Application of the projective geometry in the density mapping based on CCTV monitoring

Agata Chmielewska, Parzych Marianna,
Tomasz Marciniak, Adam Dabrowski
Poznan University of Technology,
Division of Signal Processing and Electronic Systems
Poznan, Poland
chmielewska.agata@gmail.com,
{parzych.marianna, tomasz.marciniak,
adam.dabrowski}@put.poznan.pl

Przemyslaw Walkowiak
Poznan University of Technology,
Institute of Control and Information Engineering
Poznan, Poland
przemyslaw.walkowiak@put.poznan.pl

***Abstract*—In this paper we address a problem of automatic generation of so-called density maps (heat maps) from video recordings acquired surveillance systems (CCTV – closed circuit television) in urban areas. Our software for density map estimation increments the accumulation table only in case of detecting a new moving object, resulting in a count of moving objects density map. We also applied an algorithm for projective transformation of image to achieve better accuracy of the moving object density map. We analyzed the proposed solution for different angles of camera position and next the results are compared with density maps calculated from the top view camera position.**

Keywords- behaviour analysis, density (heat) maps, intelligent video analysis, projective transformation, homography, CCTV

I. INTRODUCTION

In modern cities highly advanced systems of video surveillance are often installed. Video cameras can be found both on the streets and on/in the buildings. It should, however, be noted that most monitoring solutions are used mainly for data acquisition, which, if necessary, are manually later analyzed. Unfortunately, a detailed analysis of CCTV (closed circuit television) sequences and the development of data, requires operators and analysts to spend long hours watching recordings.

Automatic analysis of the movement of people and vehicles in cities makes it possible to obtain important information about the behavior of people, crowd, and the intensity and variability of a city traffic. This data can be widely used to improve traffic, increase safety and ergonomics of urban space. Information about traffic and its changes depending on the time of day, week, or year allows for the optimum route planning and control of traffic lights. Pedestrians traffic analysis: determining the routes chosen by them, the areas there they stop or gather, the spaces more or less crowded allows for designing space (paths and sidewalks, benches and light distribution) that responds to the people needs. Crowds density estimation in public spaces, such as stadiums, shopping malls, parks allows for increasing security by detecting and removing the narrow passages, crowded places difficult to leave, sub-optimal evacuation routes, etc.

Many commercial companies are now offering systems for video data analysis [1], whereby they cannot always make a precise processing and correct interpretation of the collected data. Simultaneously with expansion of the number of cameras in the urban environment and increased observation of citizens, we can observe rising number of algorithms for face detection [2, 3] and blurring. These algorithms allow keeping privacy of people living in urban areas.

Our solution for moving object tracking and motion analysis estimates of objects density and is designed to test large-sized scenes, for example, streets or large rooms in the buildings. The video analysis is based on an information about the trajectories of moving people and generates maps that show the people density. Our previous research about density maps were published in [4, 5] – this papers present two types of heat maps (the latter name is used because of the common heat color scale): the first one shows the people density per time unit, averaging the instantaneous density in the individual frame and the second one shows how many objects were at the specified location. This papers describe also in detail a method for choosing the adequate reference map (scale) in relation to the length of the processed video sequence and present a selection of camera perspective to achieve optimal results.

The next part of the paper presents issues related to the projective transformations. This transformation can be used in the situation, where it is not possible to place camera at a large height or in such a way, that the camera observes the scene from the top view. Visualization of the projective transformation from the side view to the top view on the parking lot was shown in Fig. 1. In this figure was marked area that will be transformed and its equivalent after transformation.

We have verified a correctness of the proposed solution by performing of two experiments. The first one on a micro scale with an artificial scene (Fig. 4 and 5) in order to efficiently examine the accuracy of the projective transformation. The second experiment was carried out at the parking lot (Fig. 1 and 8). The data from the side view was transformed to the view from the top perspective, which is preferred

by security and space planning companies. Therefore we have applied a homography (a projective transformation).

