Performance Evaluation of Reversible Integer-to-Integer Wavelet Transforms for Image Compression

Michael D. Adams, *Student Member, IEEE*, and Faouzi Kossentini, *Senior Member, IEEE*
Department of Electrical and Computer Engineering
University of British Columbia
Vancouver, B.C., Canada V6T 1Z4

Please address all correspondence concerning this paper to:
Michael D. Adams
Department of Electrical and Computer Engineering
University of British Columbia
2356 Main Mall
Vancouver, B.C., Canada V6T 1Z4
Phone: (604)822-2872, Fax: (604)822-5949
Email: mdadams@ece.ubc.ca

Performance Evaluation of Reversible Integer-to-Integer Wavelet Transforms for Image Compression*

Michael D. Adams and Faouzi Kossentini
Department of Electrical and Computer Engineering
University of British Columbia
Vancouver, B.C., Canada V6T 1Z4

Abstract

A number of reversible integer-to-integer wavelet transforms are compared on the basis of their lossy compression performance, lossless compression performance, and computational complexity. Of the transforms considered, several were found to perform particularly well, with the best choice for a given application depending on the relative importance of lossless compression performance, lossy compression performance, and computational complexity. Reversible integer-to-integer versions of numerous transforms are also compared to their conventional (i.e., nonreversible real-valued) counterparts for lossy compression. In many cases, the reversible integer-to-integer and conventional versions of a transform were found to yield results with comparable subjective image quality.

1 Introduction

There has been a growing interest in reversible integer-to-integer wavelet transforms for image coding applications. Such transforms are invertible in finite-precision arithmetic (i.e., reversible), map integers to integers, and approximate the linear wavelet transforms from which they are derived. Due largely to these properties, transforms of this type are extremely useful for coding systems requiring efficient handling of lossless coding, minimal memory usage, or low computational complexity. Furthermore, these transforms are particularly attractive for supporting functionalities like progressive lossy-to-lossless recovery of images, lossy compression with the lossless reproduction of a region of interest, and strictly lossy compression with minimal memory usage. Due to applications such as these, we can see that there is a clear need to consider not only the lossless compression performance of a particular reversible transform, but also its lossy compression performance.

In this paper, several reversible integer-to-integer wavelet transforms are compared on the basis of their lossy compression performance, lossless compression performance, and computational complexity. By using the results of such an evaluation, not only are we able to build systems with improved compression efficiency, but we can also better understand the tradeoffs

*This work was supported by both the Natural Sciences and Engineering Research Council of Canada, and Image Power Inc.

between compression efficiency and computational complexity. Reversible integer-to-integer versions of transforms are also compared to their conventional (i.e., nonreversible real-valued) counterparts for lossy compression. The objective, in this case, is to quantify any performance degradation associated with the introduction of the reversible and integer-to-integer properties.

2 Transforms

The reversible integer-to-integer wavelet transforms considered in this study were constructed using the method described in [1] and are based on the lifting framework [2] or a generalization thereof [3] [4]. In all, twelve transforms were evaluated as listed in Table 1. The transforms are all strictly 1D in nature, and based on 2-channel filter banks. Moreover, with the exception of the SPB and SPC transforms, all of the transforms are based on FIR filter banks. For convenience, the notation x/y is used to indicate that a transform is associated with a filter bank having lowpass and highpass analysis filters of lengths x and y, respectively. Similarly, the notation (x, y) is used to indicate that the analyzing and synthesizing wavelets associated with a transform have x and y vanishing moments, respectively.

