

Bikas S. Aghayev, Tarlan S. Aliyev

DOI: 10.25045/jpit.v07.i2.05

Institute of Information Technology of ANAS, Baku, Azerbaijan
depart6@iit.ab.az

PROBLEMS OF THE APPLICATION OF VIDEO ANALYTICS IN SURVEILLANCE SYSTEMS

The article analyzes the video surveillance systems (VSS) and their video analytics tools. Current state and prospects of development of the video analytics are studied. The article puts forward recommendations for selecting the best option of analytics tools and improvement of video surveillance systems.

Keywords: *information systems, audio and video information, Big Data, security systems, video surveillance systems, video analytics.*

Introduction

People realize their active labor activities by using techniques and media, which are created thanks to the gained knowledge and experiences. The development of these tools leads to the transition to new development stages of the society, and to drastic changes of material and spiritual values. Socio-cultural, political and economic evolution of the society occurs due to the rapid growth in the acquisition and the use of information [1]. Today, the information society is emerging as a form of satisfying the urgent needs of the people for access to information and its use.

The development of effective coding methods, and completion of digitization process of all the descriptive forms of information, a sharp increase in the productivity of processing facilities and its transformation from the corporate tools into mass processing ones, widespread computerization and networking, and so forth has caused a sharp increase in information production capacity and a new - VI information revolution.

Web technologies, social networks, mobile phones, new generation digital sensors, new audio-video information systems and so on have increased the volume of generated information.

One of the largest information systems, generating enormous information, is the video surveillance systems. In the 90s of the last century, simple systems with point or local architecture, set up by one or several cameras, were used, the information capacity of which was measured in megabytes, whereas, at the present, development stage of the information and telecommunication technologies, distributed video surveillance systems with hundreds and thousands of cameras, are generating terabytes (TB) of information a day. Today, even an operator cannot effectively track the data of the system consisting of a few cameras on the monitor of the work place. Therefore, the need for the automation of control functions to improve the efficiency of these systems has led to the emergence of video analytics, i.e. video content mining.

Currently, VSS is widely used in any field of activity, including in the housing sector. Therefore, selecting the best option of the system, as well as the video analytics software modules, which need to be integrated, is a quite difficult and important task for the organizations or individuals in terms of the destination parameters, functionality and economic feasibility. Currently, if we take into account the availability of the VMS solutions with different architectures and the very high prices of video analytics tools, organization managers need to have proficiency in the same field for applying these systems.

This paper analyzes the VSS systems and video analytics tools applied in these systems, and presents some recommendations for the optimal solutions.

Audio-video information systems as an object of Big Data technologies

Nowadays, the term Big Data means the architecture systems and technologies that the data processing of which is characterized by a set of features marked as “5V”. Here, the “5V” means the quality indicators associated with volume, velocity, variety, veracity and value [2, 3]. In other words, Big Data is a technology allowing for *true and valuable* (significant) data and knowledge

discovery from *various sources* (online streams or offline databases) by extremely large scale data processing acquired with high-speed that can be interpreted for certain decision making. All of these particularly refer to the online processing of unstructured data streams (streams of audio-video traffic, social networks, search engines, e-mail information, web texts, technological sensors, etc.) that do not meet structuring format of modern databases [4].

Completion of the digitization process of all descriptive forms (types) of information, and the development of efficient coding methods can be the grounds, which have led to the generation of the large amounts of data.

As a result of the use of new techniques and technologies developed during the last few decades, the volume of produced information now has reached Exabyte (EB), and according to some reports, this figure will be more than 4.5 EB [5], by the beginning of the next decade (Figure 1).