The paper is structured as follows. Section 2 describes the related work both for moving objects density maps and for projective mapping. Section 3 presents the algorithm of moving objects density estimation while section 4 described definition and estimation of projective transformation. Finally, section 5 shows experimental results and section 6 concludes the paper with the summary and outlook.

Fig. 1. Visualization of the projective transformation from the side view to the top view

II. Related Work

A. Crowd Density Maps

Estimation of distribution of objects is inter alia related to the crowd analysis. It is an issue studied by many research teams, who already presented various approaches to the topic [6]. Considering the size of the crowd on the basis of the number of people is an important issue for human safety [7]. Authors of the paper [8] divide studied motion in order to assess the density of the crowd into five different density levels. However, in this paper authors focus on solutions considering not only on the crowd density estimation, but also its visualization using maps of density.

Some of the solutions are based on the analysis of stationary scene from the video sequence. Estimation of the crowd density described in [9] is done with the use of the texture information based on the grey level transition probabilities. The crowd density estimation using texture features was extended with detection of abnormal density distribution in [10]. Authors of [11] proposed the Translation Invariant Orthonormal Chebyshev Moments method based on pattern recognition.

Different solutions are based on collecting the data about the movement of objects. The authors of [12] use optical flow for foreground objects detection. In [13] optical flow is extended with the information about objects based on head detection. In [14] crowd density is estimated on the basis on Lucas-Kanade optical flow feature tracks, and the density map is computed with the use of Gaussian kernel density. Generation of motion maps applied to the crowd monitoring system on the escalator is presented in paper [15]. A detailed analysis of the motion of objects (features) allows to estimate the density of the crowd also in the case of non-static cameras [16].

B. Inverse Perspective Mapping

Another issue that had to be resolved in our approach was a transformation of the camera view. Many CCTV systems use cameras placed in such a way, that the resulting recordings show the test area in perspective view. It is necessary to transform the camera view from the 'bird-view' to the 'top view' by inverse perspective mapping. Inverse perspective mapping is commonly used in applications for smarts cars based on the video sequences obtained from car cameras, eg. lane tracking systems [17, 18], parking assistants [19] or automatic distance determination systems [20].

There are different methods for inverse perspective mapping. Some are based on constant elements in the image like road marks [21]. Many solutions are based on vanishing point estimation: in [22] the histogram-based segmentation is used for features extraction and in [23] point estimation is based on texture analysis.

III. Density Map Generation

Process of estimation of the objects distribution (people or vehicles) consists of the object detection (this aspect is not considered here) and calculation of the object occurrence density. An illustrative density map showing the distribution of vehicles in the city parking lot is shown in Fig. 8. It makes quick distinction between the areas of high or low objects density possible.

The information about detected objects is used to generate maps that show the moving objects density. The process of calculating the map is as follows. Moving objects are detected with the use of background subtraction. Each detected moving object is described by the coordinates of its center. The Kalman filter allows expanding the system with detection of objects trajectories [5]. For object detection in the video sequence, the respective cell in the two-dimensional accumulation matrix is incremented. After processing of a given time segment of video sequence the density map is generated based on the values in the accumulation matrix, which are assigned to the chosen (e.g. the heat) color scale.

Our method of people density and traffic analysis assumes the generation of two types of the density maps [5]. The schema of the algorithm for density maps generation was shown in Fig. 2. Both types present how long different areas were occupied by moving objects. However, a difference

between them is as follows: in the case of the first type of the map (based on time) a cell in the accumulation matrix is incremented every time when a moving object is detected in a given location, while in the case of the second type of the map (based on observation of the moving objects indexes) it is incremented only once for each moving object in a given location. In this paper we show examples only from the second type of density maps – based on observation of moving objects indexes (Fig. 2).

