Intelligent Video Surveillance System Evaluation Methods

J. Sevcik, P. Svoboda, A. Paduchova

***Abstract*—**Quality of entire Intelligent Video Surveillance System plays significant role in its effectiveness. Moreover, these systems are deployed still more often as a tool for regulation of crime rate around the world. A number of cameras deployed are continuously growing and this situation has created a gap between possibilities and the reality in the quality of Intelligent Video Surveillance Systems. However, the measurement of Intelligent Video Surveillance effectiveness is relatively complex task. For this purpose the evaluation methods which utilized combination of Image Acquisition and Image Processing evaluating methods were proposed in this paper.

***Keywords*—**Intelligent Video Surveillance, Evaluation, Methods, Image Processing, Image Acquisition, Effectiveness;

I. Introduction

NUMBERof Video Surveillance System (VSS) applications is increasing constantly in last three decades. Moreover, the technological evolution have influenced all major parts of these systems, therefore the particular gap has arisen between the technological and design possibilities and shape of real installations of VSS, especially in case of its Image Properties. The first recommendation within a VSS design is to follow instructions formulated by the European Standards related to objective area nonetheless there are only a very small amount of companies, which have made efforts to design VSS in agreement with the particular standards. This situation is partially initiated by contemporary adverse economic climate in one hand and insufficient level of expertness of VSS designers on the other hand. Another reason to install obsolete VSS components is business task. A lot of VSS were designed and installed under these conditions, especially in Czech Republic. In case of large-scale VSS, like an Urban Video Surveillance System (UVSS) the importance of the problem even increases, therefore each system should be evaluated before it will be implemented into practice. The evaluations of VSS were discussed in several research papers [1]. Although particular ones discussed evaluation of the Image Acquisition process [2], the majority were aimed to evaluation of overall architecture of the system [3], including all functional blocks such as Image Acquisition, Connections, Image Processing, Activity and data management, connections with other systems, the system and data integrity. The problem of UVSS effectiveness was investigated in chosen research papers [4]. Nonetheless the recent academic activity is related to VSS intelligence. [1] As a new methods of video content analysis are proposed the new approaches to measure or evaluate them are intensively explored. VSS enhanced with Video Content Analysis (VCA) and Video Event Understanding (VEU) is the most progressive research area in last few years. Successfully deploying Video Content Analysis (VCA) solutions for urban surveillance poses significant challenges for manufacturers and system integrators. Urban surveillance is typically characterized by a very large number of cameras (thousands and more) distributed over a large area and installed in both outdoor and indoor views. From the user perspective the primary rule of the surveillance system is to provide quick, reliable and high quality access to live and recorded video streams from all cameras. VCA is considered an important but secondary functionality that is required in order to provide features such as real time alerts for predefined rules, forensic search capabilities, statistical analysis of crowd and traffic flow and more. [5] General principles of sophisticated VSS design are based on comprehensive formulation of the operational requirements. Considering this statement it is obvious that even in case of the Intelligent Urban Video Surveillance System is necessary to specify operational requirements related to Image Operational Properties. The common way how to obtain the operational requirements is qualitative expert estimation supported by purposes of the system defined by the user, however these methods are considered as insufficient chiefly because of that is impossible to compare qualitative description with reality. This research paper provides a novel approach how to measure the light conditions of particular levels of observed scene. In the problem formulation part the main features are described. Furthermore follows the problem solution section, where the novel approach to measurement is introduced.

J. Sevcik is with the Tomas Bata University in Zlin. He is now with theDepartment of Security Engineering Tomas Bata University in Zlín, nám. T.G. Masaryka 5555, CZECH REPUBLIC, (corresponding author to 774 202 346; e-mail: jsevcik@fai.utb.cz).

P. Svoboda is with the Tomas Bata University in Zlin. He is now with the Department of Security Engineering Tomas Bata University in Zlín, nám. T.G. Masaryka 5555, CZECH REPUBLIC (e-mail: psvoboda@fai.utb.cz).

A. Paduchova is with the Department of Security Engineering Tomas Bata University in Zlin (e-mail: paduchova@fai.utb.cz).