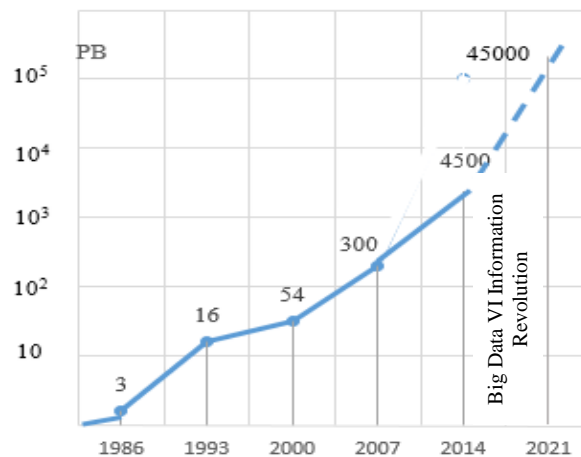


Figure 1. Data Growth

One of the resources that require large memory and computing resources is the audio-video information systems. The digitization emerges as a result of the coding techniques for the recording and the revitalization (accordingly, listening and being monitored) of this type of large-scale information. For example, 8 bits of information is enough to digitize the symbol “a” with the extended ASCII standard, while 460 Kbit (24 x 192 Kbit/sec. x 01sec.) information unit is required to code the sound “a” (of the Azerbaijani alphabet), an average duration of which is 100 m.sec., in DVD-Audio format, by discreting with 24-degree quanting and 192 KHz frequency. The required information size is 1020 megabits for replicating the description of a person uttering the sound “a” on the reflection devices (monitor, TV, screen, etc.) for 100 m.sec., with high-quality digital video standard HDMI-1.3 (2560x1440 pixels, 50 frames/sec.) (Figure 2).


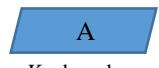
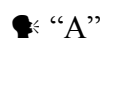

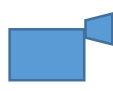
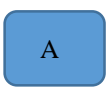
Description form	Generation technique	Code standard	volume
Text	 “A” 10111011  Keyboard	ASCII	8 Bit
Voice	 “A” 	DVD-audio	465 KBit
Video	 	HDMI-1.3	1020 MBit

Figure 2. Comparison of digital volume units of description forms of information

In practice, video recording or program shooting require the processing of very large-scale information traffic. For example, IP-AVTECH brand 16 “day/night” type video cameras are installed at the digital video surveillance system of the Institute of Information Technology of ANAS, and the institute’s territory is daily monitored. If each camera’s bit rate is averagely 2 Mbit/sec., system can record 43.8 TB of information per month of (in real terms, 3 x 3 TB HDD hard drive is sufficient for 20 days of recording) [6]. More than 1,500 video cameras of the “Intelligent traffic management system”, operating in Baku, generates more than 260 TB of information [7]. Obviously, continuous online monitoring of so much information by an operator on the monitor is very exhausting.

Video analytics as audio-video processing technologies of information systems

As mentioned above, one of the resources generating large amounts of information, which accordingly requires large memory and computing resources, is the audio-video information systems. This especially refers to video surveillance systems of video security systems.

The experience shows that if an operator tracks even only one monitor continuously and carefully, he/she will be tired in 20-30 minutes and does not “see” 45% of the information. In an hour, this “loss” increases to 80-90%, and the probability of neglecting the most important events on the screen sharply increases [8, 9]. If an operator continuously controls several monitors or a monitor divided into four segments, the effectiveness of the work reduces further. On the other hand, seeking needed video fragments from a large-scale video base (in an offline mode) is a rather time-consuming and tedious task. Therefore, automated large-scale audio-visual information processing is important, in terms of accuracy, and aims to improve the efficiency of the control functions of an operator. The need for control systems automation has caused the emergence of the video analytics - video content mining.

The video analytics, which emerged in the 90s of the past century, is described as follows: video analytics is a technology that enables to automate the information and knowledge discovery by analyzing video images received through TV cameras or video archives in online or offline mode through the computer vision methods (the term "machine vision" was used then).

Technically, the video analytics is a software and hardware that provides video content processing. The video analytics resolves the following issues [10, 11]:

- object detection;
- object recognition;
- object classification;
- object tracking.

One or more of these functions of video analytics can be used depending on the degree of complexity of security systems.

