

DOI: <http://dx.doi.org/10.21123/bsj.2020.17.2.0556>

Moving Objects Detection Based on Frequency Domain

Jalal H. Awad^{1*}

Balsam D. Majeed²

Received 30/1/2019, Accepted 16/7/2019, Published 1/6/2020



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

Abstract:

In this research a proposed technique is used to enhance the frame difference technique performance for extracting moving objects in video file. One of the most effective factors in performance dropping is noise existence, which may cause incorrect moving objects identification. Therefore it was necessary to find a way to diminish this noise effect. Traditional Average and Median spatial filters can be used to handle such situations. But here in this work the focus is on utilizing spectral domain through using Fourier and Wavelet transformations in order to decrease this noise effect. Experiments and statistical features (Entropy, Standard deviation) proved that these transformations can stand to overcome such problems in an elegant way.

Key words: Average filter, Fourier transformation, Frame difference, Median filter, Moving objects detection, Wavelet transformation.

Introduction:

In general, image processing techniques can be categorized into spatial and frequency domain techniques. Thus image filtering can be accomplished in such way. There are two choices to do image filtering, One of them is in spatial domain by convolving the image under consideration with an adequate window (basis image), while filtering in frequency domain occurs through multiplying the transformed image with an appropriate low pass filter (in this research a circle of "ones" with appropriate radius) (1). Examples of spatial domain techniques are mean (average), median, Gaussian...etc. Spatial domain techniques produce different resulted output signals (images in this research context). Filtering images in spatial domain are more computation consumer (2). Hence it is preferable to try the other techniques (frequency). Examples of such techniques are Wavelet, Fourier, Walsh...etc. Spatial domain techniques depend directly on pixel intensity levels, whereas frequency domain techniques depend on frequency coefficients (2). Transformation in spatial domain is done pixel in one domain to a pixel in other.

But in frequency domain the matter is different because every pixel in the image when it is in the spatial domain participates in producing every value in the frequency domain (1).

The whole operation is sometimes called projection which is convolution, correlation or multiplication of the original signal (image) with a basis function may be sine/cosine, high/low or Fourier, Wavelet (Haar, daubachies, ...etc.) and then taking the sum of the resulted multiplications, thus finding the transformation coefficients (1). Moving object detection is the first step for successive operations (3). One of the most difficult tasks is the detection of such objects in videos which has a great role in segmenting image frame into static and moving regions (4, 5). Such step can focus the attention just toward the moving object which leads to decreasing the computations. It may also be useful for offline videos indexing, searching, smart video data mining, community security, law enforcement (for example car excess speed) and many military applications (3). There are multiple approaches for detecting moving objects like background subtraction, optical flow and frame difference (5, 6, 7). The last is the one used in this research.

The aim of this research is to develop a technique which utilizes the two well-known Fourier and Wavelet transformations which have so important features that allow us to convert video spatial frames information with high degree of noise and redundancies into less correlated transformation

¹ Department of Computer, College of Science, Mustansiriyah University, Baghdad, Iraq.

² Department of Computer techniques Engineering, Imam Kadhim Faculty of Islamic Sciences University, Baghdad, Iraq.

*Corresponding author: jalalhameed@uomustansiriyah.edu.iq

*ORCID ID: <https://orcid.org/0000-0002-9313-6681>

coefficients (frequency domain coefficients) leading to reducing unnecessary computations as well as a flexibility in the selection of certain frequencies which composite the main frame (image) architecture elements with less noise elements that lay in some frequency ranges. So the resulting frame can be treated with frame difference technique in order to extract the mask of the moving object easily and clearly.

Materials and Methods:

Orthogonality and orthonormality of the basis functions:

Transforming an image from spatial to spectral (frequency) domain requires projecting (convolving/correlating) what are called basis images (basis functions) on the input image in order to convert it into its basis frequency components. These basis images should have two important features, namely orthogonality and orthonormality (4, 8). Orthogonality means that the basis images are perpendicular to each other which mean that their projection (inner product) is equal to zero. This feature is very important when analyzing signal (image), because the resulted zero of their inner product means that there is nothing in common between them (uncorrelated), which lead to good signal analysis. If they were not so (their inner product \neq zero) then the analysis will not achieve the required right precise analysis. If it to wonder why this should be, then the answer is that the transformation aims to analyze any complex signal (image) into weighted sum of these basis images. Whereas if these basis images were not orthogonal then the resulted transformation coefficients would include useless redundant information. Whereas orthonormality means that each basis image magnitude is equal to "1". These two features are applicable to the sine and cosine functions which are used in Fourier transformation (9).

Sine and cosine functions:

Sine and cosine functions can be clarified according to fig1 and Fig.2:

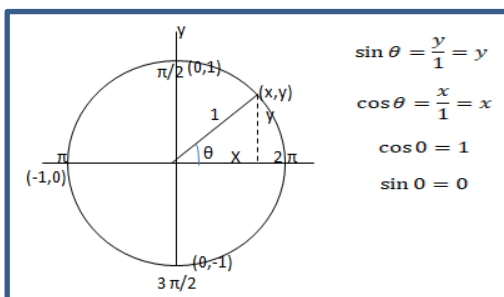


Figure 1. The unit circle used for deriving sine and cosine functions.

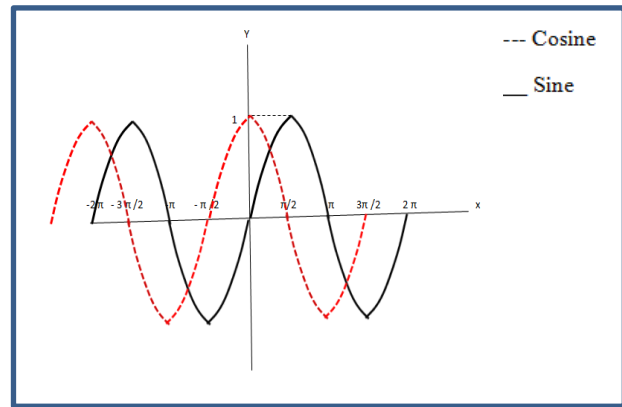


Figure 2. The sine and cosine waves for interval of 2π .

