A BINARY WAVELET-BASED SCHEME FOR GRAYSCALE IMAGE COMPRESSION★

Hong Pan, Li-Zuo Jin, Xiao-Hui Yuan, Si-Yu Xia, Jiu-Xian Li, and Liang-Zheng Xia

School of Automation, Southeast University, Nanjing, P.R.China, 210096

ABSTRACT

This paper proposes a novel grayscale image compression approach using the binary wavelet transform (BWT) and context-based arithmetic coding, namely the context-based binary wavelet transform coding algorithm (CBWTC). In our CBWTC, in order to alleviate the degradation of predictability caused by the BWT and eliminate the correlation within the same level subbands, three highpass wavelet coefficients at the same location are combined to form an octave symbol and then encoded with a ternary arithmetic coder. The conditional context of the CBWTC is properly modeled by exploiting the properties of the BWT as well as taking the advantages of non-causal adaptive context modeling. Experimental results show that the coding performance of the CBWTC is better than that of the state-of-the-art grayscale image coders except for images containing rich texture, and always outperforms the JBIG2 algorithm and other BWT-based binary coding technique.

***Index Terms*—**Binary wavelet transform, context modeling, image coding

1. INTRODUCTION

Recently, many wavelet-based image coders [1-3] have proved successful applications of the high-order entropy coding which exploits the correlations of neighboring coefficients and produces an efficient code for them. On the other hand, most of the existing wavelet filters are designed in the real field for grayscale image compression. However, the facts that these filters enlarge the range of wavelet coefficients and bring about an expansion in the alphabet size of symbols could deteriorate the compressibility of wavelet coefficients by introducing more coding passes and expending extra bits for representing sign information of wavelet coefficients. Moreover, this symbol expansion dramatically increases the "model cost" of a context-based entropy coder in the compression of grayscale images which are usually represented by only eight alphabets. There have been several attempts to generalize wavelet transform to finite fields in order to take into account the image characteristics [4-8]. Swanson and Tewfik [4] introduced a new sequency-based transform applicable to sequences over *GF* (2) and constructed a theory of BWT for finite binary images. Kamstra [5-6] generalized the BWT by making it independent on the field structure of binary numbers and including nonlinear transforms as well. As a result, the increase of freedom in choosing a particular transform can be useful in binary image compression. Law, Pan and Siu [7-8] developed a set of length-independent binary filters for the BWT and applied them to lossless image compression.

The most important feature of the BWT is the conservation of alphabet size of wavelet coefficients, which indicates that the transformed images have the same number of grayscale levels as the original images. In particular, for a K-bit grayscale image, the range of BWT coefficients is still maintained within [0, 2^K-1]. Therefore, it is reasonable to expect that the compression efficiency of the BWT coefficients could be improved in that extra bits, originally used to code sign information of the transform coefficients, are saved to code more significant coefficients. In addition, without increasing the "model cost", further compression gains could be achieved by combining the BWT with a high-order context-based entropy coding scheme. The compression complexity might be reduced as the BWT contains simple exclusive-or (*XOR*) operations only and a maximum number of eight coding passes are involved during the encoding procedure. However, the application of the BWT to grayscale image compression is not well established. There still remains some problems to solve. For example, the BWT cannot help dispersing highly correlated binary pixels into many subbands, which degrades the predictability between BWT coefficients at different subbands. As a result, when we encode BWT coefficients using a context-based arithmetic coding method, it inevitably makes the performance of context modeling worse.

