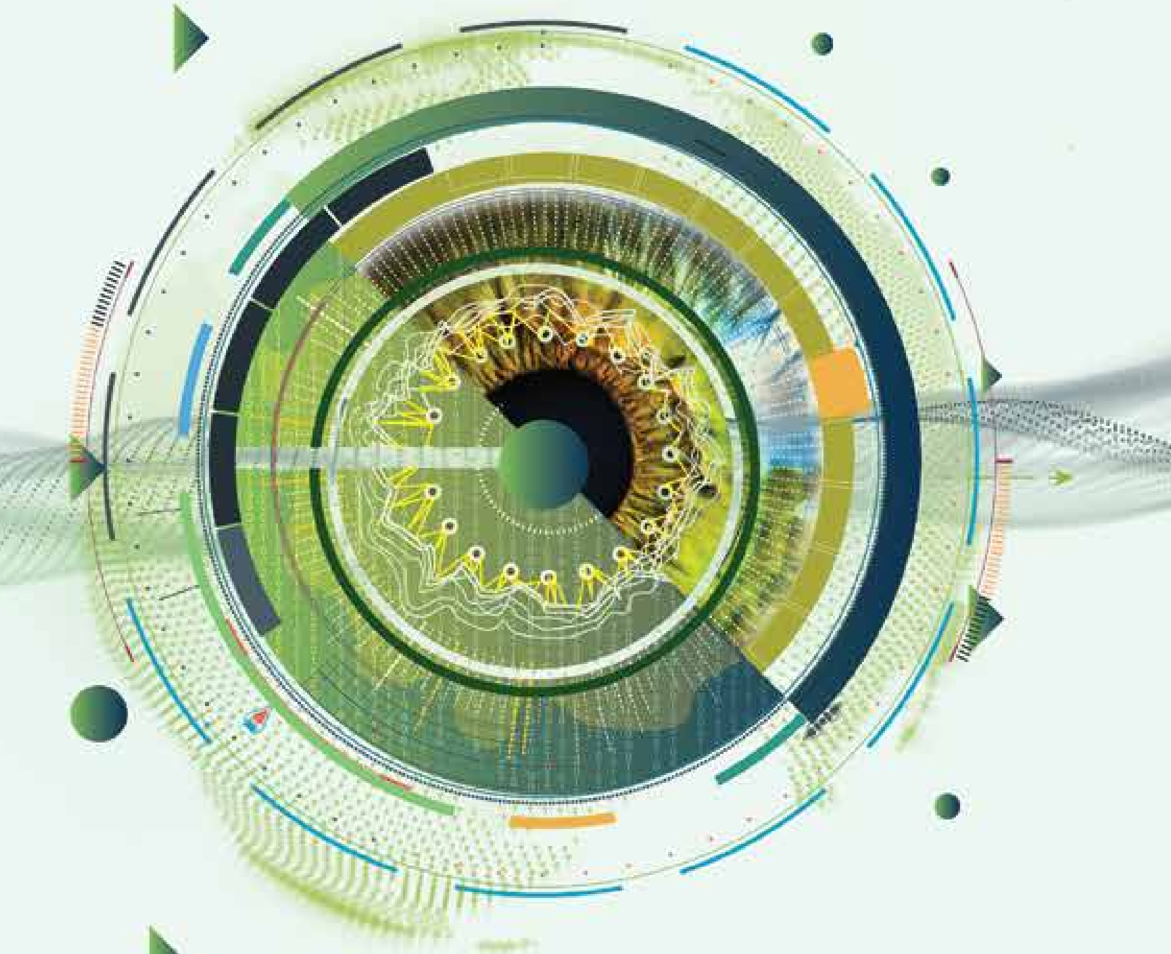
Takhmasib Dadashev



MODERN VIDEO SECURITY PLATFORMS IN THE BANKING INDUSTRY

ISBN 978-609-95007-1-3

Content

www.5ci.net

Penki kontinentai GRUPĖ

PENKI KONTINENTAI

Group of companies

We connect the continents

Penki Kontinentai Company Group is one of the leading corporations involved in banking technology, innovative payment solutions, broadband Internet, IPTV, telephony and the Internet of Things. Solutions and services by the company group are applied in 80 countries worldwide.

Foreword

Modern technologies introduced into all aspects of life have become our current reality: innovations are an essential part of our daily lives. They are guiding, monitoring and assisting people through every day's life what can be called the human experience. Availability of various mobile devices and self-service systems have advanced to such a point that business is able to gain access to virtually any information about the client. What is peculiar, however, is that the implementation of the «Know your customer» principle did not satisfy the needs of the service provider, but rather underlined the interest to obtain details that are even more personal.

The following book is considered a brief manual on surveillance, video security and data protection in banking operations. In addition, the author tries with this book to predict the tendencies that could define the development of banking technologies of the several years to follow.

The company Penki Kontinentai (Lithuania) took very active part in preparation of this book. For 25 years, the company successfully operating in the field of telecommunications, software technologies for the financial and banking sectors, in particular, for the protection of self-service devices. The subsidiary of the Penki kontinentai group, the company BS/2, specializes in outsourcing services for banks and private retailers, as well as in developing software and supplying banking equipment. BS/2 is the exclusive partner of the international concern Diebold Nixdorf - the largest banking equipment manufacturer and supplier of banking solutions worldwide.

The solutions and services provided by BS/2 gained high reputation on the global market, as the company's products present itself in 80 countries. The software include the *«.iQ»* product family - a profound complex of solutions ensuring banking self-service devices security, enhancing productivity and lowering maintenance expenses.

As such, *ATMeye.iQ* software solution designed for ATM and other self-service devices video surveillance is able to create detailed reports with photos and video on individual user's actions and produce real time alarm notifications to the security personnel, thereby implementing proactive security principles of the most sensitive component of the banking infrastructure. This software product insures reliable video surveillance of the self-service device, which leads to lowering the risks of potential fraud and vandalism cases.

What makes this book different from the fundamental work «The Bible of video surveillance» by Vlado Damyanovski [1], which mainly focuses on specialized hardware, is that this book brings more attention to information and video security software.

70 years have passed since the British writer George Orwell first described video surveillance technologies being implemented in a society based on complete control in his dystopian novel «1984» [5]. Our reader may appreciate his futuristic vision in the epigraphs preceding each chapter of the following book.

Introduction. Basic concepts, video surveillance and video security technologies

> *«The telescreen received and transmitted simultaneously… There was of course no way of knowing whether you were being watched at any moment given…»*
>
> *G. Orwell «1984»*

The goal of *video surveillance* implies visual control of a designated area using one or multiple video cameras, allowed for storage and revision of digital video data, as well as constant evaluation of the territory under surveillance in order to detect the so-called «*disturbing events*» based on changes in the environment.

Historically the two main functions of video surveillance systems have been the provision of information to the control station and further archiving of the content. Most closed-circuit television (CCTV) platform manufacturers follow this particular model.

Currently the main tendency of surveillance system development is shifting from the analog video capturing (analog cameras), displaying (television) and storage (video cassettes) to the digital platform (IP cameras, computer monitors and digital databases). Therefore, a great role in modern video surveillance system development and efficient operation play *digital video compression, storage, retrieving and transferring technologies.*

One of the part of the essential assets to current security complexes are automatic and semi-automatic surveillance systems. The *client-to-server architecture* is the base for modern distributed surveillance systems, and, as a rule, all data processing takes place on the side of the server.

Another approach is to transfer aspects of data processing to the client side: in this case, the client has access to all cameras. An example could be the digitization of video content in analog video cameras systems. It is possible to transit the data into the digital format on the server side. By doing so, all cameras should be linked to the central post via coaxial cables, or, as an alternative, an encoder may be set up to digitize information from several cameras, following the data transfer in digital format via IP.

One of the primary traditional video surveillance system drawbacks is significant reduction of rapid response possibilities in growing networks. In cases of several dozen (or even larger numbers) of inbound video streams the operator is no longer capable to objectively monitor the current situation in real time.

Research shows that after 12 minutes of work an operator exposed to a large number of video channels is no longer capable of tracking up to 45% if monitor activity. In the course of 22 minutes, this number grows to 95%.

This issue becomes particularly relevant when operating exceptionally large surveillance systems, say, if the given objective is monitoring specific objects for an entire city. For example, systematic introduction of total video surveillance in London (over 10000 cameras in a single network and more than half a million throughout the city) did not lead to a significant reduction of the incidents number or the percentage increase of crimes being solved.

Consistently grow in demand video surveillance systems that provide solutions based on a multi-level logic where most processes do not require constant active participation of a human operator. Intelligent surveillance systems as such allow not only ensuring safety, but can also become useful in solving business challenges, collecting statistics about objects being monitored etc.

Collecting precise data about the footfall of the monitored object, visitor distribution in reference to time and even the possibility to identify specific customer classes (regular customers or other groups of interests) are functions that are high in demand for many commercial organizations. In the scope of an entire city they may be implemented to monitor traffic congestion or for a variety of other fields of interest.

A modern surveillance system includes a variety of visual computing technologies, such as video surveillance technologies, optical character recognition (OCR) and *biometric technologies*.

One of the key functions of video surveillance is the identification and classification of humans. However, in order to identify a person's face in a video recording, you need to possess the general features of the particular person. That information would form the basis to confirm that the concerned person was the identified one.

On the other hand, person's regimentation without identification requires slightly less data, i.e. this procedure is less accurate from the point of video surveillance. Classification implies visually determining the person's gender, color of clothes or other parameters. Thus, identification requires a more advanced degree of video resolution in comparison to classification.

A complex intellectual video surveillance framework should include the following primary elements and program/algorithm modules:

- visual (TV, infrared (IR) and other) sensors for remote video surveillance;
- resources of distributed collection, compression, processing and transferring of digital video information through local and global networks in real time;
- automatic focus on objects of observation (people, buildings, vehicles etc.)
- automatic tracing of moving objects in the monitored area;
- biometric recognition of personnel, biometric access control to critical zones of the object of observation;
- automatic identification of vehicles, shipments or other objects with the use of distinct markings (registration numbers, bar codes, other forms of labelling);
- methods to evaluate behavioral patterns of observed objects or object groups;
- forming alarm messages to the operator to report malevolent or unusual scenarios in the surveillance zone;
- firmware necessary to implement methods and algorithms of video information collecting and processing.

The transition to digital processing, transferring, and storage of video information has become one of the essential trends of the video surveillance industry. In other words, modern video surveillance has become almost exclusively digital. Today great part of all new systems implement digital and IP technologies for coding, transferring, viewing and recording the video signal.

In fact, video surveillance has undergone such a noticeable change in the recent years that the term «*video surveillance*» could easily be substituted by the term «*IP surveillance*».

At the same time, as we notice a growth in productivity and speed of memory chips, processors and hard drives, the price of these devices gradually decreases. As a result, digital video data processing has virtually become the only rational method of processing abundant amount of good-quality video data. After all, the efficiency of data processing determines the value of a system, not the amount of input data itself.

In 2010 global share of IP cameras was roughly 20% of the entire camera market (the approximate value of which amounted to $8Bn). Nonetheless, the trend of shifting to IP cameras has been set and the technology is beginning to dominate the field.

The approach differences in constructing client-to-server video processing systems becomes noticeable in regards to modules of *video analysis*, which aim to automate the functions of a surveillance system operator. Essentially, both layouts of video analysis modules, either on the side of the server or on the side of the client, are possible. However, as corresponding machine vision technologies are developing, the data processing on the side of the client keeps gaining more user support.

Let us bring the attention to the substantial benefit of digital video, which is the possibility to check copies authenticity. We are referring to the so-called *watermarks*. This allows protecting the encoded digital information from forging, which is rather important in the video surveillance industry.

According to the predictions made by the research company *Research and Markets*, total sales revenue in the global video surveillance market is expected to grow to $75Bn by the year 2022.

It is anticipating that during the years 2017-2022, the global video surveillance market will grow in revenue by 15% annually. One of the main stimuli for such growth should be the gradual transition from analog to IP cameras. Another important factor for video surveillance industry development is the influence of «cloud technologies» which allows data storage and access in a revolutionary way. Another great influence is the immense growth's celerity in demand for video surveillance services as a means of safeguarding.

The number of objects that are being equipped with video surveillance systems is steadily growing all over the world. The Video Surveillance as a Service (VSaaS) segment should demonstrate the swiftest growth in the upcoming years. More and more applications based on this method arrive on the market, as users are able to appreciate the benefits of remote access to video system resources stored on «cloud» services. Another prevalent feature is the possibility to access live and recorded imagery not only via desktop computers but also via mobile devices. Having mobile access to surveillance data facilitates the process of monitoring rendering it more flexible.

The growing demand of video surveillance services in banking can be easily explained by the growing number of ATM attack cases. In fact, in 2017 we see the *ATM Malware-as-a-service* emerge on the black market, an offering that provides a ready to use hacking solutions (including software and hardware) for self-service devices.

The availability of all necessary malware and hacking manuals in the Darknet manifests in increasing number of hacking attempts all over the globe.

Specially designed video surveillance systems protect the property, as well as cash and personal data of bank customers. As an implementation example of such solutions, could be the major North American service provider for financial institutions and retail networks planning to equip ATM fleet of 9000 machines with *ATMeye.iQ*.

The development of the video surveillance market and the increasing number of manufacturers determine a demand in open standards that would ensure the compatibility of equipment and software produced by different manufacturers.

Two industry associations have already been formed on this premise: the *Open Network Video Interface Forum* (ONVIF) and the *Open Network Video Interface Forum* (ONVIF). Both organizations were established in 2008 with the goal being to standardize the physical security and software platform interfaces alongside compliant IP-based security systems that protect property, funds and sensitive information.

Payments.iQ

Automated payment processing system

Payments .iQ
.iQ Family Product

Payments.iQ is a software solution for arranging payments (utilities, taxes, fines), selling any types of electronic services (tickets, vouchers, etc.), automating retail banking and managing networks of self-service devices such as information or payment terminals and ATMs.

Chapter 1. Main video surveillance technologies

«They could spy upon you night and day, but if you kept your head you could still outwit them...»

G. Orwell «1984»

Modern surveillance systems allow total visual control of indoor and vicinity areas. Depending on the ways of receiving and processing information, systems can be divided in two types: *analog* and *digital*.

Analog surveillance systems, as a rule, are used in small indoor environments. Video signal transmission is carried out through a combined coaxial cable, and connecting to the *video recorder system* is done through a BNC connector.

Despite analog video cameras still being used in modern surveillance systems, the number of manufacturers who propose digital cameras for video streaming through computer networks is on the rise. Earlier, video memory devices, quad processors, video multiplexers, DVP camera circuits constituted a small part of components for digital video surveillance.

Let us note that in addition to analog CCTV cameras most existing video surveillance systems incorporate digital video recorders for monitoring and long-term storage. Besides displaying the current situation and archiving of collected video information modern video security solutions may fulfill a number of additional functions such as management and monitoring using portable devices mobile notifications, video capture and car plate number identification, forgotten items detector as well as identification of a person using face recognition.

The main qualitative parameter of a surveillance system is its technical configuration.

Current professional video surveillance systems are comprised of several technical components (cameras, lenses, transmitters, activity detectors, infrared detectors, matrix switchers, thermal imaging camera, means of video storage (video recorders/ video servers) computer systems, monitors, etc.).

Video camera resolution has always been the key measurement in evaluating the general quality of a surveillance system operation. In digital systems, however, both the quality of the video recording and the quality of video processing should be taken into consideration.

Despite the high price, digital video surveillance systems (IP) popularity is constantly growing. It is due to digital surveillance systems being significantly superior compared to standard monitoring functionality-wise.

IP video surveillance provides a solution for various practical tasks while ensuring high quality recording and zooming possibility. Its primary benefit is immense speed of data processing. Even rotary cameras of real-time monitoring are able to record peripheral events, something that analog equipment is incapable of doing.

According to the prognosis published by the analytical company HIS Markit, the number of video surveillance cameras worldwide should reach 130 million in total in 2018. For comparison, this number was less than 10 million in 2006.

1.1. History of emergence and development of video surveillance and analytics technologies

The very first technical means of video surveillance and video analytics appeared over 70 years ago [1]. The process of technological development can be divided into three stages.

First is the period between 1942 and 1970, which marks the appearance of the first highly expensive, yet quite primitive surveillance systems that consisted of cameras connected to monitors through coaxial cables. Each individual camera was connected to a separate monitor. Image quality was rather low, less than 0.3 megapixels (MP); moreover, imagery recording was technically unavailable at that time.

The creation of the multiplexer in the 1970's allowed for several images to be projected onto a single monitor and first recorders allowed video to be recorded onto magnetic tape. However, these devices had issues with video quality. In addition, technical capacity to record video in reference to specific events or distanced video to review remained impossible.

At that time point, magnetic discs as the tool of storing information were pushing out magnetic tape. The usage of magnetic discs lead to more storage space utilization in order to store higher resolution video records. Such semi-digital video surveillance systems offered some technical flexibility, as they allowed the distanced access to the video archive and to the video stream through the TCP/IP protocol. Recording audio and camera control became possible and more convenient.

Finally, the creation of digital video surveillance systems incorporating various types of IP cameras with a vast range of features, universal cabling infrastructures and complex video capturing software marks the third stage of evolution. Such systems are easily integrated; surveillance system installation no longer requires setting up additional coaxial cabling from the camera to the server.

Digital video entered the TV broadcasting industry in the early 1990's and since then it has become a new standard replacing analog TV.

Currently, two types of TV broadcasting prevail on the market: *Standard Definition* (SDTV), characterized by 4:3 screen ratio and average video quality, and *High Definition* (HDTV), which usually utilizes widescreen (16:9) screen ratio, horizontal resolution of 1920 pixels and progressive scanning. This allows 2 073 600 (1920x1080) pixel frame resolution. Framerate may differ and the letter «p» specifies it: for example 1080p30 or 1080p50.

Other HDTV formats are 1080i and 720p. Despite both of them being the same (16:9) screen ratio, 1080i displays 1920x1080 pixels in interlaced video, while 720p displays 1280x720 (921 600) pixels in progressive scanning.

HDTV televisions are based on square pixels similarly to computer monitors, thus HDTV video signals are supported in the same manner on both these devices. On the other hand, displaying high definition video with progressive scanning on a computer monitor does not require deinterlacing.

Many world countries rely on digital TV as a standard – usually in both SDTV and HDTV formats. Most casual users, however, prefer standard HDTV that has a higher resolution and screen ratio, those being similar to a widescreen cinema screen. Nonetheless, video surveillance is based on standard resolution. Therefore, we will continue by addressing questions related to digital video of standard resolution with 4:3 screen ratio.

Another resolution type that should be mentioned is 4K Ultra HD that is 4 times higher than standard HDTV 1080p. 4K Ultra HD video ratio is 16:9. 4K Ultra HD grows in demand on the market and switching to this standard in the field of video surveillance would guarantee high quality video and necessary detailing.

The *4K Ultra HD* presumes using square pixels, similarly to computer monitors. Therefore, *4K Ultra HD* video received from network video surveillance devices may be displayed on either HDTV screens or standard computer monitors. *4K Ultra HD* video using progressive scanning does not require any deinterlacing methods when processing or displaying video on computers.

1.2. Digital video surveillance

Digital video surveillance systems do not have any information transferring restrictions. In basic terms, digital data transfer has several important differences:

- Information is transferred in digital quality;
- Matrix expansion ensures high resolution video capturing in *Full HD*;
- Remote managing and setup are provided;
- A backup power supply is not required.

The possibility of remote setup is a substantial benefit that proves its worth when installing and launching a system. Moreover, the entire digital surveillance system can be set up using an already established office network.

Modern surveillance systems consist of *network cameras* (NC), *video servers* (VS), *network video registers* and *software network video registers* (SNVR) [1].

Modern cameras may be *wired* or *wireless*. Depending on input device type, they could be *analog* and *digital*.

Analog systems are using the principle of signal transferring without distorting fluidity. Digital equipment transfers a pre-segmented video signal, which is later assembled in binary coding. Contrary to analog, a digital recording may be displayable on a computer. This is a highly important benefit of digital video surveillance.

Color-wise equipment can be *black-and-white*, *color* and *with infrared backlight*.

Black-and-white digital broadcasting is known for high sensitivity and its usefulness in low light or complete darkness. Such equipment is able to focus on small details of distanced objects.

HDTV Network cameras set on progressive scanning ensure true-color and high definition even if an object is moving fast. This is a viable option when additional information is required to ensure control in such environments as, for example, in airports, passport control, casinos or highways. Such quality was impossible until video compression methods became effective enough.

Professional network cameras equipped with *Fisheye* lenses have an extra wide scope of view and range of features to meet the requirements set up by most software.

The term *bitrate* is used to define the bandwidth when measuring the effective speed of data transfer through a channel, i.e. the minimal channel size that would allow this stream to be transferred immediately and is represented by a number of bits used to transfer/process data per a unit of time. It is visualized as bits per seconds (*bps)* and its derivatives using suffixes such as kilo- (kbps) mega- (Mbps) and so on. For example, the bandwidth standard for DVD video is around 5Mbps, and for HDTV it is 10Mbps. The higher the bitrate, the higher the quality. Often bitrate is used as criteria of quality of Internet broadcasted video.

1.2.1. Popular video formats used in video surveillance

The Moving Picture Experts Group (MPEG) established in 1988 defined the standards of audio and video coding in different applied fields such as storage, distribution, and transfer of digital information.

Despite video surveillance usually uses a video signal, i.e. compression of animated frames, still imagery compression is also common. In order to differentiate the two, using the terms *animated* and *still* video imagery compression.

Let us note that a minute-long uncompressed video with 30fps frame rate, 720x576 pixel resolution and 16-bit color depth requires 1.5Gb of free disc space (not counting audio).

Therefore, the original video signal, as a rule, undergoes a certain process of compression in accordance with internationally accepted standards.

JPEG (Joint Photographic Experts Group) is one of the more popular graphic formats used to store photo and other imagery. JPEG format files use extensions such as .JPEG, JFIF, .JPG, .JPG and .JPE. .JPG and are the most common across all platforms.

The JPEG algorithm allows for both loss and lossless compression. Imagery supported cannot exceed 65535x65535 pixels in size.

One of the key drawbacks of the JPEG standard is the appearance of distinctive artefacts on recovered imagery with high compression rates – the image gets broken up into blocks of 8x8 pixels. Such an effect is especially noticeable on image areas with gradual color change. In areas of high dimensional frequency (such as contrasting contours and image boarders), it is possible to notice an appearance of distorted halos. Despite the JPEG standard (ISO/IEC 10918-1, Annex K p. K.8) recommending the usage of special filters to suppress the block effect which are highly effective, they are still rarely used. Disregarding certain drawbacks, JPEG achieved widespread popularity because of its degree of compression (which was relatively high comparing to alternatives of that time), true-color compression support and its low computational complexity.

Motion JPEG or M-JPEG is a digital video sequence consisting of still JPEG imagery. As we know, a display of 16 or more frames per second is perceived as video imagery by the human eye. Displaying 30 (NTSC) or 25 (PAL) frames per second is perceived as actual video footage. One of the key Motion JPEG benefits is the quality ensured on the level of compression chosen within the network camera or video coder settings. The higher the compression level, the lower the file size and image quality. In some cases, for example in dimmed lighting or a more complex surveyed object, file size may not only go up, but the file itself requires higher bandwidth and memory space.

Some network video devices allow choosing the maximum frame size to prevent bandwidth and memory requirements to grow. The absence of connection between separate Motion JPEG frames ensures high reliability of this format, and thus, any losses of a single frame that may happen during data transfer will not affect the quality of the entire sequence.

Motion JPEG is not a licensed compression standard. Its compatibility makes it convenient to use in applications that require the presence of separate frames in the video sequence (for investigation, for example) and a lower framerate (usually 5fps). Motion JPEG is also used in applications that require integration with systems that support exclusively the Motion JPEG format. As a result, in comparison with such standards as MPEG-4 and H.264, a higher bitrate and a lower compression level characterize Motion JPEG files.

Compression standards. There are currently many popular compression formats based on different compression algorithms available:

DV (Digital Video) is one of the first video stream compression algorithms. Development began in 1993 in collaboration by the largest video equipment manufacturers (Sony, JVC, Panasonic, Philips and Hitachi). DV format provides a 5:1 compression ratio at a high bitrate, which leads to a large size of the end file. For instance, one minute long DV video takes up around 200Mb of disk space (1 hour – 12Gb).

