Lossless Image Compression Using Super-Spatial Structure Prediction

Xiwen OwenZhao, *Student Member, IEEE*, and Zhihai HenryHe, *Senior Member, IEEE*

***Abstract*—We recognize that the key challenge in image compression is to efficiently represent and encode high-frequency image structure components, such as edges, patterns, and textures. In this work, we develop an efficient lossless image compression scheme called *super-spatial structure prediction*. This super-spatial prediction is motivated by motion prediction in video coding, attempting to find an optimal prediction of structure components within previously encoded image regions. We find that this super-spatial prediction is very efficient for image regions with significant structure components. Our extensive experimental results demonstrate that the proposed scheme is very competitive and even outperforms the state-of-the-art lossless image compression methods.**

***Index Terms*— Context-based adaptive lossless image coding (CALIC), lossless image compression, structure components, super-spatial structure prediction.**

I. Introduction

SPATIAL image prediction has been a key component in efficient lossless image compression [1], [2]. Existing lossless image compression schemes attempt to predict image data using their spatial neighborhood. We observe that this will limit the image compression efficiency. A natural image often contains a large number of structure components, such as edges, contours, and textures. These structure components may repeat themselves at various locations and scales. Therefore, there is a need to develop a more efficient image prediction scheme to exploit this type of image correlation.

The idea of improving image prediction and coding efficiency by relaxing the neighborhood constraint can be traced back to sequential data compression [3], [4] and vector quantization for image compression [5]. In sequential data compression, a substring of text is represented by a displacement/length reference to a substring previously seen in the text. Storer extended the sequential data compression to lossless image compression [6]. However, the algorithm is not competitive with the state-of-the-art such as context-based adaptive lossless image coding (CALIC) [1] in terms of coding efficiency. During vector quantization (VQ) for lossless image compression, the input image is processed as vectors of image pixels. The encoder takes in a vector and finds the best match from its stored codebook. The address of the best match, the residual between the original vector and its best match are then transmitted to the decoder. The decoder uses the address to access an identical codebook, and obtains the reconstructed vector. Recently, researchers have extended the VQ method to visual pattern image coding (VPIC) [7] and visual pattern vector quantization (VPVQ) [8]. The encoding performance of VQ-based methods largely depends on the codebook design. To our best knowledge, these methods still suffer from lower coding efficiency, when compared with the state-of-the-art image coding schemes.

In the intra prediction scheme proposed by Nokia [9], there are ten possible prediction methods: DC prediction, directional extrapolations, and block matching. DC and directional prediction methods are very similar with those of H.264 intra prediction [10]. The block matching tries to find the best match of the current block by searching within a certain range of its neighboring blocks. As mentioned earlier, this neighborhood constraint will limit the image compression efficiency since image structure components may repeat themselves at various locations.

In fractal image compression [11], the self-similarity between different parts of an image is used for image compression based on contractive mapping fixed point theorem. However, the fractal image compression focuses on contractive transform design, which makes it usually not suitable for lossless image compression. Moreover, it is extremely computationally expensive due to the search of optimum transformations. Even with high complexity, most fractal-based schemes are not competitive with the current state of the art [12].

In this work, we develop an efficient image compression scheme based on super-spatial prediction of structure units. This so-called super-spatial structure prediction breaks the neighborhood constraint, attempting to find an optimal prediction of structure components within the previously encoded image regions. It borrows the idea of motion prediction from video coding, which predicts a block in the current frame using its previous encoded frames. In order to "enjoy the best of both worlds", we also propose a classification scheme to partition an image into two types of regions: *structure regions (SRs)* and *nonstructure regions (NSRs)*. Structure regions are encoded with super-spatial prediction while NSRs can be efficiently encoded with conventional image compression methods, such as CALIC. It is also important to point out that no codebook is required in this compression scheme, since the best matches of structure components are simply searched within encoded image regions. Our extensive experimental results demonstrate that the proposed scheme is very competitive and even outperforms the state-of-the-art lossless image compression methods.

Manuscript received November 11, 2009; revised January 11, 2010. First published January 22, 2010; current version published February 24, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Lisimachos Paul Kondi.