Motivated by the desire to use the efficiency of the high-order context-based entropy coding and taking into account the predictability degradation caused by the BWT, we propose a novel BWT coding approach called Context-based Binary Wavelet Transform Coding algorithm (CBWTC) to the grayscale image compression. In our CBWTC algorithm, we attempt to solve the above problems by introducing several key techniques. First, we develop a combined arithmetic coding scheme to alleviate the degradation of predictability caused by the BWT. In this scheme, three highpass wavelet coefficients at the same location are combined to form an octave symbol and then encoded by a ternary arithmetic coder. Exploiting the fact that the inverse BWT can perfectly synthesize the four neighboring samples from the corresponding four wavelet coefficients at a higher scale, the reconstructed lowpass coefficient is used in context modeling for the combined symbol. Then, in order to match up with this combined arithmetic coder, we design a new context modeling approach where a context consists of causal neighboring samples from the three highpass subbands and causal/non-causal samples from the corresponding lowpass subband image. It should be noticed that the proposed context modeling is performed after the corresponding synthesis of lowpass subband image. Additionally, we also introduce a technique to efficiently represent large regions with the same color, which frequently appears in bi-level images.

2. GENERALIZATION OF THE BWT TO GRAYSCALE IMAGES

★ This work was supported in part by the National Natural Science Foundation of China under Grant 60805002.

The procedure of the BWT is the same as the wavelet transform conducted in real field except that the original signal and the transformed signal are reconstructed to be binary. In order to guarantee that the BWT can perform a useful invertible multiscale decomposition that inherits the important properties of the real field wavelet transform (RWT), and the forms of the inverse filters are length-independent. The lowpass filter and highpass filter must satisfy four basic constraints: the perfect reconstruction constraint, bandwidth constraint, vanishing moment constraint, and perpendicular constraint. Readers may refer to [7-8] for theory and a fast implementation of the BWT.

Originally the BWT is designed for bi-level images. A natural way to generalize the BWT to grayscale images is to decompose a multilevel grayscale image into a series of bi-level bitplane images, and then perform the BWT to each individual bi-level bitplane image. The advantage of bitplane decomposition is that it offers a viable compression scheme whereby the image can be reconstructed progressively from the most significant bitplane (MSB) to the least significant bitplane (LSB).

In order to compare the properties of BWT transformed images with those of wavelet transformed images in other fields, we performed three-level transforms using binary wavelet, Daubechies 9/7 real wavelet and integer(2,2) wavelet to a 512×512 eight-bit 'Lena' image, respectively. As expected, the range of BWT coefficients maintain within [0, 255], whereas the ranges of RWT coefficients and integer wavelet transform coefficients are [-1337.712, 6767.419] and [-191, 264], respectively. It means that we can represent BWT coefficients with eight bits, on the contrary, 14 bits including 1 bit for sign information are required to represent the above RWT coefficients. If we encode these RWT coefficients using a fifth-order context model, then we need 2^{70} tables with 4096 probabilities in each table! This number is too large to even consider implementation in real time compression. In contrast, for BWT coefficients, we need 2^{40} tables with only 256 probabilities in each table.

3. CONTEXT-BASED BINARY WAVELET TRANSFORM CODER (CBWTC)

General architecture of the proposed CBWTC algorithm is illustrated in Fig.1. First, a grayscale image is decomposed into bi-level bitplane images and BWTs are applied to these bi-level bitplane images to remove the spatial correlation, then BWT coefficients are coded in a bitplane by bitplane manner until the given rate is achieved. For each bitplane, three coding passes, i.e. *LL_pass*, *Zerotree_pass* and *CombinedSymbol_ pass*, are involved successively to output bit streams generated by three individual coders that deal with different information based on their own context models.

