IJE TRANSACTIONS C: Aspects Vol. 29, No. 3, (March 2016) 313-320

International Journal of Engineering

Journal Homepage: www.ije.ir

A New Wavelet Based Spatio-temporal Method for Magnification of Subtle Motions in Video

S. V. Shojaedini *[a], M. M. Koohi [b], R. K. Haghighi [a]

[a] *Department of Electrical Engineering and Information Technology, Iranian Research Organization for Science and Technology, Iran*
[b] *Department of Electrical, Biomedical and Mechatronics Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

PAPER INFO

Paper history:
Received 11 January 2016
Received in revised form 07 February 2016
Accepted 03 March 2016

Keywords:
Video Magnification
Spatio-temporal Analysis
Wavelet Transform
Connectivity Based Mapping

ABSTRACT

Video magnification is a computational procedure to reveal subtle variations during video frames that are invisible to the naked eye. A new spatio-temporal method which makes use of connectivity based mapping of the wavelet sub-bands is introduced here for exaggerating of small motions during video frames. In this method, firstly the wavelet transformed frames are mapped to connectivity space and then decomposed into different spatial frequency bands by applying Laplacian Pyramid to determine the pixels having more chance to be a part of a movement. Finally each candidate is partially magnified based on its time history. The performance of the proposed method is evaluated on real videos which contain several subtle motions. Parameters for performance evaluation are presented and obtained results are compared with one of the state-of-the-art video magnification methods. Increased true positive rate parallel with simultaneous decrease in false positive rate confirms the effectiveness of the proposed method in amplifying subtle motions.

doi: *10.5829/idosi.ije.2016.29.03c.05*

1. INTRODUCTION

The world is full of small spatio-temporal variations which are impossible to monitor with naked eyes. Variable color in skin happens as blood circulates [1] and human head shakes with each heart beat [2]. Video magnification is the duty to magnify such variations to show meaningful and important small motions [3, 4]. Some techniques have been used for video magnification including Lagrangian methods, Eulerian methods, phase-based methods and is Dynamic Video Magnification (DVM) methods.

In Lagrangian video magnification techniques [3], motions are estimated explicitly by extracting feature point trajectories followed by segmenting them into two stationary and moving sets. In these techniques some motion models (for example affine model) are fitted to the stationary points which register the examined sequence on a reference frame and then motions are re-estimated, scaled and added back to the registered sequence. This strategy generates the magnified output. Some other researches utilize Eulerian video magnification (EVM) to detect small motions in image sequences [4]. In this approach an input sequence is first decomposed into a multiscale stack (Laplacian or Gaussian) and then subtle variations are temporally filtered. Finally the variations scaled and added back to the input sequence, therefore a magnified output may be rendered [5]. EVM is able to amplify small motions in videos without explicitly computing optical flow. Unlike Lagrangian approaches which may only magnify motion changes, Eulerian methods are able to magnify motion as well as color changes. Unfortunately this algorithm supports only small magnification factors at high spatial frequencies; therefore it may magnify noise significantly when the magnification factor is increased.

Phase-based methods [6] are a group of the most popular techniques for video magnification. In these techniques the steerable pyramid [7] decomposes an image according to spatial scale, orientation, and position. The sub-sampling scheme in the steerable

*Corresponding Author's Email: shojadini@irost.ir (S. V. Shojaedini)

Please cite this article as: S. V. Shojaedini, M. M. Koohi, R. K. Haghighi, A New Wavelet Based Spatio-Temporal Method for Magnification of Subtle Motions in Video, International Journal of Engineering (IJE), TRANSACTIONS C: Aspects Vol. 29, No. 3, (March 2016) 313-320

pyramid keeps it away from spatial aliasing and therefore allows meaningful signal phase measurements from the coefficients of the pyramid. Such a scheme supports larger amplification factors and is significantly less sensitive to noise compared to the EVM method. As the amplification factor is increased, noise is translated rather than amplified. Unfortunately phase-based motion magnification is limited by the specific support of the complex steerable pyramid filters. The Riesz pyramid is often used for real-time phase-based motion magnification. Several investigations have used Riesz pyramid concept for real-time phase-based motion magnification. Motion-magnified videos which are produced with this representation have comparable qualities to those produced with the complex steerable pyramid. Riesz pyramid is efficiently implemented because of shared computing between bands, symmetry of the filters, and because the Riesz transform is approximated by two three tap limited difference filters. The key intuition into why the Riesz transform may be used is that it is a steerable Hilbert transformer and allows us to compute a quadrature pair that is 90 degrees out of phase with respect to the most dominant orientation at every pixel [8].