The authors are with Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO 65203 USA (e-mail: xzc33@mail.missouri.edu; HeZhi@missouri.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LSP.2010.2040925

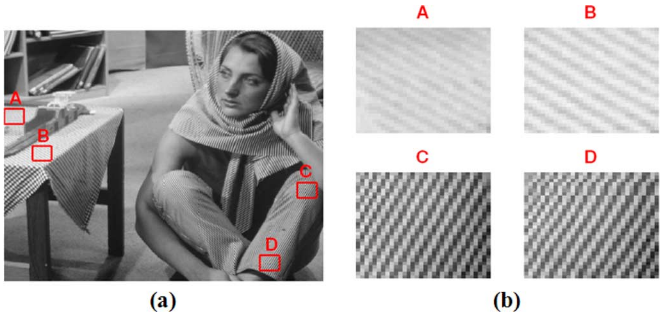

Fig. 1. (a) *Barbara* image. (b) Four image blocks extracted from *Barbara*.

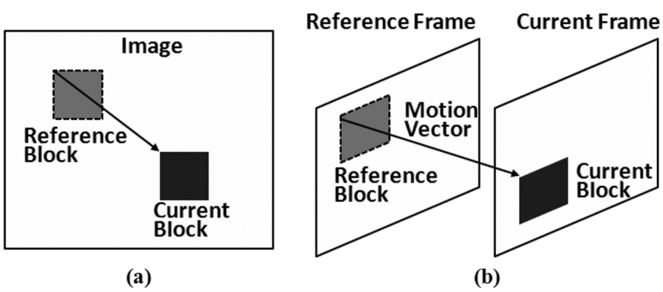

Fig. 2. (a) Super-spatial prediction. (b) Motion prediction in video coding.

II. Super-Spatial Structure Prediction for Efficient Image Compression

In this section, we explain the basic idea of super-spatial prediction and how it can be used for efficient image compression.

We observe that a real world scene often consists of various physical objects, such as buildings, trees, grassland, etc. Each physical object is constructed from a large number of structure components based upon some predetermined object characteristics. These structure components may repeat themselves at various locations and scales. For example, Fig. 1(a) shows the image *Barbara* and Fig. 1(b) shows four patches (32×32 blocks) extracted from different locations of the image. It can be seen that they share very similar structure characteristics. Therefore, it is important to exploit this type of data similarity and redundancy for efficient image coding.

The proposed super-spatial prediction borrows the idea of motion prediction from video coding, as illustrated in Fig. 2. In motion prediction, we search an area in the reference frame to find the best match of the current block, based on some distortion metric. The chosen reference block becomes the predictor of the current block. The prediction residual and the motion vector are then encoded and sent to the decoder. In super-spatial prediction, we search within the previously encoded image region to find the prediction of an image block. In order to find the optimal prediction, at this moment, we apply brute-force search. The reference block that results in the minimum block difference is selected as the optimal prediction. For simplicity, we use the sum of absolute difference (SAD) to measure the block difference. Besides this direct block difference, we can also introduce additional H.264-like prediction modes, such as horizontal, vertical, and diagonal prediction [10], as illustrated in Fig. 3. Here, block B is the current block to be encoded and block A is the reference block from the previous reconstructed image region. The best prediction mode will be encoded and sent to the decoder.

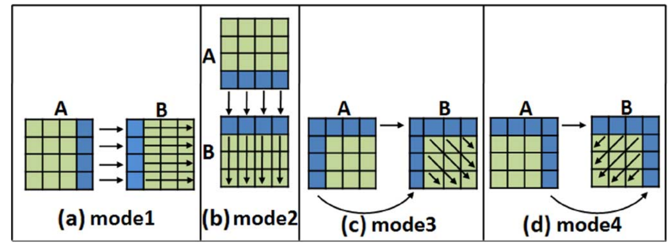

Fig. 3. Additional prediction modes.

As in video coding, we need to encode the position information of the best matching reference block. To this end, we simply encode the horizontal and vertical offsets, (dx, dy), between the coordinates of the current block and the reference block using context-adaptive arithmetic encoder.

