Wavelet-based Prediction Measures for Lossy Image Set Compression

Marc Moreau[1], **Howard Cheng**[1]
[1] Department of Mathematics and Computer Science, University of Lethbridge, Canada

Abstract—*A number of image set compression algorithms have been proposed in the literature. A key component of these algorithms is a numeric measure used to quantify how similar two images are to each other from the point of view of a compression algorithm. Since most of these image set compression algorithms use wavelet-based image compression algorithms to compress the prediction error, we propose a number of related image prediction measures based on wavelet transforms. We also show some experimental results using the proposed measures. The proposed measure performs better than previous measures proposed in the literature.*

Keywords: image set compression, wavelet transform, prediction measures

1. Introduction

Traditional image compression algorithms for individual images, such as predictive coding and transform coding, have been shown to effectively reduce coding, inter-pixel, and psycho visual redundancy within an image [8]. Image sets, however, may contain inter-image redundancy, or "set redundancy" [10], which are not reduced by these algorithms. A number of algorithms have been proposed to address this issue [1], [4], [5], [10], [11], [14], [15], [16], [17].

A key component of many of these algorithms is the use of prediction of one image from another image in the set. The choice of which image to use for predicting a particular image is often based on some prediction measure—if two images are similar enough, then the prediction error should be small and the error image will be easy to compress by a conventional image compression algorithm. This idea was used by a number of image set compression schemes based on minimum spanning trees, where the prediction measure is used as the edge weights of a graph. Assuming that the chosen edge weight accurately reflects how easy it is for an image compression algorithm to compress the difference image between two images, these image set compression schemes obtain an optimal compression scheme for a particular set of images. Furthermore, the prediction measure is also a key component in hierarchical image set compression algorithms [16], in which the prediction measure is used to cluster subsets of similar images together for more efficient compression.

In almost all of these image set compression schemes, one image is used to predict another image with simple pixel-wise subtraction. The prediction error is simply the difference image between the two images. This difference image is then compressed using a conventional image compression algorithm. In [6], a number of prediction measures in the spatial domain are studied to see how well they correlate to the actual performance of the compression algorithm on the difference images. However, the actual compression algorithms (such as JPEG 2000 [2]) typically examine an image in the transform domain. As a result, spatial domain prediction measures are not necessarily good matches to the performance of compression algorithms.

In this paper, we examine prediction measures that are computed from wavelet coefficients after a pyramid decomposition has been performed on the difference image. The use of these prediction measures should match the performance of wavelet-based image compression algorithms such as JPEG 2000, leading to better overall performance for the image set compression scheme. Furthermore, the measures we examine are more efficient than actually performing the compression itself. Experimental results will be presented.

2. Previous Works

The minimum spanning tree (MST) set mapping strategy, proposed by Nielsen and Li, is based on a graph data structure [14]. A complete graph is constructed, using images as the vertices and the root-mean-square error (RMSE) between adjacent images as the edge weights. An MST for the graph is calculated, and one image is chosen as the root. The root image and difference images represented by the edges with the lowest total cost are encoded using lossy JPEG 2000 [2]. The results of these experiments showed a clear improvement in average distortion (RMSE) between original and reconstructed images when using the MST scheme over the "traditional" scheme of compressing each image individually, especially at lower bitrates. These experiments focused on sets of similar images, and did not examine performance on sets containing dissimilar images. Other authors have applied a similar strategy to specific applications such as object movies [1], multiview video coding [13], multispectral images [19], and map images [12]. However, the image sets in these applications are known to be similar and the image sets are usually small. Further improvement and variations of this idea include the MST_a scheme [4], [5] in which the average image is included in the complete graph, and the hierarchical MST_a (HMST_a) scheme in which a clustering algorithm is applied to first group very similar images together before the MST_a scheme is applied to each cluster. Tashakkori [18] found that correlation between wavelet coefficients is higher than the correlation between original pixel values for similar images.

Most of the set image compression scheme uses RMSE to measure the similarity between two images. However, the measure does not always accurately reflect compression performance. In [6], a number of spatial domain prediction measures were studied including the RMSE, entropy, standard deviation, and an adjusted L_1 norm. It was shown that both the adjusted L_1 norm and RMSE measures predict the actual compression performance somewhat accurately for high compression ratios in many cases.