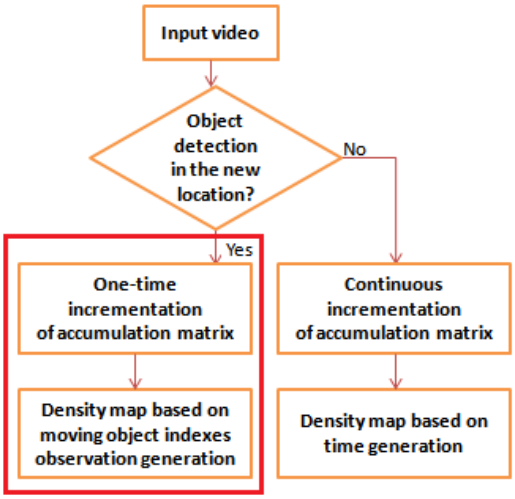

Fig. 2. Schema of the algorithm for density maps generation (based on time and based on observation of moving objects indexes)

IV. Homography

A. Definition

The projective transformation is a mapping between two planes, e.g. between image and world plane. In our case we had to find a mapping between two images of the same scene (from side and top view). It can be realized by determination of two projective transformation from one image to world plane and from world plane to second image (Figure 3).

The homography is defined as a following transformation [24]:

$$\mathbf{x}' = \mathrm{H}\mathbf{x} \tag{1}$$

where H is a block matrix of form

$$\mathrm{H} = \begin{bmatrix} \mathrm{A} & \mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix} \tag{2}$$

where linear component A, in form of 2×2 non-singular matrix, represents an affine transformation (rotation and scaling) and vector $\mathbf{t}$ represents a translation and vector $\mathbf{v} = [v_1, v_2]^T$. The v variable is only for matrix scaling purpose.

The chain of multiple transformation can be composed into one.

$$\mathbf{x}' = \mathrm{H}_1\mathbf{x} \tag{3}$$

$$\mathbf{x}'' = \mathrm{H}_2\mathbf{x}' \tag{4}$$

$$\mathbf{x}'' = \mathrm{H}_2\mathrm{H}_1\mathbf{x} = \mathrm{H}_3\mathbf{x} \tag{5}$$

Therefore, only one transformation matrix have to be computed to map between two images of the same scene.

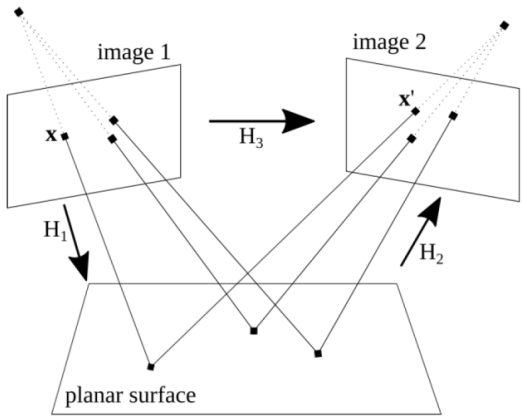

Fig. 3. The projective transformation between two images induced by world plane [24]

B. Estimation

A projective transformation matrix can be computed using Direct Linear Transformation (DLT) algorithm [24] which solves Equation 1 using a given set of four point correspondences. However, the algorithm requires that no three points are collinear on either plane.

V. Experiments

A. Projective transformation evaluation

The goal of the experiment is to show how accurate is projective transformation applied to the heat map, therefore, we prepared a testing environment with two cameras (Microsoft LifeCam Studio) looking at the scene from two different angles (Fig. 4). The angle of first camera was modified across multiple experiments from 30 degree to 50 degree (the α angle in Fig. 4) with an interval of 5 degree. The second camera was positioned in the straight down direction (the top view). The little hexbug robots were used in the experiment as moving objects in the scene. A results of the experiment is data of the same scene and movement recorded simultaneously from two different views.

An accuracy of the projective transformation of heat map was measured and compared using correlation as a similarity measure between transformed heat map recorded from side view and non-transformed heat map recorded from top view. At first, we computed correlation coefficient between whole marked regions. The results are in Table 1. The overall values of the measures drop from 0.792 to 0.671 across ascending angle of the side camera, which demonstrate difficulty of view transformation in case of when the image was taken at a large angle. In the next step, we divided each heat map into four equals sectors and compared them respectively (see Table 1).

