Improved Resolution Scalability for Bilevel Image Data in JPEG2000

Rahul Raguram, *Member, IEEE*, Michael W. Marcellin, *Fellow, IEEE*, and Ali Bilgin, *Senior Member, IEEE*

***Abstract*—In this paper, we address issues concerning bilevel image compression using JPEG2000. While JPEG2000 is designed to compress both bilevel and continuous tone image data using a single unified framework, there exist significant limitations with respect to its use in the lossless compression of bilevel imagery. In particular, substantial degradation in image quality at low resolutions severely limits the resolution scalable features of the JPEG2000 code-stream. We examine these effects and present two efficient methods to improve resolution scalability for bilevel imagery in JPEG2000. By analyzing the sequence of rounding operations performed in the JPEG2000 lossless compression pathway, we introduce a simple pixel assignment scheme that improves image quality for commonly occurring types of bilevel imagery. Additionally, we develop a more general strategy based on the JPIP protocol, which enables efficient interactive access of compressed bilevel imagery. It may be noted that both proposed methods are fully compliant with Part 1 of the JPEG2000 standard.**

***Index Terms*—Bilevel image compression, binary image compression, JPEG2000, JPIP, resolution scalability.**

I. Introduction

BILEVEL (or binary) images are often encountered in applications such as document archiving and retrieval, as well as digital libraries and facsimile, where they provide a compact means of representing black-and-white documents containing text and drawings. There exist a number of formats that specifically target the bilevel image compression task, such as the CCITT G3 and G4 fax standards, and the more recent JBIG and JBIG2 standards. The JPEG2000 standard for image compression is also capable of bilevel image compression; in fact, one of the desired features of the standard was the efficient compression of both bilevel and continuous tone image data, using a single unified compression architecture. To this end, the compression performance offered by JPEG2000 is very similar to the CCITT G4 standard [2]. Though the JBIG2 standard [8] demonstrates better compression performance for bilevel imagery, the library and archival communities have expressed interest in using JPEG2000 for their online documents due to the benefits associated with using a single compression system for all types of image data.

Manuscript received February 24, 2008; revised October 26, 2008. First published February 18, 2009; current version published March 13, 2009. This work was presented in part at the 2007 Data Compression Conference. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Bruno Carpentieri.

The authors are with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721-0104 USA (e-mail: rraguram@ece.arizona.edu; marcellin@ece.arizona.edu; bilgin@ece.arizona.edu).

Digital Object Identifier 10.1109/TIP.2008.2011455

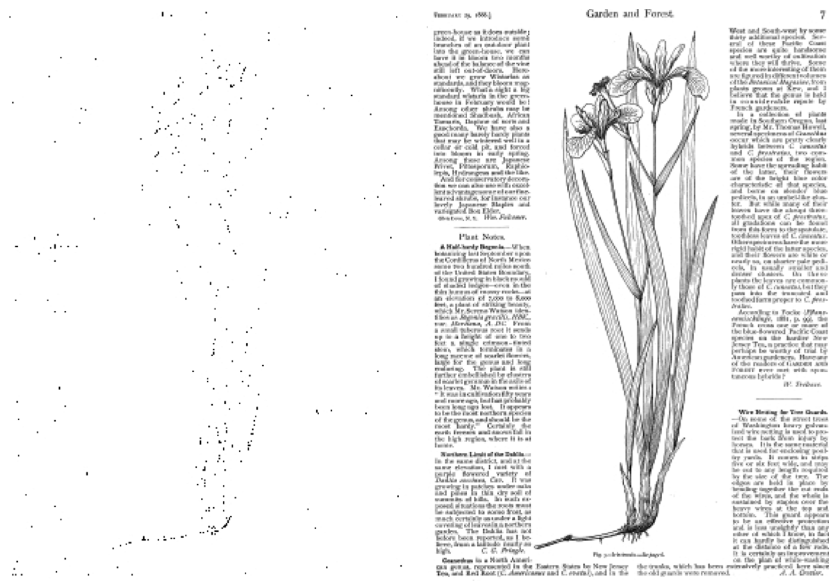
In fact, these communities have created a special interest group that meets regularly to discuss the use of JPEG2000 in library and archive applications. They have set up a web site [1] to post the minutes from these meetings, and to exchange information and experience with respect to using JPEG2000 for library and archive applications. One of the motivations for the work presented here came from our attendance at one of these special interest group meetings. At this meeting, a strong desire was expressed to use JPEG2000 for legacy binary imagery. As such imagery is often of very high resolution; the resolution scalability feature of JPEG2000 was noted to be of particular interest. It must be noted, however, that there exist certain limitations with regard to using JPEG2000 for the compression of bilevel image data. These limitations severely restrict the use of the resolution scalable features of the code-stream.

