Method of Analysis of Digital Video Integrity with Linear Computational and Low Technical Complexity

[1]Kobozieva A.A., [1]Lebedieva O.Y., [2]Savchenko V.A.
[1]Odesa National Maritime University, Odesa, Ukraine
[2] State University of Information and Communication Technologies, Kyiv, Ukraine

Abstract. The energy sector is currently undergoing a process of rapid digitalization, which leads to a significant increase in risks and vulnerabilities to cyberattacks in this area. Video sequences are desirable objects of falsification for intruders pursuing illegal goals, making the task of timely detection of their unauthorized changes critically important. However, methods for analyzing/processing digital video, as a rule, have significant computational complexity, which does not allow using them for integrity examination in real time, or require the use of additional technical means. The purpose of this work is to provide the ability to effectively examine the integrity of a video sequence received by CCTV cameras in real time under conditions of minimal technical complexity by developing a method for detecting the result of a frame substitution attack. The goal was achieved by addressing the following tasks: the justified selection of a small video frame sub-region (a key block); determining the integral characteristic of the relative contribution of the block's frequency components—the normalized separation of its maximum singular value. The most important result of the work is the substantiation of the possibility of using the properties of singular values of a single (key) small-sized block of each video frame to obtain information about the presence/absence of frame substitution attack results. The significance of the obtained result consists in the fact that the proposed method enables effective digital video integrity examination in real-time without the need for additional technical means, thereby increasing the reliability and speed of decision-making during video stream monitoring.
***Keywords*:** digital video, video integrity examining, CCTV camera, real-time, singular value.

DOI: https://doi.org/10.52254/1857-0070.2025.4-68.07
UDC: 004.056

Metodă pentru analiza integrității de video digital cu calcul linear de complexitate și cu complexitate tehnică redusă

[1]Kobozeva A.A., [1]Lebedeva O.Y., [2]Savcenko V.A.
[1]Universitatea Națională Maritimă din Odesa, Odesa, Ucraina
[2]Universitatea de Stat pentru Tehnologiile Informației și Comunicațiilor, Kiev, Ucraina

Rezumat. Sectorul energetic trece în prezent printr-un proces de digitalizare rapidă, ceea ce duce la o creștere semnificativă a riscurilor și vulnerabilităților la atacurile cibernetice în acest domeniu. Secvențele video sunt obiecte de falsificare dorite de intrușii care urmăresc scopuri ilegale, ceea ce face ca sarcina detectării la timp a modificărilor neautorizate ale acestora să fie de o importanță critică. Cu toate acestea, metodele de analiză/procesare a videoclipurilor digitale, de regulă, au o complexitate computațională semnificativă, ceea ce nu permite utilizarea lor pentru examinarea integrității în timp real sau necesită utilizarea unor mijloace tehnice suplimentare. Scopul lucrări este de a oferi capacitatea de a examina eficient integritatea unei secvențe video recepționate de camerele CCTV în timp real, în condiții de complexitate tehnică minimă, prin dezvoltarea unei metode de detectare a rezultatului unui atac de substituție a cadrelor. Scopul a fost atins prin abordarea următoarelor sarcini: selectarea justificată a unei subregiuni de cadre video mici (un bloc cheie); determinarea caracteristicii integrale a contribuției relative a componentelor de frecvență ale blocului - separarea normalizată a valorii sale singulare maxime. Cel mai important rezultat al lucrării este fundamentarea posibilității de a utiliza proprietățile valorilor singulare ale unui singur bloc (cheie) de dimensiuni mici din fiecare cadru video pentru a obține informații despre prezența/absența rezultatelor atacurilor de substituție a cadrelor. Semnificația rezultatului obținut constă în faptul că metoda propusă permite examinarea eficientă a integrității video digitale în timp real, fără a fi nevoie de mijloace tehnice suplimentare, crescând astfel fiabilitatea și viteza de luare a deciziilor în timpul monitorizării fluxului video.
Cuvinte-cheie: video digital, examinarea integrității video, cameră CCTV, timp real, valoare singulară.

Метод анализа целостности цифрового видео с линейной вычислительной и низкой технической сложностью

Кобозева А.А.[1], Лебедева Е.Ю.[1], Савченко В.А.[2]

[1]Одесский национальный морской университет, Одесса, Украина

[2]Государственный университет информационно-коммуникационных технологий, Киев, Украина

Аннотация. Энергетический сектор переживает сегодня процесс бурной цифровизации, который приводит к значительному повышению рисков, уязвимостей к кибератакам в данной сфере. С учетом того, что большая часть информации, в том числе, управляющей, представляется и используется здесь в виде цифровых видео, а видеонаблюдение представляет составную часть комплексной системы безопасности объектов энергетической инфраструктуры, видеопоследовательности являются желательными предметами фальсификации для нарушителей, преследующих противоправные цели, делая критически актуальной задачу своевременного выявления их несанкционированных изменений. Однако методы анализа/обработки цифрового видео обладают, как правило, значительной вычислительной сложностью, что не позволяет использовать их для экспертизы целостности в режиме реального времени, или требуют привлечения дополнительных технических средств. **Целью** настоящей работы является обеспечение возможности эффективной экспертизы целостности видеопоследовательности, получаемой камерами видеонаблюдения, в режиме реального времени в условиях минимальной технической сложности путем разработки метода выявления результата атаки замещения кадров **Цель была достигнута** путем решения следующих задач: обоснованного выделения подобласти кадра видео малого размера – ключевого блока; определения интегральной характеристики относительного вклада частотных составляющих блока - нормированной отделенности его максимального сингулярного числа. **Наиболее важным результатом** работы является обоснование возможности использования свойств сингулярных чисел единственного (ключевого) блока малого размера каждого кадра видео для получения информации о наличии/отсутствии результатов атаки замещения кадров. **Значимость полученного результата** заключается в том, что предложенный метод обеспечивает возможность эффективной экспертизы целостности цифрового видео в режиме реального времени без привлечения дополнительных технических средств, что позволяет повысить надежность и скорость принятия решений при мониторинге видеопотоков.

***Ключевые слова*:** цифровое видео, экспертиза целостности видео, камера видеонаблюдения, режим реального времени, сингулярное число.

INTRODUCTION

The energy sector, which is an integral part of the critical infrastructure, is currently undergoing a process of rapid digitalization, which, together with the expansion of opportunities for automation and management in this area, leads to a significant increase in risks and vulnerabilities to cyberattacks [1–3]. One of the main goals of attacks on digital information is to change it – to violate its integrity, which, if not detected in a timely manner in the energy infrastructure, can lead to catastrophic consequences, especially if the information is of a management nature.

The result of the modern development of energy systems is an urgent need for innovative methodologies to optimize their management [4]. The widely used decentralization strategy facilitates the interaction of physical systems in real time, integrates data to enable advanced analytical analysis, including forecasting. Smart grids have become an effective solution for increasing the productivity of energy systems and reducing their operating costs. However, they have not allowed us to avoid one of the main problems of information security – the problem of possible unauthorized changes to the data they use, although certain efforts are being made here and some proposed solutions, in particular, based on the use of blockchain technologies, make it possible to reduce the criticality of the problem [4], but not to solve it completely.

In recent years, video surveillance [5,6] has become an integral part of the integrated security system for critical infrastructure facilities, in particular energy infrastructure. It is a powerful mean of monitoring the situation in real time, which is an important component in making management decisions, and therefore often becomes the subject of falsifications, which today, given the level of development of network and digital technologies, the availability of powerful software environments (Adobe Photoshop, Lightworks, GIMP, etc.), can be carried out efficiently and quickly even without special training of the intruder.

The complexity of the task of protecting the video sequence received by the video surveillance system from unauthorized access, among other reasons, is also due to the fact that such access can be carried out by the intruder

remotely [7], which allows avoiding additional attention and suspicion.

All of this highlights the critical importance of ensuring the effective (and rapid) detection of video sequence tampering [7,8], particularly when the footage is obtained from surveillance cameras. This necessity, in turn, requires that the corresponding forensic methods have low computational complexity to enable their operation in real-time. Although these issues have been addressed in the global scientific literature [9–11], especially in relation to nuclear power facilities, where untimely or incorrect management decisions can lead to catastrophic consequences [12–14], there is still no satisfactory solution to the problem of detecting violations of video integrity. In particular, most existing methods for the analysis and processing of digital video (DV) are computationally intensive and therefore unsuitable for real-time applications.

