Rate-Distortion Approach to Bit Allocation in Lossy Image Set Compression

Camara Lerner, Howard Cheng
Department of Mathematics and Computer Science, University of Lethbridge, Lethbridge, Alberta, Canada
camara.lerner@uleth.ca, howard.cheng@uleth.ca

***Abstract*—In this paper we examine the problem of bit allocation in lossy image set compression. Instead of treating each image independently, image set compression algorithms examine the relationships among similar images and remove inter-image redundancies to improve compression performance. These algorithms map the original image set into a number of prediction residual images to be coded. Typically the same bit rate is used to encode each residual. We show that a rate-distortion approach based on Lagrangian optimization can lead to further improvement in image set compression algorithms.**

***Index Terms*—Image set compression, bit allocation.**

I. Introduction

Due to the need to store large volumes of images, a number of different strategies to reduce inter-image redundancies and compress sets of similar images have been proposed in the literature [2], [4], [14], [15], [16], [17], [20], [21], [22], [25], [26], [27], [28]. Many of these techniques perform well on image sets with special types of inter-image relationships, but are less effective on others. It is often not clear *a priori* which method will work best for a particular image set.

In the majority of existing image set compression algorithms, some form of prediction or compensation is used to form a residual image for every image in the set. Each of these residual images are then compressed. Different algorithms use various strategies in determining which reference image (or images) to use as the basis for prediction of each individual image, as well as in determining how the prediction is performed. Each residual image is typically compressed with the same compression setting (e.g. quality factor, bit rate, etc.).

In this paper, we examine the bit allocation problem in image set compression. Intuitively, if a residual image is easy to compress (e.g. all zero), we do not need to allocate many bits to it. Instead, these extra bits can be allocated to other residual images in order to reduce the overall distortion of the entire decompressed image set. Bit allocation strategies have been studied extensively in image compression and video compression [23], but has so far not been examined for image set compression. We will study the bit allocation for the Centroid image set compression algorithm [15], [17], the MST and the MST_a algorithm [8], [9], [10].

The paper is organized as follows. In Section II, we review the image set compression algorithms studied in this paper. Section III examines the bit allocation problem in these algorithms and describes our approach. Section IV shows the experimental results, and concluding remarks are given in Section V.

II. Image Set Compression Methods

Data compression is achieved by removing redundancy in the given data. Typical image compression algorithms attempt to remove three types of redundancy—interpixel, psychovisual, coding redundancy [12]. These redundancies are reduced by a mapper, a quantizer and an entropy coder. For sets of similar images, additional redundancy exists among the images. A set mapper is first applied to reduce this inter-image redundancy [15], [17]. The output of the set mapper is a set of images that can be processed by ordinary image compression algorithms to reduce the remaining interpixel, psychovisual, and coding redundancy.

A. Set Mapping

The centroid method of Karadimitriou and Tyler [15], [17] first computes a centroid (average) image in the set. The set mapper then predicts each image in the set by subtracting the centroid image. If all images are very similar, the prediction residual images contain mostly zeros and can be compressed very efficiently. Although the centroid method was originally designed for lossless compression, it can be adapted for lossy compression as well [8], [9], [10].

A different approach to the set mapping problem is based on minimum spanning trees [8], [9], [10]. In this approach, each image is represented as a vertex in a graph. Between every pair of vertices u and v, there is an edge whose weight represents the cost to encode image v when u is known (or vice versa). To compress the entire image set, a minimum spanning tree (MST) is computed from this graph. A special zero root image is used as the first reference image. The algorithm repeatedly chooses a target image that is connected to another reference image that has already been coded in the MST. To avoid error propagation, the prediction is based on the decompressed reference image. This is done until all images have been compressed. Decompression of the image proceeds from the root, and reconstructs the images in the same order. For certain image sets, it is beneficial to add the average image to the image set before the algorithm is applied (called the MST_A approach). When images in the set are not all very similar but there is significant similarity between pairs of images, it has been shown that MST-based set compression algorithms perform very well [8], [10]. It can also be seen that the centroid method can be viewed as a type of spanning tree method.