The unit circle of Fig. 1 as well as sine and cosine amplitude of Fig. 2 are compatible with orthonormality property of the basis functions (sin and cos in Fourier context). The "zero" result of the sine and cosine functions multiplication satisfies the orthogonality property.

Fourier transformation:

The operation of transforming the image into the frequency domain is called image analysis or decomposition which as mentioned before is done using adequate basis function. The converse operation is called reconstruction which depends on the inverse basis function which in turn is almost the same as or resemble to the one used in the decomposition step (1). Fourier transformation of two dimension function can be given with the following equations (9):

$$F(u, v) = \frac{1}{NM} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) e^{-2\pi i (\frac{ux}{N} + \frac{vy}{M})} \dots 1$$

$$e^{-iwx} = \cos(wx) - i \sin(wx) \dots 2$$

$$F(u, v) = \frac{1}{NM} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x, y) \cos\left(\frac{2\pi u}{N} x + \frac{2\pi v}{M} y\right) - i f(x, y) \sin\left(\frac{2\pi u}{N} x + \frac{2\pi v}{M} y\right) \dots 3$$

The inverse of this transformation can be given according to the following equations:

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} F(u, v) e^{2\pi i (\frac{ux}{N} + \frac{vy}{M})} \dots 4$$

$$f(x, y) = \frac{1}{NM} \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} \text{Re}(F(u, v)) \cos\left(\frac{2\pi u}{N} x + \frac{2\pi v}{M} y\right) - \text{Im}(F(u, v)) \sin\left(\frac{2\pi u}{N} x + \frac{2\pi v}{M} y\right) \dots 5$$

Where $f(x, y)$ is the function to be transformed, N, M are its dimensions, x, y are two spatial values, and u, v are the frequency coefficients (9). The 2π are used here because the

sine and cosine waves need 2π (360°) in order to accomplish one cycle as explained in the before mentioned unit circle. Thus $2\pi u$ is called angular frequency (9).

Fourier transform tends to decompose a complex signal (can be image) into “zero” frequency term (DC) and multiple sinusoidal terms (basis functions), where each sinusoid is a harmonic of the fundamental basis image (first harmonic) which has the lowest frequency. The remaining harmonics are frequency multiples of the fundamental. (1) In Fourier transformation the value of $u=0$ means the slowest frequency (first harmonic) in the signal under consideration which is wanted to be analyzed into its principal frequency components. This frequency is called the zero frequency. The low frequency components in the analyzed signal are related to a slow intensity change in the original signal and vice versa (10).

The reconstruction (inverse Fourier) of the analyzed signal can be achieved by the sum of these harmonics which are weighted through the transformation coefficients. These coefficients tell how are the magnitude and shift (phase) that each harmonic (sinusoid) has (1).

In image processing, the filters that can be used are either low pass filters for blurring or high pass filters for sharpening. In this research an ideal low pass filter (ILPF) which is a circle is used as a mask for specifying which frequencies (harmonics) to remain and which to be filtered out. The ILPF is preferable to soft the image or to use its complement for sharpening (10).

Eq.3 tells how much each harmonic of specific frequency is presented in signal $f(x)$ (image). Thus according to this equation the shift for each harmonic occurs through the time (spatial) domain using increasing variable value (x). Whereas the scaling for such harmonics is done by applying the same equation with increasing value of the frequency variable (u). Then the summation operator acts to compose the resulted harmonics (sine and cosine for example) of multiple scales (frequency) and various phases (shift) that are represented in the transformation coefficients in order to reconstruct the original signal (8). So it can be briefly concluded that any signal in the spatial (time) domain is nothing more than a linear combination (summation) of multiple harmonics with different amplitudes and frequencies (8, 9).

According to (1) the resulted coefficients will include real and imaginary parts eq.5. They both are used to infer how much the magnitude is and the phase (θ) for each sinusoid (basis function).

$$M = \sqrt{(Real)^2 + (Imaginary)^2} \dots 6$$

$$\theta = \tan^{-1} \left[\frac{Imaginary}{Real} \right] \dots 7 \quad (1)$$

The function $F(u)$ is periodic and conjugate symmetric which means that both positive and negative side of this function are symmetric, thus it is sufficient to know one value in terms of other which leads to less computations (9). The Fourier transformation of a single wave will produce spectrum with a single positive frequency which mean that such a wave include just one frequency (11).

As any function can be represented through its even and odd parts, here the cosine and sine functions represent the even and odd parts respectively.

A simple example of decomposing function $f(t)$ into its even and odd parts is given in Fig. 3.

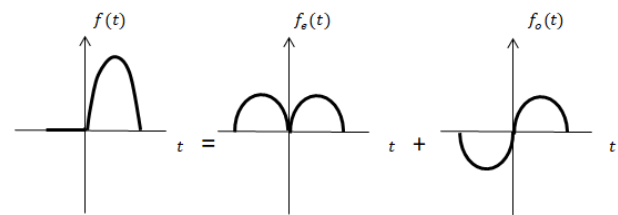


Figure 3. the even and odd component of a function $f(t)$.

Wavelet transformation:

Fourier spectrum gives global information about the signal but it doesn't give any information for that signal within specific period of time. In contrast to this time domain informs what happen within any specific time interval without global information. So it is urgent to find a method that can encompass the two types, global (frequency) and localization (time) information (1). The technique that is suitable for providing such information is then the Wavelet transformation (12). Wavelet transformation can be conducted using any of such basis functions like Daubechies, Morlet, Meyer, Mexican hat and Haar which is the first and simplest function used for Wavelet. Of course the type of such basis function is application dependent. As noted here these basis functions are the correspondences of cosine and sine in Fourier transformation.

