International Journal of Future Generation Communication and Networking
Vol. 13, No. 3s, (2020), pp. 1203–1211

Comparative Study of Motion Amplification Techniques for Video Sequences

Rajkumar Komati[1], Manoj Nagmode[2]

1. Ph.D. Scholar, Department of E&TC, MIT COE Pune,
2. Professor, Department of E&TC, GCOER Avasari, Pune

Abstract:
The purpose of this paper is to review Motion Amplification Techniques in video sequences. Various computational techniques are used to visualize, analyze and efficiently represent imperceptible variation in image and video sequences. These computational methods which are used reveal temporal changes in videos which are not visible to the naked eyes. In these techniques standard video sequence is used as an input on which spatial decomposition is applied then temporal filtering is applied. The filtered output is then amplified to expose the invisible content in the input video. This paper contains four techniques which are used for enhancement of temporal variations in the video: 1. Linear Approximation Method, Eulerian Video Magnification (EVM), 2. Phase Based Video Processing 3. Fast Phase Based Video Processing (Riesz Pyramid) and 4. Enhanced Eulerian Video Magnification (EEVM). Using these techniques, it is possible to magnify and see tiny motions and subtle color variations.
After review it is observed that, the noise level in EVM technique is more which reduced in EEVM technique using motion analysis and image warping. The other two phase based techniques supports larger amplification factor and better noise performance.
Keywords*- Motion Amplification Techniques, EVM technique.*

I. Introduction

We can capture and communicate changes in our environment or surrounding by using digital cameras, microscopic images or satellites images. Images and videos provide an important information about the time varying nature of our world. In recent times there has been a great progress in digital photography, such videos and images are now available in abundance and easy to acquire, but for understanding and analyzing the time varying processes and trends in visual data, computational methods and tools are scarce and undeveloped [4].

With modern photography and tools it is possible to acquire or capture a video of physical phenomena occurring over different scales of time. On one side modern photography supports very high-speed imaging with 6 MHz frame rates, which allows high quality capture of ultrafast events like neural activity and shock waves. On other side, time-lapse sequences can expose long term processes which spans over decades, such as melting of glaciers, evolution of different cities, and de-forestation, and have even recently become available on a planetary scale. However, tools and techniques to identify and analyze physical processes and trends in visual data are still in their early stage [4].

There are some self-explanatory motions which are too small in amplitude, below human visual spatio-temporal sensitivity. The variations which are not visible for human eyes can be useful to obtain important information. e.g. respiratory motion, the human skin colour changes with blood circulation can be used to obtain pulse rate [4][6]; high speed videos such as tiny eye movement, vibration of engines, long term physical processes like melting of glacier, growth of plants which can reveal biological and physical changes [1].

The motions which are low in spatial amplitude are difficult to see for human, can be amplified to reveal mechanical behavior [2].

ISSN: 2233-7857 IJFGCN

In this paper section 1 is Introduction. Section 2 describes block schematic of EVM method. Section 3 is literature survey of four techniques which are used for enhancement of temporal variations in the video. Section 4 is comparative study of above mentioned four methods. Section 5 explains the algorithm for EVM technique. Section 6 is discussion and section 7 is conclusion.

II. BLOCK SCHEMATIC:

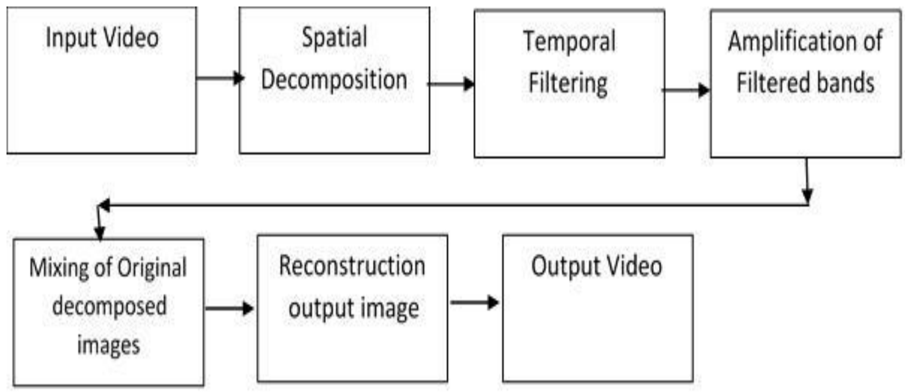


Figure 1: Generalised Block Diagram of **Eulerian Video Magnification (**EVM)

