http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 12 (2): 345-350, 2013
ISSN 1812-5638 / DOI: 10.3923/itj.2013.345.350
© 2013 Asian Network for Scientific Information

The Entropy Rate Estimation for Scalable Wavelet Video Coders using Hidden Markov Process

[1]Lizhi Zhang, [1]Zhijie Zhao, [1,2]Xuesong Jin and [1]Huadong Sun
[1]School of Computer and Information Engineering, Harbin University of Commerce, Harbin, 150028, China
[2]School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin, 150001, China

Abstract: Based on statistical investigation of image wavelet coefficients, a mathematical model of hidden Markov stochastic process was established in this study. The model parameters are related to the hidden state and observable value and an iterative calculating method of the joint distribution and the conditional entropy was proposed through investigating the inner relationship among the model parameters. Based on it, the entropy rate estimation was implemented. The proposed algorithm can be used to establish the Rate-Distortion (RD) model for the scalable wavelet video coding system and it has good performance on the analyticity and decorrelation property. Plenty of simulations were performed to compare the proposed approach with the empirical histogram method and the experiments show that the average deviation of the two methods is within 10%. This illustrates that the proposed method is more reliable.

Key words: Discrete wavelet transform, entropy, scalable video coding, hidden Markov tree, expectation-maximization algorithm

INTRODUCTION

In recent years, the demand for video application over Internet grows significantly, so more and more researchers focus their attentions on the scalability of video coding. However, this requires establishing Rate-Distortion (RD) model to guide rate allocation. Generally, there are two categories of RD analysis method. The first category is the empirical method, in which the experimental RD data are described in functional expression (Ding and Liu, 1996; Chiang and Zhang, 1997; Stuhlmuller *et al*., 2000). It is simple in terms of computation, but it is difficult to be generalized due to lack of considering the input sequence characteristics. The second category method is analytical and theoretical (Sakrison, 1968; O'Neal and Natarajan, 1997) and many approaches have been achieved in former works. For example, Mallat and Falzon (1998) provided a thorough analysis of RD performance of wavelet transforming coding in the low bit-rate region and presented the model for still-image coding. Dai *et al*. (2004) presented a closed-form operational RD model for a wide range of transformed-based coder. Wang and Van der Schaar (2006) presented an operational RD model for context coding, decorrelating the interscale dependencies. For the existing RD model, it is not good enough in terms of analyticity and the application range is often limited. Combined with the distortion analysis, the bit-rate estimated method proposed in this study can be applied to establish general RD model for video coding system and can be used widely.

The Hidden Markov Model (HMM) mainly applied to speech processing (Rabiner, 1989; Anwar *et al*., 2007) and then applied to image processing (Wang *et al*., 2005), for example, Crouse *et al*. (1996, 1997) studied image denoising, classification and segmentation. Fan and Xia (2001) developed a simple initialization scheme for the efficient HMT model training and then proposed a new four-state HMT model called HMT-2, it improved the performance of signal denoising over the two-state HMT model. Fan (2001) studied the training method of wavelet-domain HMMs and its applications to a variety of statistical image processing problems. Unlike all these research work related to image processing applications including image denoising, image segmentation, texture analysis and texture synthesis, this study focused on the application of HMM to solve the RD problems in the scalable video coding system. This study estimates the bit-rate through the entropy rate calculation of HMM. The research of the image processing shows that the wavelet-based video signal follows the mixture Gaussian distribution. The transform of an image consists of a small number of large coefficients and a large number of small coefficients and the smooth regions in the original image

Corresponding Author: Lizhi Zhang, School of Computer and Information Engineering, Harbin University of Commerce, Harbin, 150028, China

are represented by small coefficients in the coarsest resolution, while edges and ridges are represented by a few large coefficients. The Probability Density Function (pdf) of each wavelet coefficient can be roughly modeled by a two-state mixture Gaussian distribution, hence the tree-structured wavelet coefficients of video signal can be modeled by a HMM with two-hidden states, named Hidden Markov Tree (HMT). Based on HMM mathematical theory, the statistical characteristics of the wavelet video signal were analyzed and then the entropy rate was computed in this study.

