

Eulerian video magnification: a review

Haider Ismael Shahadi¹, Hayder J. Albattat², Zaid Jabbar Al-allaq³, Ahmed Toman Thahab⁴

¹Electrical and Electronic Engineering, University of Karbala, Iraq

^{2,3}Communication Techniques Engineering, Al-Furat Al-Awsat Technical University, Iraq

⁴Biomedical Engineering, University of Babylon, Iraq

Article Info

Article history:

Received Jul 19, 2019

Revised Oct 21, 2019

Accepted Nov 11, 2019

Keywords:

Eulerian video magnification

Pyramid decomposition

Spatial filters

Temporal filters

Video processing

ABSTRACT

Many important subtle changes in the environment are invisible to the naked human eyes. These subtle changes occur because of colour variations, such as blood flow in a human face that leads to face colour change, or motion variations, such as vena movement under human skin and vibration of buildings. The human eye requires optical microscopes to detect these variations. Alternatively, new technologies, such as high-speed imagery and computer processing, can be used to detect these variations. These computerised microscopes depend on computation rather than optical amplification to amplify subtle colour and motion changes in videos. The most popular technique to achieve computation-based microscope is the Eulerian video magnification (EVM). However, several challenges in EVM still need to be solved to meet the requirements of real time and video quality. This paper presents a comprehensive study of EVM methods and reviews the related literature. The strengths and drawbacks of existing works are discussed, and the important research fields and challenges in the area of EVM are concluded.

Copyright © 2020 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Haider Ismael Shahadi,
Electrical and Electronic Engineering Dept,
University of Karbala, 56001, Karbala, Iraq.
Tel.:009647814632232
Email: haider_almayaly@uokerbala.edu.iq

1. INTRODUCTION

In spite of the fact that the visual system of a human has a limited perceptibility in either spatial or temporal domains [1], technology can reveal many of the imperceptible signals that fall outside human perceptual range. Many of these signals carry important information, such as a slight difference in human skin colour according to blood circulation. Although this difference is concealed to our eyes, it can be used to estimate the heartbeat number [2-3]. Similarly, motion, which is invisible to the naked eye and has low spatial energy, can be revealed by magnification, which allows us to exploit interesting behaviour [4]. Figures 1 and 2 provide examples of magnification for colour and motion variations, respectively. Figure 1(a) shows a man's face as it appears to the naked eye, and no changes can be seen. The processed images in Figure 1(b) clearly reveal the variation in the colour of his face that is caused by his heartbeat. Similarly, Figure 2(a) shows three frames of an eye that show no change, and Figure 2(b) shows a magnification of the small pattern of the eye.

Our environment is crowded by small and significant temporal variations. Several approaches and schemes have been evolved to visualise these variations in either motion or colour, thereby resulting in the so-called computerised microscope. Computerised microscopes depend on computation rather than optical magnification to amplify subtle colour and motion changes in high-speed or ordinary videos. The success of algorithms has supported the development of modern techniques that discover unnoticeable signals in videos.

The capability to magnify unnoticeable variations in an imaging video has opened the door for applications in biology, healthcare and mechanical and material engineering.

In 2005, a research group from the laboratory of computer science and artificial intelligent in the Massachusetts Institute of Technology (MIT) proposed a video magnification algorithm based on cluster trajectories [4]. The algorithm was used to amplify subtle colour or motion changes over time, which made invisible changes visible. In [5], the authors proposed using a cartoon animation filter to create visible motion construction. These algorithms depend on a Lagrangian perspective, in which the pixel path is temporally tracked over video frames. However, these algorithms are computationally costly because they are based on accurate motion estimation. Moreover, they cannot be verified artefact-free at areas of dense motions. By contrast, Eulerian based-methods amplify the intensities of pixel variation over time in a multiple scale manner. In Eulerian methods, motion magnification does not clearly estimate motion but rather extends it through amplifying temporal colour changes at fixed positions. The Eulerian methods are similar to optical flow algorithms in using differential approximation form [6-7].

The approach of Eulerian video magnification (EVM) was firstly proposed by MIT research group in 2012 [8]. The basic methodology of EVM considers the time series of pixels and amplifies any variation in a specified interest band of temporal frequency. For example, the selected frequencies in Figure 1 consist of plausible human heart rates to amplify a temporal band. The amplification exposes the dissimilarity of redness as blood flows through the face [9-10].



Figure 1. Colour variation magnification [8]



Figure 2. Subtle motion variation magnification [9]

Since 2012, EVM has been a popular research area in several interesting applications, such as extracting the depth and velocity of hot air [11], human feeling detection [12], Android smartphone software [13], plasma physics [14], sound reconstruction from a distance by the vibration measurement of an item in a high-speed video [15], rescue [16], biology [17], mechanical engineering [18] and civil engineering [19]. EVM may also have significant potential in diagnostic and monitoring in medical applications. Most medical applications depend on the recovery of temporal features using the capabilities of colour amplification of EVM, such as measuring vital signs without touching patients [10, 20-27] and revealing otitis media in infants [28].

The EVM technique can be classified into linear- and phase-based EVM. In linear approaches, motion in video is linearly proportional to the intensity variation over a first-order expansion of Taylor series. A video sequence is considered an input, spatial decomposition is applied and frames are filtered by a temporal filter. The produced temporal region is amplified to expose hidden information. Although this procedure is simple and rapidly detects small motion variations, it easily breaks down when the magnification factor is large because the Taylor approximation becomes inaccurate. To overcome this problem, phase-based magnification (PVM) replaces the linear approximation with a Fourier decomposition by complex steerable pyramid [29]. The variations in phase of the pyramid coefficients over time are proportional to the motion in different video frames. These variations can be temporally processed and then amplified to visualise the motion. In contrast to the linear-based method, the phase-based method has a higher complexity and longer processing time, but it can support larger amplification of the motion. Eulerian

linear and phase-based approaches operate faster and suffer less noise than the Lagrangian-based video magnification. However, EVM approaches do not work well for large and arbitrary motions.

In this paper, we present a comprehensive study of EVM methods and review the latest work in this area. We also compare the existing works in terms of quality, speed and amplification factor. The rest of the paper is organised as follows. Section 2 describes EVM. Section 3 explains a linear-based EVM. Section 4 reviews the existing works that use linear EVM. Section 5 describes a phase-based EVM. Section 6 reviews the existing works that utilise phase-based EVM. Section 7 presents a brief comparison of different existing works. Section 8 elaborates the conclusions.