Table 1: Transforms

Name	Description / References
5/3	5/3, (2, 2), interpolating, [5]
TS	Two-Six Transform $\leftrightarrow$ 2/6, (3, 1), [6]
SPB	S+P Transform with Predictor B $\leftrightarrow$ IIR synthesis filters, [7]
M-9/7	9/7, (4, 2), interpolating, [8]
TT	Two-Ten Transform $\leftrightarrow$ 2/10, (5, 1), [9]
C-5/11	5/11, (4, 2), [1]
A-5/11	5/11, (2, 2), [10]
6/14	6/14, (3, 3), [11]
SPC	S+P Transform with Predictor C $\leftrightarrow$ IIR synthesis filters, [7]
T-13/7	13/7, (4, 4), interpolating, [8]
C-13/7	13/7, (4, 2), interpolating, [11]
FBI	FBI Fingerprint Transform $\leftrightarrow$ 9/7, (4, 4), [12]

Since these transforms are 1D in nature, images are handled by transforming the rows and columns in succession. Unlike in the case of conventional (linear) versions of transforms, however, the order in which rows and columns are transformed is important. That is, the inverse transform must operate on rows and columns in the reverse order from that used in the forward transform; otherwise, invertibility cannot be guaranteed.

3 Evaluation Methodology

For the most part, the work described in this paper was conducted in the context of the JPEG-2000 standardization effort. For evaluation purposes, the software implementation of the JPEG-2000 baseline coder developed jointly by SAIC and the University of Arizona [13] was employed. The test data used in our study consisted of the images listed in Table 2, all of which were taken from the JPEG-2000 test set [14]. Both lossy and lossless compression were considered in our analysis. In the lossless case, compression performance was measured in terms of the final bit rate. For lossy compression, subjective image quality was used as a performance metric.

Table 2: Test Images

Image	Size	Depth	Description
`aerial2`	2048×2048	8	aerial photograph
`bike`	2048×2560	8	collection of various objects
`cafe`	2048×2560	8	outside of cafe
`cmpnd2`	5120×6624	8	mixed text and graphics
`gold`	720×576	8	houses and countryside
`sar2`	800×800	12	synthetic aperture radar image
`target`	512×512	8	patterns and textures
`woman`	2048×2560	8	lady

Using the various transforms under consideration, each of the test images was compressed in a lossy manner at several bit rates (i.e., 0.0625, 0.125, 0.25, 0.5, 1, and 2 bpp), in addition to being compressed losslessly. Since the goal of the subjective evaluations was only to compare artifacts obtained from the transforms themselves and not tiling artifacts, the use of tiling was avoided whenever possible. More specifically, tiling was used only in the case of the `cmpnd2` image due to its large size, where a tile size of 2048×2048 was employed.

In order to perform the subjective testing, several regions from the various lossy reconstructions were selected to be scrutinized. Small regions of interest were used (instead of whole images) for two reasons:

1. Most of the test images were too large to be displayed at their original resolution on the computer screen, and displaying the images at reduced resolution hides most of the artifacts introduced by compression.

2. If the whole image is used, there might be too much information for a human observer to analyze, making an accurate and meaningful comparison difficult.

Since the human eye often cannot distinguish between a high bit rate lossy reconstruction of an image and the original, formal subjective testing was restricted to the lower bit rates under consideration. As most applications typically utilize compression ratios from 16:1 to 128:1, the bit rates corresponding to these compression ratios were of most interest. Two types of subjective evaluations were employed as described below.

The first type of subjective evaluation served to compare the image quality obtained with different reversible transforms. For a specific test image and bit rate, a particular region of the lossy reconstructions obtained with the various transforms was shown under magnification. The original image was also shown for reference purposes. Then, a human observer was asked to assign a simple integer ranking to the different results, with a value of one corresponding to best result and larger values indicating poorer results. In cases where a distinction in quality could not be made between two or more results, these results could be assigned the same rank. This process was repeated for various test images, bit rates, regions of interest within the image, and observers. The rankings were then used to determine the most effective transform from a subjective image quality standpoint.

