Volodymyr LARIN[1], Maksim PAVLENKO[2]

Supervisor: Oleksandr TYMOCHKO[3]

ULEPSZONA METODA KODOWANIA PRZEWIDYWALNYCH RAMEK Z RACHUNKIEM MECHANIZMÓW ZARZĄDZANIA W SYSTEMACH KOMPUTEROWYCH

Streszczenie: Główne trudności w pracy z wideo to duże ilości przesyłanych informacji i wrażliwość na opóźnienia w transmisji informacji wideo. Dlatego w celu wyeliminowania maksymalnej redundancji w tworzeniu sekwencji wideo stosuje się 3 rodzaje ramek: I, P i B, które tworzą grupę klatek.

Słowa kluczowe: obraz, technologie, metoda, systemy wideo, systemy komputerowe

IMPROVED METHOD OF CODING PREDICTABLE FRAMES WITH THE ACCOUNT OF MANAGEMENT MECHANISMS IN COMPUTER SYSTEMS

Summary: The main difficulties in working with video are large volumes of transmitted information and sensitivity to delays in the video information transmission. Therefore, in order to eliminate the maximum redundancy amount in the formation of the video sequence, 3 types of frames are used: I, P and B which form a frame group.

Keywords: image, technologies, method, videoinformation systems, computers systems

1. Introduction

However, taking into account the number of P-frames in the group, they make the main contribution to the total video data amount. Therefore, the possibility of upgrading coding methods for P-frames is considered on preliminary blocks' type

[1] Ivan Kozhedub Kharkiv National Air Force University, department ACS, speciality ACS, email l_vv83@ukr.net
[2] Ivan Kozhedub Kharkiv National Air Force University, department ACS, speciality ACS, email bpgpma@ukr.net
[3] Doctor of Technical Sciences, Professor, Ivan Kozhedub Kharkiv National Air Force University, department ACS, email timochko.alex@gmail.com

identification with the subsequent formation of block code structures [1].

It has been analyzed the predicted frames processing in the MPEG standard and its drawbacks has found during compressing video data.

A method for encoding P-frames with the ability to select the type of blocks processing has been developed, where the structural redundancy estimation of the block is used as a decision rule. Thus, it has been introduced an additional possibility to control the video bit rate by changing the number of I- and P-type blocks [2].

The method of coding and reconstructing predicted frames has been improved by using block coding, which unlike Huffman codes has more noise immunity and less bit and time costs when processing data blocks.

2. Justification of the requirements for video data compression systems in the computer systems

Delays in the processing and transmission of objective video monitoring information lead to the following fact [3]. The temporary standards for the transportation information duration by computer systems are not maintained both under normal conditions and during the increased load. This leads to delays in the flow of information on average by 10-15%. Also, this leads to untimely tasks fulfillment for managing and long delays in assessing the situation and making decisions [4]. Hence, it can be concluded that methods with loss of quality provide a greater compression level than methods without loss of quality. However, methods without quality loss, as well as methods with quality loss, must be used to process and transmit video information to computer networks. This need is explained by the following reasons:

1) for some practical tasks of traffic management, it is required to provide high quality images. At the same time, methods with quality loss, based on the reduction of psychovisual redundancy, can not guarantee to ensure high quality images;

2) wide use of methods without quality loss in various image formats (TIFF, GIF, BMP, ART, etc.). At the same time, these formats are characterized by an increase in the compressed volume for realistic images heavily saturated with small details [5];

3) compression methods with quality loss have a longer data encoding time than methods without (dozens of times), and in some cases this leads to a loss and in the total time for processing and data transfer;

4) developed a group of methods aimed primarily at compressing color coordinates, while arrays of series lengths are transmitted without additional coding, which reduces the image's degree compression and increases the transmission time on the communication channel;

5) in many complex compression methods, the length methods of the series are used as a subsystem for compressing the service information. Therefore, for further increase the compression ratio, it is required to compress the length of the series [6].

At the same time, existing methods without quality loss do not provide the transfer of images in the computer networks in real time. This is due to the following shortcomings of the length methods of the series and the LZW (Lempel-Ziv-Welch) method:

1. The presence of stationary background regions is not taken into account.

2. For the dynamic component:

- the amount of the simplest structural redundancy decreases (the lengths of series of the same elements decrease) for realistic images, and consequently, the effectiveness of the RLE (Run-Length Encoding) and LZW methods decreases;

- there is a problem of the color coordinates processing of the identical images series.