We observe that the size of the prediction unit is an important parameter in the super-spatial prediction. When the unit size is small, the amount of prediction and coding overhead will become very large. However, if we use a larger prediction unit, the overall prediction efficiency will decrease. In this work, we attempt to find a good tradeoff between these two and propose to perform spatial image prediction on block basis. For example, in our experiments, we set the block size to be 4×4.

We can see that, when compared to VQ-based image encoders, the proposed super-spatial prediction scheme has the flexibility to incorporate multiple H.264-style prediction modes. When compared to other neighborhood-based prediction methods, such as gradient-adjusted prediction (GAP) [1] and H.264 Intra prediction, it allows the block to find the best match from the whole image which will significantly reduces the prediction residual.

III. Image Block Classification

From the experimental results in Section V, we will see that super-spatial prediction works very efficiently for image regions with a significant amount of structure components. However, due to its large overhead and high computational complexity, its efficiency will degrade in nonstructure or smooth image regions. Therefore, we propose a block-based image classification scheme. More specifically, we partition the image into blocks (e.g. 4×4 blocks). We then classify these blocks into two categories: structure and nonstructure blocks. Structure blocks are encoded with super-spatial prediction. Nonstructure blocks are encoded with conventional lossless image compression methods, such as CALIC. CALIC is a spatial prediction based scheme, in which GAP is used for adaptive image prediction [1]. We propose to explore two classification methods. In the first method (denoted by Method_A), we compare the GAP prediction against the super-spatial prediction. Our basic idea is that, if a block can be predicted more efficiently by GAP than super-spatial prediction, we classify it into NSRs. For simplicity, we use SAD to measure the prediction performance. Otherwise, we classify it into structure regions. In the second method (denoted by Method_B), we simply perform GAP prediction on the original image and compute the prediction error for each block. If the prediction error is larger than a

Fig. 4. (a) *Barbara* image. (b) Classification result using Method_A.

Fig. 5. (a) Original *Barbara* image. (b) Nonstructure regions. (c) Structure regions.

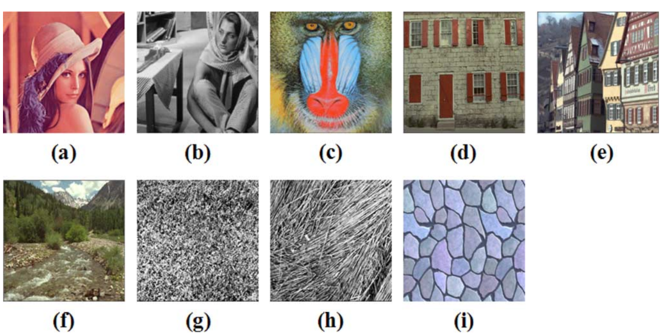

Fig. 6. Nine test images from USC and Kodak image databases.

given threshold, then it is considered as a structure block. Otherwise, it is classified as a nonstructure block. We can see that Method_A has a much higher computational complexity than Method_B since it needs to perform super-spatial prediction for each block. Fig. 4 shows a classification result for the Barbara image using Method_A. The white pixels in Fig. 4(b) indicate structure blocks.

IV. Image Coding Based on Super-Spatial Structure Prediction

Based on the super-spatial prediction and image content separation proposed in previous sections, we develop a lossless image coding system. The proposed image coding system has the following major steps. **First**, using the method described in Section III, we classify the image into NSRs and structure regions as shown in Fig. 5(b) and (c), respectively. The classification map is encoded with arithmetic encoder. **Second**, based on the classification map, our encoder switches between super-spatial prediction scheme to encode structure regions and CALIC scheme to encode NSRs.

We observe that the super-spatial prediction has relatively high computational complexity because it needs to find the best match of the current block from previous reconstructed image regions. Using the Method_B classification, we can classify the image into structure and NSRs and then apply the super-spatial prediction just within the structure regions since its prediction gain in the nonstructure smooth regions will be very limited. This will significantly reduce overall computational complexity.