For specific types of images, special set mapping methods may be more effective. For example, for satellite images with large overlap of geographic areas, adjustment for seasonal variations is needed to obtain effective prediction [25]. For multi-view images inter-view dependencies can be used to obtain better predictions [4], [18]. There are also special techniques for stereo images [6], [24].

B. Distortion Measures

An important aspect in lossy image compression is the quality of the decompressed images. The quality can be measured in a number of ways, both objectively and subjectively [29]. We will focus on objective distortion measures in this paper. These measures compare two input images and give a numerical value indicating how different the two images are.

We will examine two distortion measures in our study. The first is the well-known Root-Mean-Square-Error (RMSE) measure [12]. It is easily computable and often used to evaluate image quality, but it is well-known that it does not always correspond to human perception. The Structural Similarity (SSIM) index is a measure that appears to correspond to human evaluation better [30]. The SSIM index cannot be used directly in MST-based algorithms because it is not a metric. However, there are metrics that can be formulated based on the SSIM index [3], and this will be used in our study. In particular, we will use the D_2 metric given in [3]. We will refer to this metric as the "SSIM measure."

In lossy image compression algorithms, distortion measures are typically used to evaluate the quality of the decompressed images when images are compressed at a specific bit rate. In the spanning tree based set compression algorithms, distortion measures are also used as the edge weights in the graph—it estimates the cost to encode one image given the other. Intuitively, it is easy to compress one image given another very similar image. When the distortion measure used for edge weights is a metric (in particular, satisfies the triangle inequality), the spanning tree based algorithms are near-optimal over all algorithms based on pairwise prediction [8], [10]. This result requires that the same distortion measure be used for both edge weights and quality evaluation. Thus, this will be the case for the remainder of this paper.

III. Bit Allocation

We wish to minimize the average distortion between each original image and the decompressed image in the set at a given bit rate. This can be considered to be the best result possible when the bit rate is fixed. Equivalently, we may minimize the sum of the distortions between original images and the corresponding decompressed images in the set. This is achieved by varying the bit allocation for coding each prediction residual image to arrive at the optimal distortion. Intuitively, if a particular prediction residual is close to zero, the residual can be compressed at a lower bit rate without greatly affecting the distortion of that image. Other residuals can then be coded at a higher bit rate to minimize the sum of the distortions using the same overall bit rate.

More formally, let $S = \{I_1, \ldots, I_n\}$ be a set of n prediction residual images to be coded. If image I is compressed using b bits, the distortion between I and the decompressed image is denoted $D(I, b)$. The function $D(I, b)$ as b varies is also called the *rate-distortion curve* of image I. Let b_i be the number of bits used to compress image I_i. We want to solve the problem:

$$\text{minimize} \sum_{i=1}^{n} D(I_i, b_i), \tag{1}$$
$$\text{subject to} \sum_{i=1}^{n} b_i \leq b,$$

for some overall bit budget b. The distortion measure in our study is either the RMSE or the SSIM measure (Section II). The common case of equal bit allocation is realized when $b_i = b/n$. We call the solution to (1) the rate-distortion based bit allocation. There are a number of methods to obtain the solution of the optimization problem (1) [23]. A particularly common method is Lagrangian optimization [5].

We assume that the compression of each prediction residual is independent of each other, which is in fact false in spanning tree based image set compression algorithms because each image is predicted from a previously decompressed image. However, this assumption is commonly made to reduce computational complexity [23]. Let $\lambda \geq 0$. For each residual image to be compressed, we minimize the quantity

$$D(I_i, b_i) + \lambda b_i. \tag{2}$$

The bit rate b_i that minimizes (2) is the point on the rate-distortion curve $D(I_i, b_i)$ having slope $-\lambda$. Finding b_i in this way for each i gives the optimal bit allocation to (1) when the overall bit budget is $b = \sum_{i=1}^{n} b_i$. Unfortunately, the overall bit budget is not known *a priori*. To solve (1) for a specific value of b, one may use binary search on λ to obtain the final result. For example, if the resulting overall bit budget is too high, we may consider (2) again with a larger value of λ.

