LOSSLESS IMAGE COMPRESSION USING INTEGER TO INTEGER WAVELET TRANSFORMS

A. R. Calderbank[†] *Ingrid Daubechies*[‡] *Wim Sweldens*[*] *Boon-Lock Yeo*[**]

[†]AT&T-Labs Research, 180 Park Ave Bldg 103, Florham Park, NJ 07932
[‡]Program for Applied and Computational Mathematics, Princeton University, Princeton NJ 08544
[*]Lucent Technologies, Bell Laboratories, Rm. 2C-175, 700 Mountain Avenue, Murray Hill NJ 07974
[**]IBM Thomas J. Watson Research Center, P.O. Box 704, Yorktown Heights, NY 10598

ABSTRACT

Invertible wavelet transforms that map integers to integers are important for lossless representations. In this paper, we present an approach to build integer to integer wavelet transforms based upon the idea of factoring wavelet transforms into lifting steps. This allows the construction of an integer version of every wavelet transform. We demonstrate the use of these transforms in lossless image compression.

1 INTRODUCTION

High-fidelity images generated from studio-quality video, medical images, seismic data, satellite images, and images scanned from manuscripts for preservation purposes typically demand lossless encoding. Yet, the huge datasize prohibits fast distribution of data. There is thus a need to seek encoding methods that can support storage and transmission of images at a spectrum of resolutions and encoding fidelities, from lossy to lossless, for progressive delivery and for different end-users' needs.

In recent years, wavelet transforms have been successfully used for lossy encoding of images. The multiresolution nature of the transform is also ideal for progressive transmission. However, common wavelet filters often have floating point coefficients. Thus, when the input data consist of sequences of integers (as is the case for images), the resulting filtered outputs no longer consist of integers. Yet, for lossless encoding, it would be of interest to be able to characterize the output completely again with integers. Using dyadic coefficients with rescaling yields integer coefficients, but this largely amplifies the dynamic range of the coefficients.

We denote by $(s_{0,j})_j$ the original signal of interest, $(s_{1,j})_j$ and $(d_{1,j})_j$ the lowpass and highpass coefficients respectively after a wavelet transform. In this paper, we present constructions of wavelet transforms that map a signal $(s_{0,j})_j$ represented in integers to $(s_{1,j})_j$ and $(d_{1,j})_j$, also represented in integers. The transform is reversible, i.e., we can exactly recover $(s_{0,j})_j$ from $(s_{1,j})_j$ and $(d_{1,j})_j$. We also show applications of the proposed transforms for lossless image compression.

The oldest integer to integer wavelet transform is the S transform [1], which is an integer version of the Haar transform:

$$\begin{aligned} d_{1,l} &= s_{0,2l+1} - s_{0,2l} \\ s_{1,l} &= s_{1,l} + \lfloor d_{1,l}/2 \rfloor . \end{aligned} \tag{1}$$

Said and Pearlman [2] further proposed the S+P transform (**S** transform + **P**rediction) in which linear prediction is performed on the lowpass coefficients $s_{1,l}$ to generate a new set of highpass coefficients after an S transform. The general form of the transform is:

$$\begin{aligned} d_{1,l}^{(1)} &= s_{0,2l+1} - s_{0,2l} \\ s_{1,l} &= s_{0,2l} + \lfloor d_{1,l}^{(1)}/2 \rfloor \\ d_{1,l} &= d_{1,l}^{(1)} + \lfloor \alpha_{-1}\,(s_{1,l-2} - s_{1,l-1}) + \alpha_0\,(s_{1,l-1} - s_{1,l}) \\ &\quad + \alpha_1\,(s_{1,l} - s_{1,l+1}) - \beta_1\, d_{1,l+1}^{(1)} \rfloor . \end{aligned} \tag{2}$$

The TS-transform proposed in [3] is a special case, when $\alpha_{-1} = \beta_1 = 0$ and $\alpha_0 = \alpha_1 = 1/4$. It is an integer version of the $(3,1)$ biorthogonal wavelet transform of Cohen-Daubechies-Feauveau [4]. Said and Pearlman examine several choices for (α_w, β_1) and in the case of natural images suggest $\alpha_{-1} = 0, \alpha_0 = 2/8, \alpha_1 = 3/8$ and $\beta_1 = -2/8$. It is interesting to note that, even though this was not their motivation, this choice without truncation yields a high pass analyzing filter with 2 vanishing moments (vanishing moments correspond to the multiplicity of zero as a root in the spectrum of the filter). The S, TS and S+P transforms can all be seen as special cases of the transforms we propose in this paper.