Wavelet transformation has two recognizable functions which are Wavelet and scaling functions. It also can be recognized as having good spatial and frequency localization properties. It decomposes the image into many multi-resolution components due to the use of low and high pass filters, these components are one approximation and three different details which have different frequency components as LL (low

low), LH (low high), HL (high low) and HH (high high) respectively (1, 12, 13).

Wavelet transformation depends on the uncertainty principle which states that it is impossible to get signal that has narrow spatial and spectral domain at the same time. The law which enforces a tradeoff between these two is:

$$\text{Signal duration} \cdot \text{frequency bandwidth} \geq \frac{1}{\pi} \dots 8 \quad (11)$$

According to this, it is necessary to decide which scaling level (signal duration) is adequate for an application when using Wavelet transform. In this research prove that they both provide good balance between the global information (frequency bandwidth) and local information (signal duration).

Through the use of Wavelet transformation it is possible to reduce the noise through removing the small details which may correspond to noise without affecting the other details that are related to edges.

Morphology

It is an image processing which can be used in order to extract some ROI features like skeleton, convex hull... etc. or it can be used as a pre/post processing tool in order to enhance these extracted regions.

For the sake of such these operations, morphology uses what is called a structuring element (SE) which encompasses a set of elements with one of them as a center. This SE is used in a matter just like correlation or convolution operations which are used in spatial domains. Such that the sliding technique is also used here, by putting the SE center over the region boundaries and recursively slide over its pixels till it visits all the region pixels (14).

The two main morphology operations which other high level operations depend on are erosion and dilation.

Erosion can be used in order to shrink ROI. Mathematically it can be given by:

$$A \ominus B = \{Z \mid (B)_z \subseteq A\} \dots 9$$

Where B represents the SE, while A is the ROI.

Dilation can be used in order to enlarge regions. Mathematically it can be given by:

$$A \oplus B = \{Z \mid (B^{\wedge})_z \cap A = \emptyset \dots 10$$

where B^{\wedge} represents reflection of B about its origin.

Opening is a higher level operation which acts to eliminate the region's tiny salient and smoothing its boundary. It can be given as:

$$A \boxminus B = (A \ominus B) \oplus B \dots 11$$

As a result it consists of two consecutive operations that are Erosion for the region A by SE B followed by SE dilation of the result with the same SE (14).

Frame difference for moving object detection

Frame difference is an approach for extracting moving objects in video frames, where two consecutive video frames are subtracted pixel wise. Thus in such case if there is any moving object happened to be exist in any of these frames, will be extracted as the subtraction result (15).

Proposed algorithm

The algorithm of this work can be described through the following steps: first two consecutive frames should be read. Then they need to be filtered through a spatial or frequency domain filters. To wipe the static objects from the resulted frames, two consecutive frames subtraction should be conducted, considering remaining as moving objects. Any pixel with intensity less than predetermined threshold (*th*) must be removed. Arbitrarily multiple threshold values have been tried according to varied elected frequencies (the no. of elected frequencies increase in direct proportion with the increasing radius of the circle in case of fourier transform, and in inverse proportion with the wavelet levels) to decide which of them is adequate in each case. The primary moving object mask can be achieved by converting the resulted image into binary image. Performing an opening morphology operation (as described above) with adequate structuring element (disk shape with radius of 1 pixel) is necessary to get the final moving object mask. This mask can be used as a reference in order to circulate the corresponding position of this mask in the original frame and declare it as moving object.

In the first step deliberately two successive frames are used in order to ensure accurately no detail has been ignored. Figure 7 shows such these two consecutive frames.

Unfortunately in some cases not all video frames may be noiseless. Therefore in step tow this research suggests and implements variant spatial and spectrum (frequency) domain filters for the sake of removing such this noise. The ordinary spatial domain filters (mean and median) are used here for the purpose of performance comparison with the frequency domain filters. Threshold value selection in these spatial domain filters for noise removing depends on pixel intensity values.

However in the case of frequency domain filters the matter may be somehow different, where the consequence of the frequency transforms (Fourier/Wavelet) is frequency coefficients which requires different threshold selection approach. For example the result of applying the Fourier transform on a frame is a matrix of frequency coefficients, with low (which represents the most important

frame information) frequency components concentrated in the center of this matrix and high frequency components (frame's edges and noise) as it goes away outward this center. Therefore a circle shape mask as in Fig. 4 with adequate radius is used to extract (pick) the most important frame information (low frequencies) ignoring the others and ultimately applying the threshold on the pixels intensity values after applying the inverse Fourier transform.

Another approach is conducted by utilizing the Wavelet transform to get the frequency coefficients, then applying an appropriate threshold on these coefficients in order to preserve the most approximation sub band coefficients (low frequencies which represent the most important frame information) and neglecting some of the other sub band coefficients (some edges and noise).

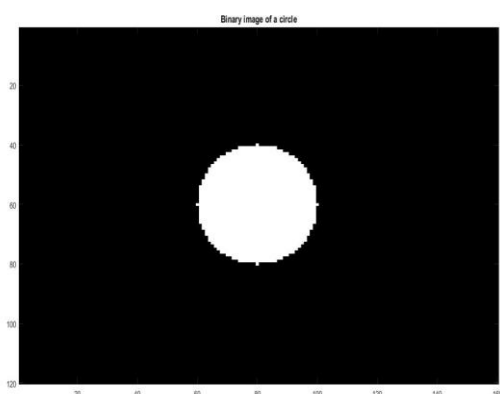


Figure 4. A circle mask used to retrieve just the low Fourier frequencies

Step three is intended to capture any tiny change in around (which indicate the presence of a moving object(s)). This can be achieved by subtracting the resulted two consecutive frames of step two. Thus the pixels in first frame will be subtracted from the corresponding pixels in the second frame, resulting in an image with just moving object(s) on approximately black background, an example is shown in Fig. 5.

