SciencePG
Science Publishing Group

Fast Dual Discrete and Complexwavelet Based Video Magnification for 3D Facial Video Identification

Gamal Fahmy[1, *], Omar Fahmy[2], Mamdouh Fahmy[1]

[1]Electrical Engineering Department, Assiut University, Assiut, Egypt

[2]Electrical Engineering Department, Bader University, Cairo, Egypt

Email address:

fahmygamal@hotmail.com (G. Fahmy)

[*]Corresponding author

To cite this article:

Abstract: Magnifying micro motion of videos that are undetectable by humans has recently been popular in many applications. This is due to its impact in numerous applications. In this paper, we explore this technique in 3D facial video identification, where we try to distinguish between real 3D facial objects in videos and 2D images of faces in a video frame sequence, and utilize this in biometric identification. We present a fast 2D Dual Discrete Wavelet Transform 2D-DWT based video magnification technique that detects micro movements by magnifying the phase differences between subsequent video frame's wavelet coefficients, at different sub bands. Next, in order to overcome shortcoming of 2D-DWT systems, 2D Dual Complex Wavelet Transform 2D-CWT has also been employed to estimate phase changes between subsequent video frames at different spatial locations of Complex Wavelets sub-bands. This latter presented CWT Technique uses the Radon Transform to detect any periodic motion in the video frames. Several simulation results are given to show that our proposed hybrid technique achieves comparable and sometimes superior performance with far less complexity when compared with recent literature in micro motion magnification, such as steerable pyramids STR and Riesz Transform RT based steerable pyramids RT-STR. Both DWT and CWT techniques are combined for 3D facial video identification. The attached videos demonstrate the superior video quality obtained by the proposed technique.

Keywords: Complex Wavelets, Micro Video Magnification, Local Phase Amplification

1. Introduction

The human visual system has far less ability to recognize spatial and temporal variations, when compared with automated computations, of frame pixel values in different videos. This is due to the fact that the human visual system has small sensitivity with the spatio-temporal changes of video signals. These micro changes of frame pixels values, whether they are intensity changes (i.e. change in brightness, change in colorings) or spatial changes (i.e. micro movement of video objects), may be invisible for the human eye. However, they reveal useful information. Hence, there is an impending need for novel techniques to exploit these changes to extract some valuable information in different engineering, nanotechnology and medical applications. Magnifying micro movements in natural videos has been investigated by several

researchers in the past decade. This is due to the fact that it reveals useful information to recognize little spatial and temporal variations of frame pixel values that are useful in numerous applications [1-5].

In video micro motion magnification we mainly analyze over time frames of videos at different pixel locations for many frequency values and magnify any changes of pixel values (which would indicate a micro motion) at any temporal position (frequency).

This temporal manipulation operation can reduce the range of frequencies that can be magnified to reduce the impact of any type of noise may exist during the camera acquisition process. According to H. Wu et al. [6], an elegant technique was presented to magnify micro brightness and color changes using an Eulerian process. It achieved outstanding performance, but it was sensitive to noise that gets amplified

during the magnification operation. According to N. Wadha et al. [7], another superior technique was presented in micro motion magnification, that replied on the utilization of complex Gabor filters that constructed Laplacian pyramids. these pyramids are well known as steerable pyramids STR. Magnification of micro movements was achieved by changing the phase of complex coefficients of the Gabor filters. Performance was outstanding but it was computationally expensive.

According to N. Wadha et al. [8], Riesz Transform RT based steerable pyramids, RT-STR, the RT was presented in phase magnification of the Laplacian pyramids presented in [7], for the same purpose of magnifying micro motion in videos. This was performed by projecting each video frame orthogonally in spatial domain, hence much computations were saved by avoiding complex Laplacian pyramids. More details about a quaternion representation that was utilized to determine the RT-STR phase values [9].

CWT was first introduced as a domain for phase magnification in video micro movement magnification [10-11]. This is due to the fact that CWT utilizes orthogonal real filters to construct analytic signals. CWT filters are also shift invariant and oriented in spatial domain. In the CWT domain, every video frame is decomposed into sub-bands. Then, the phase difference between these sub-bands coefficients of each frame and a reference frame, is multiplied by an amplification factor to get new modified real and imaginary CWT coefficients. These modified CWT wavelet coefficients are used to reconstruct the new updated video frame.

Recognizing 3D facial objects in biometric identification videos has been received much interest in recent security identification applications. It aims at distinguishing 3D facial objects in videos from 2D Facial frame sequence videos (2D face image video sequences), Figure 3(d). It is highly influential in security identification application, where the goal is to determine that the human face subject that is posing in front of a camera is a live object, rather than a printed face image that is presented in front of an authentication camera for the purpose of unauthorized access [12]. In this paper we utilized both DWT and CWT to detect 3D facial videos and identify them from 2D image sequences as well be shown in section 5.

In this paper we first present a fast DWT based video magnification technique. In this technique, the Approximate Reisz Transform of [8] is utilized to estimate the local phase difference between sub bands of wavelets decompositions of video frames for the purpose of fast micro movement magnification. This technique has been employed for 3D facial video recognition. Next, in order to overcome shortcomings of DWT systems, a CWT based technique is proposed. In this respect, a novel approach is proposed to accurately estimate the phase differences between wavelet sub band coefficients of subsequent frames of the video.

This estimation is achieved by manipulating the real and imaginary values of rows and columns of wavelets matrices of this CWT analysis sub-bands. This manipulation aims at producing exact orthogonal wavelet matrices. Then, to

calculate the orientation of every video frame, we proposed a gradient-based method. We finally for the determined orientation value, we select an appropriate value for phase amplification through the utilization of the Radon Transform. This Radon Transform is also utilized to detect periodic motions in any video frame. Simulation results of different videos, have revealed the superior quality of the micro motion magnified videos when compared with the existing approaches. Finally this CWT Radon Transform based technique has been employed for the recognition of 3D facial videos when compared with basic 2D facial image videos.

The paper is organized as follows: section II, briefly summarizes different micro movement techniques presented in recent literature such as STR and RT. Section III presents DWT based video micro motion magnification. We employed it for 3D facial identification for fast recognition. Section IV presents our novel techniques of accurate phase estimation, orientation calculation and phase gain estimation. We also utilized this approach to detect blood circulation in facial tissues of facial objects in 3D facial identification. A detailed simulation comparison between the proposed technique on CWT and DWT as well as the successful recent literature approaches STR and RT-STR is illustrated in section V. Section V also includes our simulation results with 3D face identification experiments and its recognition from 2D image sequences. Conclusion is in section VI.

