

---

# Heuristic Algorithms of Coincidence for the Estimation of Movements in Compression of Images

**Fernando José Hernández Gómez**

Department of Mathematics, Faculty of Science and Engineering, UNAN, Managua, Nicaragua

**Email address:**

fhernandez@dc.uba.ar, matematicas527@gmail.com

**To cite this article:**

Fernando José Hernández Gómez. Heuristic Algorithms of Coincidence for the Estimation of Movements in Compression of Images. *Applied and Computational Mathematics*. Vol. 9, No. 3, 2020, pp. 85-95. doi: 10.11648/j.acm.20200903.15

**Received:** September 14, 2019; **Accepted:** May 25, 2020; **Published:** June 8, 2020

---

**Abstract:** The NP-Completeness theory states that exact and efficient algorithms are unlikely to exist for the class of NP-difficult problems. One way to deal with NP hardness is to relax the optimality requirement and look for solutions instead that are close to the optimum. This is the main idea behind the approximation algorithms, which are called heuristic or metaheuristic. The problem of motion estimation is a process with a high degree of computational complexity, it requires sufficient memory space and execution time. It represents the cost of static, dynamic and video image sequence coding. The main task is to minimize the distortion rate and improve visual quality. This makes research in the field of coding, image compression and video focus on finding efficient algorithms to carry out the estimation of movement in a reasonable time. If a list of images of  $n$  elements is analyzed, there are feasible solutions. So, an exhaustive search is too slow, even for small values of the solution space. Therefore, from a practical point of view, it is crucial to have efficient and fast heuristic algorithms that avoid thorough search. In this investigation we design and implement heuristic algorithms, based on the frequency domain, which are applied on the coefficients of the discrete transform of the cosine and wavelets. Also, we propose temporal domain algorithms such as block-matching algorithms, which focus your search on the maximum coincidence of the current image with the reference one. The algorithms used during the implementation of this research work were written with the mathematical programming language MATLAB. In addition, we review the basic concepts of image processing, video, compression algorithms and motion estimation frequently used. The evaluation of the algorithms was carried out with a set of images provided by a previous acquisition system. We show the improvement of visual quality, the amount of compressed or reconstructed information and the behavior of the methods in the search for similarities between pixels or images. Finally, we contribute to the dissemination of new lines of scientific research that lead to the expansion and improvement of the study, the generation of new knowledge, since it is a young area within the Education discipline of Nicaragua.

**Keywords:** Movement Estimation, Heuristics, Discrete Cosine Transform and Wavelets, Search Algorithm

---

## 1. Introduction

Combinatorial Optimization is a branch of Optimization theory in applied mathematics and Computer Science, which is closely associated with Operations Research or Mathematical Programming, algorithm theory and computational complexity. The developed algorithms often solve instances of problems that are often difficult to find; in general, a space of solutions is explored, relatively large for them. These are achieved by reducing the size of the space, exploring and exploiting the properties to the maximum efficiently and quickly.

Within the field of Operations Research, a series of

heuristic and metaheuristic procedures have emerged to solve these problems, which are flexible when involving specific characteristics and allow interaction with the decision maker to seek approximations to the ideal solution, based on of technological development in the area of computing, generally heuristic procedures are iterative and require a large number of calculations. Heuristic is defined as a procedure in which the solution is not determined directly, but by tests, tests and retests; obtaining a high degree of confidence in quality solutions at a reasonable computational cost, although the optimal solution or its feasibility is not guaranteed, in some cases, it is not possible to establish how close it is to the optimal solution. Always in the search for

better solutions in the space of solutions and good quality, research in this field in recent years has focused its attention on the design of general purpose techniques to guide the construction of solutions from different heuristics. These techniques are called metaheuristics and they are strategies to design and / or improve heuristic procedures aimed at obtaining high performance in solving problems that are difficult to solve. The term metaheuristic was introduced in [12] by Glover, and since then many proposals have appeared for the design of better procedures to solve combinatorial problems.

By studying the Computational Complexity Theory it is possible to understand the importance of these types of problems that are mostly difficult to solve or impossible; they belong to the NP-Hard combinatorial optimization problem class, as the size of the cases grows. It is difficult to perform a simple analysis of an algorithm that determines the exact amount of time it requires to be executed, because it depends largely on the efficiency in the implementation of the algorithm and the processes of computer execution. In addition, knowing the exact time it will take to calculate or recalculate the results, it is a difficult task to find, the main task is to calculate the amount of operations that are performed according to the input data of the problem to be treated, this is known as "calculation of temporal function" and it is used in [7] by Cook from the theoretical point of view within the theory of computational complexity.

Once you have an algorithm that works in an ideal way, the next task is to define the criteria that allows you to measure its performance or behavior, that is, the efficient use of resources and the simplicity of the algorithm must be considered. A simple algorithm does not take away credit or detracts from quality, since its simplicity facilitates its maintenance, verification and efficiency. When talking about the efficient use of resources, it can be measured based on two indicators: space, amount of memory used, and the time it takes to execute. If to solve a problem P, and algorithm A, you need little execution memory in a computer with a small number of instructions compared to the rest of the known algorithms that solve the P, then it can be affirmed that A is more efficient than remaining when resolving P.

The main lines of study in the field of Artificial Vision consist on: generating intelligent and useful descriptions of: scenes, films, visual sequences on images and videos, emphasizing essential characteristics, relevant objects that appear in them. Some of the problems studied are detailed below:

Motion detection, whose main objective is to detect motion estimation in the scenes of a movie, or small films, compression of digital images.

Detection and location of moving objects. It is one of the most complicated problems for the reconstruction of movie scenes, including also the detection and prediction of future trajectories.

Obtaining 3-D properties of objects from movement. Being one of the classic problems of Artificial Vision.

A fourth study group that is related to the first: Estimation

of movement in problems of compression of images and videos.

Detection, recognition, identification, object tracking, camera calibration and stereoscopic vision, etc.

### **1.1. Motivation**

The whole idea behind the compression of images, sequences or video, consists in the coding of the Movement Estimate with the current frame with respect to the previous frame. A motion -compensated image is then created for the current frame that is constructed of blocks of the image from the previous frame. Motion vectors are sent for the blocks used for motion estimation, as well as the difference of the compensated image with the current frame, it also encodes and sends JPEG. The increasing complexity of the systems of Artificial Vision makes the possibility of easily testing different algorithms and alternative methods to carry them out, supposing a considerable saving when designing and implementing such systems of artificial vision.