In the following, let $\{LL_L\}$ and $\{HL_l, LH_l, HH_l \mid l = 1, 2,\ldots, L\}$denote the subband images generated by the BWT, where L is the transform level. LL_l, HL_l, LH_l, and HH_l are subband images at the scale l. In our coding algorithm, a zerotree map and a corresponding encoding mask are introduced to efficiently represent large mono-color regions which frequently appear in bi-level images. A zerotree map is a set of bits, where each bit corresponds to a coefficient in LL_L subband and is set to zero if all the descendents of the coefficient are zero, otherwise 1. If the bit is zero, then bits in the corresponding encoding mask for all the descendents are also set to zero so that the descendents are skipped from encoding. Zerotree map that is employed only in the lowest subband LL_L as a tradeoff between the quantity of side information to represent the map and the number of bits saved by skipping mono-color regions from encoding, is compressed with a context-based binary arithmetic coder (*Zerotree_coder*) in the *Zerotree_pass* which follows the encoding of LL_L subband with another context-based binary arithmetic coder (*LL_coder*) in the *LL_pass*. In original tree-structured wavelet coders such as EZW and SPIHT, the correlation within the same level subbands, e.g. HL_l, LH_l and HH_l, is often ignored and large redundancies are hidden in representation of these subband coefficients. In this regard, in our CBWTC, at each scale l ($l = L, \ldots,1$) , three highpass coefficients are combined to form an octave symbol and then jointly encoded by a combined ternary arithmetic coder as shown in Fig.1, where a symbol combiner combines three highpass coefficients at the same spatial location into an octave symbol. The coder then checks the corresponding encoding mask bit. If it is set to zero, the symbol is skipped from encoding. Otherwise, it is encoded by a context-based combined ternary arithmetic coder (*CombinedSymbol_coder*) in the *CombinedSymbol_pass*, conditioned on the corresponding lowpass coefficients in the LL_l subband which are reconstructed before. The reason for that can be explained by the nature of the BWT. In BWT, lowpass coefficients are obtained by sampling every other sample and highpass ones are obtained by performing *XOR* operations between the neighboring two samples [7]. As expected, there is no strong self-similarity of coefficients in highpass subbands across scales after the BWT. On the contrary, there is strong similarity in real wavelet coefficients of grayscale images, which is the motive of the study of tree-structured wavelet coders. Our research demonstrates that coefficients in highpass subbands are rather more correlated with those in the lowpass subband at the same scale. Hence, LL_l outperforms three highpass subbands at the scale ($l + 1$) in terms of the prediction of coefficients in HL_l, LH_l and HH_l.

In the CBWTC, different modeling strategies are applied to the bit streams yielded by three coders. For LL_L coefficients at the MSB, we only require to consider the intra-band correlation of the BWT coefficients. Thereby, a three-line template constructed by 10 adjacent pixels is adopted, whereas for LL_L coefficients at other bitplanes, since they contain both intra-band and inter-bitplane correlations, coefficients at the upper bitplane that are encoded previously, and the neighboring coefficients at the current bitplane are exploited in the context template to make an accurate prediction of the probability density function of the encoded symbol. For the *Zerotree_coder*, it employs the same template structures as the *LL_coder*. The context templates of *LL_coder* and *Zerotree_coder* are shown in Fig.2(a) and (b), where '*' denotes the reference coefficients and 'e' denotes the encoding coefficient.

Previous works show that a binary sample can be well predicted from its neighboring samples in the spatial domain, which is the starting point of context-based binary image compression methods [9-10]. However, the BWT degrades the predictabilities of binary samples by spreading out an edge into many subbands. As a result, the performance of a context-based binary image compression algorithm could be lowered by the BWT. In order to alleviate the degradation of predictability in the binary wavelet domain and eliminate the correlation within the same level subbands, we combine three highpass coefficients into an octave symbol using Eq.(1) and encode it with a combined arithmetic coder (i.e. *CombinedSymbol_coder*) based on a proper context model. The combined octave symbol $S(i, j)$ is defined as