Another technique is DVM [9] which contains two main components: Warping to discount large motions and Layer-based Magnification. The Warping stage seeks to remove large motions while preserving small ones. Layer-based magnification is based on decomposing image into a foreground, background through an alpha matte. An alpha matte is a piece of footage that tells the program running it exactly what is supposed to be seen through which parts. This method is capable of handling small motions within large ones and consequently shows larger amplification factors and significant reduction in artifacts over state of the art.

In this paper a new method is introduced to magnify small motions in video sequences by utilizing wavelet sub-bands. In the proposed method firstly the 2D wavelet transform is applied on the video under test and then the transformed images are mapped to connectivity space which leads to exaggeration of subtle variations in each frame. Each mapped frame is decomposed into different spatial frequency bands by applying Laplacian Pyramid to extract those pixels having more chance to include small motions (i.e. candidate pixels). Finally each candidate pixel in mapped images is partially magnified based on its time history which leads to more magnification for candidates having more chance of being a part of a movement. In the proposed method the mapping procedure leads to selecting candidates not only based on their own intensities but also due to connectivity with their neighboring pixels. Such a strategy leads to ignoring noisy pixels which logically have less connectivity with other pixels existing in their around region. This fact enables the proposed algorithm to support less sensitivity to noise even in presence of large amplification coefficients.

The paper is organized as follows. In section (2) the proposed algorithm is introduced which includes wavelet based connectivity mapping, pyramid decomposition and temporal processing. In section (3) the performance of the proposed method is evaluated in important frames of video sequence to investigate the resultant improvement due to applying the proposed scheme. In section (4), the obtained results are compared with results obtained from some existing methods using their effective parameters. Conclusion is presented in the last section of the paper.

2. MATHEMATICAL MODEL

The proposed method is composed of spatial and temporal processing. In spatial processing the video sequence is decomposed into different spatial frequency bands and then temporal processing is applied on successive sequences of frequency bands in order to magnify probable small motions.

2. 1. Spatial Processing Suppose I is a video sequence and I_t is a frame which occurs at time slot t. For this frame $I_t(x,y)$ is the brightness of a pixel which is located in row x and column y. Now suppose smoothing function $\beta(x,y)$ which satisfies the following conditions [10]:

$$\begin{aligned}&\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\beta(x,y)dxdy = 1\\&\lim_{x^2+y^2\to\infty}\beta(x,y)=0\end{aligned} \tag{1}$$

Then the scaled version of the smoothing function at scale s is defined as follow:

$$\beta_S(x,y)=\frac{1}{s}\beta\left(\frac{x}{s},\frac{y}{s}\right) \tag{2}$$

Equation (3) shows that the scaled version of the first derivative of smoothing function $\beta(x,y)$ may be used as a basic function for mother wavelet function for purposes of edge detection. The Fourier transform of this function is focused near zero and therefore acts as a noise suppressor by low-pass filtering. The mother wavelet functions may be defined as the scaled derivatives of $\beta(x,y)$ as:

$$\begin{aligned}&\mu_s^1(x,y)=s\frac{\partial}{\partial x}\beta_s(x,y)\\&\mu_s^2(x,y)=s\frac{\partial}{\partial y}\beta_s(x,y)\end{aligned} \tag{3}$$

According to the above mother wavelet functions two-dimensional wavelet transform is computed as follows [11]:

$$W_s^1 I_t(x,y) = I_t(x,y) * \mu_s^1(x,y)$$
$$W_s^2 I_t(x,y) = I_t(x,y) * \mu_s^2(x,y) \quad (4)$$

This procedure may be simply written as:

$$\begin{bmatrix} W_s^1 I_t(x,y) \\ W_s^2 I_t(x,y) \end{bmatrix} = s \begin{bmatrix} \frac{\partial}{\partial x}(I_t * \beta_s)(x,y) \\ \frac{\partial}{\partial y}(I_t * \beta_s)(x,y) \end{bmatrix} = s\nabla(I_t * \beta_s)(x,y) \quad (5)$$

The amplitude of the above terms which indicates the magnitude of the changes in image intensity is computed as [12]:

$$I_t'(x,y) = \sqrt{\left|W_s^1 I_t(x,y)\right|^2 + \left|W_s^2 I_t(x,y)\right|^2} \quad (6)$$

In the next step, important changes are extracted by applying Laplacian Pyramid [13] on $I_t'(x,y)$ as follows.