2. Background

Steerable pyramids [13] are well known of being easy to rotate in in polar domain for their frequency bands at any arbitrary orientation. This feature made them achieve superior performance in video micro movement magnification. But as a drawback, they require high computational complexity in addition to the fact that they are shift variant. In order to speed up computations with Steerable pyramids; a new image pyramid representation, the Riesz Transform RT pyramid was proposed [8, 9]. The Riesz Transform RT avoids computing the 2-D Hilbert Transform video frames in frequency domain as in steerable pyramids. It amounts to constructing approximate orthogonal projections of the frame sub band I, along the horizontal and vertical directions, namely Rx, Ry. This triple $\{I, R_x, R_y\}$ determines a spherical form as $\{A, \phi, \theta\}$. $A = \sqrt{I^2 + R_x^2 + R_y^2}$ $I = \cos\phi$, ϕ defines the local phase of this subband. This constitutes the basic step in constructing the Riesz pyramid-based video micro movement magnification process. It is based upon decomposing each video frame into multiple sub-bands, each of which corresponds to a different spatial scale. Then, we calculate the Riesz transform of each sub-band. The local phase shifts ϕ_c of these subbands are updated as $\phi_n = \phi_c + K * (\phi_c - \phi_r)$, ϕ_r is the local phase of a reference frame at the same spatial location and scale. K is the phase magnification factor. In order to speed up phase update, the quaternion technique is used [9]. Figure 1 shows a block diagram of our DWT based proposed technique. An abbreviated version of this work was presented in [14, 15], but with much less details and experiments setup and results
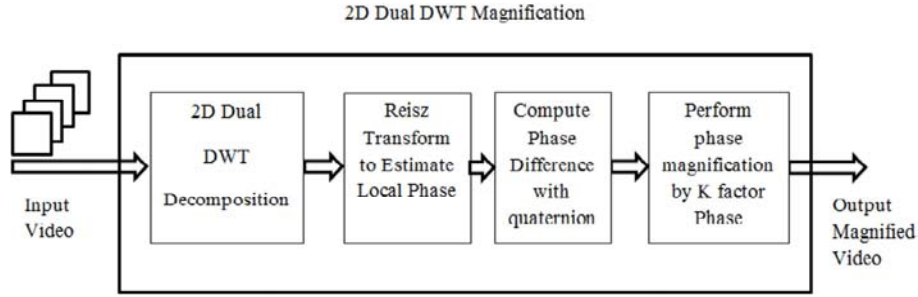
2D Dual DWT Magnification

Figure 1. Block diagram of the fast DWT based algorithm.

3. DWT Based Video Magnification

The proposed DWT and CWT-based techniques for micro movement magnification [10, 11], is based on the following observation. Given a reference video frame $f_r(x, y)$ at time t and another arbitrary frame $f(x, y)$ that is delayed from the reference frame by Δt, i.e. $f(x,y) \equiv f_r(x + \alpha_1 \Delta t, y + \alpha_2 \Delta t)$

$$f_r(x + \alpha_1 \Delta t, y + \alpha_2 \Delta t) = \sum \sum \left\{ F_r(\Omega_1, \Omega_2) e^{j(\alpha_1 \Omega_1 + \alpha_2 \Omega_2)\Delta t} \right\} e^{j(\Omega_1 x + \Omega_2 y)} \tag{1}$$

$F(\Omega_1, \Omega_2)$; $F_r(\Omega_1, \Omega_2)$ are the 2D-DFT of the oriented band pass functions of the decomposed sub-bands of the respective video frames. Thus, under constant illumination conditions, $F(\Omega_1, \Omega_2)$; $F_r(\Omega_1, \Omega_2)$ differ only by a constant phase. This means that changing the phase difference between respective sub bands by a constant amount, is equivalent to shifting these frames in time.

Now, in order to perform video micro movement magnification using DWT systems, we calculate the local Riesz basis for each DWT subband, then the local phase of every sub bands can now be determined. We note here that DWT provides real coefficients, but after this Riesz post processing the coefficients becomes complex and phase can be calculated. The DWT video micro movement technique can be designed as follows:

at time $t + \Delta t$, and using the 2-D DFT, we have

$$f(x, y) = \sum \sum F(\Omega_1, \Omega_2) e^{j(\Omega_1 x + \Omega_2 y)},$$

$$f_r(x, y) = \sum \sum F_r(\Omega_1, \Omega_2) e^{j(\Omega_1 x + \Omega_2 y)}$$

1. Decompose every frame of the video stream into n subbands, using the assigned DWT family.
2. For every sub-band, compute wavelet coefficient $\omega_{k,d}$, $k = 1,2, \ldots, n$; $d = 1,2,3$. Construct for each sub band, the Riesz basis $R_{x,kd}$ and $R_{y,kd}$. This triple $\omega_{k,d}, R_{x,kd}, R_{y,kd}$ defines the local amplitude A, local phase ϕ and local orientation θ of the spectral representation (A; ϕ; θ), where $\omega_{k,d} = A \cos\phi$, $R_{x,kd} = A \sin\theta$ and $R_{y,kd} = A \sin\phi \sin\theta$.
3. To compute the phase difference between the current frame wavelet ($\omega_{kd}^c + iR_{x,kd}^c + jR_{y,kd}^c$) and its preceding one ($\omega_{kd}^p + iR_{x,kd}^p + jR_{y,kd}^p$). We invoke the quaternion technique. This means that the quaternion phase difference is simply

$$\left(\omega_{kd}^c + iR_{x,kd}^c + jR_{y,kd}^c\right) - \left(\omega_{kd}^p + iR_{x,kd}^p + jR_{y,kd}^p\right) = r_0 + ir_1 + jr_2 \tag{2}$$

Thus, the quaternion local phase difference $\Delta\phi = cos^{-1}\left(\frac{r_0}{R_t}\right)$, $R_t = \sqrt{(r_0^2 + r_1^2 + r_2^2)}$. Similarly, its orientation $\theta = cos^{-1}(\frac{r_1}{\sqrt{r_1^2 + r_2^2}})$.

4. To perform phase magnification by a factor K; i.e $\Delta\phi_u = K\Delta\phi$, the new wavelets ω_{kd}^n are obtained as follows:

$$\omega_{kd}^n = A\cos(\phi + \Delta\phi_u) = A\cos(\Delta\phi_u)\cos\phi - A \sin(\Delta\phi_u)\sin\phi \tag{3}$$

$$\omega_{kd}^n = A\cos(\Delta\phi_u)\omega_{kd}^c - \sin(\Delta\phi_u)\left(R_{x,kd}\cos(\theta) + R_{y,kd}\sin(\theta)\right) \tag{4}$$

and all these quantities are already obtained.