2. EVM

EVM can reveal and amplify the small motions and changes in videos. EVM is applied to every level within a pyramid and not over the original pictures because the goal is to amplify the pyramid levels that contain the movement frequencies. This technique then multiplies the desired frequencies by a factor known as magnification factor α , which is defined by the user. EVM amplifies the actual motion and allows us to recognise movements that are undetectable to the naked eye. The magnified values of the desired frequency are added back to the not magnified ones of the same level to obtain the final video with exaggerated motion.

Although many methods use the same general principle as that of EVM, they differ in the way they work. One of the major differences amongst these methods is the type of pyramid utilised in the algorithm. The linear video magnification (LVM), which was presented by Wu et al. [8], applies the Laplacian pyramid decomposition technique to the input video to decompose the video sequence according to spatial frequency, which is followed by temporal filtering. The resulting output signals of this operation are then magnified by a factor and added back to the signals that are entered to the temporal filter. However, this method supports low magnification factors. To solve this problem, Wadhwa et al. [29] proposed a Eulerian method based on a technique called 'complex steerable technique' [30-31], which was inspired by phase-based optical flow [32-34]. This method accepts large magnification factors, possesses fewer artefacts and introduces less noise than LVM. However, this method requires a longer time processing due to the complexity of piping representation of the steerable pyramid. This method can be over 21 times as long as the LVM technique. In [29], the authors developed their previous work by using a new pyramid in [35-36], which they called the Riesz pyramid, to reduce execution time. Liu et al. [37] proposed a way to improve LVM after processing, which is called enhanced EVM (E2VM).

E2VM supports magnification factors with greater and less noise than LVM. The efficient motion magnification system (EMMS) method has been developed to improve processing speed [38], which depends on wavelet decomposition. This method supports large amplification factors, improves the speed of implementation and reduces noise. Section 7 (comparisons) summarises the main differences amongst these methods.

3. LINEAR-BASED EVM

Small motion amplification can be achieved through processing [4-5]. In optical flow, motion magnification can be produced via temporal processing by using the first-order Taylor series expansions [6-7]. The goal of EVM is to process the time series of colour values for each pixel in spatial domain independently by applying a standard 1D temporal signal processing to each time series to amplify a specific band of interest temporal frequencies. The input video frame is decomposed into distinct spatial frequency bands using a full Laplacian pyramid [8, 39-42]. The Laplacian pyramid is a data structure in which an image is downsampled at successively sparser densities until no further downsampling is possible. The Laplacian pyramid depends on an analysis pyramid for videos that is based on a Gaussian pyramid.

The Laplacian pyramid method for analysis video processing time has become less popular than before. [8-9] showed that this method becomes useless as the magnification factor increases because noise increases linearly with the increase in magnification factor. This method is useful in magnifying colours when the magnification factor is small. Figure 3 shows a working LVM mechanism.

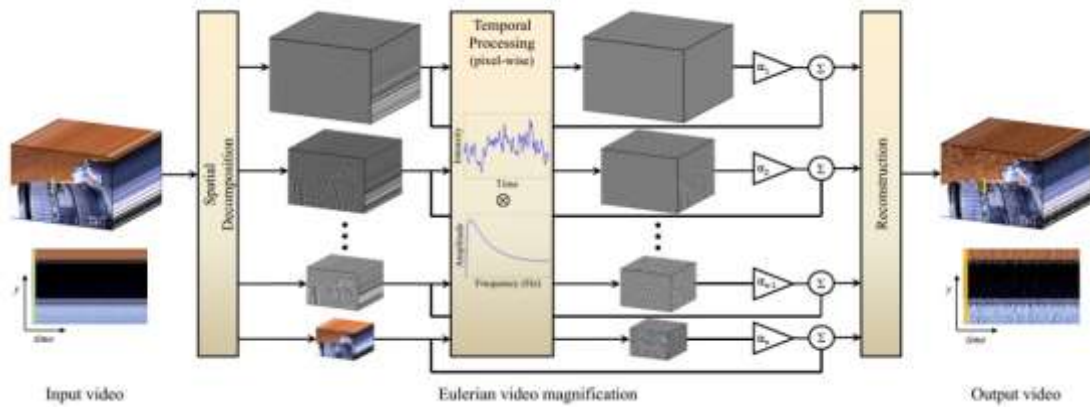


Figure 3. Overall structure of the linear-based-EVM [8]

3.1. Mathematical Analysis for Linear-Based EVM

This section explains the linear-based EVM mathematically. $I(x, t)$ denotes the intensity of the input image at a certain position x and time t . Observed intensities can be expressed with respect to a displacement function $\delta(t)$ given that translational motion is applied to the image. Therefore, $I(x, t)$ can be written as $I(x, t) = f(x + \delta(t))$ and $I(x, 0) = f(x)$. The main objective of motion magnification is to synthesise the signal, that is,

$$\hat{I}(x, t) = f(x + (1 + \alpha)\delta(t)) \quad (1)$$

where, α is an amplification numerical factor. The video frame is approximated on the basis of the expansion of the first-order Taylor series at time t . The displaced frame $f(x + \delta(t))$ in the first-order Taylor expansion around x is given by

$$I(x, t) \approx f(x) + \delta(t) \frac{\partial f(x)}{\partial x} \quad (2)$$

$B(x, t)$ is the output of a temporal bandpass filtering process for $I(x, t)$ in each x -position. The motion signal $\delta(t)$ is within the passband boundary frequencies of the temporal filter, that is,

$$B(x, t) = \delta(t) \frac{\partial f(x)}{\partial x} \quad (3)$$

$B(x, t)$ is amplified by amplification factor α and added back to the original frame ($I(x, t)$). The resulted frame is computed in accordance with,

$$\tilde{I}(x, t) = I(x, t) + \alpha B(x, t) \quad (4)$$

As shown in (2), (3) and (4) are combined, and we obtain the approximated amplified frame as,

$$\tilde{I}(x, t) \approx f(x) + (1 + \alpha)\delta(t) \frac{\partial f(x)}{\partial x} \quad (5)$$

The expansion of the first-order Taylor holds for the amplified motion signal, $(1 + \alpha)\delta(t)$, and relates to the amplification of the temporally bandpassed signal to motion magnification. As shown in (6) denotes the processed output.

$$\tilde{I}(x, t) \approx f(x + (1 + \alpha)\delta(t)) \quad (6)$$

Figure 4 demonstrates the amplification process for a single sinusoid. A cosine wave with a low frequency and a relatively small displacement ($\delta(t)$) can be approximated by an expansion of the first-order Taylor series. This procedure results in an efficient approximation for the signal at time $t + 1$.