Scalability is one of the central concepts of the JPEG2000 paradigm [2], [3]. The JPEG2000 codec is transform based, and resolution scalability is a direct consequence of the multiresolution properties of the discrete wavelet transform (DWT). A code-stream is said to be resolution scalable if it contains identifiable subsets that represent successively lower resolution versions of the original image. Since bilevel images are invariably digitized at high resolutions, this property of the code-stream is potentially very useful. Consider the case where high resolution images are being viewed by a user over a network. Typically, the image at full resolution will be too large to display on the user's monitor. By making use of the inherent scalability of the JPEG2000 code-stream, it is possible to stream only the relevant portions of the image to the client. This allows JPEG2000 content to be delivered in a manner which matches the user's display resolution.

However, for bilevel imagery, the visual quality at lower resolutions can be too poor to be of any practical use. In the following sections, we analyze the issues concerning bilevel image compression in JPEG2000 and identify two methods which may be used to improve image quality at low resolutions, thereby enabling efficient resolution scalable delivery of compressed bilevel image data. It may be noted that both of these methods maintain JPEG2000 Part 1 compliance [4]. To our knowledge, these are the first schemes that seek to optimize the JPEG2000 codec for bilevel imagery, while doing so in a Part 1 compliant fashion.

II. Bilevel Image Compression in JPEG2000

JPEG2000 can be used to efficiently code bilevel imagery, subject to suitable choices of the coding parameters. One commonly used rule of thumb while compressing bilevel imagery is the use of zero levels of DWT, in order to maximize raw coding



(a) Lowest resolution level from JPEG2000. (b) Subsampled version of original image.

Fig. 1. Poor visual quality of JPEG2000 compressed bilevel imagery.

TABLE I
JPEG2000 COMPRESSION RATIOS FOR BI-LEVEL IMAGES WITH VARYING NUMBERS OF TRANSFORM LEVELS

Image	Number of transform levels				
	0	1	2	3	4
garden2 (5088x7216)	15.2	9.1	7.6	7.2	7.1
garden3 (5088x7216)	4.0	2.7	2.6	2.5	2.5
000012 (7344x5388)	28.4	19.5	17.4	16.8	16.7
000014 (5728x7500)	7.1	5.0	4.6	4.5	4.5
000015 (11056x7492)	6.5	4.5	4.0	4.0	4.0

efficiency. In this case, the JPEG2000 block coder codes the binary valued image data using a single coding pass. However, while this results in good coding performance, resolution scalability is sacrificed since there is no multiresolution hierarchy.

In order to introduce resolution scalability, one or more levels of wavelet transform may be applied. Part 1 of the JPEG2000 standard allows two wavelet transforms, the 5/3 and 9/7, corresponding to lossless and lossy image compression, respectively. Due to practical and cultural considerations, lossless compression of bilevel imagery is of interest to the library and archive communities. Thus, we consider only the 5/3 transform here. The 5/3 transform is designed primarily for the efficient compression of continuous tone imagery. Consequently, the bit-depth expansion caused due to application of the 5/3 transform adversely impacts bilevel image compression performance. Table I shows the degradation in lossless compression efficiency as resolution scalability is introduced. This phenomenon has been previously reported in [2]. Results are reported for five different bilevel images.

In a remote browsing application, this loss in compression performance may be offset by the resolution scalable properties of the code-stream. In particular, although the compressed file size is larger, the client can now directly access only the intermediate resolution desired, which may effectively result in less data being transferred. Consider a high-resolution bilevel image compressed with zero levels of transform. Even if the client desires only a low-resolution version of this image, there are no lower resolutions directly available; thus, the image at full resolution must be transmitted to the client, where it may then be downsampled. In contrast, when the same high resolution image is compressed using multiple levels of transform, only the data for the resolution required by the client needs to be transferred, leading to more efficient transmission.

