Mansoura Journal of Computers and Information Sciences

Wavelet-based Video Enhancement Using Contrast Limited Adaptive Histogram Equalization

Fatma Faek
Computer science department
Mansoura University, Egypt
tomafayk@gmail.com

Omaima Nomir
Computer science department
Mansoura University, Egypt
omnomir@yahoo.com

Ehab Essa
Computer science department
Mansoura University, Egypt
ehab.essa@gmail.com

Ebrahim El-henawy
Computer science department
Zagazig University, Egypt
henawy2000@yahoo.com

ABSTRACT

A combination of wavelet transforms and contrast limited adaptive histogram equalization (CLAHE) techniques are used to efficiently enhance videos. The proposed technique handles the noises within video frames and enhances the resolution of the video. Lifting wavelet transform (LWT) and stationary wavelet transform (SWT) are applied to separate original frame into the low-frequency sub-bands, and the high-frequency sub-bands. Thereafter, we applied the interpolation to correct the coefficients of the high-frequency and the original frame separately. Next, Inverse Lifting wavelet transform (ILWT) is utilized for the integration of each all these enhanced sub-band. Finally, the CLAHE algorithm is applied to make the details of the frame more visible, and meaningful to generally improve the resolution of the video. The output video shows that the proposed technique enhances the quality of resolution videos under various environmental conditions, alleviates noises and avoids the over-enhancement problems.

Keywords
Video enhancement, Wavelet transforms, Interpolation, histogram equalization, CLAHE.

1. Introduction:

A digital video has become a vital part in our everyday life. It is familiar that video enhancement is considered as an active subject in computer vision. Over the last years, video enhancement has established much consideration. Ninety percent of all data volume on the internet is associated with a video or an image to enhance it [1-3]. Video enhancement is applied in many applications such as entrainment, healthcare, communication, medical imaging and surveillance [1-4].

The need for an effective video enhancement system is demanded to deal with different environmental conditions. Most cameras performance was not persuaded with light circumstances. High lights (such as front lights of cars) leads to detail loss and low signal-to-noise image that restricts the information amount transfer to user interface. Although low light produces video images with noise, in typical cameras, the electronics are simple to match that, so it is now viable for digital enhancing considerations of the video before introducing the video to user interface, accordingly it increase the information production [4, 5].
There are many problems we might face when trying to enhance low-quality video including [1-5]:

a) Low contrast: Moving objects could not be clearly extracted either from the dark background or when the colors of the extracted moving objects are in similar color with the background.
b) The ratio of signal to noise is generally very low because of high ISO (number of sensors light sensitivity for particular camera). Utilizing high ISO yields digital images with visible noise.
c) Maintenance should be performed for inter-frame coherence (e.g.in successive images, the region of moving objects could be modified smoothly.
d) Any area of the low-quality image is very important, even if it is one pixel.

While capturing video there are increasing defects, such as blocking, blur, contrast distortions, and noises. In addition, there are many problems that are difficult to control, such as the method of video capture is not professional, the unwanted environmental lightening, and the disadvantages of devices used etc. Therefore, proposing new techniques to enhance the video to improve the qualities of visual and video are needed specially with developing a novel enhancement technique for contrast that has the ability to avoid restrain noise and enhancement efficiently. Preserving edges is very important so as to magnify the resolved super image quality.

An efficient technique at low contrast that deals with various types of videos is used to enhance the visual appearance of those videos and improve the automated video processing transform representation like; segmentation, recognition, detection, and analysis of video image.

The aim of the paper is to propose a new resolution enhancement.
In this paper, two kinds of wavelet transform are used: LWT and SWT. The LWT is applying the wavelet transform by utilizing instinctive sequences of wavelet for functional and numerical analysis. LWT decomposes any given frame in four sub-bands known as high-high (HH), high-low (HL), low-high (LH), and low-low (LL). LWT has some advantages such as it holds both location information and frequency.SWT and LWT are symmetric [5, 6] but SWT does not cause a change in the frame size, so it is applied in the second step.

The interpolation is a key step to improve the results from LWT and SWT and consequently enhancing the resolution of the video. Bi-cubic interpolation is used to calculate pixel value between 4x4 neighborhoods of data points.
Here, the high-frequency sub-bands from LWT and SWT are given as the input to be interpolated by bi-cubic interpolation. At the same time, the input frame was interpolated separately to produce a smoothed version. The interpolated original

frame and modified interpolated high-frequency sub-bands are merged by using ILWT.