In order to reveal the changes in the videos, different methods are proposed, the various steps involved in revealing subtle changes in the videos or image sequences are as follows,

1. Take standard input video of the process.
2. Apply spatial decomposition on input video using Laplacian pyramid.
3. As per the application apply band pass filter to decomposed spatial bands to filter the required frequencies in each band.
4. Each filtered band is amplified with some amplification factor
5. Amplified filtered bands are mixed with original decomposed spatial input bands.
6. Reconstruction of final output video is done by adding the magnified signal to original

It uses standard video sequence as an input. Then it is decomposed into different frequency bands using Gaussian Pyramid of 3-4 levels to find out the band of interest. The video frame is passed through Low Pass Filter (LPF) and then down samples it to construct the Laplacian Pyramid [6]. Different frequency bands have different signal to noise ratio they may need to amplify with signal and disintegrating the spatial pyramid.

III LITERATURE SURVEY:

There are some computational techniques to efficiently represent, analyze and visualize both short term and long-term temporal changes in image sequences. It enhances temporal variation in videos that are imperceptible to the human naked eyes. The method uses a standard video sequence as an input, and applies spatial decomposition on the input video, followed by temporal filtering to the frames which is called "Eulerian Video Magnification" i.e. EVM. The resulting output signal is then amplified to reveal the hidden information.

There are four techniques which are used: 1. Linear Approximation Method [1], 2. Phase Based Video Processing [3], 3. Riesz Pyramid for Fast Phase Based Video Processing [4] 4. Enhanced Eulerian Video Magnification [6].

3.1 Linear Approximation Method [1]

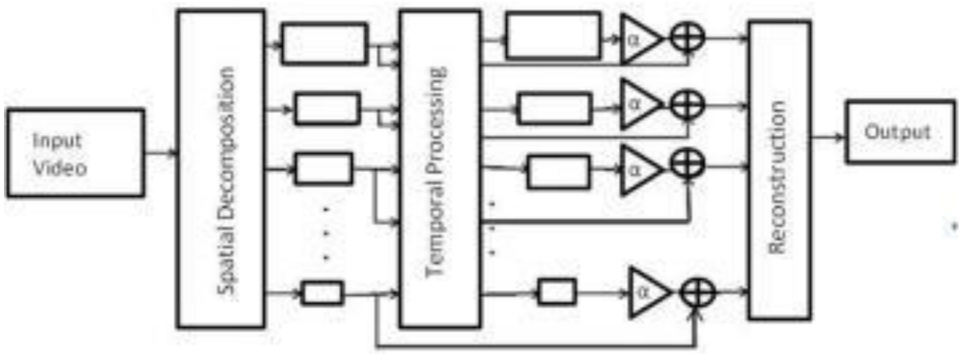


Figure 2: Linear Approximation Method Eulerian Video magnification (EVM) [1]

It uses standard video sequence as an input. Then it is decomposed into different frequency bands using Gaussian Pyramid of 3-4 levels to find out the band of interest. The video frame is passed through Low Pass Filter (LPF) and then down samples it to construct the Laplacian Pyramid [6]. Different frequency bands have different signal to noise ratio they may need to amplify with different amplification factor.Then depending on application temporal filtering is used on each spatial frequency band by performing Band Pass Filtering to extract the required frequency band. Here, the time series values at any spatial pixel in frequency band is considered, for example, to magnify baby's pulse rate we might select frequency band of 0.4Hz to 4Hz which corresponds to pulse rate of 24bps to 240 bps and applies the narrow band around that frequency band.

The extracted signals are amplified with some amplification factor (α). From [1] the limits of α is given below as,

$(1 + \alpha)\ \delta(t) < \lambda/8$

Where, δ(t)= spatial displacement function, Spatial wavelength (λ)= 2Π/ω, ω = spatial frequency. The motion magnification in temporal relies on the first-order Taylor series expansions.

