مجلة

كلية التـــراث الجامعة

مجلة علمية محكمة

متعددة التخصصات نصف سنوية

العدد الثامن والثلاثون

15 شباط 2024

ISSN 2074-5621

رقم الايداع في دار الكتب والوثائق 719 لسنة 2011

مجلة كلية التراث الجامعة معترف بها من قبل وزارة التعليم العالي والبحث العلمي بكتابها المرقم (ب 3059/4) والمؤرخ في (4/7 /2014)

Survey on Video Magnification Technology to Improve Video Quality and Recovery Sound with micro-vibrations

Khalida Ali Ahmed

Assist prof. Dr.Abdul-Wahab Sami

Assist prof. Dr.Asmaa Sadiq

Mustansiriya University

Abstract: Eulerian Video Magnification (EVM) is a technique that enhances and enlarges fine details in videos. It analyzes the temporal signal of the video, magnifying small pixel-level vibrations. By employing image processing and frequency analysis techniques, EVM improves quality and enlarges fine details. Some studies have combined EVM with virtual microphone technology to analyze micro-vibrations caused by recorded sound. These vibrations are used for sound restoration, where they are converted into an audio signal using signal processing and frequency analysis techniques. The resulting sound is then restored in the video. In this work, a literature reviews have demonstrated promising results in terms of accuracy and quality in sound restoration.

1.1 1ntrodaction

Small changes that are invisible to the human eye are observed to have an impact on the movement of pixels from their current places to different ones when examining tiny movements in subsequent video frames. Utilizing these effects, valuable information can be extracted for a range of applications [1, 2].

There are two primary categories of motion information, as illustrated in Fig 1.1. The first deals with movements like breathing in humans, and the second kind deals with signals which are further split into two sorts. Video signals, such as an electrocardiogram (ECG or EKG) drawing, are related to the first subtype [2, 3]. The second subtype deals with audio signals, as the vibrations recorded by video of sound amplifiers [4]. The same process that turns such vibrations into sound could also be applied to telephone speakers.

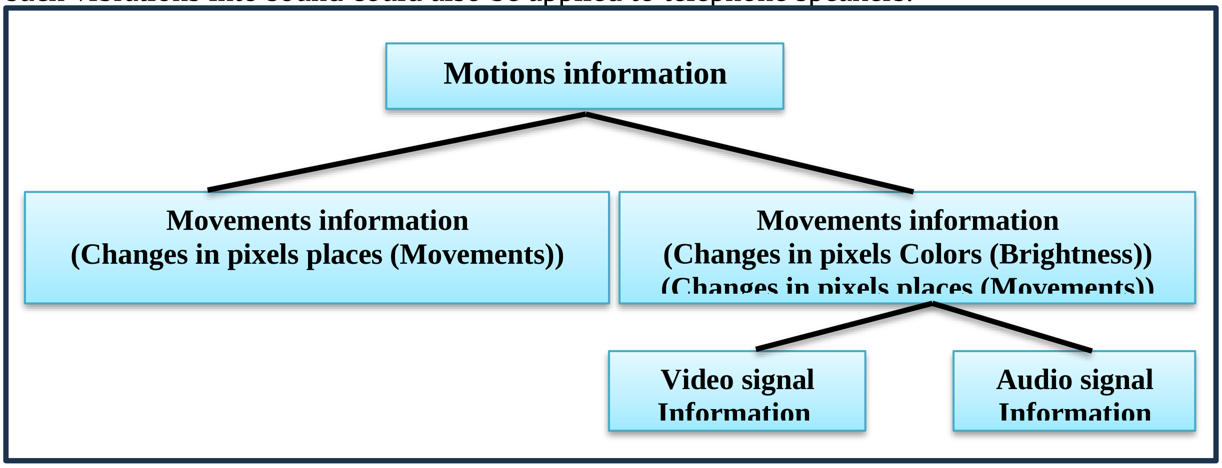

Figure (1.1): Types of information in small motions [2].