In order to solve this optimization problem, we must have access to the rate-distortion curve $D(I_i, b)$ for each residual image to be coded. The rate-distortion curve would also depend on the choice of distortion measure used. An approximation of the rate-distortion curve can be computed by compressing each image at a number of different bit rates and computing the distortion between the original image and the decompressed image. These curves are computed independently of each other, ignoring any dependencies that may be imposed by the spanning tree based compression algorithms.

IV. Experimental Results

A. Image Sets

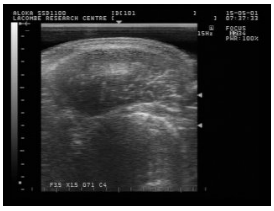
We apply our approach on four image sets that have been used in previous works to evaluate image set compression algorithms [8], [10], [21], [22]. Figure 1 shows a typical image from the four image sets. The Joe set is another webcam image set taken from a camera directed at a beach in Victoria, British Columbia [13]. The Pig set is composed of ultrasound images

TABLE I. IMPROVEMENT IN RMSE USING RATE-DISTORTION BASED BIT ALLOCATION OVER EQUAL ALLOCATION. JPEG2000 WAS USED TO COMPRESS PREDICTION RESIDUAL IMAGES.

		RMSE		
		Centroid	MST	MST_a
Joe	avg (%)	1.45	1.90	1.87
	best (%)	7.34	4.09	4.31
Pig	avg (%)	-0.37	0.50	-0.89
	best (%)	11.25	2.78	1.43
Galway	avg (%)	2.46	1.14	0.34
	best (%)	9.18	4.32	7.03
GOES	avg (%)	0.83	2.41	2.66
	best (%)	3.46	5.39	5.27

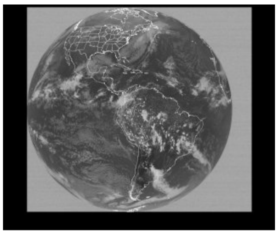
of pig rib cages. The Galway set contains webcam images from a street in Galway City, Ireland [7]. Satellite images from the GOES project [11] make up the GOES set. All the images were 8-bit gray scale images.

(a) Joe (162 images)



(b) Pig (304 images)

(c) Galway (28 images)



(d) GOES (128 images)

Figure 1. Typical images from each set.

B. Results

As in the previous studies, we use both JPEG2000 [1] and wavelet packet compression [19] to compress the prediction residual images. The experimental results are summarized in Tables I, II, III, and IV. The tables show the improvement in distortion using rate-distortion based bit allocation over equal allocation, when the overall bit budget are the same in both cases. The image sets are compressed using a number of different overall bit rates, and the average and the best improvement are shown. The bit rates tested were 0.08, 0.12, 0.16, . . . , 0.96 bpp.

Overall, we see important improvements using the our rate-distortion bit allocation method with most image sets across different combinations of distortion measures and image compression algorithms. There is less improvement in the Pig set than the other sets. That is due to the fact that most images in the Pig set are very similar to each other. As a result, the prediction residual images all have similar rate-distortion curves, so that equal bit allocation already achieves very good results. In the other image sets, the images actually form clusters so some residual images will be easier to code than others (e.g. webcam pictures taken at similar time of the day). It is also often the case that the centroid method benefits the most from our rate-distortion bit allocation method. This is expected since it is important to represent the centroid image more accurately as all other images are predicted from it. Finally, in most cases the improvement is more significant when the overall bit rate is lower—the allocation of each bit will influence the overall distortion more heavily when fewer bits are available. While the average improvement given in these results are not very significant, our rate-distortion bit allocation method is useful at improving performances of

TABLE II. IMPROVEMENT IN SSIM USING RATE-DISTORTION BASED BIT ALLOCATION OVER EQUAL ALLOCATION. JPEG2000 WAS USED TO COMPRESS PREDICTION RESIDUAL IMAGES.

