary to have a minimum number of six images. Moreover, the two-dimensional (2D) calibration field should cover a significant percentage of the area of the image and the images must be taken from different angles, different distances and from the four sides of the field (90° rotation), both horizontal and vertical layout (Figure 2). Finally, in order to yield the correct scale to the object, there should be defined some constraints like the distances between the four coded targets.

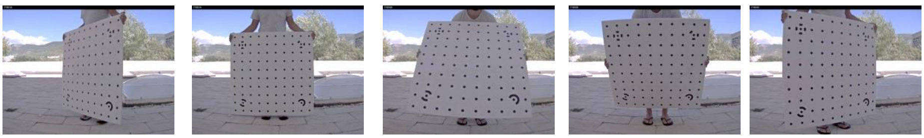

Fig. 2. The two-dimensional (2D) calibration field of Photomodeler Scanner. The images are taken from different angles, different distances and from the four sides of the field (90° rotation).

3.2 The Tracking Scheme

For performing precise and accurate 3D measurements on tracked objects, it is important to have a robust, fast and efficient Tracking Scheme. Therefore, in order to derive efficient human detection in outdoor environments, we appropriately modify the system described in [19].

Overview of the Tracking Scheme

For efficient human detection in uncontrolled outdoor environments, the system combines on the one hand adaptive background models able to capture slight modifications of the background patterns with, and on the other hand, motion-based algorithms that define with high accuracy which parts of an image should be considered as foreground/background when abrupt, non-period and sudden changes take place.

In this system, an accurate but also quick background subtraction methodology is adopted to deal with the problem of tracking over very long periods. In contrast to previous, conventional works, probabilistic mixture models like the Gaussian Mixture Models (GMMs) are adopted, exploiting geometric properties of locally connected regions. The approach described in [19] adopts a graph-based saliency map methodology so that the most important pixels are selected and these to be fed as GMM inputs.

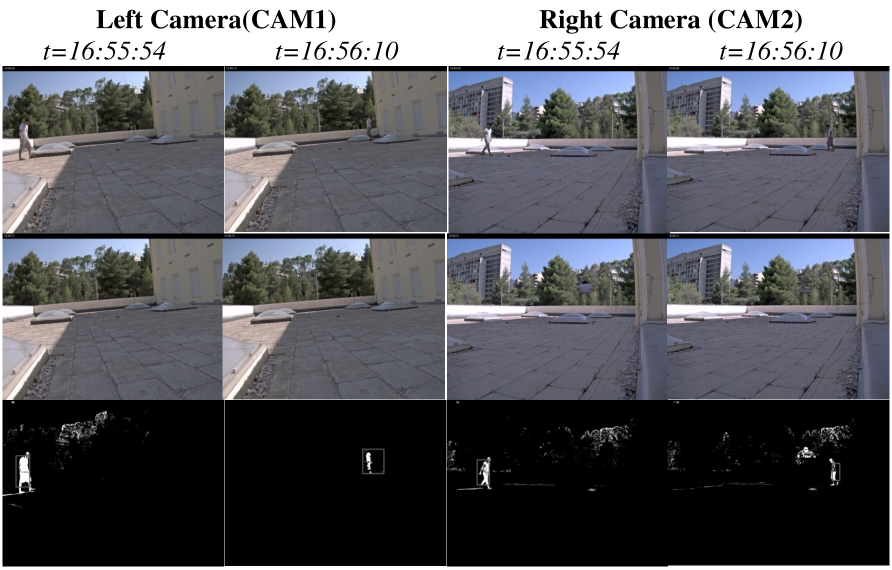

Fig. 3. Tracking results from Left and Right camera at different time. First row shows the original frame. The next two rows on the right show successful examples both for background modelling/updating and for human detection.

Having detected the moving object, it is then modelled in the foreground as a moving object. To this end, we combine optical flows techniques with the best features to track methodology. Then we define the time instances required for updating the background as well as the frames regions in which such updating is required. Shape and motion constraints are imposed to improve the analysis.

Parts of the background which are moving are also considered as foreground. First, the motion in the background is stabilized and the objects cease to move. Then, these objects automatically assigned to the background and correct background subtraction take place. The same is valid if the foreground stops moving. It will be instantaneously considered as background. However, once the foreground object starts moving again, it will be excluded from the background and correct localization takes place.