Firstly, $I_t'(x,y)$ is convolved with a Gaussian kernel which leads to its low pass filtered version. The Laplacian is then computed as the difference between $I_t'(x,y)$ and the mentioned low pass filtered image. This process is continued to obtain a set of band-pass filtered images (since each of them is the difference between two levels of the pyramid). In the rest of paper $E_l\left(I_t'(x,y)\right)$ demonstrates the value of a pixel which is located at position (x,y) in l th pyramid of $I_t'(x,y)$. A fast algorithm for generating these pyramids is given as a Pseudo code in Figure 1. The generation of pyramid involves the use of two complementary functions REDUCE and EXPAND, which the former decreases the size of image and the latter performs the reverse operation.

2. 2. Temporal Processing

After a video has been spatially processed, it is then subjected to temporal processing. For this purpose firstly the history of pyramids which have been estimated for each pixel is constructed via a time sequence as:

$$I_{tl}'(x,y,t) = \left\{E_l\left(I_0'(x,y)\right), E_l\left(I_1'(x,y)\right), \ldots, E_l\left(I_t'(x,y)\right)\right\} \quad (7)$$

The relation between temporal processing and motion magnification [12] may be demonstrated based on Taylor series expansion for $I_{tl}'(x,y,t)$ as bellow:

$$I_{tl}'(x,y,t) \approx I_{tl}''(x,y,t) + \zeta(x,y,t)\frac{\partial I_{tl}''(x,y,t)}{\partial x\,\partial y} \quad (8)$$

In above equation $\zeta(x,y,t)$ is displacement function which shows translational motion in image, such that $I_{tl}'(x,y,t) = I_{tl}''(x, y + \zeta(x,y,t))$ and therefore it may be simply written that $I_{tl}'(x,y,0) = I_{tl}''(x,y)$. Assuming that the motion signal $\zeta(x,y,t)$ is within the band pass zone, the temporal band pass filter may be constructed as:

$$U(x,y,t) = \zeta(x,y,t)\frac{\partial I_{tl}''(x,y,t)}{\partial x\,\partial y} \quad (9)$$

Begin Loop
For frame length (T)
Input set of frames as $I_t'(x,y)$
While (t<T)
Initialize primary Gaussian as
$$\left(G_0\left(I_t'(x,y)\right) = I_t'(x,y)\right)$$
For $l = 1:L$
Calculate Gaussian Kernel as:
$$w(r,c) = w(r).w(c)$$
Set weights as,
$$w(.) = \left[\frac{1}{4} - \frac{\gamma}{2}, \frac{1}{4}, \gamma, \frac{1}{4}, \frac{1}{4} - \frac{\gamma}{2}\right]$$
Do reduction by REDUCE function
$$G_{l+1}\left(I_t'(x,y)\right) = REDUCE\left(G_l\left(I_t'(x,y)\right)\right)$$
Where:
$REDUCE$: Conv (image, Gaussian kernel)
End Do
Compute $EXPAND\left(G_{l+1}\left(I_t'(x,y)\right)\right)$ **as**
$$EXPAND\left(G_{l+1}\left(I_t'(x,y)\right)\right) = 4\sum_{r=-2}^{2}\sum_{c=-2}^{2} w(r,c) G_l\left(I_t'\left(\frac{x-r}{2}, \frac{y-c}{2}\right)\right)$$
Estimate Laplacian Pyramid as:
$$E_l(x,y) = I_t'(x,y) - EXPAND\left(G_{l+1}\left(I_t'(x,y)\right)\right)$$
End For
End While
End Loop

Figure 1. Pseudo code estimating Laplacian Pyramid

Based on the above equations, amplifying the above band pass term by using amplification factor α leads to exaggeration of small motion $\zeta(x,y,t)$. Finally the amplified motion is added back to the last frame of sequence $I_{tl}'(x,y,t)$ which leads to $I_{tl}^{AM}(x,y,t)$ as its magnified version.

