Methods of information systems protection

UDC 004.932

Igor Ruban, Nataliia Bolohova, Vitalii Martovtsyki, Oleh Koptsev
Kharkiv National University of Radio Electronics, Kharkiv, Ukraine

DIGITAL IMAGE AUTHENTICATION MODEL

Abstract. The development of new technologies, the growing volume of data and the total consumption of content in the digital environment are changing the ecosystem of modern media. Data can be easily and completely duplicated. It brings great convenience to life, work, scientific research and other areas of human activity. However, information security issues have appeared that have attracted a lot of attention. Copyright ownership is an important aspect of information security, and the use of digital watermarks is an effective way to protect copyrights [1]. Copyright protection is achieved by embedding author information into the digital content. In recent years, many watermarking algorithms have been proposed [1-10].

The Internet is a free zone, where almost everything is open and everything is allowed. The Internet makes it easy to copy and replicate any pictures, texts, video and audio product, without thinking about the fact that these pictures, articles, songs have authors who have certain rights to this multimedia data.

Almost every site in the Internet contains photos and illustrations. At the same time, most of the images posted on the Internet are used without a valid license – unfortunately to authors.

To raise awareness of the extent of image copyright infringement, Copytrack regularly investigates how, where and to what extent images are used illegally. Copytrack’s 2019 Global Infringement Report consists of a statistical analysis of more than 12,000 Copytrack user profiles. Investigation of Illegal image using was based on all searches deemed illegal by individual account holders, and on website owner data based on information gathered by internal search robots. The percentages mentioned in this report refer to the number of potential copyright infringements handled by Copytrack between December 2017 and December 2018. Geographic locations were used for the analysis.

According to the 2019 Copytrack Global Infringement Report [11], Fig. 1 shows the percentage of copyright infringement in the use of images on the global Internet.

Fig. 1. Copyright Infringement Statistics by Continent [11]

Analysis of publications

A digital watermark is a special mark embedded in digital content (called a container) to protect copyrights and confirm the integrity of the document itself. A watermark can be embedded in any type of electronic document. Along with various images (photos, drawings, scanned paper documents, etc.) there are audio recordings that contain a watermark, and video (e.g., DVDs). Watermarks are actively used for placement of unique photos, videos, audio tracks in electronic form on the global Internet.

There are different methods of classification of watermarks algorithms creation. If we divide them by characteristics, there are three types of digital watermarks:
- stable (robust), such watermarks should be resistant to any influences on them [12];
- fragile, changing or collapsing with minor modification of the container [13];
- semi-fragile, resistant to some influences and unstable to others [13].

Resistant watermarks are used when it is necessary to remain the identification code, company logo and other identifiers intact with maximum distortion of the con-
After analyzing the current state of research on methods of superimposition of watermark and significant parameters of watermark, we can form the following assessment of the system of digital image authentication effectiveness:

$$EF = R \cdot \alpha_r + SR \cdot \alpha_{sr} + ER \cdot \alpha_{er} + SC \cdot \alpha_{sc} + DT \cdot \alpha_{dt},$$

where $R, R \in [0,1]$ is an estimate of the reliability of the method of embedding a watermark; $RS, RS \in [0,1]$ is an estimate of the watermark invisibility in the image; $ER, ER \in [0,1]$ is the probability of an error of the first and second kind; $SC, SC \in [0,1]$ is an estimate of the fragility of the watermark; $DT$ is the number of embedded watermarks; $\alpha_r, \alpha_{sr}, \alpha_{er}, \alpha_{sc}, \alpha_{dt}$ are significance coefficients of the corresponding parameters of the watermark method. Such coefficients are necessary because there is no universal watermark embedding method, so thanks to such coefficients you can adjust the significance of each parameter and thus influence the final effectiveness of the method for a particular watermarking task.

The process of digital image authentication with watermark is shown in Fig. 2 and consists of the following main steps:

1) defining the area for embedding;
2) generating a watermark;
3) embedding the watermark in an image fragment;
4) image preprocessing after embedding the watermarked fragment in the original image;
5) detecting the watermarked fragment;
6) extraction of the watermark from the fragment;
7) correcting errors in the watermark during extraction;
8) obtaining the label of the right holder.

Consider these steps.

**Step 1 - Definition of the area for embedding.**
Suppose there is an image $Im \in [N, M]$ in which you want to embed a watermark $W[k, l]$.

Then, in the simplest case, the procedure for determining the optimal area for embedding a watermark can be represented using the sliding window method.

---

**Digital Image Authentication Model**

Model of digital image authentication can be considered as steganographic system that transmits an encrypted identifier, which is a digital watermark.

---

**Fig. 2. Digital Image Authentication Model**
The scheme of the algorithm for determining the area for embedding is shown in Fig. 3. Analytically, the sliding window processing can be represented as follows:

\[
E(Im_{i,j}) = \mathbb{Q} \left[ q(Im_{i,j}), q(Im_{i,j+1}), q(Im_{i,j-1}) ... \right] = \mathbb{Q} \left[ q(Im_{i+k,j}), Im_{i+j} \right] \left\{ k, l \right\} \in S,
\]  

(2)

where \( E(Im_{i,j}) \) – the optimality value of the embedding area; \( Q \) – function describing the rules for evaluating the pixels in the vicinity \( S \); \( S \) – point vicinity, the set of points (pixels) surrounding the working point (usually the center pixel); \( k, l \) – the size of the sliding window, given by the \( S \) set of coordinate offsets on the abscissa axis and the ordinate axis, respectively.