Basing on the analysis of known compression methods, it can be concluded that there is a need to use methods for processing images in the computer networks without quality loss. At the same time, they do not allow to process and transmit video data in real time.

It will be analyzed the directions of developing compression methods without quality loss in order to improve the processing, transmission and video data's noise immunity in the presence of stationary background images.

Reducing the total processing time and transfer information can be done due to further improvement of existing compression methods. From the analysis of known compression methods, there follow such variants of their improvement:

1. Using processing, taking into account the stationary background presence of images. Here are the options:

- identifying motion compensation;

- using three-dimensional discrete cosine transformations.

However, this is a computationally complex processing process.

2. Using additional processing at the individual frames level, i.e. to exclude intraframe redundancy. Here the following options are possible:

1) using the series' lengths encoding. This, on the one hand, makes it possible to improve the identifying structural patterns efficiency. On the other hand, the coding efficiency of the series lengths is sharply reduced in the case of multi-degradation data processing with a high brightness difference probability;

2) as a result of image compression, by using RLE and LZW methods, structural and statistical redundancy is reduced. Therefore, the additional use of methods that reduce statistical redundancy will not significantly affect on the compression ratio. Existing element-wise coding methods, basically, exclude statistical redundancy;

3) the series lengths carry basic information about the shapes and sizes of the image objects, which is crucial for correct image recognition. Even small distortions in the lengths' values of the series lead to a partial or complete image destruction of the objects. Therefore, it is not recommended to use existing methods, which are based on the reduction of psychovisual redundancy to compress the series lengths;

4) increasing the LZW methods effectiveness is associated with an additional increase in the number of operations for coding, spent on:

- statistical characteristics calculation of image elements in each "sliding window";

- large window sizes lead to an increase in the search time of elements in the window.

In addition, the representation of the series' lengths, by uneven codes, slightly increases the compression ratio (not more than 1.5 times). But at the same time additional difficulties appear that make it difficult to implement the compression method in practice:

- additional operations are required to calculate the statistical characteristics of each block (of the order 0 ($N\log_2 N$));

- it is necessary to transmit data on statistical characteristics;

-if an error in the codeword, it is impossible to restore the whole message.

It means, that the using of existing methods, which exclude probability-statistical and psychovisual redundancy, to further enhance the effectiveness of methods without quality loss in the process of intraframe processing is inexpedient.

Therefore, it came an interest for investigate the possibility of further increasing the compression ratio of stationary background processing images, which are based on the stationary component detection of the frame in the substrate, which will allow:

- to identify the area of stationary relative to the previous frame;
- to form a binary mask of dynamic areas and thereby provide the potential for reducing structural redundancy as a result of identifying the lengths of binary series;
- to reduce the dimension of the array containing elements of dynamic objects.

Thus, it is proposed to construct a method for compressing images of a stationary background, which is based on the following mechanisms:

1) an identifying area of stationary background, which will allow to take into account the presence of interframe redundancy between neighboring frames;

2) a separated processing of the allocated stationary and dynamic components, using the operation of imposing a binary mask.

Concepts construction on the formation of stationary background's dynamic images compressed representation

The frame of the differential representation is formed on the current and previous frames in the conditions of the video information's stationary formation. This process is given by the expression:

$$e_{i,j}^{(\xi+1)} = a_{i,j}^{(\xi)} - a_{i,j}^{(\xi+1)}, \qquad (1)$$

here are

$a_{i,j}^{(\xi)}$ - (i; j) - element of the previous frame;

$e_{i,j}^{(\xi+1)}$ - (i; j) - element of the current differential-represented frame;

$a_{i,j}^{(\xi+1)}$ - (i; j) - element of the current frame in the video sequence.

The format of the differentialy-represented frame allows to distinguish two components, that describe the stationary background and dynamic objects. In order to obtain a stationary background, it is necessary to determine the positions of the elements, which belong to the dynamic component. In this case, it is necessary to take into account that the frame is formed in conditions, when the illumination of railway trains in motion changes, the video camera vibrates, as a result of fluctuations during it's fixing, distortions are detected during conversion of the analog signal to digital form.

In order to take this into account in the process of forming the stationary component of the differential-represented frame, it is proposed to use the filtering threshold ΔP. The filtering essence is to identify the elements of the differential-represented frame, which contain impulsive interference. This will determine whether the element refers to a stationary component or to a dynamic component. To reduce the number of operations for processing, it is suggested to use scalar threshold filtering. Scalar threshold filtering is performed according to the rule, where each processed element

e_{ij} is compared with the threshold value ΔP. If the value of the element is less than the threshold value, i.e. $e_{ij} \le \Delta P$. Then this element is considered an element of the stationary component $e_{ij} \in I_S$, and $e_{ij} = 0$. Conversely, if the value of the element exceeds the threshold value $e_{ij} > \Delta P$, then this element belongs to the dynamic component $e_{ij} \in I_d$.