II. State of the Art

The Evaluation of IVSS is relatively complex task. At first it is necessary to define exactly functional blocks of the system, which influence its effectiveness the most. Analyses which were realized in previous research have referred to unacceptable form of project documentation. European Standard EN 50 132 – 7 provides in its appendixes inspectional guideline for evaluation of the capability to recognize level of detail of the particular camera, but it lack needed case studies and rules for the applications. Moreover, it determines examinations methods to evaluate compliance of the system with person identification and vehicle identification number criteria. Despite of existence of these rules executing companies did not follow them. The situation has requested any rules which will regulate and guarantee quality needed. The whole problem is even more important in case of this particular area of alarms systems, because of its disposition to personal privacy. This research paper describes possible evaluation methods of particular functional properties of IVSS. As was described in section 1, it is necessary to propose evaluation method for these underestimated categories of functional properties: Image Acquisition, Image Processing and Architectural design from the perspective of whole system.

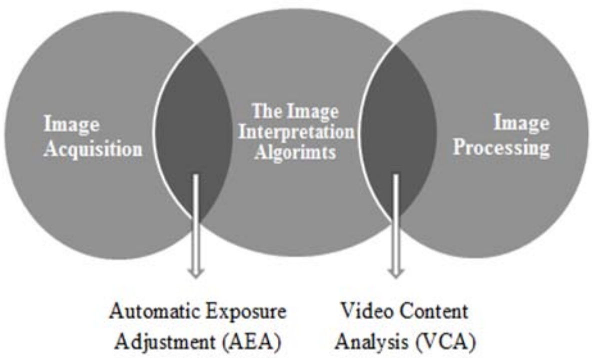

Fig.1: Image Operational Properties

III. Image operational properties

Quantitative definition of Operational Requirements (OR) within the IVSS has not been provided so far, particularly those related to the Image Functional Properties. I.As an initial step to design new evaluation approach, it is necessary to provide determination of sectional blocks of Image Functional Properties. There are to general functional sections:

- Image Acquisition,
- Image Processing;

Recently, both of these are influenced by the System Intelligence, which is realized through wide spectra of Image interpretation Algorithms.

A. Image Acquisition

The initial functional step of VSS is an Image Acquisition by the system sensing element which in this case is a Surveillance camera. The primary ability of the camera is gathering the Image information which is then transmitted through the communication interface, nonetheless the first generations of the VSS was enhanced by these major functionalities. Contemporary possibilities of the VSS advanced a lot, mostly because of the Convergence between CCTV and the area of Information and Communication Technologies (ICT). Moreover, the applicability of VSS has been increasing due to these new possibilities. For the purposes of the evaluation we could imagine Surveillance camera as a set of optical elements by which the exposed scene are scanned, whereas the field of view of the camera are depended on following parameters:

- The Image Sensor format,
- Focal length of lenses,
- Horizontal and vertical resolution of camera;

Field of view of camera is in three-dimensional (3D) characterized as a polygon, which could be divided into several segments as you can see in Figure 2. Segment marks the area between two particular distances from the camera and serves to express the particular levels of detail in the scene. Specification of concrete parameters is formulated in the European standard [6], where do particular metrics are described.

Inspect - If the OR requires the system to be capable of achieving the inspection category then the target should be reproduced at no lower resolution than 1 mm per pixel and the target should represent a screen height of 400 % PAL or equivalent identified in Figure 3.[6]

Facial Identification - If the OR requires the system to be capable of producing images suitable for identifying faces then Annex B shall be used. The test consists of 9 human faces. It is designed to assess a CCTV system's ability to identify faces. A random selection is presented towards the camera at a prescribed person percentage screen height or random selection is presented towards the camera at a prescribed person percentage screen height or which are recorded and evaluated. The OR will specify the percentage screen height at which a pass shall be achieved. [6]

Recognize - If the OR requires the system to be capable of achieving the recognition category then the target should be reproduced at no lower resolution than 8 mm per pixel and the target should represent a screen height of 50 % PAL or equivalent identified in Figure 3.[6]

Observe - If the OR requires the system to be capable of achieving the observation category then the target should be reproduced at no lower resolution than 16 mm per pixel and the target should represent a screen height of 25 % PAL or equivalent identified in Figure 3. The system should be tested

