

Hilbert–Huang transform plugin

Author: Juraj Jurco

Location: Plugins -> Analysis -> Empirical Mode Decomposition

The plugin provides Hilbert–Huang transform (HHT) of given signal as described in papers [1], [2]. HHT is composed of Empirical mode decomposition (EMD) and computation of instantaneous frequency. EMD is based on decomposition of signal into Intrinsic Mode Functions (IMF) and instantaneous frequency is computed out of those IMF functions. An IMF is defined as a function that satisfies the following requirements:

1. In the whole data set, the number of extrema and the number of zero-crossings must either be equal or differ at most by one.
2. At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

It represents a generally simple oscillatory mode as a counterpart to the simple harmonic function. By definition, an IMF is any function with the same number of extrema and zero crossings, whose envelopes are symmetric with respect to zero.

Usage



Letters at this list corresponds with letters in the image.

- A. First attach single channel to be processed
- B. **EMD Stopping criteria:** default stopping criteria is definition of maximal number of IMF functions to be computed. If e.g. maximal number of IMF is set to ten, it produces 11 output signals, where first 10 are IMF and the last one is residuals. You may define also stopping criteria by setting values drop of IMF against original signal. Default setting (by default not used) is: when 95% of values of original signal drops more than 10000x EMD decomposition will stop.

- C. **IMF Stopping criteria:** define stopping criteria for IMF decomposition. There are two different criteria implemented:
1. **Single SD:** When Standard Deviation (SD) during IMF decomposition process from the two consecutive sifting results is below this value. It is suggested to keep this value between 0.2 – 0.3, as described in [1].
 2. **2sd + alpha:** There was introduced a new criterion based on 2 thresholds θ_1 and θ_2 , aimed at guaranteeing globally small fluctuations in the mean while taking into account locally large excursions. You define α (as default it is 0.05; watch out we use it already inverted, so you set in the form “1- α ” for easier understanding in the form), θ_1 and θ_2 . It applies that $(1 - \alpha)$ values are below θ_1 and none of SD values is over θ_2 value. One can typically set $\alpha \approx 0.05$, $\theta_1 \approx 0.05$ and $\theta_2 \approx 10 \times \theta_1$ what are also default values of the plugin.
- D. You may set also maximal number of trials to generate IMF function. If this number of trials is verreached and function doesn't meet IMF criteria, searching is interrupted, last generated IMF is taken into account and process continues as it correct IMF.
- E. You may choose whether to process whole signal, if you want to export generated IMF as new channels or add them as datacaches to processed channel.
- F. When all is set, just press process and HHT transform will start.
- G. Originaly transformed signal is printed in this area.
- H. When transform is finished, you may see list of generated IMF function here.
- I. When item in H is checked, instantaneous frequency is shown in this area. The colour of the painted instantaneous frequency is equal with the color of “IMF level X” in the box H.

