WAVELET-FUZZY-STOCHASTIC KALMAN FILTERING FOR IMAGE COMPRESSION

Nastooh Avesta

Department of Information Technology
University of Turku
Turku, Finland

ABSTRACT

This paper presents a novel fuzzy stochastic Kalman filter for compression of digital images. In particular, it is shown that the state evolution of the synthesis coefficients of any Discrete Wavelet Transform (DWT), in presence of coding degradation, may be described fuzzily. The novelty of this description is that, unlike other fuzzy based methods, it does not require a predefined membership measure. The fuzzy representation is further characterized by a stochastic nominal value and an interval of uncertainty. Furthermore, traditional DCT based coding is judicially applied to the smooth regions of the DWT. It is shown that such a framework allows for an efficient coding of images.

1. INTRODUCTION

The similarity of the continuous sections of an image, in various DWT depths, has inspired plethora of image compression methods. These methods range from heuristic [1] to fractal [4] and fuzzy [2]. In this paper, we resort to the complementary aspects of fuzzy and stochastic systems to model and control the uncertainties introduced due to compression process. We will find it convenient to proceed under the umbrella of the state evolution description of the DWT coefficients. To this end, the Fuzzy Stochastic Kalman Filter (FSKF) is introduced in Section 2. Section 3 presents the FSKF algorithm along with representative examples. Finally, Section 4 concludes this paper with the summary of the current work and possible future venues.

2. FSKF

A deterministic state description of the DWT analysis and synthesis operations, with H_0 and H_1 as scale and wavelet analysis filters; F_0 and F_1 as their synthesis counterparts, is presented in (1).

$$x_k(n) = F_0^{k-1^M} x_{k-1}(n) + F_1^{k-1^M} w_{k-1}(n)$$
$$[x_{k-1}(n) \cdots x_{k-1}(n-M)]^T = H_0^{k-1^M} x_k(n+N-1) \quad , \quad (1)$$
$$[w_{k-1}(n) \cdots w_{k-1}(n-M)]^T = H_1^{k-1^M} x_k(n+N-1)$$

where $F_0^{k-1^M}, H_0^{k-1^M}$ and x_k are defined in (2). ($F_1^{k-1^M}, H_1^{k-1^M}, x_{k-1}$ and w_{k-1} are defined, similarly.) Also, F_0^{k-1} and H_0^{k-1} are defined in (3). In (1), (2) and (3) n refers to the time index and k refers to the resolution depth. N is the length of the filter and M is size of the observed signal. It is noted that this model explicitly takes into account the intra-dependencies of the M neighboring samples.

$$x_k(n) = \begin{bmatrix} x_k(n) \\ \vdots \\ x_k\left(n - N\left(\frac{M}{2}+1\right)+1\right) \end{bmatrix}_{N\left(\frac{M}{2}+1\right) \times 1}$$

$$F_0^{k-1^M} = \begin{bmatrix} F_0^{k-1}, 0, & \cdots & ,0 \\ 0, \cdots 0, F_0^{k-1}, 0, \cdots, 0 \\ & \ddots \\ 0, & \cdots & ,0, F_0^{k-1} \end{bmatrix}_{N\left(\frac{M}{2}+1\right) \times N\left(\frac{M}{2}+1\right)} \quad (2)$$

$$H_0^{k-1^M} = \begin{bmatrix} H_0^{k-1}, 0, & \cdots & ,0 \\ 0, \cdots 0, H_0^{k-1}, 0, \cdots, 0 \\ & \ddots \\ 0, & \cdots & ,0, H_0^{k-1} \end{bmatrix}_{M \times N\left(\frac{M}{2}+1\right)}$$

$$F_0^{k-1} = \begin{bmatrix} F_0^{k-1}(0), F_0^{k-1}(2), \cdots, F_0^{k-1}(N-1), & 0 \cdots 0 \\ 0, F_0^{k-1}(1), F_0^{k-1}(3), \cdots, F_0^{k-1}(N), & 0 \cdots 0 \\ 0, F_0^{k-1}(0), F_0^{k-1}(2), \cdots, F_0^{k-1}(N-1), 0 \cdots 0 \\ \ddots \\ 0, \cdots 0, & F_0^{k-1}(1), F_0^{k-1}(3), \cdots, F_0^{k-1}(N) \end{bmatrix}_{N \times N} \quad (3)$$

$$H_0^{k-1} = \left[H_0^{k-1}(0), H_0^{k-1}(1), \cdots, H_0^{k-1}(N-1) \right]_{1 \times N}$$