Audio signal information is the main topic of this work. When objects vibrate, the air becomes pressurized and rarefied, which propagates through the atmosphere and is perceived by the human ear as sound. While there are several ways for recording sounds, using a conventional microphone is the most popular. "Alternative Methods for Sound Extraction in Private Environments". In private environment in which sound recording is necessary for surveillance and security, substitute techniques were suggested to get around restrictions and guarantee efficient sound extraction. Using microphones with laser bases is one such technique [5]. Those microphones employ laser beams for detecting sound-induced vibrations in objects. It is possible to separate and record sound through examining such vibrations. Spot patterns are employed in another method for remote sound extraction [6]. This technique makes use of certain patterns that are projected onto surfaces for capturing and analyzing the vibrations brought on through sound waves. There is a possibility for extracting and recording the related sound through examining such vibrations. Sound could also be captured from high-speed images of an object through examining the vibrations that are produced when sound waves enter the air and strike the object. This method makes it possible to extract sound even in situations in which using regular microphones would not be practical or efficient [7].
These alternative methods provide viable solutions for sound extraction in private environments, enabling effective recording and surveillance capabilities. By leveraging laser-based microphones, spot patterns, or high-speed shots, sound can be extracted and recorded, enhancing security and surveillance measures. Visible Microphone (VM) technology recovers sound from silent videos by analyzing micro-vibrations of objects caused by fluctuations in air pressure generated by sound and can recover sound from objects in the video [8]. These tiny vibrations can generate a visible signal that can be detected and recorded by a high-speed camera. The vibrations can then be amplified using video processing techniques [9]. Extensive research by multiple scholars, including [10], [11], [12], [13], and [14], has contributed to the study and development of this technique.

1.2 Literature Survey

This related works is divided into two groups based on the approach used in processing videos with small vibrations. One group amplifies the subtle surface movements of the object in the video, while the other focuses on restoring the sound produced by the vibrations. The research literature is divided into these two categories, with different techniques utilized in each group:

1.2.1 Related works of small movement's magnification

Eulerian Video Magnification (EVM) approach has been employed in the majority of related work in this category [15]. In order to shorten processing times, the majority of the works aim to eliminate noise from EVM result videos. The following techniques will be examined in brief:
In 2015, Karim Al-Ghoul et al. [16] conducted a study on heart rate variability (HRV) for estimatin emotions using non-contact measurement techniques such as photoplethysmography (PPG) signal sampling of the human face through a camera. The study employed intracranial analysis, independent component analysis (ICA), or video magnification to induce subtle changes in skin color affected by PPG. The results indicate that the ICA-based method produces superior outcomes for high frequencies (HF) and low frequencies over high frequencies (LF/HF), while the AILER video amplification method performs better for low frequencies (LF). Moreover, it demonstrates a stronger correlation with absolute truth.

EVM method was created in the year 2016 by Stephanie L. Bennett et al. [17] for extracting physiological signals from thermal video. Heart rate has been extracted from medium-intensity signals as well as iPad video using EVM. With the use of broad-spectrum filter and signal processing that was enhanced for regions of interest (ROI) produced excellent results. To enhance the outcomes, a focused narrowband filter was additionally applied. The findings demonstrate the ability regarding thermal video to identify heart rates and lessen amplified noise when combined with EVM and ROI signal processing.

In the year 2017, thermal video and adaptive EVM processing were used by Stephen Bennett and colleagues [18] in an effort to identify temperature variations. In a system where known variables govern the system; this method could show heart rates. Using EVM processing, which separates the video into various frequency components, the technique analyzes thermal video data. It is possible to identify temperature variations brought on through pulsatile blood flow by concentrating on high-frequency components. To confirm the suggested approach's efficacy, the researchers ran tests on a control system. The technique's potential for non-contact heart rate monitoring is suggested by the results, which showed that it could efficiently detect temperature variations that corresponded to heart rates.

