Speeding up estimation of the Hurst exponent by a two-stage procedure from a large to small range

1. Introduction

In many fields, for example, society (Gao *et al.*, 2012; Perc, 2012; Perc, 2013; Petersen *et al.*, 2012), business (Domino, 2012; Fernández-Martínez *et al.*, 2013; Rejichi and Aloui, 2012; Rostek and Schöbel, 2013), medicine (Chang *et al.*, 2000; Chang *et al.*, 2007; Huang and Lee, 2009; Lin *et al.*, 2013) and nature (Gonçalves and Bruno, 2013; Hagerhall *et al.*, 2004; Mandelbrot, 1983; Pentland, 1984; Wang *et al.*, 2011; Zuñiga *et al.*, 2014), signals often look like self-similarity with strong long-term correlations. They look like disorganized realizations and seem to be difficult to describe them, but fortunately many of which can be explained by only one parameter, called the Hurst exponent or its corresponding fractal dimension. Therefore, how to accurately and efficiently estimate the Hurst exponent is a big issue.

Two categories of estimators for the Hurst exponent are the non-modeling class and the modeling class. The non-modeling estimators include the box-counting method (Bruce, 2001; Chen *et al.*, 1993; Jin *et al.*, 1995; Sarkar and Chaudhuri, 1992; Sarkar and Chaudhuri, 1994), the rescaled range (R/S) analysis (or the R/S statistic) (Beran, 1994), the detrended fluctuation analysis (DFA) (Peng *et al.*, 1992 and 1994; Biswas *et al.*, 2012) and the wavelet-based methods (Hansen *et al.*, 1998; Rehman and Siddiqi, 2009). On the other hand, the modeling class includes the variance method, the moving-average (MA) method, and the autoregressive (AR) method, where signals are either modeled as fractional Brownian motion (FBM) or as fractional Gaussian noise (FGN). Since modeling methods can provide physical meanings of signals and broadly speaking, their estimation accuracy is much higher than non-modeling ones, they are often adopted by engineers who engage in the field of statistical

signal processing.

The Hurst exponent can be viewed as a key feature of FBM and FGN. These two processes are related to each other; the increment process of FBM is FGN, and conversely FBM is the accumulation process of FGN. In general, FGN is used quite more frequently than FBM because it is strict-sense stationary and the behavior of its power spectral density (PSD) approaches the form of λ^{1-2H} (Lundahl *et al.*, 1986; Mandelbrot and Van Ness, 1968), where λ represents frequency; however, FBM is statistically self-similar nonstationary and its PSD changes over time (Flandrin, 1989; Lundahl *et al.*, 1986).

In practical applications, data are collected in discrete time; their corresponding types of FBM and FGN are called discrete-time fractional Brownian motion (DFBM) and discrete-time fractional Gaussian noise (DFGN), which is a regular process (Chang and Chang, 2002). DFBM and DFGN are often used to model natural and biomedical signals (Chang *et al.*, 2000; Chang *et al.*, 2007; Liu and Chang, 1997; Lundahl *et al.*, 1986). To explain the characteristics of these signals, their Hurst exponents are first estimated and then, if necessary, the corresponding fractal dimensions are calculated according to the relation $D = 2 - H$ (Falconer, 1990; Hastings and Sugihara, 1993), where D is Hausdorff dimension or box dimension.

As we know, the maximum likelihood estimator (MLE) (Lundahl *et al.*, 1986) has optimal accuracy, but it spends huge time computing the inverse of its autocovariance matrix, and it is likely to encounter computational problems, especially for H closer to 1 (Beran, 1994), hence computational inaccuracy. Therefore, the MLE generally has a higher theoretical value than its practicability.

Then the Whittle estimator (Beran, 1994; Kay, 1993; Schonhoff and Giordano, 2006; Taqqu *et al.*, 1995), an approximate MLE, was developed to reduce computational cost and even overcome the computational instability while remaining accuracy at a satisfactory level.

Other speedier estimators with different degrees of accuracy were developed, such as the variance method (Chang *et al.*, 2012; Lundahl *et al.*, 1986; Mandelbrot and Van Ness, 1968), moving-average (MA) method (Liu and Chang, 1997) and autoregressive (AR) method (Chang, 2014b; Chang and Chang, 2002; Chang *et al.*, 2014).

Recently, Chang (Chang, 2014a) introduced the Levinson algorithm (Haykin, 1989; Kay, 1988) and Cholesky decomposition (Haykin, 1989) into the MLE to raise computational efficiency, which is simply called the fast MLE. Moreover, four possible situations of two parameters (mean and variance) in the probability density function (PDF) have even been considered. Experimental results show that the fast MLE has completely surmounted computational problems and greatly minimized the computational cost. The efficiency of the fast MLE, which grows exponentially as the data size increases, will greatly raise its practicability.

Although the speed of the fast MLE has been boosted to a considerable degree, any promotion on the computational performance of the fast MLE is always popular with users, especially in the field of demanding both accuracy and efficiency. In this paper, a two-stage procedure combining two different estimators to estimate the Hurst exponent from a large to small range (simply called a two-stage estimator) will be developed. To achieve substantial benefits, the speed of the two-stage estimator should be quicker than the faster of two estimators, whereas its accuracy is still remained at almost the same level as the more accurate.

In fact, any integration of two estimators is feasible only under the condition that the first-stage estimator with accuracy as high as possible is quicker than the other and the second-stage estimator, which must be estimated by searching a finite interval for the optimal Hurst exponent, is more accurate than the other. For possibly best accuracy, the DI method (Chang, 2014b) is currently ideal for the first-stage estimator and the fast MLE (Chang,

2014a) is the best candidate for the second-stage estimator.

The rest of this paper is organized as follows. Section 2 gives a brief description of mathematical preliminaries. Section 3 introduces two Hurst exponent estimators applied in a two-stage procedure. Section 4 provides the algorithm of implementing the two-stage procedure. Section 5 discusses experimental results. Finally, Section 6 draws some conclusions from a few facts.