This system has been developed in C++ platform and by exploiting the Integrated Performance Primitives (IPP) of Intel, in order to support real-time execution of the aforementioned algorithms. The system has been evaluated in the laboratory for the accuracy of foreground detection and tracking.

3.3 Precise 3D Coordinates Estimation

As already mentioned, the capacity of estimating the 3D position of a tracked object is very important and in some cases a necessary feature for surveillance systems. Even more important is for such systems to remain precise and reliable for calculating the 3D position while the camera-to-object distance is dramatically differentiating.

The proposed system is able to compute the 3D position of a tracked human in every stereo-pair consisting of the left and the right frame. All these frames are oriented in 3D space through a bundle adjustment procedure. This process is applied in the original frames in order to find fast and reliable corresponding points between the frames of the stereo-pairs and between the whole sequences of the stereo-pairs. During this it is important to mark on the first few frames at least three points of known coordinates, contained in the overlapping area of the images. Through this georeference procedure, the computed 3D coordinates are in a real reference system, for example the World Geodetic System (WGS) 84 and could be projected onto a map. In addition, WGS 84 is the reference coordinate system used by the Global Positioning System (GPS) and that facilitates the comparison of the results.

Bundle Adjustment

Tracking cameras in outdoor environments introduce noise in the images due to luminosity conditions and abrupt weather changes. Due to this fact, image measurements are noisy and the equations of projective transformation will not be satisfied exactly.

In such cases the process seeks the Maximum Likelihood (ML) solution considering that the measurement noise is Gaussian: the main aim is to estimate projection matrices and 3D points which project to specific image position according to the equations of projective transformation for every view in which the 3D point appears. This estimation involving minimizing the reprojection error is known as bundle adjustment – it involves adjusting Camera centers, which are the perspective centers of the bundles, and 3D points through which the rays of the bundles must pass (by definition) [20]. In the above procedure, the determined parameters are the combined 3D feature coordinates, camera poses and calibration parameters.

Bundle adjustment is generally used as the final step of any feature-based reconstruction algorithm which requires high accuracy in the determination of 3D points.

In the system presented in this paper, we use this process because of the advantages over other methods. Bundle adjustment appears to be tolerant of missing data due to partial occlusions and luminosity conditions while providing the true Maximum Likelihood estimation. It allows assignment of individual covariance values to each measurement (e.g. if cameras are situated in different distances from the target) making it extremely flexible for tracking applications. It is able to gracefully handle a very wide variety of different 3D feature and camera types (points, lines, curves, surfaces and cameras), scene types (including dynamic and articulated models, scene constraints), information sources (2D features, intensities, 3D information, priors) and error models (including robust ones) [21].

Moreover, it gives precise and easily interpreted results because it uses accurate statistical error models and supports a sound, well-developed quality control methodology. Because of these, it appears as an ideal algorithm for precise computing of 3D coordinates of detected objects in uncontrolled environments. However, it requires a good initialization to be provided, and it can become an extremely large minimization problem because of the number of parameters involved. Also, as already stated, in order to avoid a self-calibration procedure which would require extra points of known coordinates, an off-line full camera calibration is necessary for initialization of intrinsic parameters of the cameras.

Finally, bundle adjustments algorithms are very efficient and appropriate for real-time applications such as human tracking because they use economical and rapidly convergent numerical methods and make near-optimal use of problem sparseness.

$$\frac{x}{-c} = \frac{r_{11}(X - X_0) + r_{12}(Y - Y_0) + r_{13}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}. \tag{1}$$

$$\frac{y}{-c} = \frac{r_{21}(X - X_0) + r_{22}(Y - Y_0) + r_{23}(Z - Z_0)}{r_{31}(X - X_0) + r_{32}(Y - Y_0) + r_{33}(Z - Z_0)}. \tag{2}$$

Equations (1) and (2) are the collinearity equations which are used as observation equations for this method and image coordinates are used as observations. The parameters (X_0, Y_0, Z_0) are the camera projection center coordinates, (X, Y, Z) the object's real coordinates, (x, y) the image coordinates of the projected point and c the focal length of the camera.