$$S(i, j) = HL_l(i, j) + 2 \times LH_l(i, j) + 4 \times HH_l(i, j) \quad (1)$$

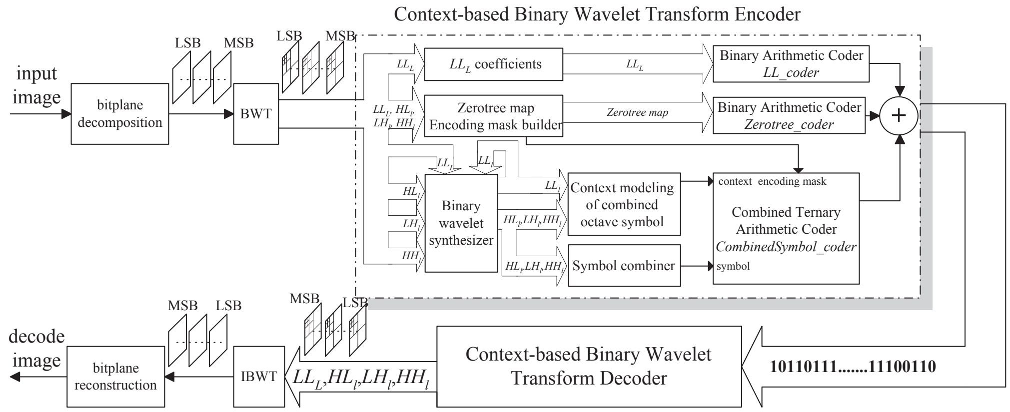
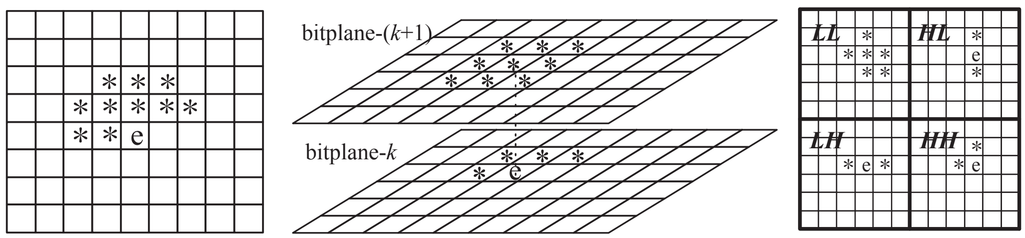


Figure 1. General architecture of the proposed CBWTC algorithm



(a) (b) (c)

(a) Context template for the *LL_coder* and the *Zerotree_coder* at the MSB, (b) Context template for the *LL_coder* and the *Zerotree_coder* at other bitplanes, (c) Context template for the *CombinedSymbol_coder*

Figure 2. Context modeling of the proposed CBWTC algorithm

where (i, j) denote the coordinates of reconstructed coefficients in LL_l subband. As mentioned before, LL_l should be reconstructed beforehand to exploit the fact that it provides a better prediction of HL_l, LH_l and HH_l than the three highpass subbands at the scale (l + 1). Thus, as demonstrated in Fig.2(c), twelve coefficients marked '*' are applied to model the contexts for S (i, j). Note that six coefficients in LL_l subband at the current bitplane are included in the context template where a lower-right coefficient is utilized to take the advantage of a non-causal prediction. Due to the relatively weak correlation of coefficients within each highpass subband, only two coefficients in each highpass subband are involved in context modeling. The choices of the two coefficients in each highpass subband are naturally adapted to the orientation characteristics of different subbands. The reference coefficient '*' in *HL* subband below the encoding coefficient 'e' and that in *LH* subband at the right of 'e' are chosen from the upper bitplane as context events. The modeling context of S (i, j) contains some future information if an octave-raster scanning order is used. In particular, this refers to the use of the abovementioned two upper bitplane coefficients from *HL* and *LH* subbands, as well as the lower-right coefficient from LL_l subband in the context modeling of S (i, j). The ability to look into the future information generally reduces the uncertainty of S (i, j), which also contributes to improving the coding efficiency of our CBWTC algorithm.