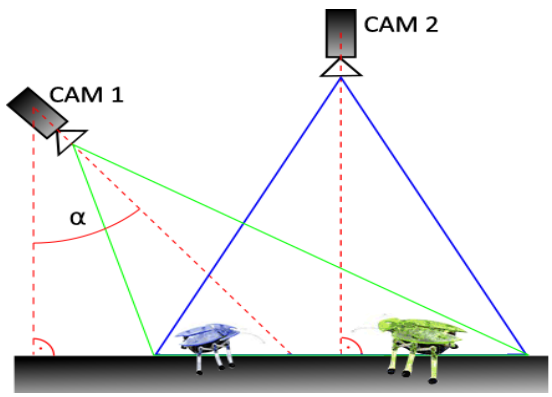

Fig. 4. Cameras' positioning during the experiment

Additionally, to show a distortion we computed cross correlation between sectors and the whole region and we found the best fitting regions. As we can see in Figure 6 and Table 2 the best fitting sectors are shifted. The values of the correlation coefficient of the best fitted sectors and whole region can be found in Table 2 and are higher than in the direct comparison (Table 1).

The coordinate system for density map images and individual sector of heat maps is defined in the upper left corner. Table 2 shows, that between transformed heat map (and their sectors) recorded from the side view and non-transformed heat map recorded from the top view there is vertical offset (Fig. 7) and horizontal offset is negligible. The average vertical offsets in the sectors from 1 to 4 are, respectively, 11.2, 7.2, 9.6, 6.4. This underlines the fact, that the areas which are closer to the camera (sectors 1 and 3) have a greater offset than regions, which are further from the camera (sectors 2 and 4).

Table. 1. Correlation coefficients between transformed heat map (and their sectors) recorded from side view and non-transformed heat map recorded from top view

Angle (degrees)	Heat map area				
	Sect. 1	**Sect. 2**	**Sect. 3**	**Sect. 4**	**Whole region**
30	0.645	0.767	0.786	0.883	0.792
35	0.711	0.855	0.723	0.814	0.795
40	0.802	0.725	0.739	0.804	0.767
45	0.686	0.624	0.766	0.682	0.672
50	0.674	0.680	0.683	0.692	0.671

Table. 2. Selected cross correlation coefficients and offsets between transformed heat map (and their sectors) recorded from side view and non-transformed heat map recorded from top view

Angle (degrees)	Heat map area	Cross correlation coefficient	Horizontal offset (X axis)	Vertical offset (Y axis)
30	Sect. 1	0.768	5	9
	Sect. 2	0.791	3	-1
	Sect. 3	0.798	0	4
	Sect. 4	0.889	-2	1
	Whole region	0.800	2	2
35	Whole region	0.825	0	6
40	Whole region	0.806	2	8
45	Whole region	0.720	2	7
50	Sect. 1	0.810	-5	14
	Sect. 2	0.825	-6	12
	Sect. 3	0.744	3	12
	Sect. 4	0.853	4	17
	Whole region	0.777	0	14