The idea of use edge-aware filtering for improving motion magnification in videos while lowering distortions and noise was investigated in 2018 by Manisha Verma and Shanmugaratnam Raman [19]. Fast local Laplacian filtering is used to improve the video, Eulerian motion augmentation is applied, and edge-aware filtering is used for maintaining subtle edges and lower noise. The motion magnification, video quality, and subtle change handling are all greatly enhanced by this suggested way.

A new framework was proposed in the year 2018 by Mithun Kumar and colleagues [20] with the goal of improving undetectable movements in videos. The framework utilizes spatial coordinates that, even with the existence of significant background motions or a steady background, could amplify target motions. A histogram of oriented gradients (HoG) with thresholding is used for defining ROI, and distinguishable points are tracked throughout time and their velocities are enhanced through a zoom factor. The framework also provides an autonomous spiral search mechanism depending on blocks to find investment returns. The efficacy of the suggested approach in magnifying extremely small movements is demonstrated by the experiment results.

In 2019 Takeda et al. [21] developed video amplification techniques to identify minute changes that are invisible to the unaided eye. Unfortunately, because imaging noise introduces irrelevant changes, these methods frequently yield noisy and incorrect results. Without requiring resources or human interaction, the suggested strategy makes use of fractional anisotropy (FA) to amplify only significant changes. In comparison to other methods, the method gets outstanding zoom results by evaluating the temporal distribution of these changes. The outcomes show how well the suggested strategy works to increase the usable video zoom range and create high-quality zoomed videos.

In 2020, Sanket Yadav et al. [22] studied video zooming methods, including phase-based Gaussian pyramid, Euler pyramid, and Laplace pyramid. Also, they used contrast parameters such as PSNR, zoom factor, execution time, color zoom, and motion to evaluate the performance of these details. Large found that while the gradient path technique is fast and sophisticated but ineffective at large magnifications, the Euler technique is fast and simple but ineffective at small magnifications. They study concluded that to enhance the effectiveness of video amplification techniques, which are not often used in fields such as Nowruz medicine, a new technology using a versatile boat should be developed.

In 2021, Rashmi Ranjan Das and colleagues [23] suggested utilizing Euler video augmentation (EVM) to identify fraudulent movies by emphasizing inconspicuous characteristics like skin pulsation and minute motions. In order to distinguish between real and false movies, the authors trained three models with features taken from Euler. Based on comparison to other approaches, it may be inferred that the proposed work proved the usefulness of their suggested technique, which shows that it can recognize phony videos with accuracy.

1.2.2 Related works of Sound Recovery Using Vibration

In this section, the micro-vibrations that appear on the object as a result of the sound from the virtual microphone in the videos have been used as a case study to restore the sound produced by those vibrations that occur on the surface of the object, and therefore it will be presented to related works dealing with micro-vibrations in clips of the video.

In 2014, Chaoyang Wang et al. [24] conducted a study demonstrating the extraction of audio information from high-speed silent video using simple and fast image processing techniques. The technology operates on the principle that sound waves cause objects to vibrate upon contact, which can be detected through image-matching processes. The technology utilizes a subset-based image-matching method to detect point movement on the object surface and the Gauss-Newton algorithm for fast and accurate image-matching. By reconstructing the original acoustic information of sound waves through the detected vibrations, this technique has proven to be highly effective and easy to implement, as confirmed by experimental verification of its robustness.

The goal of Andrew Owens and colleagues' 2016 study [25] was to predict the sound that will be produced when two objects meet. This prediction could shed light on the physical interactions in a visual scene. They demonstrated an algorithm which collects audio from silent video of individuals using drumsticks to beat and scratch objects. The system predicts audio features from videos with the use of a recurrent neural network, and after that it utilizes an example-based synthesis technique to create a waveform of such features. The researchers indicate that they were capable of tricking participants in a "fake or real" psychophysical experiment because the noises predicted through their model have been convincing. They also transmit crucial information regarding the characteristics of materials and their physical interactions.