All unauthorized changes to digital video can be divided into three main classes [7,15]: spatial (intra-frame) [16-18], temporal (inter-frame) [19] and spatio-temporal, which change the content of the video and the temporal relationship between frames. The most common for digital video generated by surveillance cameras are temporal falsifications that change the sequence of video frames [15,20], which, in turn, are divided into: insertion, deletion of frames, reordering, duplication, substitution.

In practice, among the attacks aimed at the digital video received from video surveillance cameras, one of the main ones is the frame substitution attack, which is often identified with duplication in scientific sources [21,22] (although this is not always correct, as will be clarified below), since the first three of the above-listed interframe attacks cannot be carried out in real time. This attack is the subject of further research in this article.

The substitution attack is as follows. Let's assume that there is some video sequence V_1, which we will call the original, and also, possibly, there is a DV V_2. A substitution attack involves copying a certain group of (consecutive) frames of DV V_1/V_2 (this group is further called the *replacement region*), and replacing in V_1 its original n (sequential) frames, forming a *replaced region*, while the total number of frames remains unchanged, compared to the original video, which generally distinguishes replacement from duplication. The first frame of the replaced region will be called the *entry point*. It should be noted that in practice, the replacement region is usually selected from the original digital video V_1, since in this case it has characteristics in common with the video being processed: compression ratio, acquisition conditions (video camera, weather conditions, illumination), etc., and therefore is more difficult to identify during examination.

An effective approach to detecting the result of a substitution attack on a digital image, based on spatial and temporal analysis, is presented in [19]. The integrity examination is implemented here in four stages: 1) selecting a digital image segment that is a “candidate” for being the result of substitution; 2) evaluating spatial similarity; 3) classifying duplicate frames; and 4) post-processing the results. To organize the first stage in the time domain, the difference in the histograms of two adjacent frames is considered as a key feature, which leads to significant computational complexity for the proposed method. The classifier for detecting duplicate parts of the digital image is built on the basis of the results of spatial and temporal analysis, which is its significant advantage, since this approach covers all dimensions of the analyzed object, the digital image. In [19], an attempt is made to reduce the size of the analyzed/processed part in order to reduce the overall computational complexity of the examination process by using the “from rough to fine” principle when searching for the replaced region. However, the reduction of computational costs here depends crucially on the direct selection of features used in the search for “candidates” and in practice does not allow the use of the proposed method in conditions close to real time.

The process of duplicating (replacing) frames within the boundaries of one digital video frame can occur multiple times. This fact is often not taken into account by the authors-developers of methods for detecting video forgeries. In [23], an expert method effective for multiple duplications is proposed. Here, the structural similarity index measure (SSIM) is used to assess the similarity between successive frames. The SSIM values are calculated for each successive pair of frames in the input video, forming a sequence of indices. The key point of the method is the direct definition of SSIM. This index assesses the similarity of two frames R and T in grayscale by integrally comparing their brightness, contrast, and structure: $SSIM(R,T) = l(R,T) \cdot c(R,T) \cdot s(R,T)$, where

$l(R,T)$ corresponds to the comparison of brightness by calculating the proximity of the average brightness values of two frames, $c(R,T)$ is a function for comparing the contrast of two frames, $s(R,T)$ is a structural comparison function, measures the correlation coefficient *R* and *T*. Meaning $SSIM(R,T) \in [-1,1]$. The closer the value $SSIM(R,T)$ to one, the greater the degree of similarity of frames. In the course of the method, the principle scheme "from rough to more precise" is also used. The method proposed in [23], due to the possibility of integral comparison of frame similarity, is one of the most effective modern expert methods, achieving the accuracy of frame duplication detection of 98.90%, however, it cannot be used for work in real time, since it is designed to work with a finally formed digital image, like most existing analogues.

In [24], an approach to detecting interframe video forgery is presented based on the use of convolutional neural networks CNN, consisting of four stages. In the first stage, the digital video is transformed into a sequence of matrices corresponding to its frames. The second stage is devoted to extracting frame features using two CNNs, followed by calculating the correlation between these features and the difference in correlation to find a connection between the frames. The values of the upper and lower boundaries of the parameters under consideration for identifying duplicate frames are determined in the third stage. The last stage, using the obtained threshold values, allows making a final conclusion about the originality/non-originality of the digital video and localizing the non-original area. Obviously, for efficient operation, this method must have a fully formed video, which excludes the possibility of its operation in real time.

In [25], a method for detecting duplication (substitution) of frames in digital video is proposed, based on the analysis of their texture features. The use of these features makes it possible to effectively detect without isolating small-sized frame subsequences, as, for example, in the method discussed above. For duplicated frames of digital waveforms, these features, which the authors consider to be contrast, correlation, energy, and homogeneity, are also duplicated. To obtain a quantitative characteristic of the frame texture, its discrete wavelet transform is used. The authors provide an overall high estimate of the proposed method's efficiency, which is 99.8%. However, this expert method requires a digital waveform formed in its final form to operate, which fundamentally excludes the possibility of its use in real time, leaving the problem under consideration relevant.

In [26], an approach to detecting and localizing frame duplication regions in digital waveforms is presented, based on the use of the "improved" Levenshtein distance to determine the degree of similarity of frame subsequences into which the digital waveform is divided. The degree of similarity is calculated for each obtained pair of frame subsequences. The authors claim high, but non-systematic efficiency of the proposed method, which can reach 99.5%, but it is obvious that this method is not suitable for working in real time, since it works with a fully formed video sequence, while sequentially sorting through pairs of subsequences of the original digital video, which additionally indicates the high computational complexity of the examination process.

In [27], an algorithm based on the QR decomposition of the frame matrix is proposed for detecting frame duplication. The QR decomposition allows extracting the necessary characteristics, which are compared with the characteristics of the reference frame using the Minkowski distance during the examination. Duplicate candidates are identified by randomly comparing blocks. The authors announce the high efficiency of the proposed method, in particular in conditions of multiple duplication, as well as in conditions of digital image post-processing. However, given that the computational complexity of the QR decomposition $n \times n$-matrix is defined as $\underline{O}(n^3)$, the use of such a method is impossible in real time.

In [13], an effective method is proposed for the real-time integrity analysis of digital video (DV) captured by a surveillance camera. The specific type of attack on the video sequence considered here is a 'screen overlay' – where a portion of the generated DV is replaced by a single frame that remains constant for a certain period of time. The indicator of unauthorized video modification here is the fact that the frame's matrix remains unchanged over time. To reduce computational complexity, a small sub-region of insignificant size is extracted from the current DV frame and subjected to analysis. The accuracy of the method here is highly dependent

on the size of the frame region, ranging from 76 to 100%.

An effective method for real-time DV integrity analysis is proposed in [28], where the classic frame substitution attack is considered. This work proposes the technology Secure-Pose, which exploits the pervasive coexistence of surveillance and Wi-Fi infrastructures to defend against video forgery attacks in a real-time. Although Secure-Pose achieves a high detection accuracy of 98.7%, this does not provide a definitive solution to the problem of detecting unauthorized DV modification, as it requires the presence and utilization of additional technical means.

Thus, the *purpose* of this work is to provide the ability to effective examine the integrity of the digital video obtained by video surveillance cameras in real time under conditions of minimal technical complexity by developing a method for identifying the result of a frame substitution attack of low computational complexity.

Minimal technical complexity is defined as the complete absence of additional technical equipment required for the digital video examination.

To achieve the goal, the following *tasks* are solved in the work:

1. Justification of the method for reducing the analyzed region of the digital video frame to ensure low computational complexity of the developed expert method.

2. Determination of the integral parameter of the frame subregion, informative for its selection.

3. Development of a method for selecting a frame subregion for further examination of the integrity of the digital video.

4. Justification of the choice of the parameter/parameters of the frame subregion determined as a result of solving the previous problems to identify the replacing and replaced regions of the video sequence.

5. Development of an expert method for identifying the results of a frame substitution attack, assessment of its computational complexity.

METHOD FOR IDENTIFYING KEY BLOCKS OF DIGITAL VIDEO

Recently, the General Approach to the Analysis of the State of Information Systems (General Approach) [18,29], based on perturbation theory and matrix analysis, to which the authors paid much attention on the pages of this journal in 2024, 2025, has proven itself well for solving the problems of digital content integrity examination. In this regard, General Approach is further used as a theoretical basis for the developed expert method.