Batch processing

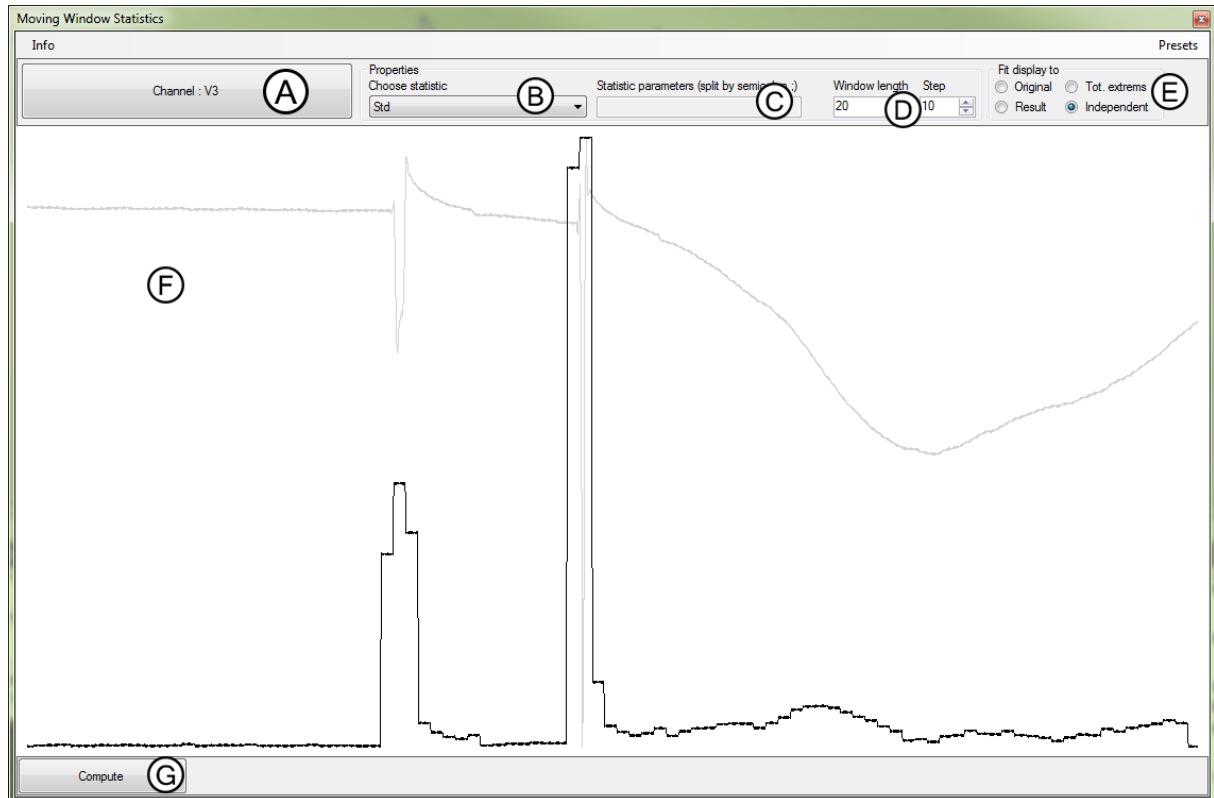
Plugin does not support batch processing.

- [1] N. Huang, Z. Shen, S. Long, M. Wu, H. SHIH, Q. ZHENG, N. Yen, C. Tung, a H. Liu, „The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proc. R. Soc. A Math. Phys. Eng. Sci.*, roč. 454, č. 1971, s. 995, 903, 1998.
- [2] G. Rilling, P. Flandrin, P. Gon, a D. Lyon, „ON EMPIRICAL MODE DECOMPOSITION AND ITS ALGORITHMS".

Moving Statistical Window plugin

Author: Juraj Jurco

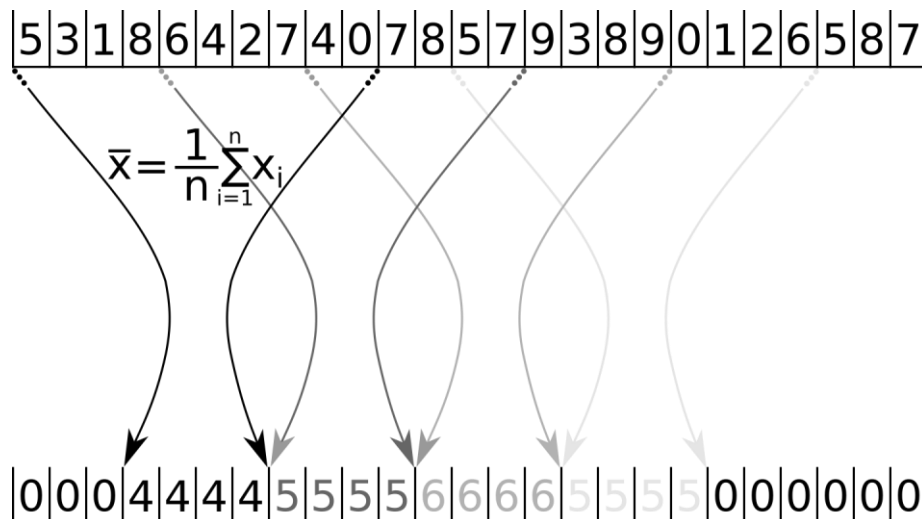
Location: Plugins -> Statistics -> Moving Window Statistics



