The Concept of Video Surveillance System Based on the Principles of Stereo Vision

Dmitry Stepanov, Igor Tishchenko
Program System Institute of Russian Academy of Sciences
Pereslavl-Zalessky, Russian Federation
mitek1989@mail.ru, igor.p.tishchenko@gmail.com

***Abstract*—The novel concept of video surveillance systems based on principles of stereovision is described. Examples of existing CCTV systems, which use stereovision tools and techniques are presented. The proposed concept allows to estimate the 3D-coordinates of the objects of interest, concept based on the positioning of the fixed and pan-tilt-zoom cameras included in the surveillance system. Positioning is performed using fixed reference points with known 3D-coordinates. Reference points are distributed over the observed area. Also, the technique of calculating the 3D-coordinates of reference points in a unified conventional coordinate system using cameras within the surveillance system.**

I. INTRODUCTION

Modern video surveillance systems use different algorithms of automatic detection, tracking and recognition of objects of interest on the images (people, vehicles, abandoned objects). In the past few years, the video surveillance systems, based on the joint use of many cameras, as well as the methods and algorithms of stereovision are developed. If the object of interest is observed in the images from two or more calibrated cameras, and relative orientation and relative position of cameras are known, then we can estimate three-dimensional position of objects of interest in some conventional coordinate system using photogrammetry equations. Also, we can estimate true object’s size and speed, perform object tracking and predict their movement. Multi-camera surveillance system increases the probability of a successful image segmentation and object tracking of an object in the field of view of any one camera due to a larger number of angles of observation.

Here are some examples of video surveillance systems based on stereo vision and using of multiple cameras, and description of some articles in this science’s field.

In 2012 KDC corporation introduced a 3D CCTV stereo camera for video surveillance systems (presented in Fig. 1). One the features of the software developed for the camera is identifying whether the customers are male or female using analyzing the average shape of thousands of male and female customers from a database [1].

In the Smart Vision Sensor system [2], a stereo camera mounted on the ceiling (Fig. 2), it has intelligent video data processing technology which imitates the human eye's visual structure, as well as spatial measurement and alarm sensor functions.

Fig. 1. 3D CCTV stereo camera

A team from the University of the West of England’s (UWE) Centre for Machine Vision in Bristol, U.K. aims to extract high-resolution 3D video information even at a distance of several hundred feet [3]. They have built a demonstrator system able to operate both at close range and at long distance, and will test it in realistic outdoor environments during night and daytime. The system uses the photometric stereo [4] – technique for estimating the depth and surface details of an object by examining the same view from different directions.

TwinEye intelligent video surveillance system [5] equipped with stereo rig consisting of two cameras, system solves the problem of detecting and tracking objects of interest, estimate their 3D-coordinates, as well as the calculation of 3D-distances and 3D-areas with images. Graphical interface of TwinEye software is shown on Fig. 3.

Smart CCTV Ltd Company uses stereo vision in its surveillance system for people counting [6].

Fig. 2. Store owner checking the real-time customer inflow data on a mobile device through ‘Smart Vision Sensor’

C:\Users\HOME\Desktop\Map\연구소 내부.png

Fig. 3. Graphical interface of TwinEye software

The article [7] describes a vision based railway platform safeguard, it uses stereo cameras and thermal cameras. System detects possible accidents on the platform and informs operators about the accidents with video information and alarms. Fig. 4 shows the installation of cameras in the platform of some metro station.

Authors of article [8] have proposed a novel approach to evaluating how effectively a closed circuit television (CCTV) system can monitor a targeted area. With 3D models of the target area and the camera parameters of the CCTV system, the approach produces surveillance coverage index, which is defined as a quantitative measure for surveillance performance. The surveillance system consists of several fixed monocular cameras. Arrangement of CCTV cameras in an underground parking lot is shown on Fig. 5.

Authors of article [9] investigate the effect of weather conditions on the comparison of images from the video surveillance system with one stereo camera.