The fundamental approaches in the proposed expert method aimed at identifying frame substitution attacks are: organizing the search for the falsification region, organized according to the principle of "from rough to more precise", as well as the idea of the spatial and temporal analysis approach proposed in [19]. The frame substitution attack, the purpose of which is most often in practice to remove objects / events unwanted by the intruder from the observed scene, is considered in real time. It is further assumed that the video surveillance camera generating the digital signal being examined is located outside the premises and is stationary, which corresponds to the nature of the use of the overwhelming majority of cameras.

The specifics of the problem under consideration will obviously force the intruder to use frames of the currently generated digital video sequence for the substitution attack, since otherwise the probability of detecting the result of the substitution will increase (including visually). But even if the substitution region belongs to the generated video sequence, then given that the camera is outside the room, weather conditions, illumination, etc. can change significantly over a short period of time, and to reduce the probability of detecting the falsification, the replacement region will obviously include successive frames (from the already generated part of the video), taken from the immediate vicinity of the entry point. Let us call this condition (A), the fulfillment of which is assumed everywhere below.

Taking into account that the digital video integrity examination should be performed in real time, in order to reduce the computational complexity of the corresponding method, it is necessary, if possible, to reduce the analyzed frame region, i.e. to select from the frame some informative subregion for the considered task. For this purpose, it is proposed to select a small-sized block in the frame, which remains relatively unchanged over the entire period of time of formation of the original digital video when moving from frame to frame: not containing (moving) objects – a block with small differences in pixel brightness values, which will be further called the background, or homogeneous, block. Changes in the formal

quantitative indicators of such region, the localization of which remains constant from frame to frame, although they will take place in the original video sequence, but they will not be associated with changes occurring in the frame scene, will be insignificant, compared to areas containing moving objects. It is obvious that the frame replacement process must lead to greater deviations in the parameters of such a block between two adjacent frames, one of which is the original and the other corresponds to the entry point, compared to two consecutive original frames. In practice, it has been established that for the information content of the analysis carried out below, it is sufficient to consider one $l \times l$-frame block where $l \in \{16,32\}$, which is then used in the development of an algorithmic implementation of the method for determining the key block of the digital signature.

In order to be able to isolate the desired block, it is necessary to determine such a quantitative parameter (parameters) that would characterize the degree of its homogeneity. It is known that blocks with minor differences in brightness values have an insignificant high-frequency (and possibly mid-frequency) component. However, even for a block of minor dimensions $l \times l$, its high-frequency component is determined by several frequency coefficients [30]. Taking into account the values of such parameters to determine the desired block will obviously be difficult, since they and their number are not fixed even for a fixed l, while the values of these frequency coefficients, and, as a consequence, their total contribution can differ significantly within the blocks of one digital video frame even in the case when these blocks have minor differences in brightness values. It is desirable here to determine one integral quantitative parameter of the block that would characterize the total relative contribution of its frequency components.

Let B be the $l \times l$ block of the frame matrix of the DV, and

$$B = U \Sigma V^T \qquad (1)$$

is its normal singular value decomposition [31], where U, V is an orthogonal $l \times l$-matrices, the columns of which $u_i, v_i, i = \overline{1,l}$ are left and right singular vectors of B, respectively, and $\Sigma = diag(\sigma_1,...,\sigma_l)$is a diagonal matrix of singular values, where: $\sigma_1 \geq ... \geq \sigma_l \geq 0$. The singular value decomposition can be written in the form of outer products [32]: $B = \sum_{i=1}^{l} \sigma_i u_i v_i^T$. According to General Approach, there is a correspondence between singular triples $(\sigma_i, u_i, v_i), i = \overline{1,l}$, of block B of the frame of the DV and its frequency components: singular triplets containing maximum/minimum/average values of the singular value carry, mainly, information about the low-frequency/high-frequency/mid-frequency components of B. As shown by the authors earlier, it is the singular values that are responsible in the matrix $\sigma_i u_i v_i^T$ for a specific main particular component of the block. Thus, l scalar parameters – singular values provide information about l^2 frequency coefficients of the block, which obviously allows, when using the first set, to reduce by an order of magnitude the total number of interesting block parameters in the process of selecting the desired one, compared to the second. The expected result of such a reduction will be no increase in the computational complexity of the developed method for selecting a frame subregion for further examination of the integrity of the digital image (task 3) when replacing the analysis of frequency coefficients with the analysis of the singular values to determine the degree of homogeneity of the block.

Let us establish the differences in the relationships between the singular values of homogeneous/non-homogeneous blocks. For $l \times l$ blocks of original digital video frames, regardless of the storage format (with losses/lossless), the following holds:

$$\sigma_1 >> \sigma_2 \geq ... \geq \sigma_l , \qquad (2)$$

in this case, the more significant the superiority σ_1 over the rest of the singular values block, the higher the relative contribution of the low-frequency component to such a block, the closer such a block is to a homogeneous one. The degree of difference σ_1 from other singular values can differ significantly for blocks of even one frame of the digital video. To reduce the effect of the dependence of the mentioned property on a specific block, it is proposed to normalize the singular values. Let the vector $\sigma = (\sigma_1, \sigma_2, ..., \sigma_l)^T$ and let us designate $\bar{\sigma}$ normalized vector of the singular values, determined in accordance with the formula:

$$\bar{\sigma} = \sigma / \|\sigma\| = \left(\bar{\sigma}_1, \bar{\sigma}_2, ..., \bar{\sigma}_l\right)^T, \quad (3)$$

where $\|\sigma\|$ is a vector norm σ. It follows from (2) that: $1 \geq \bar{\sigma}_1 >> \bar{\sigma}_2 \geq ... \geq \bar{\sigma}_l \geq 0$. The closer $\bar{\sigma}_1$ to one, the closer all other elements are $\bar{\sigma}$ to 0, the smaller the contribution of the high-frequency component to the block, the smaller the difference in pixel brightness values within this block. Finding such (homogeneous) blocks, as noted above, is an intermediate goal in finding the informative part of the frame.

The authors previously introduced the concept of normalized separation $svdgap_n(i)$ singular value σ_i, which is determined in accordance with the formula:

$$svdgap_n(i) = \min_{i \neq j} \left| \bar{\sigma}_j - \bar{\sigma}_i \right|. \quad (4)$$

From (4) it follows that:

$$svdgap_n(1) = \bar{\sigma}_1 - \bar{\sigma}_2. \quad (5)$$

The normalized separation of the maximum singular values block (5) determines the relative superiority σ_1 over all other elements of its singular spectrum without taking into account their immediate values. From (4) it follows: $0 < svdgap_n(1) \leq 1$.

From the above it follows that homogeneous blocks of the digital frame will be characterized by the ratio:

$$svdgap_n(1) \approx 1. \quad (6)$$

Thus, as an integral parameter, which is an indicator of the homogeneity of the frame block of the digital video signal, it is proposed $svdgap_n(1)$.

Example of use $svdgap_n(1)$ for the selection of homogeneous 16×16 blocks is shown in Fig.1(a,b) (here the proximity indicator $svdgap_n(1)$ to the unit in (6): $svdgap_n(1) \geq 0.99$) for the first frame of the digital video, which is considered below to demonstrate the operation of the proposed expert method. The search for homogeneous blocks was carried out among all possible $l \times l$-blocks of the frame, the receipt of which can be formally represented by successive shifts relative to some (randomly selected) block to the left, right, down, up by 1 pixel until reaching the frame boundaries, as well as among non-intersecting blocks obtained as a result of the standard partitioning of the frame matrix [33]. In the first case, the provided set of homogeneous blocks will be significantly larger (Fig. 1(a)) than in the second (Fig. 1(b)), while providing more opportunities to select one of them that will remain background for the entire video sequence, but the computational costs here will be in l^2 times more than in the second case. Indeed, for the first case, it can be assumed that each element of the frame matrix will have its own block (exceptions are the matrix elements lying on its border and close (at a distance less than l) to it), i.e., the number of blocks for $n \times n$-frame matrix will be $\approx n^2$. With standard frame splitting, the number of $l \times l$-blocks will be defined as $\left[n/l\right]^2$, where $[\cdot]$ is the integer part of the argument, i.e., it will be approximately l^2 times less. The choice of one or another option here will depend on the specific conditions and requirements for solving the problem under consideration.

In this work, a “neutral” object is used instead of an object from energy sector to demonstrate the results.