The analysis of movements in the restoration and compression of images, sequences of images and videos is one of the main problems regarding the extraction of information, characteristics and relevant clues for the subsequent analysis of its content. Object detection and segmentation are greatly benefited if there are efficient algorithms for estimating movements. It is for this reason that it is intended to obtain a good correspondence of points between images, sequences of images and consecutive images of a video and thus achieve discrimination of objects according to their movement.

### **1.2. Problem Statement**

The main advantage of the use of exact algorithms is that they guarantee to find the global optimum of any problem, but they have the disadvantage that in real problems of computational complexity NP or NP-Hard, their execution time grows exponentially as the instance number or Space of analyzed solutions. It is for this reason that heuristics that are usually quite fast will be analyzed, but the quality of the solutions found is usually quite bad. The heuristics and metaheuristics used offer a balance between both extremes. They are generic methods that offer a good solution and in many cases you get a global optimum in a reasonable computation time. Based on this, we are interested in designing, implementing and proposing new numerical methodologies, particular models, specific methods to solve it, computational algorithms for its efficient and approximate solution through hybrids and heuristics, for the most expensive of tasks: the movement estimate that corresponds to the search algorithms on the reference macroblock, focused on the analysis of frequency domain and time domain algorithms.

The development of Motion Estimation Algorithms in Image Understanding is an area of artificial vision that has made significant progress in recent decades. Among its applications can be mentioned: motion detection applied to

security (recognition of fingerprints, faces, iris and retina examinations, surveillance, etc.) of a bank or bus terminals, or airports, among others; visual tracking, detection, recognition, location of moving objects, identification of a series of local features (objects, blocks, segments, etc.) in images, including trajectory detection and prediction; obtaining properties in 2-D; Motion estimation in important problems because the compression efficiency of digital images, video, quality control and the complexity of a video encoder is determined.

The procedure consists in finding a set of allowed schemes, the one that has an associated cost and when reconstructing the complete image is good and of visual quality. It should be noted that the errors that occur frequently when reconstructing the images are: the block effect and the shadow effect, due to the amount of information analyzed. The nature of these sub-problems means that they can rarely be treated theoretically due to the large number of images that are processed in real time and the characteristics studied. Therefore, we are motivated to avoid direct calculation and consider the use of specific computers and algorithms to help us in our task. This being one of the unforeseen difficulties, which lies in the time it takes for the computer to rebuild the solution, since many of these combinatorial optimization problems belong to the NP-hard class discussed in the book of (Garey & Johnson, 1979). That is, the computation and reconstruction time that is required increases as the image size, the number of images analyzed or the characteristics under study increase. After evaluating several areas of study, it was decided to choose the topic: Heuristic Coincidence Algorithms for the Estimation of Movement in Image Compression. The algorithms addressed are the following: The frequency domain algorithms that are applied to the coefficients of the transformed ones: DCT, Wavelets and the temporal domain algorithms (block matching or of coincidences) such as: the exhaustive search algorithm, algorithm of three-step search, four-step search algorithm, diamond search algorithm, simple and efficient three-step search heuristic improvement algorithm, improved heuristic in three-step pattern search. The choice of the subject is due to the fact that it is original and of scope to our reality, since it addresses problems that provide practical results for their solution. Similarly, the focus of the subject corresponds to the discipline of doctoral studies in universities at national and international level, reflecting social and technological interest, since it revolves around one of the problems of the modern world, which is the optimization of resources for compression, transmission and storage of visual information.

The purpose of the study lies in the design and implementation of coincidence algorithms based on Motion Estimation, which require the best implementations of frequency domain algorithms based on the transformed coefficients through DFT, DCT and Wavelets, as well as the coincidence algorithms that belong to the temporal domain (FS, TTS, 4SS, DS, MHTSS, MHCS), which has been scarcely addressed in Nicaragua from a theoretical and computational point of view.

### **1.3. Justification of the Problem**

The potential of the research lies in trying to solve this theoretical - applied problem, since when designing, implementing and evaluating the coincidence algorithms for the estimation of movements in order to obtain a minimum improvement in the solution, give other alternatives; in which it is possible to obtain good algorithmic efficiency, being of social relevance since these improvements generate from the economic point great profits to the cinematographic industries, computer security system or related institutions.

Another reason that motivates us to study, is its methodological utility, since it can serve as a basis for further research in Nicaragua, since it is a subject that is not well researched and of interest to the scientific community, it will serve as support for certain courses that are could be implemented in the different universities of the country, so it is a new or null field within the scope of Education in Nicaragua, the different universities of our country do not offer a course or conferences related to the subject under study. Also, the importance and interest of the study lies in the fact that support will be left for the work carried out by both the theoretical and the experimental part, as well as its implementations in the MATLAB programming language that may be useful or supportive to the different students of the Computer Science, System, Computer Science, Mathematics, and other related careers, who are thinking about working in this area and that is affordable for them and the general public. In the same way, it has practical implications, because its results allow solving a set of problems in the area of operations research, artificial vision, various areas of knowledge and everyday life. For example, the results obtained apply to society's problems (such as the improvement of an old photograph, or reconstruction, some that are in poor condition, that is, adding certain coefficients to improve their visual quality, in the case of movie theaters projecting a larger image or movie generates certain effects of poor quality, or in the case of a video conference, etc.). In addition, it is applicable in the analysis of medical images, since it can become a support tool for diagnosis in which we need more vision and precision. In the same way, it can support us in the different fields in which it is necessary to protect evidence such as: genetics, botany, zoology, pathology, forensic medicine, diagnostic imaging, as well as institutions of the Health, surveillance and the army. It is applicable from the point of view of marketing, advertising, which wish to sell and offer their products to customers visually and effectively, etc.

## **2. Theoretical Framework**

### **2.1. State of Art**

The most expensive operation from the computational point of view and with few resources in the entire compression process is the restoration of digital images and the Estimation of Image Movement. Therefore, this field has seen the greatest activity and research interest in the last three

decades; in which the block algorithms of fundamental blocks are designed, implemented and evaluated since the mid-1980s, the recent block comparison algorithms of the year 2002, to the present. The most basic movement estimation methods are based on the difference between consecutive images; the majority of all the work on motion estimation focuses on the hypothesis of restriction of the optical flow proposed in [14, 16] by Lucas and Herbst in 2013, postulate that the intensity of a point that we follow along of his movement remains constant.