Equation (1) can be readily extended to the 2-D case, by the separable extension of the analysis and synthesis filters, as noted in (4). In (4) the S and I subscripts indicate the significant and insignificant wavelet filters, where separation of the subbands to significant, *i.e.* x_{k-1}, and insignificant, *i.e.* w_{k-1}, is based on the energy content of the respective subbands. Superscripts i to l refer to significant subbands, whereas superscripts a to d refer to insignificant subbands. N_S is the number of significant sub-bands. G_S^{k-1}, G_I^{k-1}, L_S^{k-1}, x_{k-1}, and w_{k-1} are defined in (5). $N_I = 4 - N_S$ is the

number of insignificant subbands. Furthermore, the separable synthesis filters are defined as in (6). (Analysis filters are defined in a similar manner.)

$$x_k(n) = G_S^{k-1} x_{k-1}(n) + G_I^{k-1} w_{k-1}(n)$$

$$\left[x_{k-1}^i(n) \cdots x_{k-1}^i(n-M), \cdots, x_{k-1}^a(n) \cdots x_{k-1}^d(n-M) \right]^T \quad (4)$$

$$= L_S^{k-1} x_k\left(n + N_s N^2 - 1\right)$$

$$G_S^{k-1} = \left[G_i^{k-1} | \ldots | G_l^{k-1} \right]_{N^2 \times N_S N^2} \, G_I = \left[G_a^{k-1} | \ldots | G_d^{k-1} \right]_{N^2 \times N_I N^2}$$

$$x_{k-1}(n) = \begin{bmatrix} x_{k-1}^j(n) \\ \vdots \\ x_{k-1}^l(n) \end{bmatrix}_{N_S N^2 \times 1} \, w_{k-1}(n) = \begin{bmatrix} x_{k-1}^a(n) \\ \vdots \\ x_{k-1}^d(n) \end{bmatrix}_{N_I N^2 \times 1} \, L_S^{k-1} = \begin{bmatrix} L_i^{k-1} \\ \vdots \\ L_l^{k-1} \end{bmatrix}_{N_S \times N^2}$$

(5)

$$G_0^{k-1} = F_0^{k-1^M} \otimes F_0^{k-1^M}, G_1^{k-1} = F_0^{k-1^M} \otimes F_1^{k-1^M}$$

$$G_2^{k-1} = F_1^{k-1^M} \otimes F_0^{k-1^M}, G_3^{k-1} = F_1^{k-1^M} \otimes F_1^{k-1^M} \quad (6)$$

Impairments and uncertainties are introduced in (4) due to quantization or when one attempts to perform either inter or intra band predictions of the DWT coefficients. The goal of the FSKF is to construct an implicit fuzzy membership measure, as a function of the DWT coefficients and depth, such that for a prescribed quality, the overall reconstruction quality is insensitive to the approximations of the DWT coefficients. To this end, the membership measure is described by its α-cuts [3] at different depths. These in turn are described by a nominal value and an interval of uncertainty. The effects of the quantization and prediction are then modeled by additive stochastic noise on the nominal values and expansion of the intervals of uncertainty. A Kalman filter is then developed that attempts to minimize the prediction error of the nominal value, while simultaneously minimizing the interval of uncertainty of the error signal. Furthermore, to avoid application of interval arithmetic [6], intervals are modeled as total differentials [7]. We begin by expressing the objective function (7), in terms of its nominal value, P_k^N, and its total differential, dP_k, as noted in (8).

$$\hat{x}_k = \arg\min tr(P_k), \quad (7)$$

where tr(.) is the trace operator.