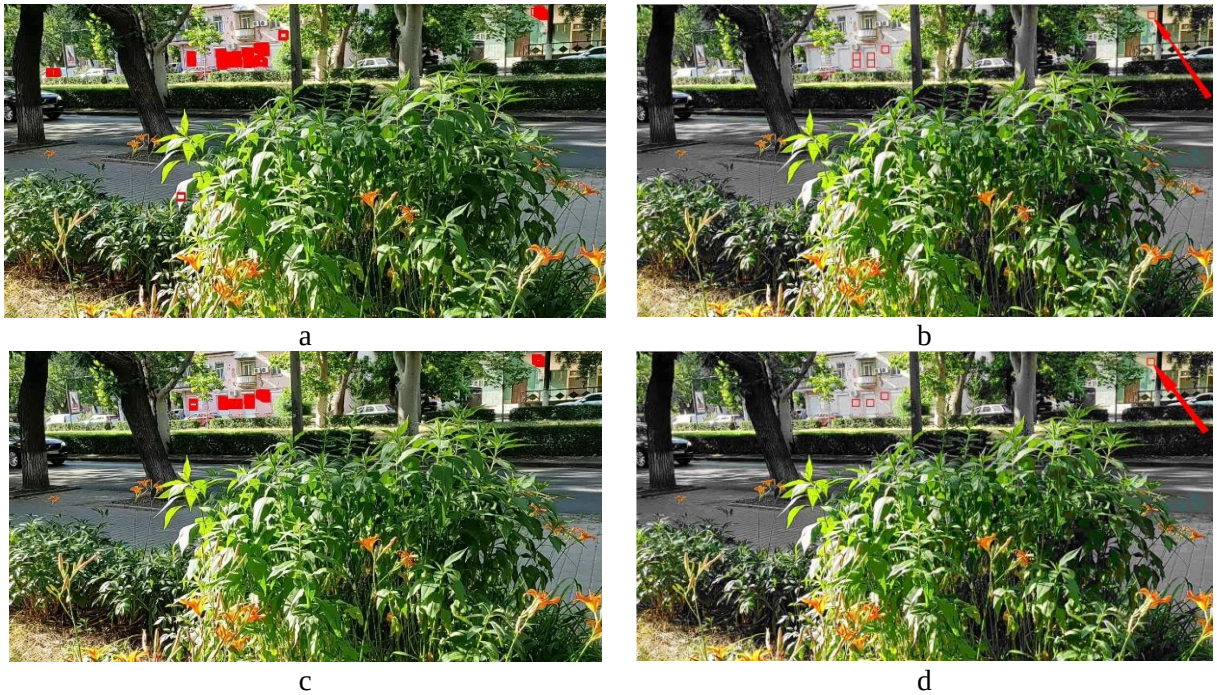
After receiving a set of homogeneous blocks of the digital video frame, it is necessary to select from them those that will remain homogeneous throughout the video, the changes of which from frame to frame will be insignificant – the *key* blocks of the video sequence. For a quantitative assessment of these changes, the correlation coefficient for the pixel brightness values was initially used as a natural indicator of the "similarity" of the blocks. However, under the conditions of the problem under consideration, such an indicator did not justify itself. Due to the fact that the video camera is supposed to be installed outdoors, changes in natural illumination, which can occur almost instantly, and other natural phenomena lead to the fact that even when the block remains homogeneous throughout the digital video, the change in the brightness of its pixels from frame to frame can be significant. This can result in a significant decrease in the correlation coefficient between the blocks of adjacent frames, falsely signaling an increase in the high-frequency component (the appearance of contours) in the block. An example is shown in Fig. 2 for one of the blocks defined in Fig. 1(b) (marked with a red arrow). This block remains uniform throughout the DV, giving "gaps" in the correlation values. It should be noted that for other of the defined uniform blocks viewed during the experiment, the correlation coefficient

could decrease as low as 0.69. Thus, using the correlation coefficient of the pixel brightness values of the blocks as a quantitative indicator of block "similarity" for determining the key blocks of the DV is inappropriate.

Unlike the correlation coefficient, the value $svdgap_n(1)$ block due to its "normalization", determining the relative correspondence between the components of the singular spectrum, is practically independent of the increase/decrease in the pixel brightness value if the block remains homogeneous. Due to this, the normalized separation of the maximum singular values block is further used as a determining parameter for identifying key blocks of the video sequence.

The search for the location of key blocks of the digital video is performed in advance for each video surveillance camera, taking into account its stationarity, before the process of direct examination of the integrity of the video sequence generated by this camera. To do this, for the identified homogeneous blocks of the frame, it is necessary to check the satisfaction of condition (6) on the remaining frames of the already existing video generated by this camera, over a certain time interval *T*.

Let $W = \{w_1, w_2, ..., w_p\}$ is an existing video sequence, and $w_1, w_2, ..., w_p$ – DV frames, which are considered as color digital images stored in accordance with the RGB or YUV scheme [33]. The coordinates of an arbitrary block of the DV frame matrix correspond to the indices of the element of this matrix, which is located in the block in the upper left corner (at the location (1,1)). Taking into account the above, the main steps of the method for determining the key blocks of the DV generated by a specific video camera are as follows.



a b c d

a, b – homogeneous blocks found during the search among intersecting and non-intersecting 16×16 blocks, respectively; c, d – recommended blocks for use in the examination of the integrity of the digital image among intersecting and non-intersecting 16×16 blocks, respectively.

Fig. 1. Results of the method for selecting key blocks for the digital image frame.

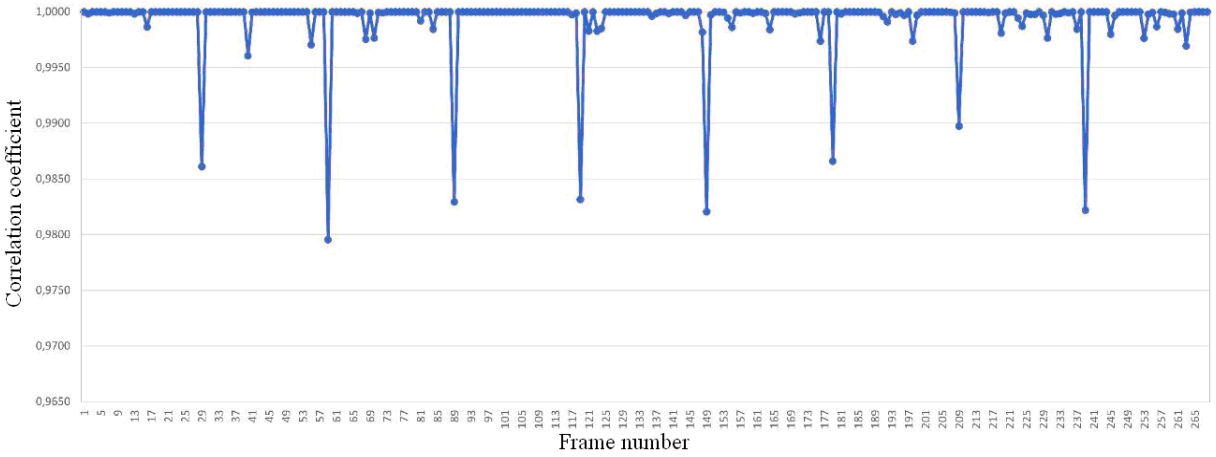


Fig. 2. Graph of the dependence of the correlation coefficient between the brightness values of pixels of pairs of blocks of consecutive digital video frames on the frame number.

Step 1. Search for homogeneous blocks in the matrix of an arbitrary frame of the digital image. For certainty, the first frame w_1 is considered. X_1 – matrix that is assigned to the frame w_1, reflecting the direct values of the pixel brightness (color component/luminance matrix).

1.1. Matrix X_1 split into intersecting/non-intersecting $l \times l$ blocks $c_i, i = \overline{1,s}$: $X_1 = \bigcup_{i=1}^{s} c_i$, each next block c_i differs from the previous one c_{i-1} by shift k pixels to the right/left/down/up; $res = \varnothing$, where res – set of coordinates of the found homogeneous blocks;

1.2. For each block $c_i, i = \overline{1,s}$, with coordinates (x_i, y_i):

1.2.1. Construct a singular value decomposition (1): $c_i = U_i \Sigma_i V_i^T$, where $\Sigma_i = diag\left(\sigma_1^{(i)},...,\sigma_l^{(i)}\right)$;

1.2.2. Define a vector $\bar{\sigma}$ (3);

1.2.3. Determine the value of the normalized separation of the maximum singular value $svdgap_n^{(i)}(1)$ of block c_i in accordance with (5);

1.2.4. *If* $svdgap_n^{(i)}(1) > th$, *then* $res = res \cup \{(x_i, y_i)\}$, where th - a parameter whose value is determined experimentally;

1.3. *If* $res = \varnothing$, *then* decrease the value of l and go to step 1.1, *else* $res = \{(x_{i_1}, y_{i_1}), (x_{i_2}, y_{i_2}),..., (x_{i_t}, y_{i_t})\}$ – a formed set of coordinates of homogeneous blocks of the frame w_1; $|res| = t$, where $|res|$ – a power of set res.

