Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 6, Issue, 11, pp.10111-10115, November, 2014

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

# IMPROVED IMAGE STABILIZATION METHOD BASED ON THE HILBERT-HUANG TRANSFORM

*,[1]Kalyani Patil and [2]Vijay Patil

[1]Department of E&TC Engg, D. N. Patel College of Engineering, Maharashtra, India

ARTICLE INFO

*Article History:*
Received 09th August, 2014
Received in revised form
23rd September, 2014
Accepted 15th October, 2014
Published online 30th November, 2014

*Key words:*
Digital Image Stabilization (DIS),
HHT, Image sequence,
Jitter motion,
IMFs, LMV and EMD.

ABSTRACT

In this paper, DIS technique is demonstrated. DIS is projected to stably eliminate the nonessential shaking phenomena in the image sequences captured by cameras without the influence caused by moving object in the image or intentional motion. Local Motion Vector estimation method is used Sum Of absolute difference (SAD) technique. Local motion vectors (LMV) of an image sequence are calculated. Image sequence of LMV is used for DIS Method, which is based on the Hilbert–Huang transform (HHT) is projected. The HHT has been successfully employed in applications for disaggregating signals into smaller portions with specific features (e.g., biomedical applications). For DIS, local motion vectors of an image sequence are measured, and they are controlled by the HHT in order to express both signals. The original signal is divided into a number of waveforms, called intrinsic mode functions (IMFs), using the process of empirical mode decomposition. Hilbert transform is useful to each IMF so that the energy content could be chosen. The real Signal is divided into a number of waveforms, called intrinsic mode functions (IMFs), using the process of empirical mode decomposition.



## INTRODUCTION

Digital image sensor, such as handheld cameras, mobile phones, and robots, are equipped with range of embedded systems which can produce image sequences. The produced image sequences covers motion caused by two different types of movements: the plane camera motion (intentional) and the undesirable trembling motion (jitter). The image stabilization procedure aims at eliminating irregular motion phenomena from image sequences in order to realize a compensated sequence that shows smooth camera movements. A variety of image processing applications needs motion-compensated image sequences as inputs. The undesirable positional vacillations of the video sequence will spot the visual excellence and impede the subsequent procedures for several applications. Vehicles equipped with vision structures use image stabilization to accomplish maximum performance in stereo image examination. In addition, unwanted motion is eliminated from motion-corrupted magnetic resonance images. Image stabilization is also integrated in video communication systems with compression codecs in order to increase efficiency. An image stabilization system was combined in a solar optical telescope using image displacements in order to abolish the jitter motion from the acquired sequence of a satellite. Finally, a video stabilization architecture for low-cost embedded systems was implemented in, which deals with the filtering of translational and rotational undesirable motion. The digital image stabilization (DIS) is the procedure of eliminating the undesirable motion effects of a moving camera to outturn a compensated image sequence by using image processing methods. Typically, an image stabilization procedure, useful after the image acquisition, it is having three main phases: firstly global motion estimation, second one is jitter motion determination, and the third one image warping. Local Motion Vectors (LMVs) are considered within smaller frame regions during the process of motion estimation. Essentially, LMVs signify the offset of specific image regions between two consecutive frames. Thus, LMVs include both the intentional and the undesirable motion of the camera. Hilbert-Huang Transform (HHT) as a signal-processing tool that adaptively decomposes non stationary signals through the procedure of Empirical Mode Decomposition (EMD) into basis purposes called Intrinsic Mode Functions (IMFs).

***Corresponding author: Kalyani Patil,***
*Department of E&TC Engg, D. N. Patel College of Engineering, Maharashtra, India.*

The HHT combines the EMD and the Hilbert spectral analysis (HSA); the HSA contains the Hilbert trans- form of each IMF produced by the EMD procedure. 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. Unlike Fourier transform, the HHT can be used to show the original signal as a function of energy, time, and frequency. Several applications have been projected using the HHT including speech analysis, health monitoring, electroencephalographic data analysis, and seismic feature production. The projected video stabilizer deals with vertical or horizontal displacements meaning that translational motions are corrected.