$$tr(P_k) = E\left[tr\left[\left(\mathbf{e}_k^N + de_k \right)\left(\mathbf{e}_k^{N^T} + de_k^T \right) \right] \right]$$

$$= E\left[tr\left[\mathbf{e}_k^N \mathbf{e}_k^{N^T} \right] \right] + E\left[tr\left[de_k de_k^T \right] \right] = tr(P_k^N) + tr(dP_k) \quad (8)$$

where $P_k = E\left[e_k e_k^T \right]$ and e_k is expressed as in (9)[5].

$$e_k = \left(I - K_k L_S^{k-1} \right)\left(G_S^{k-1} \hat{x}_{k-1} + G_I^{k-1} \hat{w}_{k-1} - x_{k-1} \right) - K_k v_k$$

$$= \left(I - K_k L_S^{k-1} \right) e_k^- - K_k v_k \quad (9)$$

where I is the identity matrix, e_k^- is the a-priori error estimate, K_k is the optimal Kalman gain, v_k is the nominal observation noise, $\hat{x}_{k-1}$ and $\hat{w}_{k-1}$ are the estimated

constituent signals at resolution level k-1. The optimal estimate, $\hat{x}_k$, is of the form (10).

$$\hat{x}_k = G_S^{k-1} \hat{x}_{k-1} + G_I^{k-1} \hat{w}_{k-1} + K_k \begin{pmatrix} x_{k-1} + v_k - \\ L_S^{k-1}\begin{pmatrix} G_S^{k-1} \hat{x}_{k-1} + \\ G_I^{k-1} \hat{w}_{k-1} \end{pmatrix} \end{pmatrix} \quad (10)$$

To obtain a convenient expression for the differential term of (8), it is noted that the differential of the posteriori error, de_k, may be expressed in terms of its a-priori error, de_k^-, as in (11).

$$de_k = abs\left(\frac{\partial\left(I - K_k L_S^{k-1} \right)e_k^-}{\partial e_k^{-T}} \right) de_k^-,$$

$$= abs\left(I - K_k L_S^{k-1} \right) de_k^- \quad (11)$$

where de_k^- is as in (12) and *abs(M)* denotes a matrix that its elements are equal to the absolute values of the elements of the matrix M.

$$de_k^- = abs\left(\frac{\partial G_S^{k-1} \hat{x}_{k-1}}{\partial \hat{x}_{k-1}^T} \right) d\hat{x}_{k-1} = abs\left(G_S^{k-1} \right) d\hat{x}_{k-1} \quad (12)$$

Noting that $abs\left(I - K_k L_S^{k-1} \right) \leq I + abs(K_k) abs(L_S^{k-1})$, an expression of for the upper bound of the differential of the posteriori error covariance matrix is obtained in (13).

$$tr(dP_k) \leq tr\begin{pmatrix} abs(K_k) abs(L_S^{k-1}) dP_k^-. \\ abs\left(L_S^{k-1^T} \right) abs\left(K_k^T \right) \end{pmatrix} + 2tr\left(abs(K_k) abs(L_S^{k-1}) dP_k^- \right) + dP_k^- \quad (13)$$

where dP_k^- is the differential of the a-priori error covariance matrix, as shown in (14).

$$dP_k^- = abs\left(\frac{\partial P_k^-}{\partial P_{k-1}} \right) dP_{k-1} = abs\left(\frac{\partial\left(G_S^k P_{k-1} G_S^{k^T} \right)}{\partial P_{k-1}} \right) dP_{k-1}$$

$$= abs\left(G_S^k G_S^{k^T} \right) dP_{k-1} \quad (14)$$

To obtain the optimum Kalman gain, both terms of (8) are differentiated with respect to K_k, equated to zero and solved for K_k. Differentiating the second term of (8) with respect to K_k, equating to zero, and solving the equation for K_k, (15) is obtained.

$$K_k^d = -diag\left(sgn(K_k^N) \right) abs(L_S^{k-1}) abs(dP_k^-)$$

$$\left[abs(L_S^{k-1}) abs(dP_k^-) abs\left(L_S^{k-1^T} \right) \right]^{-1}, \quad (15)$$

where superscript d refers to the differential, $sgn(V)$ is the sign-vector of vector V, $diag(V)$ is a square matrix, of appropriate size, whose diagonal elements are those of vector V. As for the differentiation of the first term of (8), it is noted that the DWT coefficient prediction may be performed on intra or inter bands. With respect to the former, it is seen that by introducing the memory parameter M in (1), the intra-band estimations are explicitly obtained.