This format is most common in compressing video filmed on household digital cameras and professional camcorders. Due to a low compression coefficient, the filmed media remain in very high quality while the compression process takes place in real time without requiring powerful technical components.

MPEG is virtually a family of digital data compression standards. The group under the same name developed and standardized it. The initial standard of video and audio compression was called MPEG-1. In 1993, the renowned Video CD (VCD) specification was developed in collaboration with JVC and Philips. Video CD is the standard format of compressed video and audio storage on regular compact discs.

Utilizing MPEG-1 algorithms in coding allows for up to 1.5Mbps bitrate with resolution of 352x288 pixels for PAL and 352x240 for NTSC. In 1995, the MPEG-2 standard was introduced. It was most commonly used in digital video discs (DVD) and in cable and satellite TV broadcasting. The technical capacities greatly superseded those of its predecessor: with 25fps the resolution is 720x576 pixels in the PAL system, and in NTSC it is 720x480 with 30fps, while the average bitrate is of 9.8Mbps, which is almost 7 times the bitrate of Video CD. Another undeniable benefit of MPEG-2 is its capability to save five-channel audio (*Dolby Digital 5.1* and DTS).

Alongside the MPEG-2 standard another compression standard, MPEG-3, was being developed. It was intended for audio and video encoding for high definition television at the data transfer speed of 20-40Mbps. Development of MPEG-3 was terminated when it was discovered that a modified version of MPEG-2 could be used for this purpose. The MPEG-3 standard is no longer used.

Finally, in 1998 another video compression format family, MPEG-4, was established. His purpose was the enhancing image quality at a lower bitrate. The preceding MPEG-2 standard was designed for higher bitrate; therefore, its algorithms was substantially modified. MPEG-2 is also unavailable for storing HD video with 1280x720 (720p) and 1920x1080 (1080i or 1080p) resolutions which are constantly growing in popularity.

As a rule, in surveillance systems the MPEG-4 format corresponds the standard MPEG-4 Part 2 also known as *MPEG-4 Visual*. As are all MPEG standards, MPEG-4 Visual is licensed and it therefore requires users to acquire a license for each monitoring station. The format is applicable when high image quality require and practically unlimited bandwidth and no framerate restricted.

Currently MPEG-4 is the main multimedia content compression standard. Even though it is too soon to disregard DVD, nearly all modern photo and video cameras record in HD quality.

Codec. Different standards are being used for video compression. Even with a set algorithm of data transformation, a video may be compressed using different tools or software, which results in an outcome that may vary each time. Most of the differences are defined by the *codec* – a special program that compresses (encodes) source data. Each codec uses a unique algorithm that affects both the quality and the speed of encoding.

The term «codec» is a blend of words «coder» and «decoder». It means that a codec must include not only a module of data encoding (coder), but also a module of data decryption (decoder). The latter is usually a freeware and included into popular codec packs such as *K-Lite Codec Pack* or *Windows 7 Codec Pack.*

Let us review some common codec types.

MPEG-4 Part 2 ASP is one of the first algorithms, introduced in 1999. Codecs based on it provide a rather low quality. Nonetheless, this drawback is compensated by a high processing speed and low hardware requirements. Well-known codecs based on this algorithm are the commercial DviX and its free alternative Xvid.

MPEG-4 AVC10/AVS or *H.264* is a very popular algorithm for compressing both low-resolution video and HD. The abbreviation AVC stands for Advanced Video Coding. As with the previous, this codec family consists of both free variants (x.264) and paid versions used in popular video editors.

H.264 is an open licensed standard for lossless video compression.

One could argue that it was *H.264* that lead to HDTV being implemented in video surveillance. Its effective compression ensures high resolution and framerate at a 16:9 frame ratio. *H.264* is expected to become the most common standard in the nearest future. *H.264* allows data to be compressed 80% more effectively than *Motion JPEG* and 50% more effectively than *MPEG-4*, without causing any damage to the image. On one hand, this results in diminished bandwidth and memory requirements. On the other, it allows to produce higher quality video imagery at a standard bitrate.

In the field of security video surveillance *H.264* is applicable in activities that require a higher framerate, for instance, in monitoring highways, airports and casinos, where 30/25 fps (NTSC/PAL) is the norm. The biggest saving can be achieved by lowering the bandwidth and free space designated to saving data.

Media containers and their formats. Digital information and video is stored as files, also called «media containers». Media containers may store video, audio and other streams along with metadata. At any moment data can be retrieved from the container, re-encrypted and stored in a different container (i.e. the format of the file may be changed). Multimedia containers may be of different types (formats).

Despite most containers being associated with a specific format, some may store data in different standards. For example, a file with the extension *AVI* is able to contain clips in *MPEG-1*, *MPEG-2* or *MPEG-4* formats. A codec and the compression settings make a substantial influence on the video quality. However, the media container makes a difference as well. Different types of video files have a different amount of audio tracks**,** subtitle channels, codec types used, household player compatibility etc.

The *AVI* (*Audio Video Interleave*) was first used by *Microsoft* in 1992. It can contain a combination of video and audio information compressed by different codec types. Therefore, *AVI* files may seem similar, but in reality, the content may be significantly different.

In fact, the *AVI* container is no longer up-to-date and has a number of significant flaws: its inability to store mixed video (for example, both NTSC and PAL) or alternative audio tracks, the absence of time marks and frame indexes, inability to operate subtitles properly etc. Nonetheless, a lot of Internet media content is to this day stored in this format, probably due to its versatility.

Mp4 (MPEG-4 Part 14) is one of the newer file formats for digital video and audio storage, that is part of the MPEG-4 standard.

TS and M2TS are specialized containers of HD video. Streaming digital TV *IPTV* and *DVB* use TS. M2TS is the standard container for *Blue-Ray* video that can include video and audio streams supported by the *BD-ROM* standard along with subtitles in the *PGS* graphic format.

1.2.2. Digital video recorders

Digital video recorders (*DVR*) are used for multifunctional video imagery operation (including receiving, transferring and recording of information).

Construction of surveillance systems placed in shopping venues, office environments, banks, car parks, gas stations, road checkpoints, transport enterprises etc. using video recorders.

The main types of video recorders are:

- *Digital video recorders (DVR)* are used in analog cameras. The cameras are connected to a DVR using a cable through a BNC jack. DVR receives an uncompressed video stream and saves it. The benefits of DVR are the independence of video stream processing (the recorder process and archive it) and complete functionality that is not dependent on an Internet connection.
- *Network video recorders (NVR)* are connected to IP video cameras and are used to create a surveillance network controlled by a single central point. NVR receives a signal that is already compressed. NVR benefits are software flexibility and high quality imagery received by IP cameras.
- *Hybrid video recorders* are an intermediate type of recording devices adapted for both analog and IP cameras. Analog cameras can be connected to the recorder using the BNS, while the Internet protocol (IP) cameras are connected via Internet. Hybrid video recorders combine the benefits of both DVR and NVR. Hybrid video recorders allow establishing flexible surveillance systems that may process information on two streams. The information recorded by

hybrid video recorders is stored both in local and network archives.

When choosing a video recorder one of the base parameters that should be considered is *multichannel.* Recorders should process from 4 up to 32 channels.

Digital video recorders and network TV cameras have benefited both the development of the video surveillance industry, as well as the creation of «smart» video surveillance systems. Digital video recorders have led to the line between computers, IT and video surveillance factually disappeared.

1.2.3. Video servers

A *video server* (VS) is a computer device used for receiving, storing and reproducing/rebroadcasting of a video or audio signal, as well as imagery processing (including IR images), processing of telemetry data, and managing multiple surveillance systems. From functional perspective, the video server is an advanced upgrade of a digital VCR.

Video surveillance and audio control systems implement video servers as intermediary equipment. Video servers suit as end devices if they are installed in the monitoring facility. Servers exist as serial devices, as well as in a shape of video and audio capturing boards combined with specialized software, allowing manual server setup.

In addition, video servers may serve as a core of integrated surveillance systems. In such systems servers aim to:

- manage the access control system (by implementing firmware for image processing able to recognize faces, vehicle number plates, wagons etc.);
- process audio signals and identify voices,
- analyze thermal imagery (in a scenario when heat distribution may visualize the amount of liquid fuel in a wagon tank, for instance);
- detect speeding vehicles with speed cameras etc.

In integrated surveillance systems, video servers are often replacing management devices for alert systems, including security and fire alarms. Modern video servers process certain telemetry elements, monitor cash transactions and account working hours in enterprises.

1.3. Network technologies in video surveillance

In security video surveillance solutions image quality is the key indicator ensuring certainty of the validity of the recording and participant identification in the surveilled zone. Network cameras are able to achieve better video quality and higher image resolution in comparison to analog cameras by implementing progressive scanning and megapixel technology.

Worth to notice, however, that achieving higher image quality in network surveillance systems is much easier than in analog surveillance systems.

Currently analog systems that implement DVR end up transforming information from analog into digital and back multiple times: first, the analog signal is transformed to a digital format by the camera. Then it is changed back for broadcasting. Finally, it is once again digitalized when it is being recorded. Meanwhile the quality of saved imagery is getting worse with each transformation and with the increase of the cabling length. Furthermore, the longer the transmission distance of the analog video signal, the weaker it becomes.

1.3.1 IP cameras

The key component of any surveillance system is a camera and digital video surveillance is not an exception. An *IP camera* is a digital video camera used for digital video stream transmission through Ethernet and TokenRing networks, both of which follow the IP protocol. Each network device and IP camera is assigned a proper *IP address*.

Contrary to using analog cameras, IP camera use never requires to convert digital signal to analog after a digital video image is received through a charge coupled device (CCD), or a camera based on CMOS technology (complementary metal oxide semiconductor). Such video viewing is possible through a browser installed on a computer, whereas digital imagery captured via *IP cameras* can be viewed, processed and recorded using a number of software-hardware instruments.

When implementing progressive scanning IP cameras with HDTV support are able to capture true color high quality video even of fast-moving objects. That is why this type of cameras is preferable for control operations – in such venues as airports and highways – when having more information about event details is virtually a requirement.

The use of IP systems is also called digital or network video surveillance since it is based on the idea of a computer network. Thus, digital video surveillance possesses all of the capabilities, benefits and drawbacks that a computer network responsible for video surveillance would.

1.3.2. IP video surveillance

The basic components of a *network video surveillance system* (also known as *security IP video surveillance*) are a network camera, a video coder (used for connecting analog cameras), a network, a server and a storage system, as well as video management software.

Network surveillance system use a wired or wireless IP network for video, audio and other data as a transmission environment. When using Power over Ethernet (PoE) technology within a network, it is also possible to power network surveillance devices. Let us note that this technology is unavailable for analog devices.

Computer-based network video cameras and videocoders demonstrate functionality that is unavailable for analog cameras. In an entirely digital IP video surveillance system, the imagery once digitalized within the network camera remains digital without any additional modifications, quality losses or dependency on transfer distance. Furthermore, digital imagery is easier to store than analog video cassettes.

A network video system allows reviewing or recording video via any network node, disregarding whether it is a local network or a global one. Another evident advantage to IP video surveillance is that it is easy to scale.

A network, storage systems and servers constitute a standard IT equipment set for video surveillance. The total cost of IP surveillance system setup is lower than the cost of an analog system setup.

It is common for an organization to have a pre-established IP infrastructure, used for video surveillance. Wired and wireless IP networks in general are a less costly alternative of traditional coaxial or optical cables of analog video surveillance systems. Moreover, digital video streams can be transmitted around the globe using various connection channels. Implementation of server equipment that follows industrial, as well as open standards of video recording and storage, allows lowering the equipment and system management expanses when comparing to specialized firmware expenses, such as costs of using DVR in analog surveillance systems.

When it comes to users, network video surveillance systems are actually limitless. In many cases, however, access can be limited due to the risk of unauthorized access.

IP cameras are connected to an already existing local network or any computing station of an organization. Usually, IP camera manufacturers tend to provide video recorder software as part of the package.

Other IP surveillance systems benefits are:

- Equipment is easy to connect and set up;
- IP cameras allow high quality digital video signal transferring;
- IP video surveillance can be introduced onto the basis of video cameras and network equipment that is already in use;
- Possible to connect a smartphone, a tablet or a laptop for remote viewing.

An IP camera is equipped with a *Wi-Fi* or *Ethernet* interface and connected to a local network. A web server is integrated into the camera, allowing video to be viewed from anywhere. Security is ensured, as the digital signal is encoded, minimizing the risk of unauthorized system access. Wherein IP cameras also go into the video recording mode only when movement detected in the surveyed area.

Although enhanced camera resolution puts extra strain on the network. However, most tasks can be solved using as little as 100 pixels per square meter. Moreover, many types of megapixel video cameras are simply unable to broadcast video with framerate over 10-15 fps.

Taking into consideration these factors, upon optimizing IP camera resolution, an appropriate data compression type should be chosen, say *H.264* or *MPEG*.

1.3.3 Remote video surveillance system management (VSaaS)

Video surveillance as a service (*VSaaS)* based on cloud infrastructure is an important trend for the industry. According to *IMS Research* (UK), the market value of VSaaS was globally $500 million in 2011, and the value doubled in the three years to follow. US data shows that the market value of business video surveillance tools has increased 2.5 times over the period of 2011-2016 reaching $900 million.

Unlike standard video surveillance management, the interface of a VSaaS cloud application is based on a web browser or mobile device. At the same time, VSaaS allows for more complex channel management technologies between the object of monitoring and the cloud, as well as between the cloud and the user. A video management system requires operators to be available at the monitor at all times. VSaaS users are only required to connect to the system using a browser only to react to emergency signals, analyze the video archive and make reports [11].

To introduce VSaaS the client does not have to establish an entire infrastructure of data collection, storage and processing. The customer simply sets up video cameras where the service provider covers necessary and other functions.

Depending on responsibility distribution between the client and the provider, VSaaS services are divided into *video hosting* (footage is sent directly to the data center of the provider for processing and storage) and *surveillance management services* (when footage is stored by the client and the provider is responsible only for remote monitoring). Combinations of those two are also possible, such as in instances when the client stores a backup copy of footage.

Reinstances when the client stores a backup copy of footage.lying on the VSaaS model has many benefits. Users no longer need to buy expensive DVR and NVR equipment or ensure its maintenance, security and protection from various attacks. In addition, purchasing *Video Management System* (VMS) software is no longer required as well, video stream and video analysis tools are accessed through Internet.

The evident benefits of the VSaaS model have lead to the provision of video surveillance cloud services by many world companies. Among them are the companies traditionally specializing in video camera and VMS sale, firms specializing in access control and video analytics systems, as well as numerous start-ups. *IMS Research* expect the global VSaas market to double in the next three years. The argument behind this prognosis is the range of fields where VSaaS can be implemented.

Many companies offer only basic video management functions and take extra for enhanced functionality of such solutions. Up to date only a number of providers have realized the video content analysis functionality (*Archerfish, VIAAS*) or access management integration (Brivo). More and more providers offer support for local storage devices on the client's side (cameras with SD cards or cameras connected to DVR or NAS systems). Nonetheless, such functionality is not yet ubiquitous. Usually, the average VSaaS service fee is about $5-30 per camera.

Video surveillance as a service is focused on video analysis without an operator present, while the term VSaaS usually implies only the possibility of remote video viewing and storing without any additional analysis.

For some user segments, the VSaaS service is preferable to classic solutions based on network video recorders and video management systems. The commercial VSaaS model implies that instead of the price of a hardware-software solution with no guarantee of investment return, the client pays for a specific service, such as video recording, automatic security calls, data collection and analytic report creation. The VSaaS service is scalable by stored video size, number of surveillance points and the number of system users.

Video content analysis may be considered to be a specialized coder that filters out data necessary for the user. The universal coder, such as *H.264* does not «understand» the degree of importance of each image element – therefore it is unable to effectively filter excess data in order to be able to provide VSaaS services. For example, a standard coder is unable to differentiate a human figure in the background and a snowflake in the foreground. Should a person and each snowflake be coded with equal level of idealization, the video stream would be too overloaded to transfer and save data.

Video content analysis is regarded to be the sole technology able to solve problems related to outbound user channels and video storage on clouds. Despite the emerging of more accessible video storage ways, such as *LAID (LinearArrayofIdleDisks),* storing copious amounts of video on clouds is still considered to be the most costly aspect of VSaaS.

In modern times, we notice a shift from local biometrics to cloud. Cloud biometrics supersedes local in development, mainly in financial software [10]. The results of a research conducted by the company *Acuity Market Intelligence* confirm this trend.

Such a change does not indicate that local biometrics is becoming less relevant. The technology is based on firmware installed on the input device. Specialists predict that by 2022 5.6 billion mobile biometric devices will be in use, annually processing 1.4 trillion transactions.

In local biometric, the user only confirms with a single device, whereas cloud biometrics ensure unique customer identification that is independent from the device and the platform that it is run on.

In addition, cloud biometrics can be enhanced with behavioral biometrics that ensures constant passive authentication of people in its range of use.

1.4. Review of the current state of the video security market

Today, it is possible to distinguish the following segments of the video security market:

- Network cameras;
- NVR equipment and software;
- Video servers.

Video surveillance is becoming the basis of a growing number of projects. According to the estimates by *IMS Research,* the annual growth of the global market is 12%.

The wide availability of IP cameras, constant reduction of their prices and expansion of their functionality make a positive impact on the market growth dynamic. The same impact made by cloud technologies that allow users to reduce equipment expenses. The very transition to cloud services provokes the need to process larger quantities of data using *Big Data* technological principles.

The network camera segment has been showing steadily progress with the emergence of more video content analysis solutions with enhanced functionality, while VMS and physical security management solutions offer more user-friendly interfaces.

98 million network cameras and 29 million HD CCTV surveillance cameras were sold worldwide in 2017, through professional sales channels. In addition, 400 thousand next-to-the-skin cameras were provided to various law enforcement agencies around the world [9].

The demand for certain camera types, such as panoramic view (of

180 and 360 degree view), stereoscopic and thermal is expected to grow particularly fast.

The company *Arecont Vision* envisions fully embrace the growing demand for panoramic video surveillance by using megapixel cameras with wide dynamic range of view. By combining scene elements with different exposure in a single image, wide dynamic range (WDR) technology allows maximize the number of important details both in bright and in dark areas of the scene.

The development of video content analysis and the *Internet of Things* (IoT) also influenced the growing demand for video surveillance. Video content analysis allows automatically process imagery instead of having a review by human. IoT integrates video surveillance, making it part of a «smart house» (or a «smart office»), which in its turn increases the demand for the solution [7].

In a number of countries, video surveillance has led to crime decrease. For instance, by 2020, all key public venues in China are expecting to have video monitoring. Germany issued a law that encourages stadiums, stations and shops to set up video surveillance, while police agents are allowed to wear portable cameras.

Market segments where video surveillance services are high in demand include transport, banking, utility networks, and energy provision, nonetheless, requestors often change requirements to security systems in those fields.

Nowadays organizations seek using video surveillance technologies to a greater degree than traditional video recording and real-time security monitoring.

Requestor outlook on ensuring safety is changing at its core for such large and expansive objects as parking lots, warehouses, stadiums, conference halls, fitness clubs and sports arenas. Monitoring a stadium-size object may require up to 2000 cameras and kilometers of cable. Meanwhile, modern technologies allow full coverage of a stadium using only 75 cameras.

IP system benefits for bank security are not quite evident. Having in mind the fact that banks have successfully invested into analog CCTV infrastructure, for most customer it is impossible even to think about the migration onto a network system. Nonetheless, the high-resolution analog video surveillance is a topic of interests for such clients.

Video surveillance in any weather conditions has become a reality. There are cameras able to film in backlight conditions or by direct sunlight with it having no effect on the scene detailing. Other cameras form colored imagery in almost complete darkness. Cameras can be successfully calibrated to sustain vibrations, physical hits and even fog.

Social networks and video surveillance. *IHS Technology* company experts believe that the widespread dissemination of smartphones equipped with cameras and connected to the Internet, in combination with the development of popular social networks has led to the phenomenon that the user community, for the first time in history, was able to collect a unified video footage database usable for police investigation. This is essentially crowd sourcing by definition. This trend is expected to grow during the years to follow.

According to *Cisco* expert estimates, the user traffic in the Internet has reached 1.4 zettabytes (1.4x10^21) in 2017. The entire amount of video uploaded by mobile devices between 2012 and 2017 has increased by 16 times, accounting for 2/3 of complete mobile traffic.

Video surveillance in housing security. *Perimeter cameras* can verify alarms that are triggered by security system sensors. This allows avoiding false alarms and the need to contact law enforcement about false cases. Demand for remote hosting services by house owners is also on the rise.

Security experts who participated in a survey by a US magazine SDM unanimously agree that the video surveillance market remains stable and prone to expansion. Among contributing factors is a vast variety of video software, simplified integration and some new technologies that begin to disseminate, such as content analysis, big data and other.

Nonetheless, the expansion pace may lag due to camera price reduction, market oversaturation and more even analog to IP transition.

Market drivers. One of the factors that always effects the market is compliance, especially, when end users do not comply with the standard. Vertical markets, such as utility services and healthcare have an abundance of compliance drivers.

Another major influence on the video surveillance market is technological achievement and the correlation of technologies that have never crossed paths. We are talking about the perspective of using detectors and devices in the IoT field that provides clients with effective market management tools.

Another major influence on the video surveillance market is technological achievement and the correlation of technologies that

have never crossed paths. We are talking about the perspective of using detectors and devices in the IoT field that provides clients with effective market management tools.

Video services are currently only present in portfolios of two or three major integrators and software developers – usually as a technical foundation of other systems (access and rapid response mechanisms). Another possible problem is the fact that transitioning to IP technology is almost complete. The market for IP solutions, capable of integrating with other systems, should be growing in the next several years.

Users considers already known IP camera brands that are more widespread and introducing cybersecurity protocols into network devices. In 2018, video security is one of the main directions for all product providers, integrators and end users.

According to the prognosis of *IHS Markit,* 75% of all video surveillance servers that support *deep learning* shipped globally were produced in China. It is expected that the VSaaS market should rapidly grow, increasing by over 10%. Meanwhile, the main question concerns retailers, small businesses and residential complexes. VSaaS allows using any kind of devices, such as laptops, smartphones and stationary computers for remote access to video streams stored on clouds.

As video resolution is constantly improved, its compression becomes more vital. Implementing the *H.264* standard made an impulse for mobile video popularity. The industry anxiously awaiting the *H.265* standard that should further enhance video compression effectiveness. *H.265* supersedes both *Google* and *Amazon* internal compression protocols. Another benefit of cheaper cameras is a wider video audience.

Confidentiality is required in Europe. Confidentiality and its protection is particularly vital in Europe, especially in Germany and Scandinavian countries.

Video surveillance data storage perspectives in 2018. In 2017, the VSaaS business model grew in popularity in Europe and North America. On the other hand, the evident technological progress of video cameras and other equipment cannot be disregarded. Meanwhile, service providers underline that they do not supply standard monitoring, but video content analysis, often with features of biometric data processing.

FCX.iQ

Solution for currency exchange

FCX .iQ
.iQ Family Product

FCX.iQ is the software solution allowing to carry out operations of currency exchange at Diebold Nixdorf ATMs. Furthermore, BS/2 offers expanding ATM functionality in multivendor environment.

Chapter 2. Video content analysis: hardware and software

«In general you could not assume that you were much safer in the country than in London. There were no telescreens, of course, but there was always the danger of concealed microphones by which your voice might be picked up and recognized...»

G. Orwell «1984»

Video analytics, also known as video content analysis is a methodology of applying computer analysis of video content to obtain finite data on particular surveyed objects. Video content can be defined as any series of images received either in real-time from a video camera or withdrawn from a video archive.

Applying video content analysis enhances video surveillance system effectiveness by reducing the workload of security and monitoring agents. It also allows complete video quality appreciation whilst rendering the IP camera system smarter at its core. That is achieved by applying object type recognition algorithms and recognition of behavioral scenarios in real time, as well as varying notifications and reports being available to system users.

In practice, video content analysis is performed using specialized hardware and software solutions and do not require direct monitoring by a person. Video content analysis algorithms can be integrated into various business systems including video surveillance systems.