Currently, the video analytics is mainly used in video information safety systems for security and monitoring purposes, which is a class of information systems (figure 3).

As it is noted above, the security service employees (operators) get tired, while continuously tracking the images are received through the TV cameras, lose concentration, and ultimately neglect important events. This is especially referred to video information systems with multiple cameras. Additionally, if the large areas are protected by the workers on a particular route through observing on foot or car, it is possible to avoid dangerous situations. Increasing the number of security officers can significantly improve their performance. However, this requires additional financial costs. The only way out of the situation is the application of the video analytics tools in video information systems. Video analytics operator gets rid of the obligation and responsibility of continuous observation of the information flow on the monitors received from TV cameras. In this case, the video analytics tools automatically detects important and dangerous events from the very moment and send the operator a warning signal for observation. The operator either observes the further steps, or make necessary decisions by analyzing the archived video fragments. One of the advantages of these systems is its opportunity to prevent the potential danger or damage

through detecting the event before it happens or at an early stage, and warning security guards about them. The introduction of video analytics improves the work efficiency of security systems by reducing the errors occurred during the observation process due to "human factor".

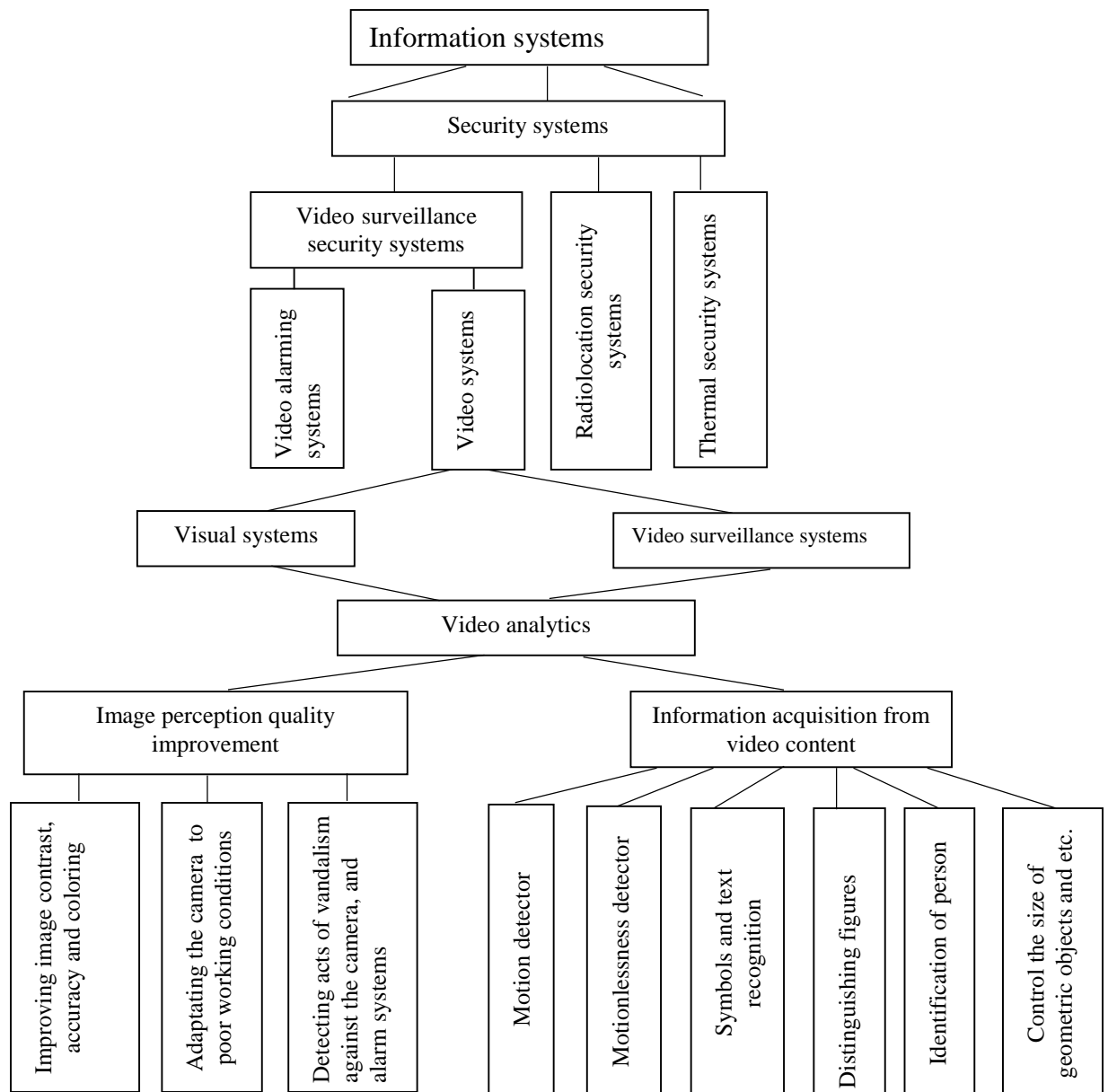


Figure 3. Classification of surveillance security systems

The video information systems, in which video analytics tools are applied, and usually divided into two great classes [12]:

- Visual Systems. In these systems, human vision apparatus is an information receiver;
- Video systems. The information generated by the video cameras of these systems, is sent to automated processing hub and processed.

The video analytics applied in the first-class systems aims at preparing the images received through the TV cameras for more qualitative perception by human vision apparatus. Video stream or basic data processing, decision-making is performed by human beings.

In the second case, solutions of video information system (detecting unauthorized access to an area, recognition of numbers or human faces, etc.) are performed by means of video analytics. The

operator checks whether the warning signals are right and wrong, and makes adequate decisions.