Finally, the output produced by the inverse wavelet transform enhanced by CLAHE. CLAHE is the most common technique in image/video processing. So, it's very effective for dealing with interesting parts and makes it seem more visible [5]. The frame is decomposed into disjoint areas, and in each area local histogram equalization is applied. Therefore, this will eliminate the boundaries between the areas by using a bilinear interpolation.

The rest of this paper is organized as follows: Sections 2 will discuss the related works. The proposed method is introduced in Section 3. Experimental results and discussions are presented in Section 4. Finally, Section 5 concludes this paper.

2. Related Works:

For the video enhancement problem, several techniques were proposed to manipulate different problems. Techniques of super resolution were introduced for video sequences as well as images. In this section, we will discuss the main related image enhancement techniques.

Lidonget al. [5] suggested a method for image enhancement by combining discrete wavelet transform (DWT) with CLAHE. The image was divided as two components: low-frequency and high-frequency by DWT then, applied the mean of weighted to the original and the reconstructed images. The experimentation results showed that the technique alleviated the over-enhancement problem, reduced noises in the image and gave high PSNR. However, the technique gave high mean square error value, and the running time was long.

Anbarjafariet al. [7] presented a video enhancement method by employing an illumination compensation technique as pre-processing and using SWT and DWT to extract the high-frequency bands of the images and applying the Vandewalle technique. Then inverse discrete wavelet transform (IDWT) was utilized to merge three enhanced sub-bands. However, the method had shortcomings like giving high mean square error value. The rate-distortion performance was low. This technique did not preserve the edges perfectly, and the running time was long.

Demirel and Anbarjafari[8] introduced a method employed dual-tree complex wavelet transform (DT-CWT) to separate the original image into various sub-bands. Then, bi-cubic interpolation was utilized to the six complex valued high-frequency sub-bands images. Also, the original image was inserted in the same time. Lastly, inverse dual-tree complex wavelet transforms (IDT-CWT) combined all of these sub-bands. The result was sharper and super-enhanced satellite images and gavethe high PSNR. However, the mean square error was high and the rate–distortion performance waslow. Vandewalle et al. [9] proposed super-resolution method using a series of low-resolution frames of the video for high-resolution frames reconstruction. This method registered a group of low-frequency aliased images; therefore reconstruct the high-resolution image by using cubic interpolation. The method showed better results but gave high mean square error value. Keren et al. [10] proposed a method to register the sequence of images captured from the camera with sub-pixel precision used in translation and rotation. The sub-pixel registration result improved the resolution and reduced the noise. The result was sharper and super-enhanced satellite images and gave the high PSNR. However, it gave high mean square error value. Wang et al. [11] compared the image enhancement based on LWT and the tradition DWT through theory and experiment. The LWT produced a better enhancement effect but the noise was increased and it gave high mean square error value.

The aforementioned techniques are good executed on some frames however; there are some of problems such as noise amplification and overstretching. Our goal of this research is to introduce a novel technique for enhancement, reduce noise and avoid over enhancement efficiently.

In this paper, we propose an efficient technique to deal with various types of videos to enhance the visual appearance of those videos and improve the transform representation for automated video processing such as video image analysis, detection, recognition, and segmentation.

3. The Proposed Method:

In this part, the main steps of the proposed technique are presented. We convert the original video into a set of frames. LWT and SWT are utilized to decompose the frames. LWT decomposes the frame into high-frequency and low-frequency sub-bands. SWT is applied to keep the size of the frame without any change. Next, bi-cubic interpolation is used to interpolate the three high-frequency sub-bands that obtained from wavelet transform. The original frames are also interpolated separately. The following step is to apply ILWT to merge the interpolated high-frequency sub-bands with interpolated original frames to obtain a high-resolution output frame. Finally, the CLAHE algorithm is applied to enhance the behavior of the frame generally. The output-enhanced frames are again converted back to video format. The proposed method steps are shown in Figure 1. The proposed method is evaluated using many video quality estimations such as structural similarity index measure (SSIM), mean squared error (MSE), peak signal-to-noise ratio, and average running time.