THE HMT MODEL OF WAVELET DECOMPOSITION IMAGE SIGNAL

Wavelet decomposition image signal: Wavelet has become increasingly popular approach to decompose a signal recent years, the wavelet transform is a decomposition that represents a signal f (t) in terms of shifted and dilated versions of special selected wavelet function ψ (t) and the functions:

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k),\ \ j,k \in Z \tag{1}$$

Form an orthonormal basis and the signal f (t) can be decomposed as (Daubechies, 1992):

$$f(t) = \sum_{j,k \in Z} w_{j,k}\psi_{j,k}(t),\ \ w_{jk} = \int f(t)\psi^*_{j,k}(t)dt \tag{2}$$

One of the most popular applications of wavelets has been to image compression, the JPEG 2000 standard use wavelets instead of the DCT to perform decomposition of the image. However, images are two-dimensional signals and the decomposition procedure can be illustrated as Fig. 1. The Discrete Wavelet Transform (DWT) decomposes an image into a multi-resolution representation and the DWT of image signal can be shown in Fig. 2. The spatial subbands LL1, HL1, LH1 and HH1 can be obtained after a two-dimensional image wavelet decomposition. The first subband provides the main information of the image and the other three

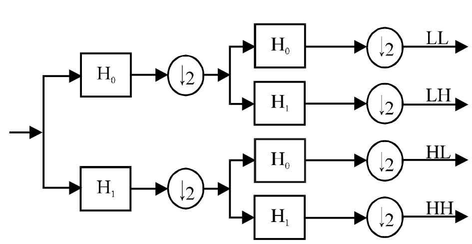

Fig. 1: Subband decomposition of an image

subbands provide the additional information of the image in the horizontal, diagonal and vertical orientations. The spatials subband LL2, HL2, LH2 and HH2 can be obtained after a two-dimensional image wavelet further decomposition to LL1. If LL2 is further decomposed, a growing quad-tree can be obtained as shown in Fig. 3 (Wainwright *et al*., 2001).

The statistical properties of image wavelet signal: The transform of an image consists of a small number of large coefficients and a large number of small coefficients and the smooth regions in the original image are represented by small coefficients in the coarsest resolution, while edges and ridges are represented by a few large coefficients. The pdf of each wavelet coefficient can be roughly modeled by a two-state mixture Gaussian distribution (Crouse *et al*., 1998). As shown in Fig. 4, f (w|s = 1) denotes the Gaussian distribution pdf of large coefficients and f (w|s = 2) denotes the Gaussian distribution pdf of small coefficients, respectively and f (w) denotes the pdf of the two-state zero-mean mixture Gaussian distribution.

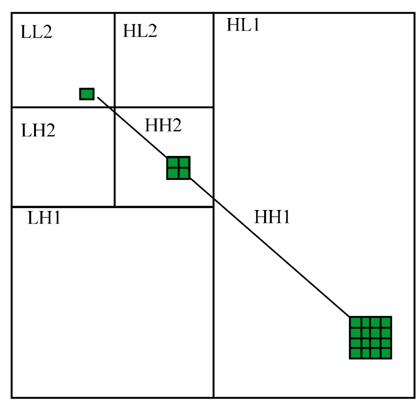

Fig. 2: The subband structure of image wavelet decomposition

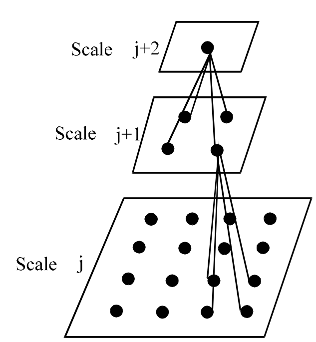

Fig. 3: The quad-tree structure of wavelet coefficient

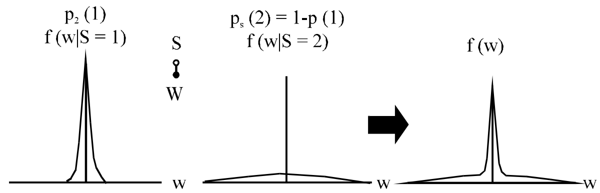

Fig. 4: A two-state zero-mean Gaussian mixture model

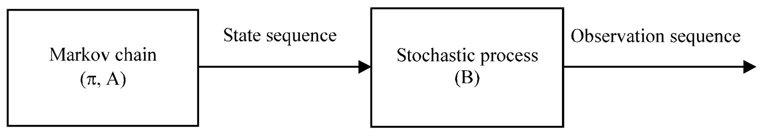

Fig. 5: A hidden Markov process

The HMT model of wavelet decomposition image signal: As the image wavelet coefficients follow two-state mixture Gaussian distribution, fitting the two-state of the mixture Gaussian distribution and the wavelet coefficient to the hidden state and the observation respectively, the wavelet image signal can be modeled as a HMM. And this HMM model can be considered as a Hidden Markov Tree (HMT) model due to the tree structure of the quad-tree wavelet coefficients.