Fig. 4. Example of the installation of stereo cameras in the platform of some metro station

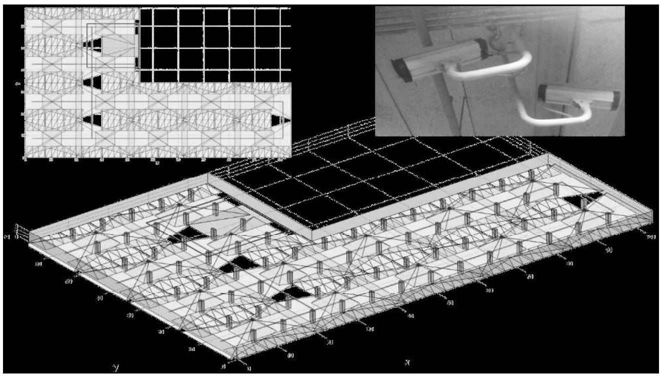

Fig. 5. Arrangement of CCTV cameras in an underground parking lot

In article [10] a first-person viewing and navigation interface for integrated surveillance monitoring in a virtual environment is proposed. It is designed for egocentric tasks, such a tracking persons or vehicles along several cameras. For these tasks, it aims to minimize the operator's 3D navigation effort while maximizing coherence between video streams and spatial context. The user can easily navigate between adjacent camera views and is guided along 3D guidance paths (Fig. 6).

In dissertation [11] visual localization algorithm in a 2D-coordinate with multiple cameras is proposed. The distance between an object and a camera is provided by a reference point. The reference point is initially a rough estimate, and author is motivated to obtain a more accurate reference point and uses an iterative process which substitutes a previously localized position with a new reference point close to a real object location. In addition, the proposed localization method has an advantage of using a zooming factor without concerning about a focal length. The effectiveness of the proposed algorithm in object position localization as well as tracking is illustrated. The proposed algorithm can be effectively applied in many tracking applications where visual imaging devices are used. But algorithm does not perform 3D-reconstruction of objects' coordinates.

In paper [12] 3D surveillance system using multiple cameras surrounding the scene is also proposed. Application is concerned with identifying humans in the scene and then identifying their postures. The cameras are fully calibrated. Object detection and interpretation are performed completely in 3D space. Using depth information, persons can easily be separated from the background and their posture identified by matching with 3D model templates. The cameras are arranged in two pairs, each of which is placed at almost equal distances surrounding the scene (Fig. 7). Each pair of cameras is used as a short-base stereo pair, which gives good depth information. The wide-line arrangement between the pairs of cameras helps resolve occlusions between objects in the scene, assuming the occlusion is absent in at least one of the short-base stereo views. The normalized sum of squared difference (NSSD) criterion is used for reconstruction the scene using stereo views obtained from two different angles. Epipolar geometry is used to speed-up the finding correspondence between features in two views. Notice that the cameras assumed to remain fixed in their positions.

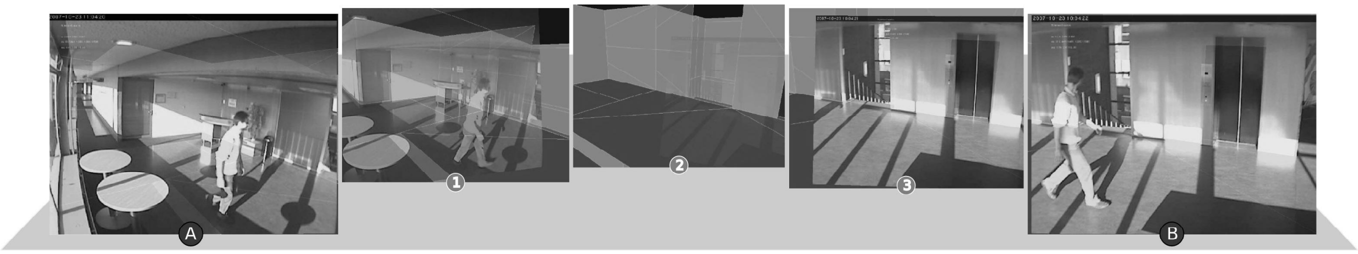

Fig. 6: View transition sequence of tracking a person between camera A and adjacent camera B in an office hallway surveillance scenario

Authors of multi-camera surveillance system, proposed in [13], introduce a cross-camera calibration approach called map-view mapping, which maps each ground point in the image ("view") to its corresponding point on a global site map ("map") - fine resolution satellite image for an outdoor application or a blueprint drawing for an indoor application. Authors assume that in each view the targets are on the same plane, called image ground plane; and the global map also represents a single plane in the world, called map ground plane. Thus for each camera, the mapping between point x in the view and the corresponding point X in the map is fully defined by a homography matrix. Also the matching line features in conjunction with the matching points was introduced. Line features can be directly specified by the user (manually), or computed from pairs of user defined calibration control points. Since all of the camera views are calibrated to the same map, the corresponding targets from multiple cameras can be naturally associated based on their map locations. As well as in previous work, cameras are fixed and static. Also, proposed method will not work of some reference points and lines don't lie on the same plane.