In 2016, Johyun Ahn and his colleagues [26] introduced Visual Microphone, a technique for recovering audio from a silent video by extracting fine motion signals from the entire image. However, using the entire image for sound restoration may introduce noise and make it unclear which part of the object is exposed to the sound wave. Their proposed patch-based framework addresses these issues by recovering sound from a sub-region (patch) in the image centered at a key point (corner). Speeches are recovered from patches centered at each key point, and then the best speech with the least noise is selected. Results show that using motion signals from a small region near a key point can improve the quality of recovered speech.

In 2017, Dashan Zhang et al. [27] utilized high-speed cameras for non-destructive testing and monitoring of offline structures. Their approach involved using a phase-based technique to extract sound-induced vibrations from video clips. Then they employed an effective Singular

Value Decomposition (SVD)-based method to detect subtle movements of pixel density in silent, high-speed videos caused by sound. The researchers validated their proposed method through simulation tests and experiments, demonstrating its potential applications in sound restoration and material analysis.
2017 research on audio recovery from videos with partial body movement was carried out by Yusuke Yasumi et al. [28]. In the case when there are noticeable movements of the body in the video due to camera shaking or wind, it could be difficult to adjust the sound. A new approach depending on immediate evolution modifications was created to address this problem. Using audio data gathered from developmental changes in the video, this approach analyzes the child's body movement. In order to get good results and precise audio tracking even when there are significant movements in the video, the researchers ran tests for recovering audio from particular visual selections.
A novel technique for polyphonic musical note tracking (NT) of string instruments with the use of video rather than audio data was suggested by Shir Goldstein et al. [29] in 2018. A camera attached on the instrument records silent video, from which the vibrations regarding each string are recovered. 1D signals collected from the video are used for computing NT of each string. The limitations of audio-based polyphonic NT could be addressed through taking into account the predicted frequencies regarding the played notes, their harmonics, and their aliases. The technique generates sheet music which could help musicians perform visually captured music and is tested on real data. The research comes to the conclusion that such visual-based NT approach could be very helpful in resolving the NT problem.

Yohei Fuse et al.'s research from 2018 [30] suggests a technique for utilizing visual objects for extracting sounds from videos. This method works particularly well for surveillance applications which require long-distance sound recording. In order to observe the vibrations brought on by the sound, the procedure involves utilizing high-speed video as well as animating each one of the pixels with a steerable pyramid that starts up. The suggested approach uses a body log-based filter for examining the vibrating component regarding each sound frequency for recovering sound and lessen the requirement for body vibration devices. The videos show off the efficiency of the approach through a series of tests.
Yohei Fuse et al.'s research from 2018 [31] suggests a method for utilizing visual elements for extracting sounds from videos. This technique is very useful for surveillance applications when sound recording over long distances is required. In order to observe the vibrations brought on by the sound, the method uses high-speed video and animate each pixel with a steerable pyramid that is started up. In order to reduce the dependence on body vibration devices, a body log-based filter is used to analyze the vibrating component regarding each sound frequency in order to recover sound. A number of tests carried out in the videos show how effective the procedure is.
An efficient technique for obtaining sound information from speckle patterns recorded through a high-speed camera was presented by Shir Goldstein et al. in 2018. To produce music with a high signal-to-noise ratio, the process entails selecting pixels with wide variances in gray-value changes over time and summing their values. Compared to conventional digital image correlation methods, it is easier to use, quicker, and less sensitive to speckle pattern