In [8] Cuevas performs a study of the estimation of block coincidence movement, which plays a very important role in the proposed video coding. The goal is to find the best match block within a search space. In addition, a new algorithm based on the optimization of the Artificial Bees Colony (ABC) is proposed to reduce the number of search locations in the BM process. In this algorithm, the calculation of the search space is drastically reduced by considering an aptitude calculation strategy that indicates when it is feasible to calculate or only estimate new search locations. Since the proposed algorithm does not consider any fixed search pattern or any other assumption of motion like most other BM approaches, there is a high probability of finding the real minimum (precise motion vector). The simulations show that the proposed method achieves the best balance over other fast BM algorithms, in terms of the accuracy of the estimate and the computational cost.

In [22] Yaakob evaluates and implements four block matching algorithms using motion estimation. In addition, the effects of the block size used to find the best algorithm between them and the most optimal are evaluated. The results suggest that among all the algorithms evaluated, ARPS has the best PSNR based on the computation time. In [9] Dufaux and Moscheni offer a classification of the proposed methods and divide them into four categories: Gradient, recursion, block comparison and frequency domain techniques. The precise monitoring of the movement for a large set of points is a task that has been addressed many times in [3] by Brox and that is still being studied by the scientific community. Also, throughout the study the block matching algorithms were reviewed and analyzed, the estimation of full search movement treated in [4] by Ratnam, in [10] a proposal of the cross-search algorithm by Ghanbari is given, the three-step search algorithm makes a proposal Li, Zeng and Liou and can see in more detail in [15], the Diamond Search algorithm is developed in [6] by Cheung, and in [17] the search is approached of Rood Adaptive Pattern proposed by Nie.

In the most recent literature we can see the articles of the authors: in [1] they work a specific focus has been on increased automation using data driven approaches, while leveraging remote sensing and real-time streaming of heterogeneous data from various resources, including unmanned aerial vehicles, surveillance cameras, and low-earth-orbit satellites. In [18] they perform an alternative grouping algorithm to group elements without considering distance as a criterion of similarity, the data was separated into different groups considering morphological operations,

with this scheme, the procedure used allows the integration of data points, which present a spatial connection and solves complex grouping problems that traditional grouping algorithms cannot solve.

In [19], the authors consider an alternative algorithm to find the best approximation of the reference macroblock to the estimate, they use as a criterion the measurement of the quality of the approximation of the estimated macroblock. The algorithm is capable of vectorizing and used for hierarchical estimation of motion vectors. In [11], they propose a fast hybrid ME algorithm to encode video frames, using an early termination technique with 18 search points compared to the default Test Zone Search (TZS) algorithm in HM16.8 which uses 77 to the diamond search pattern and 81 for the square search pattern.

In [2], they propose a method of motion-estimation-based H. 264 video coding method using the optimal search-range for video broadcasting from a studio, evaluated the performance of the method, through experiments using standard videos, obtaining a improvement of approximately 0.2–0.3 dB and 84 ms per 100 frames in terms of PSNR (peak signal-to-noise ratio) and calculation speed.

## 2.2. Image and Video Processing

This section describes the relevant ideas of image and video processing. A brief historical review is made to understand the genesis of image processing, some main authors who provided essential contributions to the development are mentioned, the basic concepts that will be taken into account throughout the investigation such as: compression and transmission of digital images, existing types of redundancy, different compression techniques, block transformation coding, image transformation and quantization processes, mathematical representation and their properties, different measures or metrics existing for the process of compression of digital images and videos. In addition, the methods of determining optical flow are described, mention is made of the Walsh-Hadamard transform, and the general aspects of the Fourier Transform and main aspects of the discrete Transform of the Cosine, the Wavelets, as well as the details of the Implement these types of methods. Also, the metrics for the estimation of movements based on block-matching algorithms (BM) are shown, in which the optimal movement vector  $(mv_x, mv_y)$  is determined from a cost function. This function indicates the degree of coincidence between the original image  $I_C$  or frame and the reconstructed image or the residual  $I_R$ , frame, the parameters  $m, n$ , represent the length and width of the image. The differences between the cost functions vary in terms of the complexity and efficiency of the metrics to be used.

### 2.2.1. Discrete Cosine Transform

The Discrete Cosine Transform (DCT), is the most widely used in image and video compression applications. Rapid or efficient algorithms of the discrete cosine transform have been developed, in [5] Chen and Smith provide an improvement factor in their computational complexity, when compared to

traditional Cosine Discrete Transform algorithms. The unique properties that the DCT possesses allow to obtain high compression rates at very low cost, since they have a good energy compaction property. The decorrelation of the coefficients is one of the fundamental tasks for compression, since the subsequent treatment of each coefficient can be performed independently, without loss of efficiency by quantifying the coefficients that are chosen visually. In general, two groups of techniques were studied: loss compression and lossless compression. Lossless coding, decompression aims to recover exactly the original image information in such a way that the compression rate is good, while compression with losses the compression process allows a loss of information that is not recovered in the decompression.

### 2.2.2. Movement Estimate Algorithms

In the estimation of movement by blocks the image is divided into rectangular blocks separated from each other, of fixed size. It is considered that there are only translations, being the model par excellence frequently used in the compression process, mainly for its simplicity and its good implementation results. One of the objectives consists in the search for the maximum coincidence of the real movement of the objects, but the strongest task is the reduction of the prediction error and the visual quality of the images. However, this presents different problems with more general movements such as a simple related transformation that includes scaling and / or rotation. The movement models can be grouped into two large classes: the parametric and non-parametric models.

### 2.2.3. Methods of Coincidence by Blocks

Block matching methods comprise a whole category of algorithms that behave similarly by dividing the image into small regions called blocks, often into squares, rhombuses, among others. These focus their search on the coincidences of the information (data, characteristics, objects, etc.) of the

current block with the reference one. The simplest idea is to divide an image into small rectangular or quadrangular blocks, of fixed size, under the assumption that each performs an independent translation. If these blocks are small enough, movements such as the rotation of large objects, or the zoom of the image, from translations of these areas or regions can be approximated, this method avoids the coding relative to the segmentation and it is only necessary to encode the Scroll vector of each block. The hypothesis on which these methods are based is that a block of size  $L \times L$  in the image  $k - l$  centered at the point  $(x_0, y_0)$  coincides with a block of the same size in the image  $k$ . This expressed mathematically would be:

$$f(x, y, k) = f(x + d_1, y + d_2, k - 1) \forall (x, y) \in B$$

In other words, each of the blocks that make up the frames of the image sequence that a counterpart has, that is, a block of the same size that represents the same portion of the scene in another image or consecutive frame of the image sequence. The algorithms designed block matching (FS, TTS, 4SS, DS, MHTSS, MHCS), focus their search on the maximum match between all or some of pixels in the current block with the reference block on the search area, based on some metric or cost function that has been addressed in depth in the previous chapter. Most of the study focuses on both the techniques of estimation and compensation of movements and the analysis of the segmentation of the images to achieve a higher level of estimation and compensation, for this the segmentation is studied by fixed blocks, by variable blocks and segmentation based on regions taking objects within the image.