Reconstruction of final output video is done by adding the magnified signal to original signal and disintegrating the spatial pyramid.

Limitations:

- In linear approximation method the noise gets amplified linearly with amplification factor.
- It can be used for small amplification factors.

3.2 Phase Based Video Processing Using Complex Steerable Pyramid [3]:

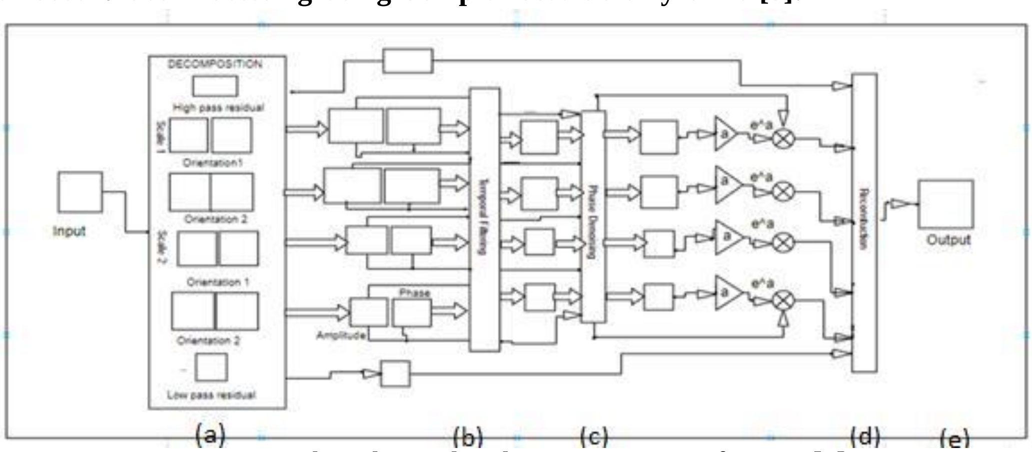


Figure 3: Phased Based Video Motion Magnification [3]

International Journal of Future Generation Communication and Networking
Vol. 13, No. 3s, (2020), pp. 1203–1211

As shown in figure 3, the different stages are,

a) Input video is decomposed and amplitude and phase is separated. b) Temporal filtering is applied at each location. c) SNR is improved by applying Phase de-noising. d) Amplification or attenuation of band pass phase is done. e) Output video reconstruction

The limitation of linear approximation is overcome by Phase based video motion processing. The Phase based video motion processing is another Eulerian approach.

- It supports larger amplification factor at all spatial frequencies.
- The Phase based video motion processing gives better noise performance than linear method.

The phase based motion processing is based on complex valued steerable pyramids [5] is inspired by phase based optical flow. The local motion in sub bands of spatial image corresponds to phase variation. The aim of this approach is to remove noise from each sub bands of image and to amplify small motions with less noise. To amplify the small motions the expansion of complex valued steerable pyramids to sub octave bandwidth pyramids is done.In order to increase spatial support of filter and motion magnification, it extends to sub octave bandwidth filters. An image is decomposed according to spatial scale, orientation and position.