DWT has been employed for 3D human face recognition (identification), as a fast online tool as will be shown in the simulations section. It is meant to be used as a pre-step before applying our next more hybrid technique of CWT based micro motion magnification with its gradient based phase direction estimation and Radon Transform based phase gain estimation components. We note here that this DWT Riesz based technique provides orthogonal projection for the

DWT subband coefficients, hence an accurate fast local phase can be determined for these coefficients, otherwise no phase information would have been possible to obtain for the original DWT real subbands. At this point, it is worth mentioning that DWT suffers from ringing artifact at edges of the produced images. It can be viewed as a singularity moving around in space [16]. To overcome this problem, it is proposed to use CWT based magnification, Figure 2.
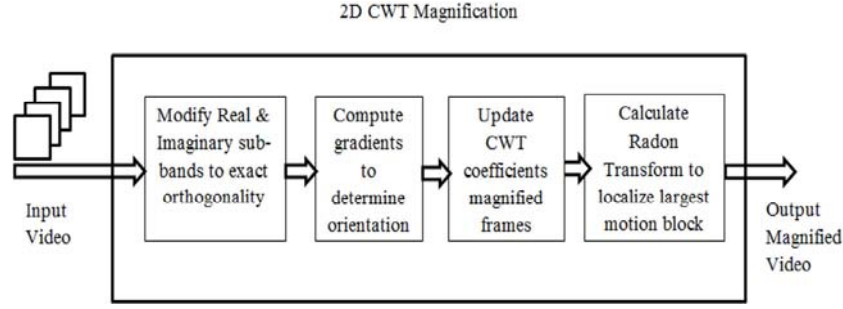
Figure 2. Block diagram of the CWT based algorithm for micro movement magnification.

4. The Proposed CWT-Based Video Magnification Technique

The one dimensional dual tree complex wavelet (1D DTCWT) transform is realized as two DWT real trees in parallel. These two transform trees must be orthogonal or biorthogonal to each other. We can construct this 1D DT-CWT with a pair of lowpass Hilbert pair filters (h_{0u}, h_{0d}) for the upper and lower DWT trees that ensures Hilbert relations. Hence, the wavelet coefficients at any decomposing level of the upper and lower trees would constitute Hilbert pair. In order to satisfy Hilbert relations, we select the first stage filter $h_0^1(n)$, and the succeeding stages filters $h_0^2(n)$ of the upper and lower trees to follow these relations

$$h_{0,d}^1(n) = h_{0,u}^1(n-1) \qquad (5)$$

$$h_{0,d}^j(n) = h_{0,u}^j(n-0.5), j = 2,3,\ldots,r$$

where r is the number of decomposition levels. The synthesis filter set pairs $g_{0u}^1, g_{0d}^1, g_{0u}^2, g_{0d}^2$ determined as a consequence of the wavelet orthogonality conditions, obey similar relations. This means that we have 4 sets of filters for 2 stages $h_{0u}^1, h_{0d}^1, h_{0u}^2, h_{0d}^2$, $g_{0u}^1, g_{0d}^1, g_{0u}^2, g_{0d}^2$. This would

construct critically sampled wavelet. This dual-tree complex wavelet, is nearly shift-invariant and is oriented in spatial domain, as it produces six sub bands at each scale, each is oriented at distinct direction.

In order to construct the two dimensional dual-tree 2D DT-CWT we follow the same structure in the regular 2D DWT, but we utilize two wavelets in each sub band. This real 2-D dual-tree wavelet is implemented using two critically sampled 2-D DWTs in parallel [17–19]. We construct these wavelets by processing 2-D dual DWT sub-bands through its columns and rows with filters of Eq. (5). We take four different combination as

$$[(h_0^1, h_0^1), (h_0^1, h_0^2), (h_0^2, h_0^1), (h_0^2, h_0^2)]$$

where $h_0 = h_{0u}$ represents the upper tree and h_{0d} represents the lower tree. We now select the upper and lower trees and then calculate the sum and difference images of the subbands of these two trees. These sum and difference images represents the wavelets for each sub band of the two trees. The sum image wavelet can represent the real part of a complex structure, while the other image (difference image) can represent the imaginary part. Hence, CWT would generate six oriented sub bands, all nearly shift invariant, at each scale for each tree.
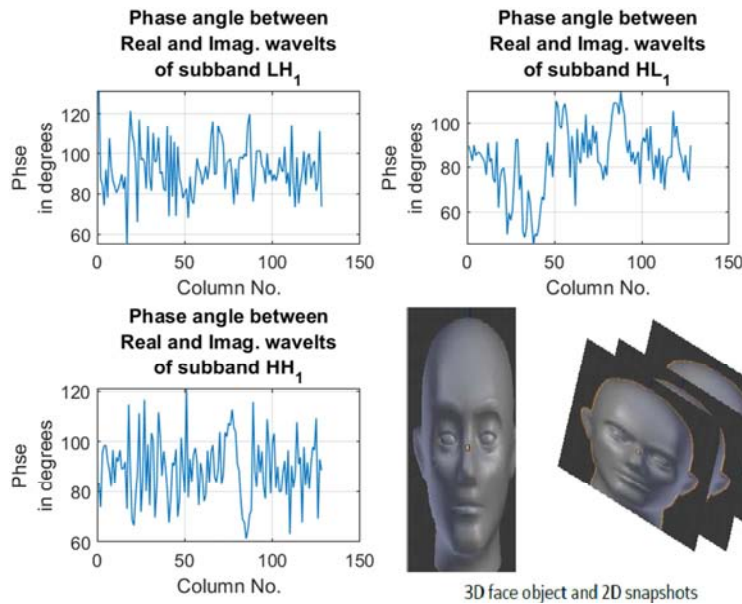

Figure 3. (a), (b), and (c). The phase between real and imaginary wavelets of the CWT decomposition LH, HL, HH sub bands of the 1st sub band of the upper tree for 256_256 Cameraman image. (d) 3D object face in comparison with 2D snapshot images.

Thus, unlike the DWT based technique described in sec. (3), the CWT based micro motion magnification approach utilizes a real and imaginary tree to estimate the phase difference $\phi_{ex} = \phi_c - \phi_r$, where ϕ_c, ϕ_r are the phases of the current frame and reference frame, at every spatial location and scale of the CWT sub band. Then, we modify the current frame phase ϕ_c to be $\phi_c + K$, where K is the phase gain. The motion magnified CWT real and imaginary wavelets are reconstructed using the modified phases ϕ_c as described in [11]. The main disadvantage of this technique is the assumption of orthogonality of the columns of the real and imaginary wavelets of the CWT decomposition at every spatial location and scale. In fact, the phase angle between the columns of these two wavelets of a CWT system, is not 90. Hence, the estimated frame phase ϕ is not always accurate and precise. Figure 3, verifies this feature for 256x256 Cameraman image, when decomposed using 10-tap orthogonal CWT filter, [10]. The figure illustrates the phase angles between the columns of the real and imaginary wavelets of the 1st sub band of the upper tree, for the LH; HL and HH sub bands. In this paper we propose to accurately estimate the local phase ϕ_c as will be presented in next section, this would enhance the micro movement magnification performance. We consider this; the main novelty of this work, Figure 2 shows a block diagram of this proposed part.