3. Approach

Our goal is to obtain a numeric measure computed from a difference image, such that the measure reflects the compression performance of the chosen compression algorithm on the difference image. For lossless compression, one can simply compress the difference image and measure the size of the compressed image. For lossy compression, we may measure the distortion (e.g. RMSE) of the reconstructed image at a particular compression ratio. Although these statistics can be obtained from applying the compression algorithm on the difference image, this can be computationally costly. This is especially important for image set compression schemes based on minimum spanning trees, as there are generally $O(n^2)$ pairs of difference images to consider for a set of n images. Instead, we aim for a measure which is computationally simpler than full compression but at the same time accurately predicts the compression performance.

Our general approach is to apply a wavelet transform to obtain a pyramid decomposition of the difference image [8], and then examine the resulting wavelet coefficients. To simplify the discussion, please refer to Figure 1 for how the different subbands are labelled. Note that the pyramid levels in different orientations at the same spatial resolution are assigned the same label.

3.1 Choice of Wavelet

In this study, we examine two different wavelets for the pyramid decomposition. The first wavelet we choose is the simple Haar wavelet [8]. In this case, the lowpass filter coefficients are $1/\sqrt{2}, 1/\sqrt{2}$ and the highpass filter coefficients are $1/\sqrt{2}, -1/\sqrt{2}$. This wavelet is chosen because of its computational simplicity. For many of the measures, the scale of the coefficients is in fact not important. In this case, we can simplify the filter coefficients further to $1, 1$ and $1, -1$. As a result, the pyramid decomposition can be performed with simple addition and subtraction.

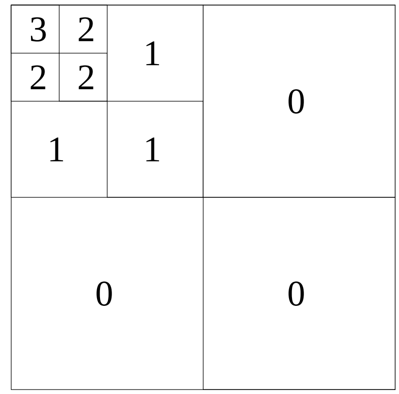


Fig. 1: Labels for subbands and levels in a 3-level pyramid decomposition.

The second wavelet we choose is the CDF 9/7 wavelet. The lowpass filter coefficients are 0.026749, -0.016864, -0.078223, 0.266864, 0.602949, 0.266864, -0.078223, -0.016864, 0.026749, and the highpass filter coefficients are 0.091272, -0.057544, -0.591272, 1.115087, -0.591272, -0.057544, 0.091272. While the CDF 9/7 wavelet transform is more computational intensive to perform, it is the same transform used in JPEG 2000 [2] and should match the performance of JPEG 2000 better.

The number of levels in the pyramid decomposition depends on the subset of wavelet coefficients required to compute the measure. For example, if only the coefficients of level 0 are needed, then only one level of pyramid decomposition is needed.

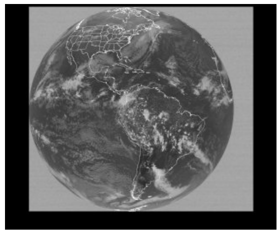
3.2 Wavelet-based Measures

3.2.1 Entropy

Since wavelet-based image compression algorithms perform compression on the wavelet coefficients, the entropy of the wavelet coefficients may be a good prediction measure for the compression performance. In this study, we examine first-order entropy, second-order entropy (1×2 blocks), and fourth-order entropy (2×2 blocks). Coefficients from a subset of levels in the decomposition are used in the computations. The entropy is computed by taking blocks of coefficients (of the appropriate size) and treating each block as a symbol when computing the histogram.

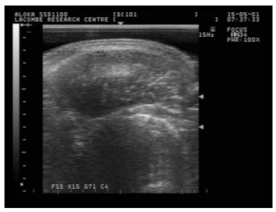
We should note that each coefficient in a level represents a square region in the image in the spatial domain. For example, a wavelet coefficient in level 1 is computed from a 16×16 block in the spatial domain, so the fourth-order entropy computed in this level in fact represents a 64×64 block in the spatial domain. Blocks from different levels in the pyramid decomposition represent blocks of different sizes in the spatial domain. As a result, we will only compute entropy of wavelet coefficients from one level only.

(a) Galway



(b) GOES

(c) Joe



(d) Pig

Fig. 2: Sample images from each set.

3.2.2 RMSE

The RMSE of the wavelet coefficients can be computed to obtain a numeric measure. Although the correlation between wavelet coefficients have been found to be higher than the correlation between original pixel values in similar images [18], we note that the wavelet transform is a linear transform on the images, so that the effectiveness of any linear prediction scheme of images (e.g. simple image subtraction) is the same whether it is done in the wavelet coefficient domain or the spatial domain. Furthermore, if the chosen wavelet is orthonormal (e.g. the Haar wavelet), the RMSE in the spatial domain and transform domain are in fact identical. This is a result of the fact that orthonormal wavelet transforms are distance-preserving.