2 WAVELET TRANSFORMS THAT MAP INTEGERS TO INTEGERS

In this section, we describe constructions of wavelet transforms that map integers to integers. The readers are referred to [5] for more details. The construction is based on writing a wavelet transform in terms of lifting [6], which is a flexible technique that has been applied to the construction of wavelets through an iterative process of updating a subband from an appropriate linear combination of the other subband.

Table 1 lists example wavelet transforms implemented as invertible integer wavelet transforms. The first set of transforms have names of the form $(N, \tilde{N})$, where N is the number of vanishing moments of the analyzing high pass filter, while $\tilde{N}$ is the number of vanishing moments of the synthesizing high pass filter. They are instances of a family of symmetric, biorthogonal wavelet transforms built from the interpolating Deslauriers-Dubuc scaling functions [6]. The next transform $(2+2, 2)$ is inspired by the S+P transform — we use one extra lifting step to build the earlier $(2,2)$ into a transform with 4 vanishing moments of the high pass analyzing filter. The idea is to first compute a

Name	Wavelet Transform	Remarks
(2,2)	$d_{1,l} = s_{0,2l+1} - \lfloor 1/2\,(s_{0,2l} + s_{0,2l+2}) + 1/2 \rfloor$ $s_{1,l} = s_{0,2l} + \lfloor 1/4\,(d_{1,l-1} + d_{1,l}) + 1/2 \rfloor$	symmetric biorthogonal interpolating
(4,2)	$d_{1,l} = s_{0,2l+1} - \lfloor 9/16\,(s_{0,2l} + s_{0,2l+2}) - 1/16\,(s_{0,2l-2} + s_{0,2l+4}) + 1/2 \rfloor$ $s_{1,l} = s_{0,2l} + \lfloor 1/4\,(d_{1,l-1} + d_{1,l}) + 1/2 \rfloor$	
(6,2)	$d_{1,l} = s_{0,2l+1} - \lfloor 75/128\,(s_{0,2l} + s_{0,2l+2}) - 25/256\,(s_{0,2l-2} + s_{0,2l+4})$ $+3/256\,(s_{0,2l-4} + s_{0,2l+6}) + 1/2 \rfloor$ $s_{1,l} = s_{0,2l} + \lfloor 1/4\,(d_{1,l-1} + d_{1,l}) + 1/2 \rfloor$	$K = 1$
(4,4)	$d_{1,l} = s_{0,2l+1} - \lfloor 9/16\,(s_{0,2l} + s_{0,2l+2}) - 1/16\,(s_{0,2l-2} + s_{0,2l+4}) + 1/2 \rfloor$ $s_{1,l} = s_{0,2l} + \lfloor 9/32\,(d_{1,l-1} + d_{1,l}) - 1/32\,(d_{1,l-2} + d_{1,l+1}) + 1/2 \rfloor$	
(2+2,2)	$d^{(1)}_{1,l} = s_{0,2l+1} - \lfloor 1/2(s_{0,2l} + s_{0,2l+2}) + 1/2 \rfloor$ $s_{1,l} = s_{0,2l} + \lfloor 1/4(d^{(1)}_{1,l-1} + d^{(1)}_{1,l}) + 1/2 \rfloor$ $d_{1,l} = d^{(1)}_{1,l} - \lfloor \alpha\,(-1/2\,s_{1,l-1} + s_{1,l} - 1/2\,s_{1,l+1})$ $+\beta\,(-1/2\,s_{1,l} + s_{1,l+1} - 1/2\,s_{1,l+2}) + \gamma\,d^{(1)}_{1,l+1} + 1/2 \rfloor$	$\alpha = \beta = 1/8, \gamma = 0$ 4 vanishing moments of high pass filter K = 1
D4	$d^{(1)}_{l} = s_{0,2l+1} - \lfloor \sqrt{3}\,s_{0,2l} + 1/2 \rfloor$ $s_{1,l} = s_{0,2l} + \lfloor \sqrt{3}/4\,d^{(1)}_{1,l} + (\sqrt{3}-2)/4\,d^{(1)}_{1,l-1} + 1/2 \rfloor$ $d_{1,l} = d^{(1)}_{1,l} + s_{1,l+1}$ $K = (\sqrt{3}+1)/\sqrt{2} \approx 1.577$	orthogonal
(9-7)	$d^{(1)}_{1,l} = s_{0,2l+1} + \lfloor \alpha\,(s_{1,2l} + s_{1,2l+2}) + 1/2 \rfloor$ $s^{(1)}_{1,l} = s_{0,2l} + \lfloor \beta\,(d^{(1)}_{1,l} + d^{(1)}_{1,l-1}) + 1/2 \rfloor$ $d_{1,l} = d^{(1)}_{1,l} + \lfloor \gamma\,(s^{(1)}_{1,l} + s^{(1)}_{1,l+1}) + 1/2 \rfloor$ $s_{1,l} = s^{(1)}_{1,l} + \lfloor \delta\,(d_{1,l} + d_{1,l-1}) + 1/2 \rfloor$ $K \approx 1.1496$	symmetric biorthogonal $\alpha \approx -1.586$ $\beta \approx -0.053$ $\gamma \approx 0.883$ $\delta \approx 0.444$