A. Exact Quadrature Real and Imaginary Wavelets CWTs
For any coefficient

For any coefficient with a specific spatial location and scale, we specify the real wavelet by X_{ij} and imaginary wavelet by Y_{ij} where

$$\begin{bmatrix} i = 1 \ Upper \ tree \\ = 2 \ Lower \ tree \end{bmatrix}$$

$$j = 1,2,3 \ for \ LH, HL \ and \ HH \ subbands\}$$

So, to accurately estimate the local phase ϕ_c, we do the following steps:

Compute $D = X_{ij}^t Y_{ij}$, $C = X_{ij}^t X_{ij}$. Construct a diagonal matrix Z with diagonal elements $z_{kk} = \frac{d_{\{kk\}}}{c_{\{kk\}}}$ where $d_{\{kk\}}$, $c_{\{kk\}}$ are diagonal elements of the matrices D and C, respectively.

Modify the imaginary wavelets

$$Y_{ij}^m = K_s \left(Y_{ij} - X_{ij} Z \right) \qquad (6)$$

K_s is a scaling matrix that makes the length of every vector of Y_{ij}^n the same as its corresponding of Y_{ij}.

Hence, the local phase is exactly estimated by

$$\phi_c = tan^{-1} \left(\frac{Y_{ij}^m}{X_{ij}} \right) \qquad (7)$$

where Y_{ij}^m is the updated/imaginary wavelet. We note here that even though the perfect reconstruction property PR of CWT is modified, but it will have minor (if any) impact on image and video quality. Table 1, shows this minor effect on image and video quality when implemented using CWT with different decomposition levels. We also note that this CWT subband modification for Exact quadrature wavelets would result in over-complete (almost) wavelet representation. This

representation is robust against noise and appropriate for all shift invariant video motion applications that can resist regular motion artifacts in video coding [16].

Table 1. *PSNR of video frames and images reconstructed without exact orthogonality.*

No. Levels	Cameraman image.	Baby video Frame 11	Crane video Frame 11
2	41.71	49.65	44.73
3	39.1	47.17	39.32
4	37.56	44.36	37.6
5	35.73	41.3	36.86

B. Phase Gain Estimation
In phase gain K estimation we measure the displacement for each video frame. The Radon Transform [20, 21], can best detect any minor change in images and video frames. The Radon Transform of a 2-D image $x(t_1, t_2)$ is defined as

$$R_x(\tau, \theta) = \int_{-\infty}^{\infty} x(t_1, t_2)\delta(t_1 \cos\theta + t_2 \sin(\theta - \tau))dt_1 \, dt_2 \quad (8)$$

Thus, $R_x(\tau, \theta)$ represents the projection of the image $x(t_1, t_2)$ with a direction angle θ with axis t_1. τ is the distance from the image center to the projection line. The Radon Transform can then best detect any slight modification in a projected image. Figure 4, demonstrated this detection ability from two successive frames (frames 1 & 2) of the Baby video, in both the vertical and horizontal directions. The Radon Transform is noticeable horizontally and vertically, in spite of the fact that the norm difference is less than 0.3. Radon Transform can also determine periodic performance if exists. That is if $x(t_1, t_2)$ exhibit a periodic changes along any specific direction with angle θ with t_1, then $R_x(\tau, \theta)$ will still have the same periodicity.

The Radon Transform is then used to measure our peak frame deviation. It is calculated with respect to the mean frame orientation. We then use the least squares estimation of the frames' gradients at every image pixel, to compute the orientation of the video frame image $I(x, y)$. It is listed as follows:

Compute the gradients $\partial_x = \frac{\partial I(x,y)}{\partial_x}$, $\partial_y = \frac{\partial I(x,y)}{\partial_y}$ at each pixel (x, y) of the image I. Then, the orientation at a specific pixel p is given by $\tan(\theta_p) = \frac{\partial_y}{\partial_x}$.

Denote the average frame orientation by θ_a. Then θ_a is determined as the least squares solution of the objective function ζ, where

$$\zeta = \Sigma\Sigma \left(\partial_y \cos(\theta_a) - \partial_x \sin(\theta_a) \right)^2 \qquad (9)$$

Where summation extends over the support of I. Minimization yields the average orientation θ to be

$$\theta_a = \frac{1}{2} \tan^{-1} \left(\frac{\nabla_y}{\nabla_x} \right) \qquad (10)$$

$$\nabla_x = \Sigma\Sigma(\partial_x^2 - \partial_y^2)$$

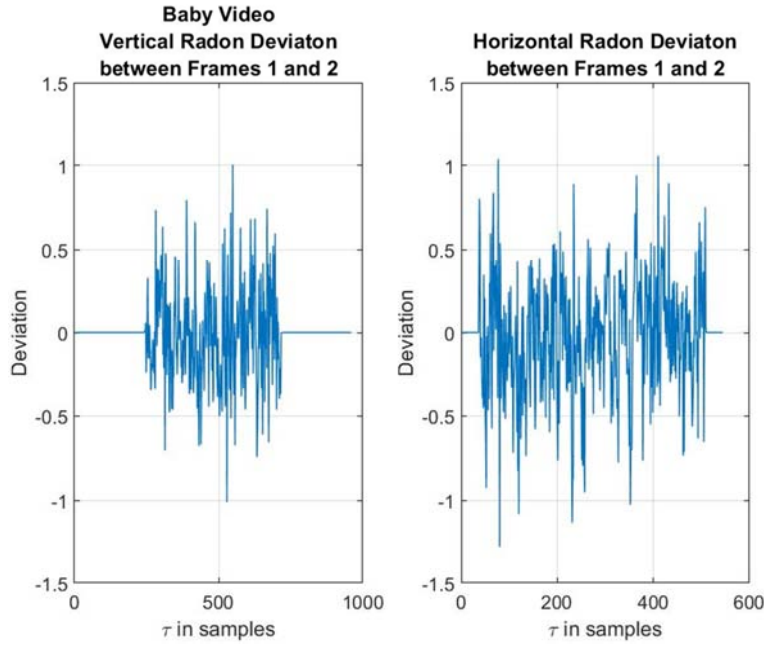$$\nabla_y = 2\Sigma\Sigma(\partial_x - \partial_y)$$

Figure 4. *The deviation of the Radon Transform in both vertical and horizontal directions, for frames 1 & 2 of the Baby video. The difference in the radon transform is noticeable while the norm difference between these two frames is less than 0.3.*

Figure 5, illustrates the orientation of the first frame of the Baby and Crane videos.

To calculate the phase gain K which would magnify motion displacement between any frame s and a reference frame r, we test the motion magnified frame using a certain phase gain value, K. We follow the next steps,

1. Compute the motion magnified frame $I_{s,magn}$ as described in sec. (4.1), for each phase gain value.

2. For each value of the phase gain vector, determine

$\Delta\theta = \theta_s - \theta_r$ and θ_r are the orientation of the motion magnified $I_{s,magn}$ and the reference frames, respectively. Finally, calculate the horizontal Radon Transform along θ_s, then select its 1st peak to determine the magnification factor.