While the discussion above might indicate the desirability of resolution scalable delivery of bilevel images, there exists a significant obstacle to the use of JPEG2000 for this purpose. Specifically, use of the reversible 5/3 transform results in rapid degradation of image quality at decreasing resolutions. This effect is shown in Fig. 1(a). The high resolution "garden2" image (5088×7216) was compressed using four levels of transform, and the lowest available resolution, corresponding to 1/16th the original resolution, was obtained by decompressing the relevant portions of the code-stream. For comparison, Fig. 1(b) shows an image at the same resolution, but obtained by applying a low-pass averaging filter to the original high resolution image, followed by downsampling. It can be seen that the JPEG2000 image has lost all detail and is unrecognizable. This drastic loss in visual quality poses a serious obstacle to the resolution scalable transmission of bilevel imagery.

Resolution reduction schemes in standards such as JBIG are carefully matched to the bilevel image compression task. For instance, JBIG uses a template-based resolution reduction scheme, involving tables defining exception rules that aim to preserve edges and lines, as well as periodic and dither patterns [5]. Wavelet transforms can be inherently ill-suited for this task, since requirements such as smoothness and vanishing moments, which are considered to be desirable in a constructed wavelet basis, may not be relevant when applied to bilevel imagery. Furthermore, the rounding steps that are introduced in the 5/3 transform to ensure reversibility can cause significant damage to bilevel images during the encoding process.

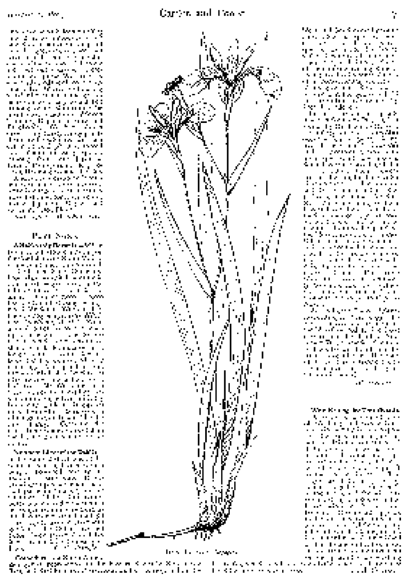


Fig. 2. Bilevel image compressed using the irreversible 5/3 transform.

In the following sections, we present two schemes that seek to overcome the above drawbacks in a JPEG2000 Part 1 compliant fashion.

III. Improving Resolution Scalability

A. Method 1: Observations Based on Rounding

Evidently, use of the reversible 5/3 transform for bilevel image compression causes substantial degradation in image quality at lower resolutions. Fig. 2 shows a low resolution bilevel image, this time compressed using an irreversible version of the 5/3 transform – i.e, without the rounding steps. By comparison with Fig. 1(a), it may be seen that for bilevel imagery, rounding results in a drastic loss of detail at low resolutions. It is, thus, worthwhile to analyze the effects of rounding in bilevel image compression.

In JPEG2000, both the reversible and irreversible transforms can be implemented using the lifting framework [6]. In a broad sense, lifting provides a means to generate invertible mappings between sequences of numbers, and the invertibility is unaffected even when arbitrary operators, which may be linear or nonlinear, are introduced in the lifting steps. This flexibility allows the use of nonlinear rounding operations in the lifting steps, in order to ensure that the transform coefficients are integers. The analysis equations for the reversible 5/3 transform, corresponding to the lifting realization, are presented below. We denote the input signal, low-pass subband signal, and high-pass subband signal by $x[n]$, $s[n]$ and $d[n]$, respectively. We also use $x_0[n] = x[2n]$ and $x_1[n] = x[2n+1]$ to represent the even and odd indexed samples of the input signal, respectively. We then have

$$d[n] = x_1[n] - \left\lfloor \frac{1}{2}(x_0[n] + x_0[n+1]) \right\rfloor \qquad (1)$$

and

$$s[n] = x_0[n] + \left\lfloor \frac{1}{4}(d[n] + d[n-1]) + \frac{1}{2} \right\rfloor. \qquad (2)$$