The filtered structure of the differential-represented frame allows you to extract dynamic and stationary components from it.

Taking into account the separate processing for the differential-represented frame, two components are formed (Figure 1).

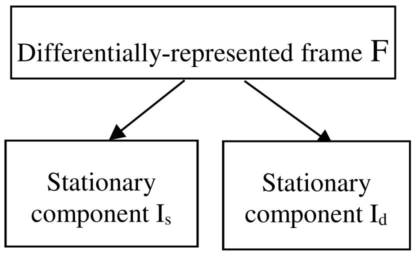

Figure 1. The scheme for the formation of two differential-represented frame components

The filtered representation of the differential-represented frame is given by the following system of expressions:

$$e'_{i,j} = \begin{cases} 0, \rightarrow \left|e_{ij}\right| \le \Delta P; \\ e_{i,j} - \Delta P, \ \left|e_{ij}\right| > \Delta P \ \& \ e_{ij} > 0; \\ e_{i,j} + \Delta P, \ \left|e_{ij}\right| > \Delta P \ \& \ e_{ij} < 0. \end{cases} \quad (2)$$

Thus, the differential representation of the frame, by using the filtering threshold, makes it possible to identify stationary background regions (stationary component) and regions containing dynamic objects (dynamic component).

Substantiation of the constructing basic principles the technology of compact video frames representation, which are presented in a differential form

The most preferable method for constructing the compact representation technology of the binary masks of frames represented in a differential form is the approach. Which is based on the identification and description of the lengths of one-dimensional binary series. A binary series is a consecutive binary elements sequence with the same value. In this case, sequences of identical binary elements are replaced by their lengths. And since the elements of the binary masks Q of the frames represented in the differential form take only two possible values 0 or 1, it is suggested to form the lengths of the binary series without indicating their level.

In order to correctly identify the alternating lengths of the series on the receiving side, it is suggested to assume, that each line in the binary mask of the frame represented in the differential form is preceded by a zero elements series of length 1. In this case, the binary series' length will be denoted by $r_{i,\varphi}$ - length binary series formed for the φ binary mask row of the frame. Formation of binary series' lengths in the acceptable restrictions has the following appearance:

1) because $q_{01}=0$ - the binary element preceding the first element of the binary mask's first line, then, if $q_{11}=0$, then $r_{i,1}=2$. Otherwise $r_{i,1}=1$;

2) assume that $r_{i,\varphi}=\gamma$. Then, if $q_{i,r(i/\phi)+\gamma}=q_{i,r(i/\phi)+\gamma+1}$, then the length of the series increases by one $r_{i,\varphi}=\gamma+1$. On the contrary, for $q_{i,r(i/\varphi)+\gamma}\neq q_{i,r(i/\varphi)+\gamma+1}$ a new series is being formed $r_{i,\varphi=1}=1$.

The following notation is used:

$r_{i/\varphi}$ - the number of binary elements of the i-th row of the binary mask of the frame preceding the beginning of the φ binary series (i.e., the number of binary elements $\sum_{\phi=1}^{\varphi-1} r_{i,\phi}$ contained in the $\Phi_{i/\varphi}$ binary series);

$\Phi_{i/\varphi}$ - the binary series quantity formed for the i-th row of the binary mask of the frame, which precede the i-th series;

$\sum_{\phi=1}^{\Phi_i-1} r_{i,\phi}+r_{i,\Phi_i}=m_M$ - the binary elements quantity in the array row of the binary mask of the frame;

Φ_i - the binary series quantity formed for the i-th row of the binary mask of the frame, $1\leq\varphi\leq\Phi_i$.

The formation of binary series stops, when the element q_{m_M,n_M} is processed.

The array's compression degree of the frame's binary mask, with the detection of the binary series' lengths, is determined by the formula:

$$\eta_M=\frac{m_M\,n_M}{L_M}\ , \tag{3}$$

where L_M is the number of bits allocated to the representation of the sequence of the binary series' lengths is formed for the differential frame's binary mask array, which is given by the expression:

$$L_M = \sum_{i=1}^{n_M} \sum_{\varphi=1}^{\Phi_i} \ell og_2 r_{i,\varphi} , \qquad (4)$$

here $\ell og_2 r_{i,\varphi}$ is the number of digits for the series representation $r_{i,\varphi}$.