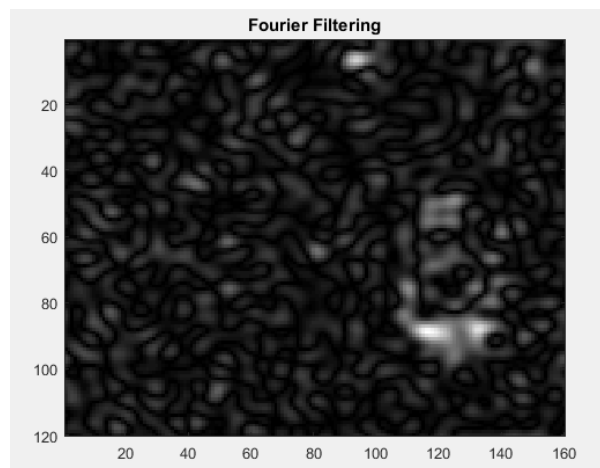


Figure 5. the resulted difference image after filtering of the two frames by utilizing low frequency Fourier coefficients.

This resulted image can be handled with a predetermined threshold in order to preserve just pixels with values higher than this threshold, Fig. 6 shows the consequence of this step after converting to binary form and application of an opening operation.

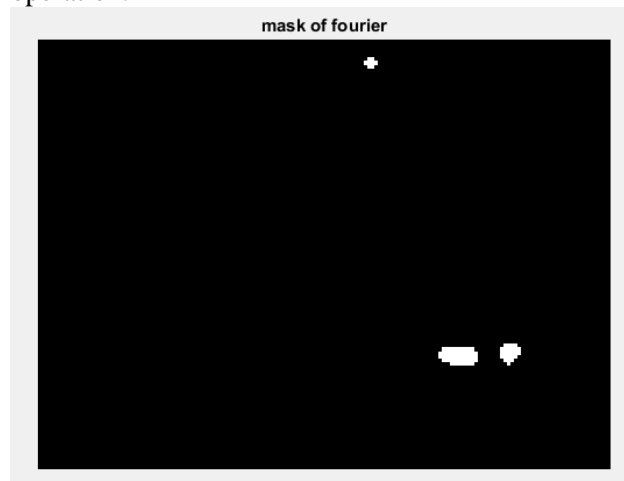


Figure 6. moving objects mask

As this mask encompass just the impact of the moving object(s). It can be superimposed over the original frame to refer to the location(s) of this/these moving object(s). It is possible to use bounding box (es) around such locations to indicate to moving object(s).

Results and Discussion:

The proposed technique depends on the difference of two successive frames in order to find moving objects. An example of such these two frames is shown in Fig.7.

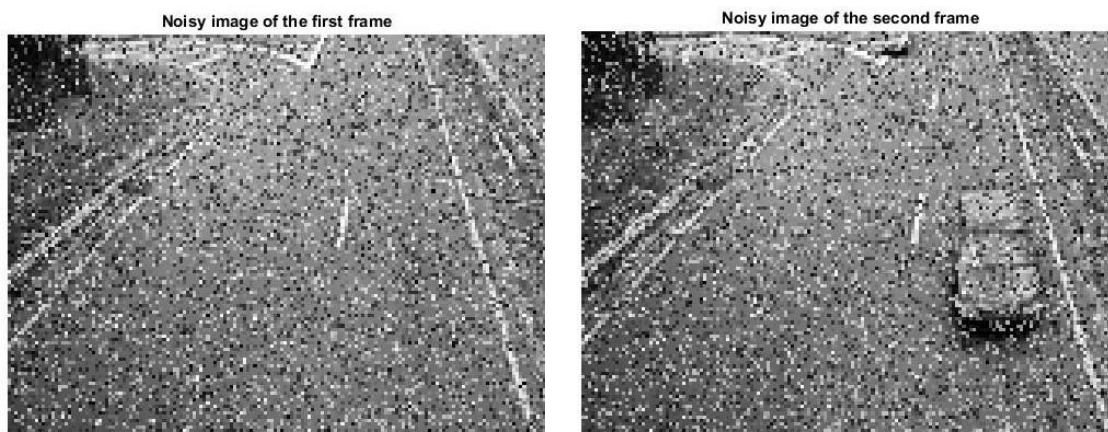
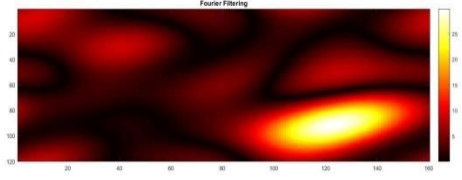

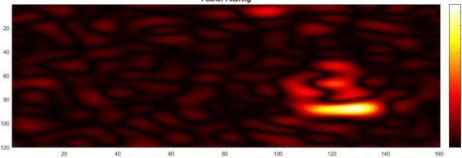

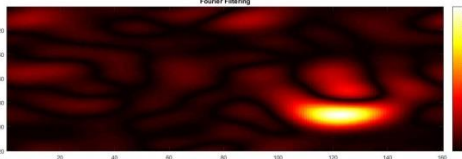

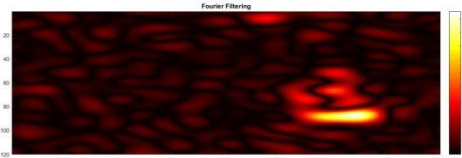

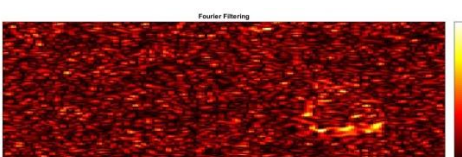
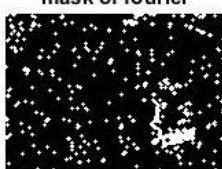