Lifting may be viewed as comprising three basic stages – split, predict, and update. In the split step, the input sequence is decomposed into its even and odd components, $x_0[n]$ and $x_1[n]$. In the next stage, the odd indexed coefficients are predicted using a combination of the neighboring even indexed coefficients. If $x[n]$ is smooth, then the predicted values will be close to the actual values; thus, a more compact representation may be obtained by replacing $x_1[n]$ by the prediction residual, $d[n]$. This sequence may be thought of as representing the extent to which the original signal fails to vary linearly with time. In terms of frequency content, these coefficients capture the high frequencies present in the original signal. In the update step, the even indexed coefficients are transformed into a low pass sequence $s[n]$, by updating $x_0[n]$ with a combination of the prediction residuals.

In view of the above interpretation, we note from (1) that $x_1[n]$ is predicted to be the average of its two neighboring samples. While the rounding operation in this step does not drastically affect grayscale imagery, it plays a much more pivotal role in the case of bilevel imagery, since it now involves making decisions between two extremes – black and white – as opposed to decisions between two neighboring gray levels.

It may be noted that while pixels in a bilevel image take on one of two values, 0 or 1, the interpretation of these values as black or white is left to the application. One common interpretation, in analogy with grayscale imagery, is to assign black the lowest value, or 0, and white the highest, or 1. All results presented above employ this convention. Alternatively, black pixels could be assigned a value of 1 and white pixels a value of 0. While examples of either convention can be found in the literature, we present an argument for using the latter interpretation when coding bilevel imagery using JPEG2000.

From (1), it may be seen that when one of the two even indexed samples $x_0[n]$ or $x_0[n+1]$ is 0 and the other 1, there exists an ambiguous case.[1] Since the lifting steps in the JPEG2000 standard employ a floor operator, we observe that for the ambiguous case, the predicted pixel value is always 0. For many instances of commonly occurring bilevel imagery, white pixels occur far more frequently than black pixels. Consequently, in the case where white pixels are assigned a value of 1, the value predicted in the ambiguous case will often be incorrect. One such instance is illustrated in Fig. 3, which shows a lifting step applied in the vertical direction along the third column. The two odd indexed samples $x_1[n-1]$ and $x_1[n]$ are predicted to be the average of their neighboring even indexed samples. It may be observed that both these pixels are incorrectly predicted to be 0, or black. Furthermore, since bilevel images invariably have a large number of edges, the ambiguous case is encountered often. When there are two successive errors in prediction, the prediction residuals $d[n-1]$ and $d[n]$ in (2) cause single pixel wide black lines to be "washed out" at lower resolutions.

A partial solution to this problem would be to use the alternate interpretation, with 0 for white and 1 for black. This serves to bias the prediction in favor of the more commonly

[1]Due to level shifting in JPEG2000, 1 is subtracted from each pixel value prior to the wavelet transform. Thus, the inputs to the transform are -1 and 0, rather than 0 and 1, respectively. For simplicity, we ignore this in our discussion without loss of validity for the cases discussed in this paper.

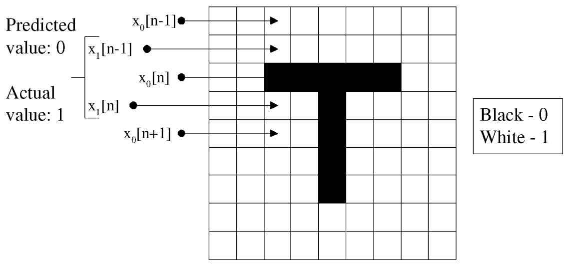


Fig. 3. Incorrectly predicted pixels in the lifting step.