2. Mathematical Preliminaries

FBM is a generalized Brownian motion, which is represented by $B_H(t)$, short for $B_H(t,\zeta)$ where t is continuous time and the outcome ζ is from some sample space, S. As defined by Mandelbrot and Van Ness (Mandelbrot and Van Ness, 1968), FBM is expressed as follows:

$$B_H(0)=b_0,$$

$$B_H(t)-B_H(0)=\frac{1}{\Gamma(H+1/2)}\left\{\int_{-\infty}^{0}\left[(t-s)^{H-1/2}-(-s)^{H-1/2}\right]dB(s)+\int_{0}^{t}(t-s)^{H-1/2}dB(s)\right\}, \qquad (1)$$

where H is the Hurst exponent, a real value lying between 0 and 1. It is obvious that FBM is equal to the ordinary Brownian motion at $H=0.5$. For ease of analysis, equation (1) is usually written in a more symmetric form (Mandelbrot and Van Ness, 1968):

$$B_H(t_2)-B_H(t_1)=\frac{1}{\Gamma(H+1/2)}\left\{\int_{0}^{t_2}(t_2-s)^{H-1/2}dB(s)-\int_{0}^{t_1}(t_1-s)^{H-1/2}dB(s)\right\}. \qquad (2)$$

Since FBM is not a stationary process which makes analysis difficult, the Wigner-Ville spectrum (WVS) of a nonstationary process is often adopted to evaluate time-dependent spectral characteristics of signals, whose formula is as follows (Flandrin, 1989):

$$f_{B_H}(\Omega,t)=\left(1-2^{1-2H}\cos 2\Omega t\right)\frac{1}{|\Omega|^{2H+1}}. \qquad (3)$$

Although FBM is nonstationary, the increment process of FBM, denoted by $B'_H(t)$ and

termed FGN, is stationary and self-similar with parameter H (Mandelbrot and Van Ness, 1968), making estimation of the Hurst exponent easier.

It is inevitable for practical applications that continuous data are sampled and then discrete data are gathered; discrete-time FBM (DFBM) are expressed by $B_H[n]=B_H(nT_s)$, where T_s is the sampling time. The increment process of DFBM, called DFGN, is denoted by $X_H[n]=B_H[n]-B_H[n-1]$.

It is beneficial for estimation that DFGN is exactly a stationary and normally distributed process with mean zero and variance $r_H[k]$, called the autocorrelation function (ACF), which is as follows:

$$r_H[k]=E\{X_H[n+k]X_H[n]\}=\frac{\sigma^2}{2}\left(|k+1|^{2H}-2|k|^{2H}+|k-1|^{2H}\right), \tag{4}$$

where $\sigma^2=\mathrm{var}(X_H[n])$ (Lundahl *et al.*, 1986; Samorodnitsky and Taqqu, 1994). The ACF provides a time-domain description of the second moment of DFGN and its behavior much approaches $k^{2H-2}=k^{-\alpha}$, $\alpha\in(0,2)$ (Samorodnitsky and Taqqu, 1994) as k gets larger. On the other hand, the PSD or the discrete-time Fourier transform of the ACF is often used as a frequency-domain description (Haykin, 1989):

$$f(\omega)=\sum_{k=-\infty}^{\infty}r_H[k]e^{-ik\omega}, \tag{5}$$

where ω is the angular frequency. Conversely, the ACF can be derived from the inverse discrete-time Fourier transform (Haykin, 1989):

$$r_H[k]=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(\omega)e^{ik\omega}d\omega,\quad k=0,\pm1,\ldots. \tag{6}$$

Based on the properties of DFGN, the PDF of DFGN is written as follows (Lundahl *et al.*, 1986):

$$p(\mathbf{x};H)=\frac{1}{(2\pi)^{N/2}|\mathbf{R}|^{1/2}}\exp\left\{-\frac{1}{2}\mathbf{x}^T\mathbf{R}^{-1}\mathbf{x}\right\}, \tag{7}$$

where $\mathbf{x} = [X_H[0] \quad X_H[1] \quad \cdots \quad X_H[N-1]]^T$ is a data set, $\mathbf{R}$ the autocovariance matrix of size $N \times N$ and $|\mathbf{R}|$ its determinant, where, $\mathbf{R} = E[\mathbf{x}\mathbf{x}^T]$. Without loss of generality, the logarithm of the PDF, also called the log-likelihood function, is adopted to estimate the Hurst exponent as follows:

$$\log p(\mathbf{x}; H) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|\mathbf{R}| - \frac{1}{2}\mathbf{x}^T\mathbf{R}^{-1}\mathbf{x}. \quad (8)$$

3. Two Estimators in a Two-Stage Procedure

In this paper, a two-stage estimator with lower computational cost than and almost the same accuracy as the fast MLE will be proposed. At the moment, the DI method (Chang, 2014b) is most suitably chosen as the first-stage estimator and the fast MLE (Chang, 2014a) as the second-stage estimator.

In the future, when there is an estimator with accuracy higher than the DI method and execution time 10 times less than the fast MLE, the new estimator can play the role of the first-stage estimator.

For ease of understanding and use of a two-stage procedure from a large to small range, two adopted estimators, the DI method and the fast MLE, will be introduced in the following subsections.

3.1. The DI method

The DI method can be directly applied in the signals of DFGN, denoted by $X_H[n]$. If a practical signal of interest is modeled as DFBM, denoted by $B_H[n]$, then its increments, $X_H[n] = B_H[n] - B_H[n-1]$, need to be calculated in advance. The following procedure is a detailed description of estimating the Hurst exponent by using the DI method.

Procedure 1: Estimating the Hurst exponent by using the DI estimator

1. Choose a suitable data size or number, N.
2. Extract a section of signal of size N, $X_H[n]$, from a realization of DFGN (or the increments of DFBM).
3. Determine the threshold for terminating the Levinson algorithm according to Algorithm 1 or 2 of Chang (Chang, 2014b).
4. Modify $X_H[n]$ into zero mean.
5. Estimate ACFs by using the following formula:

$$\hat{r}[k] = \frac{1}{N}\sum_{n=k+1}^{N} X_H[n]X_H[n-k], \quad k = 0,1,\ldots,N-1. \tag{9}$$

6. Estimate the parameter sets by using estimated ACFs via the Levinson algorithm (Haykin, 1989; Kay, 1988) until the threshold is satisfied, or the corresponding order, $\hat{p}$, is found out.
7. Estimate the slope, s, of $\{\log(\hat{f}(\omega_k))\}$ versus $\{\log(\omega_k)\}$ over $\omega_k = k\pi/M$, $k = 1, 2,\ldots, M$, (in this paper, M was chosen as 63) by using (10) and (11) below:

$$\hat{f}(\omega) = \hat{\rho}_{\hat{p}} / |\hat{A}(\omega)|^2, \tag{10}$$

where $\hat{\rho}_{\hat{p}}$ is the estimate of the white noise variance (prediction error power) and

$$\hat{A}(\omega) = 1 + \hat{a}[1]\exp\{-j\omega\} + \cdots + \hat{a}[\hat{p}]\exp\{-j\omega\hat{p}\}, \tag{11}$$

where $\hat{p}$ denotes the estimated order of the AR model.