Obviously, the main purpose of the application of video analytics in video information security systems is to process the events registered within the scope of TV cameras in the framework of the functionality of the used hardware and software, and to warn decision makers about the obtained processing results. The technical scheme and architecture of the video surveillance system, in which video analytics is applied, generally consists of the TV cameras, computers receiving video images (high-performance server), communication channels connecting these devices, data mining software adequate to the system objectives, additional devices used to increase the effectiveness of the system's performance (such as phone networks, internet access facilities, tools to warn law enforcement bodies, tools to activate the implementation mechanisms of measures against detected unpermitted events, and etc.).

Three groups of the video information systems are distinguished by their architecture solutions: Server Base analytics (SBA), Edge Videoanalysis (EV) and Hybrid base analytics (HBA) [13].

SBA systems process (coding, decoding, recording, controlling) the image received from the cameras on the server computer, installed next to the operator. Its advantages include the use of simple (cheap) TV cameras and the simplicity of the system. Its disadvantage is the need for the use of private computers, i.e. servers due to the insufficient personal computer resources (processing speed, volume of RAM) to process the video stream when the number of cameras is high and the capacity of the channels (transmission speed) is insufficient to transmit the video stream received from each camera with high-quality.

EV systems perform the pre-processing and some simple video analytics software on the microprocessors installed in the cameras (including encoding-decoding operations). The preliminary processing results and the set of search attributes are transmitted to a computer. Its advantages include: the channel load is reduced, and less powerful computers can be used. Increasing price of cameras and system, tightening requirements to operating conditions of street cameras are the disadvantages of these systems.

HBA is a combination of the previous two options. The video information processing in these systems is distributed between the camera and the central computer. For example, the video stream coding (digitization, compression) and motion detection are performed on the cameras, and decoding, monitoring of video analytics results, recording to the database, and camera management and other software are installed on the server. The hybrid system eliminates a number of imperfections of the previous versions.