		SSIM		
		Centroid	MST	MST_a
Joe	avg (%)	-0.97	0.59	0.53
	best (%)	4.84	4.83	4.80
Pig	avg (%)	0.73	1.18	1.04
	best (%)	5.02	3.84	2.93
Galway	avg (%)	1.86	0.75	0.83
	best (%)	7.00	3.50	4.74
GOES	avg (%)	0.22	1.75	1.73
	best (%)	4.38	4.69	4.60

TABLE III. IMPROVEMENT IN RMSE USING RATE-DISTORTION BASED BIT ALLOCATION OVER EQUAL ALLOCATION. WAVELET PACKET COMPRESSION WAS USED TO COMPRESS PREDICTION RESIDUAL IMAGES.

		RMSE		
		Centroid	MST	MST_a
Joe	avg (%)	-0.79	0.45	0.23
	best (%)	3.08	1.84	2.16
Pig	avg (%)	-1.69	-0.31	-1.72
	best (%)	1.96	1.68	0.90
Galway	avg (%)	1.97	2.03	2.51
	best (%)	9.27	6.85	11.57
GOES	avg (%)	2.26	4.28	4.98
	best (%)	13.11	8.36	8.15

TABLE IV. IMPROVEMENT IN SSIM USING RATE-DISTORTION BASED BIT ALLOCATION OVER EQUAL ALLOCATION. WAVELET PACKET COMPRESSION WAS USED TO COMPRESS PREDICTION RESIDUAL IMAGES.

		SSIM		
		Centroid	MST	MST_a
Joe	avg (%)	-0.54	-0.10	-0.07
	best (%)	3.67	1.92	1.97
Pig	avg (%)	-1.06	-0.12	-0.15
	best (%)	1.72	0.62	1.24
Galway	avg (%)	1.36	0.84	0.62
	best (%)	6.66	3.89	6.16
GOES	avg (%)	1.23	0.58	0.89
	best (%)	11.91	2.33	2.10

image set compression algorithms at low bit rates.

Occasionally the rate-distortion bit allocation method produces results worse than equal bit allocation. This is mainly due to the fact that our rate-distortion curves are approximated by the values at a relatively small number of fixed bit rates. Increasing the number of bit rates used in the approximation reduces the problem, but increases the run time of the optimization process.

V. Conclusions and Future Work

In this paper, we have seen that a rate-distortion bit allocation method can be used to decrease overall distortion in a lossy image set compression algorithm at fixed bit rates. In fact, the same bit allocation method can be used for other types of image set compression algorithms such as those in [4], [25].

Unfortunately, the computational costs to find the optimal solution to (1) can be high, because one needs to perform a number of compression-decompression steps at various bit rates to obtain estimates of rate-distortion curves. If we were to obtain the full rate-distortion curve for every prediction residual image in the set, it may take up to an hour to perform the optimization for relatively small image sets. In practice, however, the full rate-distortion curve is not needed. If we simply obtain the curve segments around the target overall bit rate, the final result by Lagrange optimization is a very good approximation to the optimal solution using the full curves. We have also done some preliminary work in modelling these curves so that they can be estimated more efficiently using only the edge weights. The results are promising and will be shown in a forthcoming manuscript.

Acknowledgment

The authors wish to thank the Natural Sciences and Engineering Research Council of Canada for supporting this research through its Discovery Grant and Undergraduate Student Research Award programs.

References

[1] M. Adams. JasPer project. http://www.ece.uvic.ca/~mdadams/jasper/.