For instance, one of tasks performed by applying video content analysis is defining movement parameters of an object based on a series of images of a three-dimensional scene: defining the speed of a particular image area or determining the movement trajectory.

Applying video content analysis leads to automation of the following tasks:

- *Detection* – video data is checked on its compliance with certain conditions. For example, detection based on quick calculations can be used to locate small areas of an analyzed image that would then be analyzed using interpretation algorithms that are more resource demanding.

Content may be defined, say, in terms of an image resemblance of a source image, or complex textual search criteria (ex. locating all shots taken during winter with houses without cars nearby in the background).

- Object *proximity and orientation* relative to a camera. Critical situations in the video stream (ex. groups of people, unsupervised objects, fire and smoke, etc). Should a problem occur it is required sending a message to the monitoring station and registered mobile devices, while simultaneously recording the incident and creating a protocol.
- *Tracking* – following the movement of an object, such as a car or a person, in the surveillance zone.
- *Identifying a specific unit. Identification may imply a variety of tasks* – starting with classifying an object onto target/noise, and ending with biometric verification of an object. Such data as a person's face or fingerprint can serve as parameters of biometric data.

Above-mentioned functions are performed multiple times aiming to confirm any occurring hypotheses on object quantity, location and type in the surveillance zone, as well as to resolve result bias.

When applying video content analysis, the questions such as «Where exactly should the analysis take place?» should be answered multiple times [4].

2.1. Movement detection and its analysis

Video content analysis software for surveillance cameras has several types: it can be set up on a user's camera, on an NVR device or as software provided by third parties. Locally set up applications are scaled to reduce traffic and storage space consumption. That is achieved by recording and transmitting only those video fragments that bare real value.

For instance, many retailers use video surveillance systems to detect movement in shopping venues past working hours. The system can be specifically set up to detect movement after a shop working hours, and should movement detection occur, the monitoring station will be notified immediately.

Foreground is the shot area where the events take place. These areas are highlighted using segmentation algorithms and each individual segment is analyzed individually. Such algorithms, for example, allow defining areas that correspond to area types such as «head», «eyes», «shoulders» etc. When detecting a «head» area, the algorithm will follow its framework and define eye position and other human body parts by calculating relative distance. Should no 'key' elements correspond to the «head» hypothesis, the algorithm will dismiss the task and continue with the next hypothesis. By comparing two consequential shots or the current shot with a pre-saved default shot, an incremental image can be created. Using this image (mask), forefrontmovement can be detected.

Motion Detection. Actually, all modern network cameras incorporate motion detection.

Motion Tracking. Tracking involves highlighting an object of interest and tracking its movement. To perform this task, algorithms use a current image as a point of reference and analyze any changes that take place. During comparative analysis of consequential shots, all movements taking place in the foreground can be isolated form a background.

The difference between simple motion detection and motion tracking is that the first allows detecting any changes that occur in sequential shots, whereas the latter allows detecting objects that meet certain logical conditions within a particular area.

2.1.1. Stereoscopic perception

Many applications take into account the three-dimensional nature of both the object of a scene and the image capturing process.

Such tasks as defining the location of an observed object in relation to the camera or differentiating a completely or partially obstructed object within a shot require three-dimensional information about the scene. This information may be used as a supplementary means of two-dimensional shot analysis [3].

The capabilities of the algorithm used for three-dimensional shot analysis can often be expanded by incorporating certain mathematical attributes of the process of image recording. The methods that spark our interest are those that allow us to answer questions requiring three-dimensional information about the scene or those that incorporate information about the three-dimensional scene as a supplementary means of two-dimensional analysis. Both of those tasks are resolved through the means of a single mathematical model.

As known, *perspective transformation* is the process of projecting a multitude of points onto a single point, which implies that it is impossible to unmistakably define the location of a single point on an object based on the coordinate that is relative to it.

The method of extracting additional data necessary for eliminating ambiguity based on two separate images of a single scene is called *stereoscopy*.

Despite each individual point of an object being assigned only a single definite point of an image, all points of an object that are located on a line that passes through the center of the lens share a single projection. Thus, for each point of an image there is a line in space defined by the very point of the image and the center of the lens, it is on this line that the corresponding point of the object should lie (fig. 1)

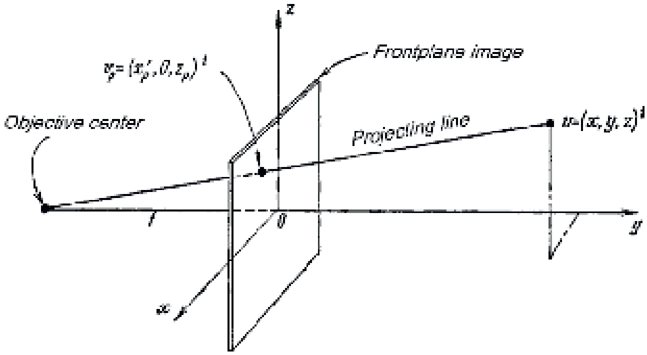


Fig 1. Model of a camera with a frontal image plane.

Therefore, two significant tasks relating to the image capturing process emerge:

- to define the location of a projection for every point of an object in a shot;
- to trace a straight line where a projection of a point of an object should lie for any given point of a shot.

Both these problems are solved through forward and reverse projective transformations [3]:

$$x_p = \frac{fx}{f+y},$$
$$y_p = 0,$$
$$z_p = \frac{fz}{f+y}. \qquad , \qquad x = \frac{x_p}{f}(y+f) = \frac{x_p}{z_p}z,$$

A simplified scheme for stereoscopic perception is depicted in fig. 2. It portrays two image planes $\boldsymbol{I_1}$ and $\boldsymbol{I_2}$, two lens centers $\boldsymbol{L_1}$ and $\boldsymbol{L_2}$ and two projection rays $\mathbf{r}_1$ and r_2 that are traced through corresponding lens centers and the V point of an object. The vector $\boldsymbol{\Delta}=\mathbf{L}_2-\mathbf{L}_1$ is called the «base vector» and its length is referred to as the «base».

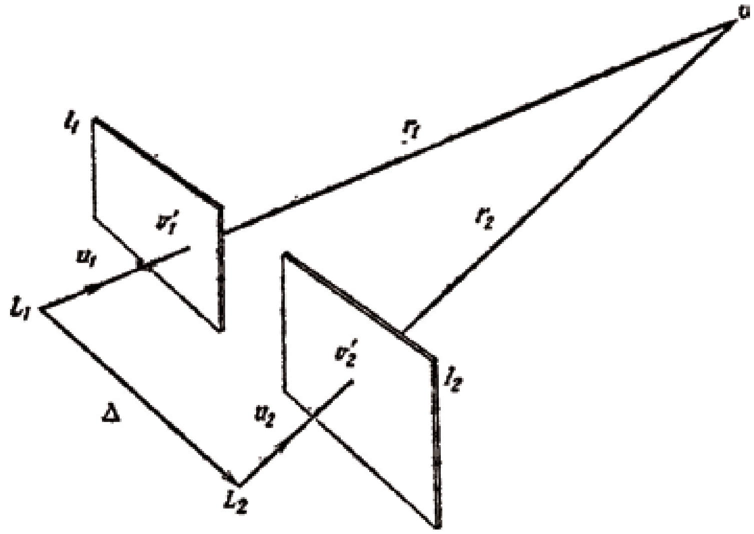


Fig. 2. The primary scheme of stereoscopic perception.

Stereoscopic calculations are conducted in two phases. Firstly, the location of two image points V_1 and V_2 corresponding to the object point V should be defined. Second, using trigonometry the intersection point for two projection rays should be found. The first task, formally known as correspondence confirmation is usually performed using one of two methods. The easiest method is to define the location of the $\boldsymbol{V}$ point projection on each image using any means available. However, sometimes it is more convenient to define the location of the $\boldsymbol{V}$ point projection on a single image and then locate it in the other image by via benchmark comparison with the first image. Referring to our example let us first define the location of the point $\boldsymbol{V_1}$. Then, as we are investigating a point (or a small area) of the plane $\boldsymbol{I_1}$, that would correspond to the point $\boldsymbol{V}$, we will use a small plane area with its center located in the point $\boldsymbol{V_1}$ as the benchmark.

Such a benchmark is then relocated around the $\boldsymbol{I_1}$ plane until a match is found.

2.1.2. Calibration

The analyzed projective transformations include a number of geometric parameters. Even a simplest of transformation cases would require knowing the distance between the image plane and the lens for it to be successful.

Any inaccuracy whilst setting the camera position may lead to undesirable errors in defining the location of a surveyed object.

Though some parameters can be measured on the spot, it is significantly more convenient in practice to define at least a minor part using the camera itself as a measurement tool.

The key idea is to assign the parameters such values that would minimize the difference the measured and calculated points of an image.

To solve real-life problems of video content analysis *external and internal calibration* of cameras is applied.

External calibration includes assigning parameters that define the general common origin a coordinate system and the location of a camera in global coordinates, camera installation height and its view angle.

In practice, in order to avoid mistakes, all cameras are calibrated simultaneously – the images are assigned certain points the coordinates of which are already known.

Perspective scaling is a very reliable method of external calibration that is based on setting size ratios of objects typical for a scene. To do so, sizes of humans, cars and technical equipment are defined manually in pixels, separately for the lower and upper parts of the image. This allows for humans and other objects of certain pixel sizes to be identified based on their geometric proportions, movement etc.

Internal calibration allows correcting image geometric distortion. Camera tends to distort not only the angles, but the size proportions as well. Each camera should undergo internal calibration separately. In order to correct geometric distortion, a special object is put into the scene meant for camera calibration.

2.2. Multi View monitoring

The accuracy of defining locations on a map highly depends on camera position. Should the cameras be slightly tilted in reference to the horizon, the bias may be very significant. Analysis errors may also lead to significant inaccuracies of automatically detected object coordinates and the true location of said object.

Currently on the market, there is a variety of multichannel monitoring means. For example, the device *Blackmagic MultiView 4* manufactured by *Blackmagic design* grants simultaneous streaming of four independent SDI signals. Each *6G-SDI* port ensures complete second synchronization allowing operating any combinations of *SD, HD,* and *Ultra HD* formats of any frame rate. The device supports *1080 HD* and *2160p Ultra HD* and is *HD* and *Ultra HD* display compatible. Operating an Ultra HD TV or monitor highlights the 1920x1080 HD possibilities. In addition, sound level indicators and window identifiers can be set up.

Another useful tool is the search engine MomentQuest developed by *AxxonSoft* that is used to quickly locate faces, car numbers or events based on set criteria in a video archive. Search and monitoring is conducted through video metadata that is calculated automatically for all objects that are caught on camera. Metadata gets stored alongside recorded video it refers to.

In order to search the face the system uses a biometric vector (brief face description) as metadata for all people in a shot. Whilst performing secondary searches the user will be able to simply upload a photo of an individual that is being searched for.

All vehicle registration numbers are stored in text. Simply, the system user is able to input a car number into search parameters.

In order to find specific events the user fills out search parameters and the system finds all archive matches in seconds.

2.3. Behavior scenarios

As a rule, automated detection of potential hazardous situations is performed based on known behavioral scenarios that are typical for such scenes in certain operational settings. Formal behavior description for its comparison with behavior scenarios is formed from data retrieved from installed detectors.

Keep in mind the paragraph 3.2.1 lists the scenarios of detecting critical or suspicious situations by implementing facial recognition in self-service devices using the *ATMeye.iQ* system.

Other examples of such module implementation is the Security *Manager and Fraud Detection of the complex software solution Vynamic Security by the company Diebold Nixdorf.*

The Security *Manager* module is used to detect scam based on template event and BIOS password management correlation. The Fraud *Detection* module uses Big Data technology and machine learning, insures tracking inconsistencies in standard program, process or user behavior scenarios. Anomaly information gets reported to security personnel in real time and any of the information, finance or property protection scenarios con be initiated.

2.4. Video analysis hardware

As a rule, modern cameras incorporate several kinds of counters and extended motion tracking algorithms. Outing results of such algorithms supersede typical motion tracking device results. Modern «smart» cameras significantly supersede the functionality of such motion detectors.

For example, the «Intellect» surveillance system developed by *AxxonSoft* incorporated three different kind of detectors: basic, situational and servicing.

Basic detectors. This group of motion detectors is used for detecting moving objects in a frame. A basic detector detects movement with no extra conditions. A tracker identifies movement and its direction, is able to follow the object even should the camera shake. An IR detector (that requires a thermal imaging camera) detects movement in the infrared spectrum and a movement direction detector detects movement in pre-set directions.

Left/lost item detector informs about an appearance a disappearance of an item in the frame. This detector allows, for example, to detect a laptop gone missing from a table, a case forgotten in a hall or an unlawfully- parked car.

Face detector is able to recognize an appearance of a person in a frame, distinguishing this person from surrounding objects.

Situational video detectors are supposed to detect specific pre-configured object movements in a frame. The operator may set certain lines, polygonal zones and periods for the system to operate based on those criteria. Situational video detectors may identify:

- object crossing a line in a specific direction;
- object crossing a polyline in a specific direction;
- movement in an area;
- object entering an area;
- object exiting an area;
- object's appearance in an area;
- object's disappearance in an area;
- object freezing in an area;
- object's presence in an area for a time period of over 10 seconds;
- item left in an area.

Any of the detectors can be set up for operations with a specific object type such as humans, cars or all objects.

Service detectors signal camera operation interruption. They allow not only identifying camera disabling attempts, but also any sort of interferences that disrupt quality image capturing.

Lens obstruction detector identifies all cases of unintentional or intentional camera lens obstruction. It is especially valuable when a camera is exposed and within reach.

Backlight detector signals cases of bright light such as of a flashlight or a laser being aimed at the camera.

Camera movement detector is triggered when a camera being spatially manipulated. It is especially valuable with cameras that are within human reach and are easy to move.

Background change detector reacts to changes in camera background. Similar to the one described above, yet this detector is meant for slightly different tasks. Whereas the movement detector reacts when the camera is being moved, the background change detector is triggered by changes surrounding the camera, such as to an effort to set up a fake background.

Focus loss detector informs about the image clarity issues due to the lens being out of focus or it becoming dirty due to accidents or tampering.

We should also name main differences of video detection of an HD system and that of a standard resolution [15]:

- Enhanced detailing of an HD scene leads to an elevated number of false triggers. The camera movements due to wind or equipment vibrating become a solid reason for image quality fluctuations. Using a digital or a mechanic stabilizer becomes mandatory;
- Megapixel cameras have a higher range and view radius. It allows for objects of a bigger scope to be captured, when comparing to SD camera systems. HD systems also require applying fundamentally different proportional scaling and optical correction algorithms, as well as implementing multidimensional algorithms of background modeling and object segmentation;

The data bandwidth of HD systems is significantly larger than that of SD systems. Most algorithms used in smart video detectors have nonlinear complexity in reference to the frame size and the processor workload is increased several times. Thus working with HD stream requires for in-depth optimization and creation of new algorithms.

2.5. Video analysis software

Video content analysis software assists monitoring the constantly growing amounts of video footage that security officers and system managers are no longer able to cover alone. Let keep in mind that a video surveillance system is only as valuable as the number of factual incidents it is able to capture.

Nowadays video content analysis is on the list of innovative technologies proposed by *video management system* (*VMS*) providers.

Software for video content analysis can be installed onto the camera or other video recorder or as third party developer software.

Video content analysis solution operation varies depending on the developer and the field of use. However, most of them run in a similar manner, when setting up software the software has to be set up and so does the alert notification system. Each time a system detects a condition matching search criteria, it will notify the operator.

For example, many companies used surveillance systems that detect movement in their office past working hours. To do so the system needs to be set up to detect movement past closing time and should movement be detected, a concerned service will be alerted.

Nowadays a variety of software is available to create a wholesome video content analysis system.

OpenCV (*Open Source Computer Vision Library*). It is a software library under the BSD license, and is therefore free for academic and commercial use. It operates on *C ++,* C, *Python, MATLAB* and *Java* interfaces and is supported by *Windows, Linux, Mac OS, iOS* and *Android* [11].

The library consists of open source code computer vision programs and machine learning on optimized *C/C++*.

OpenCV was created to establish a common infrastructure of computer vision applications and to accelerate using machine perception in commercial products. It is praised for its computing efficiency and raised interest in creating real-time applications.

Due to the inclusion of the open computing language *OpenCL,* the *OpenCV* library may benefit from hardware acceleration of a basis heterogeneous computing platform.

Globally accepted *OpenCV* has over 47 thousand user community users and the download number supersedes 14 million.

The library consists of over 2500 optimized algorithms, includes a complete set of classic and modern computer vision and machine learning algorithms. These algorithms can be used to:

- recognize faces,
- identify objects,
- classify human behavior in video,
- track camera movement,
- track moving objects,
- extracting object 3D models,
- creating point 3d clouds in stereo cameras,
- merge images to get high resolution scenery images,
- searching for image matches from an image database,

- removing «red eyes» in shots with flash,
- track eye movement, recognizing scenes and markers of augmented reality etc.

The library is widely used by companies, research groups and governmental organizations.

Along such renowned companies as *Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda, Toyota*, multiple startups are amongst *OpenCV* library users.

Implementation of *OpenCV* varies from merging several street view shots, detecting territory breaches in Israel, controlling mining operation in China, detecting accidents in public pools In Europe to launching interactive art centers in Spain and New York, providing clearance on Turkish airport landing stripes, monitoring product labels in factories all around the world and facial recognition in Japan.

Not so long ago *Microsoft* updated artificial intelligence tools used by companies to improve object and face recognition, as well as image classification. [20]. All updates are a part of the *application programming interface* (*API)* set called *Cognitive Services*, which allows developers to integrate smart functionality into their products with no prior AI operating knowledge required.

Using a paid tool called *Custom Vision Service* companies will be able to teach their their image classification systems various tasks – such as identifying bird species or distinguishing different kinds of fruit – without there being a need to create their own AI models. Models created using *Custom Vision Service* can be exported from *Microsoft* cloud services and launched on smart phones.

Face API now supports a database of 1 million faces, which enhances its face recognition abilities.

Finally, *Bing Entity Search* became available to the public. The service uses the database of the *Microsoft* search engine. It permits searching for famous people, places and items and receiving relevant information on them.

Concluding let us remark the security and video monitoring platform for self-service devices *ATMeye.iQ* that receives a lot of attention in this book.

Therefore, following is a review of several successful program solutions of video content analysis by other developers.

2.5.1. *AXIS Video Motion Detection applications* (motion detection and video recording)

Video content analysis applications *AXIS* allow conducting proactive monitoring by assisting security personnel in detecting and preventing illegal actions. These applications are able, for example, to detect offenders and automatically inform a security officer or to play an audio recording via loudspeaker. The applications are scalable, decrease storage throughput and use by only recording only relevant video.

For example, *AXIS Video Motion Detection 4* is a free video content analysis application installed on most *AXIS* network cameras and video coders. The application sends an alert message after detecting movement of objects classified as «moving», such as people or cars, in a particular area. In addition, the application allows diminishing storage space traffic and consumption as video is recorded only when movement is detected.

The application works in various lighting, both indoors and outside. It is especially valuable in places with small movement intensity – such as in office corridors, parking lots, unsupervised shopping venue areas past working hours.

Another example is the open platform *AXIS Camera Application Platform* (*ACAP*) supported by most *AXIS* cameras. The platform allows adding analysis tools and other applications depending on specific security and business needs. The applications are fully prepared to be used on Axis devices to analyze live video and video recordings. Examples of use include applications of line breach detection, counting people, identifying a car number plate etc.

2.5.2. Video content analysis on the *Axxon Next* platform

The company *AxxonSoft* is a famous developer of smart integrated security and video surveillance systems. Platforms *AxxonSoft VMS* and *PSIM* service over 240 *Safe* City civilian surveillance projects and security systems in retail networks, banks, airports, seaports, industrial facilities etc. all over the world.

Axxon Next is the next generation of open platform video surveillance management system (*VMS*) software. Surveillance systems based on *Axxon Next* are scaleable without limitations on the number of video servers, operating stations and video cameras.

The *Axxon Next* system includes video detectors of motion, background change, quality loss, left items, line breaches in a set direction, moving and stopping, object appearance etc.

Axxon platforms may connect to the camera archive to review, export and copy video onto an SD card.

The tool set for received video and audio information allow *Axxon Next* to detect various situations defined by the user. As a reaction to a detector trigger, one or several actions from the list may take place:

1. start video recording (with audio);
2. send an SMS message to one or several phone numbers;
3. send an e-mail to one or several addresses;
4. play an audio track;
5. activate a periphery device (relay);
6. initiate hazard processing mode;
7. turn on a PTZ camera in a specific direction.

When operating properly the system requires minimal personnel action while key events are all stored in the video archive and the system operator notified only in emergency cases.

The system supports over 6000 models of IP devices (including 1500 kinds of IP cameras with proprietary protocol integration and 4500 compatible ONVIF devices) as well as remote mobile access and a web interface.

Axxon Next includes a developed system of video imagery analysis that includes the following detection devices:

- motion detection (records any movement in the frame);
- background change detection (is triggered in cases of camera moving attempts);
- video quality loss recognition (is triggered when the image quality becomes lower due to blur, staining, lens blinding etc.);
- left item detection (is triggered when an unsupervised, such as a bag, box, purse etc, appears in the frame and remains still for some time);
- line breach in a set direction detection (is triggered when a moving object crosses a virtual line in a set direction, available in the platform paid version);
- movement initiation detection (records movement in a specific area);
- movement stopping detection (is triggered when a moving abject freezes in a specific area and remains still for a moment);
- object appearance detection (triggered when any object appears in a specific area);

- object disappearance detection (is triggered when an object leaves a set area or if an object within an area leaves the frame).

In additional to video detection *Axxon Next* has two audio detectors:

- noise detection (is triggered when a sound volume reaches a certain parameter);
- silence detection (is triggered when the microphone receives no signal).

Incoming video and audio data analysis tools allows *Axxon Next* detect various scenarios.

Axxon Next video detection tools are configured in the visual mode. After configuring areas, lines and other operation parameters and pressing ***Apply***, the user immediately sees the tool operating the video stream in real time in a specific area of the interface.

- object appearance detection (triggered when any object appears in a specific area);
- object disappearance detection (is triggered when an object leaves a set area or if an object within an area leaves the frame).

2.5.3. *Milestone Xprotect* software

The company *Milestone* is known [14] for its newest video management tool for enhancing IP video productivity.

Milestone XProtect is the leading *VMS* platform developed by *Milestone* that allows combining multiple IP camera and NVR models for universal scaling.

To access *Milestone XProtect VMS* on the go one may use the *iPhone Milestone apps* granting full control from various websites of IP cameras in use. Thus, there is no longer a need to monitor the events at an operator's station.

Milestone XProtect Mobile consists of three smartphone apps compatible with *iPhone, Android* and *iPad*. These apps permit remote monitoring, playing video recordings, sending pictures via e-mail and MMS and search the video archive.

Milestone VMS allows integrating a large set of *Milestone* settings, as well other hardware and software solutions.

Milestone video content analysis software has expansive functionality and in easily installed in any surveillance systems.

2.5.4. Video content analysis on the *VideoNet* platform

The Russian developer «SKAIROS» (Saint Petersburg) introduces the integrated security platform *VideoNet* in 1996.

The system has standard functions such as motion detection, distanced object detection, object counting, reacting to changes in an object's movement direction, line breach detection etc.The latest version *VideoNet 9 Prime* allows creating a professional network video surveillance system. The solution manages up to 16 IP cameras, record video, applying different settings and scenarios monitor the real time status from a computer, tablet or smartphone anywhere in the world.

VideoNet video analysis algorithms detect hazards and minimize false alarms. Only necessary information is collected, thus preserving video archive space. Metadata creation for the video stream facilitates search quires.

VideoNet incorporates smart detectors that automatically detect suspicious or dangerous events. Each camera may be set up with several detector combinations, different areas of interest with different settings for each such area.

As so, a network detector of object type identification may distinguish the following object types: human, car, bus, motorcycle, bike, dog, train, and plane.

The network detector of object type identification belongs to the category of detectors whose algorithm is based on convolutional neural networks [23].