Table 1: Wavelet transforms implemented as invertible integer wavelet transforms

$(2,2)$ yielding low pass samples $s_{1,l}$ and preliminary detail or high pass samples $d^{(1)}_{1,l}$, and then use the $s_{1,l}$ combined with $d_{1,l+n}$ $(n > 0)$ to compute $d'_{1,l}$ as a prediction for $d^{(1)}_{1,l}$. The final detail sample then is $d^{(1)}_{1,l} - d'_{1,l}$. Without truncation, we want the scheme to have 4 vanishing moments. This leads to 2 linear equations in 3 unknowns. Special cases are: (1) $\alpha = 1/6, \beta = 0, \gamma = 1/3$, (2) $\alpha = 1/8, \beta = 1/8, \gamma = 0$, and (3) $\alpha = 1/4, \beta = -1/4, \gamma = 1$. In our experiments we found that (2) works considerably better than (1) and (3), and this is the case we use when we refer to $(2+2,2)$ in Section 3. The next two transforms D4 and (9-7) [7] are built using the lifting factorization in [8]. To recover the usual normalization, a final scaling step would be needed (multiplying the lowpass coefficients by 1/K and highpass coefficients by K) for the D4 and (9-7), that can be implemented as three extra lifting steps [8]. The factorization in [8] converts every finite filter wavelet transform into lifting steps and thus allows for an integer to integer version.

3 APPLICATION TO LOSSLESS IMAGE COMPRESSION

We apply the various wavelets described in the previous section for lossless compression of digital images. For the evaluation, we use selected images from the set of standard ISO test images [9]. In this set of test images, there are natural images, computer-generated images, compound images (mixed texts and natural images, e.g., "cmpnd1" and "cmpnd2"; "us" is an ultrasound image with text on it) and different types of medical images ("x_ray", "cr", "ct", and "mri"). Separable two dimensional wavelet transforms are taken of the images.

The effectiveness for lossless compression is measured using the first order entropy. We further take into account the fact that the statistics in different quadrants of a wavelet-transformed image are different, and compute the weighted mean of the entropies in each quadrant of the transformed image. In the evaluation, we decompose each image into a maximum of five scales, the length and width permitting. The resulting bit rates are tabulated in Table 2.

We note that there is no filter that consistently performs better than all other filters on the test images. Wavelet filters with more analyzing vanishing moments generally perform well with natural and smooth images and not so with images with a lot of edges and high frequency components, such as with compound images. On the other hand, a low order filter like S-transform generally performs the worst, especially with natural images. It does a poor job in decorrelating the images. However, it performs significantly better on compound images (e.g., "cmpnd1" and "cmpnd2") than other higher order filters. Filters (4,2) and (2+2,2) have similar performances and generally perform better than other filters evaluated. For images "finger" and "ct", they perform significantly better than other filters. It is interesting to note that even though the S+P has 2 analyzing vanishing moments, it performs better than the TS which has 3 and has comparable performances with those with 4. This suggests that there are other factors besides the number of analyzing vanishing moments which affect compression efficiency. Furthermore, the $(9-7)$, which is most popularly used for lossy compression of images and which has 4 analyzing vanishing moments, generally does not perform as well as the (4,2) and (2+2,2), which have the same number of analyzing vanishing moments.