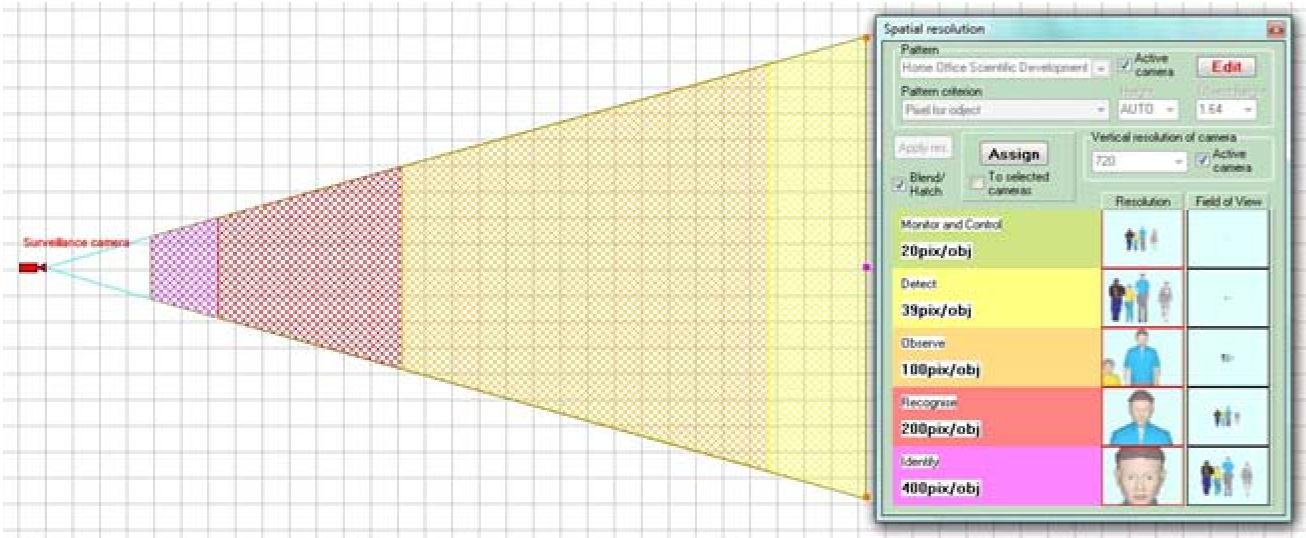

Fig.2: Fields of view

to ensure that some characteristic details of individuals can be identified, such as distinctive clothing. Individuals should be able to be clearly distinguished from each other, i.e. it should be possible to determine with a high level of confidence how many people are in the field of view. [6]

Detect - If the OR requires the system to be capable of producing images suitable for detecting the presence of an intruder then Annex E shall be used. The target should be displayed at a minimum of 10 % screen height as specified in Figure 3 and no lower resolution than 40 mm per pixel.[6]

Monitor - If the OR requires the system to be capable of producing images for crowd control or monitoring then the target should be reproduced at no lower resolution than 80 mm per pixel and no less than 5 % screen height as specified in Figure 3.[6]

Field of view and level of detail are not only parameters which should be taken into account within the VSS design. The Image Acquisition is directly depended on the level of dynamism, lighting and other environmental conditions of the scene exposed. The imaging technologies are based on the interpretation of the lighting reflected from the objects placed in different spatial levels of the scene. This generally known fact has stimulated the innovation of the Automatic Exposure Adjustment (AEA) tools. The Digital Signal Processing (DSP) units which are situated in the camera body have been used to provide necessary processing power to continuous utilization of AEA tools. Brief description of particular tool is provided in Figure 4.[7]

The question is how to evaluate the level of detail in the exposed scene influenced by spectra of environmental conditions mentioned above? Particular methods which have an ambition to partially answer this are proposed in section 3.

B. *Night vision*

Contemporaneous VSS are able to capture video sequence even in very adverse light conditions. Night vision of VSS could be realized by several methods:

- Standard additional illumination,
- Near Infrared Illumination (NIR),
- High-sensitive Image sensors.

Although the evaluation and testing of particular methods are tested by the VSS companies, there is still a lot of research space. Particularly level influence of environmental and adverse lighting conditions on the Video Content Analysis

Category	PAL	1080p	720p	WSVGA	SVGA	4CIF	VGA	2CIF	CIF	QCIF
Inspect	400	150	250	300	300	300	350	600	600	1200
Identify	100	40	60	70	70	70	85	150	150	300
Recognise	50	20	30	35	35	35	45	70	70	150
Observe	25	10	15	20	20	20	25	35	35	70
Detect	10	10	10	10	10	10	10	15	15	30
Monitor	5	5	5	5	5	5	5	10	10	15

Table 1: Level of detail specifications related to particular resolution

functions. For these purposes is also necessary to propose appropriate methods of measurement.

