

## АНОТАЦІЯ

Дібрівний О.А. Методика підвищення ефективності передачі відеопотоку при стисненні методом компенсації руху. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 123 – «Комп'ютерна інженерія». – Державний університет телекомунікацій. – Київ, 2021.

Дисертаційна робота присвячена дослідженню методики компенсації руху (motion compensation), яка використовується в більшості сучасних технологіях стиснення відеопотоку і базується на передачі кадрів у вигляді скомпенсованої міжкадрової різниці (у вигляді векторів руху окремих блоків на які розбивається зображення), з ціллю зменшення часу затраченого на стиснення відеопотоку.

Проаналізовано поточний стан та перспективи розвитку технологій стиснення відеофайлів, визначено основні алгоритми та методики, що входять до їхнього складу. Виявлено основні проблемні області таких технологій.

Розробка результуючої методики відбувалась за рахунок усунення проблем сучасних методик в двох ключових напрямках:

1. Оцінка схожості блоків зображення
2. Методи блочного пошуку схожих блоків.

На основі аналізу наукової літератури було визначено основні метрики, які використовуються для оцінки ступеня схожості зображень. Було виявлено, що в сучасних системах відеостиснення в якості основної метрики використовується SAD (Sum of absolute differences – сума абсолютних різниць), обчислення якої, при всій її простоті та невеликій обчислювальній

складності може займати від 40 до 80% загального часу кодування відеопотоку.

Для усунення даного недоліку було запропоновано використання шаблонів порівняння за характерними точками, які дозволяють зменшити обчислення метрики за рахунок зменшення кількості точок порівняння. Такий підхід зумовлений унікальністю розподілу піксельних значень на різних частинах зображення. Аналіз алгоритму оцінки схожості зображень на основі характерних точок, показав, що зменшення кількості контрольних точок призводить до різкого зменшення показника PSNR (peak signal to noise ratio – пікового відношення сигналу до шуму). Це відбулося за рахунок втрати інформації в областях з високою ентропією зображення та появи стійкої невизначеності в областях з низькою ентропією зображення.

Запропоновано для усунення втрати інформації в областях високої ентропії (на границях об'єктів та областях різкого переходу кольору) використання алгоритмів виділення країв зображення і проведення уточнення блоків кандидатів на карті країв зображення. Зручність такого підходу зумовлена тим, що процес пошуку векторів руху відбувається для відеопотоку в форматі YUV, а саме для яскравішої компоненти зображення (Y) та частковим ігноруванням кольорових компонент (людське око більш чутливе до зміни яскравості зображення чим до зміни кольору). При цьому яскравіша компонента Y є зображенням в сірих тонах. Такий підхід дозволив суттєво скоротити втрати PSNR при відносно незначних зменшеннях швидкості обчислення векторів руху.

Вдосконалено методику існуючих методів блочного пошуку, для усунення основних недоліків використання шаблонного порівняння. Було проведено аналіз сучасних методів блочного пошуку таких як: FS (full search), TSS (three step search), NTSS (New Three Step Search), DS (diamond search), MVFAST (Motion Vector Field Adaptive Search Technique), PMVFAST (Predictive Motion Vector Field Adaptive Search Technique) та виявлено їх основні недоліки. Формування методу блочного пошуку, для результуючої

методики було проведено за рахунок вдосконалення чотирьох основних частин алгоритмів блочного пошуку:

1. Попередня обробка відеокадру, шляхом обрахунку некомпенсованої різниці з попереднім кадром, та занулення векторів руху блоків, максимальне значення якої не перевищує заданий поріг (2% глибини кольору), для усунення неоднозначності в областях низької ентропії.
2. Формування розширеного набору більш точних предикатів, для скорочення часу уточнення блоків кандидатів та виключення впливу випадкових співпадінь значень пікселів на формування вектору руху.
3. Розширення умов формування порогових значень для умов ранньої зупинки пошуку.
4. Формування набору адаптивних шаблонів для уточнення блоків з яких будуть формуватися вектори руху.