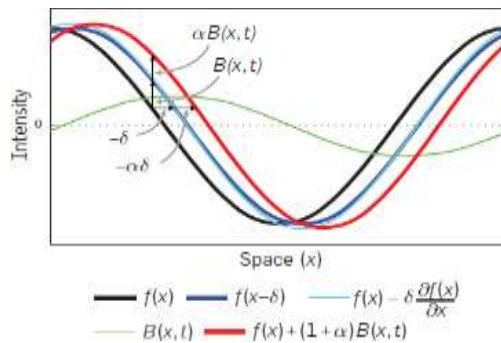


Figure 4. Illustrations of amplifying spatial translation after temporal filtering

3.2. Pyramid Decomposition

A pyramid is the process of decompositions of images in different scales spatially. The two most popular pyramid decomposition approaches are Gaussian and Laplacian. Gaussian pyramid is constructed by smoothing the original image with a Gaussian filter and then scaling it down. A Gaussian pyramid consists of a sequence of lowpass, downsampled pictures. The Gaussian pyramid is similar to a Laplacian pyramid; at a specific level of each image of the Laplacian pyramid is the difference between two corresponding neighbouring levels of the Gaussian pyramid. The smallest level is merely preserved. As a result, the difference images can be used to reconstruct the original image. However, Laplacian pyramid can be assumed to be a sequence of bandpass, downsampled images [39]. Figure 5 illustrates the Gaussian and Laplacian pyramid decompositions.

3.3. Temporal Filter

A temporal filtering process is applied to the series of the temporal pixels in each spatial band of the spatial pyramid construction to extract the interest frequency bands. A temporal bandpass filter is used to extract motions or signals that are intended to be amplified.

According to the application utilised in the algorithm, users should be able to control the frequency band interest. However, the frequency band can be automatically selected in some cases.

Filter selection also depends on the type of the application used. For example, a filter with a wide pass band is often preferred for motion magnification, whereas a narrow-passband filter is preferred for colour amplification, such as blood flow, because the latter results in less noise distortion. However, for real-time implementation for motion magnification and colour amplification, low-order IIR filters are convenient to use. Various frequency responses of several types of temporal filters are shown in Figure 6.

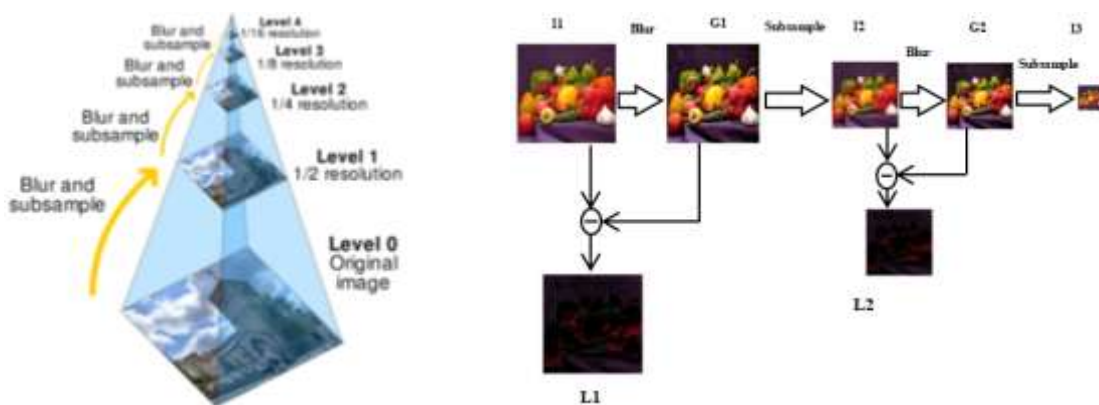


Figure 5. (a) Gaussian pyramid (b) Laplacian pyramid

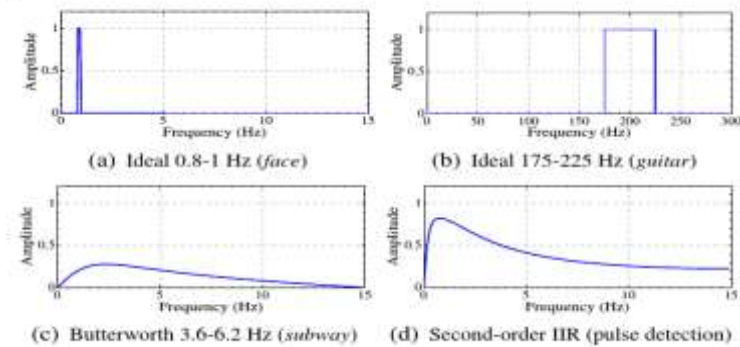


Figure 6. Temporal filters types that are used in [8-9]

4. EXISTING WORKS BASED ON LINEAR EVM

Several works in the field of EVM based on the LVM method have been published. In [22], the EVM method was used to measure the time of a pulse crossing without direct contact. The benefit of measuring pulse time crossing is to determine blood pressure, heart rate and stress. Two areas of study, namely, wrist and neck, were selected to calculate the pulse transit time. The results of this method were verified by comparing them with the results of the devices used to determine the pulse, such as the electrocardiogram; the results were close.

In [23], the authors used Eulerian method to extract the form breath wave of movement from the upper body of different breathing patterns (normal breath, hold breath, quick breath and deep breath) rapidly and accurately from the video sequence. The video must contain the head and chest area for accurate results.

In [21], Eulerian technique was used to detect neonatal diseases with limited movement, such as epileptic seizures and life-threatening apnoea events. The disease category included epilepsy and epileptic seizures in neonates. This easy-to-use method can be used in a patient's home or in a hospital, specifically in the neonatal intensive care unit, to monitor newborn without devices connected to the infant.

