Level-Successive Encoding for Digital Photography

Mehmet Celik, Gaurav Sharma*, A.Murat Tekalp
University of Rochester,
Rochester, NY
* Xerox Corporation,
Webster, NY

Abstract

We propose a level-successive encoding scheme for the compression of continuous tone images. The compressed bit stream is partitioned into individual segments corresponding to successive bits and arranged in order of decreasing significance of the bits. Bit-plane scalability is achieved by recursively partitioning the image into an LSB bit-plane and remaining higher order bits and encoding the LSB bit-plane using the remaining bits as context. This allows the scheme to achieve excellent compression performance by exploiting both spatial and inter-level correlations. We compare the proposed scheme with a number of scalable and non-scalable lossless image compression algorithms to benchmark its performance. Results indicate that the level-embedded compression incurs only a small penalty in compression efficiency over conventional lossless schemes while offering the benefit of easy bit-plane scalability.

1. Introduction

The current trend in digital cameras is toward increasingly higher resolution and bit depth for image capture. While this results in improved image quality, it also translates to increased memory requirements for storage of captured images. The use of image compression can mitigate the increase in storage space.

Algorithms that preserve the quality of the image during compression and allow for its exact reconstruction are called *lossless* compression algorithms. CALIC [1], JPEG-LS [2], and JPEG2000 [3] are among well-known lossless image compression algorithms. Among these CALIC provides best compression ratios over typical images, whereas, JPEG-LS is a low complexity alternative with competitive efficiency. The JPEG2000 standard, on the other hand, is a wavelet-based technique, which provides a unified approach for lossy-to-lossless compression. Nevertheless, none of these algorithms can provide the high compression ratios often required by commercial photography applications. Lossy compression algorithms, on the other hand, provide greater compression ratios, but they sacrifice image quality.

In several applications, it is advantageous to have *scalable compression*, where a desired level of compression may be determined after the source has been compressed. This allows flexibility in determining the data rate to meet the bandwidth, memory and processing power constraints imposed by the operating environment - with a corresponding trade-off in image quality. Scalable compression is typically achieved by generating an *embedded bit-stream* which has the property that any truncation of the (compressed) bit stream also yields an efficient compressed representation of the source in a rate-distortion sense. That is, the distortion for the truncation of the embedded stream to a particular rate should be comparable to the distortion achievable for that rate by any compressed representation, embedded or otherwise.

In this paper, we propose a specific instance of scalable compression called level-embedded compression. Level-embedded scalability refers to bit-plane scalability in the image pixel value domain. We first describe the concept of level-scalability and provide an overview of scenarios where it may be applied in digital photography. Next, we outline a new level-embedded compression algorithm for bit-plane scalability that we have recently proposed [4]. The performance of the scheme is compared against state-of-the-art compression methods and finally we end with concluding remarks.

2. Level-Embedded Compression and Applications

Level-embedded compression refers to level scalability in the image pixel domain. For an R bit image, a level-scalable compression algorithm generates an embedded bit-stream that allows scalability in the pixel-wise peak (maximum) absolute error (PAE) ranging from 0 (lossless) to $2^{(R-1)}$ (single bit thresholded representation) in suitably chosen steps. Note that, a particular instance of level-scalability is bit-plane scalability. Bit-plane scala-

bility corresponds to the case where embedding levels are chosen to be powers of 2.

A bit-plane scalable method is especially useful in applications where data is acquired by a capture device with a high dynamic range or bit-depth. A lower bit-depth representation is often sufficient for most purposes and the higher bit-depth data is only required for specialized analysis/enhancement or archival purposes. For example, a digital camera may preserve an acquired image without loss at full bit-depth of the acquisition device, but truncate later if necessary, say in order to create space for additional images. Typical change in visual quality due to bit-plane truncation is demonstrated in Fig. 1. In the figure, representation of the *Gold Hill* image is seen for 8, 6, 4, and 2 bits per pixel. Despite the degradation in quality, lower bit-depth images provide sufficient details.