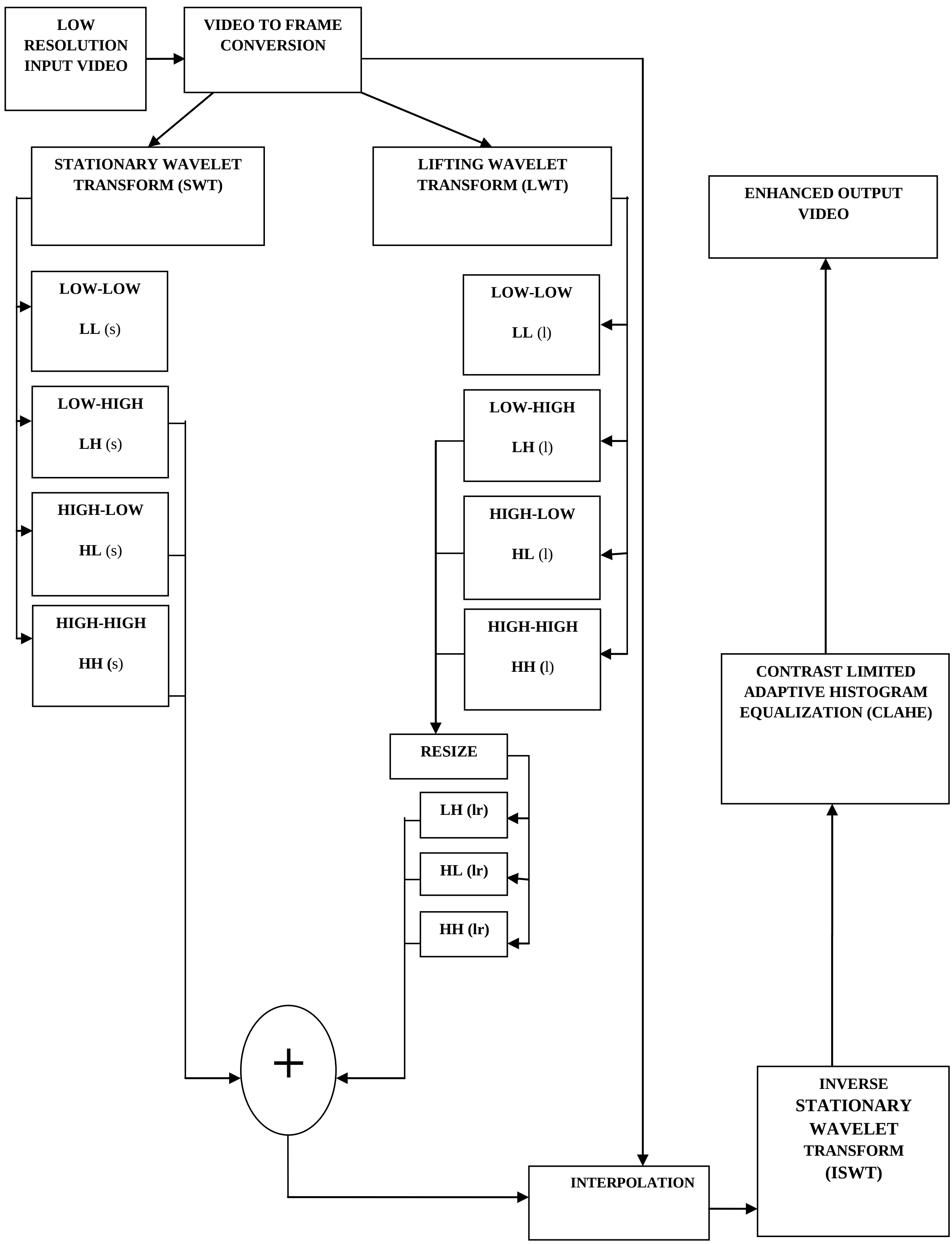


Fig 1: Details of the Proposed System with Block Diagram

3.1 Video to Frame Conversion:

The low-quality video was given as the input to enhance it. The original video is split into images (frames) using MATLAB code which converts a given video into its frames. Then, the images (frames) were transformed from RGB (three bands 24 bits/pixel) into grayscale (one band 8 bit/pixel). The input images (frames) are resized to the size 480x480 and the wavelet transform square is applied later. Square image size is needed (rows and columns should be equal). At the end, grayscale image (one band 8 bit/pixel) is converted from unsigned integer data type to double precision floating point.

3.2 Wavelet Decomposition

The LWT technique improves a given discrete wavelet transforms to obtain specific properties. Then, became an efficient technique to calculate any wavelet transforms. LWT is an efficient technique for video enhancement processing [11]. The advantages of LWT is the fast implementation of the wavelet transform and requiring simple Mathematical calculations the traditional methods that are based on discrete wavelet transform require complex Mathematical calculations and cost low of the running time.