Основним підходом для формування цих чотирьох частин методу було використання суміжних блоків не тільки поточного та попереднього кадрів, а також наступного кадру, за рахунок попередньої обробки частини блоків всіх кадрів відео послідовності (пар кадрів у випадку обробки відеопотоку в реальному часі). Такий підхід обумовлений постійністю руху об'єктів на коротких проміжках часу.

В результаті комбінації шаблонного порівняння для оцінки схожості блоків, алгоритмів виділення країв та вдосконаленого методу блочного пошуку було сформовано результуючу методику знаходження векторів руху - TFEMVFAS (Three by Frame Edge Map Motion Vector Field Adaptive Search – адаптивний алгоритм визначення векторів руху один через три за картою границь), в двох модифікаціях:

1. TFEMVFASS (- TFEMVFA speed search – швидкий пошук) – основною ціллю якої є максимальний виграш в швидкості, за

рахунок більш ранньої зупинки уточнень блоків кандидатів, наслідком чого є більш суттєві втрати значення PSNR.

2. TFEMVFAAS (- TFEMVFA accuracy search – точний пошук) – який за рахунок проведення обов'язкових уточнень по блокам на які вказує набір предикатів, що дозволяє помітно скоротити втрати PSNR, за рахунок втрат в швидкості роботи методики.

Протестовано модифікації методики згідно комплексних рекомендацій тестування міжнародної комісії MPEG по оцінці складних відеоінструментів, на прикладі 13 відеопослідовностей, з різними бітрейтами. Досліджено вплив збільшення розмірів блоків на які розбивається зображення на значення PSNR та максимального прискорення. Проведено порівняння результатів роботи запропонованої методики з сучасними методами формування векторів руху: FS, TSS, NTSS, DS, MVFAST, PMVFAST, в результаті визначено, що запропонована методика дозволяє отримати прискорення на рівні 15-17% (рівень прискорення залежить від бітрейту оброблюваної відеопослідовності: вищий бітрейт – вище прискорення) від найближчого аналога при втратах на рівні 0.3-1.2% значень PSNR (втрати PSNR залежать від розміру блоків для яких проводився пошуку, максимальна втрата для блоків 16x16 пікселів, мінімальна для блоків 64x64 пікселів) від значень еталонного методу пошуку векторів руху для модифікації A і прискорення на рівні 27-31% при втраті 1.1%-2.1% в залежності розміру блоку та бітрейту відеопослідовності, для модифікації S.

Розроблена методика знаходження векторів руху є цінною для розвитку технологій обробки, передачі та зберігання відеофайлів.

**Ключові слова:** стиснення відеопотоку, компенсація руху, оцінка схожості зображень, сума абсолютних різниць, виділення границь об'єктів, LoG, Canny, методи блочного пошуку.

## **ABSTRACT**

Dibrivny OA The method of increasing the efficiency of video stream transmission during compression by the method of motion compensation. - Qualifying scientific work on the rights of the manuscript.

Dissertation for the degree of Doctor of Philosophy in specialty 123 - "Computer Engineering". - State University of Telecommunications. - Kyiv, 2021.

The dissertation is devoted to the research of motion compensation technique, which is used in most modern technologies of video stream compression and based on frame transmission in the form of compensated interframe difference (in the form of motion vectors of separate blocks into which the image is divided), in order to reduce compression time of video stream.

The current state and prospects of development of video file compression technologies are analyzed, the main algorithms and techniques that are part of them are determined. The main problem areas of such technologies are identified.

The development of the resulting methodology took place by eliminating the problems of modern techniques in two key areas:

1. Similarity estimation of image blocks
2. Methods of block search of similar blocks.

Based on the analysis of the scientific literature, the main metrics used to assess the similarity degree of images were identified. It was found that in modern video compression systems as the main metric used SAD (Sum of absolute differences), the calculation of which, for all its simplicity and low computational complexity can take from 40 to 80% of the total coding time of the video stream.