$$I_{tl}^{AM}(x,y,t) = E_l\left(I_t'(x,y)\right) + \alpha U(x,y,t) \quad (10)$$

As all steps of the temporal processing are done in pyramid domain, i.e. as different frequency bands based on pyramid definition, therefore the magnified terms which are resulted from Equation (10) should be incorporated to reconstruct the final magnified image. Consequently in the last step of the proposed algorithm, the frequency bands are restructured to form a video in which motions are magnified, after they have been pooled and processed. The Laplacian Pyramid reconstructed back with the reverse procedure described in previous section [12, 14].

3. TESTS AND RESULTS

In order to evaluate the performance of the proposed algorithm it was applied on a set of videos containing several small motions. Table 1 depicts some important parameters corresponding to videos. More details about the dataset may be found elsewhere [4].

The proposed method was implemented using Matlab 2014a on a PC with a six-core CPU with 2.40GHz processor and 32 GB RAM.

TABLE 1. Specifications of examined videos

Specifications of input videos	Baby	Shadow
Frame. rate	30 fps	30 fps
Number of frames	301 frames	180 frames
Frame size	960*554	960*624
Duration	10 s	6 s

Figure 2. Representative frames from two examined videos (so-called baby and shadow)

Additionally, EVM [4] was implemented to compare with the proposed algorithm. Figure 2 shows two representative frames belonging to so called baby and shadow videos, respectively.

For brevity some results obtained from applying the proposed and EVM algorithms on these two videos are graphically compared in this part of article by focusing on those areas which contain small motions, but the complete statistics of the test results will be discussed in next section.

Figure 3 shows the regions including movement in four successive frames of the baby video. Figures (4-a) to (4-d) show results obtained from the magnification of theses frames by utilizing EVM method and similarly Figures (5-a) to (5-d) show magnified frames by using proposed method. It may be considered in Figure (4-a) that the main and side zipper heights obtained as 20 and 5 pixels using EVM, while the proposed algorithm gave these parameters as 22 and 13 pixels as shown in Figure (5-a). Another example may be the comparison of Figures (4-b) and (5-b) which in the former the heights of zipper have been obtained as 16 and 6 pixels while in latter these parameters have been computed as 17 and 16 pixels.

Furthermore increasing of contrast between moving and static areas has been computed as another parameter for quantifying the effectiveness of video magnification. Based on this idea, the average of contrast for EVM results has been estimated as 26% while for the proposed method it has been achieved as 31%. Two next frames also show similar results which indicate that the motion of the zipper containing area is more visible when observed after applying the proposed spatio-temporal processing.

Figures 6 to 8 illustrate the similar analysis on another video sequence, so called shadow, which contains several motions but with some different contents and specifications. Figure 6 shows original frames. Figure 7 shows that a number of branches and leaves which were invisible in original frames have become apparent by applying EVM while the proposed method has extracted these details more accurate than EVM as shown in Figure 8. Some samples for such regions have been indicated in corresponding frames of these figures by a red circle. For instance the indicated area in (8-a) shows the branches more distinguished than what appeared in (7-a). Other frames also show similar results which demonstrate that the proposed algorithm has extracted details of branches and leaves better than its alternative.

4. PERFORMANCE EVALUATION

To investigate the effectiveness of the proposed algorithm, real videos including several types of small motions were analyzed using the proposed method and then the obtained results were compared with the results obtained from EVM method by utilizing some standard parameters.

The first parameter was the similarity between the corresponding original and processed frames which is called structural similarity (SSIM) index. SSIM is defined as an image quality metric that assesses the visual impact of three characteristics of an image:

luminance, contrast and structure [15]. The measure between two windows centered at x and y of common size $N \times N$, was defined as:

$$SSIM\,(x,y)=\frac{(2\rho_x\rho_y+c_1)(2\sigma_{xy}+c_2)}{(\rho_x^2+\rho_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \tag{11}$$

In which ρ_x , ρ_y and σ_x , σ_y are the average and variance of x and y. Further $C_1=(K_1D)^2$, $C_2=(K_2D)^2$ represent two variables for stabilizing the division with weak denominator, in them D demonstrates the dynamic range of the pixel-values and $K_1=0.01$, $K_2=0.03$ by trial and error.