quality. The suggested technique successfully recovered a sound signal when tested on a variety of objects.
The viability of recovering sound from commonly utilized Closed-Circuit Television (CCTV) compression techniques that dramatically reduce the size of high-speed videos was investigated in 2018 research by S.Y. [33]. According to the research, high-quality sounds may be produced from compressed videos, suggesting that even in the absence of integrated audio input, CCTVs may be able to provide both visual and auditory capabilities for enhanced surveillance systems.
In 2018, Mohamed Amin Shabani and colleagues [34] conducted a study on extracting audio from video by analyzing local vibration patterns in various parts of the image. The study revealed that different image areas exhibit distinct vibrations in response to sound waves and proposed a method for capturing these local vibrations to enhance sound quality. Additionally, the research pinpointed a time delay in local vibrations due to the travel time of sound waves and leveraged this delay to estimate the sound's direction. They also introduced a new algorithm that speeds up the audio extraction process and achieves real-time performance in 20 kHz video.
Atsushi Yoshida et al. [35] developed a method in the year 2020 for obtaining sound from video through employing a rolling-shutter camera to record an object's surface vibrations. However, this camera has limitations that lead to gaps in the extracted sound and the inclusion of unwanted noise. To overcome this issue, the researchers suggest utilizing interpolation and noise suppression techniques to enhance the sound quality. Experimental findings indicate that the proposed approach improved the PESQ score by an average of 0.3.
In 2020, Choong, Yee Kai Teed, and others [36] conducted a study on the visual microphone technique. This technique analyzes vibrations in a video to recover sound. The study aimed to examine the impact of denouncing the video frames before sound recovery. The quality of the recovered sound was evaluated based on human intelligibility metrics and simple signal-to-noise ratio metrics. The results indicate Denoising is especially crucial for enhancing the quality of sound in noisy recorded videos.
In 2020, Ben Nassi and his team [37] introduced the Lamphone attack, which is a new technique for intercepting sound through a side channel. This method involves using an electro-optical sensor to examine the response of a suspended light bulb to sound waves, taking advantage of the bulb's slight vibrations resulting from air pressure changes. The attack can obtain speech and sing passively, externally, and in real-time. The researchers created an algorithm to separate the audio signal from the optical measurements, and they proved that Lamphone can effectively retrieve human speech and singing from a distance of 25 meters.
In 2022 [38], Ren Jun et al. conducted a study to find out the effect of changing video resolution on audio recovery using a visible microphone. The study examined the effects of upscaling, downscaling, and combining both techniques on the quality of sound recovered using neural networks. The results indicate that adjusting the video resolution does not significantly improve and may even reduce the quality of the restored audio. Therefore, it is advisable to refrain from changing the video resolution before restoring the audio using a visible microphone. Finally, they suggested the intelligent and neural algorithms may allow us to extract audio from silent videos while maintaining video quality.

Table 1.1 shows the technology used in previous studies in an attempt to clarify the most commonly used techniques for amplifying the subtle surface movements of the object captured in the video and the accuracy of the results of research dealing with restoring the sound resulting from these vibrations.

Eulerian Video Magnification (EVM) method			
Ref. No	Authers	Techniques	Feature Extraction type
[16]	Karim Al-Ghoul et al.2015	ICA (Independent Component Analysis), EVM (Eulerian Video Magnification).	Extracting the photoplethysmography (PPG) signal
[17]	Stephanie L. Bennett al.2016	Eulerian Video Magnification (EVM)	Extracting heart rate from thermal video.
[18]	Stephen Bennett and colleagues 2017	Eulerian Video Magnification (EVM)	Thermal video data analysis
[19]	Shanmugaratnam Raman and Manisha Verma 2018	edge-aware filtering.	To amplify the impact of minor adjustments
[20]	Mithun Kumar and colleagues 2018	Spatial Coordinates	To amplify the motion.
[21]	Takeda et al. 2019	the fractional anisotropy	Detecting subtle yet significant changes.
[22]	Sanket Yadav et al. 2020	Eulerian Video Magnification (EVM)	Video analysis for extracting subtle changes
[23]	Rashmi Ranjan Das and colleagues 2021	Eulerian Video Magnification (EVM)	Spatial-temporal video analysis
Sound Recovery Using Vibration			
Ref. No	Authers	Techniques	Feature Extraction type
[24]	Chaoyang Wang et al. 2014	Optical Technology.	Extracting audio information