4. EXPERIMENTAL RESULTS

In this section, compression performance of the proposed CBWTC algorithm was appraised by comparing it with the state-of-the-art grayscale image coders like the SPIHT [11] and EBCOT [2], as well as the popular binary image coder (JBIG2) and other binary coding technique based on the BWT, i.e. BWT-IAC. The result of SPIHT is from the algorithm with arithmetic coding, and the result of EBCOT is from the generic case which provides the best quality scalability. JBIG2 [10] is an emerging standard using the context-based statistical modeling and arithmetic coding for lossy and lossless binary image compression. In this experiment, JBIG2 was used directly to encode bi-level images obtained by bitplane decomposition in a bitplane by bitplane manner from the MSB to the LSB, until the given rate was reached. BWT-IAC is a primitive BWT coder, which individually codes each highpass coefficient without utilization of the zerotree map and the combined octave symbol. A five-level BWT using the group-1binary filter listed in [7] was performed in our CBWTC. For the SPIHT and EBCOT coders, the Daubechies 9/7 kernel was used and the same level of RWT decompositions was performed.

Table 1 shows the PSNR results of all coders for a set of 512×512 eight-bit grayscale images at different bit rates. The jointly encoding of three highpass coefficients using a properly designed high-order context model makes the CBWTC algorithm exhibit the best performance among all tested coders for images like 'Lena', 'Goldhill' and 'Barbara', whereas, for image of 'Mandrill', the CBWTC is inferior to SPIHT and EBCOT by a slight deficit when bit rate is lower than 0.5 bpp. On average, coefficients at higher bitplanes contribute much more to distortion reduction than those at lower bitplanes. Hence, the effectiveness of the CBWTC can only be discovered after the coefficients at the most several significant bitplanes have been transmitted. However,

Table 1. PSNR results of all tested coders for a set of eight-bit grayscale images at different bit rates

Rates (bpp)	0.125	0.25	0.5	1.0	2.0
Lena (512×512)					
SPIHT	31.11	34.14	37.25	40.46	45.13
EBCOT	31.05	34.16	37.29	40.48	45.16
CBWTC	**31.24**	**34.37**	**37.56**	**40.74**	**45.58**
JBIG2	30.02	32.84	35.85	38.72	43.51
BWT-IAC	29.73	32.68	35.63	38.49	43.20
Goldhill (512×512)					
SPIHT	28.48	30.56	33.13	36.55	42.02
EBCOT	28.47	30.59	33.25	36.59	42.17
CBWTC	**28.58**	**30.77**	**33.45**	**36.91**	**42.43**
JBIG2	27.37	29.19	31.66	35.04	40.43
BWT-IAC	27.12	29.03	31.64	34.86	40.21
Barbara (512×512)					
SPIHT	24.86	27.58	31.40	36.41	42.65
EBCOT	25.37	28.40	32.29	37.11	43.11
CBWTC	**25.41**	**28.46**	**32.42**	**37.27**	**43.34**
JBIG2	24.39	26.78	30.66	35.46	41.28
BWT-IAC	24.04	26.70	30.51	35.32	40.87
Mandrill (512×512)					
SPIHT	**21.72**	**23.27**	25.65	29.18	34.99
EBCOT	21.68	23.26	25.97	29.46	35.13
CBWTC	21.70	23.26	**26.04**	**29.53**	**35.42**
JBIG2	20.42	21.85	24.66	28.01	33.79
BWT-IAC	20.06	21.50	24.27	27.65	33.59