With respect to the latter, the nominal values of insignificant bands, *i.e.* w_{k-1}, are estimated based on the nominal values of the significant bands, x_{k-1}. The estimation error of the nominal values is then modeled as a stochastic white Gaussian noise, as in (16).

$$\hat{x}_k^N(n) = G_S^{k-1}\hat{x}_{k-1}^N(n) + G_I^{k-1}\hat{w}_{k-1}^N(n) + G_I^{k-1}i_{k-1}^N + G_S^{k-1}s_{k-1}^N$$

$$\left[\hat{x}_{k-1}^{N^i}(n)\cdots\hat{x}_{k-1}^{N^l}(n-M),\cdots,\hat{x}_{k-1}^{N^a}(n)\cdots\hat{x}_{k-1}^{N^d}(n-M)\right]^T \qquad , (16)$$

$$= L_S^{k-1}\hat{x}_k^N\left(n + N_s N^2 - 1\right)$$

where $\hat{x}_k^N$ is the nominal value of the estimated signal at depth k, i_{k-1}^N is the nominal estimation error of w_{k-1}, s_{k-1}^N is the nominal estimation error of x_{k-1} and $\hat{w}_{k-1}^N$ is obtained by minimizing $E\left[tr\left(e_{k-1}^{N^T}e_{k-1}^N\right)\right]$, *i.e.* the observable error. Noting (4), it is trivial to show that the optimal value of $\hat{w}_{k-1}^N$ is as shown in (17).

$$\hat{w}_{k-1}^N = G_I^{k-1^R}\left[\hat{x}_k^N - G_S^{k-1}\hat{x}_{k-1}^N\right] , \qquad (17)$$

where M^R is the right-inverse of matrix M. Furthermore, quantization effect is modeled by introducing noise to the observation process of (16), as noted in (18).

$$\hat{x}_k^N(n) = G_S^{k-1}\hat{x}_{k-1}^N(n) + G_I^{k-1}\hat{w}_{k-1}^N(n) + G_I^{k-1}i_{k-1}^N + G_S^{k-1}s_{k-1}^N$$

$$\left[\hat{x}_{k-1}^{N^i}(n)\cdots\hat{x}_{k-1}^{N^l}(n-M),\cdots,\hat{x}_{k-1}^{N^a}(n)\cdots\hat{x}_{k-1}^{N^d}(n-M)\right]^T \qquad , (18)$$

$$= L_S^{k-1}\hat{x}_k^N\left(n + N_s N^2 - 1\right) + G_I^{k-1}Q_{k-1}^w + G_S^{k-1}Q_{k-1}^x$$

where Q_{k-1}^x and Q_{k-1}^w are the quantization noises of processes x and w. It is easy to show that the optimal Kalman gain of the nominal process is (19).

$$K_k^N = \left(L_S^{k-1}P_k^{N^-}\right)^T\left[\begin{array}{l}L_S^{k-1}P_k^{N^-}L_S^{k-1^T} + G_S^{k-1}R_{k-1}^x G_S^{k-1^T} + \\ G_I^{k-1}R_{k-1}^w G_I^{k-1^T}\end{array}\right]^{-1} , (19)$$

where R_{k-1}^x is the variance of Q_{k-1}^x and R_{k-1}^w is defined similarly. The overall Kalman gain is the sum of (15) and (19), as shown in (20).

$$K_k = K_k^d + K_k^N \qquad (20)$$

Once the approximation signal $\hat{x}$ is obtained, a residual error is formed as in (21)

$$r_k = x_k - \hat{x}_k \qquad (21)$$

At this point, a recursive refinement process, based on FSKF, is initiated to obtain the approximation of residual error r_k, *i.e.* $\hat{r}_k$.