TABLE I
Prediction Performance Comparison on the Structure Regions

Test Image	GAP	H.264 Intra Prediction	Structure Prediction	Saving over H.264 Intra Prediction
Lena	20.2	15.9	**7.0**	**56%**
Barbara	22.7	17.8	**6.6**	**63%**
Baboon	22.0	20.5	**11.4**	**44%**
kodim01	16.0	16.0	**8.1**	**49%**
kodim08	18.3	21.8	**8.4**	**61%**
kodim13	22.2	20.8	**11.2**	**46%**
USC1.2.01	31.2	37.8	**16.4**	**57%**
USC1.2.03	31.4	31.6	**13.4**	**57%**
Floor	15.7	12.1	**2.5**	**79%**

V. Experimental Results

We have implemented the proposed super-spatial prediction scheme in CALIC [1], a very efficient lossless image compression scheme which outperforms other state-of-the-art coding methods, such as JPEG-LS and LOCO-I. In this work, we choose a block size of 4×4. The output bit stream of the proposed encoder consists of bits for the following major syntax components: image classification map, bits for NSRs, bits for prediction residual of structure blocks, addresses of reference blocks, and prediction mode.

Fig. 6 shows the test images. We first evaluate the prediction efficiency of the proposed super-spatial prediction scheme. We use Method_A to classify the image blocks. We then apply three schemes to predict the image data within the structure regions: GAP from CALIC, H.264 Intra prediction, which is a very efficient spatial prediction scheme proposed in H.264 video coding, and the proposed super-spatial prediction. We measure the SAD of the prediction residual. The results are summarized in Table I. Our super-spatial prediction scheme is able to significantly reduce the prediction error. Compared with H.264 Intra prediction, it is able to reduce the SAD by up to 79% within these structure regions.

In Table II, we compare the coding bit rate of the proposed lossless image coding method based on super-spatial prediction with CALIC [1], one of the best encoders for lossless image compression. We can see that the proposed scheme outperforms CALIC and save the bit rate by up to 13%, especially for images with significant high-frequency components. Table III shows the percentages of bits used by three major syntax components: SRs, NSRs, and overhead information (including classification map, reference block address, and prediction mode). In Table IV, we evaluate two image classification methods, Method_A and Method_B, as discussed in Section III. We can see that, the Method_B, although has low computational complexity, its performance loss is very small, when compared to Method_A. In Table V, we evaluate two super-spatial prediction methods: 1) prediction of the current block from previous reconstructed image regions and 2) prediction from previous reconstructed *structure* image regions. We can see that limiting the prediction reference to structure regions only increases

TABLE II
COMPRESSION PERFORMANCE COMPARISON WITH CALIC

Test Images	CALIC Bit Rate (bpp)	This Work Bit Rate (bpp)	Bit Rate Saving
Lena	4.097	**4.086**	0.011
Barbara	4.58	**4.471**	-0.109
Baboon	5.898	**5.71**	-0.188
kodim01	5.091	**4.998**	-0.093
kodim08	5.049	**5.008**	-0.041
kodim13	5.818	**5.638**	-0.18
USC1.2.01	6.973	**6.603**	-0.37
USC1.2.03	6.873	**6.382**	-0.491
Floor	3.855	**3.4**	-0.455

TABLE III
PERCENTAGES OF BITS OF MAJOR SYNTAX COMPONENTS

Test Image	Percentage of Bits		
	NSR	SR	Overhead
Lena	99%	0.8%	0.2%
Barbara	69%	25%	6%
Baboon	55%	36%	9%
kodim01	68%	25%	7%
kodim08	73%	21%	6%
kodim13	51%	39%	10%
USC1.2.01	18%	68%	14%
USC1.2.03	14%	71%	15%
Floor	86%	10%	4%

TABLE IV
COMPARISON BETWEEN METHOD_A AND METHOD_B IMAGE CLASSIFICATION METHODS

Image	Method_A	Method_B	Percentage of Bit Increase
Lena	4.086	4.086	0%
Barbara	4.471	4.463	-0.18%
Baboon	5.71	5.706	-0.07%
kodim01	4.998	5.011	0.26%
kodim08	5.008	5.033	0.50%
kodim13	5.638	5.671	0.59%
USC1.2.01	6.603	6.612	0.14%
USC1.2.03	6.382	6.387	0.08%
Floor	3.4	3.546	4.29%