Figure 9 shows the obtained changes in SSIM for two different videos (i.e. baby, shadow) across their successive frames. It should be noted that the computed SSIM for the proposed method was approximately 2% less than the obtained value for EVM algorithm for so called baby sequence. This gap is more when the shadow sequence was investigated in such way that the proposed method led to SSIM approximately 10% less than the obtained value for EVM.

Other evaluation parameters were True Positive Rate (TPR) and False Positive Rate (FPR) which were defined as:

$$TPR=\frac{TP}{TP+FN} \tag{12}$$

$$FPR=\frac{FP}{FP+TN} \tag{13}$$

In above equations True Positive (TP) was defined as the number of pixels which their movement was exaggerated correctly. The static pixels which had been properly unmagnified were counted as True Negatives (TN). False Positive (FP) was defined as the number of static pixels that had been incorrectly magnified.

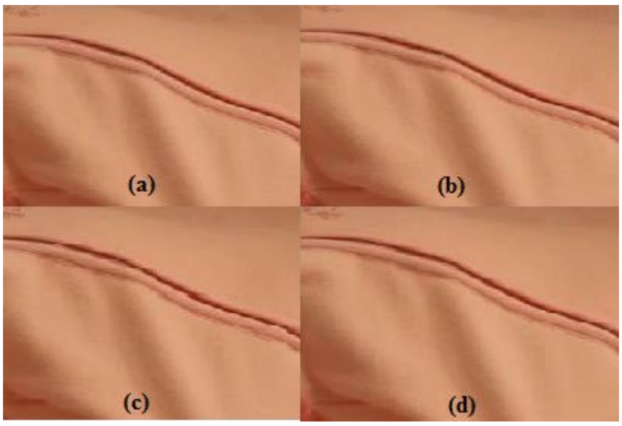

Figure 3. Four frames from the original video sequence (baby).

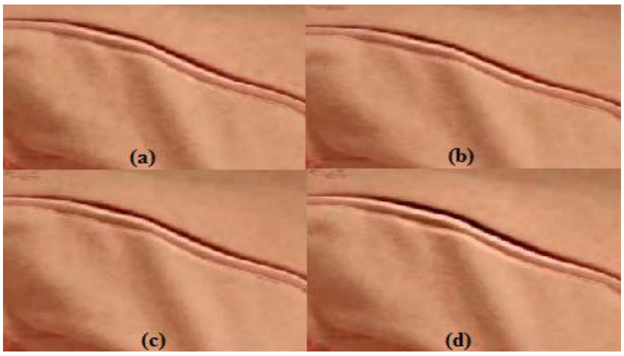

Figure 4. The same four frames with subtle motion in zipper amplified using EVM

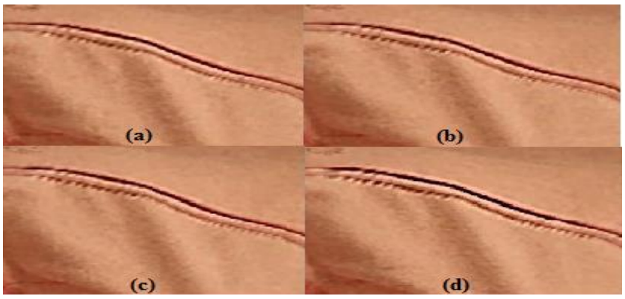

Figure 5. The same four frames with subtle motion in zipper amplified using proposed method

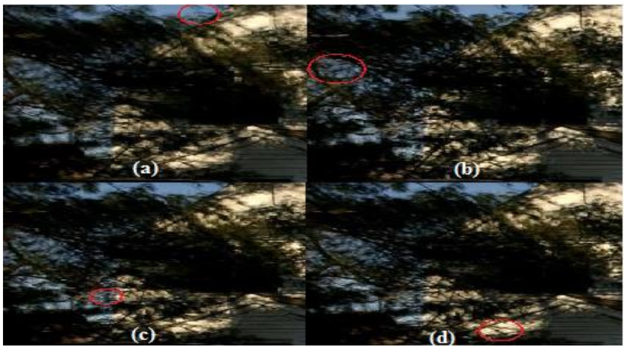

Figure 6. Four frames from the original video sequence (shadow).