We see that all of the proposed solutions for 3D-reconstruction of objects' coordinates are based on two approaches: the use small amount of stereo cameras or the use of multiple monocular or stereo cameras, which are fixed. With a small number of cameras we can't solve the problem of three-dimensional positioning of objects in large rooms or rooms with complex configuration. In the case of fixed cameras we can perform pre stereo calibration of stereo rigs formed by each pair of cameras [14]. Next, estimate of 3D-coordinates of the objects observed on the video frames is possible. But fixed cameras do not allow stable tracking of the object of interest.

Fig. 7. Experimental setup with two pairs of cameras [12]

The aim of our research is to develop a video surveillance system based on the principles of stereo vision and allows to use both fixed and PTZ cameras, to estimate the 3D-coordinates of objects of interest observed on CCTV cameras.

II. THE USE OF REFERENCE POINTS, THE PROBLEM OF EXTERNAL CALIBRATION

In our concept we use a set of fixed reference points, placed to the observed area. Points are used to position the cameras in 3D space and to estimate three-dimensional coordinates of detected objects. 3D coordinates of each reference point (in a conventional coordinate system) must be previously calculated. Such approach is not new, but we propose a method for estimate the 3D-coordinates of reference points, without using additional equipment and just using the available CCTV cameras.

Mathematical model of monocular camera ("pinhole camera") is shown in Fig. 8. Point P_0 is projected on image plane, projection has coordinates q. To calculate the coordinates of the point q, the following matrix equation are used (point q is represented in homogeneous coordinates):

$$P' = \begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = R^t\left(P_0 - T\right) = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}^t \cdot \left(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \right),$$
$$q = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z'} \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} P' = \frac{1}{Z'} MP'. \qquad (1)$$

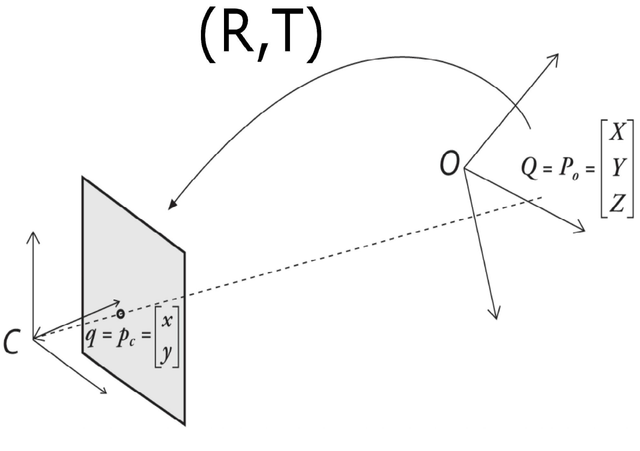

Fig. 8. Mathematical model of monocular camera

Matrix M has the intrinsic camera parameters, they are not dependent on the camera's position and orientation. The point with coordinates (c_x, c_y) is the principal point, f_x and f_y are the focal lengths, all parameters are measured in pixels.

The rotation matrix R and the parallel translation vector T are the extrinsic camera parameters. They define the orientation and position of the coordinate system attached to the camera with respect to the global coordinate system. Intrinsic and extrinsic camera parameters are estimated during camera calibration using the calibration pattern.

If the intrinsic camera parameters are known, and it is necessary to estimate the position and orientation of the camera relative to the global coordinate system, then we have the problem of the external camera calibration [14]. By solving the optimization problem (minimizing the *err* value), we need to estimate the matrix R and the vector T camera with known 3D-coordinates of a set of points (in a global coordinate system) and known coordinates of 2D-projections of these points on the image from the camera:

$$w_i\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = R^t\left(P_i - T\right), \quad q_i = M\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}, \quad R = R\left(\varphi,\psi,\theta\right),$$
$$err = \frac{1}{n}\sum_{i=1}^{n}\left\| q_i - \frac{1}{w_i} M R^t\left(P_i - T\right)\right\|^2 \to \min_{R,\ T}. \tag{2}$$

The unknown parameters are three components of the position vector and three angles that define the orientation matrix (matrix is orthonormal). There must be at least three points, the use of large number of points increases the resistance to noise when input data are noised.