However, if only coefficients from certain levels are used, the result in linear prediction may be different. Therefore, we will examine the RMSE measure only on subsets of levels of wavelet coefficients.

3.2.3 Other Measures

One may apply other spatial domain prediction measures studied in [6] on the wavelet coefficients as well (e.g. standard deviation, adjusted L_1 norm). Since the performance of these measures were similar to that of RMSE, they are excluded from this study.

4. Experimental Results

We use the same five test image sets used in previous image set compression studies (e.g. [4], [5]): Galway [3], GOES [7], Joe [9], Pig, and Combination. Sample images appear in Figure 2. The Galway set contains 28 webcam images of a street in Galway City, Ireland that were taken within a few hours under similar weather and lighting conditions. The GOES set contains 128 satellite images of the Earth. The Joe set contains 162 webcam images of an outdoor scene that were taken under varying weather and lighting conditions. The Pig set contains 300 ultrasound images of pig rib cages.

For each image set, we apply each of Haar wavelet and CDF 9/7 wavelet transform to obtain a pyramid decomposition. The RMSE, as well as first-order, second-order, and fourth-order entropy of the wavelet coefficients of selected levels are computed. These measures are used as edge weights in the computation of a minimum spanning tree in the MST_a image set compression scheme [4], [5], using JPEG 2000 to compress the difference images. The image set is compressed at a fixed compression ratio, and the average distortion between the original images and the reconstructed images are recorded. For a fixed compression ratio, a lower average distortion indicates better compression performance. Equivalently, a lower average distortion implies that storage requirement is lower when the acceptable distortion is fixed.

As we have mentioned in Section 3.2.2, the RMSE measure on all wavelet coefficients gave exactly the same results as the RMSE measure in spatial domain, as expected. Furthermore, we observed that the RMSE measure applied on different subsets of levels did not give significantly different results from those obtained from the spatial domain RMSE measure. As a result, we conclude that there is no advantage in using the RMSE measure on the wavelet coefficients instead of in the spatial domain.

For the wavelet-based entropy measures, we tested 24 different combinations of wavelet choice, level in the pyramid decomposition, and order of entropy in the computation of the measure. We also performed the same compression using spatial domain prediction measures studied in [6] for comparison. We tested compression ratios 3:1, 5:1, 7:1, 10:1, 12:1, 15:1, 20:1, 50:1, and 100:1. Instead of presenting the average distortion obtained for the image set using each measure in the MST_a compression scheme, we will only present the results concerning the best few measures out of the measures tested for each image set and groups of compression ratios.

For the entropy measures, we found that the results depend on both the compression ratio and the image sets, which is consistent with what was discovered in previous works [6]. The ranking of each measure fluctuates but the best measures tend to perform reasonably well for each type of image sets (as one of the top few measures).

We first present the results at lower compression ratios, from 3:1 up to 15:1 (Table 1). In all cases, many of the wavelet measures tested perform at least as well, and in many cases at least 10% better than all the spatial domain measures. The wavelet measures that are consistently some of the best measures all examine the entropy of wavelet coefficients in level 0 and ignore the other levels. Since coefficients in level 0 encode the finer details of the images, their encoding contributes to a significant amount of the compressed output

Table 1: Summary of best results at low compression ratios (3:1 to 15:1).

Image Set	Best Wavelet	Best Level	Entropy Order	Average Distortion	Average Distortion from spatial RMSE
Galway	CDF 9/7	0	1	0.9124	0.9265
GOES	Haar	0	4	0.6144	0.7176
Joe	Haar	0	1	0.5372	0.5943
Pig	CDF 9/7	0	2	0.3924	0.4188

at the lower compression ratios. The Galway image set is the only one in which only a small improvement is observed. It is interesting to note that the images in the set are so similar to each other that most of the measures give very similar minimum spanning trees.

For higher compression ratios, the results are somewhat less significant (Table 2). At higher compression ratios, the coefficients at the lower levels are not as important because they may be "quantized" away. We should also point out that while the relative (percentage) improvement is smaller at higher compression ratios, the absolute improvement is in fact reasonable.