Should the system detect an object of one of the categories, the video window will frame the object and present details on its type and queue number.

It is also possible to setup an automatic reaction for a specific trigger, for example, to start recording audio and video, send an SMS or an e-mail, save a shot, initiate encryption, start an application, play an audio recording, create a report, deploy an investigation squad and much more.

2.5.5. Cloud video content analysis based on *VisionLabs LUNA*

In December of 2017 the Russian company «KROK» launched its video content analysis cloud system developed on the *VisionLabs LUNA* platform in the format of a supervised service .It is accessible through the «KROK» cloud and can be used should the client lack their own IT infrastructure.

The service allows real time extracting data crucial to the business from the video stream and transferring it to other corporate services or to the very business user in shape of various reports. The solution not only highlights and detects different object types (people, transport, areas, etc.) but also is able to track their behavioral patterns and predict criteria crucial for the business. All data that the system processes is stored in the secured «KROK» cloud.

2.5.6. *WINanalyze motion tracking & analysis software*

WINanalyze is a software family developed by *Mikromak* to track movement [12] allowing movement in video analysis. *WINanalyze* can automatically track a virtually unlimited number of objects in video files and present results in different forms (graphically in a system of coordinates, speed and acceleration diagram, etc.). It supports import and export from most common applications.

WINanalyze became the first app to analyze automatically movement,able to track objects with no markers. Using icon recognition methods, it allows tracking object parts throughout the entire video sequence with no human interaction. Nowadays,

WINanalyze Tracker is widely used in movement analysis applications, such as human walk analytics, investigating body traumas, crash tests, video capturing, *General Motion Capture* etc.

WINanalyze is able to perform the following tasks:

- automatic tracking of unlimited number of objects;
- video playback (for example in *AVI, Raw* and other formats);
- linear and precise 3D calibration;
- motion analysis outputs & object trajectory tracing;
- angles and distances data collection;
- assessment of speed and acceleration (*first and second derivatives over time*);
- detecting an object's center of mass.

2.5.7. Software and hardware solutions by the company «Synezis»

The resident company of the Belarusian Park of High technologies «Synezis» developed a compact video content analysis tool for smart video processing and security monitoring.

Usually, a universal coder and video content analysis work in combination, which allows benefiting from both of those separately. In order to reduce the network channel load the video content analysis should be implemented on the side of the customer. Some types of video content analysis, such as face recognition require operating an uncompressed video stream to ensure maximum accuracy. For these reasons, a cloud infrastructure cannot be used effectively on the stage of initial processing without specialized equipment on the side of the customer.

Nonetheless, a cloud infrastructure can be effectively used in video surveillance system scaling:

- storing video and content analysis metadata;
- connecting new points of surveillance (such as new retail stores);
- applying new functionality of metadata analysis and archive searches;
- servicing a large number of customers.

It is clear, that enhancing content analysis quality should diminish the load on the network channels and cloud storage. Therefore, if content analysis accuracy indicators are available, the service provider and the customer may easily calculate the economic benefits of putting it into use.

For example, 300 high-resolution (1.2Mp) cameras are used to monitor the perimeter of a large solar thermal unit. In normal weather conditions, the data transfer rate is 1.8Gbit/s. In special conditions, such as at night, the stream may almost double up to 3,5Gbit/s. The usage of a standard motion detector allows diminishing the amount of video date up to 80% with no general changes of lighting or weather conditions between all cameras.

«Synezis» video content analysis device complies with the international *ONVIF* standard and certified with i-*LIDS*. It can be integrated into an existing security surveillance system without changing any equipment (cables, cameras, recorders, security panel monitors).

Hardware and software by «Synezis» allow:

- counting the number of customers entering a shop or a section;
- detect the number of people and their waiting time in a queue;
- monitor teller activity;
- classify objects based on color to distinguish «customer/personnel» groups;
- classify objects by size to distinguish «child/adult» groups.

2.5.8. Biometrics offers new directions for banking service development

Incorporating biometric identification technologies in client self-servicing has undoubtedly become one of the leading trends in the field of banking technologies. Image recognition technologies organically combine with video security tasks, as many self-servicing devices are initially equipped with cameras. Implementation potential of image recognition for financial institutions grows as technical characteristics of cameras are constantly enhancing and biometric technologies prosper. As for a successful bank, the sought after functionality of today becomes the new standard of tomorrow.

That is why BS/2 and VisionLabs, who are among the leaders of face recognition technologies on the market, have announced their strategic partnership in developing a line of identification services for banks in 2017. The services will be based on the complex self-service device VMS *ATMeye.iQ* and the *LUNA* platform.

«Currently banks collect immense amounts of their clients' personal data, including photo taking when signing contracts, accepting a loan or a new bank card. The images can be later used to effectively distinguish between loyal clients and blacklisted individuals. Should a suspicious person approach a self-service device and caught in the frame of a standard ATM portrait camera, the *ATMeye.iQ* system is able to notify security personnel or launch an alternative scenario (deny service, bank card capture etc.) Such integration allows us offering our clients a perfect solution for efficient video security, as this subject is in high demand», - comments the associate director of BS/2 Tomas Augucevičius.

2.6. Applying video content analysis in practice

Such tasks as protecting a head office, bank branches and 24/7 ATMs, exposing scams, monitoring the cash counting process, marketing analysis and advertisement efficiency control are among bank routine priorities.

In retail, there is also a high demand for security systems. Their task is to create «safe zones» in their venues using video surveillance and to constantly monitor product in stock and other property, as well as to prevent theft in shops. The commercial segment is followed by the market of infrastructure facilities.

Finance also makes significant input in the demand for video surveillance, as they require thorough bank branch and ATM security. New video surveillance integration possibilities emerge in the hospitality business, where resident security should be ensured by monitoring hotel corridors and halls.

The paragraph 3.2.1 lists the scenarios of implementing facial recognition in self-service devices using the *ATMeye.iQ* system.

Nowadays, due to high competition in retail, video content analysis becomes more relevant than ever. An important factor of its use in retail is that most current tasks do not require high precision of result data. As examples, we may mention marketing analysis of customer preferences with the variables being the customer count, product review time, client stream routs and intensity in the venue.

The Russian company «Videointellekt» specializes in developing security algorithms for crowded locations. They developed a software to identify customer behavior in shops. The system allows identifying a customer's intent to make a purchase and their interest for the product, as well as detecting theft intention.

It is common for video content analysis system descriptions to have the probability of a false alarm (i.e. *false alarm ratio*), which is normally between 1 and 5%.

Therefore, a stream of 100 thousand people and the minimal false alarm ratio would still produce a score of 1 thousand false cases. In other words, on average, 125 people an hour – 2 people each minute.

In practice this means, that the control service will be paralyzed by tracking false alarms within a couple of hours. The psychological aspect should also be accounted for. A security officer who receives large amounts of false alarms will become less attentive to real important cases. Quality operation requires the probability should be diminished several times.

The «Videointellekt» system analyzes the scenes with the false alarm ration of 0.3-0.6 cases per 24 hours, which is a thousandth of a percent.

Implementing convolutional neural networks in video surveillance opens up vast perspectives and grand potential for the use of this technology – from retail to «smart city» solutions. It is a huge step for situational analytics development in the field of security, such that leads from assumptions based on mathematical analysis of geometry and color characteristics of pixels to image recognition.

ATMeye.iQ

The solution for video surveillance and fraud prevention

ATMeye.iQ
.iQ Family Product

ATMeye.iQ is a comprehensive solution to improve the security level of self-service devices. It includes a video surveillance system with facial recognition functionality, as well as the sensors that react to any unlawful actions against terminals.

Chapter 3. Software solutions for modern video surveillance systems

«I understand HOW; I do not understand WHY…»

G. Orwell «1984»

Fraudsters develop various kinds of attacks aimed at ATMs and aim at different hardware elements depending on their motives.

Contrary to physical attacks aimed at different ATM elements, logical attacks aim at all ATMs of a network at the same time. This means that ATMs suffer from software threats, similarly to personal computers. Meanwhile, the potential reward for a wrongdoer is much higher.

A demonstration of logical attacks aimed at ATMs was on display during the Black Hat conference in Las Vegas in 2010. Since then, any fraud resulting in a reset ATM dispensing cash to wrongdoers got the name of «jackpotting».

The prevailing types of logical attacks are:

- *Jackpotting* – malware makes the machine distribute cash without creating a transaction. Known methods include direct malware installment onto an ATM hard drive through network access, using a computer disk or a flash drive (USB);
- *Black Box* – an external computer is connected to an ATM and later used to manipulate the machine to distribute cash;
- *Host Spoofing / Man-in-the-Middle* – tampering with server responses or extracting critical information from the network.

As ATM bandwidth increases, so does the possibility of attacks. According to the analytic company FICO, the number of hacked

ATM card readers in 2016 has increased by 30% in the US alone compaIn 2016, 25 588 ATM attacks were recorded, which is 26% more than the year before. This problem is not exclusive to the US: 10 European countries have also reported logical ATM attacks in 2016. The same year the number of hacked payment cards in the US grew by 70%. [17].

The explanation for the increase of technological attacks in recent years is the fact that scammers come up with new ways of penetrating internal processes when:

- employees or third party change their level of access;
- dishonest insiders grant access to intruders by the means of fishing or social engineering;
- personnel who operates cash benefit from process blind spots by stealing money from cassettes or while cash is in transit.

That is why financial institutions should create a layer of protection of their ATM environment to prevent all types of fraudulence.

In their first report on opposing fraud for 2018, the *European Association for Secure Transactions (EAST)* estimated the current state of affairs based on recent crime data from 18 countries that make part of the *Single Euro Payments Area (SEPA*), and 4 other countries that do not belong to SEPA [19]. Seeking to diminish risks of such kinds of attacks Europol with support from EAST EGAF published a «*Guidance and recommendations regarding logical attacks on ATMs»*.ring to 2015.

3.1. A brief overview of information security tools from lead developers

The Swiss ATM security solution supplier *TMD Security GmbH* in collaboration with the British ATM monitoring software provider *S3 Technologies Ltd*. developed a software tool *TMS ATM* for ATM security. The solution protects against logical attacks that use malware, and black boxes that serve the ATM jackpotting purpose.

TMS ATM detects unauthorized changes in profile management software or hardware, blocks unapproved USB port use, controls Windows ATM access and assists system operator to manage safely system access and BIOS passwords.

Diebold Nixdorf

BS/2
Penki kontinentai group

The company BS/2 is the official representative of the leader in banking technologies and the world's largest ATM manufacturer – Diebold Nixdorf. For over 25 years, the companies deploy technologic solutions for banking and retail around the world.

3.1.1. Solutions by *Diebold Nixdorf*

Vynamic Security complex software solution (previously known as *Terminal Security Suite*) by *Diebold Nixdorf* is a multivendor software for protection from logical and other types of attack. Consisting of four modules (*Access Protection, Intrusion Protection, Hard Disk Encryption* and *Fraud Protection*), the solution offers real-time ATM protection based on the principle of total restriction of any process or action execution.

The declared principle of software operation lies in launching only permitted and repeatedly checked processes – and no other. This so-called *whitelisting* principle seeks to protect computers of self-service devices from unauthorized use of external devices (flash drives, hard drives and other carries that may potentially contain malware).

Moreover, *Vynamic Security* imposes a set of rules based on *sandboxing technology*, when a software of specific purpose is provided a fixed set of resources and its access is regulated.

Vynamic Security also ensures that any unauthorized changes could not be made to the unique ATM «ecosystem» with its specific set of technical equipment and applications. Should an attempt to replace a hard drive that contains confidential information be noticed, the hard drive becomes spoiled (process carried out by the Hard Disc Encryption module), while one of several alert scenarios may initiate.

A module worth mentioning separately is the *Fraud Detection* module that allows tracking standard scenario deviations in programs, processes or users with the use of Big Data technology and machine learning.

As previously mentioned in the section 2.3, cases of user behavior, template deviation may trigger planned security measures to be carried out.

Diebold Nixdorf AllConnect Services deployment granted financial institutions and retailers resources and technologies necessary for a physical distribution channel to become equal to their digital equivalents in flexibility, integration possibilities and overall efficiency.

To reach this goal, *Diebold Nixdorf AllConnect Services* incorporates *IoT* infrastructure as the basis of all end service nodes that intuitively analyze data for trend prognosis and decision-making.

By applying a new multivendor software solution *ProFlex4* banks are able to modernize and optimize user interfaces of their ATMs with the latest web technology.

Furthermore, such new service concepts as withdrawing cash using a smartphone instead of a bank card or personalized user interface are also easy to provide with the help of *ProFlex4*.

Therefore, banks are able to add independently existing applications or unique user interfaces to their ATMs, or establish completely new applications. Contacting a Diebold Nixdorf or third party IT service center is also an option to outsource application creation and deployment. It is worth noting that *ATMeye.iQ* is fully compatible with ProFlex4.

3.1.2. Solutions by Ingenico Group

Ingenico Group is one of few specialized point of sale (POS) equipment manufacturers who successfully implements mainstream biometric identification.

Implementing biometric technology allowed the company to realize a series of projects with the goal of attracting new client groups to benefit from banking services.

These projects are aiming at facilitating vendor transition to multichannel sales through a wide range of smart terminals, payment services and mobile solutions that cover online shops, web and mobile channels on a global scale.

One of the latest Ingenico Group products is Axium – an open Android POS platform for retail digitalization developed according to consumer needs. It grants full access to a merchant cloud ecosystem based on Android and Ingenico Telium Tetra open operating systems.

Visually the new Axium device resembles a tablet with a cash register. Moreover, its possibilities not limited by teller services – the solution supports all Android applications.

3.1.3. Solutions by *Gemalto*

Businesses that set out to expend cloud service use often face obstacles. While trying to efficiently manage cloud identifiers and cloud access, the enterprises have to ensure user convenience and regulation compliance.

By providing automated authentication based on cloud calculations and expansive lifecycle management, the authentication management platform by Gemalto facilitate cloud operation in complex environments, reducing administration expenses and establishing a strong basis for scaling in both cloud and local Public Key Infrastructure (PKI) environments.

Based on their multifactor authentication service, *Gemalto* is offering *SafeNet Trusted Access (Identity-as-a-Service)* – an intuitive service that simplifies cloud access management with the use of single sign-on (SSO) and scenarios. *SafeNet Trusted Access* combines the SSO convenience with expanded access security leading to simplified cloud identity data management, elimination of IT personnel and user security and cloud event monitoring through a unified interface.

Gemalto Mobile Protector biometric authentication solution supports both fingerprint and face recognition using simple API interfaces convenient to developers who integrate this functionality into specialized software [18]. The solution ensures safe use of biometric data by removing the need for its storage in a data center or a server. The logged data is stored on the user's mobile device to guarantee its safety.

Gemalto presents an optimized ready-to-use UI that can be tested by banks to see how the biometric project fits their already-established UI. As a universal digital banking service package, *Gemalto Mobile Protector* is easy to introduce into the bank security lifecycle. It can also be complemented by additional products, such as *Gemalto Confirm Authentication Server* (*CAS*).

3.1.4. Solutions by SSC

Established in 2004 and currently part of the Penki kontinentai company group, the Lithuanian enterprise *Skaitmeninio sertifikavimo centras* (*SSC*) became the first company in the country to provide digital signature qualified certificate creation and other related services.

The company provides a variety of PKI technology-based services and it received state accreditation to provide qualified certificates of digital signature on the European market. Currently certificates are provided to citizens of 21 countries.

SSC is also involved in developing digital signature software, providing timestamp services and other PKI-based services, as well as developing and deploying a unified federative authentication system that would not depend on the algorithms used by CSPs.

The authentication of document content reveals any modifications after signing. In its turn, personal authentication confirms whether the individual signing a PDF document is an impostor.

Justa is a line of digital signature software products for the financial and governmental segment that includes:

- Justa WEB ID (federative authentication);
- Justa PDF Sign (PDF signature);
- *Justa Smart Forms* («smart» PDF forms).

Justa PDF Sign is a PDF file signing software that can be distributed as a service whereas special distributions can be integrated into any established IT system. *Justa PDF Sign* is the only PDF digital signature software that can be distributed as a service and personalized for each corporate user certificate or certificate package.

Signatures created using Justa PDF can be checked using the standard and free *Adobe Reader 7.0+*. Signed PDF documents provide document data integrity and allow checking who signed them. Signature check is a certificate used for digital signature that remains valid and cannot be cancelled.

PDF Sign is available in several languages including English, Russian, French, Lithuanian, Dutch etc.

The company *SSC* participates in the eTen project of the *Billing for Rent* consortium sponsored by the European Committee. Along with such companies as *Google, Apple, Microsoft, Symantec* and other is part of the international forum *CA/Browser. Annually* participates in one of the most important events on information security – the RSA conference in the US.

SSC service compliance with international standards is confirmed by the independed audit company *TÜV Informationstechnik GmbH* (*TÜViT*).

3.2. Facial recognition as a practical method of authentication and ensuring security in the banking sector

Biometric recognition is one of the most actively developing technological directions. In banking, where the issue of safe and secure client authentication is very relevant, this «biometric trend», in particular, facial recognition, has become quite in demand.

Many financial organizations already use the following means of biometric methods for personal identification, authorization and verification:

- Facial recognition using a digital image of a human face. The current image of a person is compared to the image registered in the database. This identification requires image capturing tools.
- Palm / finger vein recognition is digitizing the image of veins to consequently compare it with new input data. This method requires a specialized scanner.
- Fingerprint recognition based on a unique image of a finger. This method also requires a specialized scanner.
- Voice recognition implies encoding a voice signal into a digital format to be further verified with the database. This method requires an audio recording device.
- Iris recognition digitizes the image of the retina to confirm matches in the database. The method requires a special device to capture the retina image.

Due to face recognition functionality, clients can be provided with additional layer of securing operations on one hand, and automate a number of processes on the other.

Biometric data collection and storage are a key stage of smart banking development.

The technologies listed above can be used by bank personnel to ensure secure access, as well as on their workplaces (two-factor authentication for internal bank system access), and banking processes aimed at clients. Among those processes are:

- scoring and deciding on issuing a loan;
- online banking authentication;
- semi-automatic filling out of forms with personal information;
- additional authentication when registering on a mobile device;
- identifying fraud and looking for criminals (includes banks sharing biometric data of blacklisted persons).

Most technologies can also rely on processes connected to SSD (ATMs, kiosks, etc.). In this regard, the easiest system to deploy is facial recognition, as it does not require specialized equipment and default cameras on self-service devices should suffice. Moreover, the face recognition system can be integrated with the video monitoring system to provide multi-layer security of the very device and the transactions done through it. Below we describe possible scenarios of using facial recognition.

3.2.1. Face recognition implementation scenarios on SSD as part of the *ATMeye.iQ security system*

Through the course of years the developer of SSD security solutions, BS/2, have creates such products as *ATMeye.iQ* (a solution for enhancing SSD security), *Brancheye.iQ* (IoT expansion), *CashManagement.iQ* (for CIT authentication), *ServiceDesk.iQ* (for maintenance personnel authentication), *SmartSafe.iQ* (video control over cashier actions and self-service monitoring).

The *ATMeye.iQ* solution, discussed in the Chapter 4 of this book, provides an option of facial recognition that usually does not require installing additional equipment. Having the frontal camera of the video surveillance system is already enough.

The system is fully compatible with SSD by most global manufacturers (*Diebold Nixdorf, NCR, Hyosung etc.*).

On one hand, collecting biometric data allows to create a «whitelist» for engineers who maintain the devices and CIT officials. In case of unauthorized access, the notification will submit a report with photo and video data for the security personnel to take action.

In addition, the facial recognnition system is able to determine the gender and the approximate age of people. Thus, this function can be used to provide targeted advertising. This function is available not only on SSDs, but also virtually anywhere where clients stop for a service (fill out blanks, contracts, printing tickets etc.).

Two-factor authentication on SSDs can be used to provide an additional level of security when operating bankcards, as well as to prevent illegitimate card usage.

Within the framework of a single scenario of card issuance, the card owner should obligatorily have a picture taken at the bank branch. The photo should be taken in a well-lit room in HD. In the future when operating an SSD, after the PIN code is input, the face of the card owner should be confirmed. *ATMeye.iQ* chooses the best shot from a sequence and compares it with the image in the database. The resemblance should be set at about 90-95%.

It is important to note that the higher the level of resemblance, the lower is the list oof unlawful card use. Facial image is linked to the card number.

If the facial recognition system confirms a match with the picture from the database, the person should be granted access to the following ATM screen. Otherwise, the card will be kept with a message on the screen providing a phone number to contact for consulting. In case of an error, an ID is provided in a bank branch and the card is returned.

Blacklisting to prevent fraud

People who were involved in fraud and whose photos have been stored in the bank database after being collected by a security system can be blacklisted.

Preventing fraud on SSDs can be combined with two-factor authentication or be done without it. People who use bankcards of others can be blacklisted. As the face is detected, a best shot is chosen and is stored in the database all from a real time video stream using *ATMeye.iQ*.

Chapter 4. *«.iQ»* software product family

«Freedom is the freedom to say that two plus two make four. If that is granted, all else follows».

G. Orwell «1984»

In November 2015, *Diebold* Nixdorf ordered *Forrester Consulting* to conduct a comprehensive analysis of issues that retail banks encounter when managing their ATM fleet security. A survey of business and IT owners from around the world showed that banks would only benefit from partnerships with third party providers of security management services of ATM fleets.

The results show that an average financial institution manages from 100 to 499 ATMs on a regional level, which shows a number of significant risks that should be protected against.

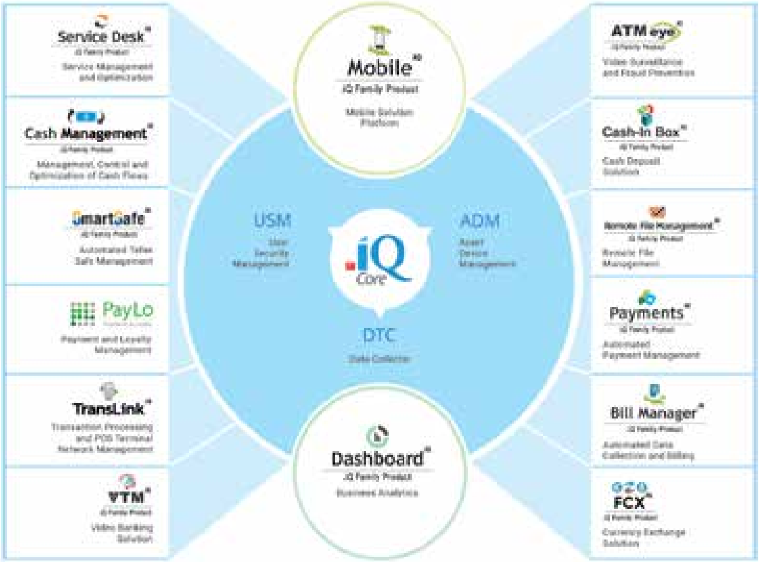


Fig. 3 «.iQ» product family

4.1. *ATMeye.iQ* – self-service security and video monitoring platform

ATMeye.iQ is a software-hardware solution for real-time self-service device security that ensures a timely response to unlawful actions taking place.

It is a part of the *«.iQ»* business management product family developed for the financial industry.

The expansion of ATM networks is a serious challenge for banks. This trend requires financial organizations to invest additionally in new personnel and significantly complicates work processes related to device technical maintenance. That is why the possibility of protecting up to 14 000 ATMs within a single network by using *ATMeye.iQ* is so relevant [15].

