Vol 06, Article 05577; July 2015
http://ijves.com
International Journal of VLSI and Embedded Systems-IJVES
ISSN: 2249 – 6556

PERFORMANCE EVALUATION OF HHT AND LOW PASS FIR FILTER FOR IMAGE STABILIZATION USING VHDL

JINSHA KRISHNAN, K.V.KARTHIKEYAN
Department of Electronics and Communication Engineering, Sathyabama University, Chennai
jinsha.krishnan@gmail.com, kandvas@gmail.com

ABSTRACT

The Digital Image Stabilization(DIS) is the process of removing the unwanted motion effects of a moving camera to produce a compensated image sequence by using image processing techniques. A technique for Digital Image Stabilization based on the Hilbert–Huang Transform (HHT) and low pass Frequency Impulse Response(FIR) filter has been used in this work. The method exploits a fundamental difference between the intentional(low frequency components) and the jitter motion (high frequency components).Initially local motion vectors of an image sequence are calculated and is then divided into a number of waveforms, called Intrinsic Mode Functions (IMFs), using the process of Empirical Mode Decomposition. Hilbert transform and FIR filter coefficients are applied to each IMF so that the energy content could be designated. Based on the basic features of the unwanted shaking phenomena (high frequencies and small power contents), intentional and jitter motions are determined, and thus, motion compensation is applied in order to eliminate possible fluctuations and produce an image sequence with smoother transitions. The Hilbert transform and FIR filter removes high frequency unwanted motion. The jitter represented in the simulation result gives that how much the image sequence should be adjusted to get stabilized image. When the image stabilization is done using FIR filter the power has been reduced.

Keywords: Digital image stabilization (DIS), Hilbert–Huang transform (HHT), Local Motion Vectors.

1. INTRODUCTION

Image stabilization has become a subject of significant interest and an active research field over the past years due to the wide use of digital imaging devices. The image stabilization process aims at removing irregular motion phenomena from image sequences in order to accomplish a compensated sequence that displays smooth camera movements.

A variety of embedded systems equipped with a digital image sensor, such as handheld cameras, mobile phones, and robots, can produce image sequences with an observed motion caused by two different types of movements: the smooth camera motion (intentional) and the unwanted shaking motion (jitter). Based on the basic features of the unwanted shaking phenomena or high frequencies and small power contents, intentional and jitter motions are determined, and thus, motion compensation is applied in order to eliminate possible fluctuations and produce an image sequence with smoother transitions**.** Hilbert Huang Transform (HHT) and low pass FIR filter are used to get stabilized images.

An image is an artefact that depicts or records visual perception, for example a two-dimensional picture, that has a similar appearance to some subject usually a physical object or a person, thus providing a depiction of it.

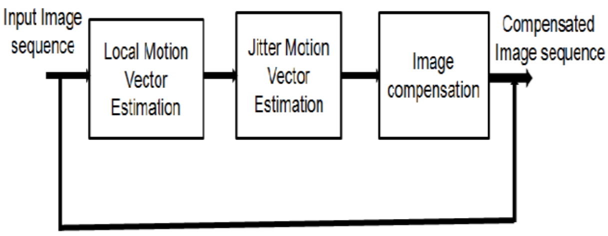

Fig.1.1: DIS system after acquiring the image sequence

Video is an electronic medium for recording copying and broadcasting of moving visual images. Video is visual multimedia source that combines a sequence of images to form a moving picture. With video cameras, camera shake causes visible frame-to-frame jitter in the recorded video. The image stabilization process aims at removing irregular motion phenomena from image sequences in order to accomplish a compensated sequence that displays smooth camera movements. A variety of image processing applications requires motion-compensated image sequences as inputs. The unwanted positional vacillations of the video sequence will affect the visual quality and impede the subsequent processes for several applications .Vehicles equipped with vision systems use image stabilization to achieve maximum performance in stereo image analysis. Imaging devices offering video recording capabilities have gained significant popularity. Image stabilization system was integrated in a solar optical telescope using image displacements in order to eliminate the jitter motion from the acquired sequence of a satellite . Finally, a video stabilization architecture for low-cost embedded systems was implemented, which deals with the filtering of translational and rotational unwanted motion [3].