Regardless of applied video analytics, one of the complex technical solution issues of any technical security systems is the organization of operative search of necessary fragments out of archived unstructured video data, namely the minimization of a search time. Since the text and audio data archiving process is carried out periodically in a chronological order as the interrupted sessions, the parameters as date (year, month, day) and time (hours, minutes, seconds) are sufficient for the search of each session data (files). In video analytics, as the information systems is uninterruptedly received from the cameras in the form of stream, the search for the necessary fragment out of thousands of archived data by its date and time, especially when the exact values of these parameters are unknown, is a very time-consuming and tedious task. For example, in Server Base analytics, though the data recorded with the HD quality indicators (high resolution systems: reading and recording matrix sensor of the system cameras generates the image with the density of less than 1 MPixels [14]) is compressed on a central server up to frequency 8 shot./seconds. However, nearly 30 000 fragments of the needed video may be reviewed (the exact timing of the incident is known exactly one hour). Currently, as a solution to minimize the search time, two-class intelligent search algorithms are used [15]:

- Detection, selection, and processing of only the events, which reflect the fragments meeting the objectives of the system, using video analytics methods, as well as their transmission

to the central computer. The server analyzes the incoming flow or archives it for later use. If the required quality indicators are not critical, the volume information can be reduced through the use of compression;

- Recording all the fragments of a shot or the scenes of detected events together with the text metadata (date, time, event type, color, motion direction, dimensions, and other sets of features) that characterize them. Depending on the degree of adequacy of the associated metadata with the description of the events, text format can reduce the search time by thousands times.

The video analytics has been emerging during the recent few years, at the first stage of which the following issues are explored [16]:

- Input flow filtration in order to reduce the size of the video to be archived, i.e., analysis or archiving of important events in accordance with the monitoring purposes;
- Identification of the set of operative search criteria, and the development of methods such as structuring archive data according to different criteria;
- Identification of characteristic templates (standards) of the set of observed events, and development of verification methods.

Currently, two-class analytic algorithms, based on different principles that are considered relatively simple, are used to resolve abovementioned video analytics issues:

- Comparison of image pixels. This class of algorithms are based on the comparison of the consecutive cadres' images of the same format with one another (cases);
- Comparison with templates. The current image is compared to the database samples (templates) that describe characteristic actions (static condition or condition series).

Video streams or database data may be the comparison object.

However, the development trend promises the development of more sophisticated and precise (high verification grade) class of algorithms in terms of functionality in coming years. For example, UNESCO International Council for Science (ICSU) [17], "Big Data Europe" project of the European Union "Horizon 2020" innovation program, and the preliminary results of other organizations and programs achieved on the data processing also forecast this. The issues, solved with practical hardware and software based on these algorithms, can be divided into 4 groups according to their destination indicators and functions [18]:

- Perimeter video analytics issues, based on relatively simple analysis algorithms. It is mainly carried out by the motion detectors that detect the fact transition of the static situation into dynamic one in the surveillance zone limited to certain perimeter, i.e., transition from motionlessness into motion (or vice versa). It is used to detect unauthorized access to a protected area, the change occurred in the observation area, the absence of something, for example, the theft and so on. Sophisticated detectors are capable to distinguish false warning signals (error), for example, to identify whether the violator is human or animal, or the movement occurred due to the trees (e.g., wind);
- Recognition video analytics. Recognition (registration) of human faces, numbers, and objects. For example, wanted criminals, missing persons, stolen vehicles;
- Service video analytics. The need for camera controlling and maintenance. For example, acts of vandalism against the camera (blocking the lens, hacking), rotating the camera toward all the four sides to observe moving objects throughout its scope, controlling focal distance of the lens, etc.;
- Behavioral video analytics. This type of analysis is carried out based on more sophisticated algorithms in the early development stages. At present, some simple and practical solutions have been developed. For example, some softwares are available to detect a large number of people gathered in the observation area, the presence (fall) of something on rails in the subway stations, the use of someone else's identification card in the final checkpoint (turnstile).