The hidden Markov process is a double stochastic process, in which one is hidden and another is observation sequence as shown in Fig. 5. The compact notation $\lambda = (\pi, A, B)$ was used to indicate the complete parameter set of the model (Rabiner, 1989), where π denotes probability distribution of the hidden state, A denotes the state transition probability distribution and B denotes the observation symbol probability distribution in given state. This HMT model can be trained using the empirical data and the model parameters can be obtained, i.e., the hidden state transition probability distribution and the mixture Gaussian distribution variance of the wavelet coefficients. With the model parameters of HMT, the entropy rate of the HMT can be estimated and this leads to bit-rate estimation of wavelet image signal.

THE JOINT ENTROPY ESTIMATION METHOD OF HIDDEN MARKOV PROCESS

A joint entropy estimation method of the stochastic process: A stochastic process is an indexed sequence of n random variables, hence by Jensen's inequality, the entropy of a stochastic process $\{X_i\}$ can be expressed as (Cover and Thomas, 1991):

$$H(X_1, X_2, ..., X_n) \le \sum_{i=1}^{n} H(X_i) \tag{3}$$

Equality holds if and only if X_i's are independent, otherwise the difference of the two terms above is the mutual information and also the compression limit of coding. Hence, the joint entropy of a process can be computed as (Cover and Thomas, 1991):

$$H(X_1, X_2, ..., X_n) = \sum_{i=1}^{n} H(X_i / X_{i-1}, ..., X_1) \tag{4}$$

where the conditional entropy can be expanded as:

$$\begin{aligned} & H(X_n / X_{n-1}, X_{n-2}, ..., X_1) \\ & = \sum_{x_{n-1} \in X_{n-1}} \sum_{x_{n-2} \in X_{n-2}} ... \sum_{x_1 \in X_1} p(x_{n-1}, x_{n-2}, ..., x_1) \\ & H(X_n / X_{n-1}, X_{n-2}, ..., X_1 = x_{n-1}, x_{n-2}, ..., x_1) \end{aligned} \tag{5}$$

where, it means that the entropy of a stochastic process can be calculated if the joint probability of the random sequences $p(x_{n-1}, x_{n-2}, ..., x_1)$ and the conditional probability $p(x_n/x_{n-1}, x_{n-2}, ..., x_1)$ are obtained. However in case of HMM, the joint probability and the conditional probability can be obtained by iterative formula based on the HMM model parameters, hence the conditional entropy can be obtained.

The basic mathematical formula about HMM: For the hidden Markov process, the compact notation $\lambda = (\pi, A, B)$ indicates the complete parameter set of model. $A = \{a_{ij}\}$, $0 \le i, j \le M$, where M denotes the number of states, $a_{ij} = p(S_{n+1} = s'/S_n = s)$, where S denotes the hidden state of HMM and the state transition probabilities a_{ij} having the following constraints:

$$a_{ij} \ge 0 \tag{6}$$

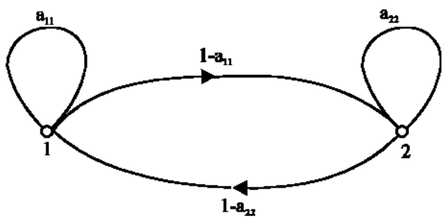

Fig. 6: The two hidden states transitions

$$\sum_{j=1}^{M} a_{ij} = 1 \quad (7)$$

Since the wavelet coefficients of an image consist of large coefficients and small coefficients, there is two hidden states. A is a 2*2 matrix and the element of the matrix corresponds to the conditional probability distribution about the two hidden states transition, The two hidden states transitions can be illustrated as Fig. 6.