In this paper, LWT was applied to determine the high-frequency segments of the frame (LH, HL, and HH). We applied one level LWT to maintain the main details of the video frames. The information losses in LWT because the down-sampling is occurred in all sub-bands, so to avoid information losses, SWT is utilized. In brief, LWT is familiar to SWT but LWT reduces the image size using down-sampling while SWT does not apply that. Therefore, the input frame will remain the same size until after decomposes to different sub-bands frame as shows in Figure 2.

Fig2: An example of lifting wavelet decomposition (left) and stationary wavelet decomposition (right).

3.3 Interpolation

Interpolation is used discrete samples to estimate the values of a continuous function. There are many of the interpolation applications such as image reduction or magnification, sub-pixel frame to edit spatial distortions, frame enhancement, and frame decompression. The interpolation techniques include the bilinear, the nearest neighbor, and the cubic convolution. The cubic convolution interpolation is used mostly in the image processing field to supply a reconstruction of a continuous function [12].

Cubic convolution interpolation in one dimension calculates the output gray level value from input coordinates by using the weighted average of the 4 closest pixels. Bi-cubic interpolation is the cubic convolution interpolation in two dimensions that calculates the output coordinates value by taking the average weight of the 16 closest pixels of the gray level value. Bilinear interpolation is generated a less quality output frame than the one generated by bi-cubic interpolation. The nearest neighbor interpolation has the disjointed appearance that is opposite what happening by bi-cubic interpolation [12, 13].

The one-dimension and two-dimension cubic convolution interpolation are explained in Figure 3.

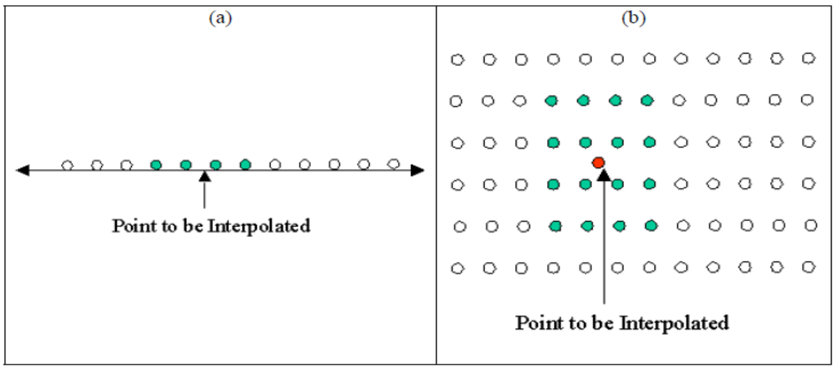


Fig3: The grid points needed for one-dimension cubic convolution interpolation and bi-cubic interpolation.

The three LWT high-frequency sub-bands for a given frame are given as input to be interpolated by bi-cubic interpolation. These high-frequency sub-bands are first resized to return it to the same frame size. As the LWT reduces the image size (down-sampling) during its processing. The high-frequency sub-bands acquired by SWT are also interpolated in order to modify the evaluated coefficients. Then, the result from these operations on LWT and SWT can be added to each other. In the same time, the original frame is interpolated. The original frame low-resolution is in the low-frequency sub-band. And then, instead of utilizing the low-frequency sub-band, we used the input frame to interpolate the low-frequency sub-band frame. Using input frame instead of the low-frequency sub-band improves the video and image quality [14] shows in Figure 4.

Fig 4: bi-cubic interpolation.

3.4 Inverse Wavelet Transform

Lastly, we used the ILWT to merge the interpolated high-frequency sub-bands and interpolated original frame to obtain in the final to a high-resolution output frame. The high-resolution output frame will have sharper edges more than the interpolated frame gained by interpolation straight of the original frame. This means, the detached high-frequency components interpolation in high-frequency sub-bands frame and utilizing the rectification is acquired by adding the result of interpolated high-frequency components sub-bands of LWT and SWT with the interpolated input frame. We keep more high-frequency components after applying the interpolation more than just the interpolated frame gained from interpolation straight of the input frame shows in Figure 5.

Fig 5: Inverse Lifting wavelets transform (ILWT).