Thus, the problem of estimate the 3D-coordinates of the object observed with two cameras is reduced to the problem of determining the position and orientation of two cameras on the global coordinate system. Two-dimensional coordinates of projection of reference points are searched on images. The images from the camera are 2D-coordinates of control points, 3D-coordinates of reference points are known, extrinsic cameras' parameters are estimated by solving the problem of external calibration.

Let M_1 and M_2 are the matrixes of the intrinsic parameters of cameras, $\left(R_1, T_1\right)$ and $\left(R_2, T_2\right)$ - the estimated extrinsic parameters of cameras, which watch the same object of interest, the coordinates of the projection of object in the image from the first camera are equal q_1, from the second camera - q_2. 3D-coordinates of the object (point P) are estimated by solving the following system of linear equations (the system is over-determined), w_1 and w_1 are unknown scale factors:

$$\begin{cases} w_1 q_1 = M_1 R_1^t\left(P - T_1\right) \\ w_2 q_2 = M_2 R_2^t\left(P - T_2\right) \end{cases} \tag{3}$$

Make the change of variables, and the system of equations is replaced by the following

$$\tilde{q}_1 = M_1^{-1} q_1,$$
$$\tilde{q}_2 = M_2^{-1} q_2. \quad .$$
$$\begin{cases} w_1 \tilde{q}_1 = R_1^t\left(P - T_1\right) \\ w_2 \tilde{q}_2 = R_2^t\left(P - T_2\right) \end{cases} \tag{4}$$

Express point P from the first equation of system:

$$P = w_1 R_1 \tilde{q}_1 + T_1 \tag{5}$$

and put it into the second equation:

$$w_2 \tilde{q}_2 = R_2^t\left(w_1 R_1 \tilde{q}_1 + T_1 - T_2\right),$$
$$w_2 \tilde{q}_2 = w_1 R_2^t R_1 \tilde{q}_1 + R_2^t\left(T_1 - T_2\right). \tag{6}$$

Make the change of variables, and the

$$\hat{q}_1 = R_2^t R_1 \tilde{q}_1,$$
$$\hat{T} = R_2^t\left(T_1 - T_2\right). \tag{7}$$

Then equation (6) is replaced by the following:

$$w_2 \tilde{q}_2 = w_1 \hat{q}_1 + \hat{T}. \tag{8}$$

Write the equation (8) in more detail:

$$w_2\begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ 1 \end{bmatrix} = w_1\begin{bmatrix} \hat{x}_1 \\ \hat{y}_1 \\ 1 \end{bmatrix} + \begin{bmatrix} \hat{T}_x \\ \hat{T}_y \\ \hat{T}_z \end{bmatrix},$$
$$\begin{bmatrix} w_2 \tilde{x}_2 \\ w_2 \tilde{y}_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} w_1 \hat{x}_1 + \hat{T}_x \\ w_1 \hat{y}_1 + \hat{T}_y \\ w_1 + \hat{T}_z \end{bmatrix}, \tag{9}$$

Therefore, $w_2 = w_1 + \hat{T}_z$. Equation (9) defines the following over-determined system of equations with respect to w_1.

$$\begin{cases} \tilde{x}_2\left(w_1 + \hat{T}_z\right) = w_1 \hat{x}_1 + \hat{T}_x \\ \tilde{y}_2\left(w_1 + \hat{T}_z\right) = w_1 \hat{y}_1 + \hat{T}_y \end{cases},$$
$$\begin{cases} \left(\tilde{x}_2 - \hat{x}_1\right) w_1 = \hat{T}_x - \tilde{x}_2 \hat{T}_z \\ \left(\tilde{y}_2 - \hat{y}_1\right) w_1 = \hat{T}_y - \tilde{y}_2 \hat{T}_z \end{cases}. \tag{10}$$

Make the change of variables:

$$a_1 = \tilde{x}_2 - \hat{x}_1,\ a_2 = \tilde{y}_2 - \hat{y}_1,$$
$$b_1 = \hat{T}_x - \tilde{x}_2 \hat{T}_z,\ b_2 = \hat{T}_y - \tilde{y}_2 \hat{T}_z. \tag{11}$$