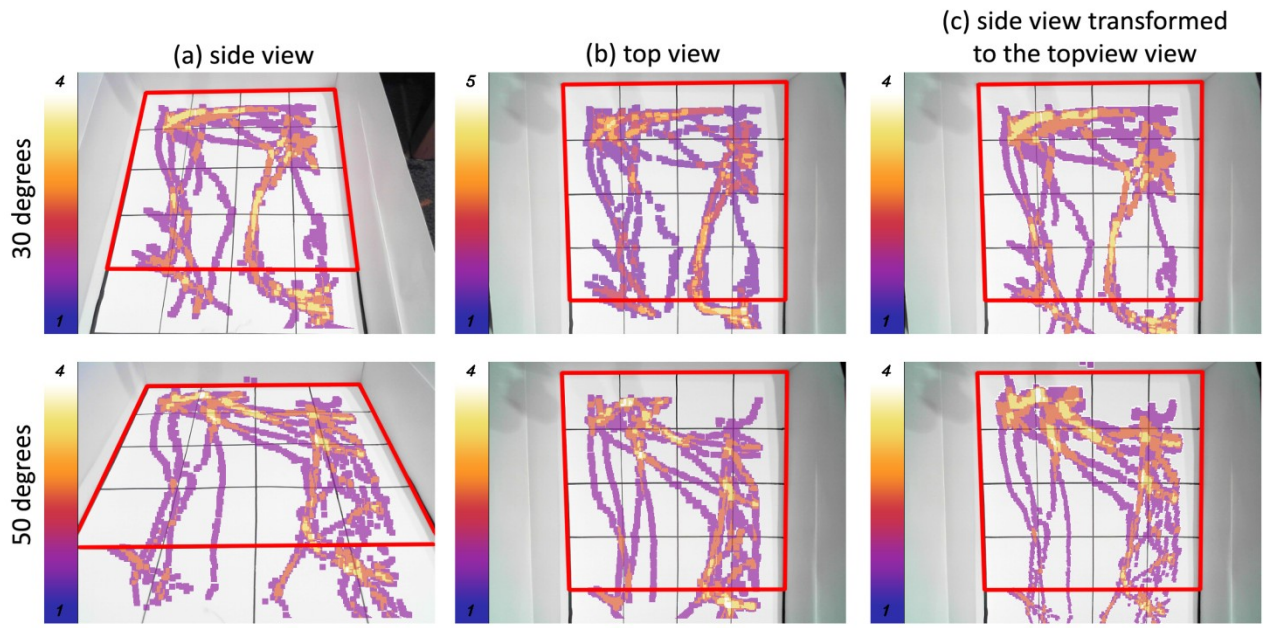

Fig. 5. From the left: heat maps from camera no 1 (side view) positioned at angle of 30 and 50 degrees; heat maps from camera no 2 (top view); heat maps from the side view transformed to the top view with the use of projective transformation

From the Fig. 7 can also be seen, that the vertical offset value increases with the increase of the α angle (side view, camera no. 1, Fig. 4), but for the video sequence resolution of 640×480 pixels, this offset can be omitted, because it is within 2.9% of error.

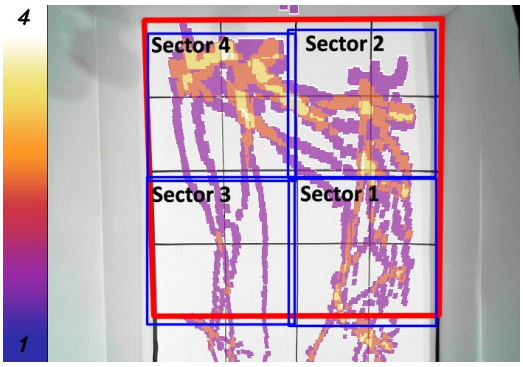

Fig. 6. Example of heat map sectors offset with highest correlation

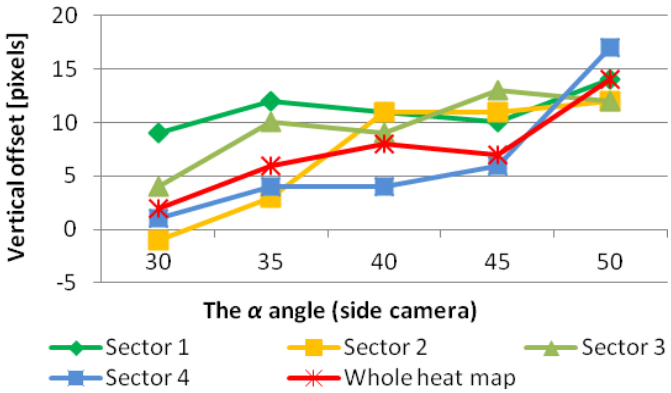

Fig. 7. Vertical offset between transformed heat map (and their sectors) recorded from side view and non-transformed heat map recorded from top view