Figure 1: *GoldHill image at different bit-depths. 8 bpp (top-left), 6 bpp (top-right), 4 bpp (bottom-left), 2 bpp (bottom-right).*

When the full bit-depth image is stored in a conventional lossless compressed stream, subsequent truncation of lower order bits requires a decompression and reconstruction of the image prior to truncation, followed by a compression of the resulting level-truncated image. If on the other hand, the compression scheme (and the corresponding bit stream) is level-embedded, the truncation can effectively be performed in the bit stream itself by dropping the segment of the stream corresponding to the truncated lower levels. The latter option is often much more desirable because of its memory and computational simplicity, which translate to lower power, time, and memory requirements.

Level-embedded compression provides similar advantages at the display end. If the display environment supports only a limited bit-depth, conventional compression methods require the decompression of the full bit-depth image and subsequent truncation. A bit-plane scalable compression method, on the other hand, processes only a part of the bit-stream which corresponds to the image truncated to a lower bit-depth that matches the capabilities of the display system. This functionality is especially important for mobile/portable devices which often have said display limitations. Level-embedded compression method saves valuable power, memory and time resources on these devices.

JPEG2000 offers scalability in resolution and distortion by allowing reconstruction of lower resolution and/or lower signal-to-noise-ratio (SNR) images. The scalability in JPEG2000 is, however, different from the scalability provided by level embedded compression. Scalability in JPEG2000 is implemented in the wavelet transform coefficient domain. Truncation of bit-planes in the wavelet transform coefficient domain often results in spatial artifacts and it does not, in general, correspond to the proposed level embedded scalability in the image pixel value domain.

As a result, in several applications, level-embedded scalability is more natural and acceptable than the scalability in JPEG2000. Document scanning applications offer a specific example, where one may require archival of complete gray level information, even though most users of the data may need only thresholded bi-level information. These dual needs are readily and efficiently met by using level embedded compression. Another example is the use of digital photography for legal evidence, where level-embedded scalability may be more acceptable because the potential for spatial artifacts in alternate scalable compression schemes may cast doubts on the veracity of photographic evidence. The bit-depth truncation in level-embedded compression is analogous to using an acquisition device with a lower resolution A/D converter and therefore likely to be more acceptable. The technique may also be applied to medical imagery which is often gathered with a high dynamic range and compressed losslessly for archival purposes. The level embedding can be beneficial if a more limited dynamic range display is available or the image is to be communicated remotely for telemedicine applications. Once again, it is sometimes preferable to truncate bit-planes instead of using alternate compression schemes that may introduce spatial artifacts and render the image useless in clinical applications. In other non-critical applications, however, the JPEG2000 scalability based on wavelet domain truncation is often superior to level truncation, because it results in a smaller visual distortion.

3. Level Embedded Compression Algorithm

In this section, we outline the proposed level-embedded compression algorithm. A preliminary paper describing the method appears in [4] and a more detailed description is scheduled for publication in [5].

The method is applied recursively by partitioning the image into two levels at each stage. We therefore describe the algorithm for two embedding levels: a base layer corresponding to the higher levels and a residual layer comprising of the lower levels. The method is subsequently generalized to multiple levels by partitioning the base layer further if necessary.

The image is separated into the base layer and a residual layer. The base layer is obtained by dividing each pixel value by a constant integer L ($B_L(s) = \lfloor \frac{s}{L} \rfloor$). L specifies the amplitude of the enhancement layer, which is the remainder, which is also called *the residual* ($r = s - L \lfloor \frac{s}{L} \rfloor$). We also call the quantity $L \lfloor \frac{s}{L} \rfloor$ as the quantized pixel, $Q_L(s)$. Note that the use of a power of 2 for L corresponds to partitioning of the images into more significant and less significant bit planes, and other values generalize this notion to a partitioning into higher and lower levels.

Since the resulting base layer representing the most significant levels of the image is coded without any reference to the enhancement layer and its statistics closely resemble those of a full bit-depth image, any lossless compression algorithm is well suited for this layer. In this paper, we use the CALIC algorithm for the base layer compression. CALIC is among the best performing lossless compression algorithms details of which may be found in [1, 6].

As the enhancement layer, or the residual signal, represents the lowest levels of a continuous-tone image, its compression is a challenging task. For small values of L, the residual typically has no structure, and its samples are virtually uniformly distributed and uncorrelated from sample to sample. Direct compression of the residual therefore is highly inefficient. However, if the rest of the image information is used as side-information, significant coding gains can be achieved in the compression of the residual, by exploiting the spatial correlation among pixel values and the correlation between high and low levels (bitplanes) of the image.