Utilizing the Radon Transform in the estimation of phase gain for CWT (Complex Wavelet Transform) based micro motion magnification would eliminate the need for the user to identify a particular range of temporal or spatial frequencies to capture any motion in it, as has been adopted in recent literature [22, 7, 6]. The horizontal Radon Transform was specifically found to be more accurate in gain estimation rather than the vertical one, due to the fact that most micro motions (Capture device shacking) tends to be in the horizontal direction. Figure 6, shows the phase gain performance of both the Baby and Crane videos. The figure shows the Radon deviation performance agrees with the phase orientation behavior. The frame under investigation is the 45th frame. The 1st frame is taken as the reference frame. Three CWT decomposition levels are used. We choose the optimum phase gain to be the gain at the 1st peak of the axial direction (orientation's direction) Radon deviations. This makes the phase gain for the Baby video is 40 and 55 for the Crane video. Simulations have verified that the phase gain K is not sensitive to the choice of the test frame s.
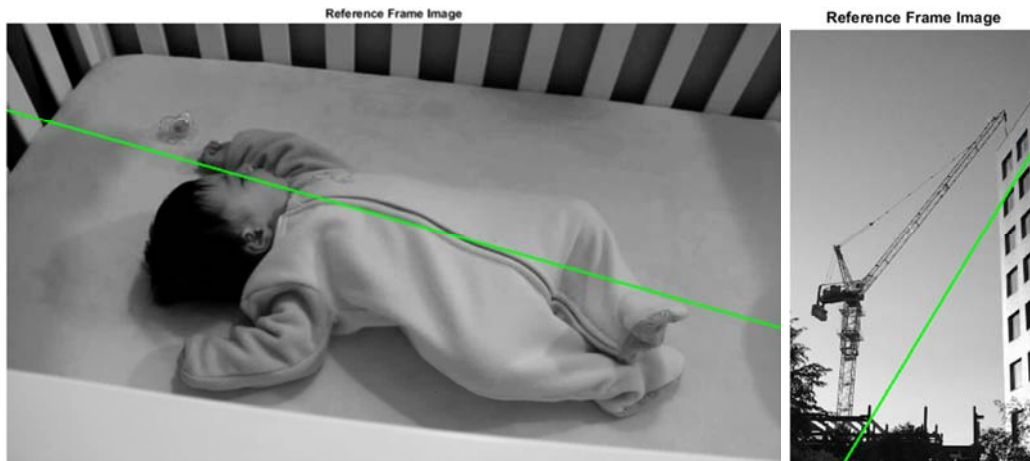

Figure 5. *Orientation of the Crane and Baby videos, first frame is displayed. We calculated frames gradients with least square minimization for every pixel.*
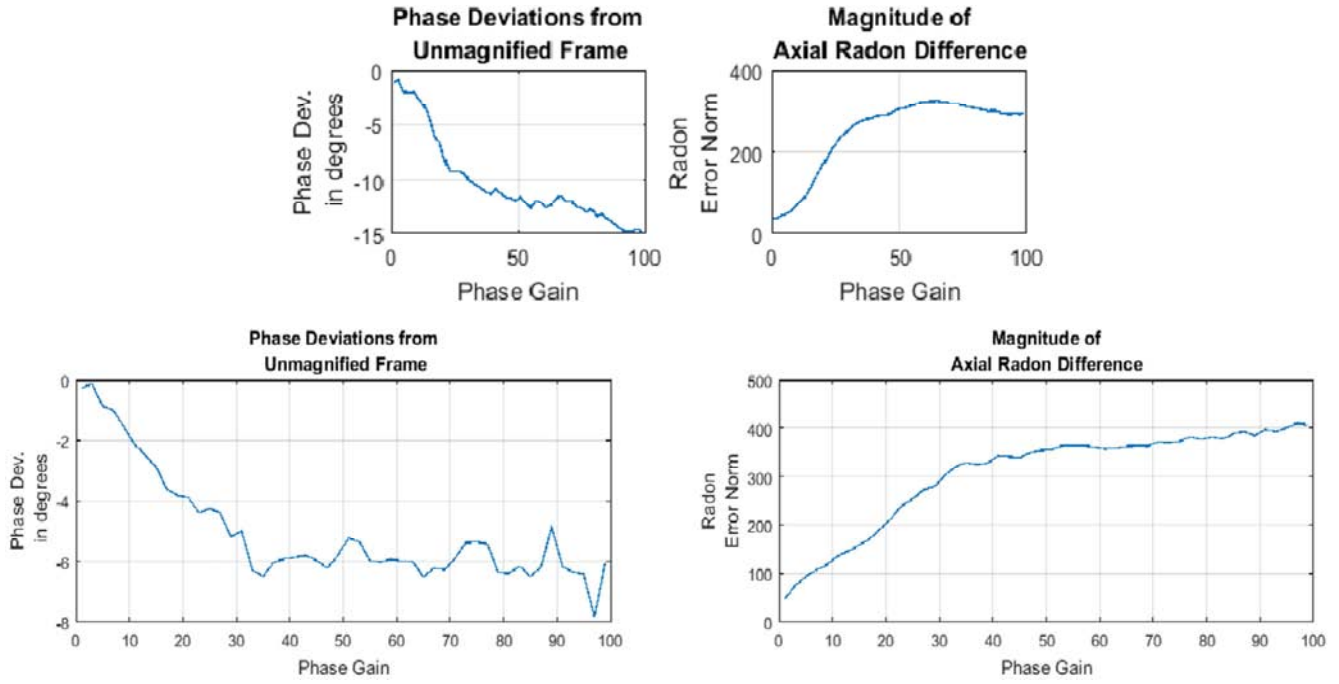
Figure 6. *The phase gain performance of Baby and Crane videos for the 45th frame. 3 CWT decomposition levels are used. The optimum phase gain is the gain at the 1st peak of the horizontal Radon deviations. It results in phase gain for the Baby video is 40 and 55 for the Crane video.*

5. Simulation Results

In this section, we implement both DWT and CWT techniques to magnify the micro movement between video frames. In the DWT technique, the wavelet family used is 'bior 2.4'. The number of decomposition levels is 3. Each sub band is processed by 3x3 Approximate Riesz Transform as described in sec. (3) to get the updated wavelet coefficient and subsequently the magnified video frame. The proposed modified CWT-based micro movement video magnification technique is implemented as follows:

1) Decompose every frame of the current video, into n levels CWT system.
2) Compute the real and imaginary wavelets $X_{i,j}$ and $Y_{i,j}$, $i = 1,2$ and $j = 1,2,3$ for the 3 sub bands of each scale of the lower and upper trees at every scale and spatial location. Update $Y_{i,j}$ to $Y_{i,j}^m$ as described by Eq. (6), calculate their modulus $R_{i,j} = \sqrt{X_{i,j}^2 + Y_{i,j}^{m2}}$
3) Evaluate the current sub band phase $\phi_c = tan\left(\frac{Y_{i,j}^m}{X_{i,j}}\right)$ as well as the phase angle ϕ_r of a reference sub band at the same spatial location and scale.
4) Modify ϕ_c to $\phi_c^m = \phi_c + K \phi_e$, only if $|\phi_e| \leq \delta \phi$. Otherwise let $\phi_c^m = \phi_c$. In this paper, we only calculate $\delta \phi$ if the *pdf* histogram of the CWT phase coefficients is below 10% of the maximum value.
5) Update the real and imaginary CWT, (and 2D DWT) coefficients of the magnified frame as

$$X_{ij}^m = R_{ij}cos(\phi_c^m) \qquad (11)$$

$$Y_{ij}^m = R_{ij} sin(\phi_{ij}^m)$$

6) Reconstruct the magnified frame using the updated wavelets in conjunction with the Inverse CWT algorithm, [17, 18].

We performed our simulation results only on the Baby and Crane videos to compare performance with standard literature results. Each video frame with the CWT technique is decomposed using a 3-level orthogonal CWT as in [17]. Frame *1* in both DWT and CWT techniques is the reference frame. We calculated a phase gain value of *K=55* for the Crane video, and *K=40* for the Baby video, sec. (4.2). The phase difference of each frame, is processed through a Butterworth filter. This band pass filter was 1st 1/150 and 25/150 cutoff frequencies. The reconstructed motion magnified video qualities are compared with qualities of videos constructed using the complex steerable pyramid technique [7], and those constructed using Riesz Transform technique [8], and the proposed hybrid technique phase gain *K*. The quality is measured by SSIM (I_{magn}, I_{org}) where I_{magn} is the motion magnified image, and I_{org} is the corresponding original one. Figure 7, shows this similarity. This verifies our claim that the proposed CWT technique achieves superior performance especially with the Crane video example. We achieved superior magnified motion for the Crane video that is free from any local vibration that was produced from current exiting literature techniques such as Riesz Transform and STR cases.
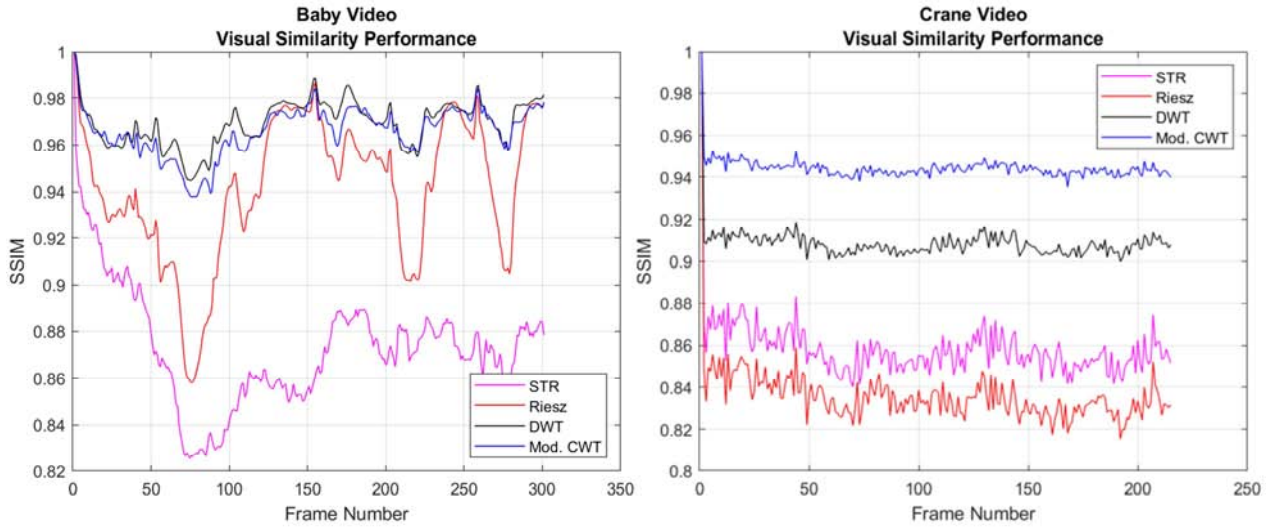
Figure 7. *The similarity between the magnified and unmagnified frames using SSIM and PSNR, measures. (a) Baby video case. (b) Crane video case. It is shown that the proposed CWT technique, competes very well the STR pyramid and Riesz micro movements techniques.*

3D Facial Video Identification

The proposed DWT based and CWT based techniques have been used for 3D human face identification. We aim at achieving clear distinction between real live 3D human facial videos and any 2D face image sequence videos, Figure 3(d). In our 3D human face object identification experiments, we captured several videos of 3D human face objects and its counter 2D face image video sequence. In our camera acquisition scenarios, we allowed the camera to vibrate horizontally within a limited rage (within 1 degree with 60 cm distance). This camera vibration is to strengthen the difference between the 3D and 2D video objects. We note here that horizontally or vertically vibrating the acquisition camera, while the object face is fixed, is similar to capturing a micro vibrating face while the camera is fixed, Figure 8. Also to simulate fake identification scenarios, we took facial videos of 2D images that are micro shacking with a similar frequency, just to distinguish between videos of real live 3D face objects that are naturally micro motioned, and videos of 2D facial snapshots that are intentionally micro moved to seem real. We applied our first DWT based technique on three different face videos, of the same subject, as a tool for fast recognition based on micro motion magnification. These 3 different micro motion face videos consisted of a real live 3D face video with natural micro motion, a video of 2D snapshot face images, fixed location, and a video of 2D snapshot face images with micro image shacking as described before. We then applied our hybrid CWT based technique for video micro motion magnification on these three videos. We detected clear difference in performance, of detecting live human tissues motion, between 3D live human face videos and 2D face image video sequences. This experiment was repeated on videos for 5 different subjects, 3 videos for each subject as described above, same clear distinction was achieved for all subjects. In order to demonstrate the micro movement magnification on videos in this manuscript we take the line in the middle of the subject face, Figure 9(d), and time graph it for all video sequence frames. It can then be displayed as in Figure 9 (a,b,c) as line