4 Experimental Results

In order to evaluate the performance of the system described above, data were collected taking into account the requirements and real-life scenarios. More specifically, videos were recorded using the stereo-configuration consisting of two synchronized AXIS P1357 Network Cameras. For these sequences, the cameras were placed on the roof of a building at National Technical University of Athens and videos were recorded. The videos were recorded with a rate of 10 fps, 1024x768 resolution, no compression and a GSD of 10mm. For this experiment, the base was chosen to be 2 meters in a direction parallel with the motion direction and cameras had a convergence angle of 20 degrees. In these sequences, a human is walking through the overlapping area of the frames of the left and right camera. Important is that the person walks at different distances from the camera in order to assess the 3D coordinates estimation performance of the applied algorithms.

4.1 Camera Calibration

A camera calibration procedure was separately performed in order to recover the intrinsic parameters of the cameras of the stereo system and compensate all the distortions introduced by the lenses and the special housing for outdoor environments. For minimizing computational complexity of the system described above, calibration was performed using photogrammetric software (Photomodeler Scanner by EOS) and a special designed 2D test field made from plexiglass. It is important to note that the developed system supports a calibration routine in C++ and thus, the aforementioned test field has on the one side a chessboard and on the other the pattern from Figure 2. In Table 1 are presented the calibration results from the left and the right camera of the stereo-configuration.

Table 1. Camera Calibration Results

Camera	c (mm)	x0 (mm)	y0 (mm)	K1	K2
Right	5.73	3.06	2.22	8.988e-003	4.328e-004
Left	5.79	3.03	2.09	7.282e-003	2.444e-004

It is important to note that the differences between the recovered parameters of the intrinsic orientation of the left and the right camera. It is apparent that the corresponding parameters between the two cameras differ by a size capable of creating problems in the accurate calculation of the 3D object's position in long distances. Therefore, based on these results the need for an accurate and full camera calibration procedure, even in tracking systems is confirmed.

4.2 3D Coordinates Estimation

For the evaluation of the performance of the 3D coordinates estimator, data were collected using the configuration described below.

Figure 4 presents the stereoscopic configuration developed for assessing the results and addressed by this paper. The two (or more) synchronized cameras are able to estimate with high accuracy the position in space of a tracked object when it is imaged in the overlapping area of the images. In more detail, this system consists of two AXIS P1357 Network Cameras which are suitable for use even in extreme outdoor conditions (rain, snow, fog etc.) and a central processor which collects and processes the frames from each camera. Frames are obtained with a resolution of 1024x768, 5.7mm camera constant and a constant horizontal overlap of 80%.