### 2.2.4. Wavelets Formulation

The wavelets used in the implementation of the different algorithms are: Haar, Daubechies 4, Symmlet 8, its functions being the following:

$$\begin{aligned}
 f(x, y) &= \sum_{i=0}^{N-1} \sum_{k=0}^{N-1} X_{ik} \Phi(x-i) \Phi(y-k) \\
 &= \sum_{i=0}^{N-1} \left\{ \sum_{k=0}^{N-1} X_{ik} \Phi(y-k) \right\} \Phi(x-i). \\
 f(x, y) &= \sum_{i=0}^{N-1} \left\{ \sum_{k=0}^{N/2-1} \tilde{c}_{i,k}^{(-1)} \Phi_{-1,k}(y) + \sum_{k=0}^{N/2-1} \tilde{d}_{i,k}^{(-1)} \Psi_{-1,k}(y) \right\} \Phi(x-i) \\
 &= \sum_{k=0}^{N/2-1} \left\{ \sum_{i=0}^{N-1} \tilde{c}_{i,k}^{(-1)} \Phi(x-i) \right\} \Phi_{-1,k}(y) + \sum_{k=0}^{N/2-1} \left\{ \sum_{i=0}^{N-1} \tilde{d}_{i,k}^{(-1)} \Phi(x-i) \right\} \Psi_{-1,k}(y). \\
 f(x, y) &= \sum_{k=0}^{N/2-1} \left\{ \sum_{i=0}^{N/2-1} LL_{i,k} \Phi_{-1,i}(x) + \sum_{i=0}^{N/2-1} LH_{i,k} \Psi_{-1,i}(x) \right\} \Phi_{-1,k}(y) \\
 &+ \sum_{k=0}^{N/2-1} \left\{ \sum_{i=0}^{N/2-1} HL_{i,k} \Phi_{-1,i}(x) + \sum_{i=0}^{N/2-1} HH_{i,k} \Psi_{-1,i}(x) \right\} \Psi_{-1,k}(y)
 \end{aligned}$$

for filtering the essential coefficients of macroblocks, respectively.

### 2.2.5. Computational Complexity

The exhaustive method requires a very high computational cost, which can prevent it from being used in real-time applications. It is for this reason that the heuristics that are usually quite fast (efficient) are designed and analyzed and the quality of the solutions found is usually quite good. The theory of computational complexity states that a decision problem can be classified as: class L, NL, P, NP, NP-Complete and NP-Hard, Co-NP. However, for this distinction it is necessary to consider the theoretical model of Turing (MT) machines. Figure 1 summarizes the above, as shown below:

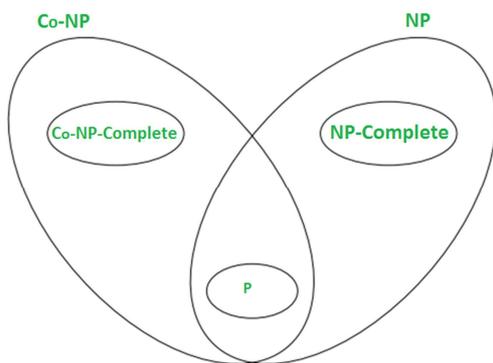


Figure 1. Computational theoretical complexity.

In addition, the following questions arise:  $P = NP$ ,  $Co-NP = NP$ ,  $P = Co-NP \cap NP$ , some of them are still open to the scientific community, being their object of study for computer scientists, computers and theoretical mathematicians.

## 3. Applied Methodology

We proceeded to design the different algorithms, the frequency domain that are based on the coefficients of the transformed: the Discrete Cosine transform and the Wavelets, these were chosen as they serve as the basis and are the ones that best adapt to the treatment of the temporal domain algorithms that respond to the problem in studies, also designing the following heuristics: the exhaustive search algorithm, three-step search algorithm, four-step search algorithm, diamond search algorithm, algorithm and improved heuristics: simple and efficient three-step search, three-step pattern search algorithm.

In figure 2 shows the methodological design of the proposed work.

The methods used in the research are: *general and pure empirical methods*, since the idea is to start from full search algorithms, heuristics already known by scientific society, apply them to different instances of the problem of compression of static, dynamic images and the movement estimation. In addition, the parameters of the models are varied in the techniques of heuristics, making hybrids of these. Also, *the synthesis analysis method* was used since it partitioned a complex problem into smaller ones, then combining the results of algorithms obtained in order to match or improve them. In addition, *abstraction and concretion methods* were used since each of the algorithmic results must be demonstrated and verified.

It should be noted and made clear that in this investigation the most used is the *algorithmic method*, since heuristic search algorithms must be implemented in the programming language and work with different computational laboratory experiments to take advantage of them.

