Digital Signal Processing ••• (••••) •••–•••

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Teager–Kaiser energy methods for signal and image analysis: A review

A.O. Boudraa [a], F. Salzenstein [b]

[a] IRENav, Ecole Navale/Arts-Métiers ParisTech, IRENav (EA 3876), BCRM Brest, CC 600, 29240 Brest, France
[b] ICube, Université de Strasbourg, 23 rue du Loess, BP 20 CR, 67037 Strasbourg Cedex 2, France

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Teager–Kaiser energy operator
Higher order differential operators
Multi-dimensional energy operator
Cross-Ψ_B-energy operator
AM–FM Model
Signal demodulation

ABSTRACT

This paper provides a review of the Teager–Kaiser (TK) energy operator and its extensions for signals and images processing. This class of operators possesses simplicity and good time-resolution and is very efficient in instantaneously estimating AM–FM signals and images. We point out the importance of the concept of energy from the point of view of the generation of the signal. More precisely, we emphasize the importance of analyzing signals from the point of view of the energy of the system needed to produce them. We show how this class of TK energy operators can be used to estimate useful features for signals and images analysis in time, space and frequency domains such as instantaneous frequency, second-order moment frequency, coherence function or spatial envelope and phase. We also show the importance of the higher derivative order of TK energy operator in terms of demodulation for both mono and multi-dimensional signals. Most of the developed tools around TK energy operators deal with real and complex-valued signals and some of them extended to multi-dimensional case. Due to their low complexity and their instantaneous-adapting nature, the class of TK energy operators offers valuable processing tools for time and frequency analysis.

1. Introduction

Signals may represent a broad variety of phenomena. In many applications, signals are directly related to physical quantities capturing energy and power in a physical system. The concept of signal energy is of primary importance in the design of continuous and discrete domain systems [1–7]. This work is interested in signals provided by sensors and thus, to the energy associated with these signals. In the real world, we always transmit signals with finite total energy $0 < E_x < +\infty$ (or with finite average power) representing the amount of energy contained in signal $x(t)$. The quantity E_x should be independent of the method used to calculate it. Engineers refer to such signals as having finite total energy, although E_x is not necessarily the *physical* energy of the signal $x(t)$. For example, the total energy of the source system modeled as a mass suspended by a spring of a constant stiffness required to produce a simple undamped harmonic oscillation is calculated by the sum of the kinetic energy of the mass and the potential energy in the spring. By studying the second order differential associated to this harmonic oscillator, it is easy to show that a simple sinusoidal varies as a function of both amplitude and oscillation frequency of the signal $x(t)$, which is quite different from simple squaring of the signal magnitude, $x^2(t)$. It is this source modeling that is used for characterizing $x(t)$ by amplitude and frequency.

In their work on non-linear speech modeling, Herbert and Shushan Teager pointed out the dominance of modulation as a process in the speech production [8,9]. Based on the Teager's work, Kaiser proposed an energy measure that includes both the amplitude and the frequency of the signal [3]. This measure is often referred to as the Teager–Kaiser (TK) energy operator. Using the conventional view of the energy, it is easy to see that two tones at 10 Hz and 1000 Hz of unit-amplitude have the same energy. However, the energy required to produce the signal of 1000 Hz is much greater than that for the 10 Hz signal [3]. Using TK definition of energy, the two tones show different energy. This definition highlights the concept of signal energy from the point of view of the generation of the signal and emphasizes the importance of analyzing signals from the energy aspect of the system needed to produce them. In the following, the theory behind TK energy operator and its different extensions as tools for time, frequency and TF analysis of signals, and for estimation of envelope and phase for images is presented. This class of operators is illustrated on simulated and real data.