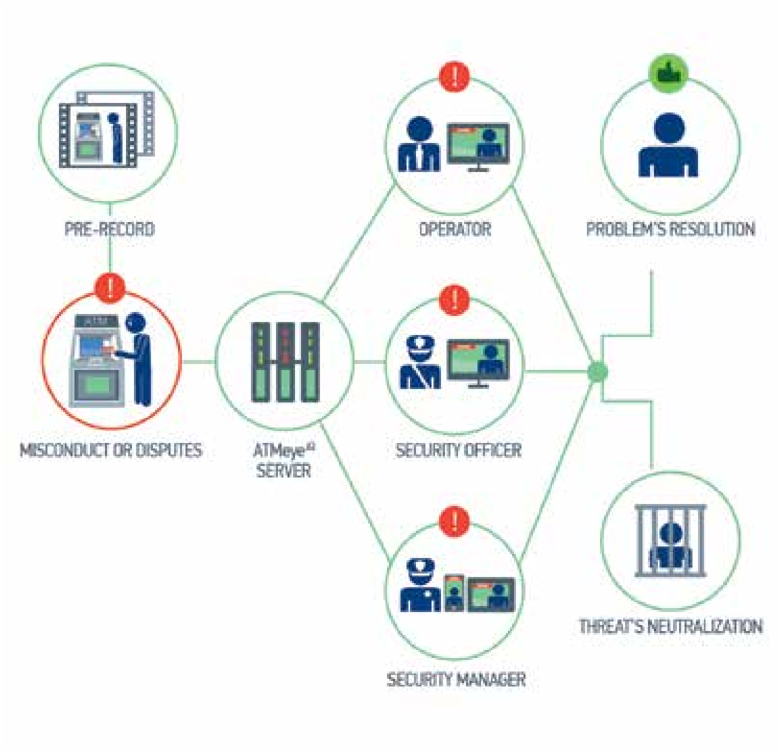


Fig. 4. ATMeye.iQ solution scheme

The *ATMeye.iQ* platform appeared on the market in 2001 and it includes a video surveillance system as well as a number of detectors that react to any unlawful actions aimed at terminals.

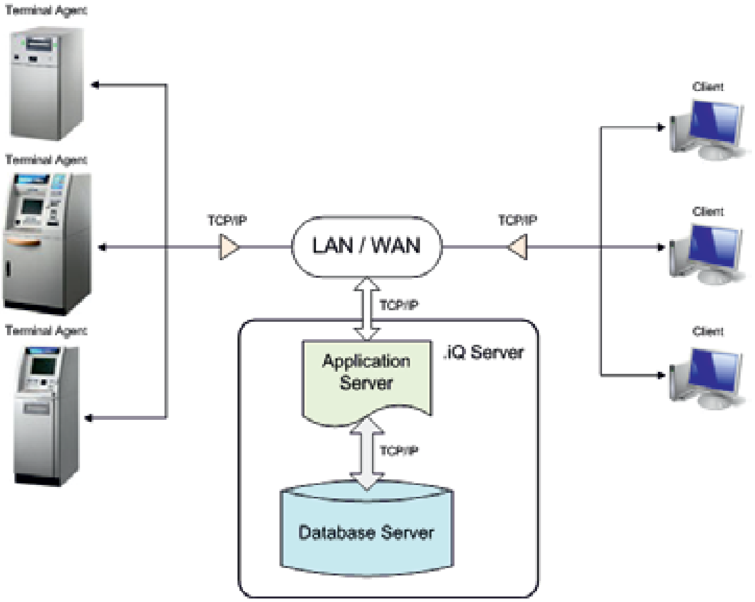


Fig. 5. The conceptual scheme of the ATMeye.iQ platform and its modules

The platform is meant for organizations that use self-service terminal networks who are interested in enhancing the security of client servicing and self-service devices.

The *ATMeye.iQ* system core functions based on client-to-server architecture, the system uses a single central database and provides access to necessary data to different user groups, ensuring information consistency. It consists of the following key components: *USM.iQ* (user and security management module), *ADM.iQ* (resource management module) and *DTC.iQ* (data collection module).

DTC.iQ is the module of *«.iQ»* integration with external systems. It provides both operational information on external system events (card introduction, cash dispensing, receiving a new service request etc.) and information on external system configuration changes (introduction of a new device, device setup address change, external company status change etc.).

To extract information from external systems the *DTC.iQ* module used configured agents that are easily adaptable to new tasks. All information provided by agents is saved in the database both in the original and ready-to-user form. External device information is filtered and the *«.iQ»* input operates only ready-to-use data.

The cryptographic protocol used in system is *TLS 1.2*. It ensures a safe connection between the client and the server that is compliant with the *PA-DSS* certificate.

In 2017, a global shift to *ATMeye.iQ* version 2.0 has been made. This version was first to provide the feature of distributing software license directly from the server. This allowed company clients and partners with a convenient tool of active software license management that enhances interaction.

Let us note that *ATMeye.iQ* is integrated with *ProView*, a software package by *Diebold Nixdorf* used for banking self-service device network monitoring, remote management, diagnostics and report generation. The software *«.iQ» Client* was developed to connect with the server (from the operator/administrator workplace) which provides a number of management interfaces. These interfaces (called *Perspectives*) are provided to *«.iQ»* users depending on their role and access. A set program consists of several modules. Amongst those – *Basic* and additional *Client* modules chosen by clients according to license type.

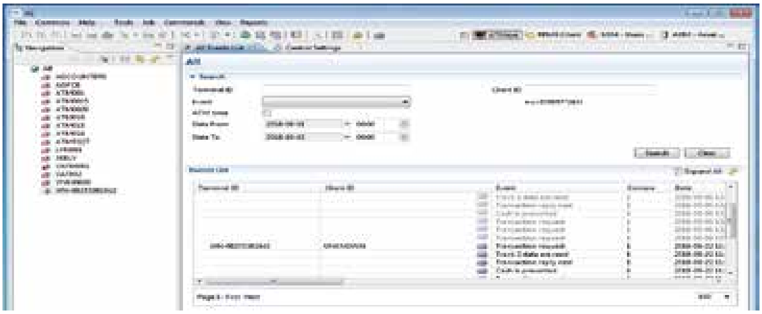


Fig. 6. Main dashboard for device and transaction monitoring in «.iQ» Client

The main dashboard of the *«.iQ»* Client application (Fig. 6) shows the central system access: the screen on the left shows the tree of current devices (ATMs etc.). Should the tree be divided in groups, the only those groups will be accessible that are allowed by the system administrator.

System modules are chosen when transitioning between USM, ADM and RFM modules and during program installation. The two basis modules are *Users* and *Assets.*

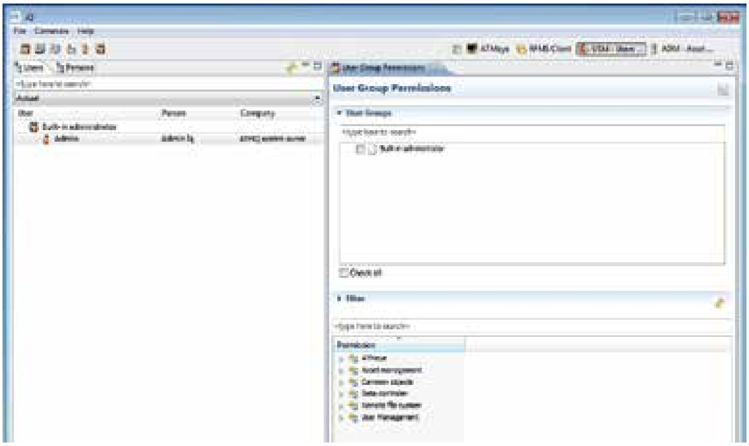


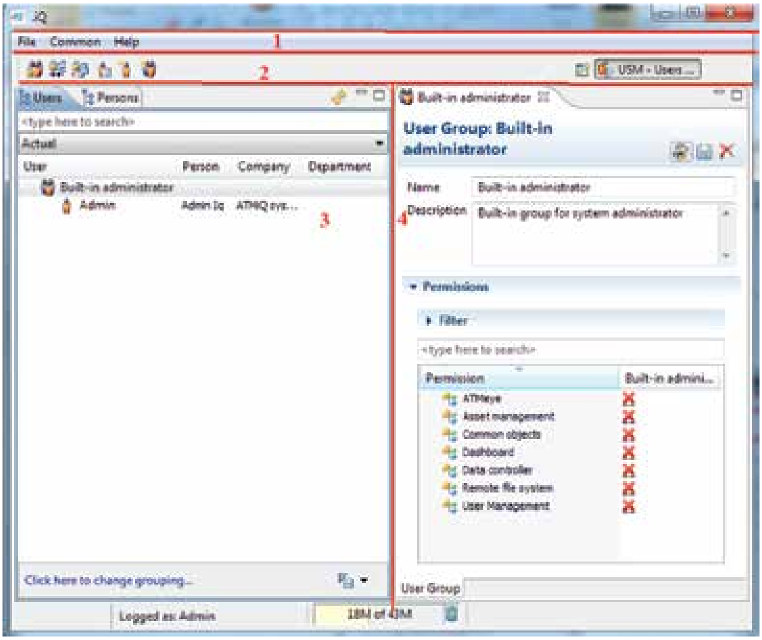
Fig. 7. Main dashboard of the USM .iQ module

***USM (Users and Security Management)* module for user and user access management.** This module provides an interface of the *«.iQ»* system user management through the «Users» perspective. This perspective functionality includes:

- user creation, deletion and pausing;
- introduction and editing of necessary information attributes about the user;
- management of *«.iQ»* module and component access rights.

User Group Permissions consists of two parts:

- list of user goups
- list of accessible object (forms, modules, functions) groups.



User	Person	Company	Department
Built-in administrator			
Admin	Admin Iq	ATMIQ sys...	

User Group: Built-in administrator

Name: Built-in administrator

Description: Built-in group for system administrator

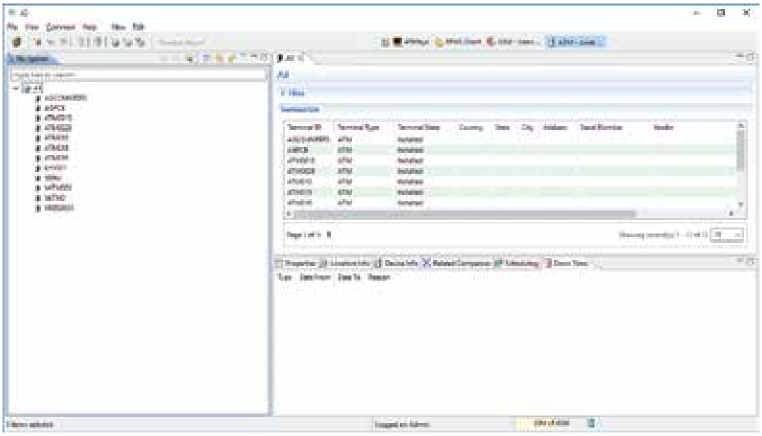
Permission	Built-in admini...
ATMeye	✗
Asset management	✗
Common objects	✗
Dashboard	✗
Data controller	✗
Remote file system	✗
User Management	✗

Fig. 8. User interface of the USM .iQ module USM .iQ.

Besides creating new user groups and their properties management, there is also a feature of cloning a specific user group. This means creating and identical user group (under a different name) with the same access rights.

The system records all user actions. When necessary, user action history can be viewed and browsed under certain search criteria (username, event, time). Reports can also be generated.

ADM **(*Asset Device Management*) module** is meant for introducing devices (ATMs, kiosks, POS terminals etc.) into the system. Beside the name of the device, the system also tracks the device address and location, its assets etc.



Terminal ID	Terminal Type	Terminal State	Country	State	City	Address	Serial Number	Vendor
[illegible]	ATM	Installed						
[illegible]	ATM	Installed						
[illegible]	ATM	Installed						
[illegible]	ATM	Installed						
[illegible]	ATM	Installed						
[illegible]	ATM	Installed						
[illegible]	ATM	Installed						

Fig. 9. Main dashboard of the ADM iQ perspective.

The list of its functions includes:

- terminal equipment adding, deletion, recovery and cloning;
- adding and editing data attributes about the terminal equipment;
- device hierarchical grouping;
- equipment access right management for users and user groups;
- adding, editing and removing a network, device list review.

In addition, one may review and edit the list of time periods, add and remove «out of order» statuses etc.

Remote File Management **module** ensures data transfer between a remote self-service device and the operator's workplace over a secured communication channel.

The *RFM.iQ* system includes:

- terminal module *RFM.iQ*, installed on each device;
- server module *RFM.iQ* with *User Security Manager* and *Asset Device Management* features.
- *«.iQ» Client* module**.**

The local panel of the system is divided in three parts:

- device tree;
- system files window for choosing a disc on a chosen device;
- system files window displaying files on the computer that is running the *«.iQ»* client.

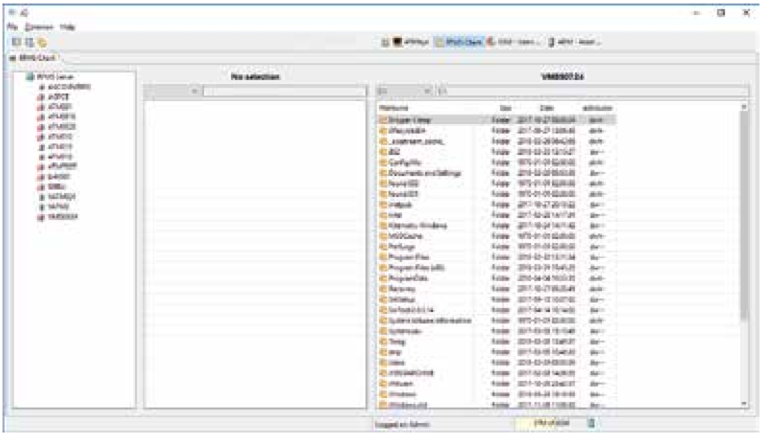


Fig. 10. The main dashboard of the RFM .iQ perspective

By using RFM technology, financial institutions are able to improve efficiency of data exchange between self-service devices management, maintenance service workflow optimization, as well as reduce expenses. After device deployment, the system completes all data exchange tasks remotely according to the set timeline. Meanwhile data can be transferred onto either a single device or the entire device network.

The *RFM.iQ* system can be used for ATMs, electronic safes, information and payment terminals, as well as computerized work environments.

The *RFM.iQ* solution is provided both as a separate product and as part of the *ATMeye.iQ* system package.

Some key features of the module include remote document search, data transfer from a remote terminal onto a database or a local file system, task planning (by dates, days of week, folders, file templates).

RFM.iQ allows certain remote real-time operations of the device: data status control, software launch, data transfer etc.

Additional features include transferring electronic journals, photos, videos and other information from a remote terminal onto the data collection server, as well as storage and archiving of said data.

ATMeye Module (***Desktop***) ensures remote management and monitoring of the ATM subsystem on the *ATMeye.iQ* platform.

To open the ATMeye module one needs to click the ***Open Module*** button on the tool panel and choose «Other» and then double-click «ATMeye» from the list.

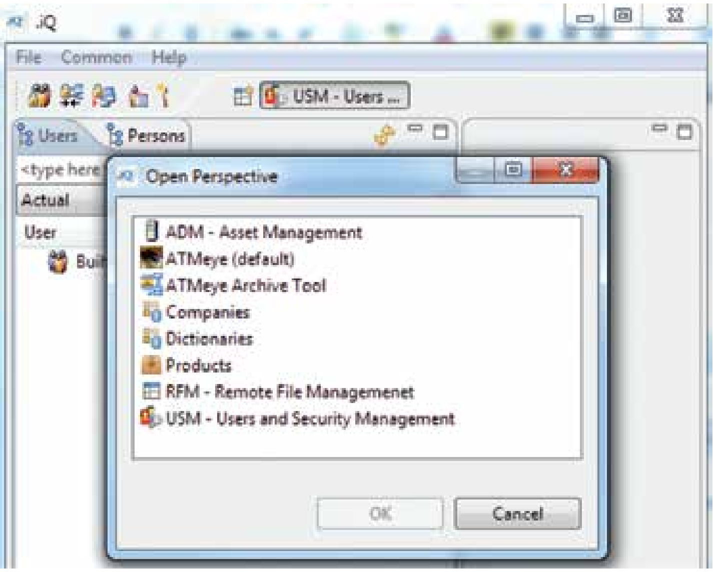


Fig. 11. Opening the ATMeye Module (Desktop) perspective.

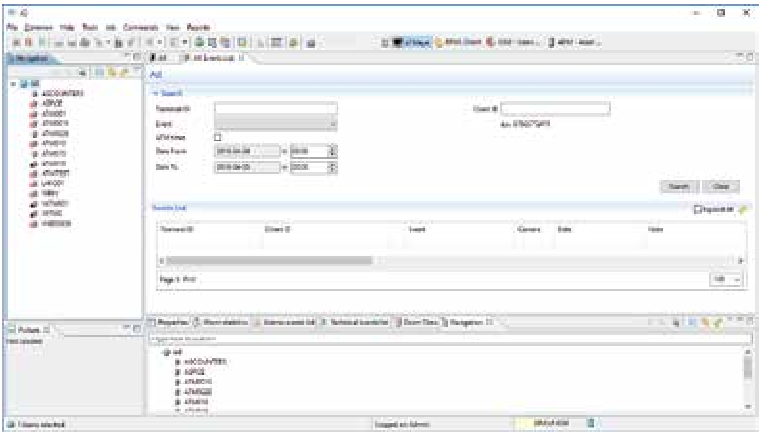


Fig. 12. ATMeye Module (Desktop) main dashboard.

4.1.1. Main features of *ATMeye.iQ*

ATMeye.iQ video monitoring system is a multivendor solution fully compatible with the latest versions of application operating systems for all models and types of self-service devices from main ATM manufacturers, which ensures equipment and bank client data protection.

ATMeye.iQ conducts real-time video monitoring and centralized management of self-service terminal networks.

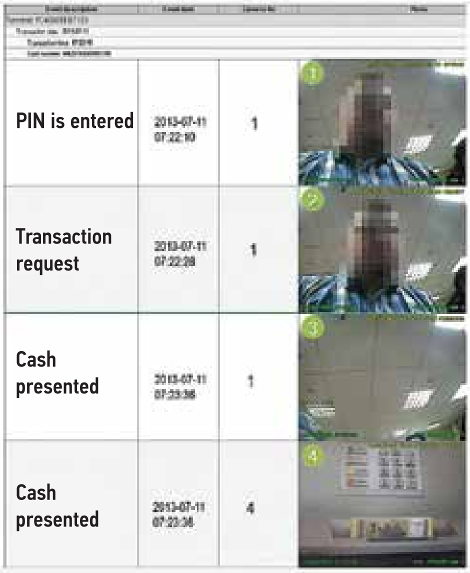
Fig. 13. ATMeye.iQ is a secure solution for self-service device protection

The solution detects and processes harmful actions (vandalism, fraud) and informs the system operator with providing all details necessary.

ATMeye.iQ functions include:

Photo and video recording of events. During any operation, the device records photos and videos of the user and user actions, which helps generating a detailed report about every transaction.

Photo and video recording after detector triggers. Triggering any detector forces the cameras to start recording, creating a complete record of the events.



Event description	Event time	Camera No	Photo
PIN is entered	2013-07-11 07:22:10	1	1
Transaction request	2013-07-11 07:22:28	1	2
Cash presented	2013-07-11 07:23:36	1	3
Cash presented	2013-07-11 07:23:36	4	4

In the report of transaction generated by ATMeye.iQ, photo 3 and 4 shows, that customer leaves ATM withaut taking money.

Fig. 14. ATMeye.iQ workflow during ATM operation.

Video recording before and after the event. Photos and video is captured before and after all transactions or other device operations to recreate the events;

Viewing video in real time. Instant access is granted to video streams and pictures of any devices;

Checking camera work order. On request, *multiple snapshot* can be made to confirm that cameras operate properly;

Support of different camera types. The system supports up to 4 internal USB and analog cameras, as well as 12 external IP cameras which can monitor the operation of the card reader and the dispenser, as well as the safe area, client face and ATM vicinity;

Camera mode setup. Cameras can work in different modes (day/night) according to a pre-set schedule, allowing to make good quality shots at all times.

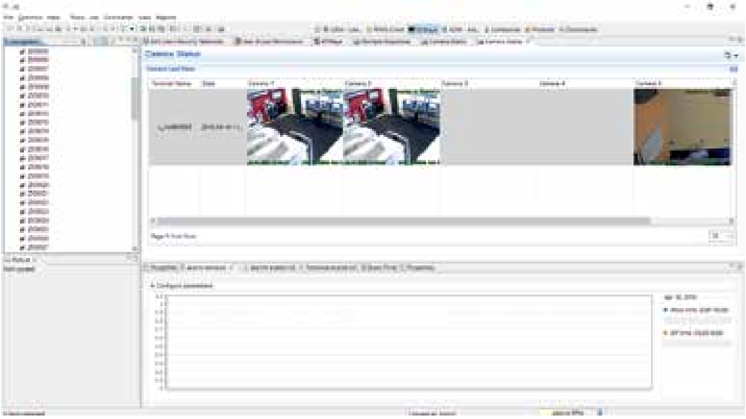


Fig. 15. Cameras status in the ATMeye.iQ system.

Camera makes from 4 to 10 shots in a second. Which provides for accurate real-time data at a low bandwidth. This allows verifying time data (tags) of events up to a tenth of a second. Photo data is stored on the ATM. Data export occurs regularly according to bank settings [29].

Supported photo/video resolution: 320x240, 352x288, 640x480, HD.

The average shot file size with 320x240 resolution (in *JPEG*) is 10-20Kb.

The system is expanded with various detectors and notification scenarios:

Great possibilities of threat detection. The device functionality can be expanded with the use of external detectors (physical strike, vibration, incline, smoke, temperature etc.) as well as with specialized devices of fraud prevention;

Real-time hazard monitoring. A responsible bank employee is immediately informed about any of the detector triggers;

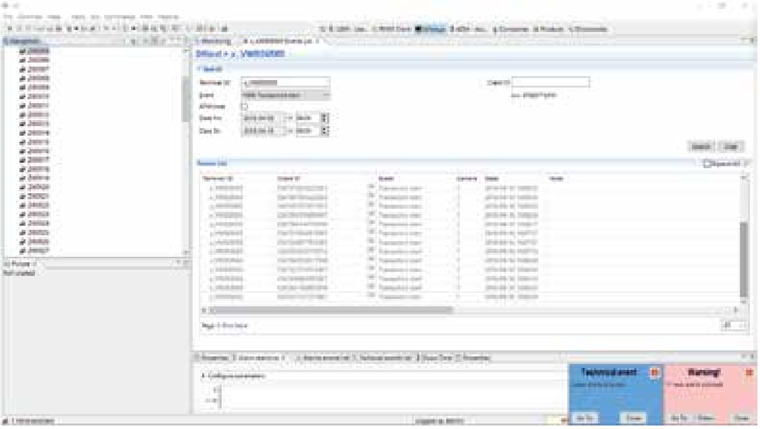


Fig. 16. Notifications about detector triggers.

Covered camera alarm. In case of camera coverage or disconnection, responsible bank employees will automatically get alarm notifications on this event to be able to prevent any possible consequences.

Mobile notifications. Emergency messages about any abnormal situations can be sent in real-time to mobile devices via the Mobile *ATMeye.iQ* application.

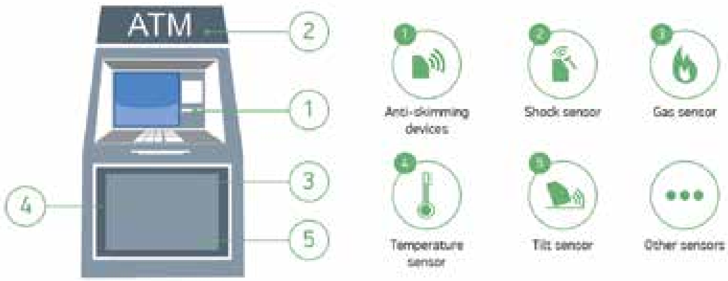


Fig. 17. Detector types for ATMeye.iQ.

Additional functions:

Remote camera configuration. The system allows configuring cameras remotely, choosing the best operation mode (day or night), setting brightness, contrast and other parameters.

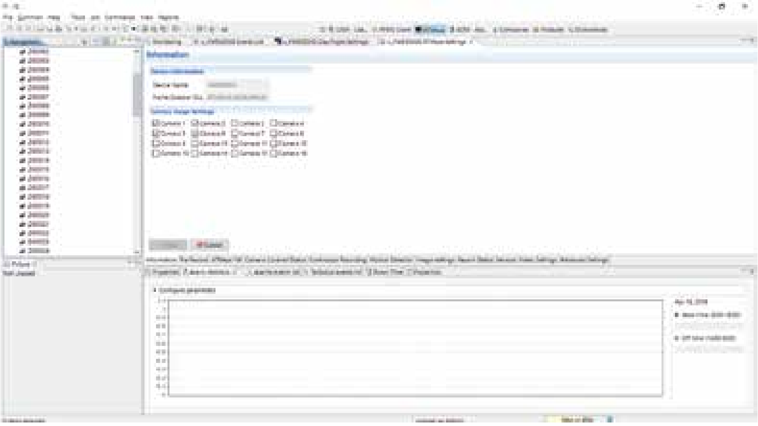


Fig. 18. Remote camera configuration interface.