Vol 06, Article 05577; July 2015
http://ijves.com
International Journal of VLSI and Embedded Systems-IJVES
ISSN: 2249 – 6556

The digital image stabilization (DIS) is the process of removing the unwanted motion effects of a moving camera to produce a compensated image sequence by using image processing techniques. This technique shifts the electronic image from frame to frame of video[1]. HHT and low pass FIR filter is applied the digital image stabilization method. Block diagram of DIS method is shown in Fig 1.1.

2. EXISTING METHODOLOGY

Image stabilization is a family of technique to reduce blurring associated with the motion of a camera during exposure. It removes irregular motion phenomena from image sequences in order to accomplish a compensated sequence that displays smooth camera movements. The project is finding out jitter and there by knowing that to which direction the image frame must be moved to get stabilized image frames.

Hilbert–Huang transform (HHT) is a signal-processing tool that adaptively decomposes non stationary signals through the process of Empirical Mode Decomposition (EMD) into basis functions called Intrinsic Mode Functions (IMFs). The HHT combines the EMD and the Hilbert spectral analysis (HSA); the HSA includes the Hilbert transform of each Intrinsic Mode Function(IMF) generated by the EMD process. The Hilbert transform of each IMF is well behaved, and the instantaneous frequency and instantaneous amplitude can be determined from the subsequent analytic signal that is formed from the IMF and its Hilbert transform. Several applications have been proposed using the HHT including speech analysis, health monitoring, electroencephalographic data analysis and seismic feature construction. Additionally, the EMD process was proposed as an independent method for a number of applications such as signal denoising. A 2-D version of EMD (bidimensional empirical mode decomposition) was first proposed in as an application for image compression as well as in for image fusion and enhancement. Finally, the process of EMD was used as an image stabilization method in nonetheless without applying any Hilbert spectra analysis. The HHT technique for analyzing data consists of two components: a decomposition algorithm called Empirical Mode Decomposition (EMD) and a spectral analysis tool called Hilbert spectral analysis. The purpose of HHT is to demonstrate an alternative method to present spectral analysis tools for providing the time-frequency-energy description of time series data. Also, the method attempts to describe non stationary data locally. The EMD algorithm is the other component to the HHT method. The algorithm attempts to decompose nearly any signal into a finite set of functions, whose Hilbert transforms give physical instantaneous frequency values. These functions are called intrinsic mode functions (IMFs).

Image stabilization using Hilbert Huang transform used to separate unwanted jitter motion from the image sequences . First, a motion estimation method is applied in order to define the LMV of an image sequence. Subsequently, the estimated LMV is decomposed into a finite number of IMFs by applying the EMD process. Each IMF is transformed using Hilbert transformation in order to define the energy content of every decomposed signal. Depending on the estimated energies, the last IMF to be considered as jitter is designated[6]. Thus, the summation of all IMFs with the lower indices up to the specified IMF from the Hilbert transform approximates the unwanted jitter motion.

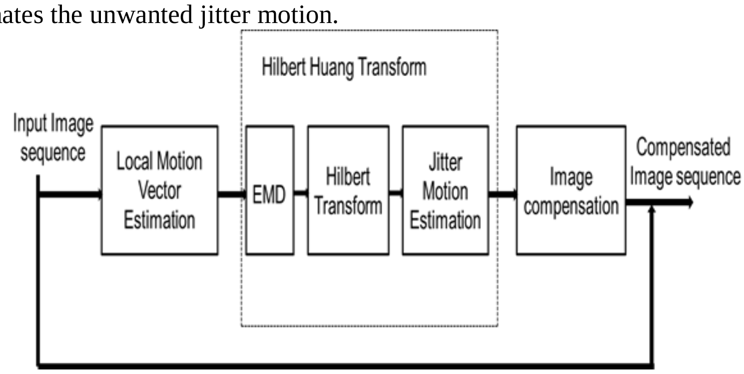

Fig. 2.1:DIS method by HHT.