B is a matrix with 2 rows and L columns, which denotes the observation symbol probability distribution in given state. L is the quantization intervals. The element of the matrix B is:

$$B[s, x] = p\,(X_n = x_n/S_n = s) \quad (8)$$

where, X denotes the observations, i.e., the wavelet coefficients and the probability distribution of the initial state is denoted by π_1:

$$\pi_n\,(x_{n-1}, x_{n-2},..., x_1) = p\,(s_n/x_{n-1}, x_{n-2},..., x_1) \quad (9)$$

where, π_n is the conditional probability distribution of the state in time n and it is associated with the past observations. The following equation can define a recursive relation for π_n, due to the Markovity. It is determined by he model parameters A and B (Rezaeian, 2006):

$$\pi_n = \eta(x_{n-1}, \eta_{n-1}) \quad (10)$$

$$\eta(x_{n-1}, \pi_{n-1}) = \frac{\pi_{n-1} D(x_{n-1}) A}{\pi_{n-1} D(x_{n-1}) \underline{1}} \quad (11)$$

where, D (x_{n-1}) is a diagonal 2*2 matrix with kth diagonal term $d_{k,k}$ (x_{n-1}) = B[k, x_{n-1}] k = 1.2, is $\underline{1} = (1,...,1)^T$ a k-dimensional vector. Eq. 8 and 9 yields:

$$\pi_n B = p\,(x_n/x_{n-1},..., x_{1-2,..} x_1) \quad (12)$$

THE BIT-RATE ESTIMATION OF DWT BASED IMAGE SIGNAL

HMT model training based on EM algorithm: As the image wavelet coefficient follows two-state mixture Gaussian distribution and can be modeled by HMT. The EM (Expectation-Maximization) algorithm is a broadly applicable method for calculating maximum likelihood estimates given incomplete data. It can be used to train HMT model of wavelet image signal and can be described as the following steps (Crouse *et al*., 1998):

Step 1: Initialization: choose initial model parameters

Step 2: E step, computing probabilities for the hidden state of the wavelet coefficients. It includes the up step, i.e., transmitting the hidden state information up the hidden Markov tree and the down step, i.e., transmitting the hidden state information down the hidden Markov tree

Step 3: M step, refresh the model parameters to achieve the expected likelihood maximum function

Step 4: Repeat the above E and M operations until convergence

The transition probability distribution of the two hidden states and the wavelet coefficients variance of the mixture Gaussian distribution are obtained after EM algorithm training. Since, it is required for the practical coding system to implement the image quality scalability, the wavelet coefficients should be processed in successive finer quantization. In this study, it takes the dead zone successive-approximation quantization scheme; hence the probability of signals in the different quantization intervals can be calculated. Together with the transition probability of the two hidden states, the entropy rate of HMM can therefore be estimated.

The joint probability distribution for the random sequence: For HMM observations, p $(x_1) = \pi_1 B$ is the initial distribution, where π_1 is the initial distribution of the hidden state, hence the joint probability distribution for the random sequence can be calculated by the following recursive formula:

$$p\,(x_n, x_{n-1},..., x_1) = p\,(x_{n-1}, x_{n-2},..., x_1)\pi_n B \quad (13)$$

Thus, for all the possible X_n, X_{n-1},..., X_1, the joint probability distribution can be obtained.

The conditional entropy for the given random sequence: The conditional entropy for the given random sequence can be calculated by:

$$\begin{aligned} &H(X_n / X_{n-1}, X_{n-2}, ..., X_1 = x_{n-1}, x_{n-2}, ..., x_1) \\ &= -\sum_{x_n \in X_n} p(x_n / x_{n-1}, x_{n-2}, ..., x_1) \log p(x_n / x_{n-1}, x_{n-2}, ..., x_1) \end{aligned} \quad (14)$$

where, p ($x_n/x_{n-1},..., x_{n-2},..., x_1$) = $\pi_n B$, HMM is a doubly embedded stochastic process with underlying stochastic process that is not observable (it is hidden), but can be observed through another set of stochastic process that produces the sequence of observations. The relationship between observation and state are not one to one mapping, the observation is a probabilistic function of the state. The mathematical deduction above shows that, π_n plays an important role among the model parameters that is the critical parameter linking the observation and the hidden state and also essential for computing joint and conditional distribution of the random sequence and hence for the conditional entropy.