It is clear, that, in the general case, expression (3) can not be used to estimate the compression degree. This is due to the fact, that the quantities values $r_{i,\phi}$ will not be equal to each other. Hence, their representation will require a different number of digits. At the same time, if the number of bits for each series is not known on the receiving side, this will lead to ambiguous reconstruction process.

3. Problem solution

The following strategies are possible for allocating the number of bits for the series lengths:

The first approach. Select the number of digits from the predefined maximum length $r_{\max}$ of the series for the binary mask of the differential frame (or for the entire bit representation of the stationary component). In this case, for the code representation of the binary series' lengths, the same number of digits (this mode is called uniform) is equal $\ell og_2 r_{\max}$. It is clear, that the compression ratio is found by the formula:

$$\eta_M = \frac{m_M \, n_M}{\Phi \;\; \ell og_2 r_{\max}} \qquad (5)$$

where Φ is the number of binary series formed for the frame's binary mask.

Here are some disadvantages:

- between the length's minimum and the preselected maximum value of the bit series, there may be a significant difference;
- the values' unevenness of the series' lengths is not taken into account;
- there is ambiguity in the choice of the quantity $r_{\max}$.

This leads to a non-rational distribution of the discharges, and as a consequence, an increasing in the initial volume for a representing the binary mask of the frame.

The second approach. It performs an uneven distribution of the number of bits under the binary series' length (the encoding mode is called non-uniform). The possible presence of the probabilities uneven distribution of the series lengths' values is taken into account. In order to implement such strategy, arithmetic coding is used. In arithmetic coding, there is no one-to-one correspondence between source symbols and code words. Instead, the whole sequence of source symbols (i.e. the entire vector of binary series) is correlated with one arithmetic code word. The code word itself sets the interval of real numbers between 0 and 1. As the number of symbols in the message increases, the interval necessary for their representation decreases, and the number of information units required to represent the interval increases. Each character in the message reduces the size of the interval in accordance with the probability of its appearance.

4. Conclusion

The magnitude of the binary mask's compression ratio of the differential frame for the proposed encoding technology depends on:

- probability of changing binary sequences;
- accuracy of a statistical model's construction of the binary series' lengths source;
- marker lengths between non-uniform code structures.

For a binary differential frame mask, these characteristics depend on the availability of dynamic objects and the number of such dynamic objects. The more saturation of objects with a dynamically changing scene, i.e. the smaller correlation between adjacent frames, the more heterogeneous and non-stationary statistical and structural characteristics of the differential frame's binary mask. In this case, the compression ratio, which is based on the non-uniform statistical coding, will be minimal.

REFERENCES

1. LARIN V., YEREMA D., BOLOTSKA Y.: The reasoning of necessity enhancing video privacy in conditions of providing the quality of the video information service provided in virtual info communication systems. Системи озброєння і військова техніка 2(2019)35. Kh. ХНУПС. 158-162. *http://www.hups.mil.gov.ua/periodic-app/article/19290.*
2. QASSIM H., VERMA A., FEINZIMER D.: Compressed residual-VGG16 CNN model for big data places image recognition. 2018 IEEE 8th Annual Computing and Communication Workshop and Conference (CCWC) 2018. DOI: *https://doi.org/10.1109/ccwc.2018.8301729.*
3. LARIN V., YEROMINA N., PETROV S., TANTSIURA A., IASECHKO M.: Formation of reference images and decision function in radiometric correlation-extremal navigation systems. Eastern-European Journal of Enterprise Technologies. 4(2018)9, 27-35. DOI: *https://doi.org/10.15587/1729-4061.2018.139723.*
4. LI L.: The UAV intelligent inspection of transmission lines. Proceedings of the 2015 International Conference on Advances in Mechanical Engineering and Industrial Informatics 2015. DOI: *https://doi.org/10.2991/ameii-15.2015.285.*
5. GONZALES R.C. Digital image processing / R.C. Gonzales, R.E. Woods. Prentice Inc. Upper Saddle River, New Jersey, 2002. – 779 p. *http://web.ipac.caltech.edu/staff/fmasci/home/astro_refs/Digital_Image_Processing_2ndEd.pdf.*
6. KHARCHENKO V., MUKHINA M.: Correlation-extreme visual navigation of unmanned aircraft systems based on speed-up robust features, Aviation 18(2014)2, 80–85. DOI: *https://doi.org/10.3846/16487788.2014.926645.*