Step 2. Search for key blocks of the DV.

2.1. An array rec of counters is introduced for each homogeneous block from the set res: $rec = (r_1, r_2,..., r_t)$. Here r_j corresponds to the block with coordinates (x_{i_j}, y_{i_j}). Originally: $r_j = 0$, $j = \overline{1,t}$.

2.2. For each block c_{i_j}, the coordinates of which are contained in res, $j = \overline{1,t}$, each frame w_i with a matrix X_i, reflecting the direct values of the pixel brightness, from $W = \{w_1, w_2,..., w_p\}$, i>1:

2.2.1. Construct the singular value decomposition (1): $c_{i_j} = U_{i_j} \Sigma_{i_j} V_{i_j}^T$, where $\Sigma_{i_j} = diag\left(\sigma_1^{(i_j)},...,\sigma_l^{(i_j)}\right)$;

2.2.2. Define the vector $\bar{\sigma}$ (3);

2.2.3. Determine the value of the normalized separation of the maximum singular value $svdgap_n^{(i_j)}(1)$ of block c_{i_j} in accordance with (5);

2.2.4. *If* $svdgap_n^{(i_j)}(1) > th$, *then* $r_j = r_j + 1$.

2.3. Among the elements of the array $rec = (r_1, r_2,..., r_t)$ find such r_j, that

$$r_j = p. \qquad (7)$$

Form a set *M* of indices *j*, for which it is performed (7).

2.4.*If* $M = \varnothing$,
then decrease the value of *l* and go to step 1.1,
else key blocks are blocks with coordinates (x_{i_j}, y_{i_j}), where $j \in M$

For the algorithmic implementation of the method, the key issue is the choice of the block size *l*: the larger *l*, the greater the probability that the result will be: $res = \varnothing \vee M = \varnothing$; the smaller *l*, the greater the computational costs of determining the key block. As a result of the computational experiment, as already indicated above, it was established that in practice the most preferable will be $l \times l$ blocks, for which $l \in \{16,32\}$. Recommended parameter values for the algorithmic implementation of the proposed method: $l = 16$, $th = 0.99$, $k \in \{1,16\}$. As a result of using such an algorithmic implementation for the DV presented in Fig. 3(a), the blocks presented in Fig. 1(c) (for k=1) and in Fig. 1(d) were recommended as key blocks.(for *k*=16).

METHOD OF DIGITAL VIDEO INTEGRITY EXAMINATION

Let us designate B_i - one of the pre-selected key blocks corresponding to the *i*-th frame of the DV. The location of the key block in the frame does not change with the change of *i*. Taking into account the General Approach, the quantitative characteristic of the state of the selected block in each *i*-frame is proposed to be the value $\sigma_1(B_i)$, since the maximum singular number of a block carries the main information about one of its main integral characteristics – energy $E(B_i)$, calculated by the formula:

$$E(B_i) = \sum_{j=1}^{l} \sigma_j^2(B_i), \qquad (8)$$

due to the relation (2), while adequately responding to changes occurring in the block matrix, in accordance with the good conditioning of the singular values [32]: $\max_{1 \le j \le l} \left| \sigma_j(B_i) - \sigma_j(B_i + \Delta B_i) \right| \le \left\| \Delta B_i \right\|_2$, where ΔB_i is $l \times l$ block disturbance matrix B_i, $\|\cdot\|_2$ – spectral matrix norm.

From frame to frame $\sigma_1(B_i)$ in the original DV will change with a greater or lesser speed due to changes in the matrix B_i,caused by changes in the environment occurring outside the closed room in the scene being filmed, while the rate of change $\sigma_1(B_i)$ at the entry point of the DV will expectedly undergo a jump. Indeed, the difference between the previous frame and the next one, corresponding to the entry point, i.e. the first frame of the replaced region, will be additionally conditioned by the difference in time when these frames are received in the original DV, which, taking into account condition (A), will depend mainly on the number of frames in the replacement regionand will be minimally equal to the time of their generation if the first frame of the replaced regionis an immediate neighbor of the last frame of the replacement one, and in the general case will be greater. Thus, the expected formal indicator of the presence of an entry point is the occurrence of a local/global extremum for the function $y(i)$ of the dependence of the rate of change $\sigma_1(B_i)$ on the frame number *i*. The rate of change $\sigma_1(B_i)$ in *i* is estimated below using the divided difference of the first order, i.e. $y(i) = \sigma_1(B_i) - \sigma_1(B_{i-1})$.

For clarity of further presentation, we will use a specific DV that has undergone a frame substitution attack, presented in Fig. 3(b), the graph of the corresponding function $y(i)$ for which is shown in Fig. 4(a).

In the practical implementation of the obtained theoretical conclusions, the following questions arise: firstly, since the function $y(i)$ is discrete, it can have local extrema at practically every point (Fig. 4), therefore it is necessary to ensure the possibility of separating such extrema that are “suspicious” of being entry point indicators, for which the principle “from rough to more precise” is further used; secondly, information is available at each moment in time – the moment of creation of a specific frame $\bar{i}$, only about previous frames with numbers $i < \bar{i}$, which will not allow mathematically determining the extremum at the point $\bar{i}$ before the moment of receiving the frame with number $\bar{i}+1$, even if the extremum will take place there; waiting for the frame $\bar{i}+1$ to be received will make the operation of the developed expert method in real time fundamentally impossible.

The solution to the issues that have arisen is implemented as follows. Let us set aside for a while the requirement for the method to operate

in real time and look at the function $y(i)$ as a whole (for clarity – Fig. 4(a)). The extremum that occurs at the entry point is, among other things, the result of a break in the correlation links between adjacent frames, so the absolute value of such an extremum cannot be small in itself, compared to other extrema. Thus, for the primary separation of extremes, an experimentally determined threshold value δ is used: a local extremum will be interpreted as one that can determine the entry point if its absolute value is greater than δ.

Let us return to the issue of implementing the possibility of the method operation in real time. The value of the extremum is determined by the value of the function $y(i)$ at the extremum point. Since the key block under consideration in the DV frames is uniform throughout the entire time of creating the video sequence, significant jumps $y(i)$ from frame to frame of the original DV are unlikely, although possible (taking into account changes in the environment, since the camera is outdoors), while the probability is very small that in the presence of a jump in the value $y(i)$ such that $|y(i)| > \delta$, this comparatively small value in absolute value will be the same or will increase in absolute value, preserving its sign, when moving to the next frame of the original video, since this would mean a significant sequential increase (decrease) in the maximum singular value for three consecutive frames. Taking into account that: in accordance with (8) and (2), the energy of the key block B_i satisfies the relation $E(B_i) \approx \sigma_1^2(B_i)$; the key block is the block with small differences in brightness values; and also taking into account the correlation of the brightness of the corresponding pixels between adjacent frames of the original digital video, a sequential multiple increase/decrease in energy from frame to frame is unlikely here. The above is illustrated by Fig.4. Thus, after a jump in value $|y(i)|$ when moving to the next frame, it will most likely decrease, giving an extremum at the previous point, and therefore the digital video frames for which are "suspicious" of being the entry point will be:

$$|y(i)| > \delta. \qquad (9)$$

However, as the graph in Fig. 4(a) illustrates, such frames do not necessarily correspond to the entry point. They require an additional quick check – clarification, which is proposed to be carried out as follows. As the next frame is formed, the maximum singular value of its key block will be stored in a one-dimensional structure – an array, the next element of which corresponds to the next frame of the digital signature. This array is formed simultaneously with the formation of the digital signature. As soon as the local extremum of the function $y(i)$ is found in accordance with (9), an additional check is carried out on the elements of the formed part of the singular values vector to find a value that coincides with the current value $\sigma_1(B_i)$ (if additional processing of the replaced region was not done) or differs by an amount not greater than a certain threshold value ε, otherwise. Finding such a value will determine the presence of frame replacement and the location of the replacement region. Note that, taking into account condition (A), the time costs for searching for the required frame of the replacement region will be insignificant, regardless of the absolute frame number *i*, if this search is organized by the elements of the formed part of the array of singular numbers in reverse order: (*i*-1)-th element, (*i*-2)-th element, etc. The maximum required for the search will be a comparison operation for the current key block B_i, which will occur in the case when B_i does not belong to the replacement region, or when the first frame of the replacement region is the first frame of the DV. In all other cases, the number of operations will be less.