In [20, 43] EVM was introduced to a new application. The method was used to measure wrist pulses. The method can also analyse the signal of the wrist pulse because vital information can be extracted from wrist pulse signals that affect a person's health status. As a result, the EVM method can be utilised to predict essential cardiovascular functions.

The authors in [44] adopted the EVM algorithm to detect and measure the soot propensity of a flame. Other authors have used EVM in face spoofing attacks because facial biometrics is vulnerable to fraud. The authors in [45] used Eulerian method for face recognition and detection of face spoofing techniques to improve the practicality of face biometrics.

Liu et al. [37] proposed an enhancement to EVM to improve the work of LVM. A pixel-level motion analyser was used to capture and amplify motion. The method uses LVM as an analyser for spatiotemporal motion. E2VM utilises image warping to magnify the temporal motion of the video based on the former motion mapping. Although the image wrapping technique provides an improved performance method to handle the frame noise in the postprocessing, the computation time required is also increased. However, the method supports larger amplification factors with less noise distortion and fewer artefacts than the LVM method.

5. PHASE-BASED EVM

The Eulerian LVM supports small factors of magnification. The main drawback of the Eulerian LVM method is the ability to amplify noise when the magnification factor is increased. The authors in [29] accordingly developed a motion-processing method that is based on complex-valued steerable technique [30-31] and inspired by phase-based optical flow [32-33] and motion without movement [41]. This method supports large amplification factors and has less noise distortion than linear-based EVM.

The phase-based EVM method does not increase spatial noise because the method modifies the phases instead of amplitudes. As a result, noise is translated rather than amplified when the amplification factor is increased. Although the computation of local motions can result from the phase variations of the complex steerable pyramids, increasing the phase variations by a multiplicative factor can amplify subtle motions. The modified video is reconstructed after amplifying the motion temporally.

The main difference between linear- and phase-based EVM is in the spatial decomposition used. The phase-based uses steerable pyramids. These pyramids result in an efficient and accurate linear decomposition for video frames in scale and orientation sub-bands. The basic functions of the transform are similar to Gabor wavelets [46], which are sinusoids windowed by a Gaussian envelope [30, 47-48].

The method links phase and motion by manipulating the phase variations in videos. The method also avoids the flow vector computation. However, this method requires high processing time. Figure 7 shows the working mechanism of PVM.

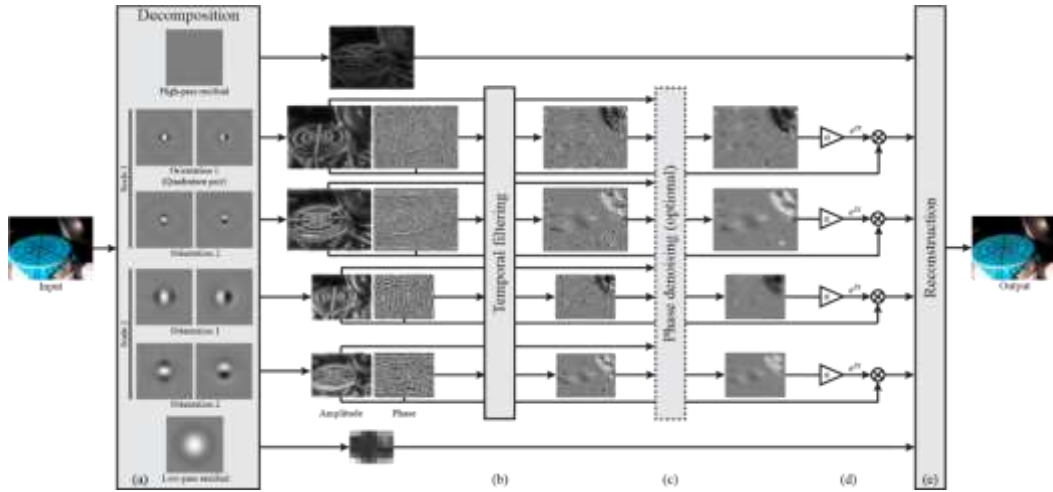


Figure 7. General structure of phase-based EVM [29]

5.1. Mathematical Analysis for Phase-Based EVM

The phase-based approach depends on complex-valued steerable pyramids given that local motions can be measured and modified. The intensity profile (f) of a 1D image under global translation over time (t), $f(x + \delta(t))$, for a movement function $\delta(t)$ is considered. A sequence is assembled with the motion modified, $f(x + (1 + \alpha)\delta(t))$, to a magnification factor (α).

The displaced intensity profile of the image $f(x + \delta(t))$ can be written as a sum of complex sinusoids using Fourier series decomposition, that is,

$$f(x + \delta(t)) = \sum_{\omega=-\infty}^{\infty} A_{\omega} e^{i\omega(x+\delta(t))} \tag{7}$$

where every band corresponds to one frequency (ω). The complex sinusoid is the band for frequency ω , that is,

$$S_{\omega}(x, t) = A_{\omega} e^{i\omega(x+\delta(t))} \tag{8}$$

The phase $\omega(x + \delta(t))$ contains motion information because S_{ω} is a sinusoid. This condition can cause motion manipulation by phase modification.

A DC balanced filter is used to filter the phase $\omega(x + \delta(t))$ temporally for separating motion in specific temporal frequencies. Only the DC component (ωx) is assumed to be removed. This assumption results in,

$$B_{\omega}(x, t) = \omega \delta(t) \tag{9}$$

The bandpassed phase $B_{\omega}(x, t)$ is multiplied by the frequency (ω), and the phase of sub-band $S_{\omega}(x, t)$ is increased by this amount to obtain the sub-band with the magnified motion.

$$\hat{S}_{\omega}(x, t) = S_{\omega} e^{i\alpha B_{\omega}} = A_{\omega} e^{i\omega(x+(1+\alpha)\delta(t))} \tag{10}$$

The result in (10) shows that $\hat{S}\omega(x, t)$ is a complex sinusoid, and the motion ratio between $\hat{S}\omega(x, t)$ and the input is exactly $(1 + \alpha)$ times. The motion-magnified video can be reconstructed by summing all the sub-bands to acquire the motion-magnified sequence $f(x + (1 + \alpha)\delta(t))$.