As a result, the system (10) would be as follows:

$$\begin{gathered} \begin{cases} a_1 w_1 = b_1 \\ a_2 w_1 = b_2 \end{cases}, \\ \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} w_1 = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \end{gathered} \tag{12}$$

The coordinates of objects in the images are always determined with an error (the main reason is the discrete images), so the problem of solving a system of linear equations is preferable to be replaced by the following optimization problem which is solved by the method of least squares:

$$\begin{gathered} F(w_1) = \left\| \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} w_1 - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \right\| \to \min, \\ F(w_1) = (a_1 w_1 - b_1)^2 + (a_2 w_1 - b_2)^2 \to \min, \\ \frac{\partial F}{\partial w_1} = 2a_1 (a_1 w_1 - b_1) + 2a_2 (a_2 w_1 - b_2) = 0, \\ 2a_1^2 w_1 - 2a_1 b_1 + 2a_2^2 w_1 - 2a_2 b_2 = 0, \\ a_1^2 w_1 - a_1 b_1 + a_2^2 w_1 - a_2 b_2 = 0, \\ w_1 = \frac{a_1 b_1 + a_2 b_2}{a_1^2 + a_2^2}. \end{gathered} \tag{13}$$

The coordinates of point P are restored by the equation (5).

The reference points must be distributed to the observed a in such a way that in any possible position and orientation of the cameras, the number of detected reference points have been sufficient to solve the problem of external camera calibration. Also, the reference point must be clearly distinguishable from each other, i.e. must be unique.

Note that only part of all reference points can be visible in the field of view of each camera. Therefore, the use of two or more identical reference points is quite acceptable, but together they can't be observed in the image from any camera. Also note that reference points must well recognizable in the images from the cameras (video frames). Appearance of reference points can be quite arbitrary, it is necessary that the reference point is a feature point in the images.

III. BUILDING A 3D-MODEL OF THE OBSERVED AREA

The main problem of the proposed approach is that it is necessary to know the three-dimensional coordinates of reference points in a global coordinate system. To accurately model the physical building, we need to acquire the sensory data of the area, such as images and laser-scanning data; then, the polyhedral model can be generated manually from stereo images or semi-automatically from point clouds [8]. But for our task, we do not need to build a complete plan of the observed territory, we only need to calculate the 3D-coordinates of the reference points.

Proposed method for estimate the 3D-coordinates of reference points consists of several steps:

1) At first, the position and orientation of the cameras is fixed. Fig. 9 shows an example of the surveillance zone, surveillance system consists of 12 cameras (cameras 9, 10, 11 and 12 are attached to the ceiling). Also the fields of view of each camera are shown.

2) Using the calibration pattern, each sensor is calibrated individually. The result is the calculated intrinsic parameters of each camera. These data are stored and used later.

3) Next, we construct a graph whose vertices correspond to the cameras, and the presence of the edge between vertices indicates the presence of a sufficient intersection of fields of view of two cameras. After that, spanning tree of this graph is calculated. An example of one of the possible spanning trees for a surveillance system from Fig. 9 is shown in Fig. 10.

4) Using the calibration pattern, we calibrate all possible stereo rigs formed by a pair of cameras that are connected by an edge in the spanning tree (get over all the edges of the tree). Let the first camera from pair has a number *i*, second camera has a number *j*. The orientations of the cameras relative to each other (matrixes $\tilde{R}_{ij}$) and the relative positions (vectors $\tilde{T}_{ij}$) are estimated. If some point has the coordinates P_1 in the coordinate system of the camera with number *i*, then in the coordinate system of the camera with number *j*, this point will have coordinates $P_2 = \tilde{R}_{ij}^t \left(P_1 - \tilde{T}_{ij} \right)$. Because cameras can be arranged at a sufficiently large distance from each other, it require a large-sized calibration pattern (example shown in Fig. 11).

5) Next, we must select the origin of a unified conventional (global) coordinate system, and choose its coordinate axes, i.e., choose a frame. First, we must select one of the cameras (let's call it the reference camera). The frame can be linked to the position and orientation of the reference camera at the time when the calibration of cameras and stereo rigs were performed.