IV. Levels and Conditions

In the previous section were defined particular functional properties of IVSS, however questions which remain are related to possibilities of its effectiveness evaluation, in accordance to the measure of operational requirements fulfillment. Determination of particular image functional properties expects appropriately sophisticated solution. Logically the method “per partes” is applicable. Simplest way is to evaluate system from basics to the most complex parts. For this purposes it is possible to comply the system into three main tiers.

- **Tier I** – (Image Acquisition) – video sequence, Automatic Exposure Adjustment.
- **Tier II** - (Elementary Image Processing) – video sequence, Video Content Analysis.
- **Tier III** – (Advanced Image Processing) – video sequence, Video Content Analysis, Video Event Understanding.

Empirical approach seems to be ideal method for the particular level of the system establishment. In the other hand it is important to consider that the area of research is realized in practice for more than twenty years. According to this fact, it is obvious that utilization of best practices of designers could have significant contribution to the IVSS evaluation guideline creation process. Gathering of the video sequence could be marked as a key function of systems at all levels. Another important factor, which is necessary to consider is reality, that the level of detail is depended on parameters of the camera merely in ideal laboratory conditions. There are a lot of factors which influence final form of the video sequence in real environment. These parameters could be divided into three main categories, considering position of system designer:

- **Fully influenceable** (including intrinsic and extrinsic calibration) – parameters of camera and lenses (optical system), position and number of cameras, functionality and configuration of video content analysis and video event understanding tools.
- **Merelyinfluenceable** – parameters of the scene (environmental conditions, lighting conditions, architectural dispositions).
- **Uninfluenceable**– attributes related to object of observation (purely depended on the particular application).

Not all of parameters mentioned above are measureable in certificated way. Some of them are different within the particular application what makes them totally individual. In other words we could consider these parameters as a group of technical and technological preferences and conditions of the exposed locations. Firstly, it is necessary to establish evaluation methods for purposes of both, the influenceable parameters and conditions. In following section the detailed description of existing evaluating methods is provided.

V. Evaluation Methods

It is important to consider, that evaluation methods vary in accordance with level of the tested system. In case of Tier I, a primary functional property is Image Acquisition. Exact methods of measurement are contained in European standard EN 50 132 -7. Whole process is based on comparison of measured levels of detail against these specified in operation requirements establishment process. For measurement is used evaluation pattern, as could be seen on Fig. 2.

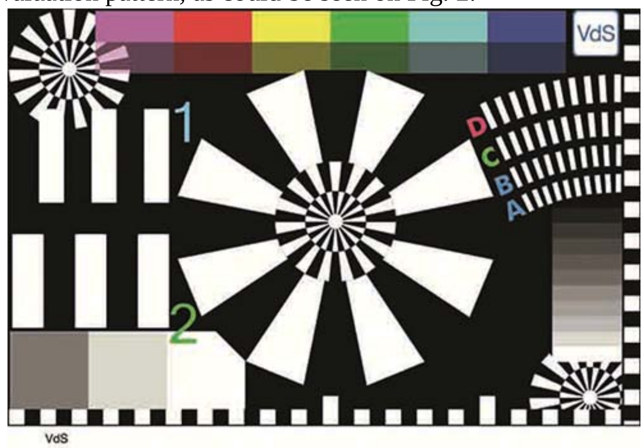

Fig. 3: Evaluation pattern for level of detail measurement

Measurement of the set optical characteristics is in laboratory conditions realized thru Modulation Transfer Function (MTF) [6]. Evaluation under real conditions utilize identical principal, but in subjective perspective. Both objective and subjective measurements are based on divergence of subject of observation contrast between subject and image planes. The measurement of contrast alternation relation on spatial frequency of black and white grid is provided in case of MTF. In real scenes is utilized Image Pattern illustrated on Fig. 2.

In the novel approach which was proposed, the action of particular effect within the standard level of detail measurement methods are considered. These effects are lighting and environmental conditions. The quantitative measurement of lighting is proposed by use of color histograms. The main difference is that the light reflected is measured instead of the light level falling on the scene. Nonetheless, the comparison between these two methods will be provided, to establish the more appropriate one. The whole method is consisted of three main parts.

- Level of detail evaluation
- Lighting level measurement
- Evaluation of Environmental conditions

A. Level of detail evaluation

As a method of level of detail evaluation was chosen ROTAKIN, the rotating target board. This method is very appropriate to measure dynamic scenes.