5.2. Spatial Filtering: Complex Steerable Pyramids

The steerable pyramid [30-31, 48, 50] is considered an over complete transform given that an image is decomposed according to spatial scale, orientation and position. Gabor wavelet [52] is one of the basic functions of the transform, including steerable sinusoids windowed by a Gaussian envelope.

A 2D discrete Fourier transform computation is used for all of the video frames, and spatial filters with different sizes and orientations are subsequently applied. This procedure produces a linear multiscale and -orientation image decomposition, which is called a steerable pyramid. An array of compound numbers is at each level of the pyramid and for each element; the amplitude and phase are computed. The complex-valued pyramid permits measurement of local amplitude and phase, which are exploited to process motion. Figure 8 shows the construction of the steerable pyramid.

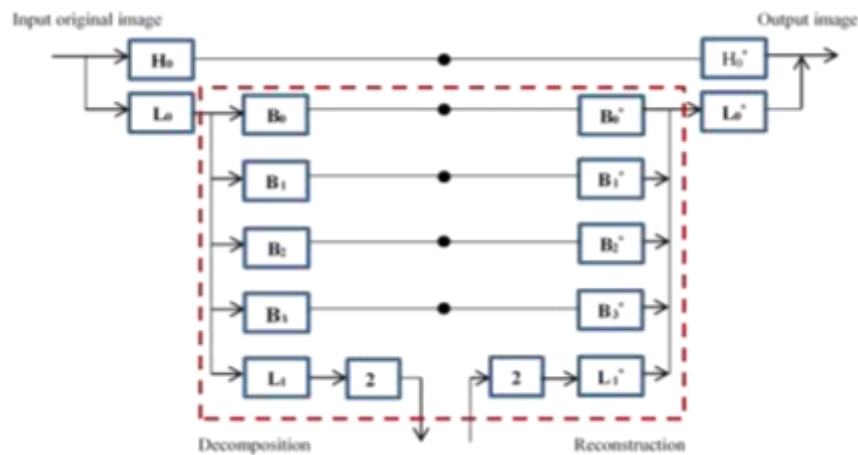


Figure 8. Steerable pyramid construction

5.3. Temporal Filter

This process aims to isolate motions in specific temporal frequencies to be magnified. When the pyramid is constructed, the phase on each spatial scale and orientation are isolated. Their differences are then determined, and a process of temporal bandpass filtering is applied. The key to improve the filter performance is to amplify either colour alterations or local phase variations, which can be achieved by maximising the signal-to-noise ratio of the temporal variations. Temporal and spatial filtering is applied to the variations to eliminate components that correspond to noise and maintain components that correspond to the signal for improving the signal-to-noise ratios. Temporal filtering can also isolate a signal of interest because various motions occur at different temporal frequencies

Narrowband linear filters in the temporal domain improve the signal-to-noise ratios for motions, such as respiration and vibrations, that particularly occur in the narrow range of frequencies. These filters can also separate object motions [53] that correspond to various frequencies, such as pipe vibration, which vibrates at a certain set of modal frequencies. Each model frequency has a different vibration spatial pattern. Video magnification can amplify these motions that correspond to a range of temporal frequencies and reveal these spatial patterns. Figure 9 shows a single frame from each motion-magnified video with its theoretically expected shape [54].

Objects in local image patches move consistently; otherwise, it is considered noise. Oversmoothing operations can reform white noise into reasonable motion signals. As a result, verification of these amplified signals is vital, and many experiments have been conducted to compare the visual motion signal with signal recorded by accurate devices. The results of these experiments verify the reality of these motions [19, 29, 55-56].

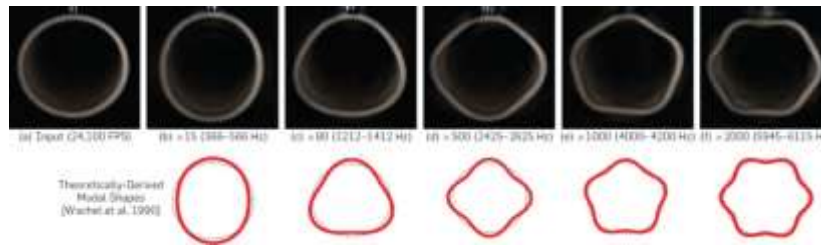


Figure 9. Isolating different types of spatial motions with temporal filtering [57]

5.4. Amplifying the Colour

The phase-based formulation removes low-amplitude, short motions, whereas larger amplitude motions remain to pass through. Motion attenuation is controlled by setting the amplification factor (α) to a negative value in the range of $(-1, 0)$. If (α) has a value of -1 , then all of the phase changes over time within the desired frequency band zeros-out. However, this result differs for every frame because coefficient amplitudes are changing over time. This condition is similar to other motion processing, such as the denoising presented in [58] and the video deanimation in [59]. However, these processing operations can be carried out efficiently when motions in the scene are small.

The linear magnification method jointly amplifies motions and colour changes for a particular input video; therefore, small face motions can become larger and visible when amplifying colour changes corresponding to the pulse. This feature may be undesirable in the application. This method can remove these undesirable motions. A similar colour amplification can result in that of the linear method entirely with steerable pyramids.

6. EXISTING WORKS USING PHASE-BASED EVM

In [12], the researcher used Eulerian motion magnification (EMM) to determine the feelings of others, which are transverse and sudden, for a short period. Movement magnification can enhance the detection rate of hidden emotions for facial expressions, such as tense, disgust, happiness, surprise or repression.

In [16], the researchers used EMM to detect vital signs for people who are alive. However, they either faint or are infected in dangerous environments and places that are difficult for rescuers to access, such as battlefields or terrorist acts, which represent finding survivors a challenge.

In [17], the EVM technique was used to visualise the movement of microorganisms, such as bacteria, that are invisible to the human eye.

In [18-19], a method was presented to measure displacement and vibration and identify typical characteristics for a particular structure to detect any damage. A camera was used as a sensor for monitoring, and the resulting video was analysed using EMM to magnify small movements; such vibrations and displacement in structures are evident for engineers. The results obtained from the camera were validated by comparing the results obtained from the camera with the results obtained by the devices used to measure vibration and displacement. The results extracted from the video were accurate.