B. *Application of the projective transformation*

The second experiment concerned the generation of density maps from CCTV data. Figure 1 (top part) shows illustrative image from video sequences used during tests for moving objects density estimation. Recordings were carried out against shopping center under real-life conditions. The camera field of view includes a parking lot, an entrance to the shopping center and other retail outlets. The shopping center in the day of recordings was opened from 6AM to 9PM. For this reason eight one-hour length video recordings were registered, starting from 6AM with an hour interval between each video sequence.

Figure 8 shows the density map based on observations of moving objects indexes from the shopping center parking lot for the whole day. Because each relevant moving object is tracked and only one path for each object is generated, the number of moving objects passed along the path can be determined. The maximum value in this density map is 871 – this value (indicated together with some others in Fig. 8) means that in this place there were 871 relevant moving objects. From the numbers of moving objects passing through a given point (Fig. 8) we can e.g. conclude about evacuation ways, marketing statistics, etc.

The next step is to transform a heatmap to showing the data from top view. However, to perform this step a homography matrix had to be computed first. Therefore, a set of four corresponding points were selected for each view of scene. The computations were performed by the OpenCV library [25]. The results of the projective transformation can be seen in Fig. 8.

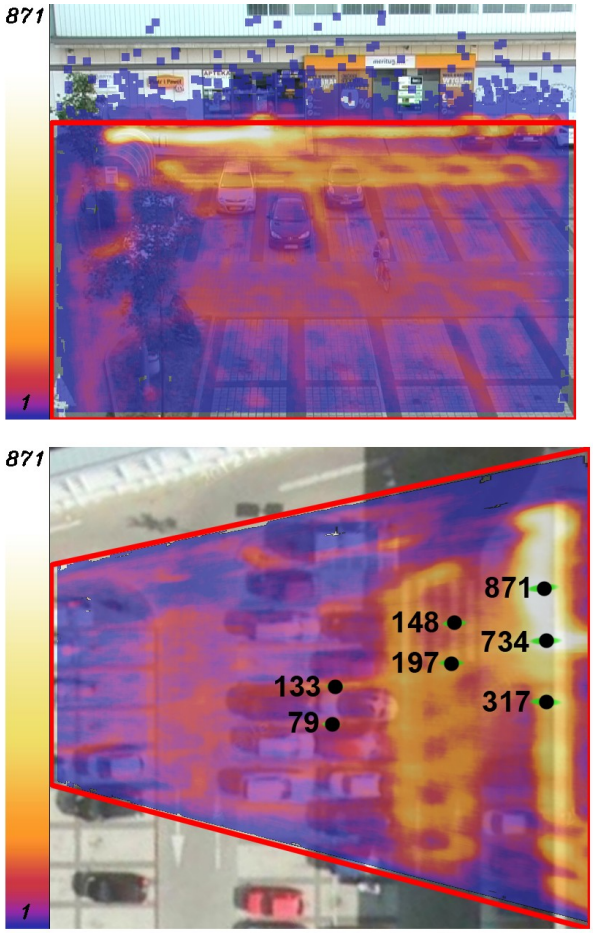

Fig. 8. Moving objects density map based on observation of moving objects indexes from the whole-day long video sequence; side view (top part of the image) of the scene and top view (bottom part of the image) with the marked numbers of moving objects

VI. Conclusions

Intelligent video analysis in CCTV systems allows for a definite improvement of the people safety in smart cities [26-28]. Commercial solutions are already available for the analysis of human behavior, crowd density estimation, or people counting [29]. Furthermore, it should be noted, that there is a need for improvement of the existing analysis methods of and for the development of new algorithms.