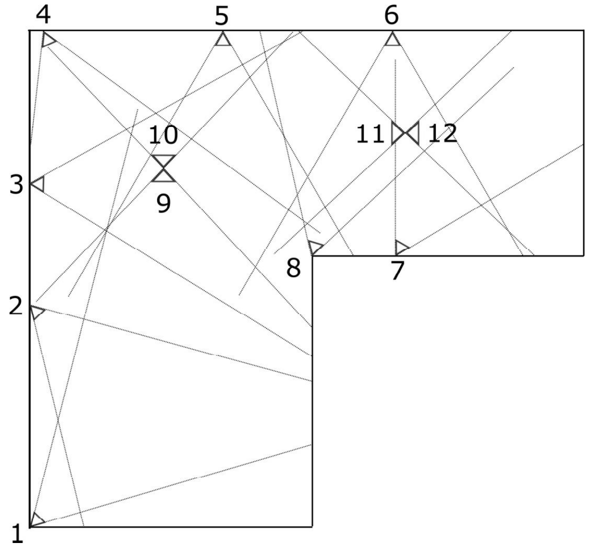

Fig. 9. Example of placing of video cameras

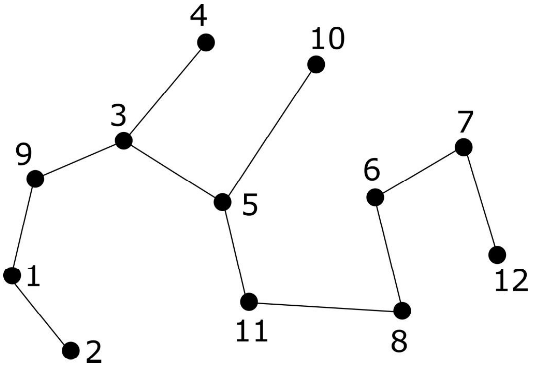

Fig. 10. Graph showing the overlapping between fields of view of the cameras

There is another approach: we can take the three-dimensional calibration pattern and located it in a certain way in the field of view of reference camera. For example, the pattern can be positioned on a wall or in a corner of the room. The frame is linked to the position and orientation of the calibration pattern. The set of keypoints are chosen on the pattern, their 3D-coordinates are measured.

Further, the photo of calibration pattern is performed by reference camera. Images of keypoints are detected on photo (manually or automatically). Using the set of pairs "2D-coordinates of the keypoint on the photo - 3D-coordinates of the keypoint in the conventional coordinate system" the external calibration problem is solved (for reference camera). After that, the pattern can be removed. An example of such calibration pattern with the selected conventional coordinate system is shown in Fig. 12.

6) Position and orientation of the all other cameras (relative to the conventional coordinate system) are calculated. All vertices of spanning tree are iterated, starting with the reference camera. Let the orientation and position of camera with number *i* are already calculated and equal R_i и T_i respectively. If some point has the coordinates P_0 in the conventional coordinate system, then in the coordinate system of the camera with number *i*, this point will have coordinates $P_1 = R_i^t\left(P_0 - T_i\right)$.