In summary, the entropy measures on wavelet coefficients perform better than spatial domain measures in most cases. While the best choice of parameters depends on many factors, entropy measures based on wavelet coefficients in level 0 consistently outperform spatial domain measures. Since we are interested only in level 0 wavelet coefficients, we only have to apply the highpass filter in the wavelet transform only once, significantly reducing the amount of computation required for these measures.

5. Conclusion

In this paper, we introduced a family of wavelet-based prediction measures to determine how effectively a difference image can be compressed. By examining the entropy of only the wavelet coefficients in level 0, we obtain a measure that outperforms previous spatial domains when applied to the MST_a image set compression scheme. Furthermore, the wavelet transform only has to be applied once to obtain the coefficients in level 0, so that the computational cost of these measures is not high.

We have not examined the effect of quantization on the calculation of these measures, especially at the higher compression ratios. We may wish to quantize the wavelet coefficients first before computing the entropy. Since quantization is changed implicitly by varying compression ratio, the amount of quantization would depend on the compression ratio. This may lead to more accurate prediction measures, at the expense of slightly higher computational costs.

References

[1] C.-P. Chen, C.-S. Chen, K.-L. Chung, H.-I. Lu, and G. Tang. Image set compression through minimal-cost prediction structures. In *IEEE International Conference on Image Processing*, pages 1289–1292, 2004.

[2] The JPEG Committee. JPEG 2000. `http://www.jpeg.org/jpeg2000/index.html`.

[3] Galway.net webcams. http://www.galway.net/about/webcam/.

[4] B. Gergel. Automatic compression for image sets using a graph theoretical framework. Master's thesis, University of Lethbridge, 2007.

[5] B. Gergel, H. Cheng, C. Nielsen, and X. Li. A unified framework for image set compression. In *Proceedings of the 2006 International Conference on Image Processing, Computer Vision, & Pattern Recognition*, pages 417–423, 2006.

[6] B. Gergel, A. Schmieder, X. Li, and H. Cheng. A study of prediction measures for lossy image set compression. In *Proceedings of the 2008 International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV'08)*, pages 69–74, 2008.

[7] GOES project science. `http://rsd.gsfc.nasa.gov/goes`.

[8] R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice Hall, 3rd edition, 2007.

[9] JoeTourist Weather Station. http://wx.joetourist.net/webcam/Default.htm.

[10] K. Karadimitriou. *Set Redundancy, the Enhanced Compression Model, and Methods for Compressing Sets of Similar Images*. PhD thesis, Louisiana State University, 1996.

[11] K. Karadimitriou and J. M. Tyler. The centroid method for compressing sets of similar images. *Pattern Recognition Letters*, 19(7):585–593, 1998.

[12] P. Kopylov and P. Franti. Compression of map images by multilayer context tree modeling. *IEEE Transactions on Image Processing*, 14(1):1–11, January 2005.

[13] D.-X. Li, W. Zheng, X.-H. Xie, and M. Zhang. Optimising interview prediction structure for multiview video coding with minimum spanning tree. *Electronics Letters*, 43(23):1269–1271, November 2007.

[14] C. Nielsen and X. Li. MST for lossy compression on image sets. In *Proceedings of the Data Compression Conference*, page 463, 2006.

[15] C. Nielsen, X. Li, and K. Abma. Methods of grouping similar images for compression coding. In *Proceedings of The 2005 International Conference on Computer Vision*, pages 93–99. CSREA Press, 2005.

[16] A. Schmieder, H. Cheng, B. Gergel, and X. Li. Hierarchical Minimum Spanning Trees for Lossy Image Set Compression. In *Proceedings of the 2008 International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV'08)*, pages 57–63, 2008.

[17] A. Schmieder, H. Cheng, and X. Li. A study of clustering algorithms and validity for lossy image set compression. In *Proceedings of the 2009 International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV'09)*, pages 501–506, 2009.

[18] R. Tashakkori. *Medical Image Set Compression Using Wavelet and Lifting Combined With New Scanning Techniques*. PhD thesis, Louisiana State University, 2001.

[19] S. R. Tate. Band ordering in lossless compression of multispectral images. *IEEE Transactions on Computers*, 46(4):477–483, April 1997.

Table 2: Summary of best results at high compression ratios (20:1 to 100:1).

Image Set	Best Wavelet	Best Level	Entropy Order	Average Distortion	Average Distortion from spatial RMSE
Galway	CDF 9/7	0	1	4.223	4.230
GOES	Haar	0	1	3.258	3.314
Joe	CDF 9/7	0	1	4.159	4.261
Pig	CDF 9/7	0	2	1.066	1.124