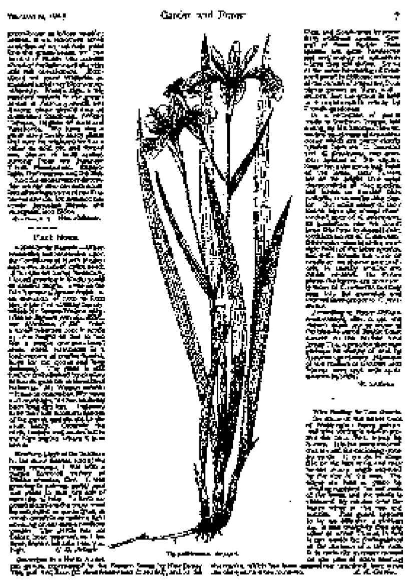


Fig. 4. Improved visual quality by assigning white – 0, black – 1.

occurring pixel value, which is now 0. It may be noted that for the former interpretation, the identical decompressed image could be obtained by replacing the floor operator in (1) with a ceiling operator (see the Appendix). Since modifying the nature of the lifting steps would make the encoder noncompliant with the standard, use of the alternate interpretation is preferable since it maintains JPEG2000 Part 1 compliance. For bilevel imagery with a white-1, black-0 assignment policy, this effect may also be achieved by marking the input bilevel image samples as "signed" data, with a bit-depth of 1 bit per sample. This alters the interpretation of the pixel values by JPEG2000, yielding an identical result. The results of this strategy are shown in Fig. 4. It may be seen that much more detail is retained, particularly in the textual regions, as compared to Fig. 1(a). As noted in [2] (Section 16.3), the use of 0 for white pixels and 1 for black pixels results in a loss in compression performance. This is because the procedure tends to decrease the length and frequency of insignificance runs and consequently, the efficiency of the JPEG2000 significance coding run mode. Even though lossless compression rates are decreased, the visual quality of the reduced resolution bilevel image improves considerably.

It must also be noted that the above discussion assumes that bilevel images are predominantly white, with fewer black foreground pixels. This is, indeed, true for a large class of bilevel imagery. In cases where this assumption is not true, it may be preferable to use 1 for white pixels, and 0 for black. This may be observed in Fig. 5, which compares the two cases for a bilevel image containing a large region of "halftone" material. Fig. 5(a) shows an image compressed using 0 for white and 1 for black, while Fig. 5(b) uses the opposite assignment. While Fig. 5(a) retains more detail in the textual region, the halftone region appears "blacked out." In contrast, it may be seen from Fig. 5(b) that the halftone region retains slightly more detail. Thus, a more effective strategy would be to design a scheme that adapts to the local nature of the bilevel image and assigns pixel values of 0 or 1 accordingly. We note that Part 6 of the JPEG2000 standard defines a Mixed Raster Content (MRC) model, that may be of use in this regard [7].

B. Method 2: Using the JPIP Protocol

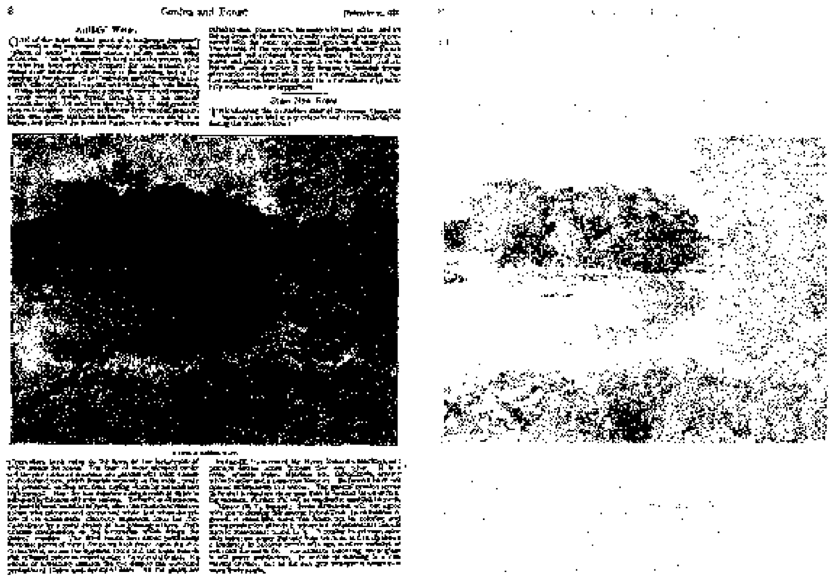
While the JPEG2000 standard offers many features that support interactive access of compressed imagery, Part 1 of the standard describes only the core coding system and syntax for the code-stream. While it is indeed possible for a client to interact remotely with image content by intelligently accessing appropriate byte ranges from the compressed file, the JPIP protocol seeks to standardize client/server interaction in an efficient and intelligent manner.