The proposed density estimation (based on indexed objects observation density maps) gives the possibility of precise monitoring of the public areas.

The projective transformation of image from side view camera positioning can be very useful in areas where, is no possibility to mount the camera directed vertically downward.

The shift of the best fitted sectors is related to the method of detection and tracking of objects in the algorithm for heat map generation. Both heat maps (side and top view) were generated using the center of the object found in the scene. However, the location of the found center of the same object is different in the image made from side and from top (for this view is very close to the real center of the object). To prevent this shift, and to increase accuracy of the projective transformation, a modification of the detection and tracking algorithm is needed to correctly recognize a center of objects.

ACKNOWLEDGMENT

This work was supported partly by the 04/45/DSPB/0136 and 09/93/DSPB/512 project means and partly by the Scholarship for Ph.D. students specializing in major strategic development for region Wielkopolska, Poland, Sub measure 8.2.2 Human Capital Operational Program, co-financed by the European Union under the European Social Fund.

REFERENCES

[1] Retail Surveillance, Heat Maps http://retail-surveillance.com/merchandising/solutions/diy/heat-maps.htm, access: 02/12/2015

[2] T. Marciniak, R. Weychan, A. Chmielewska, A. Dabrowski, "Influence of Pose Angle on Face Recognition from Very Low Resolution Images", *IEEE NTAV/SPA 2012, New Trends in Audio and Video / Signal Processing Algorithms, Architectures, Arrangements and Applications,* pp. 177–181, 2012

[3] T. Marciniak, A. Chmielewska, R. Weychan, M. Parzych, A. Dabrowski, "Influence of low resolution of images on reliability of face detection and recognition," *Multimedia Tools and Applications (2015), No. 74*, pp. 4320-4369, 2015

[4] M. Parzych, A. Chmielewska, T. Marciniak, A. Dabrowski, A. Chrostowska, M. Klincewicz, "Automatic people density maps generation with use of movement detection analysis", *International Conference on Human System Interaction,* pp. 26-31, 2013

[5] M. Parzych, A. Chmielewska, T. Marciniak, A. Dabrowski, "New Approach to people Density Estimation based on Video Surveillance", *IEEE Int. Conf. on Signal Processing Algorithms, Architectures, Arrangements, and Applications,* pp. 94-99, 2014

[6] B. Zhan, D. Monekosso, P. Remagnino, S. Velastin, and L. Xu, "Crowd analysis: A survey," Machine Vision and Applications, Vol. 19, No. 5, pp. 345–357, 2008

[7] P. Karpagavalli, A.V. Ramprasad, "Estimating the density of the people and counting the number of people in a crowd environment for human safety", International Conference on Communications and Signal Processing, pp. 663-667, 2013

[8] H. Fradi, Xuran Zhao, J. Dugelay, "Crowd density analysis using subspace learning on local binary pattern", IEEE International Conference on Multimedia and Expo Workshops, pp. 1-6, 2013

[9] A. Marana, M. Cavenaghi, R. Ulson, F. Drumond, "Real-Time Crowd Density Estimation Using Images," Advances in Visual Computing, Vol. 804, pp. 355-362, 2005

[10] Xinyu Wu, Guoyuan Liang, Ka Keung Lee, Yangsheng Xu, "Crowd Density Estimation Using Texture Analysis and Learning," Bulletin of Advanced Technology Research, Vol. 3, No. 5, pp. 35–41, 2009

[11] H. Rahmalan, M. Nixon, J. Carter, "On Crowd Density Estimation for Surveillance," The Institution of Engineering and Technology Conference on Crime and Security, pp. 540–545, 2006

[12] M. Patzold, R. H. Evangelio, T. Sikora, "Counting people in crowded environments by fusion of shape and motion information," IEEE Int. Conf. on Advanced Video and Signal Based Surveillance, pp. 157–164, 2010

[13] Cao Lijun, Huang Kaiqi, "Video-based crowd density estimation and prediction system for wide-area surveillance," China Communications, Vol. 10, No. 5, pp. 79–88, 2013