2.1 Local Motion Vectors (LMV)

An image stabilization process, applied after the image acquisition, consists of three main stages: global motion estimation, jitter motion determination, and image warping [1]. The first processing stage is dedicated to the determination of parameters that designates the entire camera motion, meaning that displacements between either blocks or features in subsequent frames are determined. On the contrary, local motion vectors (LMVs) are calculated within smaller frame regions during the process of motion estimation[5]. Motion compensation is an algorithmic technique employed in the encoding of video data for video compression, for example in the generation of MPEG-2 files[4]. Motion compensation describes a picture in terms of the transformation of a reference picture to the current picture. The The reference picture may be previous in time or even from the future. When images can be accurately synthesised from previously transmitted/stored

Vol 06, Article 05577; July 2015
http://ijves.com
International Journal of VLSI and Embedded Systems-IJVES
ISSN: 2249 – 6556

images, the compression efficiency can be improved. Essentially, LMVs represent the offset of specific image regions between two consecutive frames. Thus, LMVs include both the intentional and the unwanted motion of the camera. This process dominates over the other two stages in terms of time complexity. Therefore, the main objective of many proposed methods for image stabilization is the improvement of the performance of the motion estimation in terms of complexity and accuracy. The full-search block matching method using the mean absolute difference and mean square error (MSE) criteria is considered to be the optimal solution for the block motion estimation, since the most accurate and reliable results are ensued against the required processing time. Nevertheless, the appropriate segregation between the wanted movement of the camera and the jitter is considered to be the most crucial stage of the entire process. In this stage, accurate separation between the estimated LMVs should be performed. Each LMV may represent only a global motion vector (GMV), a moving-object motion vector, or even an error vector. Irregular conditions or mixture between global motion, moving-motion, and unwanted jitter motion could produce such an error vector. Finally, after the discrimination between the intentional and the jitter motion, the image sequence is compensated in the final stage of the process, where a stabilized version of the sequence is produced.

2.2 Hilbert Huang Transform

Hilbert Huang Transform is combination of EMD and Hilbert Transform

2.2.1 Empirical Mode Decomposition (EMD)

In general, the EMD separates non stationary data into locally non overlapping time-scale components. The signal decomposition process dissociates the original signal into a set of complete and almost orthogonal components, namely, IMFs. An IMF is a function that satisfies the following two conditions.

- The number of extrema and the number of zero crossings must either equal or differ by at most one in the whole data sets.
- The mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero at every point.

To begin the EMD, a function or signal is decomposed as follows.

First, local extrema are identified. Local maxima are connected by cubic spline to form the upper envelope. The exact same procedure is followed for the local minima to produce the lower envelope. All the data should be included between the upper and the lower envelope. If the mean of the upper and lower envelopes is designated as m_1 and the difference between the data x(t) and m_1 is the first component h_1 is equation 2.1, then

$$h_1(t) = x(t) - m_1(t) \quad (2.1)$$

$$m1(t) = [U(t) + L(t)]/2 \quad (2.2)$$

where $U(t)$ and $L(t)$ are the local maxima and the local minima, respectively. Equation 2.2 gives the mean.

Essentially, $h_1(t)$ is supposed to be an IMF, except that some error might be introduced by the spline curve fitting process. Therefore, a sifting process is repeated multiple times. The sifting process serves two purposes: Riding waves smaller waves that seem to ride bigger waves are eliminated, and the signal is transformed into a more symmetric form about the local zero-mean line.

In the second round of sifting, h_1 is treated as the data or the first component. Then, a new mean is computed with the same former procedure, is given in equation 2.3. Considering that the new mean is m_{11}, then

$$h_{11}(t) = h_1(t) - m_{11}(t). \quad (2.3)$$

After repeating the sifting process up to k times, h_{1k} is labelled as an IMF, is given in equation 2.4.

$$h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t) \quad (2.4)$$

Let $h_{1k} = c_1$ be the first IMF from the data. c_1 should contain the finest scale or the shortest period component of the data. The process to generate one IMF may be considered as an inner loop, as shown in Fig. 2.2 Then, $c_1(t)$ is removed from the rest of the data so that the residual is calculated in equation 2.5

$$r_1(t) = x(t) - c_1(t) \quad (2.5)$$

where r_1 is the residue and it contains information on longer period components. The residue is treated as the new data and subjected to the same sifting process (start of outer loop in Fig. 2.2).The aforementioned procedure is repeated in order to obtain all the subsequent r_w functions as given in equation 2.6.

$$r_w(t) = r_w - 1(t) - c_w(t) \quad (2.6)$$

where $w = 2, 3, .., n$. A pictorial depiction of the EMD process is shown in Fig. 2.2.