[25]	Andrew Owens and his colleagues 2015	Recurrent Neural Network.	The generation of sound filaments into a sound wave.
[26]	Johyun Ahn and his colleagues 2016	Visual Microphone	Sound is extracted from each patch by converting motor signals into audio signals
[27]	Dashan Zhang et al. 2017	Singular Value Decomposition (SVD).	Extracting sound-induced vibrations from silent videos.
[28]	Yusuke Yasumi et al.2017	Momentary phase variations	Extract sounds from subtle movements in the video.
[29]	Shir Goldstein et al. 2018	Optical detection to extract music information	Extracting string vibrations from silent video and calculating the note tracking for each string separately
[30]	Yohei Fuse et al.2018	complex steerable pyramid	Extracting audio from vibrations captured in video footage.
[31]	Yohei Fuse et al.2018	Complex Steerable Pyramid.	To extract sounds from videos
[32]	Shir Goldstein et al.2018	Variance-based Sound Recovery "VSR".	To extract audio data from dispersed point patterns
[33]	S.Y. et al. 2018	Sound Recovery from CCTV Video Compression Techniques.	Audio recovery from video compression techniques
[34]	Mohamed Amin Shabani and colleagues 2018	Local Visual Microphones	Extracting audio from video and analyzing local vibration patterns.
[35]	Choong, Yee Kai Teed, and others 2020	Visual Microphone	Reduce noise before analyzing vibrations and recovering audio from video
[36]	Ben Nassi and his team 2020	Remote photoelectric sensor	Extracting sound from light bulb vibrations with a remote photoelectric sensor
[37]	Atsushi Yoshida et al.2021	Sound quality improvement through Interpolation	Extract audio signals from video by imaging vibrations caused by sound waves

[38]	Ren Jun et al. 2022	Visual Microphone	Study of video resolution change during audio extraction from body vibrations in video using an optical microphone

1.7 Conclusion

In this research, the amplification method was demonstrated and applied to video clips using the Eulerian video amplification (EVM) technique. The study showcased the effectiveness of this method in addressing gaps in the academic literature and produced favorable results across various metrics. Amplification was also utilized in conjunction with Virtual Microphone (VM) technology, as well as in related research. However, there is a lack of literature on analyzing micro-vibrations on the surface of an object using a visible microphone and restoring audio from these vibrations while minimizing noise interference on individual video frames before performing audio restoration using a virtual microphone. The results of the study were supported by several video and audio processing techniques and can be applied across a range of fields, including security, surveillance, medical, engineering, and physics.

1.8 References

[1] Loay.K.Abood and Ahmed.A.Shkara, "Optimal Path Routing," in M.Sc. Thesis, Department of Computer Science, University of Baghdad, Baghdad- Iraq, 2012.

[2] Ahmed A. Shkara, "Magnification of the Small Video Motions," in D.Sc. Thesis, Department of Computer Science, University of Baghdad, Baghdad- Iraq, 2018.

[3] Manab Kumar, "Electrocardiogram Signal Analysis for Heartbeat Pattern Classification," in M.Sc. Thesis, Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela- INDIA, May, 2015.

[4] Nirupam Roy, Romit Roy Choudhury, "Listening through a Vibration Motor," the 14th Annual International Conference on Mobile Systems, Applications, and Services, pp. Pages 57-69, June 26 - 30, 2016.

[5] Gergely Takács, Jakub Otčenáš, Ján Vachálek and Boris Rohal'-Ilkiv, "Modal response-based technical countersurveillance measure agai nst laser microphones," Journal of Vibroengineering, vol. 18, no. 5, p. 3369–3382, 2016.