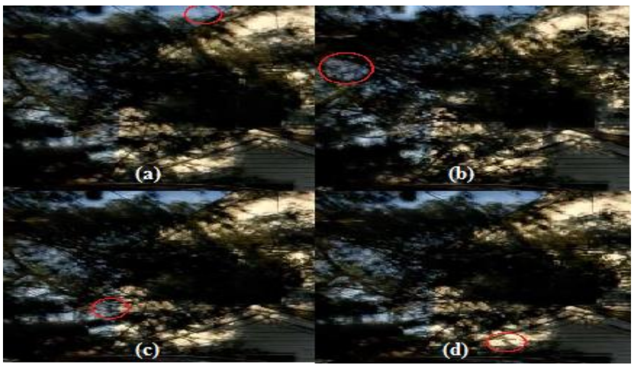

Figure 7. The same four frames with subtle motion amplified using EVM

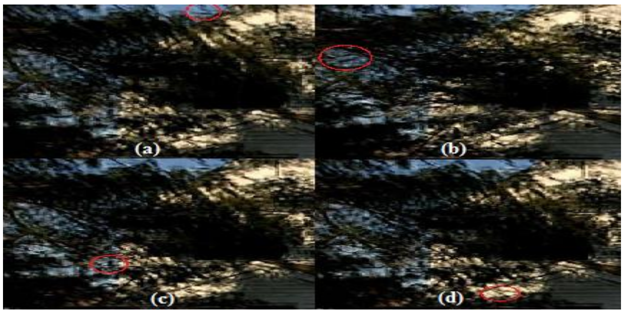

Figure 8. The same four frames with subtle motion amplified using proposed method

For computing Equations (12) and (13), firstly several frames were inspected to mark moving regions as ground truth to compare the automatic methods (for instance the zipper area in video so-called baby).Then the regions which were exaggerated using automatic methods were justified with the above ground truth. By using this strategy, the changes of detection rate versus false detection rate (e.g. ROC curve) were obtained for baby and shadow videos as shown in Figures (10-a) and (10-b). For better interpretation of results FPR=5% and TPR =95% were considered as typical acceptable thresholds for false detection and detection rates which led to Table 2 for baby sequence and Table 3 for shadow sequence. By exploiting these tables better performance of the proposed method was proven compared to EVM algorithm in such way that the proposed algorithm achieved to TPR values 2.2% and 3.6% better than EVM for two above sequences, in presence of false detections equal to 5%. In similar way the proposed method reached to FPR value which was 17 and 19% better than its alternative in presence of at least 95% of detection rate.

TABLE 3. The obtained results from examining algorithms on shadow video

Algorithm	TPR (v.s.FPR=5%)	FPR (v.s.TPR=95%)
EVM	92.3%	20%
Proposed Method	95.9%	1%

TABLE 2. The obtained results from examining algorithms on baby video

Algorithm	TPR (v.s.FPR=5%)	FPR (v.s.TPR=95%)
EVM	93.8%	20%
Proposed Method	96%	3%

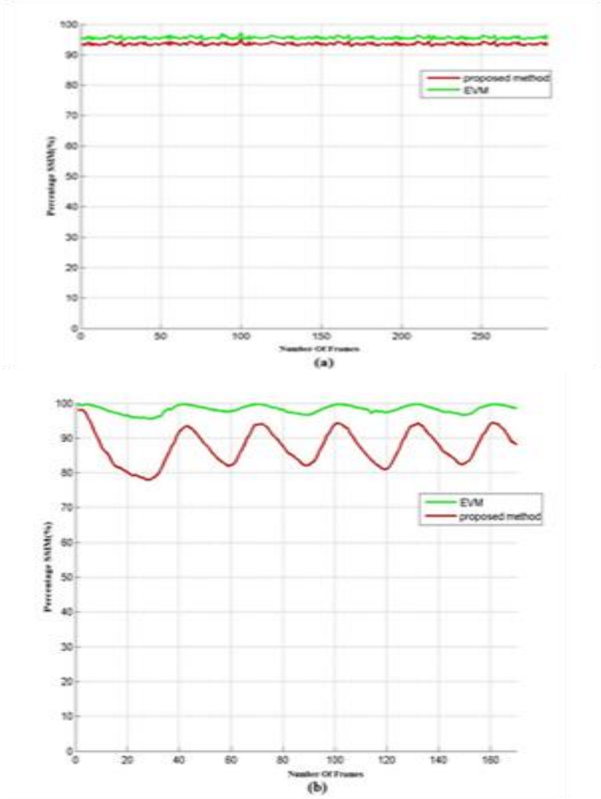

Figure 9. SSIM v.s. frames obtained for (a) video so-called baby and (b) video so-called shadow. Proposed method (line-red) and EVM (line-green)