Using the JPIP protocol, the interaction between client and server is carried out through requests made by the client, which identify the current focus window of the client-side application. Rather than describing the focus window in terms of low-level code-stream constructs, these requests contain information regarding the client's spatial region of interest, resolution and image components of interest. The description of the focus window in terms of its geometric attributes is much more intuitive, and the server receives a representation of the end-user's ultimate interests, rather than a client's translation of those interests into JPEG2000 code-stream elements [9].

JPIP requests are composed of a sequence of "name = value" pairs. A basic JPIP request typically contains the name of the target file, together with a description of the focus window. When transmitted over a text-based transport protocol such as HTTP, the name and value fields are ASCII strings, with requests being separated by the "&" character. For instance, the following request refers to a file called images/garden2.jp2, at a resolution whose full size ($fsiz$) is 2544×3608 (columns $\times$ rows), which corresponds to half the resolution of the original image. The request is for a square region of 600×600 pixels ($rsiz$), which is located at an offset of 1000 pixels from the left and 1200 pixels from the top of the image (roff)

$$\begin{aligned}&\mathtt{target} = \mathtt{images/garden2.jp2\%fsiz} = 2544{,}3608\\ &\&\mathtt{rsiz} = 600{,}600\&\mathrm{roff} = 1000{,}1200.\end{aligned}$$

An important feature of the JPIP protocol is that image decompression and rendering are separated from client-server communication. JPIP specifies a means of interacting with JPEG2000 data, and mechanisms for communicating compressed image data and metadata between a client and server. However, JPIP does not specify how the client application should process/display this transmitted data. This enables us to build a client-side application that intelligently generates



(a) White - 0, Black - 1. (b) White - 1, Black - 0.

Fig. 5. Comparison of the two assignment policies.

requests for appropriate windows of interest and then processes the data received in order to improve the quality of the low resolution bilevel imagery. Details of the postprocessing operations are thus abstracted from the actual client-server communication process.

For our experiments, we used Kakadu v5.1 [10], which includes a JPIP client/server implementation compliant with the ISO JPIP final committee draft. The Kakadu image viewer initially attempts to render an image at a resolution that matches the display resolution of the user's terminal. The user can then pan and zoom as desired, and the image is re-rendered based on the data available in the cache along with further data received from the server, based on the new focus window parameters. It may be noted that in the case where the image is compressed with zero levels of transform, the image is displayed at full resolution. Even if the user desires only a low resolution image, the image at full resolution must be transferred to the client. This is a significant drawback for the interactive access of compressed bilevel imagery.

In order to design an improved scheme, we note the following points.

- If an image is compressed using D levels of transform, it is possible to request the image at any of the $D+1$ available resolutions. We refer to the low-pass subband LL_{D-r} as resolution r, where $r=0$ corresponds to the lowest available resolution and $r=D$ corresponds to the original image resolution. As seen in Section II, for bilevel imagery, the quality of the low resolution decompressed images can be far too poor to be of any practical use. However, our experiments indicate that for an assignment policy of white-1, black-0, image quality is usually acceptable for $D-2 \leq r \leq D$, but deteriorates rapidly when lower resolutions are viewed. This range was found to be stable over various categories of bilevel image content (text documents, line drawings, newspaper scans, etc.).
- Research on human perception has shown that if text is generated on a display with grayscale capability, then visual clarity may be improved by using shades of gray in rendered text [5]. The use of grayscale values forms an additional cue to the human visual system, which uses this extra information to compensate for the inaccurate visual data due to low resolution. In other words, the visual quality of a subsampled bilevel image may be enhanced by applying a suitable low-pass filter, retaining the intermediate gray-levels produced, and then downsampling by the required amount. This procedure is sometimes referred to as *scaling to gray*.
- Using the resolution scalable features of the JPEG2000 code-stream, we do not need to transfer the image at full resolution in order to scale down. For instance, to view the image at resolution r, we can request the server to send data corresponding to any resolution r', where $r < r' \leq D$, and then filter and downsample the image to the required resolution. Our experiments show that in order to obtain good image quality when viewing resolution r, where $0 \leq r < D-2$, we can request data corresponding to resolution $D-2$, and then scale-to-gray by the appropriate amount.
- Compressing bilevel imagery using $D > 0$ levels of wavelet transform results in an increase in file size over the zero level case (Refer to Table I). However, since we can access intermediate resolution levels, we may transfer smaller subsets of the compressed file. In order for the proposed scheme to be attractive, it must improve the quality of the bilevel image substantially while at the same time, the amount of data transferred must be significantly less compared to transferring the image at full resolution.