8. Estimate the Hurst exponent, H, through the following formula (Chang, 2014b):

$$\hat{H} = c_3\hat{s}^3 + c_2\hat{s}^2 + c_1\hat{s}^1 + c_0, \tag{12}$$

where $c_0 = 0.4990$, $c_1 = -0.4355$, $c_2 = 0.0441$ and $c_3 = 0.0236$.

9. Calculate, if necessary, the fractal dimension using $\hat{D} = 2 - \hat{H}$ (Falconer, 1990).

Note that although formula (5) can give a direct and easy way to estimate the PSDs, the DI method estimated the PSDs through an AR model because the AR model possesses good characteristics. Kay (Kay, 1988) has summarized some reasons for the AR spectral estimator better than the Blackman-Tukey (BT) one. First, the AR spectral estimator has higher resolution than the BT one because of an implicit extension of the estimated ACFs. The BT spectral estimator directly truncates the ACFs and then appends the ACFs by zeros, thereby smearing spectral estimates. However, the AR spectral estimator extrapolates the ACFs according to an iteration formula, and thus the resultant spectral estimates are closer to the true ones.

3.2. The fast MLE

Considering two parameters of mean and variance exist in the PDF, as well as each parameter is known or unknown, there are four cases considered. In addition, for unknown mean there are two ways of estimation: the mean by the MLE and the sample mean. For ease reference, they are rewritten as follows.

3.2.1. Case 1: Known mean (displacement) and known variance

Suppose the mean of the data set $\mathbf{x}$ is zero. The equivalent formula for maximization is as follows:

$$\max_{H}\left\{-\log\left|\overline{\mathbf{R}}\right|-\sigma^{-2}\mathbf{x}^{T}\overline{\mathbf{R}}^{-1}\mathbf{x}\right\}, \tag{13}$$

where σ^2 is the variance known to users and

$$\mathbf{R}=\sigma^{2}\overline{\mathbf{R}}, \tag{14}$$

where $\overline{\mathbf{R}}$ can be called the normalized autocovariance matrix.

3.2.2. Case 2: Known mean (displacement) and unknown variance

For signals modeled as DFBM, this is the more general case than Case 1. For estimation of the Hurst exponent, the increments of signals must be calculated, and thus the obtained

increment process is DFGN with zero mean, but its variance is unknown to users. The equivalent formula for maximization is as follows:

$$\max_{H}\left\{-\log\left|\overline{\mathbf{R}}\right|-N\log\left(\frac{1}{N}\mathbf{x}^{T}\overline{\mathbf{R}}^{-1}\mathbf{x}\right)\right\}. \tag{15}$$

Then the estimate of variance is calculated by

$$\hat{\sigma}^{2}=\frac{1}{N}\mathbf{x}^{T}\overline{\mathbf{R}}^{-1}\mathbf{x}. \tag{16}$$

3.2.3. Case 3: Unknown mean (displacement) and known variance

Suppose measurement data are an expression of the form: $\mathbf{z}=\mathbf{x}+\boldsymbol{\mu}$, where $\mathbf{x}$ is modeled as DFGN with zero mean and $\boldsymbol{\mu}$ is a column vector of scalar μ, i.e., $\boldsymbol{\mu}=[\mu \quad \mu \quad \cdots \quad \mu]^{T}$. According to two estimators of μ, the following two approaches are used to estimate the Hurst exponent.

Approach 1:

The equivalent formula for maximization is as follows:

$$\max_{H}\left\{-\log\left|\overline{\mathbf{R}}\right|-\sigma^{-2}\left(\mathbf{z}-\hat{\boldsymbol{\mu}}\right)^{T}\overline{\mathbf{R}}^{-1}\left(\mathbf{z}-\hat{\boldsymbol{\mu}}\right)\right\}. \tag{17}$$

where σ^{2} is known to users and

$$\hat{\mu}=\frac{1}{\left\|\overline{\mathbf{A}}\right\|_{s}}\sum_{k=0}^{N-1}\left\|\overline{\mathbf{a}}_{k}\right\|_{s}z_{k}, \tag{18}$$

where $\overline{\mathbf{A}}=\overline{\mathbf{R}}^{-1}$, $\overline{\mathbf{a}}_{k}=\left[\overline{a}_{0k} \quad \overline{a}_{1k} \quad \cdots \quad \overline{a}_{(N-1)k}\right]^{T}$, $\left\|\overline{\mathbf{a}}_{k}\right\|_{s}=\sum_{i=0}^{N-1}\overline{a}_{ik}$ and $\left\|\overline{\mathbf{A}}\right\|_{s}=\sum_{k=0}^{N-1}\left\|\overline{\mathbf{a}}_{k}\right\|_{s}$.

Approach 2:

For ease coding, the sample mean for μ, denoted as follows, is often adopted:

$$\hat{\mu}=\frac{1}{N}\sum_{k=0}^{N-1}z_{k}. \tag{19}$$

3.2.4. Case 4: Unknown mean (displacement) and unknown variance

For signals modeled as DFGN, this is the more general case than Case 3. No mean and

variance are known to users in advance. Like Case 3, suppose measurement data are the form $\mathbf{z} = \mathbf{x} + \boldsymbol{\mu}$ and its unknown variance is σ^2. According to two estimators of μ, the following two approaches are used to estimate the Hurst exponent:

Approach 1:

The equivalent formula for maximization is as follows:

$$\max_{H}\left\{-\log\left|\overline{\mathbf{R}}\right| - N\log\left(\frac{1}{N}(\mathbf{z}-\hat{\boldsymbol{\mu}})^T\overline{\mathbf{R}}^{-1}(\mathbf{z}-\hat{\boldsymbol{\mu}})\right)\right\}, \tag{20}$$

where $\hat{\mu}$ is given by (18). Then the estimate of variance is calculated by

$$\hat{\sigma}^2 = \frac{1}{N}(\mathbf{z}-\hat{\boldsymbol{\mu}})^T\overline{\mathbf{R}}^{-1}(\mathbf{z}-\hat{\boldsymbol{\mu}}). \tag{21}$$

Approach 2:

Likewise, the sample mean for μ is used other than the MLE for μ.