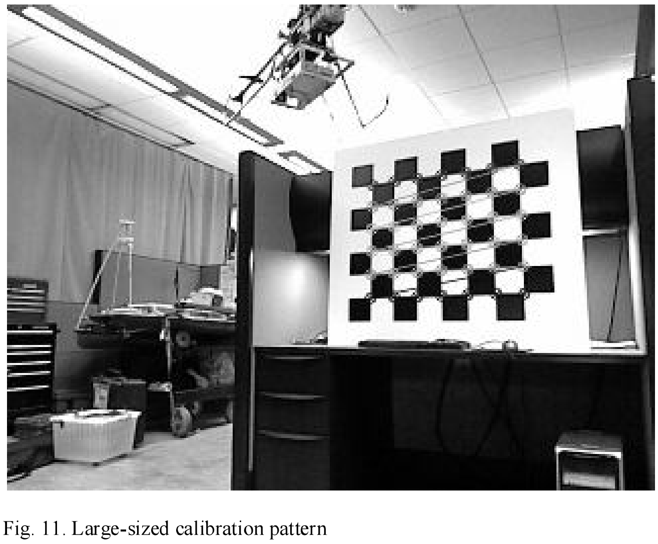

Fig. 11. Large-sized calibration pattern

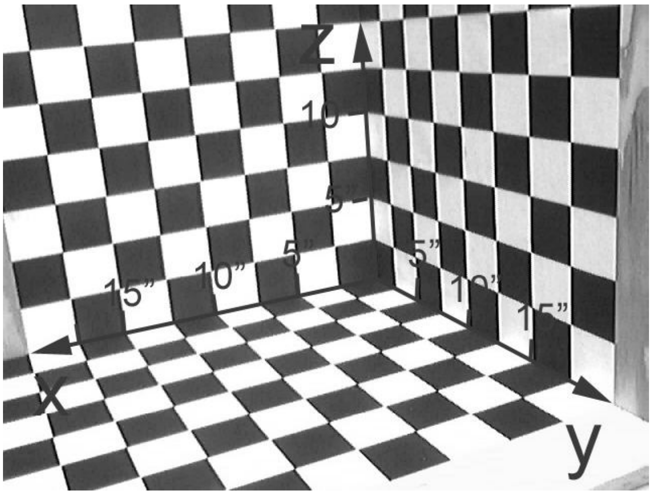

Fig. 12. Example of three-dimensional calibration pattern

Fields of view of cameras with number *i* and *j* are intersected. Calculated relative orientation and relative position between cameras are determined by matrix $\tilde{R}_{ij}$ and vector $\tilde{T}_{ij}$ respectively. Then the orientation and position of the camera with number *j* in the global coordinate system (matrix R_j and vector T_j) is given by following equations:

$$\begin{gathered}
\begin{cases} P_1 = R_i^t\left(P_0 - T_i\right) \\ P_2 = \tilde{R}_{ij}^t\left(P_1 - \tilde{T}_{ij}\right) \end{cases}, \\
P_2 = \tilde{R}_{ij}^t\left(R_i^t\left(P_0 - T_i\right) - \tilde{T}_{ij}\right), \\
P_2 = \left(R_i\tilde{R}_{ij}\right)^t \cdot \left(P_0 - \left(R_i\tilde{T}_{ij} + T_i\right)\right), \\
R_j = R_i\tilde{R}_{ij}, \\
T_j = R_i\tilde{T}_{ij} + T_i.
\end{gathered} \tag{14}$$

7) Reference points are detected on images from cameras (using some algorithm). Knowing the position and orientation of each camera, we estimate 3D-coordinates of the reference points in the conventional coordinate system.

8) The object of interests (if they are presented on observed territory) are detected on images from cameras (with some algorithm of localization and recognition of objects). Using its 2D-coordinates on images from two cameras, we estimate its 3D-coordinates in the conventional coordinate system. Information on the 3D-coordinates of reference points and 3D-coordinates of objects of interest is stored and used later.

III. ESTIMATION OF 3D-COORDINATES OF THE OBJECTS OF INTEREST

The method of estimation the 3D-coordinates of objects detected by the video surveillance systems consists of two steps. All cameras can be moving (change its position and orientation using servo) for object tracking tasks.

1) Reference points are detected on images from cameras. Using known 3D-coordinates of reference points, external calibration problem is solved for each camera: position and orientation of cameras in the conventional coordinate system are estimated.

2) The objects of interests are detected on images from cameras. Using their 2D-coordinates on images (each object must be observed on two cameras or more), we estimate their 3D-coordinates in the conventional coordinate system.

IV. CONCLUSION

The concept of video surveillance systems based on the principles and methods of stereo vision, as well as on the use of fixed and PTZ cameras has described. Concept includes a technique for estimation of 3D-coordinates of the objects observed on the images from cameras. The technique is based on a preliminary determination of the position and orientation of each camera in a global coordinate system by solving the problem of external camera calibration. This problem is solved using the fixed reference points distributed across the observed territory. Also a methodology of calculating of 3D-coordinates of reference points is described.

Currently, an experimental model of multi-camera surveillance system based on the proposed concept is under development in the Program Systems Institute of Russian Academy of Sciences. We develop our own solutions, which we have called “sensor computing node”. This solution combines a camera and Raspberry Pi microcomputer. An example of a stationary sensor computing node is shown in Fig. 13, movable nodes is under development.