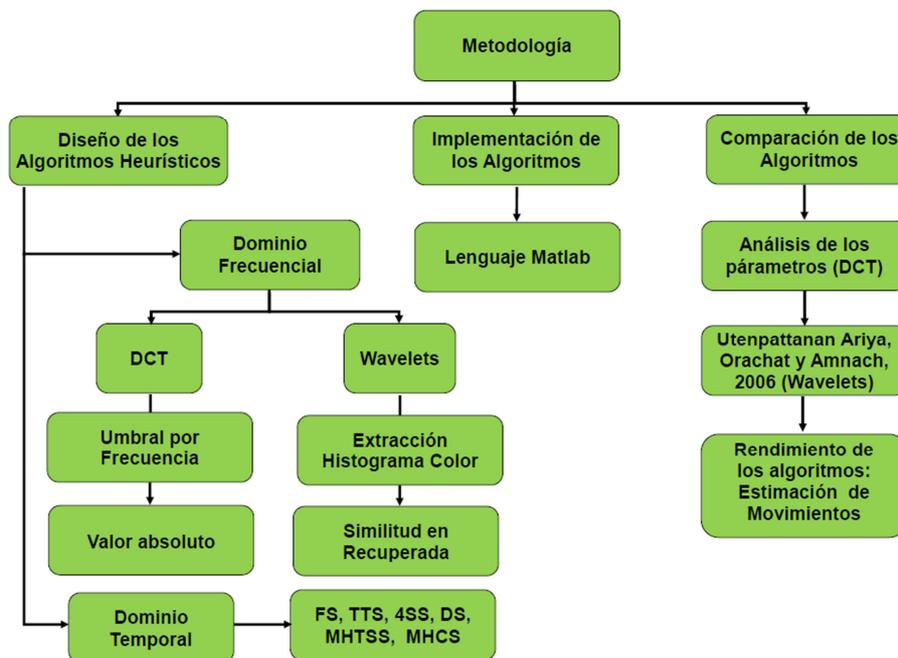


Figure 2. Scheme of the methodology used.

## 4. Results and Discussion

### 4.1. Image and Video Processing

For the implementation of the algorithms there are two alternatives, the first is to start from the theory, from the mathematical bases and program them in their entirety; the second option is to use some of the libraries or specific functions of *Matlab software*, since the mathematical development as we could see throughout the previous chapters is very extensive and the coding in software programs would take us a long time, necessary to be able to achieve the objectives of the thesis. The functions used were implemented and coded in the Matlab software that will have as input a matrix-type variable that represents an image in gray (black and white) or color levels with their respective levels. In the case that a color image is used and needed in its gray levels, the Matlab function *rgb2gray()* will be used. To this input matrix, which being gray levels will be two-dimensional, the *DCT-2D* will be calculated using the *dct2()* function, finally its coefficients will be graphically represented by the *mesh()* function to measure performance when it is not achieved. Appreciate the visual quality. To perform the selection

of *DCT* coefficients by frequency or based on their absolute value, the functions *dct\_ufrec\_modif* and *dct\_valor\_absoluto\_modif* that uses the percentiles instead of dividing the range of the coefficients of the *DCT* into equal intervals were implemented. We observe the block or frame effect, a very pronounced difference that originates when dividing the image into blocks of different sizes ( $2 \times 2, \dots, 8 \times 8$ ), when applying the *DCT* to each block and reconstructing the image, it appears pixelated over the entire image, regardless of whether we apply a uniform quantizer to the entire image, or if we select coefficients based on their frequency or based on their absolute value. It should be noted that by taking 16% and 25% of the *DCT* coefficients, this same artifact or effect is still appreciated when reconstructing the image, but this fact can only be appreciated if we look closely at the image. For example, we present the image of Lena, satellite and reconstructed tomato taking 25% of the *DCT* coefficients, which is equivalent to taking only the coefficient *D*(0,0) of each block of size  $2 \times 2$  and taking *DCT*(*i*, *j*)  $\forall 1 \leq i, j \leq 3$  in each  $8 \times 8$  block. We enlarge a certain area of the images to better appreciate the pixelated difference of the images under study, as shown in Figure 3.

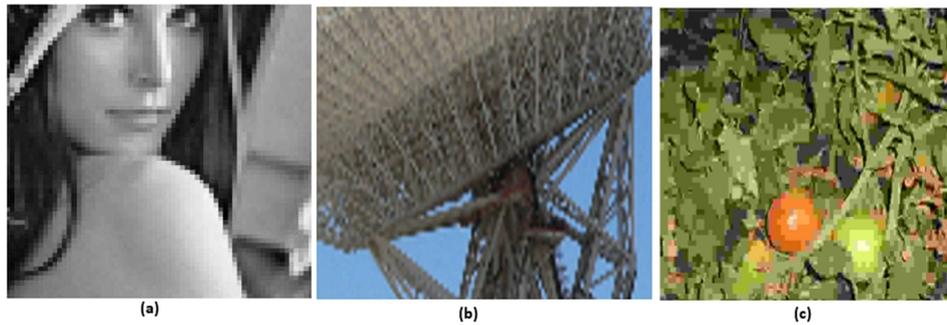
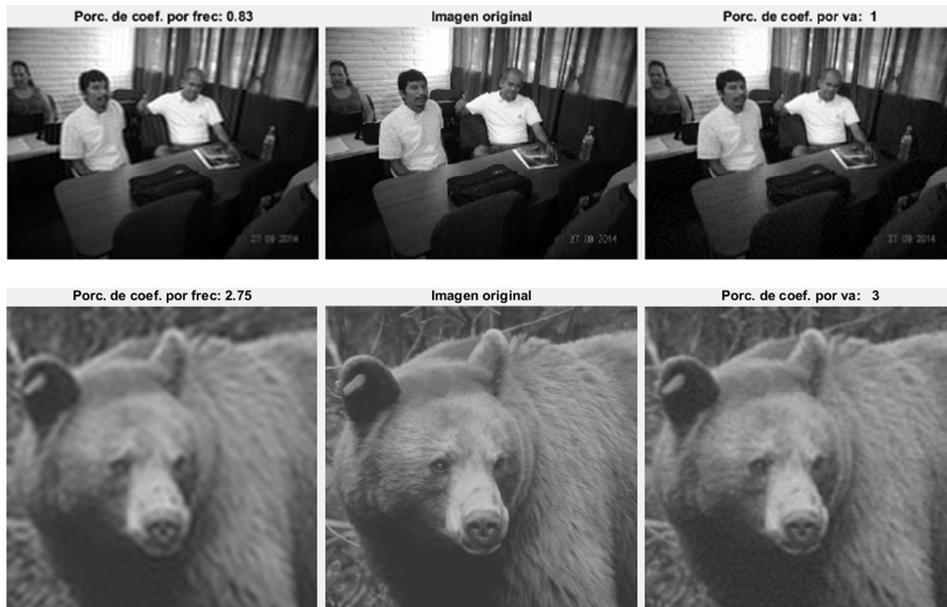


Figure 3. (a) Lena. (b) Satellite. (c) Tomato.

In the same way, we present an experiment that consisted of selecting different images (FerPrueba1 with approximately 1%, bear with approximately 3%, Amber with approximately

5% and 20% and FerinBuenosAires with approximately 13% of the selected coefficients, as shown in the figure 4, respectively.