The second type of subjective evaluation was used to compare the quality of the lossy reconstructions obtained with reversible integer-to-integer and conventional versions of the same transform. For a specific test image, bit rate, and transform, a particular region of interest in the lossy reconstructions obtained with the reversible and conventional versions of a transform were shown under magnification. Again, the original image was also shown for reference purposes. Then, the observer was asked to rank the two lossy reconstructions relative to one another. The ranking was done on an integer scale from -2 to 2. The values -2, -1, 0, 1, 2 correspond to the result obtained with the reversible version of the transform being significantly worse than, slightly worse than, the same as, slightly better than, and significantly better than the result obtained with the conventional version, respectively. This process was repeated for numerous transforms, images, bit rates, and observers.

4 Lossless Compression Performance

Each of the test images in Table 2 was compressed in a lossless manner using the various transforms under evaluation, yielding the results shown in Table 3. For each image, the best result has been highlighted. As one would expect, image content is an important factor influencing transform effectiveness. Clearly, for images with a relatively greater amount of high-frequency content (e.g., `cmpnd2` and `target`), the 5/3 transform yields the best results. For smooth images (e.g., `woman`, `aerial2`, etc.), the T-13/7 transform is most effective, followed closely by the A-5/11, C-5/11, M-9/7, and C-13/7 transforms.

Table 3: Lossless compression results

	Bit Rate (bpp)											
Image	5/3	TS	SPB	M-9/7	TT	C-5/11	A-5/11	6/14	SPC	T-13/7	C-13/7	FBI
`aerial2`	5.243	5.251	5.234	5.221	5.233	5.223	5.226	5.229	5.263	**5.217**	5.222	5.233
`bike`	4.415	4.436	4.416	4.406	4.425	4.407	4.401	4.429	4.453	**4.397**	4.402	4.456
`cafe`	5.194	5.215	5.191	5.177	5.201	5.178	**5.175**	5.217	5.228	**5.175**	5.186	5.220
`cmpnd2`	**2.540**	2.606	2.655	2.753	2.732	2.770	2.686	2.855	2.787	2.775	2.797	2.906
`gold`	4.476	4.501	4.496	4.476	4.496	4.471	**4.467**	4.490	4.537	4.473	4.477	4.518
`sar2`	7.610	7.625	7.626	7.615	7.627	7.616	**7.610**	7.627	7.662	7.612	7.616	7.611
`target`	**2.129**	2.234	2.237	2.256	2.314	2.247	2.213	2.355	2.249	2.250	2.243	2.546
`woman`	4.394	4.406	4.381	4.366	4.386	4.367	4.372	4.385	4.408	**4.362**	4.368	4.420
Mean†	5.222	5.239	5.224	5.210	5.228	5.210	5.209	5.230	5.259	**5.206**	5.212	5.243

†Mean over smooth images (i.e., excluding `cmpnd2` and `target`).

5 Subjective Lossy Compression Performance

Each of the images in Table 2 was compressed in a lossy manner at several bit rates using the various transforms under evaluation. Subjective testing was then undertaken as specified in Section 3. In order to make the testing more practical, the evaluation was split into two stages. The same testing strategy was used in each stage. Only the transforms under evaluation in each case differed.

In the first stage, the 5/3, TS, SPB, M-9/7, TT, SPC, T-13/7, and FBI transforms were compared. After obtaining the rankings from the individual subjective tests, these results were averaged together to yield the overall rankings given in Table 4. From these results, we see that the transforms fall into three distinct performance categories. The FBI, 5/3, T-13/7, and M-9/7 transforms perform best, followed by the TS and TT transforms, followed by the SPB and SPC transforms. These numerical results generally agree with comments made by participants during the subjective tests—namely, that there are three distinct performance categories into which the transforms fall. Some examples of the lossy reconstructions obtained in this stage of testing can be found in [3].