We further evaluate the wavelet filters by attaching an entropy coder to compute the *actual* bit rate. We use the

Images	S	TS	(2,2)	S+P	(4,2)	(2+2,2)	$D4$	9-7
air2	5.13	4.91	4.82	4.81	4.77	**4.76**	5.09	4.98
bike	4.36	4.28	4.19	**4.18**	4.20	4.21	4.37	4.31
cafe	5.69	5.54	**5.41**	5.42	**5.41**	**5.41**	5.63	5.51
cats	3.69	3.53	3.47	3.43	**3.42**	**3.42**	3.60	3.47
cmpnd1	**2.25**	2.84	2.79	2.97	3.31	3.36	3.04	3.45
cmpnd2	**2.41**	2.96	2.79	3.01	3.28	3.33	3.12	3.42
cr	5.40	5.25	**5.20**	5.24	5.22	5.22	5.28	5.22
ct	5.54	4.63	4.50	4.30	**4.15**	4.16	4.96	4.36
faxballs	1.61	1.31	**1.08**	1.41	1.36	1.17	1.54	1.97
finger	6.24	5.71	5.49	5.48	**5.35**	**5.35**	5.85	5.45
gold	4.27	4.10	4.05	4.08	4.04	**4.03**	4.19	4.14
graphic	3.18	2.82	2.60	2.67	**2.56**	**2.56**	3.08	3.00
hotel	4.30	4.18	**4.03**	4.10	4.06	4.04	4.25	4.18
mri	6.59	6.16	6.02	**5.90**	5.91	5.91	6.26	5.97
tools	5.84	5.80	**5.69**	5.73	5.73	5.72	5.88	5.81
us	**3.64**	3.79	3.69	3.79	3.87	3.85	3.95	4.26
water	2.46	2.45	**2.42**	2.47	2.45	2.44	2.46	2.50
woman	4.87	4.67	4.57	4.54	**4.53**	4.54	4.78	4.64
x_ray	6.42	6.13	**6.06**	6.09	**6.06**	**6.06**	6.18	6.08

Table 2: Bit rates of transformed images in entropies.

Images	TS	(2,2)	S+P	(4,2)	(2+2,2)	$D4$	9-7
air2	4.34	4.29	4.28	**4.24**	**4.24**	4.60	4.52
bike	3.83	3.78	**3.76**	3.78	3.79	3.94	3.88
cafe	5.07	4.98	4.98	**4.95**	4.98	5.17	5.07
cats	2.60	2.56	2.53	**2.52**	**2.52**	2.67	2.57
cmpnd1	2.34	**2.31**	2.38	2.59	2.62	2.40	2.64
cmpnd2	2.46	**2.33**	2.45	2.58	2.61	2.49	2.64
cr	5.21	**5.17**	5.22	5.19	5.19	5.25	5.20
ct	3.96	3.92	3.70	**3.61**	3.65	4.30	3.80
faxballs	1.12	**1.02**	1.20	1.17	1.07	1.51	1.59
finger	5.73	5.56	5.51	**5.41**	5.42	5.90	5.50
gold	3.99	3.94	3.99	3.95	**3.94**	4.11	4.09
graphic	3.05	2.91	2.93	**2.87**	2.89	3.34	3.29
hotel	3.97	**3.90**	3.95	3.91	**3.90**	4.10	4.07
mri	5.99	5.90	**5.77**	5.80	5.81	6.13	5.89
tools	5.59	**5.49**	5.53	5.52	5.51	5.65	5.60
water	1.83	**1.79**	1.85	1.81	1.80	1.85	1.88
woman	4.24	4.17	4.15	**4.14**	4.15	4.37	4.28
us	3.17	**3.13**	3.16	3.23	3.22	3.38	3.53
x_ray	5.99	**5.92**	5.94	**5.92**	**5.92**	6.05	5.95

Table 3: Exact bit rate of test images.

entropy coder of Said and Pearlman [2]. In each quadrant, the wavelet coefficients are visited in a scanline order; for each coefficient, a context C_j is derived based on the values of its adjacent coefficients which have already been visited and its parent coefficient. A total of 25 different contexts are used. The statistics of the context model is adaptively updated as more pixels are visited and the model is reset after each quadrant is visited. An arithmetic coder is then used to code the coefficients according to their contexts. The results are tabulated in Table 3. The numbers are derived from compressed file sizes. We found that the relative strength of each filter over the others in actual coding is similar to that computed using the entropies. In actual coding, wavelet filters (4,2) and (2+2,2) generally outperform other filters for natural and smooth images. Also, filters with more analyzing vanishing moments perform poorly with compound images. We also found that numbers computed from entropy computations provide a good indication of the actual performance of a wavelet filter. We also observe that the use of simple context modeling of the wavelet coefficients as in [2] cannot provide a significant enough gain in actual coding when the performance of the wavelet filter is poor in the first place. This points to the importance of using "good" wavelet filters and necessitates the search for such filters.