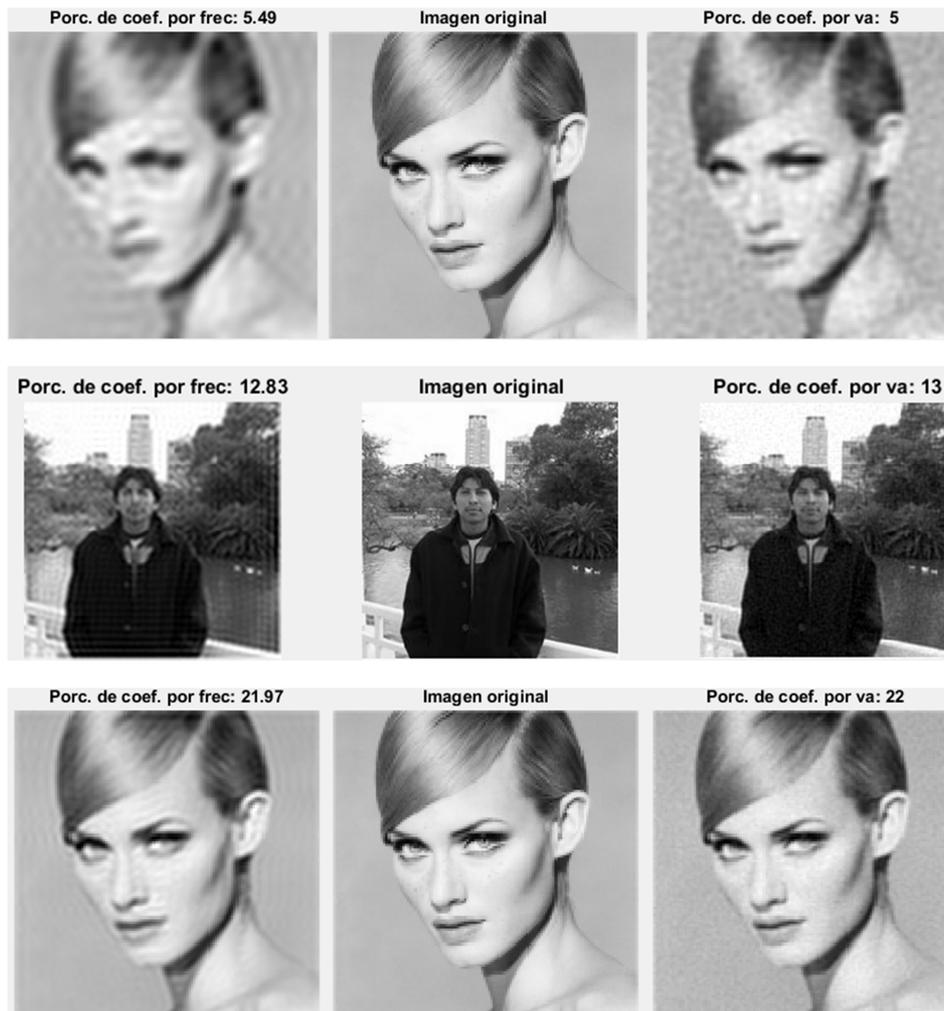


Figure 4. Set of test images.

Before carrying out the tests, we conjectured that this difference would be significant in images that present high frequencies, since in the selection of the coefficients by frequency these they would be discarded and cause the reconstructed image to lose those details.

#### 4.1.1. Application of the Discrete Cosine Transform

For the implementation of the algorithms there are two alternatives, the first is to start from the theory, from the mathematical bases and program them in their entirety; the second option is to use some of the libraries or specific functions of *Matlab software*, since the mathematical development as we could see throughout the previous chapters is very extensive and the coding in software programs would take us a long time, necessary to be able to achieve the objectives of the thesis. The functions used were implemented and coded in the Matlab software that will have as input a matrix-type variable that represents an image in gray (black and white) or color levels with their respective levels. In the case that a color image is used and needed in its gray levels, the Matlab function *rgb2gray()* will be used. To this input matrix, which being gray levels will be two-dimensional, the *DCT-2D* will be calculated using the *dct2()* function, finally its coefficients will be graphically

represented by the *mesh()* function to measure performance when it is not achieved. Appreciate the visual quality. To perform the selection of *DCT* coefficients by frequency or based on their absolute value, the functions *dct\_ufrec\_modif* and *dct\_valor\_absoluto\_modif* that uses the percentiles instead of dividing the range of the coefficients of the *DCT* into equal intervals were implemented. We observe the block or frame effect, a very pronounced difference that originates when dividing the image into blocks of different sizes ( $2 \times 2, \dots, 8 \times 8$ ), when applying the *DCT* to each block and reconstructing the image, it appears pixelated over the entire image, regardless of whether we apply a uniform quantizer to the entire image, or if we select coefficients based on their frequency or based on their absolute value. It should be noted that by taking 16% and 25% of the *DCT* coefficients, this same artifact or effect is still appreciated when reconstructing the image, but this fact can only be appreciated if we look closely at the image.

#### 4.1.2. Compression Achieved by Applying the DCT

Everything analyzed in the previous items, allowed us to conclude that the compression achieved using the selection by absolute value would be greater than or equal to that achieved using the frequency selection, since, we can

consider that all the coefficients of the *DCT* are of the same order (Perhaps the only coefficient of greater value with respect to the others is the first coefficient  $D(0,0)$ , it is one of the main reasons why it should be included in any selection that is made, either by frequency or by absolute value).

#### 4.1.3. Progressive Transmission of the DCT

For the analysis of this section we attach *six files.m*, it is the extension, which when executing the implementations of the algorithms produce videos. These files are:

*TransProgresivAmberSurfc.m*, *TransProgresivbear.m*,  
*TransProgresivFerRudy.m*, *TransProgresivLena.m*,  
*TransProgresivMandril.m* y *TransProgresivPuente.m*.

In the first five experiments we managed to observe how the image is better reconstructed as the percentage of coefficients of the *DCT* increases. As we mentioned, the reconstruction is faster by selecting the coefficients of the *DCT* by absolute value, but in both cases what we observe is that the image is being reconstructed from the thickest details to the finest details, that is, at the beginning we go obtaining as an outline of the image and as *DCT* coefficients are added the details appear, as seen above all in *TransProgresivMandril.m* and *TransProgresivLena.m* where the finer details are the hairs of the mandrel mustache, as well as the Lena's hat feathers appear as final details in the reconstruction of the image. The justification for this is direct in frequency reconstruction, since we choose from the lowest frequencies to the highest frequencies and precisely being the finest details in the reconstruction are provided by the higher frequencies. Another point of interest that we observe is that from a not so high percentage of coefficients the changes in the reconstruction of the image are imperceptible to our eye, because from a certain percentage not very high of the coefficients, it is possible to reconstruct an approximation to the original image. However, the image is modified even if our eye is not able to capture these details.