[6] Zeev Zalevsky, Yevgeny Beiderman, Israel Margalit, Shimshon Gingold, Mina Teicher, Vicente Mico, and Javier Garcia"Simultaneous remote extraction of multiple speech sources and heart beats from secondary speckles pattern," Optics express, [e-journal], vol. 17, no. 24, p. 21566–21580., 2009.

[7] ZiyiChen, ChengWang, Chaohong Huang, Hongyan Fu, Huan Luo and Hanyun Wang,"Audio signal reconstruction based on adaptively selected seed points from laser speckle images, " Audio signal reconstruction based on adaptively selected seed points from laser speckle images, " Optics Communications,p. 6-13,2014.

[8] Abe Davis, Michael Rubinstein,1 Neal Wadhwa,Gautham J. Mysore, Fredo Durand and William T. Freeman," "The visual microphone: Passive recovery of sound from video," Association for Computing Machinery, vol. 33, no. 4, p. 1–10, 27 July 2014.

[9] Hao-Yu Wu, Michael Rubinstein,Eugene Shih, John Guttag, Fredo Durand and William Freeman , "Eulerian video magnification for revealing subtle changes in the world," Association for Computing Machinery, vol. 31, no. 4, pp. 1-8, 1 July 2012.
[10] Neal Wadhwa, Michael Rubinstein,Fredo Durand ´and William T. Freeman1, "Quaternionic representation of the riesz pyramid for video magnification," International , 2014.
[11] Michael Rubinstein, "Analysis and visualization of temporal variations in video: Massachusetts Institute of Technology," in Thesis: Ph. D., Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2014.
[12] Khalida Ali Ahmed, "Fingerprint Identification Using WAVELET Transform," Journal of College of Education, Volume , Issue 3, Pages 56-70, 2009.
[13] Khalida Ali Ahmed, "Modeling combinational circuits using neural network," Journal of College of Education, Volume , Issue 1, Pages 457-468, 2020.
[14] Kalda A. Ahmed, "Aerial images Segmentation by based on the segmentation methods fusion", Journal of College of Education, Pages 17-25, 2022.
[15] Hao-Yu Wu, "Eulerian Video Processing and Medical Applications," in M.Sc. Thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, June 2012.
[16] KARIM ALGHOUL, SAEED ALHARTHI, HUSSEIN AL OSMAN, AND ABDULMOTALEB EL SADDIK,"Heart Rate Variability Extraction from Videos Signals: ICA vs. EVM Comparison," IEEE Access, vol. 5, pp. 4711 - 4719, 20 March 2017.
[17] Stephanie L. Bennett, Rafik Goubran and Frank Knoefel, "Adaptive Eulerian Video Magnification Methods to Extract Heart Rate from Thermal Video," IEEE International Symposium on Medical Measurements and Applications (MeMeA), 08 August 2016.
[18] Stephanie Bennett; Tarek Nasser El Harake; Rafik Goubran; Frank Knoefel, "Adaptive Eulerian Video Processing of Thermal Video: An Experimental Analysis," IEEE Transactions on Instrumentation and Measurement (Volume: 66, Issue: 10, October 2017, vol. 66, no. 1, 07 April 2017.
[19] Manisha Verma & Shanmuganathan Raman, "Edge-Aware Spatial Filtering-Based Motion Magnification," Springer Nature Singapore Pte Ltd, 2018.
[20] Mithoon Kumar, Tilendra Choudhary and M. K. Bhuyan , " Small Motion Magnification using Automated RoI Selection and Spatial Co-ordinate Approach," International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), March 2018.
[21] Shoichiro Takeda, Yasunori Akagi, Kazuki Okami, Megumi Isogai and Hideaki Kimata, "Video Magnification in the Wild Using Fractional Anisotropy in Temporal Distribution," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
[22] Sanket Yadav,Prajakta Bhalkare,Sayali Shingde and Usha Verma, "Performance Analysis of Video Magnification Methods," Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020.
[23] Rashmiranjan Das, Gaurav Negi, Alan F. Smeaton, , "Detecting Deepfake Videos Using Euler Video Magnification," in s Masters in Data Analytics from, 2021.
[24] Zhaoyang Wang, Hieu Nguyen, Jason Quisberth," Audio extraction from silent highspeed video using an optical technique", Society of Photo-Optical Instrumentation Engineers (SPIE), November 2014.