Table 4: Overall subjective rankings from first stage of testing

	5/3	TS	SPB	M-9/7	TT	SPC	T-13/7	FBI
Rank	2.51	4.47	6.05	2.97	4.57	6.88	2.88	2.36

The second stage of subjective testing included some of the best transforms from the first stage and several new transforms, resulting in the following set of transforms: 5/3, M-9/7, C-5/11, A-5/11, 6/14, T-13/7, and C-13/7. After conducting the subjective tests, the rankings were averaged together with the results shown in Table 5. In this round of tests, the general consensus amongst the test participants is that all of the transforms are very close in terms of subjective quality, with the possible exception of the 6/14 transform which performs slightly worse. In other words, although the average rankings may differ by a large amount in this stage of the testing, it is important to understand that even large differences correspond to very minute differences in subjective image quality.

Table 5: Overall subjective rankings from second stage of testing

	5/3	M-9/7	C-5/11	A-5/11	6/14	T-13/7	C-13/7
Rank	1.84	2.30	2.34	2.15	2.75	1.77	1.83

Considering the results from both stages of the subjective testing together, we conclude that the FBI, 5/3, T-13/7, C-13/7, C-5/11, and A-5/11 transforms have the best performance, followed by the 6/14, TS, and TT transforms, followed by the SPB and SPC transforms.

The SPB and SPC transforms performed worst due largely to their tendency to introduce ringing artifacts. This tendency can be attributed to the fact that both transforms are associated with filter banks having IIR synthesis filters. The 6/14 transform often introduced more ringing artifacts than some of the other transforms, leading to its relatively poor performance. In the case of the TS and TT transforms, blockiness often resulted. This can be attributed to the rougher synthesis scaling and wavelet functions associated with these transforms.

6 Reversible Integer-to-Integer Transforms Versus Conventional Transforms for Lossy Compression

As part of our work, subjective testing was undertaken to compare the quality of lossy reconstructions obtained with reversible and conventional versions of the same transform. For this purpose, a subset of the transforms listed in Table 1 was used—namely the 5/3, TS, SPB, M-9/7, TT, SPC, T-13/7, and FBI transforms. Using reversible and conventional versions of each transform, the test images listed in Table 2 were compressed in a lossy manner at several bit rates. The lossy reconstructions obtained with the two different versions of each transform were then compared. This testing was performed as described previously in Section 3.

After taking all of the individual subjective test results and averaging them together, the results in Table 6 were obtained. These numerical results indicate the relative performance of the reversible and conventional versions of the transforms. A negative value corresponds to better performance in the conventional transform case. From the experimental results, we can see that in most cases there is only a small bias favoring the conventional version of a transform. Thus, there is a very slight performance penalty associated with using reversible integer-to-integer transforms in place of their conventional counterparts. This said, however, it is important to point out that this difference in performance is very marginal. In most practical applications, this difference is not likely to be of concern.

Table 6: Difference in subjective performance between reversible integer-to-integer and conventional versions of transforms

	5/3	TS	SPB	M-9/7	TT	SPC	T-13/7	FBI
Rank	-0.20	-0.41	-0.80	-0.29	-0.36	-0.77	-0.37	-0.24

From the above subjective analysis and some other objective testing detailed in [3], we found that three factors affect the performance of reversible integer-to-integer transforms relative to their conventional counterparts:

1. Whether the underlying filter bank has IIR filters.
2. The depth of the image (in bpp) being compressed.
3. The bit rate used for compression.

In the calculation of reversible integer-to-integer transforms, intermediate results are quantized to integer values in numerous places. If the underlying filter bank has IIR filters, this very coarse quantization occurs inside feedback paths. Thus, quantization errors can accumulate indefinitely[1]. For this reason, there is a much bigger difference between reversible and

[1] Moreover, it may even be difficult to show that the resulting nonlinear system is stable.

conventional versions of a transform in the case where the transform is associated with a filter bank having IIR filters. In the case of the transforms studied, the SPB and SPC transforms are associated with such filter banks. From the subjective results above, it is evident that these two transforms suffer the greatest degradation in performance as a result of the reversible and integer-to-integer properties.

As the depth of the image increases, the effects of quantizing intermediate results to integer values has less effect. Although for 8 bpp images there is a clear difference in PSNR performance, there is little difference in the case of 12 bpp images (e.g., `sar2`) as long as the transforms are based on FIR filter banks.