**Stage 2 - Generation of the watermark.**

Let \( W', I', K', B' \) be the set of possible watermarks, containers (the form of watermark representation) of keys and hidden identifiers of right holders, respectively. Then the generation of the watermark can be represented as:

\[
F: I' \times K' \times B' \rightarrow W, W = F(I, K, B),
\]

(2)

where \( W, I, K, B \) – are elements of the corresponding sets. Generally speaking, the function can be arbitrary, but in practice, the robustness requirements of the watermark impose certain restrictions on it. Thus, in most cases, \( F(I, K, B) = F(I + \alpha, K, B) \), i.e., a slight change in the container does not change the hidden IDs of the right holders. The function is usually composite:

\[
F = T \circ G, \quad \text{where } G: K' \times B' \rightarrow C' \quad \text{and} \quad T: C' \times I' \rightarrow W',
\]

(3)

that is, the watermark depends on the properties of the container. The function \( G \) can be implemented using a cryptographically secure pseudorandom number generator with \( K \) as the initial value.

To improve the robustness of the watermark, authors use interference-resistant codes such as BCH codes, convolutional codes. The optimality function is usually composite:

\[
T(C, I_0) = T(C, I_W) = T(C, I_W^\ast),
\]

(4)

**Step 3 - Generation of the watermark.**

The process of embedding the watermark \( W(i, j) \) in the original image \( I_0(i, j) \) can generally be described as a superposition of two signals:

\[
e: I' \times W' \times L' \rightarrow I_W, \quad I_W(i, j) = I_0(i, j) \oplus L(i, j)W(i, j) p(i, j),
\]

(5)

where \( L(i, j) \) - watermark-embedding mask, which takes into account the characteristics of the human visual system, serves to reduce the visibility of the watermark; \( p(i, j) \) - projecting function;

Sign \( \oplus \) denotes the superposition operator, including, addition, truncation, and quantization.

The projecting function carries out the "distribution" of the watermark over the image area. Using of it can be considered as an implementation of parallel channel information spacing. In addition, this function has a certain spatial structure and correlation properties used to counteract geometric attacks.

**Step 4 - Image preprocessing after embedding the watermarked fragment in the original image.**

At this stage, the image fragment \( I_W \) is inserted into the original image \( I_m \). Since \( I_W \) is a modification of the original fragment \( I_0 \), which serves as a container for the watermark. When embedding it into the original image \( I_m \), the watermarked fragment will differ from the general distribution of the image brightness \( I_0(i, j) + \alpha \), where \( \alpha \) is a modification factor that occurs during the embedding process. To eliminate this drawback, after embedding a fragment with a watermark in the original image, anti-aliasing is applied along the edges of the embedded fragment to eliminate the visibility of the presence of a watermark in the image.

**Step 5 - Detect the watermarked fragment.**

The task of detection is to find objects (watermark) with certain properties on the image, and if the objects are
detected, to determine their coordinates on the image plane. The basic principle of object detection on the image is to compare the brightness function of the image with some "reference" - a fragment of the brightness field containing the desired object. When implementing the detection procedure, the standard is sequentially moved along the image field, and at each position, its similarity to the real brightness function on the fragment is investigated. Complete coincidence of the standard and the image, as a rule, does not happen due to noise and distortion, as well as because usually there is no complete information about the shape and structure of the object, (we have to use the standard, only approximately describing the object).

Since watermarks are hidden and different in structure in the image in such a way that they are less visible, it is best to use neural networks for detection and localization.

**Step 6 - Extracting the watermark from the fragment.** The extraction uses the inverse of the operation in step 3 and depends on the embedding method.

**Step 7 - Correcting errors in the watermark during extraction.** Since it is recommended to use interference-resistant codes, such as BCH codes, convolutional codes, to improve robustness, this step is the correction of distortions in the watermark.

**Step 8 - Obtaining the label of the copyright holder.** This step is the reverse of step 2 and depends on the method used to generate the watermark.

There is a probability that the decoder will not detect an existing watermark and a probability of falsely finding a watermark in an empty container (false alarm probability). Decreasing one probability leads to increasing the other probability.

Reliability of the decoder is characterized by the probability of false detection. This model for verifying the authenticity of a digital image is designed to minimize the likelihood of both errors, since each of them can lead to denial of service.

**Conclusion**

This article proposes a model for reliable verification of digital image authenticity with a high degree of protection and parameters for assessing the effectiveness of such systems. Reliability is achieved because the watermark is hidden not in the whole image, but in its fragment, which is most suitable for hiding the image, as well as for using anti-noise codes as a watermark. Based on the current state of watermarking methods, it is recommended to use modern algorithms and architectures of convolutional neural networks to ensure a high degree of security.

**REFERENCES**


Надійшла (received) 23.11.2020
Прийнята до друку (accepted for publication) 10.02.2021
Модель аутентифікації цифрових зображень

І. В. Рубан, Н. Н. Бологова, В. О. Мартовицький, О. О. Концев

Анотація. Розвиток нових технологій, зростання обсягів великих даних і тотальне споживання контенту в цифровому середовищі змінюють екосистему сучасних медіа. Дани можна легко і повністю дублювати, внаслідок чого виникає потреба в надійних методах підтвердження ідентичності контенту. Основними головними критеріями надійної комунікації є безпека, інформаційна безпека, надійність та безпека. Для цього можна використовувати цифрові водяні знаки (ЦВЗ), які дозволяють скривати індивідуальні ідентифікатори в постійній частині зображення, матуючи високу кібербезпеку.

Ключові слова: цифровий водяний знак; аутентифікація; авторське право; відкриті архітектури; кібербезпека.