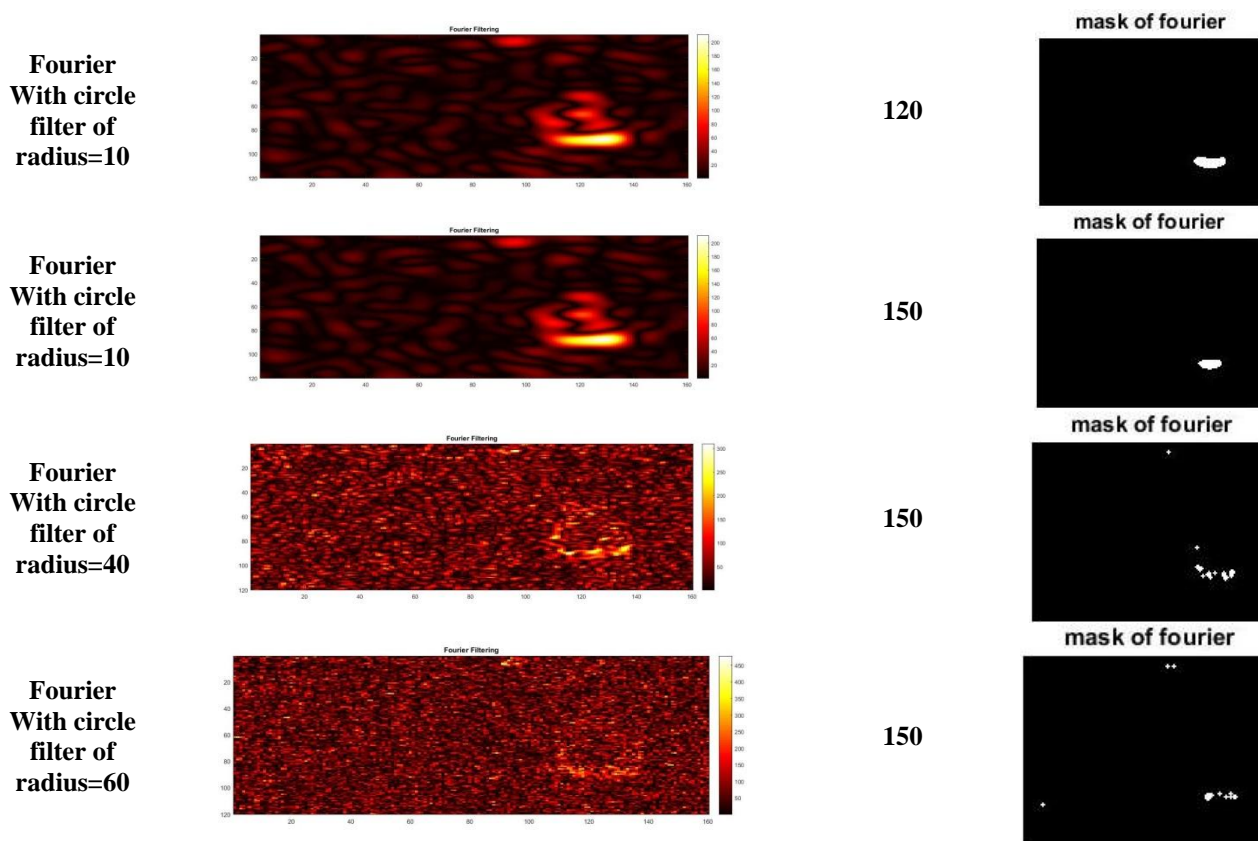
Figure 7. two successive noisy frames with moving objects (cars)

In Fourier transformation, an ILPF (circle) is used in order to filter out some of unwanted high frequencies (which may include noise). There is a

positive relation between the circle area and the ratio of the remaining (unfiltered) frequencies and hence the cutoff threshold as shown in Table 1.

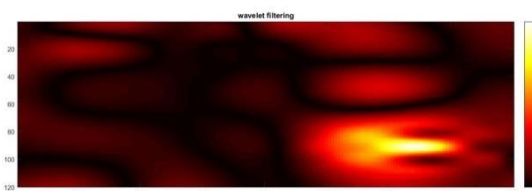
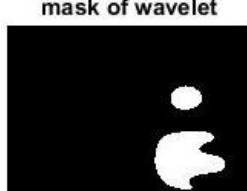
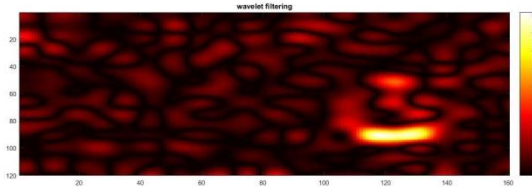
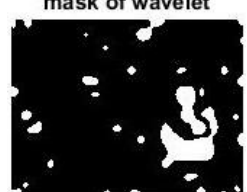
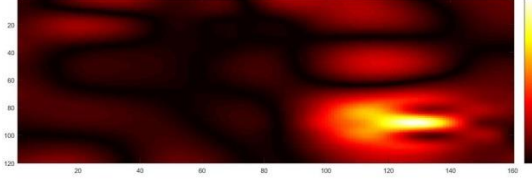

Table 1 filtering with Fourier transformation using different thresholds and limited number of elected frequencies. Where hot color map reflects the frame pixels intensity values after picking (selection) of the appropriate frequency coefficients and applying inverse frequency transform (in case of Fourier and wavelet), and frame pixels intensity values after applying filter (in case of mean and median filters).

Filtering method	Hot color map	Threshold	Mask
Fourier With circle filter of radius=2		20	
Fourier With circle filter of radius=10		30	
Fourier With circle filter of radius=5		55	
Fourier With circle filter of radius=10		55	
Fourier With circle filter of radius=40		55	

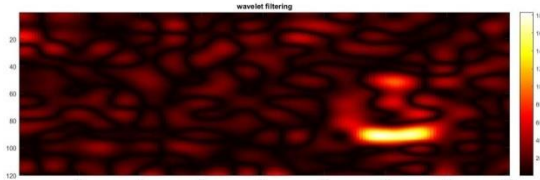


A similar behavior is found through using the Wavelet transform but with an inverse relationship between the Wavelet decomposition level and the cutoff frequency (threshold) as shown in Table 2.

Table 2. filtering with Wavelet transformation using different thresholds and limited number of elected frequencies. The threshold is to determine the elected values.