Finally, the golden section search (Schilling and Harris, 2000) is appropriately used to find out the maxima of (13), (15), (17) and (20) and then their corresponding optimal Hurst exponents, as adopted by Chang (Chang, 2014a). The following procedure is a detailed description of estimating the Hurst exponent by the fast MLE.

Procedure 2: Estimating the Hurst exponent by using the fast MLE

1. Choose a suitable data size or number, N.
2. Extract a section of signal of size N, $X_H[n]$, from a realization of DFGN (or the increments of DFBM).
3. Choose a suitable error tolerance, t, for the golden section search. In general, $t = 0.001$ is enough, and thus it is chosen in this paper.
4. Select an appropriate case from Case 1 to Case 4 according to whether two parameters (mean and variance) are known or unknown.
5. Use the golden section search to find out the optimal Hurst exponent through efficiently

implementing (13) for Case 1, (15) for Case 2, (17) for Case 3 and (20) for Case 4 (Chang, 2014a). For the normal or one-stage procedure, the search interval is between 0 and 1.

6. Calculate, if necessary, the fractal dimension using $\hat{D} = 2 - \hat{H}$.

4. A Two-Stage Procedure from a Large to Small Range

Since the Hurst exponent is a real value lying between 0 and 1, the search interval for the golden section search is $(0,1)$. Suppose an acceptable error tolerance is t, then the required number of iterations can be calculated using the equation $(b-a)r^{k-1} = t$ (Chang, 2009), where a is the upper limit of the search interval, b the lower limit and $r = (\sqrt{5}-1)/2 \approx 0.6180$ whose reciprocal, $\phi = 1/r$, is called the golden ration. That is, $k = \lceil (\ln t - \ln(b-a))/\ln r \rceil + 1$, where $\lceil x \rceil$ denotes the smallest integer larger than and equal to x. For example, when $a = 0$, $b = 1$ and $t = 0.001$, the number of iterations, k, is 16. Therefore, the number of function evaluations is 17. Obviously, if the search interval is narrowed down, the number of function evaluations will be reduced. To obtain a reliable estimation interval containing the true Hurst exponent, another estimator for the Hurst exponent will be adopted in the first stage. A qualified first-stage estimator should contain at least two features: one is that the execution time must be 10 times less than that of a second-stage estimator; the other is that the closer its accuracy gets to that of the second-stage estimator, the better the estimator is. For possibly best accuracy, the DI method (Chang, 2014b) is currently the most suitable first-stage estimator and the fast MLE (Chang, 2014a) is the best second-stage estimator.

Procedure 3: Estimating the Hurst exponent by using a two-stage procedure

1. Choose a suitable data size or number, N.

2. Extract a section of signal of size N, $X_H[n]$, from a realization of DFGN (or the increments of DFBM).
3. Use Steps 3 to 8 of Procedure 1 to estimate the Hurst exponent in the first stage. The estimated Hurst exponent is denoted as $\hat{H}_1$.
4. Select a reliable interval, I, centered at the $\hat{H}_1$ according to the case and the data size. In general, the interval decreases as the data size increases.
5. Calculate the interval for the golden section search as (a, b), where $a = \hat{H}_1 - 0.5I$ and $b = \hat{H}_1 + 0.5I$.
6. Use Steps 3 to 5 of Procedure 2 to estimate the Hurst exponent in the second stage. The estimated Hurst exponent is denoted as $\hat{H}$.
7. Calculate, if necessary, the fractal dimension using $\hat{D} = 2 - \hat{H}$.

For users to easily apply the two-stage estimator to practical signals, a detailed procedure is provided as follows:

Procedure 4: Applying a two-stage estimator to practical signals

1. Input a practical signal for analysis; suppose that its data length is L.
2. Model the signal as DFBM or DFGN. If DFBM is adopted, then the signal is viewed as $B_H[n]$ and its increments of DFBM, called DFGN, are calculated via $X_H[n] = B_H[n] - B_H[n-1]$, which is a zero-mean realization and hence no estimation of mean is necessary. In general, Case 2 is used except that variance is known to users in advance. If DFGN is adopted, the signal is viewed as $X_H[n]$. Thus, Case 4 is used except that mean and variance are known to users in advance.
3. Set the error tolerance, t, for the golden section search as 0.001.
4. Select the window size for estimation, N, also called the data size in this paper.

5. Select the shift size, S, the distance between the current window and the next window.

6. Calculate the required number of estimation, $K = \left\lfloor \frac{L-N}{S} \right\rfloor + 1$, where $\lfloor x \rfloor$ denotes the maximum integer smaller than or equal to x.

7. Initiate $k = 0$.

8. Execute $k = k + 1$, if $k > K$, then terminate, otherwise take the *k*th window from the input signal, denoted as $X_H[n]$.

9. Use Steps 3 to 6 of Procedure 3 to estimate the Hurst exponent.

10. Calculate, if necessary, the fractal dimension using $\hat{D} = 2 - \hat{H}$ (Falconer, 1990).

11. Go to Step 8.

12. Plot the figure of the estimated Hurst exponents ($\hat{H}$) versus the time sequence (K) and then explain the characteristics of the input signal.

5. Results and Discussion

For a comprehensive comparison of the fast MLE and the proposed two-stage estimator, a wide range of Hurst exponents and data sizes were considered, including $H = 0.01$, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95 and 0.99 (totally, 13 Hurst exponents) as well as $N = 128$, 256, 512, 1024, 2048 and 4096 (totally, 6 types of data sizes).

For long-term dependency and fine correlation structure of simulated signals, the algorithm proposed by Lundahl et al. (Lundahl *et al.*, 1986) was adopted. First, 100 realizations of white Gaussian noise for each data size were produced by a Gaussian random generator, and then 100 realizations of DFGN for each Hurst exponent were made through the algorithm mentioned above.