3.5 CLAHE Enhancement:

The classical histogram equalization (HE) can display the intensities efficiently, but HE may be caused over enhancement. If there are high peaks in the histogram, the output image will have noise and disruption [5].
To overcome these defects, The CLAHE [6] is a HE-based image enhancement technique that divided the frame into series of sub-bands and non-overlapped, then improves all sub-bands separately and employs an interpolation step to minimize the block artifacts. The output which produced of the inverse wavelet transformed will be enhancing CLAHE.

CLAHE is an efficient algorithm to improve the local details of the frame. CLAHE is one of the most common techniques in video/image processing because it's very efficient for dealing with interesting parts and make it seems more visible.
The frame is decomposed into disjoint areas, and in each area is applied local histogram equalization. Therefore, this will eliminate the boundaries between the areas by using a bilinear interpolation [6].

The technique major goal is the point transformation detection interior of local large window assuming the intensity value interior. It is the local distribution representation stoical of the total frame intensity value. Using the intensity gradual variation between the edges and image centers leads to make the local window unaffected. The distribution of point transformation is centralized on the window average intensity and the point transformation distribution covers the entire frame intensity range [14].
As an output of this adaptive histogram equalization, the dark region in the input frame that was poor illumination will become good illumination it is brighter in the output frame, even bright region that was good illumination keeps or decreases it. So, total frame illumination remains the equivalent shows in Figure 6.

Fig 6: The output frame

4. Experimental Results

4.1 Evaluation Metrics:

MSE

MSE measures the difference between pixels of the original image y and the enhanced images x without considering the correlation between the neighboring pixels. MSE between these two images is defined as:

$$\mathbf{MSE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2 \qquad \textbf{(1)}$$

Where N is the total number of pixels, and x_i and y_i are the values of the i-th pixels in x and y, respectively.

PSNR

PSNR is the most used to measure the quality of a video, image and sound files in dB (decibels).

$$\mathbf{PSNR} = \mathbf{10} \times \mathbf{log_{10}} \frac{(255)^2}{\mathbf{MSE}} \qquad \textbf{(2)}$$

SSIM

SSIM is a measure that calculates the quality of the image. The SSIM requires two frames from the same video capture an original frame and an enhanced frame. The enhanced frame is usually compressed. In generally, SSIM measures the difference between two identical frames. SSIM unlike PSNR, SSIM is based on visible structures in the frame while PSNR isn't considered as indicator of frame quality degradation.

4.2 Results and Discussion:

Intel ® core™ i5 M 460 @ 2.53GHz RAM 4.00 GB 1366 *768(32bit) (60Hz) Generic PnP Monitor The proposed technique is evaluated using a dataset of eight video sequences [15], namely "Mother daughter", "Akiyo", "Foreman", "Container", "Bowing", "Coastguard", "Pamphlet", and "Silent". For each video, the numbers of frames are 300 frames, the length of each is 10 seconds, with a size of 1.08 MB and its format is AVI. The low-resolution video sequences have the size of 128 × 128.
Our experimentations are implemented using MATLAB version 2012, running on.

The proposed technique is compared to some of the various enhancement techniques such as Vandewalle et al. [9] which examined a frequency domain technique to registers a group of low-frequency aliased images, Therefore, it reconstructs the high-resolution image. Keren et al. [10] proposed an approach to register the sequence of images captured from the camera, and Anbarjafariet al. [7] applied an illumination compensation technique then a registration step and using DWT, SWT, and the Vandewalle technique on the output. Finally, all these sub-bands have been collected using IDWT.

Table 1: The PSNR of the proposed method compared to various enhancement techniques on the same dataset.

resolution enhancement technique		**psnr (db) value for different sequences**			
Method	**reconstruction**	**"mother daughter"**	**"akiyo"**	**"foreman"**	**"container"**
Vandewalle [9]	**Interpolation**	25.30	28.46	27.99	24.1
	Iterated Back Projection	27.88	32.05	31.18	25.2
	Robust SR	26.98	31.6	30.26	24.47
	Structure Adaptive Normalized Convolution	28.96	32.99	33.47	26.39
Keren [10]	**Structure Adaptive Normalized Convolution**	28.64	32.98	33.26	26.16
Gholamreza [7]	**DWT and SWT-based resolution enhancement with illumination compensation**	34.18	38.44	38.90	33.19
The proposed technique LWT and SWT-based resolution enhancement with CLAHE algorithm		**35.64**	**42.96**	**43.56**	**37.26**

Table**2** shows the average PSNR values compared to other enhancement techniques for the video. As noticed the proposed technique is better than different other techniques because the average PSNR has high value than other values.