For convenience of description of the proposed expert method the following notations are used. Let $F = \{f_1, f_2, ..., f_n\}$ – be the DV already formed by the video surveillance camera at the given moment of time from successive frames $f_1, f_2, ..., f_n$, and let B_i – be the key block f_i of size $l \times l$. For convenience of presentation let us denote D – a one-dimensional array, the *i*-th element of which $D(i)$ is defined as: $D(i) = \sigma_1(B_i)$. Let the set *res* be used here to collect frames of the DV that form the replacement region. Initially $res = \varnothing$. Before the start of the examination process *find* - the indicator of the presence of replaced frames is brought to the state: $find = false$.

The main steps of the proposed expert method for identifying the replaced region of the DV are as follows.

Step 1 (*initialization*). $i = n + 1$.

Step 2. For a frame f_i obtain $\sigma_1(B_i)$ by singular value decomposition (1) of the key block B_i: $B_i = U_i \Sigma_i V_i^T$, $\Sigma_i = diag(\sigma_1(B_i),...,\sigma_l(B_i))$.

Step 3. *If* $find = false$

Then calculate $y(i) = \sigma_1(B_i) - \sigma_1(B_{i-1})$.

If $|y(i)| > \delta$,

Then

If among the elements $D(j), j = \overline{1,i-1}$ there is such a $D(j_0)$, that $|D(j_0) - \sigma_1(B_i)| \le \varepsilon$, *Then* $(res = res \cup f_i) \& (find = true)$, output numbers j_0, i of replacement and replaced frames *Else* $find = false$.

Else

If among the elements $D(j), j = \overline{1,i-1}$ there is such a $D(j_0)$, that $|D(j_0) - \sigma_1(B_i)| \le \varepsilon$, *Then* $(res = res \cup f_i) \& (find = true)$, output numbers j_0, i of replacement and replaced frames *Else* $find = false$.

Step 4. $D(i) = \sigma_1(B_i)$.

Step 5. The video sequence is formed $F = \{f_1,..., f_n, f_{n+1}\}$; $n = n+1$. Transition to step 1.

It is obvious that at the moment of determining the entry point in real time it is fundamentally impossible to determine the entire falsified region, since it has not yet been fully formed, but the signal about the presence of falsification gives here the opportunity to take emergency measures to prevent/properly react to illegal actions that obviously take place in reality. The proposed method will determine the replaced (non-original) region of the video when this region is fully formed, but information about the numbers of replaced (and replacement) frames is supplied to it sequentially in real time. The final form of the replacement region, if this is of interest, will be formed in *res*.

The algebraic implementation of the proposed method corresponds to the following parameter values: $l = 16$, $\delta = 5$ (established experimentally), the parameter value ε will be determined in accordance with those disturbing effects that can be applied as post-processing to the DV frames after their duplication. However, it should be noted here that, due to the assumed real-time process, additional post-processing of frames is unlikely in practice.

Let us now illustrate in detail the use of the proposed method for the falsified DV (Fig. 3(b)). The DV did not experience any additional disturbing effects, except for frame replacement ($\varepsilon = 0$). The graph of the function $y(i)$, $i = \overline{139,268}$, for the block, the location of which in the frame is indicated in Fig.1(d) by the red arrow, is shown in Fig. 4(a). In reality, the replacement region is frames from 155 to 185, and the entry point is frame 205. During the operation of the method, until the entry point is detected in this falsified video, only one "suspicious" 179-th frame is analyzed, for which $|y(i)| > \delta$, for which there are no repeating values for i<179, which excludes it as the entry point. The next value: $|y(204)| > \delta$, which defines the 205-th frame as the first replaced frame (entry point) (i=205, $j_0 = 155$) due to the coincidence of $y(205) = y(155)$. After the entry point is fixed, the subsequent frames of the replaced region are also determined. In the graph $y(i)$ (Fig. 4(a)) the parts corresponding to the replacement and replaced regions of the DV are highlighted in red and green, respectively.

An important distinctive feature of the proposed expert method is that it is able to detect the result of frame replacement and duplication not only in the case when there is one replacement and one replaced region, but also in cases when one replacement region is used several times, creating several replaced ones, and also when there are several replacement and several replaced regions within one digital image. It is important to note here that if several replaced regions of the digital image corresponding to one replacement region are located close to each other, then the entry point may not be the first to satisfy the condition $|y(i)| > \delta$, which will not prevent its detection, since when moving to this entry point, the value of the pointer to the falsified frames will be determined as $find = true$. All of the above is confirmed by the graph of the function $y(i)$ (Fig.4(b)), constructed for the DV shown in Fig.3(c), where one replacement region corresponds to two replaced ones: frames from 205 to 235 and from 238 to 268.

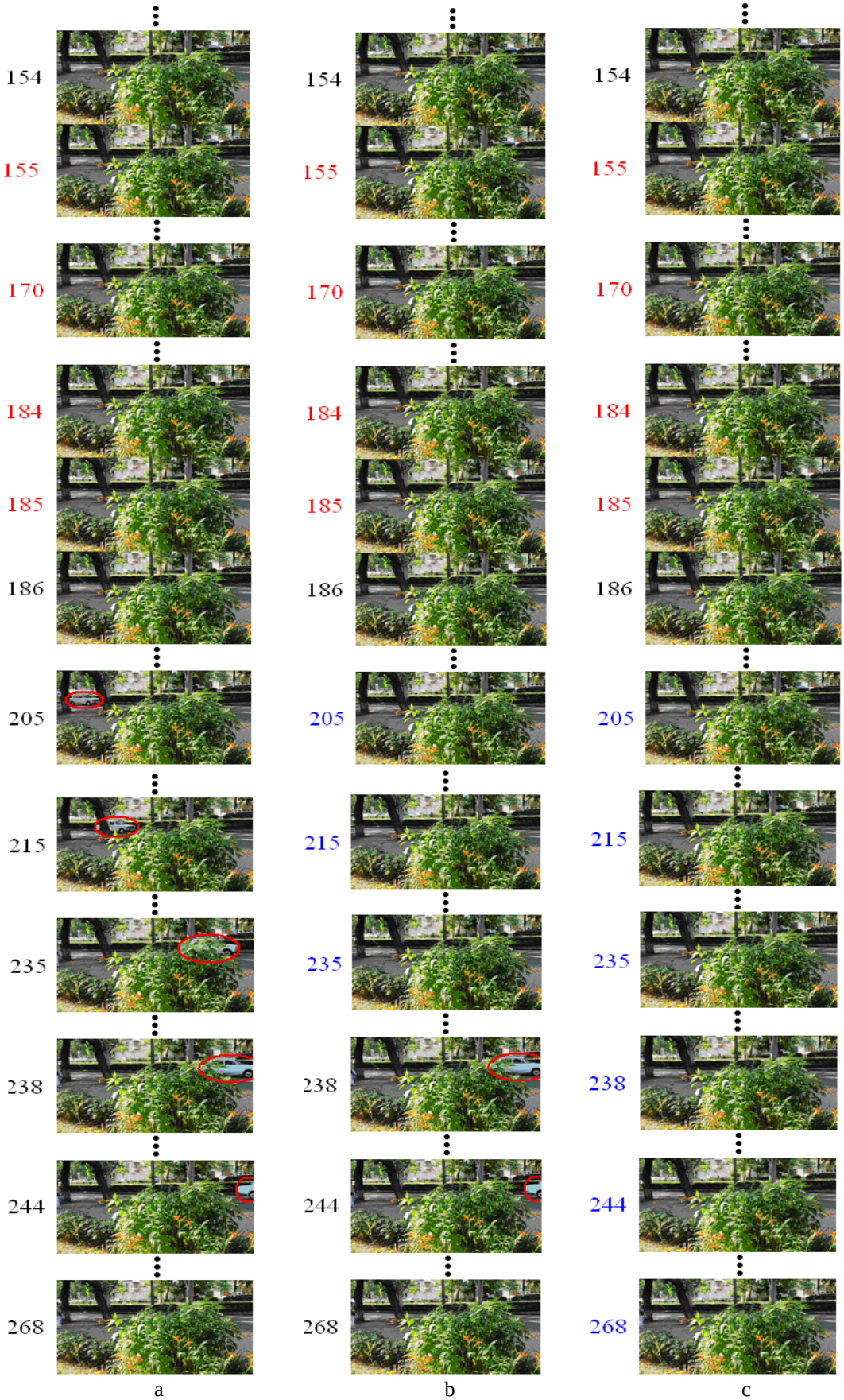


Fig.3.Original and counterfeit DV: a – original DV; b – DV resulting from frame substitution (one replacement region– one replaced region); c – DV resulting from frame substitution (one replacement region– two replaced regions).