In [35], a new method of phase-based EVM was proposed. The authors presented an image pyramid decomposition that uses Riesz transform (R-T). The new Riesz pyramids of a video frame decompose an approximate R-T using two three-tap finite difference filters, and invertible octave bandwidth filters are applied. These authors claimed that their method can be utilised for real-time subtle motion visualisation. This method is considerably faster than complex steerable and maintains video quality. However, the Riesz pyramid fails to maintain the power to the input signal, which can cause side artefacts.

In [38], a new method called EMMS was suggested to solve problems, such as increased processing time and noise and reduced video quality. The method is based on wavelet decomposition, Chebyshev bandpass filter and image denoising and can extract nondetectable motion information from input videos. This method has surpassed previous methods in terms of reducing processing time and eliminating noise. The method is also considerably faster than previous methods, has higher video quality and reduces noise. EEMS may also be passable for use in video-processing applications in real time.

7. COMPARISONS

In this section, we summarise the main differences amongst the magnification methods in Table 1. In Table 2, the differences in video quality are presented for different methods. Four source videos (baby1,

baby2, guitar and camera video) are used in this comparison. They are available on the MIT website through the webpage (<http://people.csail.mit.edu/mrub/vidmag/>). The quality is measured by computing peak signal-to-noise ratio [60], mean squared error and the other quality metrics that are shown in the first column of Table 2. Table 3 shows the differences amongst the magnification methods in the processing time to the video sources. Table 3 shows the differences amongst the magnification methods in the execution time by using the same four source videos that are used in Table 2. All the execution time measurements are achieved on the basis of the same computer and MATLAB software version. In each test, 100 frames are taken from each video source.

Table 1. Comparison of The EVM Methods

Method	Decomposition	Magnification	Overcompleteness	Noise	Exact for
LVM [8]	Laplacian pyramid	Supports small magnification factor	$\frac{3}{4}$ image	Magnified	Linear ramps
Enhanced EVM [37]	Laplacian pyramid	Supports medium magnification factor	$\approx 15\% - 20\%$ slower than LVM	Minimised via postprocessing	Linear ramps
Phase-based method [29]	Complex steerable pyramids	Supports large magnification factor ² with suboctave bandwidth filters	$2k/(1-2^{-\frac{2}{n}})^{-1}$ (3-4 times slower than LVM with octave bandwidth, 2 orientations)	Translated	Sinusoid
Fast phase-based method [35]	Riesz pyramid	Supports large magnification factor	$\approx 20\% - 80\%$ faster than phase-based method (octave bandwidth, 2 orientations)	Translated	Sinusoid
EMMS [38]	Wavelet pyramid	Supports large magnification factor	$\approx 60\% - 70\%$ faster than LVM	Removed	Linear ramps

Table 2. Comparison Amongst Eulerian Methods on Average Video Quality Metrics for Baby 1 Video, Baby 2 Video, Guitar Video and Camera Video

Video quality	The Method				
	LVM [8]	Enhanced EVM [37]	Phase-based method [29]	Fast phase-based method [35]	EMMS [38]
Peak signal-to-noise ratio (PSNR)	29.3105	33.85575	26.746025	26.788	36.16575
Mean squared error (MSE)	95.81475	29.617	175.03425	156.79375	19.63335
Structural similarity (SSIM)	0.894125	0.96295	0.87055	0.88645	0.96895
Multiscale structural similarity (MS-SSIM)	0.916525	0.983875	0.9498	0.963825	0.992525
Gradient similarity measure (GSM)	0.98855	0.9967	0.991675	0.9928	0.99815
Feature similarity (FSIM)	0.9167	0.986925	0.93245	0.940775	0.9911
Spectral residual similarity (SRSIM)	0.95085	0.993225	0.967125	0.97145	0.994725
Noise quality measure (NQM)	17.92975	24.42125	20.68125	19.28175	29.80925
Weighted signal-to-noise ratio (WSNR)	26.87175	33.9785	29.8235	27.24025	40.238
Visual signal-to-noise ratio (VSNR)	19.3675	28.829	19.6445	21.27025	34.66275
Visual information fidelity (VIF)	0.517775	0.673725	0.4149	0.482525	0.72445

Table 3. Comparison of the Magnification Methods in Terms of Processing Time Measured by Seconds

Input Videos	Methods				
	LVM [8]	E2VM [37]	Phase-based method [29]	Fast phase-based method [35]	EMMS [38]
baby1	85.67	101.32	213.17	117.45	50.68
baby2	82.52	97.84	211.74	115.73	48.54
camera	100.76	120.05	243.18	130.24	59.29
guitar	19.77	24.15	33.84	26.37	12.31

8. CONCLUSIONS

The EVM approach can be considered an efficient tool in the field of subtle variations in motion and colour of source videos. This tool represents a computer-based microscope that can be used in several applications in different fields, such as mechanical, material and civil engineering and biology and

healthcare. EVM has two main categories, linear and phase based. Each category has its own strengths and drawbacks. Linear-based EVM is simple and fast but fails to process large values of amplification factor because amplifying noise results in poor quality and unclear magnification. By contrast, phase based-EVM can amplify subtle motion by large amplification factor; however, it is an intensive algorithm. In conclusion, several challenges, such as real-time suitability, magnified noise reduction and significant and subtle motions in source videos, must be solved in EVM through scientific research.