Filtering method	Hot color map	Threshold	Mask
<p>Wavelet 5 levels</p>		<p>20</p>	<p>mask of wavelet</p> 
<p>Wavelet 3 levels</p>		<p>30</p>	<p>mask of wavelet</p> 
<p>Wavelet 5 levels</p>		<p>30</p>	<p>mask of wavelet</p> 

Wavelet 3 levels

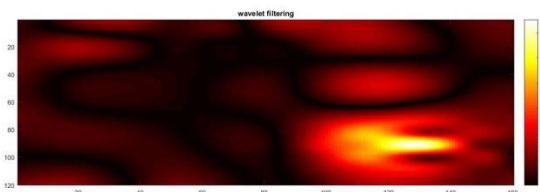


55

mask of wavelet



Wavelet 5 levels

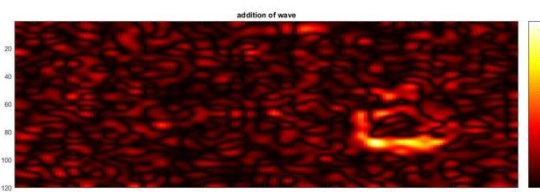


55

mask of wavelet

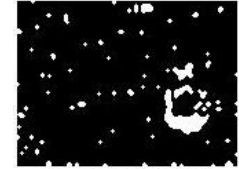


Wavelet 2 levels

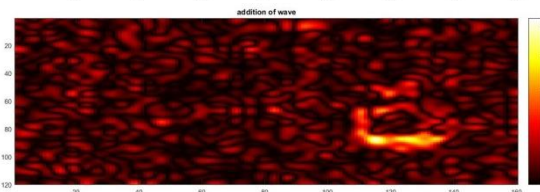


55

mask of wavelet



Wavelet 2 levels

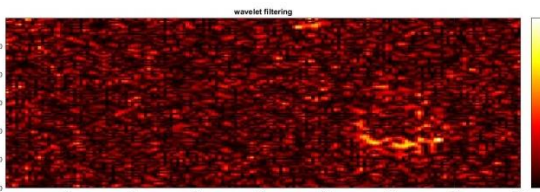


120

mask of wavelet



Wavelet 1 level

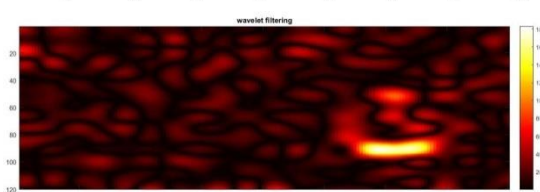


150

mask of wavelet



Wavelet 3 levels

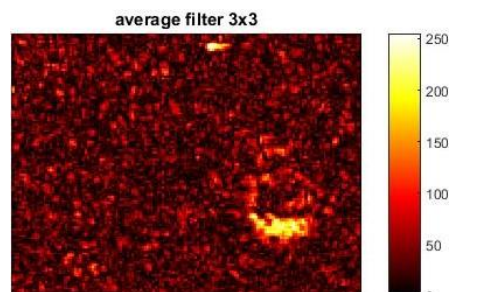


150

mask of wavelet

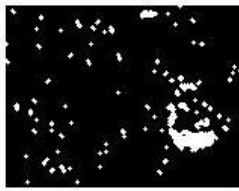


Mean 3x3

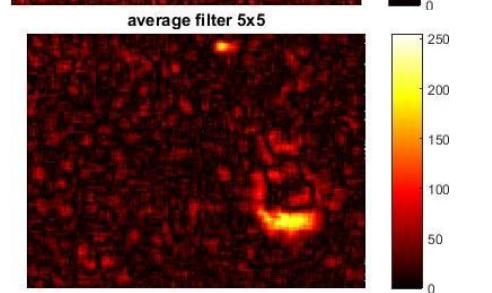


55

mask of average (mean) 3x3



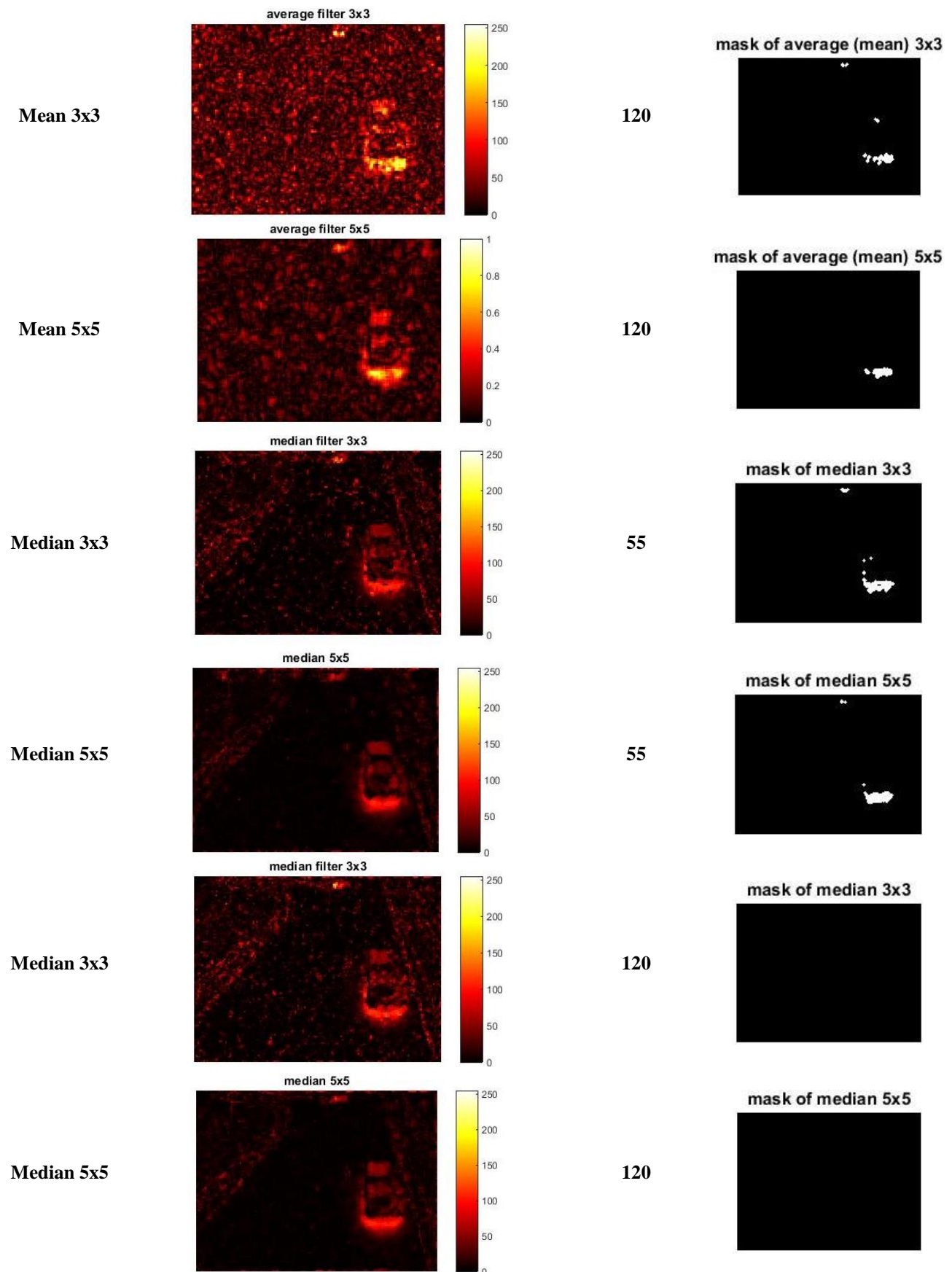
Mean 5x5



55

mask of average (mean) 5x5





Entropy can be defined as the amount of information and noise that exist in the signal, or it is the number of times that the system can be ordered differently. As long as the entropy is high, the

system instability is also high as well as there will be lower system harmony. So in image for instance as the pixels values are close to each other, the entropy scalar value will be lesser and vice versa.

Image entropy can be calculated using the following formula: (14)

$$E = - \sum_{k=0}^{L-1} P(r_k) \log_2 P(r_k) \dots 12$$

Where r_k is the K_{th} intensity value, $P(r_k)$ is the probability of occurrence of this intensity level.