2. Teager–Kaiser energy operator

In its continuous form, TK energy operator, noted Ψ_c, when operating on continuous-time signal $x(t)$ is given by [3]

E-mail addresses: boudra@ecole-navale.fr (A.O. Boudraa), f.salzenstein@unistra.fr (F. Salzenstein).

https://doi.org/10.1016/j.dsp.2018.03.010

$$\Psi_c[x(t)] \triangleq \left(\frac{dx(t)}{dt}\right)^2 - x(t)\left(\frac{d^2x(t)}{dt^2}\right) = \dot{x}^2(t) - x(t)\ddot{x}(t) \quad (1)$$

where $\dot{x}(t)$ and $\ddot{x}(t)$ are the first and the second derivative of $x(t)$ with respect to time t respectively. When Ψ_c is applied to signals generated by a simple harmonic oscillator (mass-spring oscillator of constant stiffness), it can track the oscillator's energy (per half unit mass) which is proportional to the squared product of the oscillation amplitude and frequency. For narrowband signal $x(t)$ and under realistic conditions,[1] $\Psi_c[x(t)]$ approximately estimates the energy of the source producing the oscillation $x(t)$. Using, for example, the backward approximation of the time derivatives (Eq. (1)), the discrete-time counterpart of Ψ_c becomes

$$\Psi_d[x(n)] = \frac{x^2(n-1) - x(n)x(n-2)}{(\Delta t)^2} \quad (2)$$

where $x(n) = x(n\Delta)$ with $\Delta t = 1/f_s$ and $n \in \mathbb{N}$; f_s is the sampling frequency. Equation (2) is scaled and centered yielding

$$\Psi_d[x(n)] = x^2(n) - x(n-1)x(n+1) \quad (3)$$

The energy operator spans three adjacent samples of $x(t)$ and is still a very local property of $x(t)$. This input–output relation is called by Kaiser "Teager's algorithm" [3]. For a tone $x(n) = A\cos(\omega n + \phi_0)$, TK operator yields

$$\Psi_d[x(n)] \approx A^2\omega^2 \quad (4)$$

with $\omega < \pi/4$. This property is also valid for damped sinusoids of the form

$$x(n) = r^n\cos(\omega n + \phi_0) \quad (5)$$

The operator $\Psi_d[x(n)]$ approximates the product of squared amplitude and frequency of $x(n)$ and can be termed as frequency-weighted energy. The non-linear operators Ψ_c and Ψ_d were developed by Teager [9,10] and introduced by Kaiser [3,11]. Note that $\Psi_d[x(n)]$ is independent of the initial phase of $x(t)$, symmetric and capable of responding very quickly to changes in amplitude and frequency of $x(t)$ [3]. Furthermore, it is robust even when $x(n)$ passes through zero, as no division operation is required. The operator $\Psi_d[x(n)]$ offers excellent time resolution because only three samples $(x(n-1), x(n), x(n+1))$ are required for the energy computation at each time instant $n\Delta t$, so it has good adaptability to the instantaneous changes in $x(t)$. This why TK operator is well adapted, for example, for measuring formant modulation. The response of $\Psi_d[x(n)]$ is nearly instantaneous and this operator can be easily implemented in DSP processors, due to its low computation cost and extremely low requirements of memory storage.

The TK operator has been discussed from the mathematical point of view. For example, it has been characterized algebraically by giving expressions of the outputs corresponding to different types of combination (e.g. linear, algebraic) of input signals and classified the root and pre-constant signals of the operator [12]. The matrix framework of the operator has been introduced in [13] by interpreting it as the determinant of a Toeplitz matrix containing the signal and its derivatives:

$$\Psi_c[x(t)] = \begin{vmatrix} \dot{x}(t) & x(t) \\ \ddot{x}(t) & \dot{x}(t) \end{vmatrix}; \quad \Psi_d[x(n)] = \begin{vmatrix} x(n) & x(n+1) \\ x(n-1) & x(n) \end{vmatrix}$$

The determinant is time-invariant for a signal with constant frequency. If such matrix is generalized to an $M \times M$ Toeplitz matrix by adding delayed $x(n)$ up to $x[n \pm (M-1)]$, the determinant is also time-invariant [13,14]. This time-invariant property can be exploited using probe tones to uncover the amplitudes and frequencies of multiple sinusoids [13]. Using this matrix framework, the output of the TK operator is interpreted as the measured energy corresponding to the square of the eigenvalue of its underlying energy matrix, a notion analogous to that seen in quantum mechanics [15]. An extension of this matrix framework to generalized TK operator has been introduced in [16].