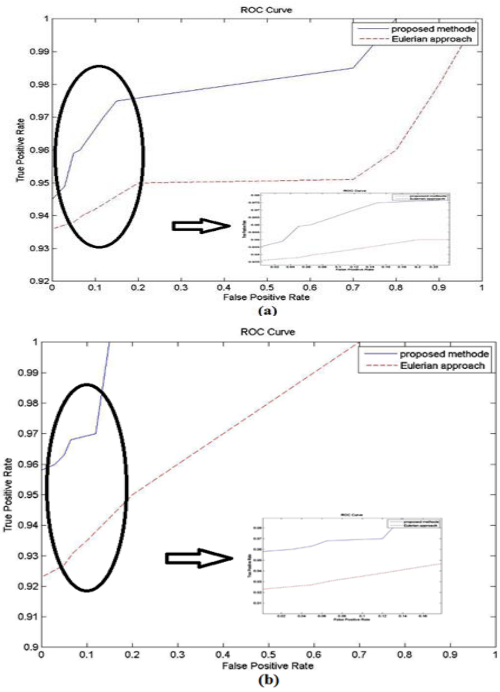

Figure 10. ROC curves obtained for (a) video so- called baby and (b) for video so-called shadow. Proposed method (solid line-blue), EVM (red dashed-line).

5. CONCLUDING REMARKS

In this paper a new method was introduced for exaggerating subtle motions in video using wavelet based spatio-temporal processing of image sequence. This procedure was started by selecting some candidates to be member of moving regions due to their connectivity in wavelet mapped plane. Then Laplacian pyramid decomposition followed by estimating a displacement function was utilized to amplify the small motions. This connectivity-based scheme enables the proposed algorithm to highlight subtle motions in video without incorporating noisy pixels which logically have less connectivity with the around pixels.

To evaluate the performance of the proposed algorithm, it has been applied on several video frames containing small movements. The performance of the proposed algorithm was also compared with EVM approach in terms of three effective parameters including SSIM, TPR and FPR. Numerical comparison showed better performance of the proposed algorithm in detecting and amplifying small motions compared to its alternative. It was observed that the proposed algorithm has magnified subtle motions at least 2.2% and better than EVM method in presence of FPR equal to 5%. Furthermore it was shown that false detection rate of the proposed algorithm has been at least 17% better than those which had been obtained for EVM. This considerable improvement in FPR initiates from the noise suppression nature of the proposed algorithm which has been explained in body of paper. Finally structural similarity index belonging to the proposed algorithm have been obtained minimally 2% better than those obtained by the EVM for the examined sequences.

Consequently, we believe that the proposed method may be used as a suitable alternative for magnifying motions in video sequences which include weak movements.