B. Lighting level measurement

As it was mentioned above, the original method to measure light conditions will be tested. The Rotakin test will be executed under the influence of different light conditions and the final interpretation of the scene will be evaluated within the range of particular levels of detail. The evaluation of effect of AEA tools will be measured and evaluated as well.

C. Environmental conditions effect evaluation

Raining, snowing and the strong wind have also significant effect on the final appearance of the image. These aspects should be evaluated as well, the scale of intensity of each of them have to be proposed in order to specify requirements, how to oppose these conditions. The test is based on measurement and evaluation of particular conditions effect on the capability to provide defined level of detail.

Tiers II and III, utilize completely different methods of evaluation. Firstly, it is necessary to specify functional properties of these tiers. Main function is to recognize specified activity in the scene, through video content analysis and video event understanding algorithms. The objective is testing effectiveness of these algorithms to provide support for the operators observing big amount of monitors. The most exploited and effective method for evaluating the algorithms is called mapping [7]. The mapping is based on compliance evaluation between real semantic information related to the scene and the information generated by particular algorithm. The packages of video sequences called datasets were developed for the purposes of video analytics algorithms evaluation. These evaluation annotations are a part of groundtruth, which represents complex testing tool. The rate of algorithm effectiveness is quantitatively expressed by appropriate metrics. The most used are following:

- Precision,
- Recall,
- F-measure.

Completely effective system matches high value in the metrics mentioned above.

Precision expresses amount of relevant generated data within all generated data. This ratio is illustrated via equation n.1.

$$P=\{RD\}/\{OD\} \qquad (1)$$

Recall is function which demonstrates relation between relevant generated data within all real relevant data. This dependency is visualized through equation n.2.

$$R=(\{RD\})/(\{ARD\}) \qquad (2)$$

Effectiveness of entire system is than illustrated by relation of these metrics, which forms precision-recall curves. The algorithms are than divided into two main categories. Precision or recall oriented. In order to get comprehensive result the F-measure is utilized. F-measure is formulated by following equation:

$$F=(\{P\}*\{R\})/(\{P\}+\{R\}) \qquad (3)$$

Nonetheless, there exists relatively wide scale of evaluation metrics. For example: Sequence Frame Detection Accuracy (SFDA), Multiple Object Detection Accuracy (MODA) a Multiple Object Detection Precision (MODP) [8].

VI. CONCLUSION

The area of the VSS Image Functional properties is quite complex. The novel approach how to evaluate these systems has been proposed, but it will be necessary to test it and produce some relevant data, before this complex evaluating approach will advance to next stage, the image processing process. This initial step of my research briefly recapitulated contemporary research on the field and establishes the evaluating methods. The future research will be predominately about testing and upgrading. These methods have an ambition to evaluate contemporary Intelligent VSS and their EVU tools.

REFERENCES

[1] Péter L. Venetianer, Hongli Deng, Performance evaluation of an intelligent video surveillance system – A case study, Computer Vision and Image Understanding, Volume 114, Issue 11, November 2010, Pages 1292-1302, ISSN 1077-3142.W.-K. Chen, *Linear Networks and Systems* (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[2] Aldridge JJ994, CCTV Operational Requirements Manual JSDB Publication 17/94, ISBN 1 85893 335 8.

[3] Ra□ty, T.D., "Survey on Contemporary Remote Surveillance Systems for Public Safety," Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, vol.40, no.5, pp.493, 515, Sept. 2010.E. H. Miller, "A note on reflector arrays (Periodical style—Accepted for publication)," *IEEE Trans. Antennas Propagat.*, to be published.

[4] Pavlidis, I.; Morellas, V.; Tsiamyrtzis, P.; Harp, S., "Urban surveillance systems: from the laboratory to the commercial world," Proceedings of the IEEE , vol.89, no.10, pp.1478,1497, Oct 2001

[5] Ashani, Z., "Architectural Considerations for Video Content Analysis in Urban Surveillance," Advanced Video and Signal Based Surveillance, 2009. AVSS '09. Sixth IEEE International Conference on, vol., no., pp.289, 289, 2-4 Sept. 2009 doi: 10.1109/AVSS.2009.112.