## HHT METHOD

Let first, the data are decomposed into a finite number of IMF modules as a preprocessing stage. Hilbert transform is useful to all resulting IMFs in order to build the energy-time distribution. HHT having the modes: 1. EMD process 2. Hilbert Transform procedure 3. Jitter Motion Vector Estimation process.

## EMD PROCESS

EMD splits non stationary facts into locally non overlapping time-scale modules. EMD Algorithm: The sifting procedure is as follows: 1) Recognized the local extrema (maxima, minima) of the signal. 2) Join the maxima with an interpolation function, generating an upper envelope about the signal. 3) Join the minima with an interpolation function, producing a lower envelope about the signal. 4) Estimate the mean of the upper and lower envelopes. 5) From original signal subtract the local mean. Repeat on the residual. 6) The residue is treated as the new records and subjected to the same sifting procedure (start of outer loop). Local extrema are recognized. Local maxima are linked by cubic spline to form the upper envelope. The exact same method is followed for the local minima to build the lower envelope. If the mean of the upper and lower envelopes is chosen as $m1$ and the difference between the data $x(t)$ and $m1$ is the first element $h1$, then

$$h1(t)=x(t)-m1(t) \quad \ldots\ldots(1)$$

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

Where $U(t)$ and $L(t)$ are the local maxima and the local minima, respectively.

In the second round of sifting, $h1$ is preserved as the data or the first element. Then, a new mean is calculated with the same former technique. Considering that the new mean is $m11$, then

$$h11(t)=h1(t)-m11(t). \quad \ldots\ldots(3)$$

After repeating the sifting method up to $k$ times, $h1k$ is categorized as an IMF, meaning

$$h1k(t)=h1(k-1)(t)-m1k(t). \quad \ldots\ldots(4)$$

Let $h1k = c1$ be the first IMF from the records. $c1$ should cover the finest scale or the shortest period element of the records. $c1(t)$ is removed from the rest of the data so that the residual is calculated

$$r1(t)= x(t) - c1(t) \quad \ldots\ldots(5)$$

where $r1$ is the residue and it covers facts on longer period modules. The aforementioned method is repeated in order to gain all the subsequent $rw$ functions as follows:

$$rw(t)= rw-1(t) - cw(t) \quad \ldots\ldots(6)$$

where $w = 2, 3, \ldots, n$. A pictorial depiction of the EMD. The IMF modules must retain adequate physical sense of both amplitude and frequency modulations which can be attained by controlling the value of the sum of the difference calculated from two sequential sifting effects as

$$SD = \frac{\sum_{t=0}^{T} |h_{k-1}(t) - h_k(t)|^2}{\sum_{t=0}^{T} h_{k-1}^2(t)} \quad \ldots\ldots\ldots\ldots\ldots (7)$$

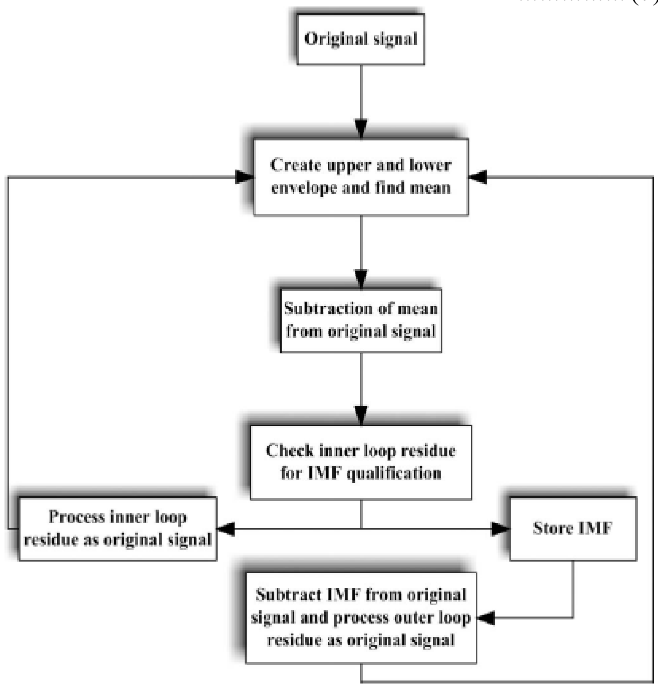


**Fig. 1. Flow for EMD Process**

where $T$ signifies the total number of the samples and $k$ signifies the iteration number of the sifting procedure. The whole EMD procedure is terminated if any of the following standards is satisfied:

1) when the residue $rw$ is a function with one unique extremum;
2) when the residue $rw$ becomes a monotonic function from which no IMF can be extracted.

Summing (5) and (6) yields the following calculation:

$$x(t) = \sum_{j=1}^{w} c_j + r_w \quad \ldots\ldots\ldots (8)$$

which designates wholeness, in that the sum of the IMFs and the residue improves the original signal. $cj$ is the $j$th IMF, and $w$ is the number of sifted IMFs. $rw$ can be interpreted as the overall trend of the original signal.

## HILBERT TRANSFORM PROCESS

The Hilbert transform of a real-valued function $x(t)$, which belongs to $Lp$, is given by

$$H\left(x(t)\right) = y(t) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau \qquad \text{………..(9)}$$

where $P$ represents the Cauchy principal value. The function $x(t)$ and its Hilbert transform $y(t)$ form an analytic signal $z(t)$ specified by

$$z(t) = x(t) + iy(t) = \alpha(t) e^{i\theta(t)} \qquad \text{……(10)}$$

where $\alpha(t)$ and $\theta(t)$ signify the instantaneous amplitude and instantaneous phase, respectively. $\alpha(t)$ and $\theta(t)$ are determined by the following calculations:

$$\alpha(t) = \sqrt{x^2(t) + y^2(t)} \qquad \text{……(11)}$$

$$\theta(t) = \tan^{-1}\left(\frac{y(t)}{x(t)}\right). \qquad \text{………(12)}$$

By explanation, the instantaneous frequency can be calculated as the derivative of the phase function given by

$$f(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt}. \qquad \text{..........(13)}$$

From (10) and (13), after Hilbert transformation, each IMF can be characterized by

$$c_j = \mathrm{Re}\left[\alpha_j(t) e^{i\int 2\pi f_j(t)dt}\right] \qquad \text{………..(14)}$$

and thus, the original data $x(t)$ can be mended as

$$x(t) = \mathrm{Re}\sum_{j=1}^{w} \alpha_j(t) e^{i\int 2\pi f_j(t)dt}. \qquad \text{……..…(15)}$$

Equation (14) gives both the amplitude and frequency of all element as a function of time, while (15) gives a frequency–time distribution of the amplitude, which is named the Hilbert spectrum $H(\omega, t)$.

**Proposed procedure**

**A. LMV Estimation**

First, a motion estimation technique is useful in order to express the LMV of an image sequence. Consequently, the estimated LMV is decomposed into a finite number of IMFs by applying the EMD technique. Each IMF is transformed using Hilbert transformation in order to express the energy content of every decomposed signal. Depending on the estimated energies, the last IMF to be measured as jitter is designated. Thus, the summation of all IMFs with the lower indices up to the identified IMF from the Hilbert transform approximates the undesirable jitter motion. Finally, the image arrangement is compensated according to the calculated sum in direction to construct a stabilized image flow.