As the price of the modern data mining software is high enough, simple (cheap) video analytics modules (functions) are usually used in small and medium-scale video systems, or the observation and analysis functions are carried out autonomously by operators. For example, above-mentioned intelligent traffic management system fulfills a number of simple analytical analysis functions: identifying the speed of the vehicle, its state license, detection of movement at the red lights, as well as crossing prohibited line and so on. Whereas, detection of seat belt use, driver's car ownership or the right to control and so on is impossible thus far.

Video surveillance system of the Institute of Information Technology

Video surveillance system of the Institute of Information Technology of ANAS is a centrally controlled video system with a video class (figure 4). This system does not use video analytics tools for the data processing received from the IP-network cameras of different constructions without analysis functions, i.e., it is a centrally controlled system. The operator controls 16-quadrant monitor. There are no any video data compression or coding functions. No motion detector and personnel filtration is used.

Working hours are daily, continuous, three-shift and controlled by an operator. Incoming data stream is archived in 16-channel recording device, and transmitted directly to the monitor. Warning signals about a dangerous situation are not generated for the operators. The search is only carried out by date and time attributes. The lack of video data compression, but other high-parameters (bitrate - 2 Mbit/sec., frequency - 25 shot/sec., codec - H.264) of the system generate HD quality images (resolution of 1.2MPixel).

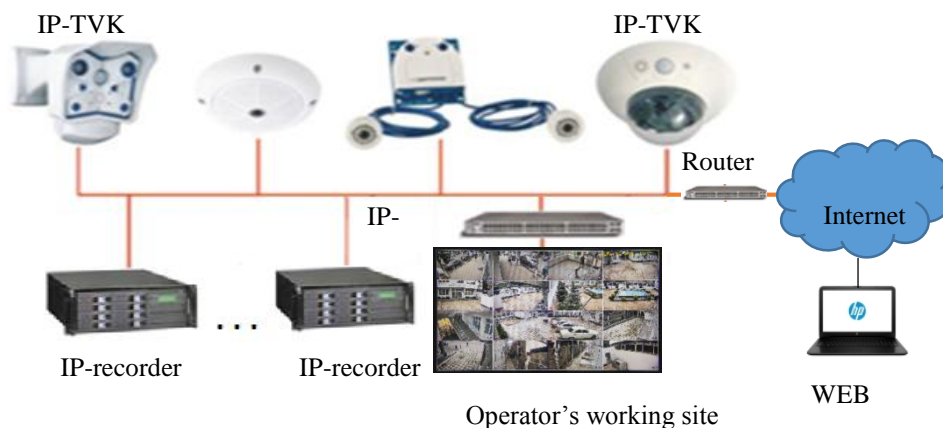


Figure 4. Summarized architectural scheme of the video-surveillance system of IIT of ANAS

Obviously, such a simple system, as noted above, is not able to perform control functions effectively (especially during non-working hours). Therefore, in order to improve the operating efficiency of the system, the following improvements are recommended to be fulfilled.

- modules of video compressions and motion detectors to be applied to the incoming flow to reduce the volume of memory and data;
- video analytics applications to be used to improve the control quality;
- checkpoint system and video surveillance system to be integrated (for example, to detect whether the activated checkpoint identification card belongs to the user);
- intelligent search application modules to be introduced to accelerate the search of video archive fragments, remaining static date and time search attributes.

Conclusion

The article illustrates that, during the last 25-30 years, the development of effective coding methods and technologies, the completion of the process of digitization of all descriptive types of data, the shift of processing facilities from corporative to massive ones, along with the rise in the productivity of computing techniques, the expansion of computerization and networking, and other factors have sharply increased the data generation.

It is noted that, along with the equipment and technologies developed in the last few decades, such as web technologies, social networks, mobile phones, a new generation of digital sensors and so on, audio-video information systems, as well as video surveillance systems also generate a large amounts of information. As the volume of incoming video data received from the numerous TV cameras of video surveillance systems is very large, an operator gets tired while continuously tracking these images during a certain period, he/she loses concentration, and is likely to omit the most important events. It is explained that the video analytics - intelligent processing of video content emerged due to the need for the automation of control functions of these systems, and overall improvement of the efficiency of the system.