REFERENCES

- [1] G. R. Poornima and S. C. P. Kumar, "Efficient H . 264 Decoder Architecture Using External Memory and Pipelining," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, no. 3, pp. 995-1002, 2018.
- [2] M. Z. Poh, D. J. McDuff, and R. W. Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation," *Opt. Soc. Am.*, vol. 18, no. 10, pp. 10762-10774, 2010.
- [3] W. Verkruyse, L. O. Svaasand, and J. S. Nelson, "Remote plethysmographic imaging using ambient light," *Opt. Express*, vol. 16, no. 26, pp. 21434-21445, 2008.
- [4] C. Liu, A. Torralba, W. T. Freeman, F. Durand, and E. H. Adelson, "Motion magnification," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 519-526, 2005.
- [5] J. Wang, S. M. Drucker, M. Agrawala, and M. F. Cohen, "The cartoon animation filter," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 1169-1173, 2006.
- [6] B. D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision.," in *International Joint Conference on Artificial Intelligence*, 1981, vol. 81, pp. 121-130.
- [7] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, no. 1-3, pp. 185-203, 1981.
- [8] H. Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 1-8, 2012.
- [9] M. Rubinstein, "Analysis and Visualization of Temporal Variations in Video," PhD thesis, Massachusetts Institute of Technology, 2014.
- [10] H.-Y. Wu, "Eulerian video processing and medical applications," Master's thesis, *Massachusetts Institute of Technology*, 2012.
- [11] T. Xue, M. Rubinstein, N. Wadhwa, A. Levin, F. Durand, and W. T. Freeman, "Refraction wiggles for measuring fluid depth and velocity from video," in *Computer Vision-ECCV 2014*, vol. 8691. Springer, pp. 767-782, 2014.
- [12] A. C. Le Ngo, Y. H. Oh, R. C. W. Phan, and J. See, "Eulerian emotion magnification for subtle expression recognition," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing-Proceedings*, 2016, pp. 1243-1247.
- [13] P. B. Chambino, "Android-based implementation of Eulerian Video Magnification for vital signs monitoring," Master's thesis, Do PORTO, 2013.
- [14] D. A. Ryan, "Visible imaging of global MHD on MAST," *IEEE Trans. Plasma Sci.*, vol. 42, no. 10, pp. 2556-2557, 2014.
- [15] A. Davis, M. Rubinstein, N. Wadhwa, G. J. Mysore, F. Durand, and W. T. Freeman, "The visual microphone: passive recovery of sound from video," *ACM Trans. Graph.*, vol. 33, no. 4, pp. 1-10, 2014.
- [16] C. Ordóñez, C. Cabo, A. Menéndez, and A. Bello, "Detection of human vital signs in hazardous environments by means of video magnification," *PLoS One*, vol. 13, no. 4, pp. 1-15, 2018.
- [17] M. Kayani, M. M. Riaz, A. Ghafoor, and N. Iltaf, "An efficient Eulerian video magnification technique for microbiology applications," *Radioengineering*, vol. 26, no. 1, pp. 316-322, 2017.
- [18] A. Sarra, Z. Mao, C. Niezrecki, and P. Poozesh, "Vibration-based damage detection in wind turbine blades using Phase-based Motion Estimation and motion magnification," *J. Sound Vib.*, vol. 421, pp. 300-318, 2018.
- [19] J. G. Chen, N. Wadhwa, Y. J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk, "Modal identification of simple structures with high-speed video using motion magnification," *J. Sound Vib.*, vol. 345, pp. 58-71, 2015.
- [20] X. He, R. A. Goubran, and X. P. Liu, "Wrist pulse measurement and analysis using Eulerian video magnification," in *3rd IEEE EMBS International Conference on Biomedical and Health Informatics*, Las Vegas, NV, USA, 2016, pp. 41-44.
- [21] L. Cattani *et al.*, "Monitoring infants by automatic video processing: A unified approach to motion analysis," *Comput. Biol. Med. Elsevier*, vol. 80, pp. 158-165, 2017.
- [22] X. He, R. A. Goubran, I. Fellow, X. P. Liu, and I. Senior, "Using Eulerian Video Magnification Framework to Measure Pulse Transit Time," in *2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*, 2014, pp. 2-5.
- [23] Y. Zhang and F. Shang, "Noncontact Extraction of Breathing Waveform," in *International Power, Electronics and Materials Engineering Conference (IPEMEC 2015)*, 2015, pp. 788-793.
- [24] U. Rubins, J. Spigulis, and A. Miscuks, "Application of colour magnification technique for revealing skin microcirculation changes under regional anaesthetic input," in *Biophotonics-Riga 2013*, 2013, vol. 9032, pp. 1-5.
- [25] J. H. G. M. Klaessens, M. van den Born, A. van der Veen, J. Sikkens-van de Kraats, F. A. M. van den Dungen, and R. M. Verdaasdonk, "Development of a baby friendly non-contact method for measuring vital signs: first results of clinical measurements in an open incubator at a neonatal intensive care unit," in *Proceedings of SPIE - International Society for Optical Engineering*, 2014, vol. 8935, pp. 1-7.
- [26] M. Kamphuis, F. de Jongh, J. Hilderink, R. Vaartjes, J. Goorhuis, and B. Thio, "Analysis of video-amplified body parameters in pediatric sleep," *Eur. Respir. J.*, vol. 44, no. Suppl 58, p. P3274, 2014.