Encryption and data security. The system allows configuring cameras remotely, choosing the best operation mode (day or night), setting brightness, contrast and other parameters.

Centralized file management. The integration with *RFM.iQ* allows the system to transfer files (photos, videos, logs, software updates) between remote self-service devices, administrator workstation and collection server.

Remote software updates. Software updates is done remotely and centrally for the whole SSD network without the need to physically come to the SDD.

Task scheduling. The system allows transferring data and archiving photos, videos and logs according to the predefined schedules or preset algorithms.

Monitoring functionality:

One common interface for all events. The system gathers all information from the whole fleet of SSD and presents it in an intuitive way on a single screen, allowing you to get a complete picture of the operation of the entire system.

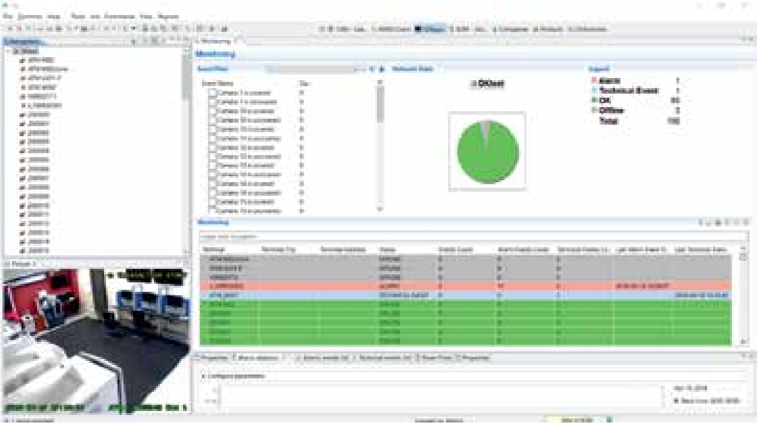


Fig. 19. Common monitoring in ATMeye.iQ

Proactive notifications system. Event notifications are sent to the responsible staff through various delivery channels thus avoiding routine scanning of the incoming events stream and reducing the reaction time.

Easy search capabilities. A system operator is able to find a particular self-service device using the device tree sort and look up capabilities in an easy way. An exact self-service device can be find by branch, city, region and other parameters.

Particular transaction search. The system allows searches based on the card number, type of event, date of transaction and other parameters.

SSD status check. A system operator receives information about current operating status of all self-service devices in the network. It allows him to react faster in case of technical problems

Card capture or service denial. A self-service device if needed might capture a blacklisted bankcard. Responsible staff will be notified about such an event by an instant message

Masking the bank card number. Key card information (card number) is masked in *ATMeye.iQ* system, ensuring the complete safety of the client's personal data.

Automated workplaces. The *ATMeye.iQ* system has several preconfigured type of users with the role based access rights and predefined unique functions allowing to work with and control the whole system while maintaining highest possible security level of operations.

The system administrator has the ability to manage other users' access rights, manage the licenses, import data from other systems, backup and restore information from the archives, monitor network status, schedule network activities, get comprehensive statistical reports.

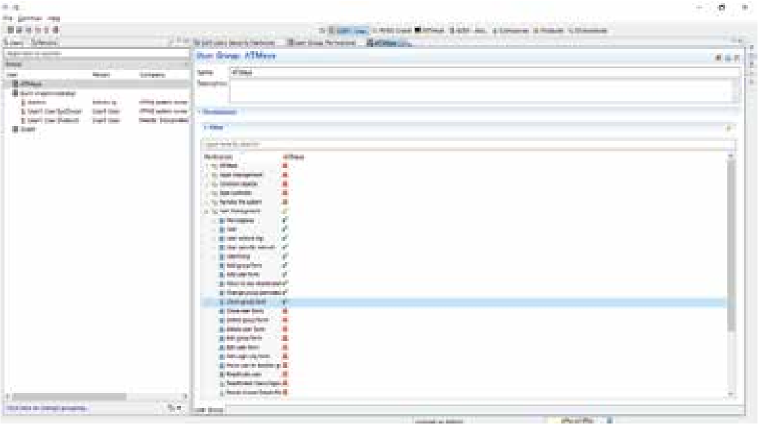


Fig. 20. Role and access rights management interface in ATMeye.iQ.

The system operator has all the necessary rights and functionality to obtain the full list of transactions together with the information about client actions, retained cards or retracted cash as well as operational status for any self-service device. The operator is able to connect to monitoring devices and supervise the venues and areas around a particular ATM. Face recognition technology allows the system to identify people within the frame, and with access to the database, immediately inform the operator about a blacklisted person is sight. Thus, the efficiency of security services is greatly augmented [29].

The security officer can receive instant notifications on all alarm events that is harmful for SSD: hack attacks, vandalism or tampering attempts, and many others. At the same time, the system provides possibilities to generate various reports with statistics of such events for analysis.

ASM.ATMeye.iQ is a result of the *ATMeye.iQ* system and anti-skimming device integration. This innovative anti-fraud solution works with any types of anti-skimmers and logical security solutions (fraud-monitoring systems).

ASM.ATMeye.iQ provides the following functions:

Real time notifications. If an installation of a set of skimming tools is detected, the *ATMeye.iQ* system in real time sends out notifications of this event to the desktop computers or to the mobile devices of bank employees. At the same time, cameras record criminal actions, and photographs and video recordings are sent to the security officers.

Detailed reports. The system provides reports about all attempts to install skimming devices for complete self-service device network on request or based on predefined schedule. It allows analyzing the overall picture of network security.

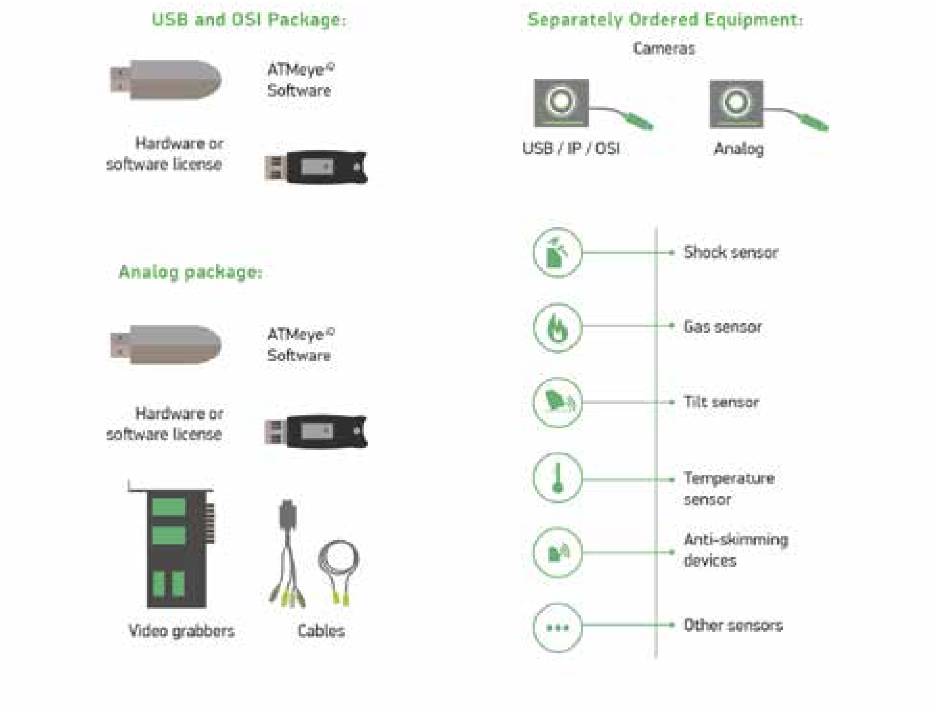


Fig. 21. ATMeye.iQ supply package.

BS/2 provides to its clients:

- The sale and rental of necessary equipment packages that include frame grabbers, various detectors and other components on demand.
- Deployment, support and updates of server and network hardware and software for the *ATMeye.iQ* and *Mobile ATMeye.iQ app*.

In chapter 2.5.8 we already discussed the importance of biometric client identification being one of the key trends in modern banking technology development. Facial recognition solutions are often easily compatible with video security, as most devices have video cameras by default. In this sense the partnership between BS/2 and VisionLabs, one of the leaders in facial recognition technology developer on the world market, in developing a line of services of facial identification based on the *ATMeye.iQ* system and the LUNA platform is quite insightful.

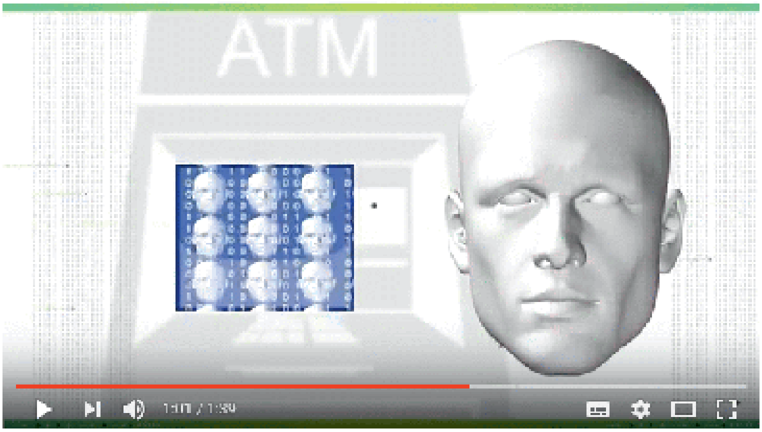


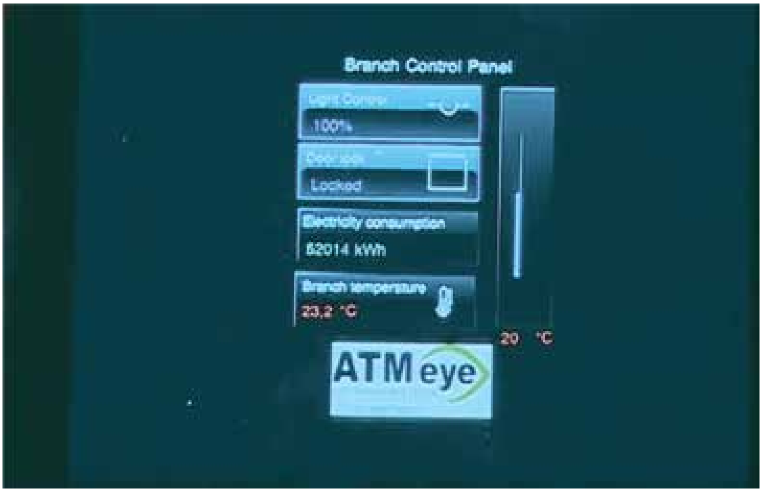
Fig. 22. Applying facial recognition technology in ATMeye.iQ for ATM protection

On one hand, financial institutions aim to demonstrate as many targeted advertisements to the client while providing everyday services, on the other, enable the client with additional services (bill payment feature, exchange currency etc.). For achieving said goal, in 2017 the system *ATMeye.iQ* was integrated with the current *ProFlex 4.0* version by *Diebold Nixdorf*.

4.2. *Brancheye.iQ* for bank branch security and video monitoring

ATMeye.iQ can connect remote financial service channel management with bank affiliate security, real time process control and lifecycle monitoring of heating and conditioning systems, electricity networks etc.

Brancheye.iQ is a solution developed to conduct video monitoring of bank branches, offices and other venues.



Branch Control Panel

Light Control
100%

Door lock
Locked

Electricity consumption
62014 kWh

Branch temperature
23,2 °C

20 °C

Fig. 23. Brancheye.iQ main dashboard

Brancheye.iQ enhances the level of bank branch security (break in prevention, vandalism, system errors), which is also effective in customer service monitoring. It can be integrated with other building management systems. Introducting *Brancheye.iQ* diminishes operational and maintenance expenses.

Main *Brancheye.iQ* functions include:

- flexible system of video and photo event search and archive administering;
- timely reaction to hazard notifications, photo and video stream analysis;
- automatic alerts of suspicious events;
- report analytical data generation;
- user rights and access management;
- easy integration with other IT systems;
- multivendor;
- Operation on all device types – smartphones, tablets, web;
- Remote monitoring;
- Fixing errors in the communication system;
- Time-based system management;
- Support of up to 16 different video information sources;
- Building control system management;
- Smart monitoring of electricity, gas, water, heating and conditioning;
- Facial recognition.

Cash Management.iQ

Automation, optimization, management and control of cash circulation

Cash Management.iQ

.iQ Family Product

Cash Management.iQ is a software designed to solve the problem of efficient cash distribution at cash collection and withdrawal points: ATMs and other self-service devices, as well as in bank vaults and branches, mail and retail networks.

4.3. *Cash Management.iQ* – cash flow optimization, management and control

Cash Management.iQ belongs to the *«.iQ»* product family. It is used for automating processes related to cash distribution between all cash processing points (ATMs, electronic teller, storage vaults, cash settlement units, payment terminals, information kiosks etc.).

This multivendor solution consists of four modules and it ensures that an optimal amount of cash is present throughout the network and while optimizing cash flow processes.

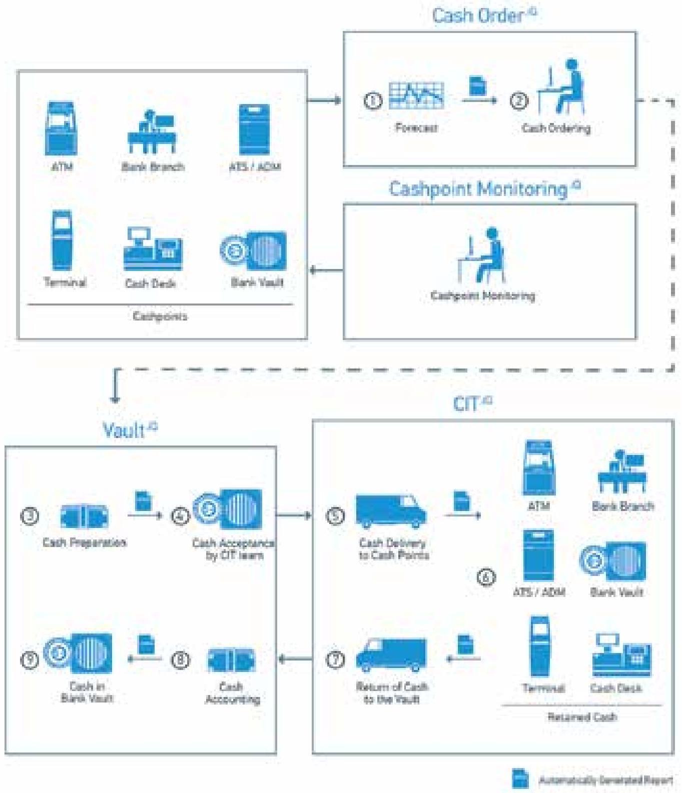


Fig. 24. Cash Management.iQ workflow

Using the *Cash Order.iQ* modules one can create cash flow prognosis, i.e. analyze the need for cash, plan and process CIT orders of SSD and bank branches, as well as set limits for replenishment and remainder amounts – for this task in particular the prognosis mechanism based on neuron networks is used.

Thus, CIT operation timeline for each individual event based on cash flow statistics.

The *Cashpoint Monitoring.iQ* module is used for cash balance monitoring in real time for different cash dispensing points with amount and currency precision. Status can be checked for the whole network or a single device.

The *Vault.iQ* module serves to monitor cash balance in vaults with amount and currency precision. In addition, CIT operation requests for replenishment and cash takeout can be filed. The module ensures control over the vault status at the start and the end of an operational day.

The *CIT.iQ* module is used to form CIT brigades, manage their work, optimize routes, and control cash delivery.

Another business analytics module *Dashboard.iQ* is used to generate reports for every stage of the process. Depending on the business process and enterprise needs the entire *Cash Management.iQ* service packet can be installed, as well as various modules combinations separately.

Thus, the *Cash Management.iQ* solution:

- reduces the expenses for cash distribution support and CIT operation generation and logistic planning;

- controls the state of customer service nodes in the network in real time and processes alert scenarios;
- facilitates cash distribution workflow;
- promptly reacts to demand changes in cash banking.

Other features include quality and compliance control, CIT operation price optimization, cash flow automation and document management, risk management (finding a compromise between security and workflow), business analytics, quality parameter trend analysis.

Implementing the tools described above enhances cashpoint availability and reduces the number of violations.

The *return on investment* (ROI) calculation for each individual case of deployment is done in account with influential factors, such as price of CIT operations, frozen cash, cash insurance etc.

SmartSafe.iQ

Management of Automated Teller Safes

SmartSafe .iQ

.iQ Family Product

The **SmartSafe**.iQ system allows managing ATSs and additional peripheral equipment, as well as automation of operations, related to the cash deposit and withdrawal in the bank offices, mail and other organizations.

4.4. *SmartSafe.iQ* for teller workplace automation and administering

SmartSafe.iQ integrated with a core banking system is used to manage electronic tellers of various manufacturers, as well as administering cashpoint.

Another benefit of *SmartSafe.iQ* is centralized control over all cash operations due to the system's centralized management.

To enhance security by adding video surveillance functionality the system is integrated with *ATMeye.iQ*. The video monitoring solution monitors electronic teller and safe locations and prepares photo reports of separate events (cash operations).

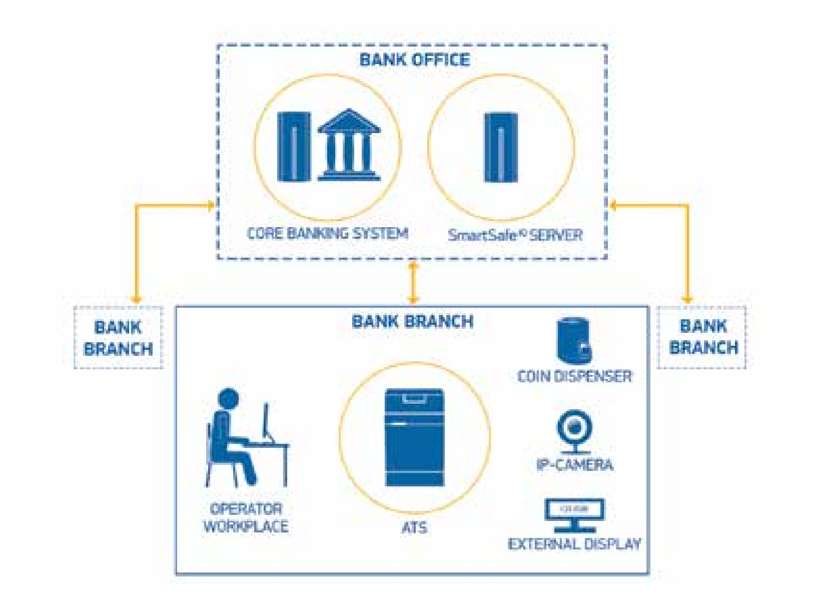


Fig. 25. SmartSafe.iQ scheme

Multivendor solution *SmartSafe.iQ* is compatible with devices by world's largest manufacturers such as *Wincor Nixdorf, CTS, DoCash, Glory, De La Rue, Talaris, Vertera*, and the list is constantly expanded. Integration with a new device on average takes 2-3 months.

SmartSafe.iQ supports simultaneous operation of several world currencies and electronic teller operation of 128 note denominations. Data transmission security is ensured by using the *SSL* protocol. The system is able to internally generate security certificates or have them uploaded from external systems.

For banks that seek to completely automate electronic teller workplaces, *SmartSafe.iQ* also supports coin dispenser operations, letting the device dispense both bills and coins.

The possibility to unite teller workplaces in a common computer network grants personnel the feature of quickly switch between safes and merge them with other periphery (coin dispensers, external displays, cameras, card readers etc.) using the *SSL* protocol, forming a flexible configuration according to one's needs.

An additional module for business analytics generates all necessary *Dashboard.iQ* reports in a user-friendly format.

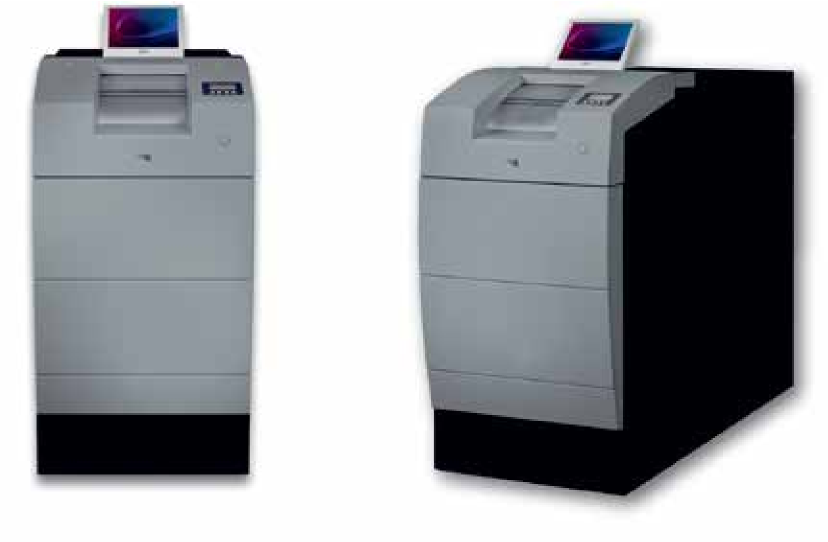


Fig. 26. Electronic teller by Diebold Nixdorf.

SmartSafe.iQ solution is performing the following key tasks:

- automated client cash withdrawal and deposit;
- automated CIT withdrawal and deposit operations;
- procedure control and monitoring during the day;
- real time cash flow monitoring;
- report generation;
- administering software and hardware.

Specifically for banking, the system provides:

- cash withdrawal and deposit;
- payment for goods and services;
- withdrawal from outside bank accounts;
- providing information from external information and billing systems;
- counting stacks of cash;
- exchanging cash for larger/smaller note denominations.

Reports that the system generates may include electronic teller cash flow, CIT operations, time frames, teller cash flow by period, cash leftover, equipment working condition with up to node accuracy etc.

The system supports current teller balance monitoring. The interface and all reports are available in multiple languages.

4.4.1. *Mobile SmartSafe.iQ* - a mobile application for cash operations

Mobile SmartSafe.iQ is a new self-service payment infrastructure based on *automated teller safes* (ATS) and smartphones. Main benefits of *Mobile SmartSafe.iQ* are workflow convenience, enhanced functionality, high level of security, reducing cashpoint price and reducing the workload of bank branch personnel. The end user of *Mobile Smart Safe.iQ* is offered a spectrum of cash services (cash deposit and withdrawal, bank account top up, payments, currency exchange etc.).

The system combines recycler ATM, payment kiosks and currency exchange functionality available through a user-friendly smartphone UI.

Mobile SmartSafe.iQ allows:

- providing the clients with a convenient environment for financial operations via smartphone;
- adding innovative features, such as augmented reality-based user guidelines, cash withdrawal under the obligations of the mobile network provider;
- reducing self-service device installation costs;
- diminish the workload of bank personnel by redirecting a group of clients to self-service terminals;
- reduce electronic teller expenses due to bulk manufacturing;
- reconsidering electronic tellers to become mass servicing devices;
- develop the electronic teller-servicing infrastructure.

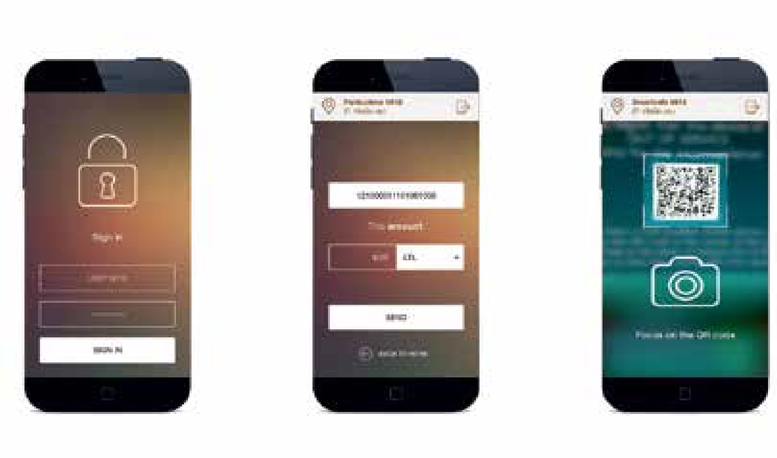


Fig. 27. Examples of Mobile SmartSafe.iQ user interfaces.