High standard deviation values means that there is a high dispersion between these values. According to image concepts, taking the low frequencies means that it is intended to take the uniform (close intensity values) regions. Therefore the consecutive frames subtraction results small values due to the subtraction of these corresponding uniform regions in both under consideration frames. So the selected threshold must be small in order to be suitable for such these before mentioned resulted subtraction values. But in case of taking low and high frequencies, which means that there will be uniform regions in addition to abrupt changes causing the subtracted values to hold both small and high values which affect the threshold selection toward higher values. Standard deviation can be given by (14):

$$\sigma = \sqrt{\sum_{k=0}^{L-1} (r_k - m)^2 P(r_k)} \dots 13$$

Where m is the mean intensity value.

All the used filter types tend to decrease differences between neighbor pixels values. This is to diminish noise just like blurring does. This may sound to decrease the Std, but in fact this at the same time create groups (blocks of pixels with each of such groups having pixels sharing the same value). Thus a lot of the original frame pixels values to disappear causing increasing Std value. This Std value tend to increase whenever these regions (groups) increase. For example vector A=[1 2 3 4] has a Std of 1.2910 which is lower than the Std of vector B=[1 1 4 4] that has a Std of 1.7321. this behavior tend to converse when these regions grow up more and more, because in this case a lot of pixels will have the same value in each region which in turn means low Std. high Stds give better flexibility in the selection of a threshold from a wider range of values that the low Stds which limits (narrows) this range. Figure 8 shows the normal distribution of standard deviation.

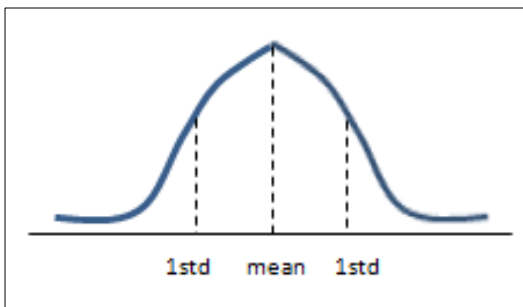
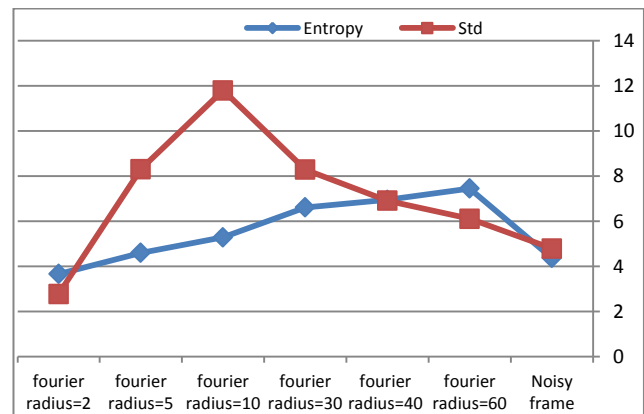
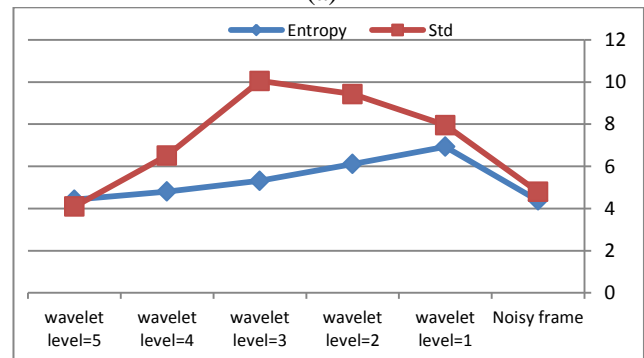