- A. First attach channel(s) to be processed. Preview window (F) shows only the first attached channel
- B. Choose statistic you would like to compute
- C. Some statistics has optional parameters (see below). Before any statistic is computed you have to set those parameters
- D. Specify Window Size and Step
- E. Select how preview display has to be specified
- F. Grey signal is original signal from the first selected channel. Black signal is a preview of selected statistic over the first attached channel.
 - a. NaN and infinite values: When the result of the statistic is NaN or infinity, black link is missing at this place.
- G. When you press compute, selected statistic with specified parameters is computed over all selected channels. Computed statistic is stored in the channels datacache.

Window computation

Sliding window of specified size and step moves over given signal and selected statistic is computed from values of this sliding window. Example below shows computation process of the sliding window of size 10 with the step 4. For the sake of simplicity it is shown on MEAN statistic.



Definition of variables and statistics

- **Window Size:** length of area from which the statistic is computed. For some statistics this area has to be larger than 1. Window size cannot be lower than “Step” variable.
- **Step:** length of the jump of sliding window
- **Statistic:** Supported statistics are:
 - **Min:**
 - Requires NO parameter
 - Mathematical expression: $Y = \min(A)$
 - **Max:**
 - Requires NO parameter
 - Mathematical expression: $Y = \max(A)$
 - **Sum:**
 - Requires NO parameter
 - Mathematical expression: $Y = \sum_{i=1}^n a_i$
 - **Product:**
 - Requires NO parameter
 - Mathematical expression: $Y = \prod_{i=1}^n a_i$
 - **Power Sum:**
 - Required 1 parameter:
 - p – power
 - Mathematical expression: $Y = \sum_{i=1}^n a_i^p$
 - **Normalized power sum:**
 - Requires 2 parameters:
 - p – power
 - norm – normalization factor, currently you can set only constant value. Maybe in the future some other statistic may be more appropriate.
 - Mathematical expression: $Y = \sum_{i=1}^n (a_i + norm)^p$
 - **Mean:**
 - Requires NO parameter
 - Mathematical expression: $Y = \frac{1}{n} \sum_{i=1}^n a_i$
 - **Variance:**
 - Requires NO parameter

- Mathematical expression: $Y = \frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^2$
- **Unbiased variance:**
 - Requires NO parameter
 - Mathematical expression: $Y = \frac{1}{n-1} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^2$
- **Standard Deviation:**
 - Requires NO parameter
 - Mathematical expression: $Y = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^2}$
- **X-th moment (1)**
 - Requires 1 parameter:
 - X – moment we would like to compute
 - Mathematical expression: $Y = \frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^x$
- **Unbiased x-th moment (1)**
 - Requires 1 parameter
 - X – moment we would like to compute
 - Mathematical expression: $Y = \frac{1}{n-1} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^x$
- **Kurtosis**
 - **Kurtosis – g2**
 - Requires NO parameter
 - Mathematical expression: $Y = \frac{4^{th} \text{ moment}}{\text{variance}^2} = \frac{\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^4}{\left(\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^2 \right)^2}$
 - Description:
 - **Kurtosis – G2**
 - Requires NO parameter
 - Mathematical expression: $Y = \frac{(n-1) \times ((n+1) \times \text{kurtosis} + 6)}{(n-2) \times (n-3)} =$