Let us define the main features and benefits of *Mobile SmartSafe.iQ*.

Providing the following services to bank clients:

- cash withdrawal and deposit;
- payment for goods and services;
- currency exchange with the possibility to transfer the difference to a bank account;
- *Cash Anywhere* (withdrawal from outside bank accounts);
- providing information from external information and billing systems;
- counting stacks of cash;
- exchanging cash for larger/smaller note denominations.

On the other hand, the application ensures full accountability to bank personnel (the basic lists of reports can be expanded on demand), including safe reports, CIT reports, operation reports for a set timeframe etc.

The application has the following monitoring functions:

- safe network monitoring;
- safe and server availability on the message cannel;
- current safe balance;
- equipment working condition status with single network node accuracy.

Cash-in Box.iQ

Cash Deposit Optimization

Cash-In Box.iQ

.iQ Family Product

The solution **Cash-In Box.iQ** comes in the form of a reliable deposit machine, integrated with the IT system of the bank serving the outlet. It comes equipped with effective and convenient tools for monitoring and controlling the operation of all connected self-service devices.

4.5. *PayLo* payment and loyalty program management solution

Developed and supported by *ASHBURN International* (Penki Kontinentai group), the modular system of payment scenario and loyalty management *PayLo* is a universal tool of payment and loyalty scheme management that can be integrated into tellers or realized on individual payment terminals [28].

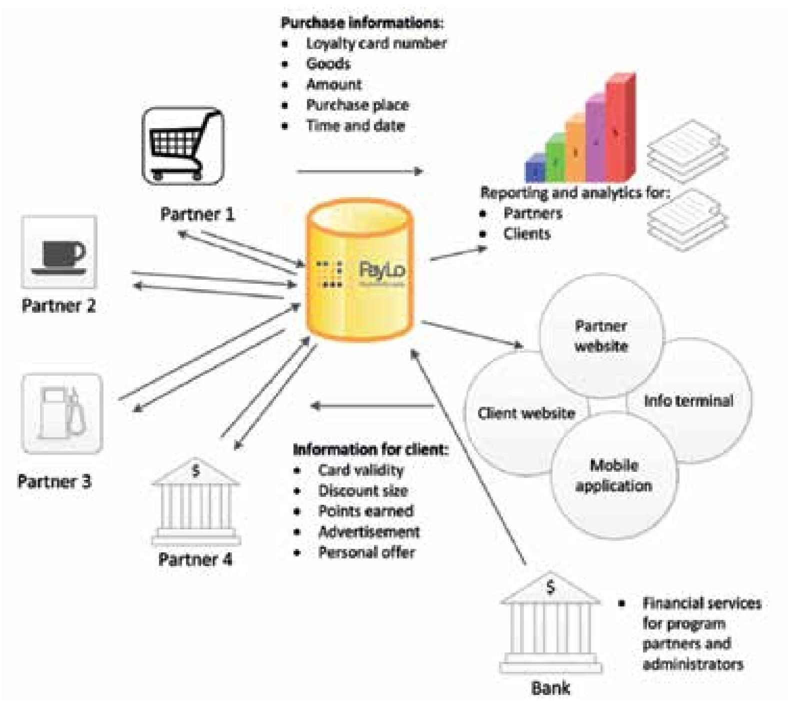


Fig. 28. Software cross-loyalty logical scheme.

The *PayLo* system is oriented at retailer needs in creating loyalty programs. The system includes all stages of loyalty program servicing, starting with the initial identification phase co loyalty scheme composition, loyalty operation processing and analyzing data collected by the software.

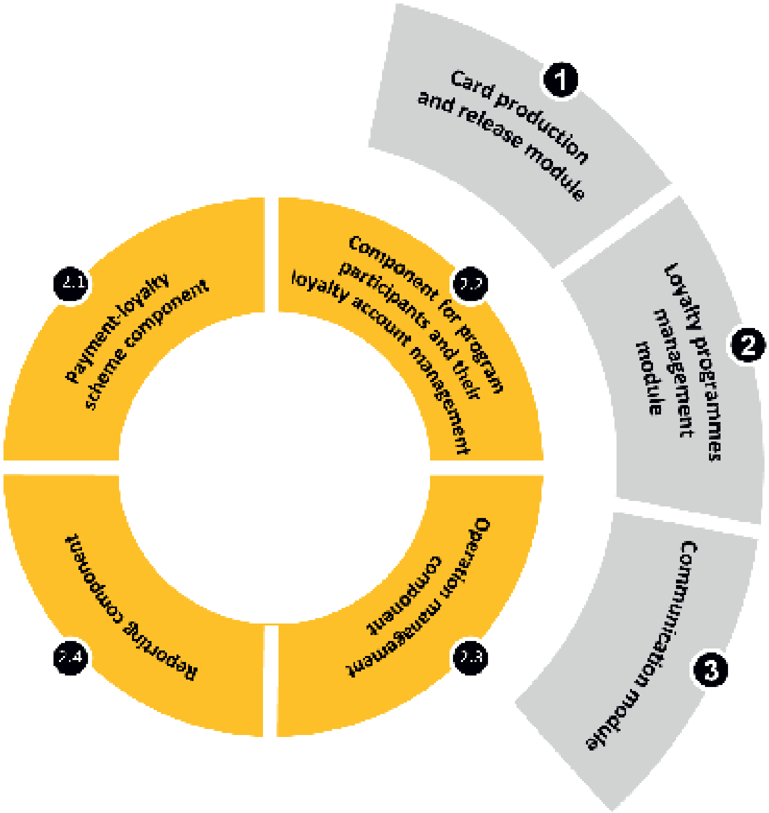

1
Card production and release module
2
Loyalty programmes management module
3
Communication module
2.1
Payment-loyalty scheme component
2.2
Component for program participants and their loyalty account management
2.3
Operation management component
2.4
Reporting component

Fig. 29. Components and modules of the PayLo solution.

Loyalty program participants may manipulate loyalty schemes using an Internet user interface.

Some major *PayLo* user interface features include:

- Introducing and editing loyalty program participants in the system;
- Loyalty scheme setup and editing;
- Accessing information about loyalty scheme components;
- Real time loyalty operation monitoring.

The *PayLo* system is comprised of target components and modules, thus including all stages of loyalty program introduction and servicing (Fig.29).

Payment and loyalty scheme management component is meant for creating and management of loyalty schemes, and realizing loyalty logics.

Program participant and their loyalty account management component is responsible for establishing a database of program participants (natural/legal persons), creating new payment accounts and loyalty accounts, card blocking, setting up credit limits, attributing loyalty schemes etc.

Operation management component ensures the authentication and correction of payment and loyalty operations, payment history collection.

Report component is meant for generating periodic and special reports on different sections using data stored in the system that determine user demographics, activity, behavior etc.

Card manufacturing and issuance module is responsible for card manufacturing and issuance (from form scanning to card distribution).

Purchase management module manages complex payment and loyalty program scenarios that incorporate the provider's range of products and services.

Communications module is a marketing tool for sustaining direct contact with loyalty program participants via SMS and e-mail.

The devices for customer identification in points of sale may vary:

- *EFTPOS terminals*;
- *NFC scanners* (integrated/connected to EFTPOS terminals);
- Cash registers with magnetic stripe card readers;
- Card readers working in combination with mobile devices (smartphones and tablets);
- «Virtual card» identification via mobile device (smartphone or tablet).

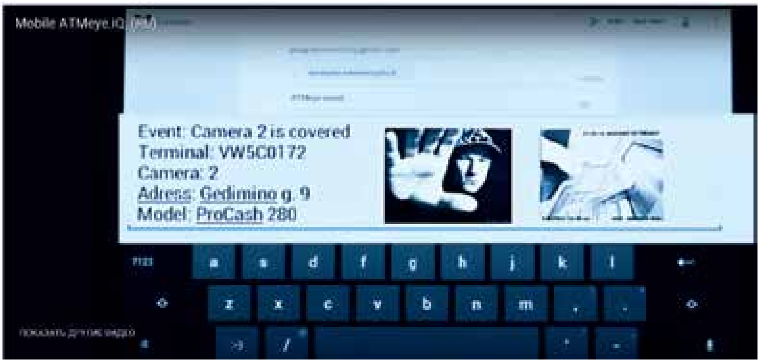
Using *PayLo* allows for many ways of client identification. Among those are payment and loyalty cards (magnet strips, bar code, QR code, and NFC).

Identification can also be performed via «virtual card» on a smartphone or tablet. In addition, *PayLo* provides an opportunity to use facial recognition in some loyalty programs.

ASHBURN International follows the highest-level physical and virtual data security standard *PCI DSS Level 1* that insures the safety of international payment card information. The standard includes strict requirements for technical equipment, software, internal processes of the company, IT architecture, software development process and documentation.

4.6. *Mobile ATMeye .iQ* mobile solutions platform

Mobile ATMeye.iQ is an application for a smartphone or a tablet that allows instant notifications about any suspicious actions on SSDs or in their proximity.



Event: Camera 2 is covered
Terminal: VW5C0172
Camera: 2
Adress: Gedimino g. 9
Model: ProCash 280

Pic. 30. Mobile ATMeye iQ application: notification on camera coverage of a specific terminal.

Mobile ATMeye iQ provides a set of secure tools for taking action in cases of emergency:

- inquiry on current device and component status;
- viewing photos and videos in real time, as well as data before or after the event.

Mobile ATMeye iQ ensures the following functions:

- informing bank personnel via SMS, e-mail or other channels;

- calling security services;
- activating any process that was pre-programmed into the device;
- turning on the alarm;
- stopping and launching an application on an SSD;
- SSD restart.

The application is available in the *Google Play* store via the following link:

https://play.google.com/store/apps/details?id=com.bs2.android.atmeye

VTM.iQ

Solution for videobanking

VTM.iQ
.iQ Family Product

The **VTM**.iQ solution for remote banking effectively in reduces the costs of providing various banking services, improves their quality and strengthens customer loyalty.

4.7. *VTM.iQ* remote banking service solution

VTM.iQ is a completely new technical outlook on financial services that allows completely re-imagining client service.

VTM.iQ is meant for remote banking and payment operations. It includes a unique *CuRie* (*Customer Rich Experience*) technological concept, an electronic cash recycling teller machine *CS 6060* by *Diebold Nixdorf,* as well as software solutions and system integration by *BS/2*.

VTM.iQ is easy to integrate with main banking systems (BIS, distanced service center, «Processing»). It provides an additional client service channel aimed at advanced computer users, allowing them to save time and individually (or with the assistance of a service center agent) perform virtually any banking operations, participate in online video consulting on *VTM.iQ* as well as services offered by the bank, get personalized information on bank services, order and instantly receive a bank card, pay loans etc.

The solution incorporates different means of personal identification, (using a bank card, ID scan, fingerprint). The solution allows signing a bank contract, open an account, file a loan and receive the amount on a bank account, credit card account or in cash, check account balance, top up a bank or card account, withdraw cash from a bank or card account, perform cash payments, pay utilities, phone bills, transportation bills, buy tickets, vouchers, block a card in case of a loss or theft, change a card's PIN number and so on.

VTM.iQ is a symbiosis of several functional systems that provides an interactive approach to customer service.

Security and cash management optimization is ensured with such technologies as HD video (for contacting customer center specialists, fingerprint scanning, signature digitization, contact and contactless card identification, ATM-like servicing and video surveillance integration).

Alongside traditional banking functionality, *VTM.iQ* offers remote customer service consulting. Cooperating with an operator ensures safer, more convenient and professional banking, insurance and administrative servicing, at the same time ensuring greater customer satisfaction.

Personnel workflow is optimized by applying the «center of competence» approach. Using *VTM.iQ* the bank is able to provide quality service disregarding the bank branch location and local personnel. In fact, *VTM.iQ* reduces the need for bank branches in general, and reduces customer service-related expenses.

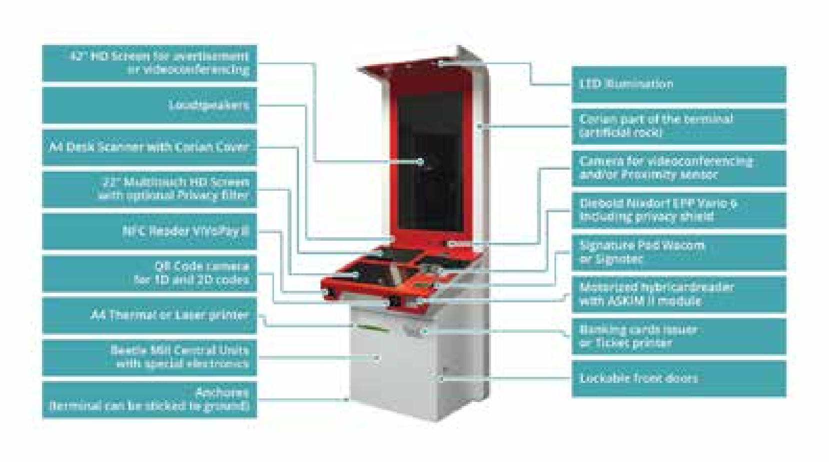

42" HD Screen for avertisement or videoconferencing
Loudspeakers
A4 Desk Scanner with Corian Cover
22" Multitouch HD Screen with optional Privacy filter
NFC Reader ViVoPay II
QR Code camera for 1D and 2D codes
A4 Thermal or Laser printer
Beetle Mill Central Units with special electronics
Anchores (terminal can be sticked to ground)
LED Illumination
Corian part of the terminal (artificial rock)
Camera for videoconferencing and/or Proximity sensor
Diebold Nixdorf EPP Vario 6 including privacy shield
Signature Pad Wacom or Signotec
Motorized hybricardreader with ASKIM II module
Banking cards issuer or Ticket printer
Lockable front doors

Fig. 31. Main components of the VTM.iQ computer appliance.

VTM.iQ is a perfect solution for the current generation. It also ensures convenient and easu service for people with disabilities.

VTM.iQ is enabled with impressive security measures: HD video cameras, security software solutions (*Anti-Skimming, Encrypted PIN-pad*), that cover any potential blind spots in a security system. These measures improve system availability and help gain client trust.

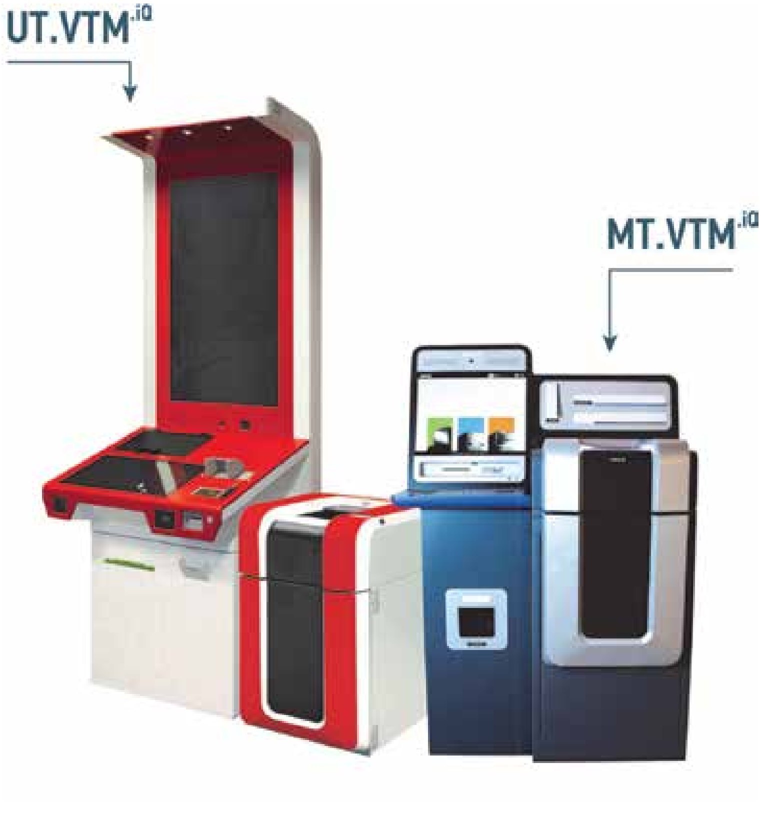


Fig. 32. Multifunctional VTM.iQ terminals.

The coverage of service zones is greatly fortified by setting up the device in bank branches, 24h self-service points, residential and commercial premises, airport and station lounges, hotels and rural areas. *VTM.iQ* is a solution for bringing banks closer to customers.

4.8. Prospects of applying video content analysis of *ATMeye.iQ* in «smart houses» and «smart cities»

A number of megacities are in the process of deploying «smart city» projects that imply creating a hybrid automated system for solving all urban technical tasks. The system should include a complex of software and hardware means (video cameras, resource management devices, emergency voice communications etc.) and organizational means of ensuring security, as well as urban object management.

«Smart» effectiveness in a modern city requires such an environment that would allow facilitating control and accelerate reacting to threats. Risks grow due to global instability, leading to city center protection incorporating detectors that connect via IoT. A modern city is a complicated multi-level structure composed from systems such as transport, telecommunications, electricity and water networks, and other that intertwine and interact.

Therefore, an advanced information system is required to ensure control over all systems, ensuring city infrastructure security, retrieving and archiving information on important events and transmitting this information to any services involved.

Currently a number of large home security companies operate on the market. Among those – *Comcast, Time Warner Cable* and *AT&*T. According to expert opinion, this is caused by security being the dominant direction of a «smart house» industry.

According to the research by *Parks Associates* on the condition of the competitor environment in housing security of the US, «smart house» devices make up 42% of new projects of security system installment in rural housing.

That being said, interactive services and means of remote control make part of 70% of currently installed home security systems. Overall, such services are are used by about 60% of US families who have broadband Internet and receive professional monitoring services. Among home users of broadband networks in the US, 23-25% are equipped with a security system and around 21-23% receive professional monitoring services. Around 50-60% of «smart house» equipment is connected to security and are bought as a part of the entire house system.

In Europe the environment is different due to a large portion of security equipment being out of date, and information and communication projects in this region, as a rule, are oriented at pushing the current technology use limits while trying to modernize the equipment and resolve any system coordination conflicts.

In 2021, largely due to the interest of governments in advanced countries to utilize the latest security technology, the global market for urban security solutions is expected to overcome $20 billion.

The wide range of possibilities of the *ATMeye.iQ* platform, reviewed in this book, provide new possibilities for governmental and civil institutions to effectively detect threats and fight consequences.

4.9. Innovative solutions by *BS/2* as the company's key to success on the global market

It is vital to note that the solution *ATMeye.iQ* first received the *ATMIA* (*ATM Industry Association*) *Best ATM Security Technology* award in *2002*

In 2017 *ATMeye.iQ* was granted a golden medal on the traditional «Product of the year» contest organized by the Lithuanian Confederation of Industrialists. For BS/2 this became their second award, as the first «Product of the year» medal was granted to *ATMeye.iQ* in 2008.

During the annual partner convention of the world's largest banking equipment manufacturer *Diebold Nixdorf*, the company *BS/2* received the prestige *Special Achievement Banking 2017* award for their innovative solution in Azerbaijan. The company set up a fully-automated bank branch for one of the local banks. The solutions incorporated included the computer appliance *VTM.iQ* that runs on an electronic cash recycling teller machine *CS 6060* by *Diebold Nixdorf,* as well as software solutions and system integration by *BS/2*. This computer appliance was the first in its kind not only in Azerbaijan but in the whole CIS region. Since then the solution has also been deployed in Kazakhstan.

In early 2011 the company BS/2 accepted the renown IT process management practices *ITIL V3 (Information Technology Infrastructure Library)* and successfully underwent audit and received the compliance certificate of the *ISO 27001* standard, which proves the high quality of service that the company provides.

During the *Diebold Nixdorf* partner convention that took place in

2014, the company BS/2 was praised for their activity in Georgia with the *Best Service Banking 2013/2014* award. And the *Special Achievement Banking 2013/2014* award was granted to the company for the impressive sales results in Moldova. Another reward, *Special Achievement Banking 2013/2014* was shared with the Azerbaijani company *Komtec* for successful operation in the banking sector of Azerbaijan.

In 2016 the company received the *Best Banking Solution 2016* award from *Diebold Nixdorf* for successfully realizing a project of modernizing and optimizing an ATM fleet of one of the largest banks in the Baltics.

Previously BS/2 has also been awarded *Diebold Nixdorf / Wincor Nixdorf Best Banking Solution 2016, Special Achievement Banking 2013/2014, Best Banking Service 2013/2014,Best Banking Service 2012/2013* and a number of other prestige awards.

BS/2 is a member of the international association *ATMIA* that represents the international banking equipment industry.

The company employs over 300 highly qualified specialists, has 7 affiliates in Azerbaijan, Estonia, Georgia, Kazakhstan, Kyrgyzstan, Latvia and Uzbekistan.

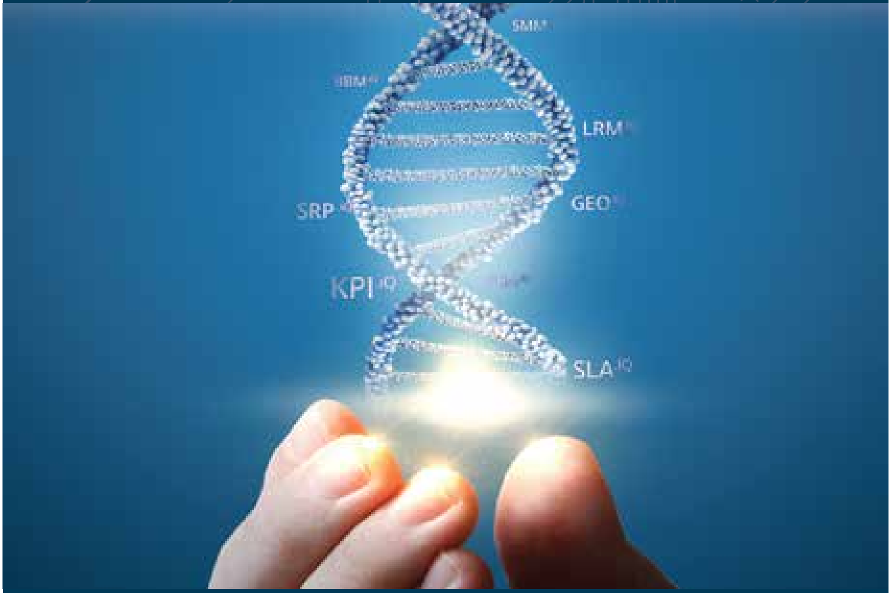
BS/2 operates in 80 countries, with over 800 clients and partners around the world. Those include:

- Banking and financial institutions;
- Retailers;
- Gas stations;

- Postal services;
- Other companies (horse racetracks, pawnshops, casinos etc.)...

ServiceDesk.iQ

Service management and optimization



Service Desk .iQ
.iQ Family Product

Service Desk.iQ is a solution for the service processes automation for banks and retailers equipment, responsible for the opening, distribution, execution and closing of client requests. This tool helps to organize the work of the service company personnel and format the reports.

Annex 1. Deep learning and convolutional neural networks for video security

One of the ways the term «intelligence» can be interpreted is «an ability to make right choices in accordance to set criteria» (for example, the survival instinct in the animal world). Modern computers already possess a certain level of intellect due to programmers who created software. Therefore, computers are able to perform actions that humans find useful (thus, we speak of computers making «right choices»). Nonetheless, the wide range of tasks that humans and animals are able to easily perform are still out of reach for computers. A vivid example is recognizing a person's accent and identifying a person in photos.

In order to make more significant decisions knowledge needs to be set in a form that would be integrated.

Most of the tasks belong to the «artificial intelligence» (AI) category and include a number of perception and control tasks.