[2] S. Ait-Aoudia and A. Gabis. A comparison of set redundancy compression techniques. *EURASIP Journal on Applied Signal Processing*, 2006:1–13, 2006.

[3] D. Brunet, E.R. Vrscay, and Z. Wang. A class of image metrics based on structural similarity quality index. In *Image Analysis and Recognition (ICIAR) 2011, Part I*, number 6753 in Lecture Notes in Computer Science, pages 100–110, 2011.

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

[5] Hugh Everett. Generalized lagrange multiplier method for solving problems of optimum allocation of resources. *Operations Research*, 11(3):399–417, 1963.

[6] T. Frajka and K. Zeger. Residual image coding for stereo image compression. *Optical Engineering*, 42(1):182–188, 2003.

[7] Galway.net. http://www.galway.net/webcam/images.shtml.

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

[9] B. Gergel, H. Cheng, and X. Li. A unified framework for lossless image set compression. In *Data Compression Conference*, page 448, 2006.

[10] 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 (IPCV'06)*, volume II, pages 417–423, 2006.

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

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

[13] JoeTourist Infosystems. http://www.JoeTourist.net.

[14] G. Jomier, M. Manouvrier, and M. Rukoz. Storage and management of similar images. *Journal of the Brazilian Computer Society*, 6(3):13–25, 2000.

[15] K. Karadimitriou. *Set redundancy, the enhanced compression model, and methods for compressing sets of similar images*. PhD thesis, Louisiana State University, 1996.

[16] K. Karadimitriou and J. M. Tyler. Min-max compression methods for medical image databases. *ACM Special Interest Group on Management of Data*, 26(1):47–52, 1997.

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

[18] P. Merkle, K. Müller, A. Smolic, and T. Wiegand. Efficient compression of multi-view video exploiting inter-view dependencies based on H.264/MPEG4-AVC. In *IEEE Intl. Conf. on Multimedia and Expo (ICME'06)*, pages 1717–1720, 2006.

[19] F. Meyer, A. Averbuch, and J. Strömberg. Fast adaptive wavelet packet image compression. *IEEE Transactions on Image Processing*, 9(5):792–800, 2000.

[20] Y. S. Musatenko and V. N. Kurashov. Correlated image set compression system based on new fast efficient algorithm of Karhunen-Loeve transform. In C.-C. J. Kuo, S.-F. Chang, and S. Panchanathan, editors, *Proceedings of The International Society for Optical Engineering: Multimedia Storage and Archiving Systems III*, pages 518–529, October 1998.

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

[22] C. Nielsen, X. Li, and K. Abma. Methods of grouping similar images for compression coding. In *Proceedings of the International Conference on Computer Vision (WCAC'05)*, pages 93–99, Las Vegas, June 20–23, 2005.

[23] A. Ortega and K. Ramchandran. Rate-distortion methods for image and video compression. *Signal Processing Magazine, IEEE*, 15(6):23 –50, nov 1998.

[24] M. G. Perkins. Data compression of stereopairs. *IEEE Trans. on Communications*, 40(4):684–696, 1992.

[25] V. Trivedi and H. Cheng. Lossless compression of satellite image sets using spatial area overlap compensation. In *Image Analysis and Recognition, 8th International Conference, ICIAR 2011, Part II*, pages 243–252, 2011.

[26] M. Vassilakopoulos and Y. Manolopoulos. Dynamic inverted quadtree: a structure for pictorial databases. *Information Systems*, 20(6):483–500, 1995.

[27] M. Vassilakopoulos, Y. Manolopoulos, and N. Economou. Overlapping quadtrees for the representation of similar images. *Image and Vision Computing*, 11(5):257–262, 1993.

[28] J. Wang and H. Yan. Form image compression using template extraction and matching. In *Proceedings of Visual Information Processing*, December 2000.

[29] Z. Wang and A.C. Bovik. *Modern Image Quality Assessment*. Morgan & Claypool Publishers, 2006.

[30] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.