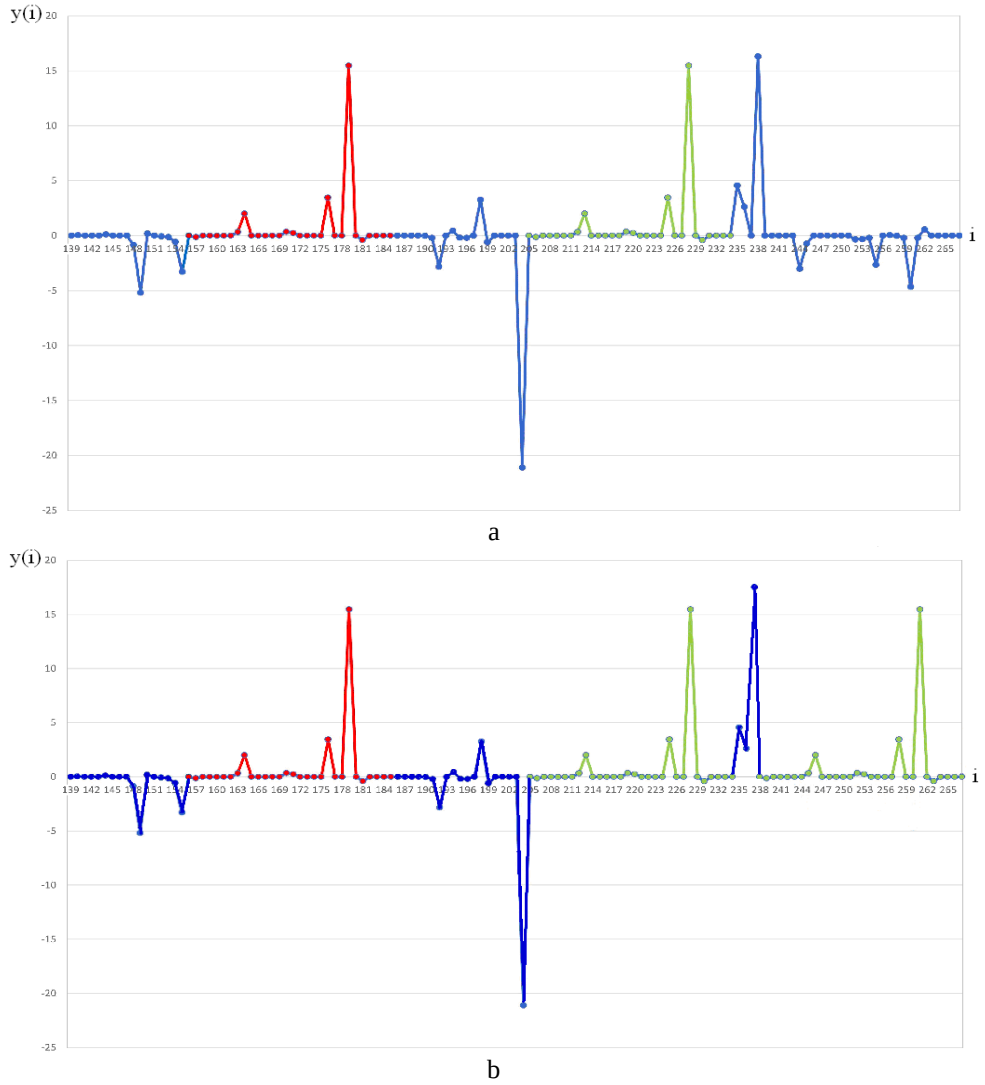


Fig. 4. Graphs of the dependence of the rate of change of the maximum singular value of the 16×16 block of DV frames (Fig. 3), depending on the frame number: a – graph for DV (Fig. 3(b)); b – for DV (Fig. 3(c)).

Table 1

The results of the comparative analysis of the proposed method's effectiveness

Method	ACC, %	Limitations of Applicability
[13]	76-100 (depending on the size of the analyzed frame sub-region: for a 16×16 block – 81%)	Constant-in-Time Substitution Area
[28]	98.7	Presence of synchronized video camera and Wi-Fi signals
New	96.4	-

The computational complexity of the proposed expert method is insignificant: the main computational contribution here will be due to the calculation of the singular value of one key $l \times l$-block in each frame. Thus, the work with each specific frame does not depend on its size and with some approximation can be estimated as $K=const$ operations, then the computational complexity of the method will be proportional to the number of frames n: $O(n)$ operations, which makes the proposed method potentially suitable for work in real

time. The authors of the article found only one analog in open sources with similarly insignificant computational complexity: the [13] method, which is subsequently used in the comparative analysis.

To confirm the possibility of the developed expert method operating in real time, the total time of the method operation with a digital video recording with a specific number of frames was determined during a computational experiment using a non-professional computer (specifications: ASUS Zenbook UX5401EA; processor: 11th Gen Intel(R) Core (TM) i7-1165G7 @ 2.80GHz; RAM: 16.0 GB). Thus, for a digital video recording containing 270 frames, the time of its examination was 00:00:17.8771138, thereby providing a processing frequency of 15 frames per second (fps). According to experts working with CCTV cameras, the frequency of the digital video recording generated by cameras usually fluctuates within the range of 1–30 fps [34]. In this case, the frequency of the digital video recording is selected using a compromise between the requirements for video quality and the weight of the corresponding archive. Since the quality of a video sequence created at the maximum frame rate (30 fps) is practically no different from a video at 15 fps, but at the same time having memory requests that are almost 2 times greater than the second option, the optimal frame rate for almost any video surveillance object is considered to be 10–15 fps [34].

Thus, the developed expert method is capable of identifying the results of a substitution attack, duplication of digital video frames in real time at frequencies practically used in the operation of video surveillance cameras, even under conditions of using non-professional computers, which gives it an advantage over the absolute majority of existing analogues.

The results of the comparative analysis of the proposed method's effectiveness—quantified using the ACC metric [28]—against modern analogs designed for real-time DV integrity analysis are presented in Table 1. Although the developed method is somewhat less effective than the analog [28], it does not require, unlike [28], any additional technical means to perform the analysis, which is its significant advantage.

CONCLUSION

Today, the issues of cybersecurity of critical infrastructure, in particular the energy sector, are among the most important and relevant for ensuring its continuous operation. Most of the information, including control information, when organizing the functioning of critical infrastructure is presented and used today in the form of digital videos, which provide the most complete visual representation of the situation taking place, and therefore are desirable objects of falsification for violators pursuing illegal goals. In this regard, the protection of digital video, in particular, received by video surveillance cameras, from unauthorized access, timely detection of the results of their falsification, if this still happened, is one of the conditions for the normal functioning of critical infrastructure.

This paper proposes a method for identifying the results of temporary attacks of substitution (duplication) of digital video frames. During the development of the method, the concept of a key block of a video sequence was introduced as a block with an insignificant difference in pixel brightness values – homogeneous, remaining such in all digital video frames. The quantitative indicator of the block homogeneity was the normalized separation of its maximum singular value. The proposed expert method is based on the analysis of the rate of change of the maximum singular values of key blocks of video sequence frames. It functions under conditions of minimal technical complexity.

The computational complexity of the developed method does not depend on the size of the digital video frames, but is defined as $\underline{O}(n)$, where n – is the number of frames, which makes it possible to use it to conduct an examination of a video sequence, in particular, received by a video surveillance camera in real time, which has been confirmed in practice. The obtained effect was achieved due to the reasonable possibilities of reducing the analyzed frame region to one small block.

Using the method will increase the reliability and speed of decision-making during video stream monitoring.

References

[1] Röttinger R. Cyberattacks on renewable energies: "how hackers are threatening the energy transition and which technologies can protect

us". *European Journal of Engineering and Technology*, 2024, vol. 12, no.1, pp. 42-49.

[2] Kyrylenko O.V., Denysiuk S.P., Blinov I.V. Digital transformation of the energy industry: current trends and task. *The Proceedings of the Institute of Electrodynamics of the National Academy of Sciences of Ukraine*, 2023, 6, 5-14.