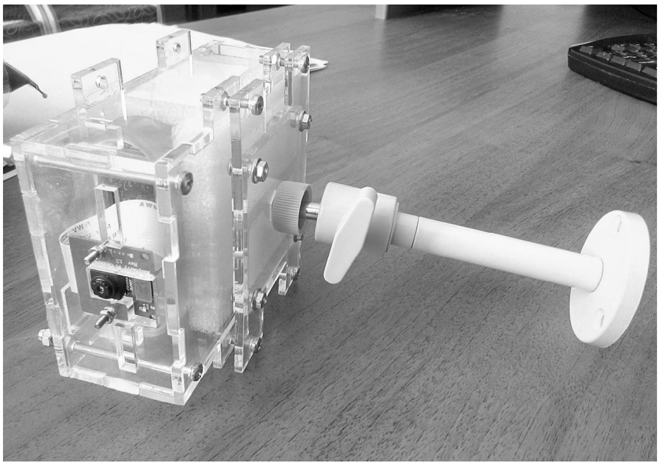

Fig. 13. Example of stationary sensor computing node

ACKNOWLEDGMENT

This work was supported by the Ministry of Education and Science of the Russian Federation (RFMEFI60714X0088) agreement for a grant on “Development of methods and means of processing and intelligent image analysis and flow of data obtained from a set of stationary and mobile sensors, using high-performance distributed computing for the tasks of monitoring the indoor placement and adjacent outdoor territories”.

REFERENCES

[1] Korea IT Times, KDC Develops the World’s First 3D CCTV, Web: http://www.koreaittimes.com/story/21899/kdc-develops-world%E2%80%99s-first-3d-cctv.

[2] LG CNS official blog, Korea’s First CCTV with 3D Video Analysis Function from LG CNS. Web: http://www.lgcnsblog.com/features/koreas-first-cctv-with-3d-video-analysis-function-from-lg-cns.

[3] FOX News, Could 3D technology revolutionize surveillance? Web: http://www.foxnews.com/tech/2014/09/04/could-3d-technology-revolutionize-surveillance.html.

[4] University of Wisconsin-Madison, Photometric Stereo. Web: http://pages.cs.wisc.edu/~csverma/CS766_09/Stereo/stereo.html.

[5] buyKOREA, twinEye - Analytical 3D CCTV Solution. Web: http://www.buykorea.org/product-details/twinEye-Intelligent-CCTV-Analytical-3D-cctv-Solution-Stereo-CCTV--3025835.html.

[6] Smart CCTV – Intelligent Video Systems, People counting. Web: http://www.smartcctvltd.com/traffic-products-and-surveys/video-analytics/people-counting.

[7] S. Oh, J. Won, G. Kim, “Development of a vision based railway platform safeguard”, *in Proc. 9th World Congress on Railway Research.*, May. 2011, 10 p.

[8] K. Choi, I. Lee, “CCTV Coverage Index Based on Surveillance Resolution and Its Evaluation Using 3D Spatial Analysis”, *Sensors* vol.15(9), 2015, pp. 23341-23360; doi:10.3390/s150923341.

[9] K. I. Moon, M. W. Pyeon, S. B. Lee, D. R. Lee, "Setting of the Operating Conditions of Stereo CCTV Cameras by Weather Condition", *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*. vol.32(6), Dec. 2014. pp. 591-597, doi: 10.7848/ksgpc.2014.32.6.591.

[10] G. de Haan, J. Scheuer, R. de Vries, F. H. Post, "Egocentric Navigation for Video Surveillance in 3D Virtual Environments", *3D User Interfaces*, *IEEE Symposium on*, 14-15 March 2009, pp. 103-110, doi: 10.1109/3DUI.2009.4811214.

[11] K.-S. Park, *Robust Object Detection and Localization for Real-Time Autonomous Surveillance Applications.* PhD dissertation, Stony Brook University, 2008.

[12] A. K. Mishra, B. Ni, S. Winkler, A. Kassim, "3D Surveillance System Using Multiple Cameras", *in Proc. of SPIE 6491, Videometrics IX*, 29 Jan. 2007, 8 p., doi: 10.1117/12.703449.

[13] Z. Zhang, A. Scanlon, W. Yin, L. Yu, P. L. Venetianer, "Video Surveillance using a Multi-Camera Tracking and Fusion System", *Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications - M2SFA2 2008*, Oct. 2008, 13 p.

[14] G. Bradski, A. Kaehler, *Learning OpenCV.* O'Reilly Media., 2008.