For each realization, the error tolerance for the golden section search is set to be 0.001. For the fast MLE, the search interval is $(0, 1)$, and thus it needs 16 iterations, totally 17

function evaluations (Chang, 2009).

The purpose of the DI method (the first-stage estimator) is to locate a reliable interval containing the true Hurst exponent. To understand how the interval affect the accuracy of the fast MLE (the second-stage estimator), five intervals were considered, including 0.0469, 0.0760, 0.1229, 0.1990 and 0.3219, which can be computed as $(b-a)$ via the formula $(b-a)r^{k-1}=t$ with $r=(\sqrt{5}-1)/2$ and $t=0.001$ under iteration k from 9 to 13, or function evaluations from 10 to 14.

Tables 1–4 show the accuracy of the fast MLE and two-stage estimator for Case 1 (known mean and variance), Case 2 (known mean and unknown variance), Case 3 (A1) (unknown mean and known variance) and Case 4 (A1) (unknown mean and variance), each value representing the mean of mean-squared errors (MSEs) of 100 realizations over 13 Hurst exponents, where A1 denotes that the mean is estimated by the MLE instead of the sample mean (A2). Since the results by A2 are similar to those by A1, they are left out for conciseness.

To more clearly distinguish the error between the fast MLE and the two-stage estimator under a given interval, Tables 5–8 list the results of error percentage according to the formula: $(E_2-E_1)/E_1$, where E_2 stands for the error of the proposed two-stage estimator, E_1 for that of the one-stage estimator (the fast MLE).

In these tables, the values in purple represent that error percentage is negative values, and thus the two-stage estimator is superior to the fast MLE, the ones in black represent that error percentage belongs to $O(10^{-1})$, green $O(10^{-2})$, orange $O(10^{-3})$, blue $O(10^{-4})$, and finally, blue and bold $O(10^{-5})$.

Generally speaking, the results with error percentage belonging to $O(10^{-3})$ can be viewed as almost the same as those of the fast MLE. Therefore, in the more general case for

description of signals modeled as DFGN, Case 4, the two-stage estimator can save up to 41.18% (7/17, 17 ↓ 10) the computational time of the fast MLE for six data sizes while remaining almost the same accuracy as the fast MLE; in the more general case for description of signals modeled as DFBM, Case 2, the two-stage estimator can save about 35.29% (6/17) for five larger data sizes (except for 128); Cases 1 and 3 can save about 29.41% (5/17) for three larger data sizes, about 23.53% (4/17) for data size being 256 and 17.65% (3/17) for data size being 128.

From a data-size perspective, these tables also show that time saving gets higher as the data size gets larger. In addition, it is larger data sizes that we worry about whether or not the estimator can meet the demand of real-time estimation. Therefore, the amount of time saving in larger data sizes is more valuable for applicability than that in smaller data sizes.

For real efficiency comparison, all results were run under the same computing environment. (1) Hardware: a computer of Intel® Core(TM) i7-2600 processor, up to 3.40GHz and a RAM of 8.00GB (7.89 GB available); (2) operating system: Windows 7 Professional Service Pack 1; (3) programming software: MATLAB R2011b 64-bit (win64); (4) optimization algorithm: golden section search with threshold 0.001.

Since the computational time difference among four cases is minor, only the efficiency comparison of Case 4 (A1) is plotted in a log-log scale, as shown in Fig. 1. As explained previously, the number of function evaluations for Case 4 (A1) only needs 10 times other than the original 17 times. Therefore, when larger data sizes are considered, the computational time of data size 2048 decreases from 2.27 seconds to 1.46 seconds for each estimate of the Hurst exponent; the computational time of data size 4096 also drops from 8.32 seconds to 5.46 seconds for each estimate. It can be expected that the time saving is enormous for massive estimates. More importantly, time saving contributes to raising the chance of making systems real-time responses.

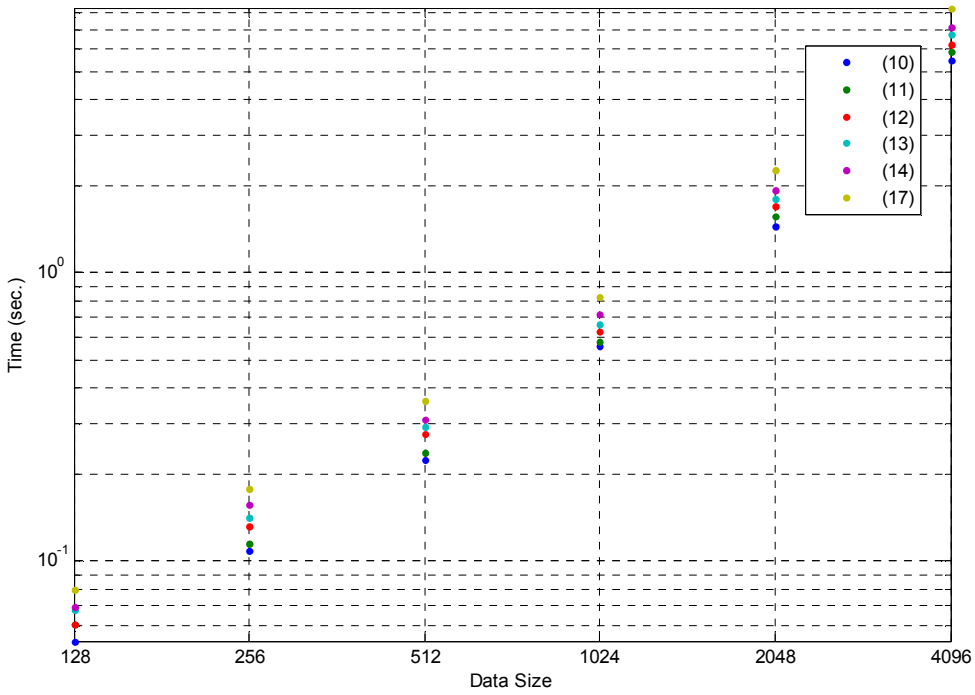

Fig. 1. Efficiency comparison of Case 4 (A1) for six data sizes 128, 256, 512, 1024, 2048 and 4096, each data size being performed from 10 to 14 (two-stage estimator) and 17 (one-stage estimator) function evaluations.