Terminating criteria are applied to the sifting process for IMFs since allowing sifting to go beyond a certain point may smooth out important signal variations and features that arise from the natural dynamics of the system.

Vol 06, Article 05577; July 2015
http://ijves.com
International Journal of VLSI and Embedded Systems-IJVES
ISSN: 2249 – 6556

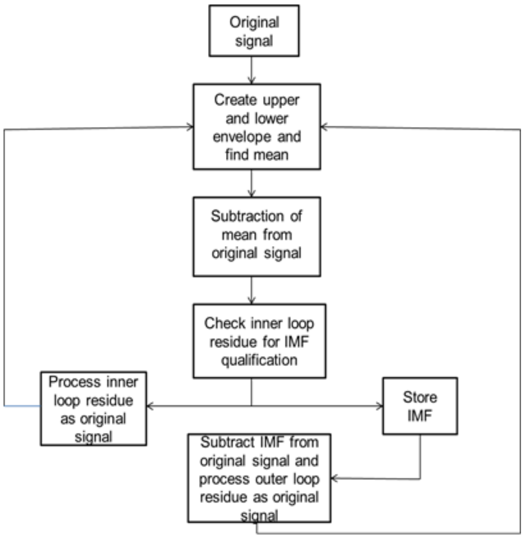

Fig.2. 2: Flowchart of EMD process.

The IMF components must retain enough physical sense of both amplitude and frequency modulations which can be achieved by limiting the value of the sum of the difference computed from two sequential sifting results as given in equation 2.7.

$$\frac{\sum_{t=0}^{T} |h_{k-1}(t)-h_k(t)|^2}{\sum_{t=0}^{T} h^2_{k-1}(t)} \quad (2.7)$$

where T represents the total number of the samples and k represents the iteration number of the sifting process.
A value of SD between 0.2 and 0.3 is usually preferable based on experimental analyses performed by Huang. An alternate stopping criterion, proposed by Huang, is to check if the number of zero crossings is equal to or differs by at most one from the number of the total extrema.
The entire EMD process is terminated if any of the following criteria is satisfied:
1) when the residue r_w is a function with one unique extremum
2) when the residue r_w becomes a monotonic function from which no IMF can be extracted.
Summing (2.5) and (2.6) yields the following equation 2.8:

$$x(t) = \sum_{j=1}^{w} c_j + r_w \quad (2.8)$$

which indicates completeness, in that the sum of the IMFs and the residue recovers the original signal. cj is the jth IMF, and w is the number of sifted IMFs. r_w can be interpreted as the general trend of the original signal.

3. PROPOSED METHODOLOGY

Image processing using VHDL is a new concept. In this work an image stabilizer is designed in VHDL. For stabilization Hilbert Huang Transform and low pass FIR filter coefficients are used.
Image stabilization does not prevent motion blur caused by the movement of the subject or by extreme movements of the camera. Image stabilization is only designed for and capable of reducing blur that results from normal, minute shaking of a lens due to hand-held shooting.
The project is finding out jitter and there by knowing that to which direction the image frame must be moved to get stabilized image frames. In the proposed system instead of applying Hilbert transform, low pass FIR filtering is used.