Conditional entropy estimation: Since already obtained the joint probability distribution for the random sequence p ($x_{n-1}/x_{n-2},..., x_1$) at time n, together with H ($X_n/X_{n-1}, X_{n-2}...,$ X_1 = $x_{n-1}, x_{n-2},..., x_1$), the conditional entropy can be obtained as:

$$\begin{aligned} &= \sum_{x_{n-1} \in X_{n-1}} \sum_{x_{n-2} \in X_{n-2}} \cdots \sum_{x_1 \in X_1} p(x_{n-1}, x_{n-2}, ..., x_1) \; H(X_n / X_{n-1}, X_{n-2}, ..., X_1 = x_{n-1}, x_{n-2}, ..., x_1) \\ &= \sum_{x_{n-1} \in X_{n-1}} \sum_{x_{n-2} \in X_{n-2}} \cdots \sum_{x_1 \in X_1} p(x_{n-1}, x_{n-2}, ..., x_1) \left[-\sum_{x_n \in X_n} p(x_n / x_{n-1}, x_{n-2}, ..., x_1) \; \log p(x_n / x_{n-1}, x_{n-2}, ..., x_1) \right] \end{aligned} \quad (15)$$

This conditional entropy is the entropy of the observations under all the possible observation sequences. It has decorrelated the residue dependencies and it's the entropy of an each tree of the image DWT signal. With that, it is easy to estimate the entropy of an entire image frame.

THE EXPERIMENTAL RESULTS AND ANALYSIS

The above EM and HMT bit-rate estimation algorithm were implemented in the MATLAB experimental environment and the proposed approach was compared with the empirical histogram method. For the histogram based method to estimate the mutual information between the random variable X and Y, it first partition the range X and Y into intervals with the respective resolution and then calculate the joint and marginal probabilities. Hence the estimated mutual information can be shown as follows (Cover and Thomas, 1991) and this is the bit-rate limited bound that could be compressed:

$$I(X;Y) = E_{XY}\left[\log \frac{p(x,y)}{P(x)P(y)}\right] = \sum_x \sum_y P(x,y) \log \frac{p(x,y)}{P(x)P(y)} = D\left(P(x,y) \| P(x)P(y)\right) \quad (16)$$

Table 1: The entropy rate estimation of the proposed algorithm (HMT) versus the histogram based method (HIST)

	Suzie		Foreman		Mother and Daughter		News	
T	HIST	HMT	HIST	HMT	HIST	HMT	HIST	HMT
T_0	0.0798	0.0779	0.1366	0.1921	0.0847	0.0713	0.3287	0.2646
T_1	0.2632	0.2547	0.4177	0.3895	0.2755	0.2371	0.7696	0.7768
T_2	0.6454	0.4648	0.9378	0.9267	0.7496	0.7010	1.3742	1.5092
T_3	1.1967	1.2741	1.6022	1.6128	1.4082	1.5461	2.0163	2.1111
T_4	1.8846	1.8575	2.2255	2.0325	2.1041	2.1652	2.5723	2.5858

T: Quantization step, Values are expressed in bits wavelet

For the experiment, the standard "Foreman, Suzie, Mother and Daughter, News" test sequences were chosen, the t+2D of MCTF scheme was performed with a three level temporal decomposition and a Group of Frames (GOP) is 8. Then a three level spatial decomposition to the LLL frame was performed. Applying the EM algorithm to train the model of the LLL frame, the model parameter can be obtained and applying the entropy estimation algorithm, the entropy rate can be estimated. For the quantization scheme, it takes a quantization step size sequence T_0, T_1,...,T_{n-1}, satisfying T_i, $T_i/2$ and $2T_0$,>$|X_{max}|$, where $|X_{max}|$ is the maximum value of the wavelet coefficients.

Table 1 shows the LLL frames entropy rate estimation experimental results of the proposed algorithm versus the histogram based method in case of different quantization step. It shows that the algorithm is accurate and the average deviation of the two methods is within 10%, for example, in case of the "Suzie" test sequence, the entropy rates are 0.0798 and 0.0779 bits per wavelet coefficient when the quantization step size T taking T_0 = $|X_{max}|/2$ and the deviation is about 2.4%. While the quantization step size T taking T_4 = $|X_{max}|/32$, the entropy rates are 1.8846 and 1.8575 bits per wavelet coefficient and the deviation is about 1.4%. This entropy rate estimation has great significance to the RD model construction for the video coding system.