Table 1: Accuracy of the fast MLE and two-stage estimator for Case 1

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	1.77E-03	6.91E-04	3.56E-04	1.75E-04	7.59E-05	3.85E-05
0.0760 (11)	1.53E-03	6.05E-04	3.15E-04	1.56E-04	7.20E-05	3.73E-05
0.1229 (12)	1.33E-03	5.63E-04	2.93E-04	1.51E-04	7.12E-05	3.73E-05
0.1990 (13)	1.21E-03	5.52E-04	2.90E-04	1.50E-04	7.12E-05	3.73E-05
0.3219 (14)	1.18E-03	5.52E-04	2.90E-04	1.51E-04	7.12E-05	3.73E-05
1.0000 (17)	1.18E-03	5.52E-04	2.90E-04	1.51E-04	7.11E-05	3.73E-05

S.I. (F.E.): Search Interval (Function Evaluation)

Table 2: Accuracy of the fast MLE and two-stage estimator for Case 2

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	2.01E-03	8.85E-04	4.74E-04	2.65E-04	1.20E-04	6.61E-05
0.0760 (11)	1.85E-03	8.58E-04	4.67E-04	2.59E-04	1.20E-04	**6.57E-05**
0.1229 (12)	1.75E-03	8.50E-04	4.65E-04	2.59E-04	1.19E-04	6.58E-05
0.1990 (13)	1.72E-03	8.50E-04	4.65E-04	2.59E-04	1.19E-04	6.58E-05
0.3219 (14)	1.72E-03	8.50E-04	4.64E-04	2.59E-04	1.20E-04	6.58E-05
1.0000 (17)	1.72E-03	8.50E-04	4.64E-04	2.59E-04	1.19E-04	6.57E-05

S.I. (F.E.): Search Interval (Function Evaluation)

Table 3: Accuracy of the fast MLE and two-stage estimator for Case 3 (A1)

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	1.87E-03	7.45E-04	3.62E-04	1.80E-04	7.91E-05	3.82E-05
0.0760 (11)	1.66E-03	6.57E-04	3.19E-04	1.61E-04	7.51E-05	3.70E-05
0.1229 (12)	1.49E-03	6.11E-04	2.97E-04	1.55E-04	7.44E-05	3.70E-05
0.1990 (13)	1.39E-03	5.99E-04	2.94E-04	1.55E-04	7.44E-05	3.71E-05
0.3219 (14)	1.36E-03	5.99E-04	2.94E-04	1.55E-04	7.43E-05	3.70E-05
1.0000 (17)	1.36E-03	5.99E-04	2.94E-04	1.55E-04	7.44E-05	3.71E-05

S.I. (F.E.): Search Interval (Function Evaluation)

Table 4: Accuracy of the fast MLE and two-stage estimator for Case 4 (A1)

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	2.51E-03	1.12E-03	5.10E-04	2.98E-04	1.38E-04	6.79E-05
0.0760 (11)	2.50E-03	1.14E-03	5.18E-04	2.97E-04	1.38E-04	**6.76E-05**
0.1229 (12)	2.50E-03	1.14E-03	5.19E-04	2.97E-04	1.38E-04	6.77E-05
0.1990 (13)	2.51E-03	**1.15E-03**	5.18E-04	2.97E-04	1.38E-04	6.77E-05
0.3219 (14)	2.51E-03	1.15E-03	**5.18E-04**	2.97E-04	1.38E-04	6.77E-05
1.0000 (17)	2.51E-03	1.15E-03	5.18E-04	2.97E-04	1.38E-04	6.76E-05

S.I. (F.E.): Search Interval (Function Evaluation)

Table 5: Error percentage between the fast MLE and the two-stage estimator for Case 1

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	4.99E-01	2.53E-01	2.31E-01	1.64E-01	6.69E-02	3.08E-02
0.0760 (11)	2.94E-01	9.67E-02	8.72E-02	3.80E-02	1.22E-02	-8.88E-05
0.1229 (12)	1.24E-01	2.03E-02	1.12E-02	4.29E-04	1.78E-03	-4.10E-04
0.1990 (13)	2.52E-02	6.31E-04	4.87E-04	-5.81E-04	1.65E-03	9.82E-04
0.3219 (14)	1.57E-04	4.68E-04	4.24E-04	-3.04E-04	9.13E-04	-2.22E-04

S.I. (F.E.): Search Interval (Function Evaluation)

Table 6: Error percentage between the fast MLE and the two-stage estimator for Case 2

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	1.68E-01	4.05E-02	2.01E-02	2.44E-02	7.65E-03	5.72E-03
0.0760 (11)	7.60E-02	9.51E-03	4.90E-03	2.69E-03	1.32E-03	**8.35E-05**
0.1229 (12)	1.79E-02	2.54E-04	9.69E-04	6.64E-04	1.21E-04	1.87E-03
0.1990 (13)	1.49E-03	1.52E-04	1.01E-03	9.92E-04	-5.06E-04	1.16E-03
0.3219 (14)	-1.13E-04	1.33E-04	-3.22E-05	-4.59E-05	7.36E-04	1.13E-03

S.I. (F.E.): Search Interval (Function Evaluation)

Table 7: Error percentage between the fast MLE and the two-stage estimator for Case 3 (A1)

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	3.69E-01	2.44E-01	2.31E-01	1.59E-01	6.37E-02	2.85E-02
0.0760 (11)	2.16E-01	9.59E-02	8.49E-02	3.74E-02	1.04E-02	-3.04E-03
0.1229 (12)	9.49E-02	1.93E-02	1.06E-02	9.36E-04	-5.86E-05	-5.15E-03
0.1990 (13)	2.16E-02	6.94E-04	-2.36E-04	-1.52E-04	7.58E-04	-1.38E-03
0.3219 (14)	2.45E-04	-1.45E-04	-6.89E-04	2.11E-04	-7.85E-04	-3.37E-03

S.I. (F.E.): Search Interval (Function Evaluation)

Table 8: Error percentage between the fast MLE and the two-stage estimator for Case 4 (A1)