3. FSKF CODEC ALGORITHM

The operation of the compressor may be broken into two parts: the encoder and the decoder.

Encoder
1. Perform 2-D DWT on the current input

2. Based on energy content of the resulting subbands, identify the significant and insignificant subbands
3. For the significant subbands identify the disparity ratio, *i.e.* the ratio of the difference between the maximum and minimum values of the subband to the total energy of the subband.
4. If the disparity ratio is higher than a prescribed value, then perform 1-3 on each significant subband
5. Else employing (10), (15), (17), (19) and (20) obtain the optimal estimate $\hat{x}_k$, using the quantized DWT coefficients
6. Obtain r_k, as in (21), and perform 1-6
7. Repeat till maximum depth is reached
8. At the maximum depth, obtain the Discrete Cosine Transform (DCT) of the subbands and use the quantized coefficients for reconstruction.

The motivation for step 8 is based on the observation that as the depth of the DWT tree increases, the subbands become smoother. Therefore, it is judicial to model such subbands based on their smoothness, *i.e.* using DCT, as oppose to their disparities, *i.e.* using FSKF.

The decoder receives the topology of the tree, nominal value of each subband along with its range of variation, quantized coefficients of the elements that could not have been estimate and the quantized DCT coefficients of the maximum depth.

Decoder
1. Using the DWT topology, quantized coefficients at depth k-1, (10), (15), (17), (19) and (20) reconstruct the subbands at depth k
2. Repeat till depth 1 is reached.

Simulation Results

In this section, we provide a number of typical simulations to verify the applicability of the FSKF in image compression. In particular, comparison with similar, highly efficient, DWT based method, *i.e.* SPIHT [8] is provided. The results are tabulated in Table 1. The setup is based on the 8 tap Daubechies wavelet with maximum depth of 7. 1-4 bit quantization is performed at all depths, depending on the variance of the subband. The disparity threshold is set to 10^{-3}. It is evident, from Figure 1 to Figure 6 that at extremely low bit rates, which include all side information, the FSKF achieves a high quality reconstruction.

Table 1 SPIHT and FSKF Comparison

Image Name	Bit Rate [bpp]	SPIHT PSNR [db]	FSKF PSNR [db]
Lenna	0.103	24.38	25.88
Barbara	0.133	22.00	26.20
Baboon	0.165	20.61	23.66

Figure 1.   FSKF Compression

Figure 2.   SPIHT Compression

Figure 3.   FSKF Compression

Figure 4.   SPIHT Compression

Figure 5.   FSKF Compression

Figure 6.   SPIHT Compression

4.   CONCLUSION

In this paper, the utility of the fuzzy stochastic modeling of the DWT, for the purpose of image compression was demonstrated. It was noted that at extremely low bit rate such an approach allows for a high quality reconstruction. Furthermore, unlike other fuzzy based methods, the membership measure is not predefined and is customized to a particular image and required PSNR. The future work will concentrate on application of the smoothing Kalman filter in FSKF.

REFERENCES

[1] Said Amir and Pearlman A. William, "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," IEEE Trans. on Cir. and Sys., Vol. 6, No. 3, pp. 243-250, June 1996

[2] Chacon M.I., Aguilar L., Delgado A., "Definition and applications of a fuzzy image processing scheme," Digital Signal Processing Workshop, pp. 102 – 107, 2002

[3] D. Dubois and H. Prade, Fuzzy Sets and Systems, Academic Press, 1980

[4] Davis G. M., "A wavelet-based analysis of fractal image compression," IEEE Trans. on Image Proc., Vol. 7, No. 2, pp. 141-154, Feb. 1998

[5] Frank L. Lewis, Optimal Estimation, John Willey & Sons, 1986

[6] Ramon E. Moore, Interval Analysis, Prentice-Hall, 1966

[7] John R. Taylor, An Introduction to Error Analysis, University Science Books, 1997

[8] http://www.cipr.rpi.edu/research/SPIHT/spiht0.html