3.3 Fast Phase Based Video Processing Using Riesz Pyramid [4]:

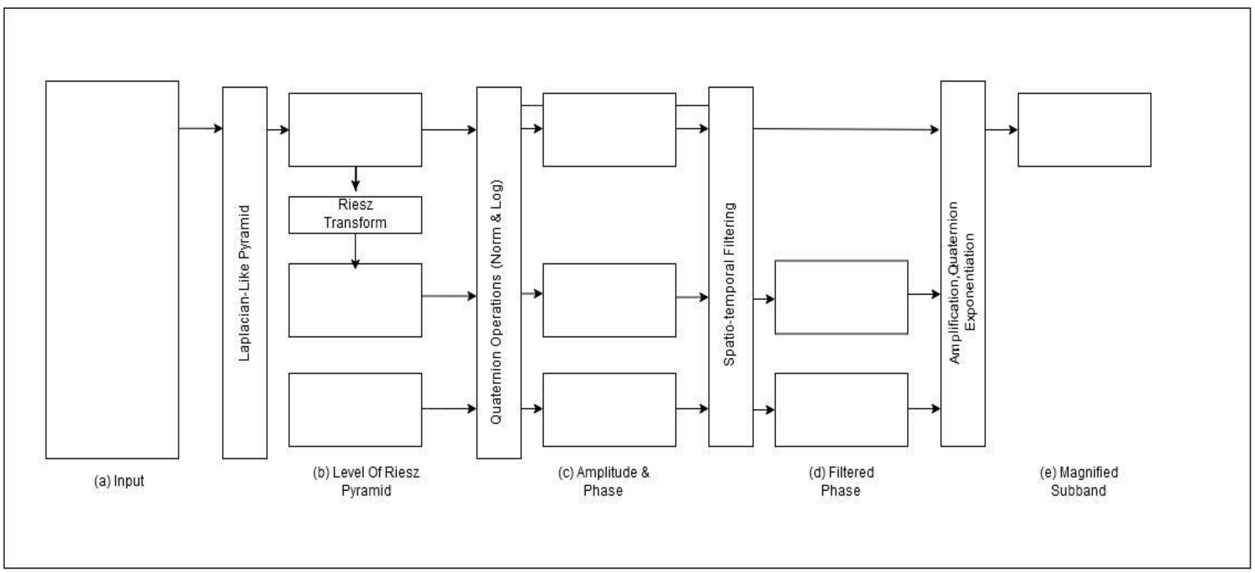


Figure 4: Riesz Pyramid for Fast Phase Based Video Processing [4]

The different stages involved in fast phase based video processing as shown in figure 4 are, (a) Input video (b) Input video is decomposition using a Laplacian-like pyramid (only one level is shown). The Riesz transform is used to produce the Riesz pyramid. (c)The quaternion norm is used to compute the amplitude (top row) and the quaternion logarithm is used to produce the quaternionic phase (bottom rows). (d) The quaternionic phase is filtered spatio-temporally and then to produce a motion magnified subband, same quantity is used to phase-shift the input Riesz pyramid level (e). These subbands are collapsed to produce a motion magnified video.

Riesz pyramid yields the fastest phase based method in spatial domain.

The results show that this method is four to five times faster than 8 orientations complex steerable pyramid [11, 14]. Also it is less expensive to implement than complex steerable pyramid.

Limitations:

ISSN: 2233-7857 IJFGCN

The Riesz pyramid transform method does not maintain the power of an input signal which can cause minor artifact.