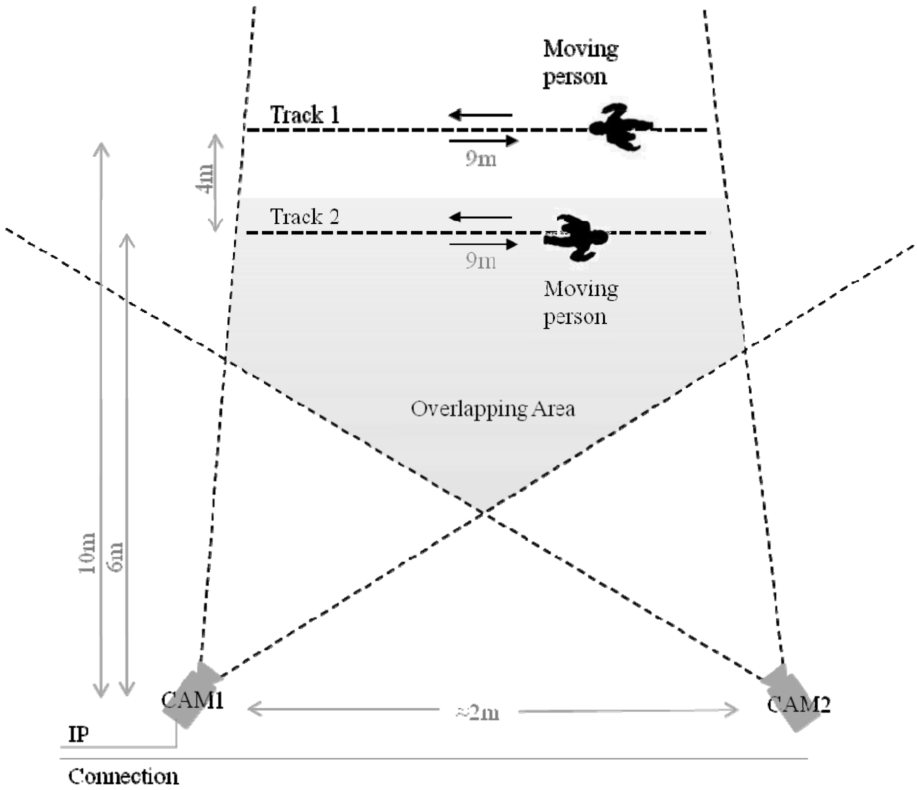

Fig. 4. System's setup for applying stereoscopic measurements and evaluate the results

As already stated, a person was walking in different directions in two different tracks. Track 1 was at a distance of 10m from the two cameras and had a length of 9m. Track 2, was at a distance of 6m from the stereo-configuration and had also a length of 9m. This set up was adopted in order to assess the performance and the accuracy of the system at different distances between the detected target and the cameras. In addition, the human was walking till the edge of the overlapping area in order to evaluate the calibration results (much larger distortions) and the accuracy of the adjustment procedure.

It is important to note that Track 1 and Track 2 were chosen to be a straight line in order to test how the system's results are fitting on these lines. It is clear that in these configurations the main issue is the depth determination which in our case equals with the camera-target distance.

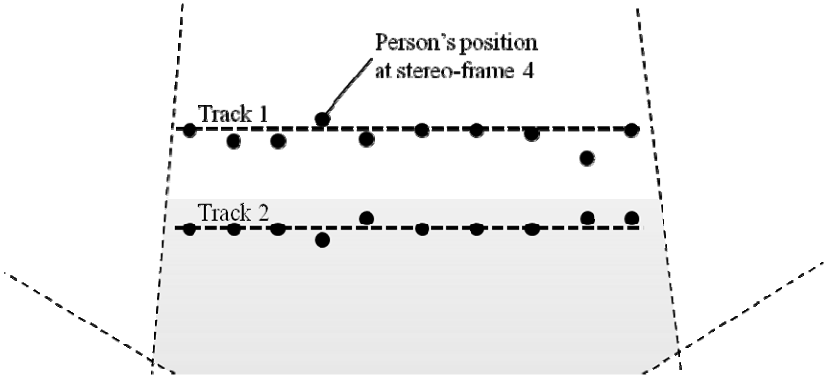

Fig. 5. Person's positions at every stereo frame pairs for Track 1 and Track 2

Table 2. Person's position at every stereo-frame in the form of world coordinates

Track	Frame	X	Y
1	1	11.30	3.06
	2	12.06	3.57
	3	12.88	3.90
	4	13.66	4.58
	5	14.98	5.41
	6	15.43	5.65
	7	16.57	6.13
	8	17.10	6.67
	9	18.62	7.01
	10	19.03	7.66
2	1	13.54	-0.21
	2	14.85	0.59
	3	14.93	0.75
	4	15.95	1.15
	5	16.64	1.67
	6	17.10	2.27
	7	17.99	2.48
	8	18.71	2.74
	9	19.53	3.33
	10	19.93	3.75

In Figure 5 are graphically presented the resulting positions at every stereo-frame. In Table 2 the person's positions are also presented at every stereo-frame in the form of world coordinates. As it is observed, the system described in this paper works properly and accurate for different moving directions and different cameras-to-object distances. In more detail, our system achieves absolute accuracy of less than 10cm in both control tracks. Important is that even these inaccuracies, are assigned mainly to the movement of the body of the walking person and the low resolution image. Moreover, the largest errors are observed on the edge of the overlapping area of the frames which is quite normal.

Future Work

Our future work includes the testing of the described system over different luminosity conditions and weather changes. In parallel with this, important is to test and evaluate the results absolutely by comparing the resulting trajectory from our system with accurate GPS RTK (real time kinematic) measurements having no less than 1cm accuracy.