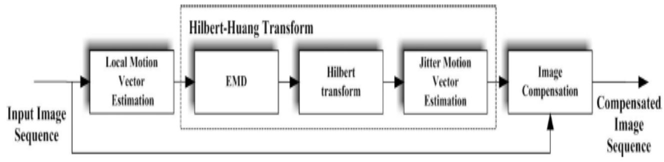


**Fig. 2. DIS method with HHT**

LMV estimation procedure since the first stage of the proposed technique is in common with the matching stage of a conventional DIS algorithm. Block-matching motion estimation is considered to be the most commonly used set of procedures for optical flow estimation, such as video coding. The current frame $ft$ is first separated into blocks of $M \times N$ pixels. Thus, the same motion vector is allocated to all pixels within the block. This motion vector is valued by searching for the best match block in a larger search window of $(M + 2dmx) \times (N + 2dmy)$ pixels centered at the same location in a reference frame $ft-\Delta t$, where $dmx$ and $dmy$ are the maximum allowed motion displacements in the horizontal and vertical directions, respectively. The most widely used matching criteria for lock-based motion estimation are the following: 1) sum of absolute difference (SAD); 2) MSE; and 3) matching pixel count. For LMV estimation, a specific image region is determined from every frame of the image sequence in order to identify its motion vector with respect to its reference. A square block of size $N \times N$ pixels is usually measured. The intensity value of the pixel at coordinate $(n1, n2)$ in the frame $k$ (current frame) is known by $s(n1, n2, k)$ where $(0 \le n1, n2 \le N - 1)$.

For the proposed technique, the SAD equivalent criteria were used; thus, the SAD value of a candidate block of $k - 1$ frame at a displacement $(i, j)$ in the reference frame is known by

$$SAD(i,j) = \frac{1}{N^2} \sum_{n1=0}^{N-1} \sum_{n2=0}^{N-1} \left[s(n_1, n_2, k) - s(n_1 + i, n_2 + j, k-1)\right]. \qquad \text{………..(16)}$$

The displacement vector is defined by

$$[d_1, d_2] = \arg \underbrace{\min}_{i,j} [SAD(i,j)] \qquad \text{…………..(17)}$$

where argmin is a task that results in the position where the $SAD(i, j)$ is minimized.

**B. EMD Process for LMVs**

The resulting vertical LMV between the successive frames is considered to be a time-varying signal signified as $x(t)$. The original signal $x(t)$ is actually a combination of the required camera motion and the trembled camera movement, known as jitter. By applying the EMD procedure to the $x(t)$, a finite number of IMFs with specific features will be created. A simplified application of the EMD method is given hereinafter. The mean envelope is designed (2) and subtracted from the initial signal $x(t)$. Depending on the resulting signal, the sifting procedure is repeatedly implemented until the stopping criterion is fulfilled, and thus, the resulting signal is signified as an IMF. The EMD procedure is terminated when the last signified IMF becomes a monotonic function.

### C. Designating the Unwanted Camera Motion

The EMD procedure splits the initial LMV signal into a number of sub signals with specific features. Lower IMF indices indicate higher frequencies. EMD process divides the initial signal into a finite number of sub signals based on their frequencies, the last IMF that includes jitter motion components. The power of each IMF is proportional to the amplitude of each sample and is given by

$$P_i = \sum_{t=0}^{K} \alpha_i^2(t) \quad \text{……….(18)}$$

Where $\alpha i$ represents the calculated amplitude of an IMF's sample, $i = 1, \ldots, w + 1$ denotes the corresponding IMF, and variable $t$ represents the time step, meaning the number of a frame of the image sequence. For $i = w + 1$, variable $\alpha i$ represents the amplitude of the residue of the signal. The IMF with the higher index and the lower energy content corresponds to the last IMF which includes jitter components. Thus, the required threshold is defined by

$$d = \arg\min[P_i] \quad \text{……(19)}$$

where $i = 1, \ldots, w + 1$. Arg min denotes a function that results in the position where $Pi$ is minimized.

The summation of all the IMFs up to the indicated threshold $d$ designates the undesirable camera motion, and thus, jitter can be calculated by (20).

$$X_J(t) = \sum_{i=1}^{d} c_i(t) \quad \text{………(20)}$$

The sum of the remained IMFs as well as the residue defines the intentional camera motion, as (21) designates.

$$X_G(t) = \sum_{i=d}^{w} c_i(t) + r_w \quad \text{……….(21)}$$

where $XJ\ (t)$ and $XG(t)$ denote the jitter and the intentional camera motion, respectively.