3.4 Enhanced Eulerian Video Magnification (E2VM) [6]

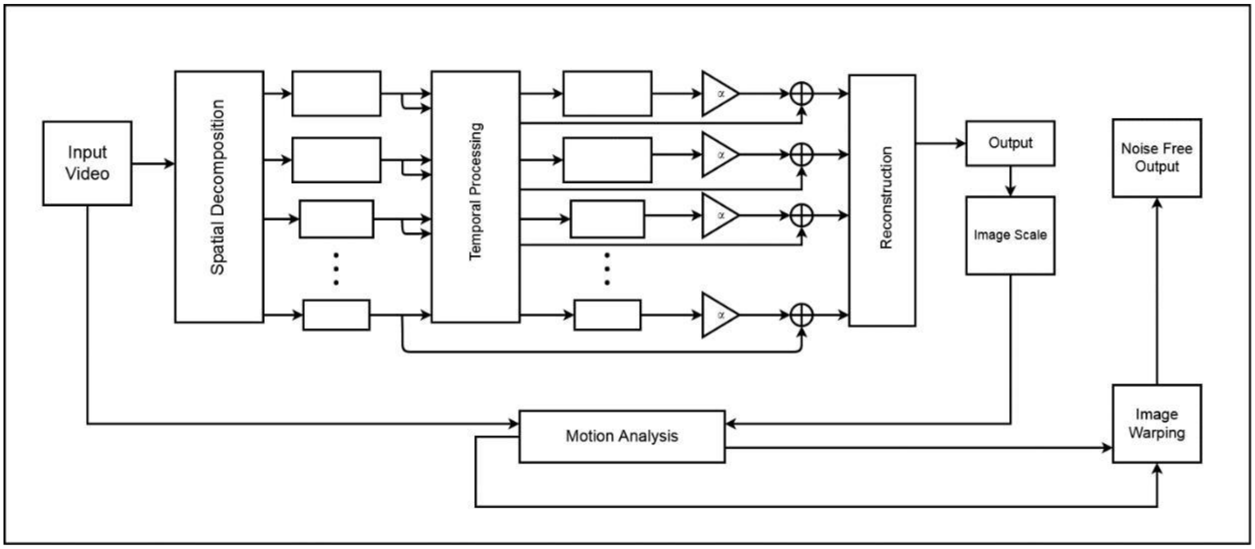


Figure 5: Enhanced Eulerian Video Magnification (E2VM) [6]

In this method, the difference between input video and EVM processing unit is taken to calculate pixel level motion mapping. The video motion is amplified by using image warping. In this, some frame pixels are warpped across motion mapping direction. The motion mapping is sub-sampled to boost the video processing speed to get a spares grid. This spares grid image warping of E2VM allocates a better way to handle noise of EVM with only small resolution image difference calculation.

Though it does not increase additional computation time but still has some advantages like support for larger amplification factors with fewer artefacts and less noise compared to EVM.

Using negative motion mapping, E2VM method can also be used to remove small motions of video for applications, such as motion de-noising [1], de-animating and video stabilization.

The EVM method uses, MATLAB code, Six core processor, 32 GB RAM, 640 X 480 videos resolution, Videos of 45 fps and Computation time of the order of 280s to 300s.

IV. COMPARATIVE STUDY:

//K;; Parameters	Methods			
	Phase Based Magnification	Fast Phase Based Magnification	Linear Approximation Method	E2VM
Video	Complex	Riesz	Laplacian	Laplacian

ISSN: 2233-7857 IJFGCN

Decomposition				
Amplification Factor (A) Limits	αδ(t) < λ n/4	αδ(t) < λ n/4	(1 + α)δ(t) <λ/8	(1 + α)δ(t) <λ/8
Noise Level	Minimized	Minimized	Magnified	Minimized
Suitable for	Sinusoid [3]	Sinusoid [3]	Linear ramp [1]	Linear ramp [1]

Main difference between methods is as shown in Table 1.

Table 1: Comparison of different methods

n - Number of filters per octave for each orientation.

K - Number of orientation bands

V. MOTION MAGNIFICATION IN EVM [1]

Let the image intensity at position x and time t be denoted as I(x,t).

In EVM, the intensities are observed with respect to a displacement function δ(t). Since the image undergoes translational motion, such that,

I (x, t) = f (x + δ(t)) and I (x, 0) =f (x). ------(1)

Assuming the image f (x + δ(t)) at time t can be approximated by using first order Taylor series expansion about x as,

$I(x, t) \approx f(x) + \delta(t)\ (\partial f(x)/\partial x)$ ----------(2)

Let $\mathbf{B}(\mathbf{x}, \boldsymbol{t})$- temporal band pass filter output Assuming the motion signal, δ(t), is within passband of temporal band pass filter. Then we have,

$B(x, t) = \delta(t)(\ \partial f\ (x)/\partial x)$----------- (3)

Now we amplify the band pass signal with some amplification factor α and add it back to original signal I (x, t) to obtain the processed signal as,

ĩ(x, t) = I (x, t) + αB(x, t) ------------------(4)

Combining Eqs. (2), (3), and (4), we have,

$\tilde{\imath}\ (x, t) \approx f(x) + (1 + \alpha)\ \delta(t)\ (\partial f\ (x)/\partial x)$-----(5)

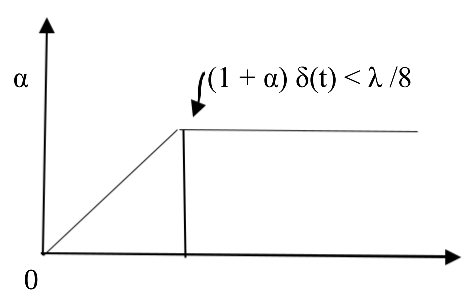
The first-order Taylor expansion holds the amplified larger perturbation, (1 +α)δ(t). Hence motion magnification can be related to the amplification of the temporally band passed signal.