S.I. (F.E.)	128	256	512	1024	2048	4096
0.0469 (10)	3.56E-04	-2.14E-02	-1.67E-02	3.52E-03	6.77E-04	4.53E-03
0.0760 (11)	-5.48E-03	-6.32E-03	1.21E-04	-4.23E-04	-7.03E-04	**4.55E-05**
0.1229 (12)	-4.43E-03	-2.28E-03	9.98E-04	-1.29E-03	-1.52E-03	6.39E-04
0.1990 (13)	-6.29E-04	**8.40E-05**	1.23E-04	-4.85E-04	-1.76E-03	7.42E-04
0.3219 (14)	-2.33E-04	2.46E-04	**5.30E-05**	-1.23E-03	-2.12E-03	1.14E-03

S.I. (F.E.): Search Interval (Function Evaluation)

6. Conclusions

A trade-off between accuracy and efficiency of estimation must be made anyway. Generally, a more accurate estimator needs more computational time; a more efficient estimator has less accurate performance. A perfect estimator, both more accurate and efficient than another, is very welcome, but it is difficult to exist except that certain helpful structures of computation, as in the fast MLE, are discovered to speed up the computation.

Normally, users choose the most accurate estimator to evaluate parameters in the absence of a time limit; otherwise they might choose the quickest estimator with acceptable accuracy. Obviously, no estimator is suitable for all situations. How to balance between accuracy and efficiency mainly depends on individual needs and resources available. However, a quicker estimator with almost the same accuracy as the original one is always welcomed by users.

In this paper, a two-stage estimator through an integration of the DI method and the fast MLE is proposed. The two-stage estimator makes full use of the advantages of two different kinds of estimators to achieve the most effective combination. Relatively, a quicker estimator with reliable accuracy is suitable for the first stage, and a more accurate estimator is chosen in the second stage.

Experimental results show that the proposed two-stage estimator can save up to 41.18% the computational time of the fast MLE in the more general case of DFGN while remaining almost the same accuracy as the fast MLE. In the more general case of DFBM, the two-stage estimator can also save about 35.29% except for data sizes not larger than 128. In the future, when another estimator has more accurate than the DI method and its speed is at least 10 times quicker than that of the fast MLE, it can play the role of the DI method, thereby saving more computational time.

In addition, the proposed two-stage estimation procedure is also a novel and valuable idea. It can be expected that other fields of parameter estimation can apply the concept of the two-stage estimation procedure to raise computational performance while remaining almost the same accuracy as the more accurate of two estimators.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Beran, J. (1994), *Statistics for Long-Memory Processes*, Chapman & Hall, New York.
2. Biswas, A., Zeleke, T.B. and Si, B.C. (2012), "Multifractal detrended fluctuation analysis

in examining scaling properties of the spatial patterns of soil water storage," *Nonlinear Processes in Geophysics*, Vol. 19, pp. 227-238.

3. Bruce, E.N. (2001), *Biomedical Signal Processing and Signal Modeling*, John Wiley & Sons, New York.
4. Chang, S., Li, S.-J., Chiang, M.-J., Hu, S.-J. and Hsyu, M.-C. (2007), "Fractal dimension estimation via spectral distribution function and its application to physiological signals," *IEEE Transactions on Biomedical Engineering*, Vol. 54, No. 10, pp. 1895-1898.
5. Chang, S., Mao, S.-T., Hu, S.-J., Lin, W.-C. and Cheng, C.-L. (2000), "Studies of detrusor-sphincter synergia and dyssynergia during micturition in rats via fractional Brownian motion," *IEEE Transactions on Biomedical Engineering*, Vol. 47, No. 8, pp. 1066-1073.
6. Chang, Y.-C. (2009), "N-Dimension Golden Section Search: Its Variants and Limitations," *The 2nd International Conference on BioMedical Engineering and Informatics* (BMEI2009), Tianjin, China, October 17-19.
7. Chang, Y.-C. (2014a), "Efficiently implementing the maximum likelihood estimator for Hurst exponent," *Mathematical Problems in Engineering*, vol. 2014, Article ID 490568, 10 pages.
8. Chang, Y.-C. (2014b), "An efficient estimator of Hurst exponent through an autoregressive model with an order selected by data induction," *Bio-Medical Materials and Engineering*, Vol. 24, No. 6, pp. 3557–3568.
9. Chang, Y.-C. and Chang, S. (2002), "A fast estimation algorithm on the Hurst parameter of discrete-time fractional Brownian motion," *IEEE Transactions on Signal Processing*,

Vol. 50, No. 3, pp. 554-559.

10. Chang, Y.-C., Chen, L.-H., Lai, L.-C. and Chang, C.-M. (2012), "An efficient variance estimator for the Hurst exponent of discrete-time fractional Gaussian noise," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Vol. E95-A, No. 9, pp. 1506-1511.
11. Chang, Y.-C., Lai, L.-C., Chen, L.-H., Chang, C.-M. and Chueh, C.-C. (2014), "A Hurst exponent estimator based on autoregressive power spectrum estimation with order selection," *Bio-Medical Materials and Engineering*, Vol. 24, No. 1, pp. 1041-1051.
12. Chen, S.S., Keller, J.M. and Crownover, R.M. (1993), "On the calculation of fractal features from images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp. 1087-1090.
13. Domino, K. (2012), "The use of the Hurst exponent to investigate the global maximum of the Warsaw Stock Exchange WIG20 index," *Physica A*, Vol. 391, No. 1-2, pp. 156-169.
14. Falconer, K. (1990), *Fractal Geometry: Mathematical Foundations and Applications*. John Wiley & Sons, New York.
15. Fernández-Martínez, M., Sánchez-Granero, M.A. and Trinidad Segovia, J.E. (2013), "Measuring the self-similarity exponent in Lévy stable processes of financial time series," *Physica A*, Vol. 392, No. 21, pp. 5330-5345.
16. Flandrin, P. (1989), "On the spectrum of fractional Brownian motions," *IEEE Transactions on Information Theory*, Vol. 35, No. 1, pp. 197-199.
17. Gao, J., Hu, J., Mao, X. and Perc, M. (2012), "Culturomics meets random fractal theory Insights into long-range correlations of social and natural phenomena over the past two

centuries," *Journal of the Royal Society Interface*, Vol. 9, No. 73, pp. 1956-1964.