[25] Andrew Owens, Phillip Isola, Josh McDermott, Antonio Torralba, Edward H. Adelson, William T. Freeman," Visually Indicated Sounds", Conference on Computer Vision and Pattern Recognition (CVPR) 2016 IEEE.
[26] Juhyun Ahn, Yong-Joong Kim,Daijin Kim," Patch-Based Visual Microphone for Improving Quality of Sound", 23rd International Conference on Pattern Recognition (ICPR), 2016.
[27] Dashan Zhang, Jie Guo Yi Jin ,Chang’an Zhu," Efficient subtle motion detection from high-speed video for sound recovery and vibration analysis using singular value decomposition-based approach", Opt. Eng. 56(9), 094105 (2017).
[28] Yusuke Yasumi, Tetsuya Takiguchi, Yasuo Ariki," Visual Sound Recovery Using Momentary Phase Variations", Physics, 2017.
[29] Shir Goldstein, Yael Moses," Guitar Music Transcription from Silent Video", Computer Science, Published in BMVC 2018.
[30] Yohei Fuse, Yusuke Yasumi and Tetsuya Takiguchi," Sound Recovery Using Vibration Modes of the Object in a Video ", 12-15 November 2018.
[31] Yohei Fuse, Yusuke Yasumi and Tetsuya Takiguchi," Sound Recovery Using Vibration Modes of the Object in a Video ", 12-15 November 2018, Hawaii.
[32] Ge Zhu, Xu-Ri Yao, Peng Qiu, Waqas Mahmood, Wen-Kai Yu, Zhi-Bin Sun, Guang-Jie Zhai, Qing Zhao, "Sound recovery via intensity variations of speckle pattern pixels selected with variance-based method," Optical Engineering, vol. 57, no. 2, 2018.
[33] S. Y. Lee, W. S. Yap, Y. C. Hum, B. H. Kwan, B. M. Goi, Y. K. Tee, Z. Jin, "Study the Effect of Commonly Used Video Compression Techniques on Sound Recovery via Negligible Object Vibrations for Visual Surveillance System," he 2nd International Conference on Advances in Image, p. 111–115, 2018.
[34] Mohammad Amin Shabani, Laleh Samadfam," Local Visual Microphones: Improved Sound Extraction from Silent Video", Computer Vision and Pattern Recognition, Submitted on 29 Jan 2018.
[35] Ren-Jun Choong, Wun-She Yap, Yan Chai Hum and Yee Kai Teed." Improving the Quality of Sound Recovered Using the Visual Microphone with Frame-wise Image Denoising Preprocessing, "journal citation and DOI. Published under licence by IOP Publishing Ltd,2020.
[36] Ben Nassi, Yaron Pirutin, amd B. Zadov," Lamphone: Real-Time Passive Sound Recovery from Light Bulb Vibrations," Published in IACR Cryptology ePrint ,Physics, 2020.
[37] Atsushi Yoshida, Hiroki Shindo, Koichi Terano, Takahiro Fukumori,Takanobu Nishiura,"Interpolation of Acoustic Signals in Sound Capture with Rolling- shuttered Visual Camera," Forum Acusticum, pp. 39-45, Submitted on 27 May 2021.
[38] REN-JUN, WUN-SHE, YAN CHAI, YEE KAI, " A Study on the Effect of Video Resolution on the Quality of Sound Recovered using the Visual Microphone," Published 18 March 2022, Physics, Proceedings of the 2022 4th International Conference on Image, Video and Signal Processing