motion with respect to frame number. We can then take a horizontal slice of Figure 9 (a,b,c) to display the micro motion/vibration of a specified pixel across all frames, as in Figure 8. Our proposed DWT and CWT based techniques of micro motion magnification can clearly distinguish between 3D live facial videos with their live tissues micro motion and 2D rigid facial image videos, as in Figure 8, with more enhanced performance with the CWT based technique (170 simulation time for CWT on average compared to 23 sec for DWT). Figure 8 shows a cross section of the time line motion of the face middle line video in Figure 9(d). In Figure 8, the solid fixed line shows a specific pixel for this face middle line for a 2D facial snapshot video without any magnification or manipulation. It is normal to expect it in a fixed position. The dotted blue line (shacking), which is close to the fixed line, shows this same pixel in the middle line for a video of a 2D snapshot image with camera shacking as mentioned above. Also in Figure 8, the star (asterick) line shows this same middle line for a 3D facial video after DWT magnification of micro movement. Finally also in Figure 8, the solid black line with a significant motion represents this same pixel for the face middle line for a 3D facial video after CWT magnification of micro movement. Figure 9 shows a comparison of a middle line face video between a fake 2D image face video (Figure 9,a) and a real live 3D facial video (with both DWT and CWT magnifications, Figure 9 b,c). The graph in Figure 9 (a), (b) and (c) displays the face middle line across 260 frames for each of the videos described here. Figure 9(a) shows this face time-line for basic 3D face video without any magnification or manipulation. Figure 9(b) shows this same figure after DWT micro movement magnification, sec 3. Figure 9(c) shows this same figure after CWT magnification, sec 4. Clear distinction was found for all videos for this middle face line, Figure 9 (b) and (c). This time line deviation of Figure 9 is our method of measuring the magnification factor. Table 2 shows this clear distinction for another 4 different subjects, where the standard deviation is listed for each graph for each subject.
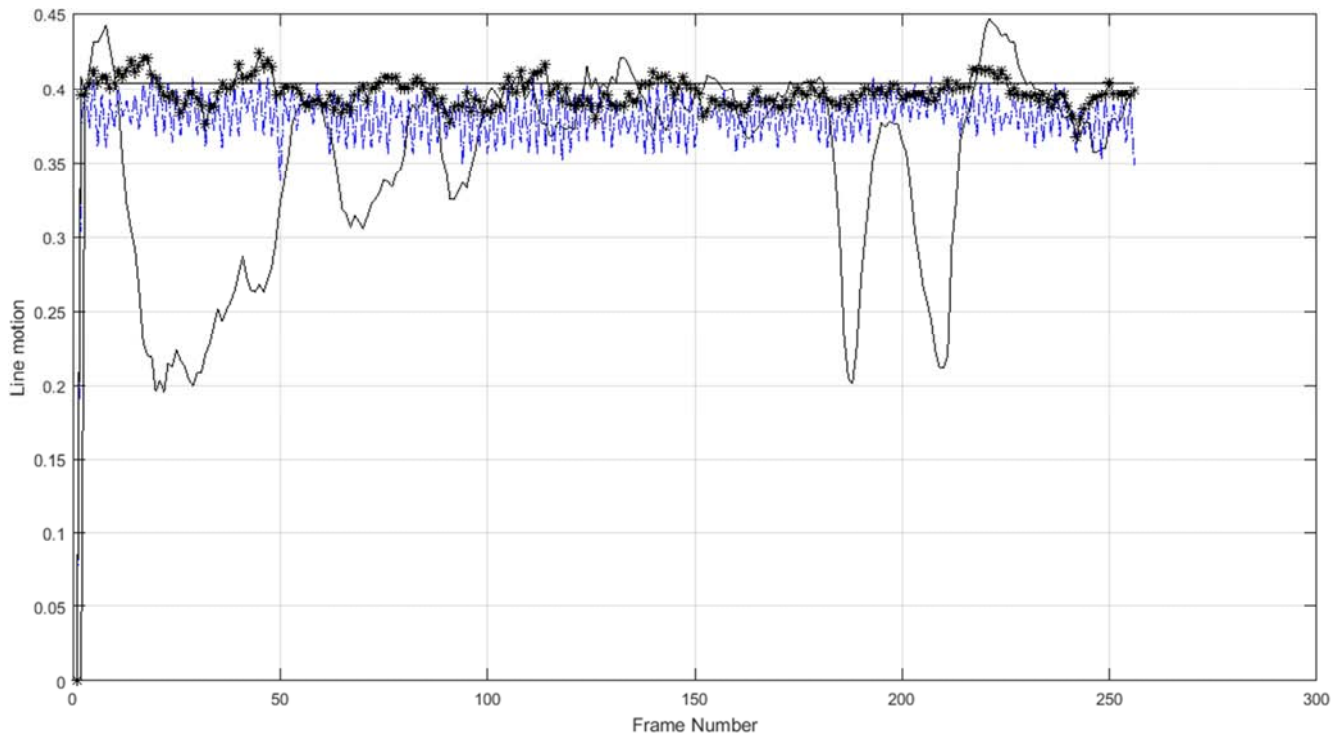
Figure 8. *Detection time line pixel motion across all frames. This correlates well with periodic tissues movement in 3D facial videos. (a) Without motion magnification with natural camera vibration/motion (b) after DWTmotion magnification (c) Without magnification or camera movement (d) after CWT magnification.*
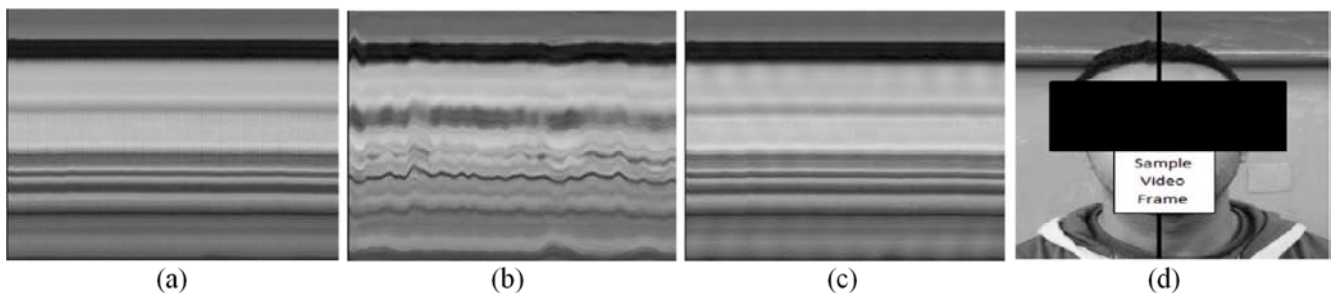

Figure 9. *An Example of time line motion across all frames for a middle face video line. This is performed between a 3D live facial video and a 2D image sequence fake video, (a) Without motion magnification with fixed camera motion (b) after DWT magnification (c) after CWT magnification (d).*

Table 2. *Distinction of Live periodic micro motion for 3D facial videos.*

Subject no.	2D Still image Video	3D facial video after DWT magnification	3D facial video after CWT magnification
2	0.023	0.036	0.069
3	0.023	0.039	0.072
4	0.022	0.035	0.067
5	0.021	0.042	0.077

We note here that the micro movement magnified in Facial videos, are mainly due to the live periodic blood circulation, this live periodic motion should be consistent with heart beep rate that can be detected from subjects with more sophisticated medical equipment. It is also normal to have zero movement with 2D snapshot image video sequences with or without camera shacking. In, [6], research has been conducted with magnification of micro changes of videos but with brightness changes, unlike our own technique direction which focuses on magnifying micro movements similar to most recent literature, [4, 5, 7]. It is our belief that the combination of both direction could lead to significant detection of live skin objects.