CONCLUSIONS

Based on statistical investigation of image wavelet coefficients, in this study, a HMM model has been trained with EM algorithm and then the model parameters have been obtained. Hence the joint probability and the conditional entropy of the given observation sequence has been iteratively calculated and then a bit-rate estimation approach has been obtained based on HMM model. The algorithm presents an effective and analytical method to the RD model construction for the scalable video coding system. Moreover, the algorithm can also be used to solve the problem of the channel capacity estimation of the communication system.

ACKNOWLEDGMENTS

This study is supported by National Natural Science Foundation of China (60772157), Project of Heilongjiang Province Education Department (12511125) and The natural science foundation of Heilongjiang Province (F201114). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

Anwar, W., X. Wang, L. Li and X. Wang, 2007. Hidden Markov model based part of speech tagger for Urdu. Inform. Technol. J., 6: 1190-1198.

Chiang , T. and Y.Q. Zhang, 1997. A new control scheme using quadratic rate distortion model. IEEE Trans. Circuits Syst. Video Technol., 7: 246-250.

Cover, T.M. and J.A. Thomas, 1991. Element of Information Theory. John Wiley and Sons, Inc., New York, NY., USA.

Crouse, M.S., R.G. Baraniuk and R.D. Nowak, 1996. Hidden Markov models for wavelet-based signal processing. Proceedings of the 13th Asilomar Conference on Signals Systems and Computers, Volume 2, November 3-6, 1996, USA., pp: 1029-1035.

Crouse, M.S., R.G. Baraniuk and R.D. Nowak, 1997. Signal estimation using wavelet-Markov models. Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, April 21-24, 1997, Munich, Germany, pp: 3429-3432.

Crouse, M.S., R.D. Nowak and R.G. Baraniuk, 1998. Wavelet-based statistical signal processing using hidden Markov models. IEEE Trans. Signal Process., 46: 886-902.

Dai, M., D. Loguinov and H. Radha, 2004. Rate-distortion modeling of scalable video coders. IEEE Int. Conf. Image Process., 2: 1093-1096.

Daubechies, I., 1992. Ten Lectures on Wavelets. SIAM, Philadelphia, USA.

Ding, W. and B. Liu, 1996. Rate control of MPEG video coding and recording by rate-quantization modeling. IEEE Trans. Circuits Syst. Video Technol., 6: 12-20.

Fan, G., 2001. Wavelet domain statistical image modeling and processing. Ph.D. Thesis, University of Delaware, USA.

Fan, G. and X. Xia, 2001. Improved hidden Markov models in the wavelet-domain. IEEE Trans. Signal Process., 49: 115-120.

Mallat, S. and F. Falzon, 1998. Analysis of low bit rate image transform coding. IEEE Trans Signal Process., 46: 1027-1042.

O'Neal, J. and T.R. Natarajan, 1997. Coding Isotropic Images. IEEE Trans. Inf. Theory, 23: 697-707.

Rabiner, L.R., 1989. A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE, 77: 257-286.

Rezaeian, M., 2006. Hidden Markov process: A new representation, entropy rate and estimation entropy. http://arxiv.org/pdf/cs/0606114.pdf

Sakrison, D.J., 1968. A geometric treatment of the source encoding of a Gaussian random variable. IEEE Trans. Inform. Theory, 14: 481-486.

Stuhlmuller, K., N. Farber, M. Link and B. Girod, 2000. Analysis of video transmission over lossy channels. IEEE J. Sel. Areas Commun., 18: 1012-1032.

Wainwright, M.J., E.P. Simoncelli and A.S. Willsky, 2001. Random cascades on wavelet trees and their use in modeling and analyzing natural imagery. Applied Comput. Harmon. Anal., 11: 89-123.

Wang, M. and M. Van der Schaar, 2006. Operational rate distortion modeling for wavelet video coders. IEEE Trans. Signal Process., 54: 3505-3517.

Wang, Y., K.F. Loe, T. Tan and J.K. Wu, 2005. A dynamic hidden Markov random field model for foreground and shadow segmentation. Proceedings of the 7th IEEE Workshops on Application of Computer Vision, January 5-7, 2005, Breckenridge, CO., USA., pp: 474-480.