2.1. Multiresolution Teager–Kaiser energy operator

To resolve two closely spaced tones (AM signal), an extended version of TK operator has been developed by introducing a lag parameter, k, in the expression (3) as follows [17]

$$\Psi_{d_k}[x(n)] = x^2(n) - x(n-k)x(n+k) \quad (6)$$

where $k \in \mathbb{N}$ is a resolution parameter. For a dual tones with frequencies f_1 and f_2, an optimal choice of k parameter enhances the difference $(f_1 - f_2)$ or the sum frequency $(f_1 + f_2)$ [17]. For a sinusoidal input $x(n) = A\cos(\omega n)$, the operator output is given by

$$\Psi_{d_k}[x(n)] \approx (A\sin(\omega k))^2 \quad (7)$$

where $k = \pi/2\omega$ $(\omega = 2\pi f/f_s)$ gives the maximum output. Thus, each frequency has its own optimum resolution parameter which maximizes the relation (7). The resolution parameters M_d and M_s are optimally chosen to enhance respectively the sum and the difference frequencies. Approximate relations of the optimum lag parameters for $(f_1 - f_2)$, M_d, and $(f_1 + f_2)$, M_s, are given by [17]

$$M_d = \left\lfloor \frac{0.5 f_s}{f_1 + f_2} + 0.5 \right\rfloor; \quad M_s = \left\lfloor \frac{i \times f_s}{f_1 + f_2} + 0.5 \right\rfloor \quad (8)$$

where $\lfloor . \rfloor$ is the floor function and $i \in \mathbb{N}^*$. Larger values of i produce greater signal enhancement, but can introduce longer delay [17]. We illustrate the interest of $\Psi_{d_k}[x(n)]$ operator using a signal of two closely spaced tones, $x(t)$:

$$x(t) = \sin(2\pi f_1 t) + \sin(2\pi f_2 t) + n(t) \quad (9)$$

where $t \in [0, 1]$, $n(t)$ is a white Gaussian noise of signal to noise ratio (SNR) of 26 dB, $f_1 = 435$ Hz, $f_2 = 400$ Hz and $f_s = 8$ kHz. The parameter i is set to 2. Using relations (8) we get $M_d = 5$ and $M_s = 19$ (Figs. 1(a)–(c)). To identify the frequency components at $f_1 - f_2 = 35$ Hz and $f_1 + f_2 = 835$ Hz, a frequency analysis of $\Psi_{d_k}[x(n)]$ and $\Psi_d[x(n)]$ operators is performed by calculating their power spectral density (PSD). $\Psi_{d_k}[x(n)]$ operator is applied to $x(t)$ with $k = 1$ (classical TK operator), $k = M_d$ and $k = M_s$. Both Ψ_d and Ψ_{d_5} operators reveal a strong frequency component at the difference frequency of 35 Hz, but not corresponding component at the summation frequency of 835 Hz. Note that the spectral peak at 35 Hz is more prominent in the case $k = 5$ compared to result provided by TK operator $(k = 1)$. This result shows the interest to combine TK operator with a lag parameter $k \neq 1$, even in the presence to noise, to enhance the difference frequency. The spectrum of $\Psi_{d_{19}}$ shows a clear peak at 835 Hz and none at 35 Hz. The use of large value of the lag parameter accentuates the summation frequency. This example confirms the findings reported in [17] and show that changing the lag parameter k can result in enhancing the summation or difference frequencies of two component tones. Furthermore, noise with SNR = 26 dB does not degrade the prominence of the summation or the difference frequency component. To

[1] The amplitude and the frequency of the signal do not vary too fast (rate of change) or too greatly (range of value) with time compared to carrier frequency.