18. Gonçalves, W.N. and Bruno, O.M. (2013), "Combining fractal and deterministic walkers for texture analysis and classification," *Pattern Recognition*, Vol. 46, No. 11, pp. 2953-2968.
19. Hagerhall, C.M., Purcell, T. and Taylor, R. (2004), "Fractal dimension of landscape silhouette outlines as a predictor of landscape preference," *Journal of Environmental Psychology*, Vol. 24, No. 2, pp. 247–255.
20. Hastings H.M. and Sugihara G. (1993), *Fractals: A User's Guide for the Natural Sciences*, Oxford University Press, Oxford.
21. Haykin, S. (1989), *Modern Filters*, Macmillan, New York.
22. Hansen, A., Simonsen, I. and Nes, O.M. (1998), "Determination of the Hurst exponent by use of wavelet transforms," *Physical Review E*, Vol. 58, pp. 2779-2787.
23. Huang, P.-W. and Lee, C.-H. (2009), "Automatic classification for pathological prostate images based on fractal analysis," *IEEE Transactions on Medical Imaging*, Vol. 28, No. 7, pp. 1037-1050.
24. Jin, X.C., Ong, S.H. and Jayasooriah, (1995), "A practical method for estimating fractal dimension," *Pattern Recognition Letters*, Vol. 16, No. 5, pp. 457-464.
25. Kay, S.M. (1988), *Modern Spectral Estimation: Theory & Application*, Prentice-Hall, New Jersey.
26. Kay, S.M. (1993), *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, New Jersey.
27. Lin, P.-L., Huang, P.-W., Lee, C.-H. and Wu, M.-T. (2013), "Automatic classification for

solitary pulmonary nodule in CT image by fractal analysis based on fractional Brownian motion model," *Pattern Recognition*, Vol. 46, No. 12, pp. 3279-3287.

28. Liu, S.-C. and Chang, S. (1997), "Dimension estimation of discrete-time fractional Brownian motion with applications to image texture classification," *IEEE Transactions on Image Processing*, Vol. 6, No. 8, pp. 1176-1184.
29. Lundahl, T., Ohley, J., Kay, S.M. and Siffert, R. (1986), "Fractional Brownian motion: A maximum likelihood estimator and its application to image texture," *IEEE Transactions on Medical Imaging*, Vol. 5, No. 3, pp. 152-161.
30. Mandelbrot, B.B. (1983), *The Fractal Geometry of Nature*, W. H. Freeman and Company, New York, 1983.
31. Mandelbrot, B.B. and Van Ness, J.W. (1968), "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, Vol. 10, No. 4, pp. 422-437.
32. Peng, C.K., Buldyrev, S.V., Goldberger, A.L., Havlin, S., Sciortino, F., Simons, M., and Stanley, H.E. (1992), "Long-range correlations in nucleotide sequences," *Nature*, Vol. 356, pp. 168-170.
33. Peng, C.K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E. and Goldberger, A.L. (1994), "Mosaic organization of DNA nucleotides," *Physical Review E*, Vol. 49, pp. 1685-1689.
34. Pentland, A.P. (1984), "Fractal-based description of natural scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-6, No. 6, pp. 661-674.
35. Perc, M. (2012), "Evolution of the most common English words and phrases over the centuries," *Journal of the Royal Society Interface*, Vol. 9, No. 77, pp. 3323-3328.

36. Perc, M. (2013), "Self-organization of progress across the century of physics," *Scientific Reports*, Vol. 3, No. 1720.

37. Petersen, A.M., Tenenbaum, J.N., Havlin, S., Stanley, H.E. and Perc, M. (2012), "Languages cool as they expand: Allometric scaling and the decreasing need for new words," *Scientific Reports*, Vol. 2, No. 943.

38. Rehman, S. and Siddiqi, A.H. (2009), "Wavelet based hurst exponent and fractal dimensional analysis of Saudi climatic dynamics," *Chaos, Solitons and Fractals*, Vol. 40, pp. 1081-1090.

39. Rejichi, I.Z. and Aloui, C. (2012), "Hurst exponent behavior and assessment of the MENA stock markets efficiency," *Research in International Business and Finance*, Vol. 26, No. 3, pp. 353-370.

40. Rostek, S. and Schöbel, R. (2013), "A note on the use of fractional Brownian motion for financial modeling," *Economic Modelling*, Vol. 30, pp. 30-35.

41. Samorodnitsky, G. and Taqqu, M.S. (1994), *Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance*, Chapman & Hall, New York.

42. Sarkar, N. and Chaudhuri, B.B. (1992), "An efficient approach to estimate fractal dimension of textural images," *Pattern Recognition*, Vol. 25, No. 9, pp. 1035-1041.

43. Sarkar, N. and Chaudhuri, B.B. (1994), "An efficient differential box-counting approach to compute fractal dimension of image," *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 24, No. 1, pp. 115-120.

44. Schilling, R.J. and Harris, S.L. (2000), *Applied Numerical Methods for Engineers: Using MATLAB and C*, Brooks/Cole, New York.

45. Schonhoff, T.A. and Giordano, A.A. (2006), *Detection and Estimation Theory and Its Applications*, Prentice-Hall, New Jersey.

46. Taqqu, M.S., Teverovsky, V. and Willinger, W. (1995), "Estimators for long-range dependence: An empirical study," *Fractals*, Vol. 3, No. 4, pp. 785-798.

47. Wang, Y.-Z., Li, B., Wang, R.-Q., Su, J. and Rong, X.-X. (2011), "Application of the Hurst exponent in ecology," *Computers and Mathematics with Applications*, Vol. 61, No. 8, pp. 2129-2131.

48. Zuñiga, A.G., Florindo, J.B. and Bruno, O.M. (2014), "Gabor wavelets combined with volumetric fractal dimension applied to texture analysis," *Pattern Recognition Letters*, Vol. 36, pp. 135-143.

Biographical Details

Yen-Ching Chang was born in Changhua, Taiwan, ROC, in 1966. He received the B.S. degree in electrical engineering from National Taiwan Institute of Technology, Taipei, Taiwan, in 1991 and the M.S. and Ph. D. degrees in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1993 and 2002. Since 2003, he has been with Department of Medical Informatics, Chung Shan Medical University, Taichung, Taiwan, where he is currently a professor. His research interests include statistical signal processing, image processing and soft computing.