To eliminate this shortcoming, it was proposed to use comparison templates based on characteristic points, which allow to reduce the calculation of the metric by reducing the number of comparison points. This approach is due to the unique distribution of pixel values in different parts of the image. When evaluating the performance of the algorithm for estimating the similarity of images based on

characteristic points, it was determined that reducing the number of control points leads to a sharp decrease in PSNR (peak signal to noise ratio), due to loss of information in areas with high image entropy and the emergence of persistent uncertainty in areas with low image entropy.

To eliminate the loss of information in areas of high entropy (at the boundaries of objects and areas of sharp color transition), it was proposed to use algorithms for selecting the edges of the image and refining the candidate blocks on the map of the edges of the image. The convenience of this approach is due to the fact that the process of finding motion vectors is for the video stream in YUV format, namely for the luminance component of the image (Y) and partial ignoring of color components (human eye is more sensitive to changes in image brightness than to color change). Y is essentially a grayscale image (image edge selection algorithms work with grayscale images). This approach has significantly reduced the loss of PSNR with relatively small reductions in the speed of calculation of motion vectors.

The next step in the formation of the resulting technique is to improve existing methods of block search, to eliminate the main disadvantages of using a template comparison. An analysis of modern block search methods such as: FS (full search), TSS (three step search), NTSS (New Three Step Search), DS (diamond search), MVFAST (Motion Vector Field Adaptive Search Technique), PMVFAST (Predictive Motion Vector Field Adaptive Search Technique), and identified their main shortcomings was performed. The formation of the block search method for the resulting technique was carried out by improving the four main parts of block search algorithms:

1. Pre-processing of the video frame, by calculating the uncompensated difference with the previous frame, and zeroing the motion vectors of the blocks, the maximum value of which does not exceed a specified threshold (2% color depth), to eliminate ambiguity in low entropy areas.
2. Formation of an extended set of more accurate predicates, to reduce the time of refinement of candidate blocks and to exclude the influence

of random coincidences of pixel values on the formation of the motion vector.

3. Extension of the conditions for the formation of threshold values for the conditions of early search.

4. Forming a set of adaptive templates to specify the blocks from which the motion vectors will be formed.

The main approach for the formation of these four parts of the method was to use adjacent blocks not only of the current and previous frames, but also the next frame, by pre-processing part of the blocks of all frames of the video sequence (pairs of frames in real time video stream processing). This approach is due to the constant movement of objects over short periods of time.

As a result of a combination of template comparison for estimating block similarity, edge selection algorithms and an improved block search method, the resulting method of finding motion vectors - TFEMVFAS (Three by Frame Edge Map Motion Vector Field Adaptive Search) was formed), in two versions:

1. TFEMVFASS (TFEMVFA speed search) - the main purpose of which is the maximum gain in speed, due to the earlier stop of refinements of blocks candidates, resulting in more significant loss of PSNR value.

2. TFEMVFAAS (TFEMVFA accuracy search) - which is due to the mandatory refinement of the blocks indicated by the set of predicates, which can significantly reduce the loss of PSNR, due to losses in the speed of the method.

Both modifications of the method were tested according to the comprehensive testing recommendations of the international commission MPEG for the evaluation of complex video tools, on the example of 13 video sequences, with different bitrates. The effect of increasing the size of the blocks on which the image is divided on the PSNR value and the maximum acceleration was also investigated. Comparison of the results of the proposed method was carried out with modern methods of motion vectors: FS, TSS, NTSS, DS, MVFAST, PMVFAST, which determined that the proposed method allows to obtain acceleration at 15-17% (the level of acceleration depends on the bitrate of the

processed video sequences: higher bitrate - higher acceleration) from the nearest analogue at losses at the level of 0.3-1.2% of PSNR values (PSNR losses depend on the size of the blocks for which the search was performed, maximum loss for 16x16pixel blocks, minimum for 64x64pixel blocks) from the values of the reference vector search method motion for modification A and acceleration at the level of 27-31% with a loss of 1.1% -2.1% depending on the block size and bitrate of the video sequence, for modification S.

The developed method of finding motion vectors is valuable for the development of technologies for processing, transmission and storage of video files.

***Keywords:*** *video stream compression, motion compensation, image similarity estimation, sum of absolute differences, object boundaries selection, LoG, Canny, block search methods.*