We can visualize this with the Matlab *surf* function that graphs the relationship between each pixel in the image and its gray value. For example, we propose to watch the video generated by the file *TransProgresivAmberSurfc.m*. When observing the reconstruction of the image we see that from a certain moment our eye does not capture differences between the reconstructed images and the original one, the *surf* function allows us to observe that there are still changes. In addition, it was seen that when taking an area of interest and enlarging the image (zoom), the details are observed on a large scale; pronouncing the block effect to a large extent according to the size of blocks chosen. As well as the shadow effect.

#### 4.2. Restoration of the Images in the Wavelet Domain

The algorithm used to calculate Haar's binary descriptor consists of a series of steps that are mentioned below:

The *color histogram* is obtained using *RGB* color space and the *Haar coefficients* of the histogram are determined by using the wavelet transform.

Haar coefficients are *Binarized* and Hamming distance is

applied.

In the case of using the color histograms of the query, in the different experiments it was possible to see that the light tones of the image correspond to the values of the right histogram. It was observed that if we take into account the original image to the test group, the algorithm finds the original image and locates the most similar to it. In addition, we affirm that:

Two images compared through their global color histograms may not be semantically related despite sharing a similar color distribution.

The binary descriptors of the Haar Wavelet of the images of interest are compared with the query using the Hamming distance. The most similar are retrieved and sorted according to the distance to the query.

#### 4.3. Heuristic Algorithms of Coincidences for the Estimation of Movements

The implementation of motion estimation algorithms is one of the key tasks of the image compression process and sequence of them, it takes us a good percentage in the processing time. These are developed with the purpose of reducing computational complexity, due to their efficiency and good performance. The proposed algorithms are modified based on the search strategy, they are considerably improved by the implementation of heuristic algorithms and their hybrids, providing better results for *PSNR* without increasing the number of calculations.

#### 4.4. Test Sequence for the Set of Algorithms

The experiments performed for the analysis of the performance of the described algorithms consisted of coding one of sequences of *bus\_cif\_frame* test images, selected from the set of images available in the Video Trace Library [20] (ARIZONA STATE UNIVERSITY), with certain specific characteristics. To perform the performance comparisons, measurements of the number of search points and the error in the estimation and compensation of generated movements were taken. In addition, the conditions of the experiments are described, in which 150 frames of the selected video sequences of the same page are used.

The different executions of the algorithms, take a block size of  $16 \times 16$  pixels and with a search parameter  $p = \pm 7$ , since it presents us with better yields and visual quality in the reconstruction. The evaluation of the improved heuristics according to the *MHCS* cross-search pattern implemented gives us an average performance, with respect to the rest of the algorithms, and behaves well, both computationally and in terms of visual quality, in the different experiments performed. This is observed in figure 5.

In addition, it is appreciated that the *ES* algorithm is the one that uses the highest computational cost, according to the metric used. The *TSS*, *4SS* and *DS* heuristic algorithms use approximately the same execution time, that is, they behave similarly, the metaheuristics *MHTSS* and *MHCS* work on average and their implementations are on average better with

respect to the other algorithms due to the improvements of the varied parameters.

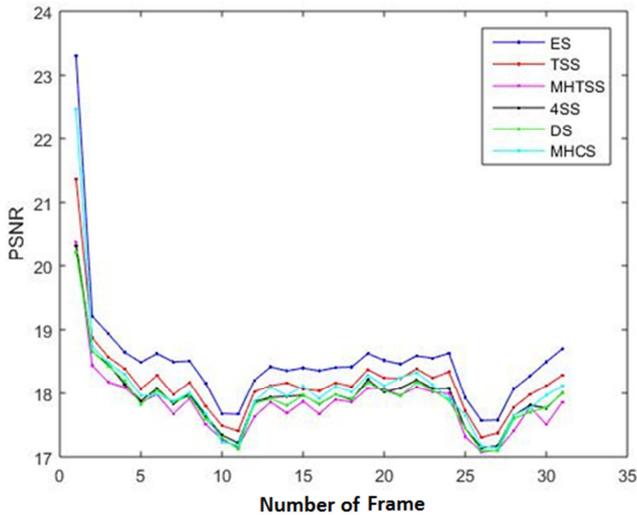


Figure 5. Movement Estimation Algorithms.

## 5. Conclusions and Future Perspective of the Research

### 5.1. Conclusions

In this investigation, we deal with the estimate of available 2-D movement. To this end, two main lines of study were raised: those that are developed on frequency domain methods, using static textures and the implementation of temporal domain methods using dynamic textures.

A unified study of all compression measures was reviewed and conducted, focusing our interest in the discrete transformation of cosine and Wavelets, essential for the experimental part of the motion estimation algorithms, and their variants in heuristics. The formal model was developed, allowing to characterize the heuristics, observing the common points that the different compression and search algorithms have in the reference macroblocks, and focusing mainly on the improvement of visual quality, the amount of information compressed or reconstructed and the behavior of the methods in the search for similarities between pixels or images. The algorithms are generic enough to address the existing models and their variants, but at the same time allow to control details of low level, which are especially interesting in the specific treatment of heuristics. Also, the theoretical results of the mathematical theory essential for the design and implementation of the Matching Algorithms for the Estimation of Motion in image compression were structured.

The design and implementation of the codes of the main functions and procedures of the frequency domain algorithms (DCT, Wavelets) and Heuristics (FS, TTS, 4SS, DS, MHTSS, MHCS) of Coincidence for Movement Estimation in image compression, written in the Mathematical Programming Mathematics language. The general results show that when

filtering the coefficients of the DCT based on absolute value it is better than the coefficients selected by frequency, independent of the size of the macroblock used. Different runs of both methods were performed, and their possible combinations, these experiments serve as a basis to affirm that an absolute value selection would give better visual results since they eliminate high frequencies and leave the lowest ones. In addition, we could observe that by varying the parameters and dividing the images into blocks of different sizes, it allowed us to decide that the  $8 \times 8$ , blocks can be a good size, since we obtained a pixelated smaller than when dividing the image into blocks of size More smalls. For sizes  $\geq 8$ , the visual quality and compression level of the analyzed images considerably improves.