Figure 8. Std distribution shape

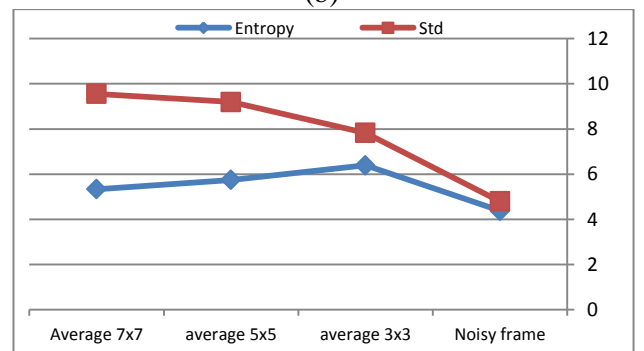
As the circle shrink for Fourier and the decomposition level increases for Wavelet, the elected frequencies decrease, which mean that the selected frequencies will be the low frequencies where there is no abrupt changes in the values. So the resulted frames won't include a lot of edges as well as noise. Then the difference of the two consecutive filtered frames will include growing unified values regions leading to entropy decreasing as well as the thresholding value and vice versa as shown in Fig 9. Then the whole matter is a tradeoff between all of these things.



(a)



(b)



(c)

Figure 9. Entropy & Std with increasing encompassed frequencies for (a) Fourier transform, (b) Wavelet transform, (c) Average filter.

Acknowledgement:

It is our pleasure to express our appreciation and thanks for Computer Science Department/ College of Science/ University of Mustansiriyah/ Baghdad/Iraq and Department of Computer techniques Engineering/Imam Kadhim Faculty of University Islamic Sciences for the valuable assistance and encouragement to accomplish this research

Authors' declaration:

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are mine ours. Besides, the Figures and images, which are not mine ours, have been given the permission for re-publication attached with the manuscript.
- Ethical Clearance: The project was approved by the local ethical committee in Almustansiriyah University.

References:

1. Umbaugh S. Digital Image Processing and Analysis. Book. CRC press, 2011; 2nd ed.
2. Naik A, Barot N, Brahmabhatt R, Dahiya V. Comparison Between Spatial and Frequency Domain Methods. IJERMT. 2015; 4(12):45-50.
3. Dedeo glu, Y. Moving Object Detection, Tracking and Classification for Smart Video Surveillance. MSc Thesis, Bilkent University, 2004.
4. Dong L, Ganesh S. Minimum Delay Moving Object Detection, IEEE CVPR. 2017; 4250-4259.

5. Pranali A, Ajay A. Review on Automatic Fast Moving Object Detection in Video of Surveillance System. IJSRST. 2017; 3(3):545-549.
6. Pavankumar K, Satone M. Moving Object Detection Survey using Background Detection Methods. IRJET. 2017; 4(5): 1836-1838.
7. Shilpa, Prathap L, Sunitha R.A Survey on Moving Object Detection and Tracking Techniques. IJECS. 2016; 5(5): 16376-16382.
8. Ayush B, Yonina E. Sampling and Super-resolution of Sparse Signals Beyond the Fourier Domain. JLCF. 2018; 67(6): 1508-1521.
9. Svoboda T, Kybic J, Hlavac V. Image Processing Analysis and Machine Vision a Matlab Companion. Book, Thomson/West, 2008.
10. Shaikh S, Choudhry A, Wadhvani R. Analysis of Digital Image Filters in Frequency Domain. IJCA. 2016; 140(6): 12-19.
11. Sonka M, Halvák V, Boyle R. Image Processing, Analysis, and Machine Vision. Book. Thomson/West 2008; 3rd ed.
12. Dipalee G, Siddhartha C. Discrete Wavelet Transform for Image Processing. IJETAE. 2015; 4(3): 598-602.
13. Narjes K, Rania F, Mohamed B. A Fast Selective Image Encryption Using Discrete Wavelet Transform and Chaotic Systems Synchronization. ITC. 2016; 45(3): 235-242.
14. Gonzalez R, Woods R. Digital Image Processing. Book, Pearson Education. Inc., 2008; 3rd ed.
15. Sharma R, Gupta S. A Survey on Moving Object Detection and Tracking Based On Background Subtraction. OJIDDS. 2018; 2018(1): 55-62.

تحسس الكائنات المتحركة بالاعتماد على مجال التردد

بلسم ضياء مجيد²

جلال حميد عواد¹

¹ قسم الحاسوب، كلية العلوم، الجامعة المستنصرية، بغداد، العراق.
² قسم هندسة تقنيات الحاسوب، كلية الإمام الكاظم للعلوم الإسلامية الجامعة، بغداد، العراق.

الخلاصة:

في هذا البحث ، يتم استخدام تقنية مقترحة لتحسين أداء تقنية فرق الاطار لاستخراج الكائنات المتحركة في ملف الفيديو. أحد العوامل الأكثر فعالية في انخفاض الأداء هو وجود ضوضاء ، والتي قد تتسبب في تحديد هوية الكائنات المتحركة بشكل غير صحيح. لذلك كان من الضروري إيجاد طريقة لتقليل تأثير الضوضاء هذا. يمكن استخدام المرشحات التقليدية المتوسطة والوسيطية في معالجة مثل هذه المواقف. ولكن هنا في هذا العمل ، ينصب التركيز على استخدام النطاق الطيفي من خلال استخدام تحويلات فورييه والموجات من أجل تقليل تأثير الضوضاء. أثبتت التجارب والإحصاءات (الانتروبيا ، الانحراف المعياري) أن هذه التحويلات قادرة على التغلب على مثل هذه المشاكل بطريقة أنيقة.

الكلمات المفتاحية: تحسس الاجسام المتحركة، تحويل الموجات، مرشح المتوسط، تحويل فوريير، فرق الاطار، مرشح الوسيط.