ĩ (x, t) ≈ f (x + (1 + α)δ(t)) -------------(6)

Above equation shows that temporal processing magnifies motion-the spatial displacement $\delta(t)$ of local image f(x) at time t, has been amplified with magnitude of (1+ α).

(1 + α) δ(t) < λ /8 Eq. (7)

The above equation gives the largest motion amplification factor, α, compatible with accurate motion magnification of a given video motion δ(t) and image structure spatial wavelength, λ



λ

Figure 5: Amplification factor, α, as function of spatial wavelength λ**,** for amplifying motion.

The amplification factor is fixed to α for spatial bands that are within our derived bound (Eq.7), and is attenuated linearly for higher spatial frequencies. Above effect was also exploited in the earlier work of Freeman et al. [3] to create the illusion of motion from still images.

VI DISCUSSION:

Linear approximation method based on EVM, magnify the minute colour and invisible motions by using spatio-temporal process to enhance both invisible spatial signals and colour changes in the video. This technique is simple and fast but, it has some limitations. It linearly magnifies noise. Phase based video magnification which uses complex steerable pyramid, analyses the local phase over time at different scales and orientation for processing small motions and amplifies the temporal phase difference in corresponding band. Phase based method gives noise free results and high quality photo realistic videos with amplified motion. Riesz pyramid uses fast phase based video processing. The speed in phase based video magnification is achieved by this method without reduction in quality. The real-time phase based motion magnification is implemented by Riesz pyramid.
E2VM presents an efficient and noise free motion magnification method by image warping and spatio-temporal filtering.These methods used the database as shown in Table-2 to extract the useful information from the input videos.

SN	Filename	Time(s)	Width and Height	Size (MB)	FPS
1	Baby.mp4	10	360x544	1.75	30
2	Baby2. mp4	29	640x352	4.58	30
3	Face. mp4	10	528x592	1.56	30
4	Face2. wmv	9	432x192	1.14	29.97
5	Guitar. mp4	10	960x624	2.68	30
6	Shadow. mp4	6	640x352	1.90	30
7	Wrist. mp4	29	1032x392	28.4	29.97
8	Camera.wmv	33	1152x728	3.73	30
9	Subway.wmv	8	1288x368	11.9	29.97

Table-2 Database of video files

Eulerian Video Magnification (EVM) aids in the development solutions by allowing for the extraction of physiological signals from video data. Thermal video in conjunction with EVM can be used to extract physiological measures, particularly heart rate. An adaptive EVM approach ca be used to amplify the

ISSN: 2233-7857 IJFGCN

desired signal, while avoiding noise amplification. Results shows that physiological signals like heart rate can be revealed using thermal video, in conjunction with the adapted EVM method and ROI post-processing limiting the noise signal in the output. [18]
Eulerian Video Magnification (EVM) can be used in a multi-modal selective passband search approach, using passbands which are predefined, and the use of intelligent data fusion of the three different modalities provided by the Intel RealSense RGB-D camera. It can be shown that the effectiveness of using the color, depth, and near infrared streams to obtain a consensus heart rate estimate under various lighting conditions and subject poses. Results shows that the fusion of estimate of heart rate acquired from each modality is effective and robust to different environmental conditions. [19]

VII Conclusion:

The methods for observing the hidden information embedded in the minute invisible motions in a video or image sequence are reviewed in this paper. The method that takes a input video and amplifies subtle color changes and unseen variations is presented. Four techniques are reviewed 1. Linear Approximation Method (EVM), 2. Phase Based Video Processing 3. Fast Phase Based Video Processing (Riesz Pyramid) and 4. Enhanced Eulerian Video Magnification (EEVM). After review it is observed that, the noise level in EVM technique is more which is reduced in EEVM technique using motion analysis and image warping. The other two phase based techniques supports larger amplification factor and better noise performance. Using these techniques, it is possible to magnify and see tiny motions and subtle color variations
.