As the bit rate is lowered, the difference in compression performance between reversible and conventional versions of transforms decreases. At a sufficiently low bit rate, the quantization of transform coefficients becomes so coarse that any errors introduced by quantizing intermediate results to integers are masked by quantization of the coefficients themselves.

7 Computational Complexity

All of the reversible transforms under consideration are calculated using only fixed-point arithmetic. In particular, only integer addition/subtraction, multiplication, and division operations are required. Since all filter coefficients are approximated by dyadic rational numbers, all division operations can be implemented as bit shifts. Assuming a one-level wavelet decomposition (in one dimension), the number of addition, multiplication, and shift operations required per two input samples for each transform is given in Table 7. Two sets of numbers are provided for each transform[2]. The first set shows the number of operations in the straightforward implementation of the transform. The second set of numbers in parenthesis indicates the number of operations required if all multiplications are converted to shift and add operations. This complexity metric is of particular interest for hardware implementations or software implementations on architectures where integer multiplications are more costly than addition and shift operations. In the case where both sets of numbers are the same, only one is given.

Notice that although some of the transforms have equal operation counts, some require more multiplications than others. In particular, note that the 5/3, TS, C-5/11 and A-5/11 transforms are truly multiplierless (i.e., their underlying filters all have coefficients that are powers of two).

Clearly, the 5/3 and TS transforms require the least computation, and the FBI transform requires the most. The remaining transforms lie somewhere in between, with the SPB, M-9/7, TT, C-5/11, and A-5/11 transforms having roughly the same complexity, and the 6/14, SPC, T-13/7, and C-13/7 transforms having approximately the same complexity.

8 Memory Usage

The amount of memory needed by a transform-based coder can be strongly influenced by the amount of memory required to store transform coefficients. In many image coding systems, transform coefficients are represented as real numbers. Typically, in such systems, 32 bits

[2]Symmetry in filter coefficients was considered in the calculation of these values.

Table 7: Computational complexity

Transform	Adds	Shifts	Multiplies	Total
5/3	5	2	0	7
TS	5	2	0	7
SPB	7 (8)	3 (4)	1 (0)	11 (12)
M-9/7	8 (10)	2 (3)	1 (0)	11 (13)
TT	7 (9)	2 (5)	2 (0)	11 (14)
C-5/11	10	3	0	13
A-5/11	10	3	0	13
6/14	10 (12)	3 (5)	1 (0)	14 (17)
SPC	8 (10)	4 (7)	2 (0)	14 (17)
T-13/7	10 (14)	2 (4)	2 (0)	14 (18)
C-13/7	10 (14)	2 (4)	2 (0)	14 (18)
FBI	12 (24)	4 (14)	4 (0)	20 (38)

of memory are needed to store each coefficient. Transforms which map integers to integers can improve significantly upon this situation. Although some intermediate results are still real-valued, such results are only needed transiently. All stored values, including the final transform coefficients, are integers. In fact, experimental results suggest that for many integer-to-integer wavelet transforms, 16-bit integers provide sufficient dynamic range for representing the transform coefficients (at least, for 8 bpp images). By using 16 bits of memory per coefficient instead of 32, memory requirements can be reduced by half.

One might even be able to use a smaller number of bits to represent transform coefficients in some cases. Using a smaller number of bits, however, does increase the risk of numerical overflow. Fortunately, the invertibility of lifting-based transforms can be made to hold even in the presence of overflow. Provided that occurrences of overflow are infrequent, lossless compression performance is not seriously affected. In the case of lossy compression, however, one needs to be much more cautious about overflow. Overflow has the potential to cause very disturbing artifacts in reconstructed images, especially if the overflow affects coefficients associated with basis functions having large support.