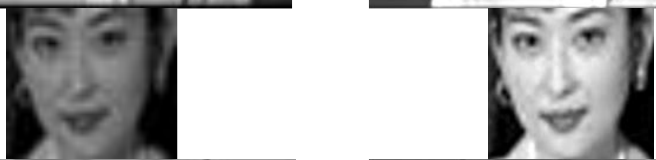
Table 3: The average performance of the proposed method using PSNR, mean square error, average running time and SSIM on 8 videos

the average of performance values	**mother daughter**	**akiyo**	**foreman**	**container**	**bowing**	**coastguard**	**pamphlet**	**silent**
PSNR	**35.64**	**42.96**	**43.56**	**37.26**	**30.17**	**41.56**	**42.21**	**44.72**
MSE	**0.0030**	**0.0017**	**0.0025**	**0.0003**	**0.0412**	**0.0011**	**0.0036**	**0.0005**
Time	**0.00162**	**0.0027**	**0.00032**	**.00543**	**0.0036**	**0.0007**	**0.0002**	**.0004**
SSIM	**0.367**	**0.367**	**0.367**	**0.367**	**0.366**	**0.252**	**0.242**	**0.273**

Table 3 shows the average of performance values such PSNR, mean square error, average running time and SSIM for eight video sequences [15] namely "Akiyo", "Mother daughter", "Foreman", "Container", "Bowing", "Coastguard", "Pamphlet", and "Silent". Figures 8–15 show frames that are selected randomly from "Akiyo", "Mother daughter", "Foreman", "Container" ,"Bowing", "Coastguard", "Pamphlet", and "Silent" videos for the visual analysis.

Input frame **output frame**

Fig8: Akiyo



Frame:1

Fig 9: Mother daughter

Frame: 1

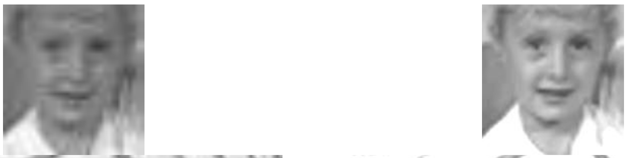


Fig 10: Forman

Frame: 1

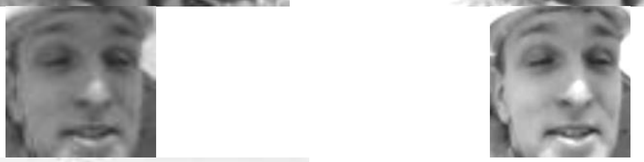


Fig 11: Container

Frame: 1

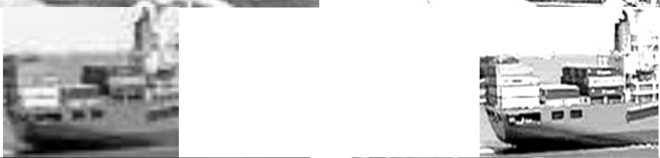


Fig 12: Bowing

Frame: 1

Fig 13: Coastguard

Frame: 1

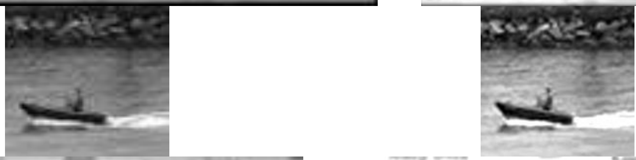


Fig 14: Pamphlet

Frame: 1

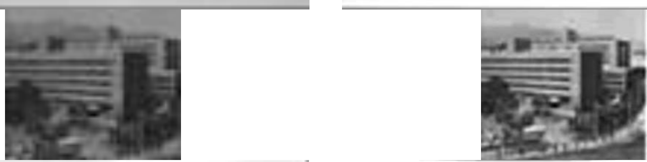
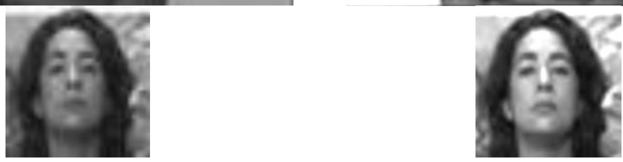


Fig 15: Silent

Frame: 1



As we noted from figures 8:15the noise is enhanced by the proposed technique and became more specific and clear. This occurs as the proposed technique enhances only the low-frequency segment and makes no changes to the high-frequency component that includes most of the original image noise. It is observed that some bright parts details of enhanced images using CLAHE are lost due to over enhancement that is improved in our technique as we showed in the figures.