3.1 FIR filter

In signal processing a Finite Impulse Response (FIR) filter is a filter whose impulse response is of finite duration, because it settles to zero in finite time. This is in contrast to infinite impulse response (IIR) filters, which may have internal feedback and may continue to respond indefinitely (usually decaying). The impulse response of an Nth-order discrete-time FIR filter i.e., with a Kronecker delta impulse input lasts for N + 1 samples, and then settles to zero.
FIR filters can be discrete-time or continuous-time, and digital or analog.The output y of a linear time invariant system is determined by convolving its input signal x with its impulse response b. For a discrete-time FIR filter, the output is a weighted sum of the current and a finite number of previous values of the input. The operation is described by the following equation 3.1, which defines the output sequence y[n] in terms of its input sequence x[n]:

$$Y[n] = B_0x[n]+B_1x[n-1].....+B_nx[n-N]$$
$$= \sum_{i=0}^{N} Bix[n-i] \quad (3.1)$$

Vol 06, Article 05577; July 2015
http://ijves.com
International Journal of VLSI and Embedded Systems-IJVES
ISSN: 2249 – 6556

where:

- x[n] is the input signal,
- y[n] is the output signal,
- b_I are the filter coefficients, also known as tap weights, that make up the impulse response,
- N is the filter order; an Nth-order filter has (N+1) terms on the right-hand side. The x[n-i] in these terms are commonly referred to as taps, based on the structure of a tapped delay line that in many implementations or block diagrams provides the delayed inputs to the multiplication operations. One may speak of a 5th order/6-tap filter, for instance.

Video consist of number of frames. For analysis and understanding in simple way only two frames are considered here.

Steps followed are:

(1) Local motion vector estimation
(2) Finding displacement
(3) Finding minimum and maximum displacement
(4) Finding mean (IMFs)
(5) Low pass FIR filter
(6) Finding jitter.

3.1.1 Local Motion Vector estimation and displacement

Motion estimation is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence. It is an ill-posed problem as the motion is in three dimensions but the images are a projection of the 3D scene onto a 2D plane. The motion vectors may relate to the whole image (global motion estimation) or specific parts, such as rectangular blocks, arbitrary shaped patches or even per pixel. The motion vectors may be represented by a translational model or many other models that can approximate the motion of a real video camera, such as rotation and translation in all three dimensions and zoom.

Steps can be simply explained as:

- The current frame is first divided into blocks of $M \times N$ pixels. Here is it taken as 32X32.
- The algorithm then assumes that all pixels within the block undergo the same translational movement.Thus, the same motion vector is assigned to all pixels within the block.
- So a reference block is selected, shown in Fig.3.1.
- The block is compared with the nearby blocks and displacement vector is calculated. Fig 3.2 shows the comparison

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

Fig. 3.1:Reference block

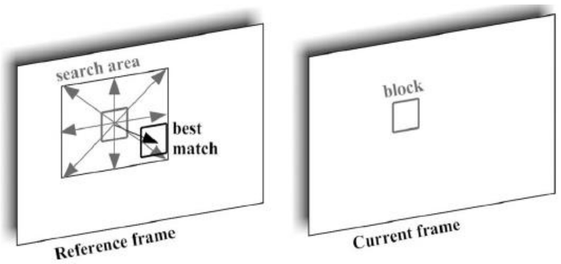

Fig.3.2: Comparing reference block with nearby blocks

3.1.2. EMD process for LMV

In this process, IMF is calculated.

Steps are:

- Displacements calculated from comparison of blocks, are compared each other
- Minimum displacement and maximum displacement is calculated
- Thus local maxima and minima is obtained
- From this mean is calculated
- This mean is considered as IMF.

3.1.3. Designating the unwanted motion

This step includes applying frequency impulse response to the IMFs and finding jitter.The EMD process divides the initial LMV signal into a number of sub signals with specific features. Lower IMF indices indicate higher frequencies, according to Huang. Nevertheless, in order to identify both jitter and intentional motion, the last IMF which includes the high frequencies of the jitter motion must be defined. The IMF is multiplied with the FIR filter coefficients. Low pass FIR filter with sampling frequency 25Hz and with 1Hz to 10Hz pass band The

International Journal of VLSI and Embedded Systems-IJVES
ISSN: 2249 – 6556

image stabilization using Hilbert and FIR filter removes high frequency unwanted motion. This low frequency is jitter or shake of the camera. After passing through low pass fir filter, the result is stored in jitter out. According to jitter output it is understood that to which direction the camera should be moved to get stabilized image.