for images like 'Mandrill' that contain rich texture structures and detailed elements, there exists more transform coefficients at high bitplanes than other types of images, which explains the performance degradation of the CBWTC at low bit rates. Among all tested bit rates, the average PSNR improvement of the CBWTC over EBCOT is 0.175dB for all images. Unlike the EBCOT that should output additional bits to code sign information of the wavelet coefficients, the bit stream yielded by the CWBTC contains no sign bits. Thus, sign bits, comprising 20% of the total output bits by the EBCOT coder on average, are saved to represent more significant coefficients that lead to better distortion reduction. Among all tested bit rates, the average PSNR improvement of the CBWTC over SPIHT is 0.397dB for all images. The SPIHT algorithm employs a structure-based model in the significance coding of the unknown pixels, whereas in our CBWTC, the significance coding is established on a context model. Compression advantages of the CBWTC over SPIHT can be interpreted by the introduction of a high-order context-based BWT compression scheme which achieves high coding efficiency without increasing extra "model cost" and the fact that information provided by the known pixels in significance coding is not sufficiently utilized in the structure-based model. The CBWTC algorithm outperforms the JBIG2 at all bit rates. The reason can be explained by the fact that JBIG2 is mainly dedicated to the compression of binary documents which contain text data, halftone data and generic data. Usually, a binary document image has different characteristics from the bi-level bitplane images of a grayscale image. Consequently, the dictionary-based coding technique involved in the JBIG2 can not expect to provide a satisfactory compression performance when used to code bi-level bitplane images. For all tested images, the BWT-IAC always shows the worst performance, which illustrates the fact of degradation of predictability caused by the BWT.

5. CONCLUSION

Motivated by the potential gains of the high-order context-based entropy coding when applied to grayscale image compression, we examine the BWT as an alternative approach to the RWT in order to bypass some of the implementation difficulties of the high-order context-based entropy coding. Furthermore, we extend the BWT to grayscale image compression and propose a novel CBWTC algorithm. By jointly representing three highpass coefficients using a combined octave symbol and encoding it with a context-based ternary arithmetic coder, the CBWTC not only minimize the degradation of predictability caused by the BWT, but also achieve additional compression gains by eliminating the correlation in the same level subbands. The proposed arithmetic coding is based on the context properly modeled by exploiting the nature of the BWT as well as taking the advantage of non-causal prediction capability of the BWT. A zerotree map is also introduced as the side information to effectively compress the large mono-color regions which frequently appear in bi-level images. Experimental results show that the CBWTC offers better performance than the state-of-the-art grayscale image coders on average, and always outperforms the JBIG2 algorithm and other BWT-based binary coding technique.

6. REFERENCES

[1] N.Zhang, and X.L.Wu, "Lossless compression of color mosaic images", *IEEE Trans. Image Processing*, vol.15, no.6, pp.1379-1388, 2006.

[2] D.Taubman, "High performance scalable image compression with EBCOT", *IEEE Trans. Image Processing*, vol.9, pp.1158-1170, 2000.

[3] Y.Zhang, and D.A.Adjeroh, "Prediction by partial approximate matching for lossless image compression", *IEEE Trans. Image Processing*, vol.17, no.6, pp.924-935, 2008.

[4] M.D.Swanson, and A.H.Tewfik, "A binary wavelet decomposition of binary images", *IEEE Trans. Image Processing*, vol.5, pp.1637-1650, 1996.

[5] L.Kamstra, "The design of linear binary wavelet transforms and their application to binary image compression", *Proc. IEEE Int.Conf. on Image Proc., ICIP'03*, pp.241-244, 2003.

[6] L.Kamstra, "Nonlinear binary wavelet transforms and their application to binary image compression", *Proc. IEEE Int.Conf .on Image Proc., ICIP'02*, vol.3, pp. 593-596, 2002.

[7] N.F.Law, and W.C.Siu, "A filter design strategy for binary field wavelet transform using the perpendicular constraint", *Signal Processing*, vol.87, no.11, pp.2850-2858, 2007.

[8] H.PAN, W.C. Siu, and N.F.Law, "Lossless Image Compression Employing Binary Wavelet Transform", *IET Image Process*, vol.1, no.4, pp.353-362, 2007.

[9] S.Xiao, and C.G.Boncelet, "On the use of context-weighting in lossless bilevel image compression", *IEEE Trans. Image Processing*, vol.15, no.11, pp.3253-3260, 2006.

[10] *"JBIG2 Final Draft International. Standards"*, ISO/IEC JTC1/SC29/WG1 N1545, Dec. 1999.

[11] A.Said, and W.A.Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical trees", *IEEE Trans. Circuits and Systems for Video Technology*, vol.6, pp.243-250, 1996.