[14] H. Fradi, J. Dugelay, "Crowd Density Map Estimation Based on Feature Tracks", IEEE 15th International Workshop on Multimedia Signal Processing, pp. 40-45, 2013

[15] H. Sharif, N. Ihaddedene, C. Djeraba, "Crowd Behavior Monitoring on the Escalation Exits," Int. Conf. on Computer and Information Technology, pp. 194–200, 2008

[16] T. Senst, V. Eiselein, I. Keller, T. Sikora, "Crowd analysis in non-static cameras using feature tracking and multi-person density", IEEE International Conference on Image Processing, pp. 6041-6045, 2014

[17] A.M. Muad, A. Hussain, S.A. Samad, M.M. Mustaffa, B.Y. Majlis, "Implementation of inverse perspective mapping algorithm for the development of an automatic lane tracking system", IEEE Region 10 Conference TENCON, Vol. 1, pp. 207-210, 2004

[18] Dongwook Seo, Kanghyun Jo, "Inverse Perspective Mapping based road curvature estimation", IEEE/SICE International Symposium on System Integration, pp. 480-483, 2014

[19] Chien-Chuan Lin, Ming-Shi Wang, "Topview Transform Model for the vehicle parking assistance system", International Computer Symposium, pp. 306-311, 2010

[20] S. Tuohy, D. O'Cualain, E. Jones, M. Glavin, "Distance determination for an automobile environment using Inverse Perspective Mapping in OpenCV", Signals and Systems Conference, pp. 100-105, 2010

[21] Hua Li, Mingyue Feng, Xiao Wang, "Inverse perspective mapping based Urban road markings detection", EEE 2nd International Conference on Cloud Computing and Intelligent Systems, Vol. 3, pp. 1178-1182, 2012

[22] M. Nieto,L. Salgado, F. Jaureguizar, J. Cabrera, "Stabilization of Inverse Perspective Mapping Images based on Robust Vanishing Point Estimation", IEEE Intelligent Vehicles Symposium, pp. 315-320, 2007

[23] Daiming Zhang, Bin Fang, Weibin Yang, Xiaosong Luo, Yuanyan Tang, "Robust inverse perspective mapping based on vanishing point", International Conference on Security, Pattern Analysis, and Cybernetics, pp. 458-463, 2014

[24] R. I. Hartley, A. Zisserman, "Multiple View Geometry in Computer Vision, 2nd ed.", Cambridge University Press, pp. 39-43, ISBN: 0512540518, 2014

[25] A. Bradski, "Learning OpenCV, computer vision with opencv library; software that sees", 2008

[26] J. Balcerek, A. Chmielewska, A Dabrowski, D. Jackowski, A. Konieczka, T. Marciniak, P. Pawlowski, "Recognition of threats in urban areas by means of the analysis of video sequences", *Multimedia Communications, Services and Security 2010,* pp. 41-48, 2010

[27] A. Chmielewska, A. Dabrowski, A. Namerla, P. Pawlowski, R. Weychan, M. Stankiewicz, „Comparison of NI Lab VIEW and NI Vision Builder AI environments in fast prototyping of video processing algorithms for CCTV using smart camera", *Electronics – constructions, technologies, applications,* No. 5, pp. 72-76, 2011

[28] A. Chmielewska, R. Weychan, T. Marciniak, A. Dąbrowski, M. Hartwich, M. Owczarczak, "Fast prototyping for Video Monitoring Systems with the Use of DSP Module", *International journal of Electronics and Telecommunications,* Vol. 59, No. 4, pp.375-381, 2013

[29] T. Marciniak, A. Dabrowski, A. Chmielewska, M. Nowakowski, "Real-Time Bi-Directional People Counting Using Video Blob Analysis", *IEEE New Trends in Audio and Video / Signal Processing Algorithms, Architectures, Arrangements and Applications*, pp. 161-166, 2012