4.RESULTS AND DISCUSSION

Here simulation output waveform of existing and proposed Image stabilization methods are given. The design is simulated by using the simulation tool ModelSim. The version used is ModelSim 6.3g**.**

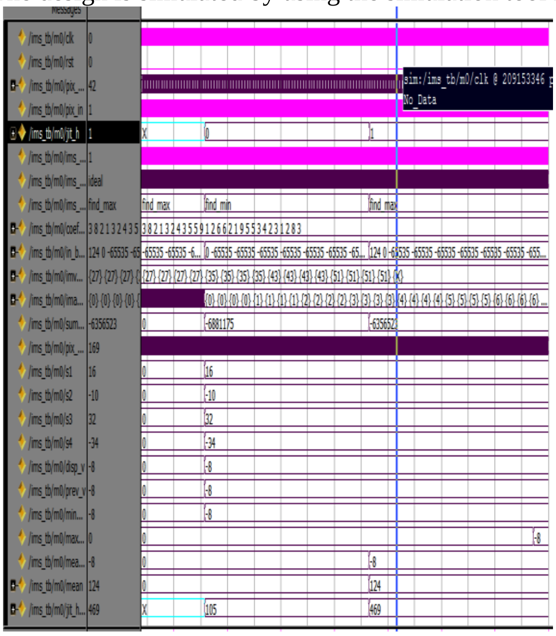

Fig.4.1: Simulation output of image stabilizer using HHT

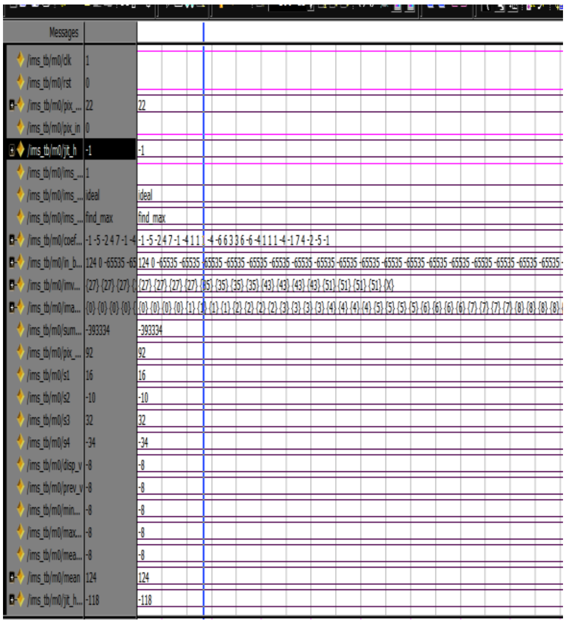

Fig.4.2:Simulation output of image stabilizer using low pass FIR filter

Fig. 4.1 describes the output waveform of the image stabilization using Hilbert Huang Transform. The simulation result gives to which direction image sequence should be adjusted to get stabilized image in the waveform jit_h=1 implies jitter output is 1, the image should be shifted to left. Fig.4.2.describes the output waveform of the image stabilization using Low pass FIR filter.

The simulation result gives to which direction image sequence should be adjusted to get stabilized image. Here jit_h= -1 implies that jitter output is -1, the image should be shifted to right. The performance analysis of existing image stabilization using hht and proposed image stabilization using low pass fir filter is done using XILINX 8.1. Comparison of performances in terms of power, area, memory and gate count in design is also done. Table 4.1 and 4.2 gives gate count and the power analysis.

Table 4.1: Gate count analysis of both methods

Methodologies	Gate count
HHT IMAGE STABILIZATION	82,855
FIR IMAGE STABILIZATION	67,079

Table 4.2: Power analysis of both methods

Methodologies	Power consumption(mW)
HHT IMAGE STABILIZATION	56.11 mW
FIR IMAGE STABILIZATION	56.01 mW

From Table 4.2 it can be seen that power of then image stabilization using low pass filter have been reduced by 0.10mw. Table 4.1 gives the area analysis of the image stabilization using HHT and FIR .It can be seen that gate count of the image stabilizer using low pass FIR filter has been reduced.