The proposed method for the compression of the enhancement layer has three main components: *i)* prediction, *ii)* context modeling and quantization, *iii)* conditional entropy coding. The overall structure and components for the enhancement layer compression scheme are inspired by the CALIC algorithm [1] but adapted to the special case of encoding an enhancement layer rather than a full image. The prediction component is aimed at decreasing the redundancy in the enhancement layer data by exploiting cor-

relations both with the already decoded base layer and with available spatial neighbors. The context modeling stage allows the prediction to adapt to locally varying statistics in the image and also enables the same adaptation for the conditional entropy coding. Finally, the conditional entropy coding is an adaptive arithmetic coder that estimates and exploits context dependent probability models to encode the information losslessly into the smallest number of bits. The algorithm is presented below in pseudo-code. Full details of the algorithm can be found in [5].

1. $\hat{s}_O$ = Predict Current Pixel();
2. d, t = Determine Context D,T($\hat{s}_O$);
3. $\dot{s}_O$ = Refine Prediction($\hat{s}_O, d, t$);
4. θ = Determine Context $\Theta(\dot{s}_O)$;
5. If ($\theta \geq 0$),
    Arithmetic Encode/Decode Residual(r_O, d, θ);
  else,
    Arithmetic Encode/Decode Residual($L - 1 - r_O, d, |\theta|$);

As indicated earlier, the algorithm is generalized to multiple levels by repeated decomposition of the base layer. In the first stage, the image is separated into a base layer B_1 and an enhancement layer r_1 using level L_1. In the second stage, the base layer B_1 is further separated into a base layer B_2 and enhancement layer r_2 using a (potentially different) level L_2. The process is continued for additional stages as desired. Each enhancement layer r_i is compressed using the corresponding base layer B_i, and last base layer B_n is compressed as earlier.

4. Experimental Results

We evaluated the performance of the proposed scheme using four 512×512 pixel, 8-bit gray-scale images, namely *Mandrill, Barbara, GoldHill* and *Lena*.

Although the algorithm works for arbitrary values of the embedding level L, in order to allow comparison with bit-plane compression schemes, here we concentrate on bit-plane embedded coding, which corresponds to using $L = 2$. Furthermore, the recursive scheme outlined in the previous section is used to obtain multi-level embeddings with more than one enhancement layer, each consisting of a bit-plane. The number of enhancement layers, i.e. embedded bit-planes, is varied from 1 through 7. One (1) enhancement layer corresponds to the case where the LSB-plane is the enhancement layer and 7 MSB-planes form the base layer. Likewise, seven (7) enhancement layers correspond to a fully scalable bit-stream, where all bit-planes can be reconstructed consecutively, starting with the most significant and moving down to the least significant. As indicated earlier, in each case, the corresponding base layer is compressed using the CALIC algorithm.

In Table. 1, the performance of the proposed algorithm is compared with that of state-of-the-art lossless compression methods. The methods included in this benchmarking include the regular (non-embedded) lossless compression methods: CALIC, JPEG2000, JPEG-LS; and embedded compression using JBIG (independent bit-planes), gray-coded JBIG -"JBIG(gray)", and the level-embedded scheme proposed in this paper. The different level embeddings are denoted as L.E. 1, L.E. 2, ... , L.E. 7 for the cases corresponding to 1, 2, ... 7 enhancement layers. In our experiments, CALIC provided the best compression rates for non-embedded compression. Therefore, in Table. 1, we list results for all non-embedded schemes and the level-embedded scheme proposed here as the percentage increases in bit-rate with respect to the CALIC algorithm.