The article classifies the video surveillance system, included in the class of video security systems, and identifies main objective and scope of the video analytics. The authors conditionally divide these issues into two groups: resolved issues and the issues to be resolved in the near future. The problem of minimizing the time of video fragment search in the database is investigated, for this purpose, two groups of used intelligent search algorithms are compared.

The article provided the brief information about the “Intelligent traffic management system” of Baku city. In addition, it explained the architectural solution of video-surveillance system of the Institute of Information Technology of ANAS and offered some recommendations to improve the efficiency of the system through its intellectualization.

The organization managers and others may benefit from this article to get brief information for the selection of the system and optimal video analytics application modules to be integrated to the system, in terms of its performance, functionality and economic feasibility.

References

1. Moiseev N.N. Information Society: Opportunities and Reality // Information Society, M.: ACT, 2003, pp. 248–251.
2. Leney. D. 3D data management: Controlling data volume, velocity and variety. Technical report META Group, 2001. <http://blogs.Gartner.com>
3. Baaziz A., Quoniam L. How to use Big Data technologies to optimize operations in Upstream Petroleum Industry / International Journal of Innovation, 2013, vol.1, no.1, pp.19–29.
4. Big Data. <http://www.tadviser.ru/index.php> / Article: Video surveillance (world market)#2012_.D0.B3.D0.BE.D0.B4
5. Imamverdiyev Y.N. great potential and challenges of Big Data technologies // Information Society Problems, 2016, No1, pp.23–34.
6. Agayev B.S. Problems of big video information processing resources / proceedings of the republican scientific-practical conference “Big data: opportunities, multidisciplinary problems and perspectives”, Baku, Information Technology, 2016, pp.49-53.
7. “Intelligent traffic management system” of Baku city. www.niim.az/aboutBakuITS.do
8. Orlov S. Video analytics practice // “Journal of Network Solutions/LAN”, 2015, No03. www.osp.ru/lan/2015/03/13045265/
9. Kucherenko O. Video Analytics algorithms and detectors // Protection Technologies, 2013, No5. <http://www.tzmagazine.ru/jpage.php?uid1=1000&uid2=1117&uid3=1138>
10. Lyubimov M. Applying analytics to integrated security systems // Protection Technologies, 2014, No1. <http://www.tzmagazine.ru/jpage.php?uid1=1181&uid2=1182&uid3=1200>

11. Orlov S. Video analytics: problems and solutions // “Journal of Network Solutions/LAN”, 2014, No 6. www.osp.ru/lan/2014/06/13041879/
12. Collins R.T, Lipton A.J., Kanade T., and etc. A System for Video Surveillance and Monitoring. Robotics Institute, CMU-RI-TR-00-12, VSAM Final Report. http://www.ri.cmu.edu/pub_files/pub2/collins_robert_2000_1/collins_robert_2000_1.pdf
13. Ptitsyn N.V. Video Motion Detectors: driving force of the market of cameras // Security Systems 2009, No 6. www.secuteck.ru/articles2/videonabl/videodetektory-dvijeniya
14. Kucherenko O. IP and HD-SDI video monitoring: Myths and Reality//Protection Technologies, 2014, No1. http://tzmagazine.com.ua/upload/tmp/1_2014i.pdf
15. Favorskaya M.N., Pakhirka A.I., Shilov A.S., Damov M.V. Methods of Search for motion in video sequences // Bulletin of the Siberian State Aerospace University named after Academician Reshetnev M.F. <http://cyberleninka.ru/article/n/metody-poiska-dvizheniya-v-videoposledovatelnostyah>
16. Peng C. Video analytics: Content analysis in security and surveillance applications. EETimes, 2007. www.eetimes.com/document.asp?doc_id=1273765
17. International Council for Scient. www.icsu.org
18. Gagvani N. Introduction to video analytics. EETimes, 2008. http://www.eetimes.com/document.asp?doc_id=1273834