[3] Bobrov Y. Influence of digital technologies and digitalization in the sectors of energy demand and supply. *Scientific Notes of «Krok» University*, 2018, 4, 222-230.

[4] Boi B., Esposito C., Seo J.T. Preventing data tampering in Smart Grids: A blockchain-based digital twin framework. In: Gervasi O. et al. (eds) Computational Science and Its Applications – ICCSA 2024 Workshops. ICCSA 2024. Lecture Notes in Computer Science, vol. 14822. Springer, Cham. https://doi.org/10.1007/978-3-031-65318-6_10

[5] Wawrzyniak N., Hyla T., Popik A. Vessel Detection and Tracking Method Based on Video Surveillance. *Sensors*, 2019, vol. 19, no. 23, 5230.

[6] Guruh Fajar Shidik at al. A Systematic Review of Intelligence Video Surveillance: Trends, Techniques, Frameworks, and Datasets. *IEEE Access*, 2019, vol 7, pp.170457-170473.

[7] El-Shafai W. at al. A comprehensive taxonomy on multimedia video forgery detection techniques: challenges and novel trends. *Multimedia Tools and Applications*, 2024, vol. 83, pp. 4241-4307.

[8] Xuan Hau Nguyen, Thai Son Tran. Video Forgery Detection: State-of-The-Art Review. *International Journal of Recent Research in Mathematics Computer Science and Information Technology*, 2022, vol. 9, Issue 1, pp. 1-9.

[9] Piza E.L. at al. CCTV surveillance for crime prevention: A 40-year systematic review with meta-analysis. *Criminology & Public Policy*, 2019, vol 18(1), pp. 135-159.

[10] Abdulla Salem Al Ashkhari, Norain Ismail. Framework for Assessing the Impact of CCTV Surveillance Systems on Crime Prevention. *Academic Research in Business and Social Sciences*, 2024, pp.1747-1757.

[11] AL Ashkhari and Ismail. Effectiveness of CCTV Surveillance System on Crime Prevention: A Proposed Framework. *International Journal of Business Society*, 2024, vol. 8(4), pp. 904 -913.

[12] Li L., Sun FY., Chen ZT., Tian YJ. Discussion of Intelligent IP Camera Application in Nuclear Power Plant Video Monitoring System. In: Xu Y., Sun Y., Liu Y., Gao F., Gu P., Liu Z. (eds) *Nuclear Power Plants: Innovative Technologies for Instrumentation and Control Systems*. SICPNPP 2020. Lecture Notes in Electrical Engineering, vol 779. Springer, Singapore.

[13] Bobok I., Kobozeva A., Maksymov, Maksymova O. Checking the integrity of CCTV footage in real time at nuclear facilities. *Nuclear & Radiation Safety*, 2016, no.2, pp. 68-72.

[14] Song J.-G., Lee J.-W., Lee C.-K. A Cyber Security Risk Assessment for the Design of I&C Systems in Nuclear Power Plants. *Nuclear Engineering and Technology*, 2012, vol. 44, no. 8, pp. 919–928.

[15] Habeeb R., Manikandan L.C. A Review: Video Tampering Attacks and Detection Techniques. *International Journal of Scientific Research in Computer Science Engineering and Information Technology*, 2019, vol. 5, no. 5, pp. 158-168.

[16] Timothy D.P., Santra A.K. Detecting Digital Image Forgeries with Copy-Move and Splicing Image Analysis using Deep Learning Techniques. *International Journal of Advanced Computer Science and Applications*, 2024, vol. 15, no. 5, pp. 1299-1306.

[17] Lebedeva H. Metod lokalizatsii i identifikatsii original'noi i klonirovannoi oblastei izobrazheniya [Localization and identification method of original and cloned image areas]. *Informatyka ta Matematychni Metody v Modelyuvanni – Informatics and Mathematical Methods in Simulation*, 2014, vol. 4, no. 1, pp. 76–84. (in Ukrainian)

[18] Bobok I.I., Kobozeva A.A., Grygorenko S.M. Method for detecting of clone areas in a digital image under conditions of additional attacks. *Journal of Signal Processing Systems*, 2020, vol. 92, pp. 55–69.

[19] Lin G.-S. and Chang J.-F. Detection of Frame Duplication Forgery in Videos Based on Spatial and Temporal Analysis. *International Journal of Pattern Recognition and Artificial Intelligence*, 2013, vol. 26(07), pp. 1-27.

[20] Himani Sharma at al. An Ontology of Digital Video Forensics: Classification, Research Gaps & Datasets. *Proceedings of International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*, 2019. DOI: 10.1109/ICCIKE47802.2019.9004331.

[21] Vakaliuk T. at al. Vulnerabilities and Methods of Unauthorized Gaining Access to Video Surveillance Systems. In: *CPITS 2023: Workshop on Cybersecurity Providing in Information and Telecommunication Systems*, 2023, pp. 174-181.

[22] Nagothu D. et al. A study on smart online frame forging attacks against Video Surveillance System. Proc. SPIE 11017, Sensors and Systems for Space Applications XII, 110170L (28 May 2019); https://doi.org/10.1117/12.2519005 .

[23] Mohiuddin S., Malakar S., Sarkar R. Duplicate Frame Detection in Forged Videos Using Se-

quence Matching. *Proceedings of Third International Conference «Computational Intelligence in Communications and Business Analytics» (CICBA)*. Santiniketan, India, January 7–8, 2021. P. 29-41.
[24] Kumar V., Kansal V., Gaur M. Multiple Forgery Detection in Video Using Convolution Neural Network. *Computers, Materials & Continua*, 2022, vol. 73(1), pp. 1347-1364.
[25] Li L. at al. Frame Duplication Forgery Detection in Surveillance Video Sequences Using Textural Features. *Electronics*, 2023, vol. 12(22), 4597.
[26] Ren H. at al. Frame Duplication Forgery Detection and Localization Algorithm Based on the Improved Levenshtein Distance. *Scientific Programming*, 2021. Special Issue: Scientific Programming Approaches to Deep Learning for Source Code Transformation.
[27] Loukhaoukha K. Frame duplication forgery detection and localization based on QR decomposition and Minkowski distance. *Journal of Forensic Sciences*, 2025, vol. 70, iss. 4, pp. 1359-1374.
[28] Huang Y. et al. Forgery Attack Detection in Surveillance Video Streams Using Wi-Fi Channel State Information. *IEEE Transactions on Wireless Communications*, 2022, 21(6), 4340-4349. doi: 10.1109/TWC.2021.3129188
[29] Kobozieva A., Bobok I., Kushnirenko N. Steganalysis method for detecting LSB embedding in digital video, digital image sequence. *Proceedings of the 11th Information Control Systems & Technologies (ICST-2023)*. Odesa, Ukraine, 2023. P. 78–90.
[30] Laptiev O.A. at al. The method of spectral analysis of the determination of random digital signals. *International Journal of Communication Networks and Information Security (IJCNIS)*? 2021, vol 13, no 2, pp.271-277.
[31] Bergman C., Davidson J. Unitary embedding for data hiding with the SVD. Available at: https://dr.lib.iastate.edu/entities/publication/bb2b5041-1c92-4ff5-b7f4-ff73c3483eed (accessed 23.09.2022)
[32] Demmel J. *Applied Numerical Linear Algebra.* SIAM, 1997. 430 p.
[33] Gonzalez R.C., Woods R.E. *Digital Image Processing*. Pearson: Upper Saddle River, USA, 2018.
[34] Frame rate for video surveillance, 2016. Available Online: https://stronghold.com.ua/ru/blog/chastota-kadrov-frames-per-second-fps-dlja-videonabljudenija-opisanie-i-vybor-optimalnogo-znachenija (accessed 20 March 2025).

Information about authors.

Alla Anatoliivna Kobozieva
Doctor of Technical Science, Professor. Department of Cybersecurity and Information Protection. Odesa National Maritime University.
Research interests: Steganography, Mathematics in information security
Email:alla_kobozeva@ukr.net
ORCID:0000-0001-7888-0499

Vitalii Anatoliyovych Savchenko
Doctor of Technical Science, Professor. Department of Cybersecurity Management. State University of Information and Communication Technologies. Research interests: Cybersecurity, Information Protection.
E-mail:savitan@ukr.net
ORCID: 0000-0003-4933-7379

Olena Yuriivna Lebedieva
Candidate of Technical Sciences, Associate Professor. Department of Cybersecurity and Information Protection. Odesa National Maritime University. Research interests: Information security
Email: whiteswanhelena@gmail.com
ORCID:0000-0001-5459-0251