Thanks to AI computers are able to learn from their experience and perceive the world in terms of hierarchy – each concept is being determined by its relations to simpler terms. By collecting knowledge from experience this approach avoids the need for an operator to define all formal knowledge a computer needs to perform a task.

But how does one achieve «artificial intelligence» status? The answer lies in using data and examples to establish exploitation knowledge – i.e. machine learning.

Deep learning is a specialized kind of machine learning that teaches computers to do things humans do naturally, that is, learn from an example. Recently deep learning stepped into the foreground of attention due to the precision of results. Previously such results were considered unachievable. In deep learning a computer model learns to qualify objects directly form imagery, texts or audio.

The basis for this approach is the hierarchy of terms that allows forming difficult concepts based on simple ones. Fig. 33 depicts the architecture of a multi-layered module. That is the reason this approach was titled to «deep» learning for AI.

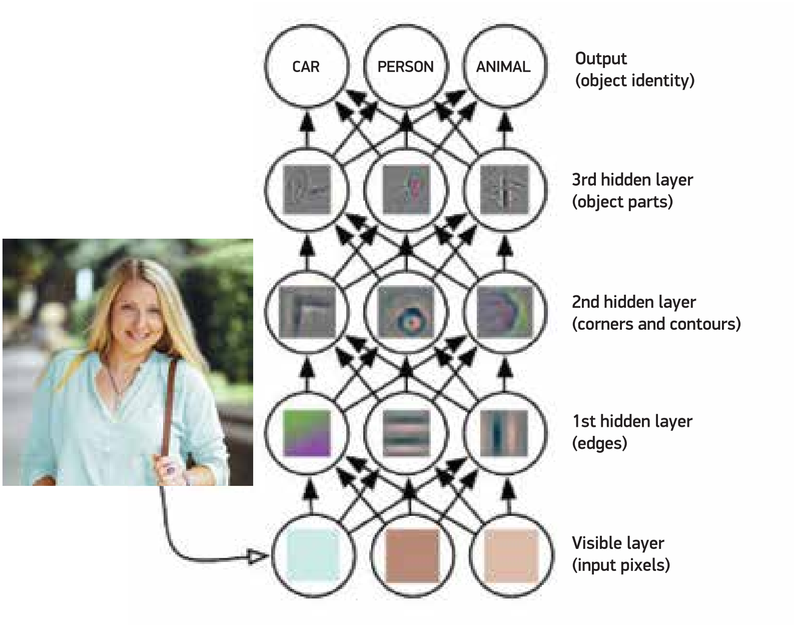


Fig. 33. Deep learning model architecture.

Most deep learning methods use neuron network architecture.

The term «deep» usually refers to the number of hidden neuron network layers. Traditional neuron networks consist from only 2-3 hidden layers, whereas deep networks may have up to 150.

Deep learning models use large amounts of marked data and neuron network architectures that learn the functions **without the need to establish the features of classified objects.**

Convolutional neural networks (CNN) use dimensional correlation by applying a local scheme of neuron connections between different layers. In other words, the input of hidden blocks in the «m» layer are connected to the subset of blocks in the «m-1» layer, the blocks of which have spatially-related receptor fields (Fig. 34).

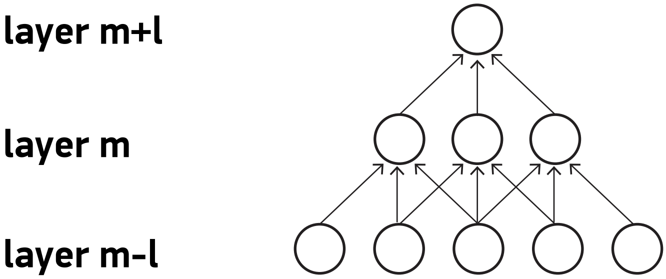


Fig. 34. Neuron networks organized into layers consisting of interconnected nodes.

Such networks may have tens and hundreds of hidden nodes. For the purpose of visualization, let us imagine that the layer «m-1» is the entrance to the eye retina (fig. 24). Here blocks in the «m» layer have receptor fields measuring to the width of 3 in the retina, therefore, they are connected to only 3 proximate neurons on the retina. Blocks in the «m+1» layer have the same connection to the preceding layer. It is considered that the width of their receptor field in relation to the preceding layer is also 3, however, the width of their receptor fields in relation to the entrance is larger, i.e. measuring 5. It is important that each block does not react to the changes outside of its receptor field from the side of the retina. Therefore, such an architecture guarantees that the educated «filters» produce the strongest reaction to the dimensional image input in the system.

At the same time, having such a variety of layers leads to nonlinear «filters» that become more «global» (i.e. react to a bigger area of pixels). For example, the device in the hidden layer «m+1» can code a nonlinear function of the width 5 (in terms of pixel space).

In addition, in CNN each filter is replicated around the entire scope of vision. These replicated blocks use the same parametrization (vector of scales and displacement) and create a map of features.

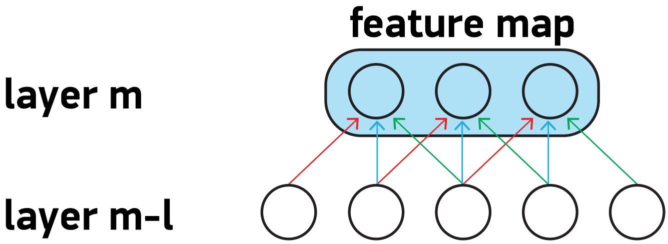


Fig. 35. Feature map of a multilayer neural network.

Fig. 35 depicts three hidden blocks that belong to a single feature map. The scales of the same colors are distinguished by color in such a way to be same in each of them.

CNNs remove the need of defining features for manual classification. CNNs operate by extracting these features directly from imagery.

Specific features do not require preliminary learning - they are defined by the neural network during image processing. Such automatic feature extraction makes deep learning models quite precise for computer vision – such as object classification.

CNN learn to detect various features of an image using tens and hundreds of hidden layers. Each hidden layer augments the complexity of defined features. For example, the first hidden layer can learn to distinguish object edges, and the last one would detect more complex shapes directly relating to the shape of the object that requires recognition.

Let us note that the workflow of machine learning begins with manual extraction of image features. These features are then used to create a model that would classify objects in an image.

Thanks to the deep learning process specific features are extracted automatically. In addition, deep learning performs «cross-cutting learning» that provides the network with unprocessed data and a task is set (such as classification) and the network learns to perform the task automatically.

A key benefit of deep learning network is the fact that the results continue to improve upon imputing more data.

In most applications deep leaning uses a method based on transfer learning – a process that includes a precise setup before launching a model.

For this purpose a pre-established network can be used and new data of previously unknown categories can be input. After making some changes to the network, a new task can be performed, for example, distinguishing only people or only ATMs among other objects in a shot. A great benefit is reducing the amount of data needed (i.e. several thousands instead of millions of images), therefore the time of computing is reduced to hours or even minutes.

In addition to object recognition that identifies a type of an object on a photo or video, deep learning can also be used for object detection. Therefore, deep learning finds natural ways of implementing it in video security of different sectors.

Annex 2. *Glossary*

Bitrate – the speed of data transition expressed in data units (bites) transferred per time unit (second). It is measured in Kbps and Mbps.

Video content analysis (also known as video analytics) is a methodology of applying computer analysis of video content to obtain finite data on particular surveyed objects.

Video camera is a device that translates an optical image into an electric signal. Video cameras can be wired or wireless, or, according to input type, analog or digital.

Dome camera the most common type of video cameras with a spherical shape.

Video server is a device (server) meant for receiving, storing, playing and broadcasting a video or audio signal.

CCTV camera is a camera for monitoring a building perimeter, its close vicinity and other outside objects.

IP-camera is a stationary camera that has a built-in IP server, a network interface and a connection to *LAN / WAN / Internet*.

Quad processor is a device that allows transmitting four images from different cameras onto a single screen split into four segments.

Switcher is a device for transmitting a signal from several cameras onto a single or several outputs (monitors).

Video multiplexer is a system of video recording and management with expanded functionality to record several (up to 16) video channels form different cameras onto a single cassette, playing such cassettes and processing alarm signals.

Video surveillance is a process performed with the use of technical means that serves to visually control secured and monitored objects.

Movement detector is an electronic block or software that tracks the current image from a video camera and triggers an alarm should the image change.

IR camera is a camera with infrared lighting (that allows «seeing» in the dark).

Frame is a single full shot. In the format of progressive scaling 2:1 of the *RS-170* and *CCIR* standards, a frame is composed from two separated fiends of 262.5 or 312.5 lines that alternate with the frequency of 60 or 50Hz, which allows forming a complete frame with the frequency of 30 or 25Hz. In video cameras that support progressive scaling each shot is projected line by line and does not come one after another. As a general rule, the frequency of the process is also 30 or 25Hz.

Codec is a software that encodes analog signal into digital and manipulates digital signals to make them transferable via channels with lower bandwidth.

In communication engineering the word «codec» usually stands for «coder/decoder». Codecs are used in integral micro schemes or chips that convert analog audio and video into digital format. A codec is also able to convert a digital signal into analog. A codec combines analog-to-digital and digital-to-analog conversion in a single chip.

Matrix switcher is a device of shaping several image sequences from several cameras in any required order, as well as depicting camera numbers and house numbers where cameras are set up, manage alarm signals, current time and date, instructions to the operator etc.

Media container is a file with stored digital information and video.

Pixel is the basic discreet element of an image.

High-speed camera is a camera with a greater viewpoint that allows tracking several control zones simultaneously.

Image resolution is a parameter that determines the degree of granulation of a digital image. The higher the resolution, the higher is the level of details. Resolution is characterized by the number of pixels horizontal (width) to vertical (height), for example, 320x240. Another way of expressing resolution is by naming the whole number of pixels.

Spectral resolution is the maximum number of television channels that are distinguished within a camera output signal in minimal modulation depth being 10%.

Network Video Recorder (NVR) is used to work in IP video surveillance systems. In comparison to common digital video recorders (DVR), which is only capable of recording, storing and playback of video information, NVR collect compressed data via Ethernet. Data may be received from analog or IP video cameras, connected through special adapters. The main specificity of NVR is their limitation in working with a limited number of IP camera models, as currently the standardization of network exchange interfaces is not common.

Hub is used to connect several devices to a network. It processes data transmission for all connected devices, contrary to a switcher that only transfers data with a device the data is intended for.

Telecrane is a fictional device of receiving and transmitting a video signal, described in the dystopian novel «1984» by George Orwell (real name Eric Arthur Blair (1903-1950)).

Standard Definition Television (SDTV) is a variety of televised broadcasting standards, the parameters of which are chosen in accordance to the viewing distance, equal to six heights of a viewed image, Standard definition systems are based on scanning standards 625/50 (576i) and 525/60 (480i) that exist since the 1940, when TV reached the mass audience. There is both analog and digital standard definition television, however the term SDTV mostly refers to digital TV. It is characterized by the 4:3 aspect ratio and average video quality.

High Definition Television (HDTV) is a television system with resolution capability which is almost doubled in comparison to SD.

Thermal imaging camera is a device that functions in the heat spectrum (invisible for the human eye), which allows seeing objects that cannot he seen using a night vision device or video cameras.

Focal distance is the distance between the optical center of a lens and a focal plane (CCD) of a camera.

CIF (Common Intermediate Format) **format** is the basic format of video image transmission used in video surveillance that depends on the system resolution.

H.264 compression format is used for achieving a high degree of compression whilst ensuring good quality.

JPEG compression format is a compression algorithm for rich color still imagery. An image in the JPEG format is a bitmap image with the extensions .JPG or .JPEG. When creating an image in JPEG one can define the level of compression. The lower the compression level, the higher is the image quality and the larger is the file size.

M-JPEG compression format is a shot-per-shot compression method the main feature of which is compressing each individual shot of the stream using the JPEG image compression algorithm.

Motion JPEG compression format is a network video compression and extraction technology. It ensures low delay and stable image quality disregarding its dynamics and complexity. Video quality is defined by the level of compression which, in its turn defines the file size and thus, the transmission speed. *Motion JPEG* video stream can be easily broken down into separate high definition images.

MPEG-1 compression format is a digital audio and video compression standard group developed by *Moving Picture Experts Group*.

MPEG-2 compression format is a digital audio and video compression standard mostly used in broadband broadcasting, including satellite and cable TV.

MPEG-4 compression format is an international standard used mostly in compressing audio and video. Its main fields of application are the internet, compact discs, communication devices (videophones) and TV broadcasting.

Framerate is the frequency of the video stream updates. It is measured in frames per second.

Progressive scanning is the method of forming frames applied in television.

Aspect ratio is the width to height ration of an image. Standard aspect ratio for TV and computer screens is 4:3, whereas HDTV applies the 16:9 format.

AVI (Audio Video Interleave) is a media container that supports simultaneous playback of audio and video.

Bitmap is a data file that is essentially a rectangular pixel grid. It defines the place and color of each pixel (bit) on the screen. GIF and JPEG are file type examples of files containing bitmaps.

Client/server is a process that defines the interaction of two computer programs when one of the programs (client) sends a service request to the other (server), that should manage the request. As a rule, several client programs usually refer to a single server.

CCTV (closed circuit television). The term is often used as a synonym of «video surveillance».

Coaxial cable is a standard means of analog video signal transfer in closed circuit television systems. It is also used in domestic cable TV.

DSP - Digital Signal Processing

DV (Digital Video) one of the first algorithms of video data compression.

DVR - Digital Video Recorder

GIF (Graphics Interchange Format) is one of the more common graphic file formats. There are two versions of this format: 87a and 89a. The 89a version supports animations.

H.264 (or **MPEG-4 part 10**) new generation digital video compression format that ensures higher resolution than Motion JPEG and MPEG-4 at the same bitrate and bandwidth, in other words, high quality video of a low transmission speed.

HTTP (**Hypertext Transfer Protocol**) is a set of rules of transferring files (text, graphic, audio, video and/or other media files) online. The HTTP protocol is the highest degree protocol in the *TCP/IP* protocol family.

HTTPS (**Hypertext Transfer Protocol over SSL**) is a network protocol used by browsers and web pages to encrypt pages requested by the client and those returned by the server. Encrypted data exchange is possible due to using the *HTTPS* protocol, issued by a certification organization, which guarantees server validity.

IEEE 802.11 is a standard family for wireless networks. The *802.11* standard supports transferring data at the speed of 1 or 2 Mbps in the range of 2.4GHz. The *IEEE 802.11b* standard supports transferring data at the speed of up to 11 Mbps at the range of 2/4Ghz, while the standard *802.11g* allows achieving the speed of up to 54 Mbps in the range of 5GHz.

IP protocol (Internet Protocol) ensures the delivery of data packages to a specified address. Since IP is a protocol with no connection organization, which means that there is no connection between end points, packages can be sent using different routes, which requires that they be delivered in the right order. Upon package delivery, anther protocol - TCP (Transmission Control Protocol) sorts them in the correct order.

IP address is the address of the IP network used by computers and devices connected to it. The IP address allows all connected computers and devices to find each other and exchange data. To avoid conflict, each IP address needs to be unique. An IP address can be fixed, or it can be automatically assigned by the *DHCP* protocol. An IP address consists of four groups of decimal digits that are separated by dots, for example, 123.36.43.15.

LAN (**Local Area Network**) is a group of computers and other devices using the same resources and existing on a limited territory.

MPEG (**Moving Picture Experts Group**) develops digital video and audio compression standards. Is a department of the *International Organization for Standardization (ISO*).

Multicast is the technology of optimizing bandwidth based on simultaneous delivery of one data stream to several users of the network.

NTSC (**National Television System Committee**) is a system of colored analog coding used in TV broadcasting in Japan, the USA and other countries in the Americas. Within the NTSC system each frame of the video signal consists of 525 lines with the update frequency of 30fps.

PAL (**Phase Alternating Line**) is a system of colored analog coding used in Tvbroadcasting in Europe and elsewhere. In the PAL system each frame consists of 625 lines with the update frequency of 25fps.

Proxy server serves as the medium between the users of a working station and the Internet. In ensures security and administers cache services. A proxy server connected to the gateway server or its component successfully separates the external network from the local.

PTZ camera is a modern automated surveillance camera renown for its high speed of changing angles.

Server is a computer program that provides services to other computer programs on one or several computers. The computer that runs the server software often is called the server. A server can contain an unlimited number of server and client programs. A web server is a computer program that delivers requested HTML pages or files to the client (browser)

SNMP (**Simple Network Management Protocol**) is a part of the IP protocol package according to the regulation of the Internet Engineering Task Force (IETF). This protocol supports device monitoring, and connecting to a network to notify the administrator about any detected issues.

SSL/TLS (**Secure Socket Layer / Transport Layer Security**) the two protocols are cryptographic protocols that ensure safe data exchange online. Usually the SSL protocol is used along with HTTP, thus leading to establishing the HTTPS protocol which is often used for digital finance operations online. The SSL protocol uses open key certificates to confirm server authenticity.

TCP (**Transmission Control Protocol**) is used along with the IP protocol to transfer data packages between computers within a network. While IP ensures the package delivery itself, TCP tracks individual packages that compose data blocks and composes the file once all packages are delivered.

UDP (**User Datagram Protocol**) is a protocol for exchanging data with limitations on the IP protocol. UDP is an alternative to the TCP protocol. The benefit of UDP is that it does not require all packages to be delivered in a specific order. In fact, some packages can be skipped is the network is overloaded. This is especially convenient when transmitting video in real time, as there is no need to defile outdated information twice.

Unicast is data exchange between a single sender and a single recipient in the network. For each new user a new connection is established.

Virtual Private Network (VPN) allows creating secure tunnels between network nodes. Only devices that possess the correct «key» are able to operate in a VPN.

VOP (Video Object Plane) is a single shot in the *MPEG-4* video stream. There are several VOP types. I-VOP is the entire image frame. P-VOP encodes a difference between images where appropriate. Otherwise it encodes the entire image.

WAN (**Wide-Area-Network**) is a network similar to a local one, but of a much greater geographic span.

Web Server is a program that allows web browsers to obtain files from computers connected to the Internet. Web servers receive file requests from a web browser and upon retrieving, it sends it to the client. The main function of a web server is providing pages to other remote computers that are constantly connected to the Internet. It also manages server access and tracks server access statistics.

WDR (**Wide Dynamic Range**) is the difference in lighting between the darkest and the lightest points of a frame; a significant contrast between the bright and the dark areas.

References

[1.] Владо Дамьяновски. Библия видеонаблюдения, 3rd ed.: Translated from English. – М.: Секьюрити Фокус, 2017. – 422 p.: fig. («Энциклопедия безопасности» series)

[2]. У. Прэтт. Цифровая обработка изображений: Vol.. 1- 2. Translated form English - М.: Мир, 1982.

[3]. Richard O. Duda, Peter E. Hart and David G. Stork. Pattern Classication and Scene Analysis (2nd ed.), 1995

[4]. Торстен Анштедт, Иво Келлер, Харальд Лутц. Видеоаналитика: Мифы и реальность. – М., Секьюрити Фокус, 2012 - 174 p.

[5]. George Orwell. “1984”. Penguin Publishing Group. 2003 reissue

[6]. *ATMeye.iQ*. http://www.atmeye.com/ru/o-produkte/

[7]. Рост рынка видеонаблюдения подпитывается обилием больших данных // Security News 12.01.2018. http://www.secnews.ru/foreign/23589.htm#ixzz56zIohYGu

[8]. Tim A. Scally State of the Market: Video Surveillance 2018 https://www.sdmmag.com/articles/94822-state-of-the-market-videosurveillance-2018?v=preview

[9]. Top Video Surveillance Trends for 2017. By the IHS Markit video surveillance group; https://cdn.ihs.com/www/pdf/TEC-Video-Surveillance-Trends.pdf

[10]. Видеоаналитика из облака на базе VisionLabs LUNA. http://www.tadviser.ru/index.php/Продукт:Крок:_173

[11]. OpenCV (Open Source Computer Vision Library), https://opencv.org/

[12]. Motion Analysis Systems, https://winanalyze.com/

[13]. Tango Concepts. https://developers.google.com/tango/overview/concepts

[14]. Milestone XProtect,

https://www.videosurveillance.com/manufacturers/milestone.asp

[15]. 4 разработки, сделавшие *ATMeye.iQ* лучшим решением для видеобезопасности банкоматов в 2017 году. http://www.bs2.lt/ru/novosti-o-produktah/4-razrabotki-sdelavshie-atmeyeiq-luchshim-resheniem-dlya-videobezopasnosti-bankomatov-v -2017- godu /

[16]. 2017 ATM AND SELF-SERVICE SOFTWARE TRENDS. https://nmgprod.s3.amazonaws.com/media/filer_public/37/84/37842221-ce8f-4726-b55c-438901cf2fc0/kal_guide_2017_final.pdf

[17]. Jill Jaracz. Managing ATM Security: Layered Approaches for 21st Century Issues https://www.atmmarketplace.com/whitepapers/managing-atmsecurity-layered-approaches-for-21st-century-issues/

[18]. Mobile Protector. https://www.gemalto.com/financial/ebanking/sdk/mobile-protector

[19]. EAST publishes 1st ATM crime update of 2018. https://www.atmmarketplace.com/news/east-publishes-1st-atmcrime-update-of-2018/?utm_source=AMC&utm_medium=email&utm_campaign=Week+In+Review&utm_content=2018-03-09

[20]. Microsoft improves its AI face and image recognition tools. https://venturebeat.com/2018/03/01/microsoft-improves-its-ai-faceand-image-recognition-tools/

[21]. Bengio, Y.; Courville, A.; Vincent, P. (2013). "Representation Learning: A Review and New Perspectives". IEEE Transactions on Pattern Analysis and Machine Intelligence. 35 (8): 1798–1828. arXiv:1206.5538 . doi:10.1109/tpami.2013.50.

[22]. Schmidhuber, J. (2015). "Deep Learning in Neural Networks: An Overview". Neural Networks. 61: 85–117. arXiv:1404.7828. doi: 10.1016/j.neunet.2014.09.003. PMID 25462637.

[23]. Convolutional Neural Networks (LeNet). Convolutional Neural Networks (LeNet) - DeepLearning 0.1 documentation. DeepLearning 0.1. LISA Lab. http://deeplearning.net/tutorial/lenet.html

[24]. Convolutional Networks and Applications in Vision. Yann LeCun, Koray Kavukcuoglu and Cl´ement Farabet, Computer Science Department, Courant Institute of Mathematical Sciences, New York University {yann,koray,cfarabet}@cs.nyu.edu

[25]. BS/2 receives 2nd Product of the Year award for *ATMeye.iQ* solution. https://www.atmmarketplace.com/news/bs2-receives-2ndproduct-of-the-year-award-for-atmeyeiq-solution/?utm_source=Email_marketing&utm_campaign=emnaAMC12222017&cmp=1&utm_medium=html_email

[26]. Role of Video in Transforming Retail Banking and Wealth Management. Angus Hislop, David Morland https://www.cisco.com/c/dam/en_us/about/ac79/docs/fs/Video-in-Retail-Banking_IBSG_0418FINAL.pdf

[27]. Дадашев, Т.М., Паукштис, С.С. Горизонты телевидения нового столетия, Вильнюс, 2011.

[28]. Решения для управления сценариями платежей и лояльности. http://www.ashburn.eu/ru/produkty/paylo

[29]. Даниель Фуксон. Современные возможности видеомониторинга: что должна уметь «умная» система. Журнал «Расчеты и операционная работа в коммерческом банке», №2/2018, http://futurebanking.ru/reglamentbank/article4993?access_key=ktg3k

Takhmasib Dadashev

Currently an Associate Professor of the Moscow Institute of Physics and Technology (Russia). Holds a degree of Candidate of physical and mathematical sciences.

Is the author of over 40 research articles and 6 books on innovative technologies, facial recognition and programming. Amongst those – «FrontPage 98», «Java in action. Microsoft Visual J++ 6», «Horizons of new-century television».

Since 1997 is employed as the editor-in-chief of the publication «Penki Kontinentai» that specializes in issues relating to technological development of banking and retail.

In 2005-2007 was invited as the scientific expert to the «Branch Optimizer» project carried out by BS/2 in collaboration with Wincor Nixdorf (Germany). Also took part in a number of BS/2 projects including «ASOMIS».

9 786090 126080

ISBN 978-609-01-2608-0

2018