6. Conclusion and Future Work

In this work, we proposed a novel video enhancement technique that generates sharper and the high-resolution video frames. The proposed technique used the LWT with SWT and CLAHE. The LWT and SWT decompose the input frame into low-frequency and high-frequency components. The high-frequency sub-bands obtained by LWT and the high-frequency sub-bands obtained by SWT have been interpolated by bi-cubic interpolation. At the same time, the original frame is also interpolated. Afterward, all these frames have been combined using ILWT to obtain super-resolution video frames. Lastly, we applied CLAHE to enhance the quality of the video generally. The proposed technique has been tested on low-resolution videos, where their PSNR and visual results show the superiority of proposed technique over the conventional video enhancement techniques. The experimental results show the better performance of the proposed technique, which reduces the time-consuming and can alleviate noises and over-enhancement problem. The future work includes applying the proposed method on medical images in order to improve its resolution and facilitate the diseases diagnosis.

REFERENCES:

References:

[1] YunboRao and Leiting Chen, "A survey of video enhancement techniques", Journal of Information Hiding and Multimedia Signal Processing, volume3, number1, pp. 71–95,January 2012.

[2] AbhilashGhatol andSourabh Sharma, “Video image enhancement and object tracking", Journal for Technological Research in Engineering, Volume 1, number8,pp. 1–15, April-2014.

[3] Agaian, Blair and Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy", IEEE Trans. Image Processing, volume 16, number 3, pp. 741-758, February2007.

[4] YunBo Rao, Lin and Chen, "Image-based fusion for video enhancement of nighttime surveillance", Opticalengineering Letters, volume 75, number18, pp.11221-11239, September 2016. [5] Huang Lidong, Zhao Wei, Wang Jun. and Sun Zebin" a combination of contrast limited adaptive histogram equalization and discrete wavelet transform for image enhancement", IET Image Process, Volume 9, number 10, pp. 908–915, September 2015.

[6] Pizer, S.M., Amburn, E.P., Austin, J.Dand et al, "Adaptive histogram equalization and its variations", Computer Vision. Graph and image process, vol. 39, no. 3, pp. 355–368,September 1987.

[7] GholamrezaAnbarjafari, Sara Izadpanahi and Hasan Demirel, "Video resolution enhancement by using discrete and stationary wavelet transforms with illumination compensation", Springer-VerlagLondon, Volume9, Number1, pp. 87–92,January2015.

[8] Demirel and Anbarjafari, "Satellite image resolution enhancement using complex wavelet transform", IEEE Geosciences and Remote Sensing Letter, Volume 7, number 1, pp. 123–126, September2010.

[9] Vandewalle, Süsstrunk, and Vetterli, "A frequency domain approach to registration of aliased images with application to super resolution" EURASIP Journal on Applied Signal Processing, Volume 2006, Article ID 71459, pp.1–14, May2006.

[10] Keren, Peleg and Brada, "Image sequence enhancement using subpixel displacements", IEEE Computer Soc. Conf. Computer Vision Pattern Recognition, Ann Arbor, MI, pp. 742–746 , June1988.

[11] Xiu-bi Wang, "Image enhancement based on lifting wavelet transform ", in Proceeding. International Conf. on Computer Science & Education, China, pp. 739- 741, September2009.

[12] G.Anbarjafari and H. Demirel, "Image super resolution based on interpolation Of wavelet domain high frequency sub-bands and the spatial domain input image", ETRI Journal, Volume 32, number 3, pp.390–394, June2010.

[13] Kian Kee Teoh and Haidi Ibrahim, "Investigation on Several Basic Interpolation Methods for the Use in Remote Sensing Application", in Proceeding. IEEE conf. on Innovative Technologies in Intelligent Systems and Industrial Applications, Cyberjaya, Malaysia ,pp.253 - 257, August2008.

[14] Kim, T.K., Paik, J.K., Kang and B.S, "Contrast enhancement system using spatially adaptive histogram equalization with temporal filtering", IEEE Transactions on Consumer Electronics, vol. 44, no. 1, pp. 82–86, February1998.

[15] Xiph.org Test Media, http://media.xiph.org/video/derf/ , retrieved on October 2010.

[16] [Online]. Available at http://r0k.us/graphics/kodak/, accessed August 2010.