6. REFERENCES

1. Rubinstein, M., Wadhwa N., Durand F., Freeman, W. and Eugene, H., "Revealing invisible changes in the world", MIT Computer Science and Artificial Intelligence Lab (CSAIL), Vol. 339, (2013), 518-519.
2. Balakrishnan, G., Durand, F. and Guttag, J., "Detecting pulse from head motions in video", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, (2013), 3430-3437.
3. Wu, H.-Y., Rubinstein, M., Shih, E., Guttag, J.V., Durand, F. and Freeman, W.T., "Eulerian video magnification for revealing subtle changes in the world", MIT Computer Science and Artificial Intelligence Lab, (2012).
4. Fuchs, M., Chen, T., Wang, O., Raskar, R., Seidel, H.-P. and Lensch, H.P., "Real-time temporal shaping of high-speed video streams", ***Computers & Graphics***, Vol. 34, No. 5, (2010), 575-584.
5. Wadhwa, N., Rubinstein, M., Durand, F. and Freeman, W.T., "Phase-based video motion processing", ***ACM Transactions on Graphics (TOG)***, Vol. 32, No. 4, (2013).
6. Simoncelli, E.P. and Freeman, W.T., "The steerable pyramid: A flexible architecture for multi-scale derivative computation", Proceedings of the International Conference on Image Processing, Vol. 3, (1995).
7. Wadhwa, N., Rubinstein, M., Durand, F. and Freeman, W.T., "Riesz pyramids for fast phase-based video magnification", in Computational Photography (ICCP), IEEE International Conference on, (2014), 1-10.
8. Elgharib, M.A., Hefeeda, M., Durand, F. and Freeman, W.T., "Video magnification in presence of large motions", ***Qatar Computing Research Institute***, (2015), 4119 - 4127.
9. Guojin, C., Miaofen, Z., Yongning, L.I. and Wanqiang, W., "Research on image focusing method and system with wavelet filtering", International Conference on Optoelectronics and Image Processing, (2010), 236 - 239.
10. Hao, Y., Changshun, L. and Lei, P., "An improved method of image edge detection based on wavelet transform", International Conference on Computer Science and Automation Engineering, Vol. 3, (2011), 678-681.
11. Burt, P.J. and Adelson, E.H., "The laplacian pyramid as a compact image code", Communications, IEEE Transactions on, Vol. 31, No. 4, (1983), 532-540.
12. Do, M.N. and Vetterli, M., "Frame reconstruction of the laplacian pyramid", in Acoustics, Speech, and Signal Processing, Proceedings.(ICASSP'01), IEEE International Conference on, Vol. 6, (2001), 3641-3644.
13. Sarode, L. and Mandaogade, N., "Video motion magnification using spatio-temporal algorithm", ***International Journal of Computer Applications***, Vol. 96, No. 9, (2014).
14. Wang, Z., Simoncelli, E.P. and Bovik, A.C., "Multiscale structural similarity for image quality assessment", in Signals, Systems and Computers, Conference Record of the Thirty-Seventh Asilomar Conference on, Vol. 2, (2003), 1398-1402.

A New Wavelet Based Spatio-temporal Method for Magnification of Subtle Motions in Video

S. V. Shojaedini [a], M. M. Koohi [b], R. K. Haghighi [a]

[a] *Department of Electrical Engineering and Information Technology, Iranian Research Organization for Science and Technology, Iran.*
[b] *Department of Electrical, Biomedical and Mechatronics Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

PAPER INFO

Paper history:
Received 11 January 2016
Received in revised form 07 February 2016
Accepted 03 March 2016

Keywords:
Video Magnification
Spatio-temporal Analysis
Wavelet Transform
Connectivity Based Mapping

چکیده

بزرگنمایی ویدئویی، روشی محاسباتی برای آشکارسازی آن دسته از تغییرات ضعیف در فریم‌های ویدئویی است که به طور معمول با چشم غیر مسلح قابل مشاهده نمی باشند. در این مقاله یک روش جدید زمانی- مکانی برای بزرگنمایی حرکات کوچک در فریم‌های ویدئویی معرفی می شود که با استفاده از نگاشت مبتنی بر پیوستگی زیرباندهای تبدیل موجک عمل می نماید. در این روش، ابتدا فریم‌های حاصل از اعمال تبدیل موجک، به فضای پیوستگی نگاشته شده و سپس با استفاده از هرم لاپلاسین به باندهای مختلف فرکانسی مکانی تجزیه می شوند تا آن دسته از پیکسل‌هایی که شانس بیشتری برای متحرک بودن دارند، مشخص گردند. در نهایت هر یک از این پیکسل های کاندیدا براساس سابقه زمانی خود تقویت می‌شود. عملکرد روش ارائه شده، بر روی ویدئوهای واقعی که شامل حرکات ضعیف متنوعی هستند، ارزیابی می گردد. پارامترهای لازم برای ارزیابی عملکرد ارائه شده و نتایج بدست آمده با یکی از روش‌های استفاده شده برای بزرگنمایی ویدئویی مقایسه می گردند. افزایش نرخ آشکارسازی صحیح به موازات کاهش در میزان نتایج کاذب، سودمندی روش پیشنهادی را در تقویت حرکات ضعیف تایید می‌کند.

***doi**: 10.5829/idosi.ije.2016.29.03c.05*