TABLE V
IMPACT OF SEARCH RANGE OF SUPER-SPATIAL PREDICTION

Image	Search in Encoded Image Domain	Search in Encoded Structure Regions	Percentage of Bit Increase
Lena	4.086	4.086	-0%
Barbara	4.471	4.506	+0.78%
Baboon	5.71	5.736	+0.46%
kodim01	4.998	5.041	+0.86%
kodim08	5.008	5.083	+1.50%
kodim13	5.638	5.687	+0.87%
USC1.2.01	6.603	6.633	+0.45%
USC1.2.03	6.382	6.415	+0.52%
Floor	3.4	3.666	+7.82%

the overall coding bit rate by less than 1%. This implies most structure blocks can find its best match in the structure regions.

VI. CONCLUSION

In this work, we have developed a simple yet efficient image prediction scheme, called super-spatial prediction. It is motivated by motion prediction in video coding, attempting to find an optimal prediction of a structure components within previously encoded image regions. When compared to VQ-based image encoders, it has the flexibility to incorporate multiple H.264-style prediction modes. When compared to other neighborhood-based prediction methods, such as GAP and H.264 Intra prediction, it allows the block to find the best match from the whole image which significantly reduces the prediction residual by up to 79%. We classified an image into structure regions and NSRs. The structure regions are encoded with super-spatial prediction while the NSRs are encoded with existing image compression schemes, such as CALIC. Our extensive experimental results demonstrated that the proposed scheme is very efficient in lossless image compression, especially for images with significant structure components. In our future work, we shall develop fast and efficient algorithms to further reduce the complexity of super-spatial prediction. We also notice that when the encoder switches between structure and NSRs, the prediction context is broken and it will degrade the overall coding performance. In our future work, we shall investigate more efficient schemes for context switching.

ACKNOWLEDGMENT

The authors would like to thank X. Wu for generously providing the source codes of CALIC.

REFERENCES

[1] X. Wu and N. Memon, "Context-based, adaptive, lossless image coding," *IEEE Trans. Commun.*, vol. 45, no. 4, pp. 437–444, Apr. 1997.

[2] M. J. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS," *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1309–1324, Aug. 2000.

[3] J. Ziv and A. Lempel, "A universal algorithm for sequential data compression," *IEEE Trans. Inform. Theory*, vol. 23, no. 3, pp. 337–343, May 1977.

[4] J. Ziv and A. Lempel, "Compression of individual sequences via variable-rate coding," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 5, pp. 530–536, Sep. 1978.

[5] V. Sitaram, C. Huang, and P. Israelsen, "Efficient codebooks for vector quantization image compression with an adaptive tree search algorithm," *IEEE Trans. Commun.*, vol. 42, no. 11, pp. 3027–3033, Nov. 1994.

[6] J. Storer, "Lossless image compression using generalized LZ1-type methods," in *Proc. IEEE Data Compression Conference (DCC'96)*, pp. 290–299.

[7] D. Chen and A. Bovik, "Visual pattern image coding," *IEEE Trans. Commun.*, vol. 38, no. 12, pp. 2137–2146, Dec. 1990.

[8] F. Wu and X. Sun, "Image compression by visual pattern vector quantization (VPVQ)," in *Proc. IEEE Data Compression Conf. (DCC'2008)*, pp. 123–131.

[9] [Online]. Available: http://wftp3.itu.int/av-arch/video-site/0005_Osa/a15j19.doc

[10] T. Wiegand, G. J. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, Jun. 2003.

[11] Y. Fischer, "Fractal image compression," in *SIGGRAPH'92 Course Notes*.

[12] B. Wohlberg and G. de Jager, "A review of the fractal image coding literature," *IEEE Trans. Image Process.*, vol. 8, no. 12, pp. 1716–1729, Dec. 1999.