In the experimental part of the Wavelets we could see that in the different executions of the implemented algorithms, that by including the original image to the test database the implemented algorithm finds it in step 1 or 2 with a minimum distance, or samples the one that is at a minimum distance. In addition, two images compared through their global color histograms may not be semantically related despite sharing a similar color distribution. The binary descriptors of the Haar Wavelet of the images of interest are compared to the query image using Hamming distance. Therefore, the most similar images are retrieved and sorted according to the distance to the query.

The third set of experiments performed in this work, focuses on the implementation and evaluation of the performance of the set of heuristic matching algorithms for estimating fast search movements, in the context of MPEG-JPEG standardized encoders. To conclude, the results show that the algorithms that have a higher performance are the improved heuristic MHCS and MHTSS because they are more efficient and because of their visual quality, and to a lesser extent in increasing order of computational complexity we have the heuristic algorithms: DS, 4SS, TS, and ES. Based on the analysis, these are considered to be the best on average.

### 5.2. Research Perspective

As future work, it would be very useful for both lines to have a more efficient implementation of the heuristics, to work even with the variation of the parameters and the performance measures, to think about the design of metaheuristics to achieve better results. In addition, you should take advantage of the hardware boom in parallel or the use of computer clusters to streamline performance processes, or in the case of working with personal computers and with a good computing power. Since one of the disadvantages observed in the line of research treated, is that these search algorithms are very demanding in terms of computational performance, from the software used that requires a good cache capacity and a good processor for its execution, since they make them impracticable for real-time applications, when the image or sequence has a lot of information to analyze. The work of the codes in  $C++$  is open to improve their performance.

---

## References

- [1] Ahmed, Z., Hussain, A. J., Khan, W., Baker, T., Al-Askar, H., Lunn, J.,... & Liatsis, P. (2020). Lossy and Lossless Video Frame Compression: A Novel Approach for High-Temporal Video Data Analytics. *Remote Sensing*, 12 (6), 1004.
- [2] Boonthep, N., & Chamnongthai, K. (2019). A Method of Motion-Estimation-Based H. 264 Video Coding Using Optimal Search-Range. *Wireless Personal Communications*, 1-18.
- [3] Brox, T., Bruhn, A., Papenbergh, N., and Weickert, J. (2004). "High accuracy optical flow estimation based on a theory for warping". *Computer Vision-ECCV*, 25-36.
- [4] CH, C. S., Ratnam, J. V. K., and Student, P. G. "Comparison of Fast Block Matching Algorithms for Motion Estimation".
- [5] Chen, W. H., Smith, C. H., and Fralick, S. C. (1977). "A fast computational algorithm for the discrete cosine transform". *IEEE Transactions on communications*, 25 (9), 1004-1009.
- [6] Cheung, C. H., and Po, L. M. (2002). "A novel cross-diamond search algorithm for fast block motion estimation". *IEEE transactions on Circuits and Systems for Video Technology*, 12 (12), 1168-1177.
- [7] Cook, S. A. (1971, May). "The complexity of theorem-proving procedures". In *Proceedings of the third annual ACM symposium on Theory of computing* (pp. 151-158). ACM.
- [8] Cuevas, E., Zaldívar, D., Pérez-Cisneros, M., Sossa, H., Osuna, V. (2013). "Block matching algorithm for motion estimation based on Artificial Bee Colony (ABC)", *Applied Soft Computing Journal* 13 (6), pp. 3047-3059.
- [9] DUFAUX, F. (2010). "Motion Estimation Techniques for Digital TV: A Review and a New Contribution".
- [10] Ghanbari, M. (1990). "The cross-search algorithm for motion estimation (image coding)". *IEEE Transactions on Communications*, 38 (7), 950-953.
- [11] Gogoi, S., & Peesapati, R. (2019, December). A Hybrid Motion Estimation Search Algorithm for HEVC/H. 265. In 2019 IEEE International Symposium on Smart Electronic Systems (iSES)(Formerly iNiS) (pp. 129-132). IEEE.
- [12] Glover, F. W., and Kochenberger, G. A. (Eds.). (2006). "Handbook of metaheuristics". (Vol. 57). Springer Science & Business Media.
- [13] Hernández Gómez, F. J. (2017). "Algoritmos Heurísticos de Coincidencia para la Estimación de Movimientos en Compresión de Imágenes". (Doctoral dissertation, Universidad Nacional Autónoma de Nicaragua, Managua).
- [14] Herbst Evan, Ren Xiaofeng, and Fox Dieter. (2013) "Rgb-d flow: Dense 3-d motion estimation using color and depth". In *Robotics and Automation (ICRA)*, IEEE International Conference on, pages 2276-2282. IEEE, 2013.
- [15] Li, R., Zeng, B., and Liou, M. L. (1994). "A new three-step search algorithm for block motion estimation". *IEEE transactions on circuits and systems for video technology*, 4 (4), 438-442.
- [16] Lucas, B. D., & Kanade, T. (1981). "An iterative image registration technique with an application to stereo vision". In *IJCAI*, pages 674-679, 1981.
- [17] Nie, Y., and Ma, K. K. (2002). "Adaptive rood pattern search for fast block-matching motion estimation". *IEEE Transactions on image processing*, 11 (12), 1442-1449.
- [18] Ortega-Sánchez, N., Cuevas, E., Pérez, M. A., & Osuna-Enciso, V. (2020). Clustering Data Using Techniques of Image Processing Erode and Dilate to Avoid the Use of Euclidean Distance. In *Applications of Hybrid Metaheuristic Algorithms for Image Processing* (pp. 187-203). Springer, Cham.
- [19] Pyko, N. S., Pyko, S. M., Mikus, O. A., & Filippov, B. S. (2020, January). Motion Vectors Estimation for Video Encoding Based on the Macroblock Signatures. In 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus) (pp. 1417-1420). IEEE.
- [20] Reisslein, M. (2005-2011). "Video Trace Library". Arizona State University, [online] Available: <http://trace.eas.asu.edu>.
- [21] Yaakob, R., Aryanfar, A., Halin, A. A., and Sulaiman, N. (2013). "A comparison of different block matching algorithms for motion estimation". *Procedia Technology*, 11, 199-205.