- [27] G. Balakrishnan, F. Durand, and J. Guttag, "Detecting pulse from head motions in video," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2013, pp. 3430-3437.
- [28] M. Janatka *et al.*, "Examining in vivo tympanic membrane mobility using smart phone video-otoscopy and phase-based Eulerian video magnification," in *Medical Imaging 2017: Computer-Aided Diagnosis*, 2017, vol. 10134, pp. 1-7.
- [29] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, "Phase-based video motion processing," *ACM Trans. Graph.*, vol. 32, no. 4, p. 80, 2013.
- [30] J. Portilla and E. P. Simoncelli, "A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients," *Int. J. Comput. Vis.*, vol. 40, no. 1, pp. 49-71, 2000.
- [31] E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger, "Shiftable Multiscale Transforms," *IEEE Trans. Inf. THEORY*, vol. 38, no. 2, pp. 587-607, 1992.
- [32] D. J. Fleet and A. D. Jepson, "Computation of component image velocity from local phase information," *Int. J. Comput. Vis.*, vol. 5, no. 1, pp. 77-104, 1990.
- [33] T. Gautama and M. M. Van Hulle, "A phase-based approach to the estimation of the optical flow field using spatial filtering," *IEEE Trans. Neural Networks*, vol. 13, no. 5, pp. 1127-1136, 2002.
- [34] B. Alibouch, A. Radgui, M. Rziza, and D. Aboutajdine, "Optical flow estimation on omnidirectional images: An adapted phase based method," in International Conference on Image and Signal Processing(ICISP 2012), 2012, vol. 7340 LNCS, pp. 468-475.
- [35] N. Wadhwa, M. Rubinstein, F. Durand, and W. T. Freeman, "Riesz pyramids for fast phase-based video magnification," in Computational Photography (ICCP), 2014 IEEE International Conference on, 2014, pp. 1-10.
- [36] N. Wadhwa, "Revealing and analyzing imperceptible deviations in images and videos," PhD thesis, Massachusetts Institute of Technology, 2016.
- [37] L. Liu, L. Lu, J. Luo, J. Zhang, and X. Chen, "Enhanced Eulerian video magnification," in Image and Signal Processing (CISP), 2014 7th International Congress on, 2014, pp. 50-54.
- [38] A. Al-Naji, S. H. Lee, and J. Chahl, "An efficient motion magnification system for real-time applications," *Mach. Vis. Appl. Springer*, vol. 29, no. 4, pp. 585-600, 2018.
- [39] P. J. BURT and E. H. ADELSON, "The Laplacian Pyramid as a Compact Image Code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532-540, 1983.
- [40] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. Burt, and J. Ogden, "Pyramid methods in image processing," *RCA Eng.*, vol. 29, no. 6, pp. 33-41, 1984.
- [41] W. Freeman, E. H. Adelson, and D. Heeger, "Motion without movement," in *Computer Graphics (ACM)*, 1991, vol. 25, no. 4, pp. 27-30.
- [42] J.-Y. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corp.*, vol. 5, pp. 1-9, 2001.
- [43] P. G. Pansare and M. P. Dale, "Magnification Of Wrist Video For Heart Rate Measurement," vol. 5, no. 1, pp. 111-114, 2017.
- [44] A. J. Rojas, A. Fuentes, H. O. Garcés, L. E. Arias, J. Cuevas, and J. Pino, "Soot propensity by image magnification and artificial intelligence," *Fuel Elsevier*, vol. 225, pp. 256-265, 2018.
- [45] S. Bharadwaj, S. Member, T. I. Dhamecha, and S. Member, "Face Anti-spoofing via Motion Magnification and Multifeature Videolet Aggregation," in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014, pp. 1-12.
- [46] K. Sudhakar and P. Nithyanandam, "An Accurate Facial Component Detection Using Gabor Filter," *Bull. Electr. Eng. Informatics*, vol. 6, no. 3, pp. 287-294, 2017.
- [47] W. Jiang, T.-Z. Shen, J. Zhang, Y. Hu, and X.-Y. Wang, "Gabor wavelets for image processing," in *Computing, Communication, Control, and Management, 2008. CCCM'08. ISECS International Colloquium on*, 2008, vol. 1, pp. 110-114.
- [48] D. Barina, "Gabor Wavelets In Image Processing," in *arXiv preprint arXiv:1602.03308*, 2016.
- [49] W. Freeman, T. and E. Adelson, H., "The Design and Use of Steerable Filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 9, pp. 891-906, 1991.
- [50] E. P. Simoncelli and W. T. Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in Image Processing, 1995. Proceedings., International Conference on, 1995, vol. 3, pp. 444-447.
- [51] W. T. Freeman and E. H. Adelson, "Steerable filters for early vision, image analysis, and wavelet decomposition," in [1990] Proceedings Third International Conference on Computer Vision, 1990, pp. 406-415.
- [52] G. Ramkumar and E. Logashanmugam, "Multimodal verge for scale and pose variant real time face tracking and recognition," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 13, no. 2, pp. 665-670, 2019.
- [53] K. Anuradha and N. R. Raajan, "Strategy for Foreground Movement Identification Adaptive to Background Variations," *Int. J. Electr. Comput. Eng.*, vol. 8, no. 6, pp. 4258-4264, 2018.
- [54] J. C. Wachel, S. J. Morton, and K. E. Atkins, "Piping Vibration Analysis.," in Proceedings of the 19th turbomachinery symposium, 1990, pp. 119-134.
- [55] J. G. Chen, N. Wadhwa, Y.-J. Cha, F. Durand, W. T. Freeman, and O. Buyukozturk, "Structural modal identification through high speed camera video: Motion magnification," in *Topics in Modal Analysis I*, Volume 7, 2014, pp. 191-197.
- [56] J. G. Chen, N. Wadhwa, F. Durand, W. T. Freeman, and O. Buyukozturk, "Developments with motion magnification for structural modal identification through camera video," in *Dynamics of Civil Structures*, Volume 2, Springer, 2015, pp. 49-57.

- [57] N. Wadhwa *et al.*, "Eulerian Video Magnification and Analysis," *Commun. ACM*, vol. 60, pp. 87-95, 2017.
- [58] M. Rubinstein, C. Liu, P. Sand, F. Durand, and W. T. Freeman, "Motion denoising with application to time-lapse photography," in *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference on, 2011, pp. 313-320.
- [59] J. Bai, A. Agarwala, M. Agrawala, and R. Ramamoorthi, "Selectively de-animating video.," *ACM Trans. Graph.*, vol. 31, no. 4, pp. 61-66, 2012.
- [60] N. O. Alsrehin and A. F. Klaib, "VMQ : an algorithm for measuring the Video Motion Quality," *Bull. Electr. Eng. Informatics*, vol. 8, no. 1, pp. 231-238, 2019.

BIOGRAPHIES OF AUTHORS



Haider Ismael Shahadi received his B.ESc degree in information engineering from the university of Baghdad-Iraq in 2001, his master's degree in Electronic and Communication Engineering from the University of Baghdad-Iraq in 2004, and his Ph.D. in Electronic and Communication Engineering from the Tenaga National University-Malaysia in 2014. Currently, he is an assistant professor at the University of Kerbala-Iraq. His research interests include digital signal and multimedia processing, data security, FPGA design and implementation and embedded systems, IOT systems, and smart systems.



Zaid Jabbar Al-allaq received the Bachelor of Engineering in communication Techniques from the Engineering Technical College, Al-Furat Al-Awsat Technical University, Najaf, Iraq, in 2007, He is currently working toward the M.S. degree in the Department of communication engineering, University of Al-Furat Al-Awsat Technical, Engineering Technical College - Najaf.



Hayder J. Albattat received a B.A.Sc. degree in Electrical Engineering Basra University, Iraq, M.Sc.s. in Electrical Engineering (Electronics and communication –Image Processing) Electrical engineering Department in Basra University, Iraq, and a Ph.D. degree in Electrical Engineering (Electronics and communication) – University of Basra, Iraq. He is currently a Asst. Professor in Electronic and Communication at Najaf Technical College / the Dpt. of Communication - Al-Furat Al-Awsat Technical University.