References:

[1] Wu H, Rubinstein M, Shih E, et al., “Eulerian Video Magnification for Revealing Subtle Changes in the World.” ACM Transactions on Graphics., 31(4): 65. 2012
[2] Liu, C., Torralba, A., Freeman, W. T., Durand, F., Adelson, E. H., “Motion magnification” ACM Trans. Graph. 24, 519 –526, 2005
[3] Wadhwa, N., Rubinstein, M., Durand, F.,Freeman, W., “Phase-based Video Motion Processing”, ACM Trans. Graph. 32, 4, Article 80 July 2013.
[4] N. Wadhwa, M. Rubinstein, F. Durand, and W. Freeman, “Riesz pyramid for fast phase-based video magnification” In Computational Photography (ICCP), 2014 IEEE International Conference on. IEEE, 2014.
[5] Simoncelli E. P., Freeman W. T, “The steerable pyramid: a flexible architecture for multi-scale derivative computation”, In Proceedings of the 1995 International Conference on Image Processing (Vol. 3), IEEE Computer Society, Washington, DC,USA, ICIP’95, 3444, 1995.
[6] Le Liu , Lu L, Jingjing Luo ,Jun Zhang, Xiuhong Chen., “Enhanced Eulerian video magnification.” 7th International Congress on Image and Signal Processing (CISP), oct.2014.
[7] Verkruysse, W., S vaasand, L. O., Nelson, J.S., “Remote plethysmographic imaging using ambient light”, Opt. Express 16, 26, 21434–21445, 2008.
[8] Poh. M. Z., M. Ccduff, D. J., Picard, R. W., “Non-contact, automated cardiac pulse measurements using video imaging and blind source separation.”, Opt. Express 18, 10,10762–10774, 2010.
[9] Fuchs, M., Chen, T., Wang, O., Raskar, R.,Seidel, H. P., Lensch, H. P., “Real-time temporal shaping of high speed video streams.”, Computers & Graphics 34, 5, 575–584, 2010.
[10]Shou-Ling Peng and Gkrard Medioni, “Spatio-temporal Analysis for Velocity Estimation of Contours in an Image Sequence with Occlusion.” Pattern Recognition, 1988., 9th International Conference on , Nov 1988 ,pp.236 - 241 vol.1.
[11] R. C. Bolles and H. H. Baker, “Epipolar-Plane Image Analysis: A Technique for Analyzing Motion Sequences”, Proceedings of the Third Workshop on Computer Vision: Representation and Control, October 1985, pp. 168-178.
[12] Wang, J., Drucker, S. M., Agrawala, M.,Cohen, M. F., “The cartoon animation filter”, ACM Trans. Graph.25, 1169–1173.

ISSN: 2233-7857 IJFGCN

[13] Horn, B., Schunck, B., "Determining optical flow", Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A. 17, 1-3, 185– 203.
[14] Burt P., Adelson E., "The Laplacian pyramid as a compact image code", IEEE Trans., Comm. 31, 4, 532–540, 1983.
[15] M. Felsberg and G. Sommer, "The monogenic signal", IEEE Trans. Signal Process. 49(12):3136–3144, 2001.
[16] Wadhwa, N., Rubinstein, M., Durand, F., Freeman W. T. "Quaternionic Representation of Riesz pyramid for fast phase-based video magnification",IEEEInternational Conference on In Computational Photography (ICCP), 2014.
[17] Lokesh Sarode , N.N.Mandaogade, "Video Motion Magnification using Spatio-Temporal Algorithm", International Journal of Computer Applications (0975 – 8887), Volume 96– No.9, June 2014
[18] Stephanie L. Bennett, Rafik Gaubrann Frank Knoefel, "Adaptive Eulerian Video Magnification Methods to Extract Heart Rate From Thermal Video", IEEE International Symposium on Medical Measurements and Applications (MeMeA), 15-18 May 2016, ISBN: 978-1-4673-9172-6, Benevento, Italy
[19] Yasmina Souley Dosso, Amente Bekele, James R. Green," Eulerian Magnification of Multi-Modal RGB-D Video for Heart Rate Estimation", International Symposium on Medical Measurements and Applications (MeMeA), eISBN: 978-1-5386-3392-2, 11-13 June 2018, Rome, Italy