The classical approach to lossless compression is decomposed into two steps: spatial decorrelation and entropy coding of the decorrelated signals. The decorrelating steps have always been performed in the spatial domains and they involve some form of non adaptive or adaptive prediction of a pixel values based on previously visited pixels (see [10] and the references therein for comparisons of state-of-the-art spatial-domain predication techniques). In fact, we can view wavelet transforms that use two lifting steps such as those built from Deslauriers-Dubuc scaling functions as decorrelation steps. The high pass coefficients are the decorrelated components. The lifting step computes the difference between a true coefficient and its prediction and has the following form:

$$\begin{aligned} d_{j+1,l} &= s_{j,2l+1} - \\ &\quad \lfloor f(s_{j,2l-2L}, \cdots, s_{j,2l-2}, s_{j,2l}, s_{j,2l+2}, \cdots, s_{j,2l+2L}) \rfloor. \end{aligned} \tag{3}$$

We are thus predicting a pixel at an odd location $2l+1$ using pixels at even locations from both sides of $2l+1$. This is in contrast to the spatial domain approach to prediction, which has the following one-sided form:

$$d_j = s_j - \lfloor g(s_{j-J}, \cdots, s_{j-3}, s_{j-2}, s_{j-1}) \rfloor. \tag{4}$$

The price to pay for prediction based on (3) is to retain knowledge of the remaining pixels at the even locations. Similar prediction steps can be performed on these remaining pixels, but because the distances between adjacent pixels are larger than before, correlations between adjacent pixels tend to be lower. The dual lifting step in generating the $s_{j+1,l}$ smooths this set of remaining pixels before the lifting step on the coarser level is performed.

The big advantage of using wavelet transform to represent images is multiresolution representation, which lossless compression methods based on spatial-domain prediction (see [10] and the references therein for comparisons of state-of-the-art spatial-domain predication techniques) cannot offer. Using wavelet transforms that map integers to integers permits lossless representation of the image pixels and easily allows the transmission of lower resolution versions first, followed by transmissions of successive details. Such a mode of transmission is especially valuable in scenarios where bandwidth is limited, image sizes are large and lossy compression is not desirable. Examples are transmission of 2-D and 3-D medical images for telemedicine applications and transmission of satellite images down to earth.

Valuable lessons could also be learned from the recent advances in spatial-domain prediction techniques [10]. Adaptation of prediction methods is made based on local statistics of the pixels. In the wavelet transform approach to decorrelating an image, we could also use an adaptive scheme in deciding the use of wavelet filters on a pixel by pixel basis. As we have seen, there is not a wavelet filter that performs uniformly better than the others. Thus, the activities in some small neighborhood should be used in determining the type of filters to use. The use of adaptive wavelet filters in lossless and multiresolution representation of images warrants further investigation.

REFERENCES

[1] V. K. Heer and H-E Reinfelder, "A comparison of reversible methods for data compresssion", *in Medical Imaging IV*, pp. 354–365. Proc. SPIE 1233, 1990.

[2] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy image compression", *IEEE Trans. Image Process.*, vol. 5, pp. 1303–1310, September 1996.

[3] A. Zandi, M. Boliek, E. L. Schwartz, and M. J. Gormish, "CREW lossless/lossy medical image compression", Technical Report CRC-TR-9526, 1995.

[4] A. Cohen, I. Daubechies, and J. Feauveau, "Biorthogonal bases of compactly supported wavelets", *Comm. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.

[5] A. R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers", Technical report, Department of Mathematics, Princeton University, 1996.

[6] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets", *Journal of Appl. and Comput. Harmonic Analysis*, vol. 3, pp. 186–200, 1996.

[7] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Series in Appl. Math., Vol. 61. SIAM Journ. of Math. Anal., Philadelphia, PA, 1992.

[8] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps", Technical report, Bell Laboratories, Lucent Technologies, 1996.

[9] ISO/IEC JTC 1/SC29/WG1, "Call for contributions - lossless compression of continuous-tone still pictures", ISO Working Document ISO/IEC JTC1/SC29/WG1 N41, Mar. 1994.

[10] N. Memon, V. Sippy, and X. Wu, "A comparison of prediction schemes proposed for a new lossless image compression standard", *in IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 309–312, May 1995.