6. Conclusion

This paper first describes a fast DWT based micro movement magnification technique for online 3D facial video identification from fake 2D facial image videos. We then present a CWT based technique for video micro movement magnification that is based on an orientation gradient and a Radon Transform based CWT technique to detect live periodic

motion of live facial tissues in 3D facial object identification. Our proposed techniques in both DWT and CWT approaches achieves superior performance with micro motion magnified videos, when compared to existing literature results. Our proposed CWT system identifies regions with largest motion because of the CWT orientation features with high precision. Further, applying Radon Transform of the magnified video frames enables us to measure biomedical and identification features just from the videos. This work has been supported mainly from Alexander Von Humboldt foundation, Germany.

References

[1]   Jue Wang, Steven M Drucker, Maneesh Agrawala, and Michael F Cohen, "The cartoon animation filter," in Proceeding of ACM Transactions on Graphics (TOG). ACM, 2006, vol. 25, pp. 1169–1173.

[2]   Ming-Zher Poh, Daniel J McDuff, and Rosalind W Picard, "Non-contact, automated cardiac pulse measurements using video imaging and blind source separation.," Optics express, vol. 18, no. 10, pp. 10762–10774, 2010.

[3]   Martin Fuchs, Tongbo Chen, Oliver Wang, Ramesh Raskar, Hans-Peter Seidel, and Hendrik PA Lensch, "Real-time temporal shaping of high-speed video streams," Computers & Graphics, vol. 34, no. 5, pp. 575–584, 2010.

[4]   Yichao Zhang, Silvia L. Pintea, and Jan C. van Gemert, "Video acceleration magnification," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 2017.

[5]   Tae-Hyun Oh, Ronnachai Jaroensri, Changil Kim, Mohamed Elgharib, Fr'edo Durand, William T. Freeman, and Wojciech Matusik, "Learning-based video motion magnification," in The European Conference on Computer Vision (ECCV), September 2018.

[6]   H. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman, "Eulerian video magnification for revealing subtle changes in the world," Proceeding of ACM Transactions on Graphics, Volume 31, Number 4 (Proc. SIGGRAPH) 2012.

[7]   Neal Wadhwa, Michael Rubinstein, Fredo Durand, and William T. Freeman, "Phase-based video motion processing," Proceeding of SIGGRAPH, ACM Transactions on Graphics 2013, vol. 32, no. 4, pp. 80, 2013, https://doi.org/10.1145/2461912.2461966.

[8]   Neal Wadhwa, Michael Rubinstein, Fredo Durand, and William T. Freeman, "Riesz pyramids for fast phase based video magnification," Proceeding of IEEE International Conference on Computational Photography (ICCP), 2014, DOI: 10.1109/ICCPHOT.2014.6831820.

[9]   Neal Wadhwa, Michael Rubinstein, Fredo Durand, and William T. Freeman, "Quaternionic representation of the riesz pyramid for video magnification," MIT Computer Science and Artificial Intelligence Laboratory, Technical report, 2014, URI http://hdl.handle.net/1721.1/86300.

[10] Gamal Fahmy, Omar Fahmy, and Mamdouh Fahmy, "Micro movement magnification in video signals using complex wavelet analysis," Journal of IET on Image Processing, Vol. 11, Issue 11, November 2017, p. 986 – 993, DOI: 10.1049/iet-ipr.2017.0049.

[11] Omar Fahmy, Gamal Fahmy, and Mamdouh F. Fahmy, "A new video magnification technique using complex wavelets with radon transform application," Journal of Signal, Image and Video Processing, June 2018, https://doi.org/10.1007/s11760-018-1306-9. vol. 12, pp. 1505–1512.

[12] Andrea F. Abate, Michele Nappi, Daniel Riccio, and Gabriele Sabatino, "A survey: 2d and 3d face recognition," in Pattern Recognition Letters, Pages 1885-1906, Volume 28, Issue 14, 15 October 2007, https://doi.org/10.1016/j.patrec.2006.12.018.

[13] Eero P Simoncelli and William T Freeman, "The steerable pyramid: A flexible architecture for multi-scale derivative computation," in Image Processing, 1995. Proceedings., International Conference on. IEEE, 1995, vol. 3, pp. 444–447.

[14] G. Fahmy, M. Fahmy, and O. M. Fahmy, "Complex wavelet transform cwt based video magnification for 3D facial video identification," in 2020 International Conference on Computing, Electronics Communications Engineering (iCCECE), 2020, pp. 75–80.

[15] Gamal Fahmy, Omar Fahmy, and Mamdouh Fahmy, "Fast enhanced dwt based video micro movement magnification," in Proceeding of IEEE International Symposium on Signal Processing and International Technology, ISSPIT, Dec. 2019, Ajman UAE, DOI: 10.1109/ISSPIT47144.2019.9001874.

[16] Til Aach and Hartmut Fuhr, "Shift variance measures for multirate lpsv filter banks with random input signals," IEEE Transaction on Signal Processing, vol. 60, no. 10, pp. 5125–5134, 2012, DOI: 10.1109/TSP.2012.2205683.

[17] I. Selesnick, R. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform," Signal Processing Magazine, IEEE, vol. 22, no. 6, pp. 123–151, 2005, DOI: 10.1109/MSP.2005.1550194.

[18] N. Kingsbury, "A dual-tree complex wavelet transform with improved orthogonality and symmetry properties," in Proceedings of 2000 IEEE International Conference on Image Processing. DOI: 10.1109/ICIP.2000.899397, 2000, vol. 2, pp. 375–378.

[19] J. Ng A. A. Bharath, "A steerable complex wavelet construction and its application to image denoising," Trans. on Image Processing, vol. 14, no. 7, pp. 948–959, 2005.

[20] Mark JT Smith and Alen Docef, A study guide for Digital Image Processing, Scientific Publishers, 1997.

[21] M. Pattichis C. Carranze, D. Lllamocca, "Fast and scalable computation of the forward and inverse discrete periodic radon transform," IEEE Transaction on Image Processing, vol. 25, no. 1, pp. 119–133, 2016, DOI: 10.1109/TIP.2015.2501725.

[22] Mohamed Elgharib, Mohamed Hefeeda, Fredo Durand, and William T. Freeman, "Video magnification in presence of large motions," in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.