[6] EN 50 132-7. Alarm system - CCTV surveillance systems for use in security applications - Part 7: Application guidelines. B - 1000 Brussels: Management Centre: Avenue Marnix 17, 2011.W.-K. Chen, Linear Networks and Systems (Book style). Belmont, CA: Wadsworth, 1993, pp. 123–135.

[7] SKALICKÝ, Petr. Aplikacesignálovýchprocesorů. Vyd. 1. Praha: Vydavatelství ČVUT, 2005, 136 s. ISBN 8001026477.

[8] Aleksandra Karimaa (2011). Efficient Video Surveillance: Performance Evaluation in Distributed Video Surveillance Systems, Video Surveillance, Prof. Weiyao Lin (Ed.), ISBN: 978-953-307-4368, InTech,

[9] LUKÁŠ, L., HROMADA, M. Resilience as Main Part of Protection of Critical Infrastructure, In: INTERNATIONAL JOURNAL of MATHEMATICAL MODELS AND METHODS IN APPLIED SCIENCES, Issue 1, Volume 5, p. 1135-1142, 2011, ISSN: 1998-0140, dostupnétiež z WWW: http://www.naun.org/journals/m3as/J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility (Periodical style)," *IEEE Trans. Electron Devices*, vol. ED-11, pp. 34–39, Jan. 1959.

[10] HROMADA, M., LUKAS L., Multicriterial Evaluation of Critical Infrastructure Element Protection in Czech Republic,In. International Conferences, ASEA and DRBC 2012, Held in Conjunction with GST 2012, Jeju Island, Korea, November 28-December 2, 2012. Proceedings, pp 361-368, ISBN 978-3-642-35266-9

[11] HROMADA, M., LUKAS L., Conceptual Design of the Resilience Evaluation System of the Critical Infrastructure Elements and Networks in Selected Areas in Czech Republic, The twelfth annual IEEE Conference on Technologies for Homeland Security (HST '12), will be held 13-15 November 2012 in Greater Boston, Massachusetts. Pp. 353-358, ISBN 978-1-4673-2707-7

[12] NECESAL Lubos, HROMADA Martin, "Physical Protection Systems and Critical Infrastructure Protection in the Czech Republic" 1st International Conference on Automatic Control, Soft Computing and Human-Machine Interaction (ASME '12), Recent Researches in Automatic Control, Systems Science and Communications, Portugal, Porto, pp. 149-153, July 2012

[13] P. N. Belhumeur, J. P. Hespanha, and D. J.Kriegman, "Eigenfaces vs. Fisherfaces: RecognitionUsing Class Specific Linear Projection," IEEE

Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, 1997.
[14] Anagnostopoulos, C and et all, License Plate Recognition From Still Image and video Sequences: A Survey in proceedings of IEEE Transactions on Intelligent Transportation Systems. 2008.
[15] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, Moving target classification and tracking from real-time video, in Proc.IEEE Workshop Applications of Computer Vision, 1998, pp. 8–14.
[16] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, Moving target classification and tracking from real-time video, in Proc.IEEE Workshop Applications of Computer Vision, 1998, pp. 8–14.
[17] VALOUCH, Jan. Integrated Alarm Systems. In Computer Applications for Software Engineering, Disaster Recovery, and Business Continuity. Series: <http://www.springer.com/series/7899>Communications in Computer and Information Science, Vol. 340, 2012, XVIII.
[18] CAPUTO, Tony C. Digital video surveillance and security. Boston: Butterworth-Heinemann/Elsevier, c2010, 333 s. ISBN 18-561-7747-5.
[19] LUKÁŠ, L., HROMADA, M. Simulation and Modelling in Critical Infrastructure Protection, In: INTERNATIONAL JOURNAL of MATHEMATICS AND COMPUTERS IN SIMULATION, , Issue 1, Volume 5, p. 386-394, 2011, ISSN: 1998-0159, dostupnétiež z WWW: http://www.naun.org/journals/mcs/

J. Sevcikwas born in 1986. In 2011 completed a master's degree in Security technologies, systems and management at the Tomas Bata University in Zlin, where he attends postgraduate studies. The object of her interest is in the Effective design of the Video Surveillance System and the Security assessment of the environmental conditions influencing the object.

With support of the European Regional Development Fund under the project CEBIA-Tech No.CZ.1.05/2.1.00/03.0089 and by Grant No. /FAI/2012/2110 from the IGA (Internal Grant Agency) of Tomas Bata University in Zlín.