Acknowledgements. The research leading to these results has been supported by European Union funds and National funds (GSRT) from Greece and EU under the project JASON: Joint synergistic and integrated use of eArth obServation, navigatiOn and commuNication technologies for enhanced border security funded under the cooperation framework.

References

1. Avidan, S.: Support Vector Tracking. IEEE Transaction on Pattern Analysis and Machine Intelligence 26(8), 1064–1072 (2004)
2. Lepetit, V., Lagger, P., Fua, P.: Randomized trees for real-time keypoint recognition. In: Proc. IEEE CVPR, vol. 2, pp. 775–781 (2005)
3. Doulamis, A., Ntalianis, K., Doulamis, N., Kollias, S.: An Efficient Fully-Unsupervised Video Object Segmentation Scheme Using an Adaptive Neural Network Classifier Architecture. IEEE Trans. on NNs 14(3), 616–630 (2003)
4. Stalder, S., Grabner, H., Van Gool, L.: Beyond semi-supervised tracking: Tracking should be as simple as detection, but not simpler than recognition. In: Proc. of IEEE ICCV, pp. 1409–1416 (2009)
5. Grabner, H., Leistner, C., Bischof, H.: Semi-supervised on-line boosting for robust tracking. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part I. LNCS, vol. 5302, pp. 234–247. Springer, Heidelberg (2008)
6. Tang, F., Brennan, S., Zhao, Q., Tao, H.: Co-tracking using semi-supervised support vector machines. In: Proc. of IEEE ICCV, pp. 1–8 (2007)
7. Yu, Q., Dinh, T.B., Medioni, G.G.: Online tracking and reacquisition using co-trained generative and discriminative trackers. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part II. LNCS, vol. 5303, pp. 678–691. Springer, Heidelberg (2008)
8. Cehovin, L., Kristan, M., Leonardis, A.: An adaptive coupled-layer visual model for robust visual tracking. In: Proc of IEEE ICCV, pp. 1363–1370 (2011)

9. Xing, J., Ai, H., Lao, S.: Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1200–1207 (2009)
10. Huang, C., Wu, B., Nevatia, R.: Robust Object Tracking by Hierarchical Association of Detection Responses. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part II. LNCS, vol. 5303, pp. 788–801. Springer, Heidelberg (2008)
11. Zhang, L., Yuan, L., Nevatia, R.: Global data association for multi-object tracking using network flows. In: Proc. of IEEE CVPR, pp. 1–8 (2008)
12. Henriques, J.F., Caseiro, R., Batista, J.: Globally optimal solution to multi-object tracking with merged measurements. In: Proc. of IEEE ICCV, pp. 2470–2477 (2011)
13. Wan, D., Zhaou, J.: Stereo Vision Using Two PTZ Cameras. Computer Vision and Image Understanding 112(2), 184–194 (2008)
14. Hart, J., Scassellati, B., Zucker, S.W.: Epipolar Geometry for Humanoid Robotic Heads. In: Proc. of 4th International Cognitive Vision Workshop, pp. 24–36 (2008)
15. Kumar, S., Micheloni, C., Piciarelli, C.: Stereo localization using dual PTZ cameras. In: Jiang, X., Petkov, N. (eds.) CAIP 2009. LNCS, vol. 5702, pp. 1061–1069. Springer, Heidelberg (2009)
16. Tsai, R.Y.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE Int. Journal Robotics and Automation 3(4), 323–344 (1987)
17. Heikkilä, J., Silven, O.: A four-step camera calibration procedure with implicit image correction. In: CVP 1997 (1997)
18. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. on PAMI 22(11), 1330–1334 (2000)
19. Kokkinos, M., Doulamis, N.D., Doulamis, A.D.: Local Geometrically Enriched Mixtures for Stable and Robust Human Tracking in Detecting Falls. Int. J. Adv. Robot Syst. 10, 72 (2013), doi:10.5772/54049
20. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press
21. Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzgibbon, A.W.: Bundle adjustment – A modern synthesis. In: Triggs, B., Zisserman, A., Szeliski, R. (eds.) ICCV-WS 1999. LNCS, vol. 1883, pp. 298–372. Springer, Heidelberg (2000)