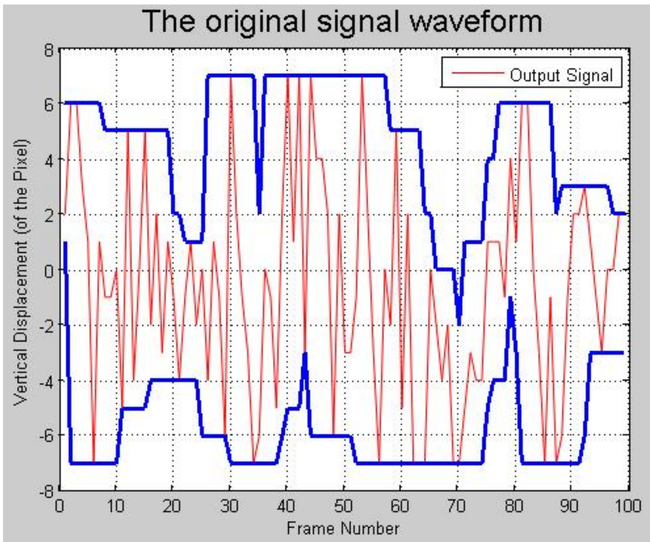


**Fig. 3. LMV signals with upper and lower envelopes**

## Conclusion

DIS process based on the HHT has been offered. LMV and EMD, which is first block of HHT has been offered. By using Image sequences accomplish the meaning of two essential motion such as horizontal and vertical LMV. EMD effects also achieve such as upper and lower envelope, mean and first IMF. To confirm the effectiveness of the projected DIS technique, the experimental outcomes of projected system are offered in this article. The horizontal displacements and vertical displacements are existing; the procedure for horizontal motions is accurately the identical. In order to estimate the performances of the method, three different image sequences are managed. In this paper, three image orders are used to discover out IMF and horizontal, vertical LMV. The upper and lower envelope are produced using the local maxima and local minima of initial signal. Experimental results have confirmed that the proposed method can successfully decompose the two camera motions, and therefore, the image sequence can be effectively remunerated since the jitter motion has been well-defined.

## REFERENCES

Bosco A., A. Bruna, S. Battiato, G. Bella, and G. Puglisi, "Digital video stabilization through curve warping techniques," *IEEE Trans,* May 2008.

Cardani B., "Optical image stabilization for digital cameras," IEEE Control Syst. Mag., vol. 26, no. 2, pp. 21–22, Apr. 2006.

Chen H., C. Liang, Y. Peng, and H. Chang, "Integration of digital stabilizer with video codec for digital video cameras," *IEEE Trans. Circuits Syst. Video Technol*., vol. 17, no. 7, pp. 801–813, Jul. 2007.

Konstantinos Ioannidis and Ioannis Andreadis "A Digital Image Stabilization Method Based on the Hilbert–Huang Transform" *IEEE Transactions On Instrumentation And Measurement*, Vol. 61, No. 9, September 2012.

Lu Z., B. Liu, and L. Shen, "Speech endpoint detection in strong noisy environment based on the Hilbert–Huang transform," in Proc. *IEEE Int. Conf. Mechatron. Autom*., 2009, pp. 4322–4326.

Sheng-Che Hsu, Sheng-Fu Liang, and Chin-Teng Lin" A Robust Digital Image Stabilization Technique Based on Inverse Triangle Method and Background Detection" *IEEE Transactions on Consumer 336 Electronics*, Vol. 51, No. 2, May 2005.

Sheng-Che Hsu, Sheng-Fu Liang, Kang-Wei Fan, and Chin-Teng Lin, Fellow, Ieee" A Robust In-Car digital Image Stabilization Technique"IEEE Transactions On Systems, Man, And Cybernetics—Part C: *Applications And Reviews,* Vol. 37, No. 2, March 2007.

Zhang Y., M. Xie, L. Mao, and D. Tang, "An adaptive digital image stabilization based on empirical mode decomposition," in Proc. *Int. Conf. Commun. Circuits Syst.,* 2010, pp. 327–330.

*******