$$\frac{n-1}{(n-2) \times (n-3)} \times \left((n+1) \times \frac{\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^4}{\left(\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^2 \right)^2} + 6 \right)$$
 - **Kurtosis – b2**
 - Requires NO parameter
 - Mathematical expression:

$$Y = \frac{4^{th} \text{ moment}}{(\text{unbiased variance})^2} = \frac{\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^4}{\left(\frac{1}{n-1} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i \right)^2 \right)^2}$$
 - **Summary of the differences between the formulasⁱ**
 1. Using **g2**, a normal distribution has a kurtosis value of 3 whereas in the formulas involving a correction term -3 (i.e. G2 and b2), a normal distribution has an excess kurtosis of 0.
 2. **G2** is the only formula yielding unbiased estimates for normal samples (i.e. the expectation of G2 under normality is zero, or $E(G2) = 0$).
 3. For **large samples**, the difference between the formulas are negligible and the choice does not matter much.

4. For **small samples from a normal distribution**, the relation of the three formulas in terms of the mean squared errors (MSE) is: $mse(g2) < mse(b2) < mse(G2)$. So $g2$ has the smallest and $G2$ the largest (although only $G2$ is unbiased). That is because $G2$ has the largest variance of the three formulas: $Var(b2) < Var(g2) < Var(G2)$.
5. For **small samples from non-normal distributions**, the relation of the three formulas in terms of bias is: $bias(G2) < bias(g2) < bias(b2)$. In terms of mean squared errors: $mse(G2) < mse(g2) < mse(b2)$. So $G2$ has the smallest mean squared error and the smallest bias of the three formulas. $b2$ has the largest mean squared error and bias.
6. For **large samples ($n > 200$) from non-normal distributions**, the relation of the three formulas in terms of bias is: $bias(G2) < bias(g2) < bias(b2)$. In terms of mean squared errors: $mse(b2) < mse(g2) < mse(G2)$.

○ **Skew**

- Requires NO parameter

- Mathematical expression: $Y = \frac{mean^3}{variance^3} = \frac{\left(\frac{1}{n} \sum_{i=1}^n a_i\right)^3}{\left(\frac{1}{n} \sum_{i=1}^n \left(a_i - \frac{1}{n} \sum_{i=1}^n a_i\right)^2\right)^3}$

○ **Min/Max slope:** Finds minimal and maximal value in the window and computes slope between those two values.

- Requires NO parameter

- Mathematical expression:

- $[min.t, min.v] = min(A)$; $[max.t, max.v] = max(A)$; where t coordinate represents temporal position of value v in the array A .

- $Y = \begin{cases} 0 & R.t - S.t = 0 \\ \frac{max.y - min.y}{max.x - min.x} & else \end{cases}$

○ **Min/Max slope degree**

- Requires NO parameter

- Mathematical expression:

- $[min.t, min.v] = min(A)$; $[max.t, max.v] = max(A)$; where t coordinate represents temporal position of value v in the array A .

- $Y = \begin{cases} 0 & R.t - S.t = 0 \\ \tan\left(\frac{max.y - min.y}{max.x - min.x}\right) & else \end{cases}$

○ **Two percentile diff (2)**

- Requires 2 parameters: k_1 and k_2

- Mathematical expression:

- $P_{k_1} = percentile(k_1, A)$; $P_{k_2} = percentile(k_2, A)$;
- $Y = P_{k_2} - P_{k_1}$

○ **Signal Energy**

- Requires NO parameter

- Mathematical expression: $Y = \sum_{i=1}^n a_i^2$

○ **Signal Power**

- Requires NO parameter

- Mathematical expression: $Y = \frac{1}{2n+1} \sum_{i=1}^n a_i^2$

Batch processing

Plugin does not support batch processing.

ⁱ Source: <http://stats.stackexchange.com/questions/61740/differences-in-kurtosis-definition-and-their-interpretation>