From the table, it is apparent that JPEG-LS and JPEG2000 offer fairly competitive performance to CALIC with only modest increases in bit rate. Nonetheless, just like CALIC these methods are not bit-plane scalable. JPEG2000 provides resolution and distortion scalability but not bit-plane scalability. In its default mode, JBIG provides bit-plane scalability, however at a significant loss of coding efficiency (almost a 35% increase in bit rate over CALIC, on average). The level embedded compression scheme does significantly better than JBIG in this mode. The performance of JBIG is significantly improved when pixel values are gray-coded prior to separation into bit-planes. This corresponds to the row labeled "JBIG(gray)" in the table. A one-to-one relation between the base layer bits at any level and corresponding gray-coded bit-planes allows for level-embedded construction–with additional processing–in JBIG(gray) algorithm. In this case, JBIG's performance is comparable to the proposed method. Nevertheless, the proposed scheme offers additional flexibility wherein the number of embedded levels can be limited to improve compression efficiency. For a small number of embedding levels the penalty is quite small with up to 4 enhancement layers requiring under 8% increase in bit-rate over CALIC.

In Table. 1, we also see that the proposed method incurs a penalty which increases roughly linearly for each image with increase in the number of enhancement layers (embedded bit-planes). In a hypothetical application, where 2 bit-planes are embedded, for instance, to truncate 8-bits to 6-bits in a digital camera, the increase in bit-rate is 3% on the average. This number is quite competitive with the non-scalable JPEG-LS and CALIC algorithms in view of the added functionality. It is also better than the corresponding rate for the JPEG2000 algorithm. When all bit-planes are embedded the penalty increases to 14%. This is approximately equal to the JBIG(gray) and is considerably worse than the JPEG2000, where alternate scalability

Table 1: *Performance of level-embedded compression scheme against different lossless compression methods. Percent increase with respect to CALIC is indicated.*

	Image	Mand	Barb	Gold	Lena	Avg.
	Comp. Method	Best loss-less compression rate (Baseline)				
	CALIC (bpp)	5.66	4.42	4.58	4.08	4.36
		Percent Increase in bit-rate *wrt* baseline				
Regular	CALIC	0.0	0.0	0.0	0.0	0.0
	JPEG2000	4.0	4.6	4.6	5.2	4.8
	JPEG-LS	2.8	6.2	1.8	3.4	3.8
Level-Embedded	JBIG	26.2	36.3	33.6	39.7	36.5
	JBIG(gray)	11.2	17.6	13.7	15.8	15.0
	L.E. 1	0.2	1.4	0.7	1.1	0.8
	L.E. 2	0.9	4.1	2.2	3.7	2.6
	L.E. 3	2.2	6.3	4.7	5.8	4.5
	L.E. 4	3.4	10.1	6.6	8.6	6.9
	L.E. 5	5.3	14.0	8.5	11.7	9.5
	L.E. 6	6.6	17.5	10.7	14.4	11.9
	L.E. 7	7.6	20.0	12.6	17.5	13.9

is provided. The degradation at higher levels of embedding is not a major concern because most applications of level-embedded compression are likely to require only a small number of embedded bit planes.

The rate distortion performance of the level-embedded scheme is compared against the rate-distortion performance for JPEG2000 in Fig. 2 for the peak absolute error (PAE) distortion metric.

$$D(PAE) = max||s_i - \hat{s_i}|| \qquad (1)$$

where $\hat{s_i}$ is the reconstructed value at pixel position i. Note that, for the proposed scheme the reconstruction level is selected as the mid-point of the quantization interval. For instance, if only the most significant bit is received and its value is zero then actual value of the pixel lies in $[0, 128)$ (for an 8-bpp image) and $\hat{s} = 64$ is selected as the reconstruction value. From Fig. 2 we can see that the proposed scheme offers better rate vs PAE distortion performance than JPEG2000, with particularly significant benefits in the near lossless region where only a small number of embeddings allowing truncations of LSBs are required.