9 Conclusions

In the case of lossless compression, the smoothness of an image is an important factor influencing transform effectiveness. For smooth images, the T-13/7 transform is most effective, followed by the M-9/7, A-5/11, C-5/11, and C-13/7 transforms as a group. For images with a greater amount of high frequency content, the 5/3 transform performs best.

For lossy compression, several transforms were found to yield the best subjective image quality. More specifically, the 5/3, M-9/7, C-5/11, A-5/11, T-13/7, C-13/7, FBI transforms were found to perform best.

Considering both the lossy and lossless compression results together, we conclude that the

5/3 and T-13/7 transforms are the most attractive without regard to computational complexity. For applications requiring low computational complexity, however, the 5/3 transform is the clear winner overall. It performs well for both lossless and lossy compression and also has the lowest computational complexity of the transforms under consideration. If a low complexity transform is specifically needed for smooth images, the M-9/7, C-5/11, and A-5/11 transforms are likely the best choices due to their good lossless and lossy performance. Unfortunately, the FBI transform, although very good for lossy compression, is significantly higher in complexity than all of the other transforms, and performs poorly for lossless compression.

10 Acknowledgments

The authors would like to thank Igor Kharitonenko at the Motorola Australian Research Centre (Botany, NSW, Australia) for providing some additional subjective testing results to supplement our own. Also, the authors would like to express their gratitude to all of the participants in the subjective testing for their time and effort.

References

[1] A. R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers," Tech. Rep., Department of Mathematics, Princeton University, Aug. 1996, Available from `http://cm.bell-labs.com/who/wim/papers/integer.ps`.

[2] W. Sweldens, "The lifting scheme: A new philosophy in biorthogonal wavelet constructions," in *Proceedings of SPIE: Wavelet Applications in Signal and Image Processing III*, 1995, vol. 2569, pp. 68–79.

[3] M. D. Adams and F. Kossentini, "Performance evaluation of several reversible integer-to-integer wavelet transforms in the JPEG-2000 baseline system (VM release 0.0)," ISO/IEC JTC 1/SC 29/WG 1 N 866, June 1998, Available from `http://www.ece.ubc.ca/~mdadams`.

[4] M. D. Adams, "Reversible wavelet transforms and their application to embedded image compression," M.A.Sc. thesis, Department of Electrical and Computer Engineering, University of Victoria, Victoria, Canada, Jan. 1998, Available from `http://www.ece.ubc.ca/~mdadams`.

[5] D. Le Gall and A. Tabatabai, "Sub-band coding of digital images using symmetric short kernel filters and arithmetic coding techniques," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1988, pp. 761–764.

[6] J. D. Villasenor, B. Belzer, and J. Liao, "Wavelet filter evaluation for image compression," *IEEE Trans. on Image Processing*, vol. 4, no. 8, pp. 1053–1060, Aug. 1995.

[7] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy compression," *IEEE Trans. on Image Processing*, vol. 5, no. 9, pp. 1303–1310, Sept. 1996.

[8] G. Strang and T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1996.

[9] M. J. Gormish, E. L. Schwartz, A.F. Keith, M. P. Boliek, , and A. Zandi, "Lossless and nearly lossless compression of high-quality images," in *Proc. of SPIE*, Mar. 1997, vol. 3025, pp. 62–70.

[10] M. D. Adams and F. Kossentini, "Low complexity reversible integer-to-integer wavelet transforms for image coding," in preparation, Nov. 1998.

[11] B. Berthelot, E. Majani, and P. Onno, "Transform core experiment: Compare the SNR performance of various reversible integer wavelet transforms," ISO/IEC JTC 1/SC 29/WG 1 N 902, June 1998.

[12] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Processing*, vol. 1, no. 2, pp. 205–220, Apr. 1992.

[13] SAIC and University of Arizona, "JPEG-2000 VM0 software," ISO/IEC JTC 1/SC 29/WG 1 N 840, May 1998.

[14] "JPEG-2000 test images," ISO/IEC JTC 1/SC 29/WG 1 N 545, July 1997.