CONCLUSION

The digital image stabilization is the process of removing the unwanted motion effects of a moving camera to produce a compensated image sequence by using image processing techniques. DIS methods outperform other image stabilization methods by means of hardware requirements and flexibility, since they are hardware

Vol 06, Article 05577; July 2015
http://ijves.com
International Journal of VLSI and Embedded Systems-IJVES
ISSN: 2249 – 6556

independent. Video is an electronic medium for recording copying and broadcasting of moving visual images. Video is visual multimedia source that combines a sequence of images to form a moving picture. With video cameras, camera shake causes visible frame to frame jitter in the recorded video.

In this work, digital image stabilization method based on the HHT and low pass FIR filters has been used. The method acquires the required data from an image sequence in order to achieve the definition of the two fundamental motions, the intentional and the trembled camera motion. The method exploits a fundamental difference between the intentional and the jitter motion. The intentional motion of a camera is characterized by a smooth transition between two consecutive frames during the entire image sequence. Thus, it includes lower frequency components. In addition, this motion exhibits higher pixel displacements compared to the jitter motion in terms of amplitude leading to higher energy levels. The jitter motion is characterized by a trembled movement of the camera leading to higher frequency components. These two motions are inserted in the image sequence due to both the camera movement and a non steady motion of the camera. Thus, image sequences captured with handheld cameras, mobile phones, or even with a camera of a mobile robot can be processed to produce a stabilized sequence with smoother frame transitions. The effectiveness of the method relies on the fact that the jitter signal is defined based on its two principal features, high frequencies and low-energy content. Therefore, the definition of both the jitter and the intentional movement is straightforward.

The Hilbert transform and FIR filter removes high frequency unwanted motion. The jitter represented in the simulation result gives that how much the image sequence should be adjusted to get stabilized image. When the image stabilization is done using FIR filter the power has been reduced.

Both the image stabilizers are coded using VHDL and simulated using Modelsim 6.3g and output is obtained. Here random pixel values are selected and given as the input and at the output the direction to which the image sequence to be moved is obtained. Future work is hardware implementation of the image stabilizers and integrating in cameras.

ACKNOWLEDGEMENTS

We express our profound gratitude to Honourable Chancellor of Sathyabama university Col.Dr.Jeppiar. We feel greatly indebted to express our sincere gratitude to our Directors Directors Dr. Marie Johnson and Dr.Mariazeena Johnson for providing necessary facilities for the completion of this paper. We thank all those who have inspired and motivated us to make the work a success.

REFERENCES

[1]. Amanatiadis A. A. and Andreadis I.(2010), “Digital image stabilization by independent component analysis,” IEEE Trans. Instrum. Meas., July.vol. 59, no. 7,pp. 1755–1763.

[2]. Atkinson.D, Hill D. L. G., Stoyle P. N. R., Summers P. E., and Keevil S. F.(1997), “Automatic correction of motion artifacts in magnetic resonance images using an entropy focus criterion,” IEEE Trans. Med. Imag., Dec vol. 16, no. 6, pp. 903–910.

[3]. Auberger.S and Miro.C(2005), “Digital video stabilization architecture for low cost devices,” in Proc. 4th Int. Symp. Image Signal Process. Anal., 2005, pp. 474–479.

[4]. Baglietto P, Maresca M, Migliaro A, and Migliardi. M(1995)., “Parallel implementationof the full search block matching algorithm for motion estimation,”in Proc. Int. Conf. Appl. Specific Array Process., pp. 182–192.

[5]. Battiato.S, Bruna A. R., and Puglisi G.(2010), “A robust block based image/video registration approach for mobile imaging devices,” IEEE Trans. Multimedia, Nov.vol. 12, no. 7, pp. 622–635.

[6]. Bosco A, Bruna A, Battiato S, Bella G, and Puglisi G(2008), “Digital video stabilization through curve warping techniques,” IEEE Trans Consum.Electron., May.vol. 54, no. 2, pp. 220–224.

[7]. Burke.B.E, Reich R. K., Savoye E. D., and Tonry J. L.(1994), “An orthogonaltransfer CCD imager,” IEEE Trans. Electron Devices, Dec.vol. 41, no. 12, pp. 2482–2484.