5. Conclusions

We present a level-embedded lossless image compression method and outlined its applications. In legal and medical imaging applications where spatial artifacts are undesirable, level-embedded compression provides the efficient means for truncation of high bit-depth images for near lossless compression. Experimental results comparing the method with state-of-the-art lossless compression methods indicate that level scalability is achieved with only a
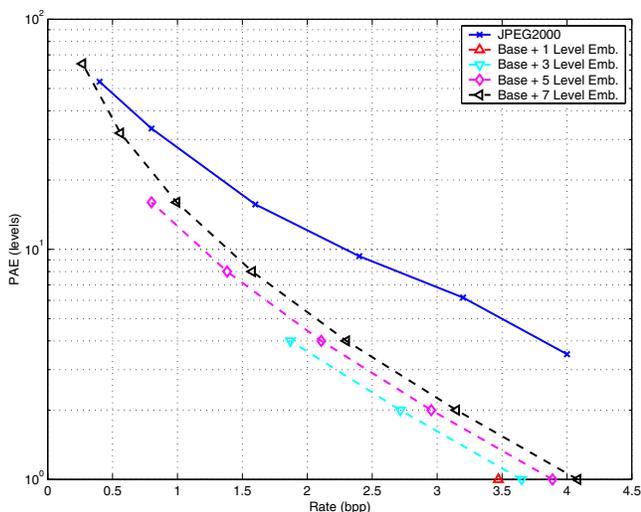
Figure 2: *Comparison of Rate-Distortion $|| \cdot ||_\infty$ performance (averaged over all images) for proposed level-embedded scheme and JPEG2000.*

small penalty in the compression efficiency over regular (non level-embedded) compression schemes.

References

[1] X. Wu, "Lossless compression of continuous-tone images via context selection, quantization, and modelling," *IEEE Trans. on Image Proc.*, vol. 6, no. 5, pp. 656–664, May 1997.

[2] ISO/IEC 14495-1, "Lossless and near-lossless compression of continuous-tone still images- baseline," 2000.

[3] ISO/IEC 15444-1, "Information technology–JPEG 2000 image coding system–part 1: Core coding system," 2000.

[4] M. U. Celik, A. M. Tekalp, and G. Sharma, "Level-embedded lossless image compression," in *Proc. of IEEE ICASSP*, Hong Kong, April 2003, accepted for presentation.

[5] M.U. Celik, G. Sharma, and A.M. Tekalp, "Gray-level-embedded lossless image compression," *to appear in EURASIP Image Comm.*, July 2003.

[6] X. Wu and N. Memon, "Context-based, adaptive, lossless image codec," *IEEE Trans. on Communications*, vol. 45, no. 4, pp. 437–444, April 1997.

Biographies

Mehmet Utku Celik received the B.Sc. degree in electrical and electronic engineering in 1999 from Bilkent University, Ankara, Turkey and the M.Sc. degree in electrical and computer engineering in 2001 from the University of Rochester, Rochester, NY.

Currently, he is a Research Assistant and Ph.D. candidate in the Electrical and Computer Engineering Department, University of Rochester. His research interests include digital watermarking and data hiding—with emphasis on multimedia authentication—image and video processing, and cryptography. Mr. Celik is a member of the ACM and the IEEE Signal Processing Society.

Gaurav Sharma received the B.E. degree in electronics and communication engineering from University of Roorkee, India in 1990, the M.E. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, in 1992, and the M.S. degree in applied mathematics and Ph.D. degree in electrical and computer engineering from North Carolina State University (NCSU), Raleigh, in 1995 and 1996, respectively.

Since 1996, he has been a Member of Research Staff in Xerox Corporation's Research and Technology division located in Webster, NY. He is also involved in teaching in an adjunct capacity at the Electrical and Computer Engineering Departments at the Rochester Institute of Technology, Rochester, NY. His research interests include image security and watermarking, color science and imaging, signal restoration, and halftoning. Dr. Sharma is a member of Sigma Xi, Phi Kappa Phi, Pi Mu Epsilon, and is the 2003 Chair for the Rochester Chapter of the IEEE Signal Processing Society.

A. Murat Tekalp received the M.S. and Ph.D. degrees in electrical, computer, and systems engineering from Rensselaer Polytechnic Institute (RPI), Troy, New York, in 1982 and 1984, respectively.

From December 1984 to August 1987, he was a Research Scientist at Eastman Kodak Company, Rochester, New York. He joined the Electrical and Computer Engineering Department, University of Rochester, Rochester, NY, in September 1987, where he is currently an endowed Distinguished Professor. His current research interests are in the area of digital image and video processing, including image restoration, video segmentation, object tracking, content-based video description, and protection of digital content. At present, he is the Editor-in-Chief of the EURASIP Journal on Image Communication. He authored Digital Video Processing (Englewood Cliffs, NJ: Prentice-Hall, 1995). He is the Founder and First Chairman of the Rochester Chapter of the IEEE Signal Processing Society.