The proposed algorithm is formalized in Fig. 6. The block diagram of the scheme is shown in Fig. 7. It may be seen from Fig. 7 that the JPIP client is unaware of the postprocessing operations. In particular, the JPIP client knows only that data for resolution $D-2$ is being served. The task of modifying focus

While session is active
    Determine the client's window of interest in terms of focus box parameters. These parameters define the resolution of interest r, and spatial region of interest.
    if $r \geq D-2$
        Do not modify focus window parameters
    else if $r < D-2$
        Modify focus window parameters: (Refer to Figure 8)
            Scale the dimensions of the spatial region of interest
                $rsiz' = rsiz * 2^{(D-2-r)}$
                $roff' = roff * 2^{(D-2-r)}$
            Modify the resolution level parameters
                $fsiz' = fsiz * 2^{(D-2-r)}$
        Filter the received data using appropriate low-pass filters.
        Downsample by a factor of $2^{(D-2-r)}$
    end if
    Display
End while

Fig. 6. Algorithm for method 2.

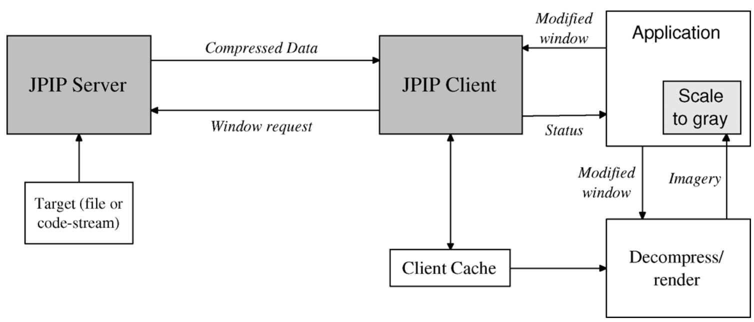


Fig. 7. Block diagram for the JPIP scheme.

window parameters and downsampling as required is carried out by the application.

The results obtained using the proposed method are shown in Fig. 9. The white-1, black-0 pixel assignment policy has been used. As can be seen, there is a dramatic improvement in the visual quality of the low-resolution decompressed images, as compared to Figs. 1(a) and 5(b), which employ the same assignment policy. Additionally, there is significant improvement even over Figs. 4 and 5(a), which employ the white-0, black-1 policy. It may also be noted that this method produces excellent quality for both the text and halftone regions of a compound bilevel image. Our experiments indicate that a simple averaging filter produces results of good quality. If desired, more sophisticated low-pass filters that possess good properties for downsampling, may be used. One such filter, commonly used in the graphics community, is the Lanczos filter [11], which produces slightly smoother images.

Table II shows the data savings for the above method. Since the proposed scheme uses data from resolution $D-2$ to generate all lower resolutions, only two levels of wavelet transform are required. The amount of data transfer for the JPIP scheme is listed in the third column of Table II. For comparison, the second column lists the amount of data required to transfer the image at full resolution D, and then downsample. It may be observed that the proposed scheme needs, on average, 85% less data in order to achieve comparable image quality.

The results shown in Table II demonstrate data savings using the white-1, black-0 assignment policy. Since the alternate interpretation of white-0, black-1 causes more detail to be retained at lower resolutions, the JPIP scheme may benefit from this assignment policy. Our experiments indicate that when viewing any resolution r, where $0 \leq r < D-2$, it is sufficient to request data corresponding to resolution $r+1$ (as opposed to resolution $D-2$). At low resolutions, this could result in increased data savings.

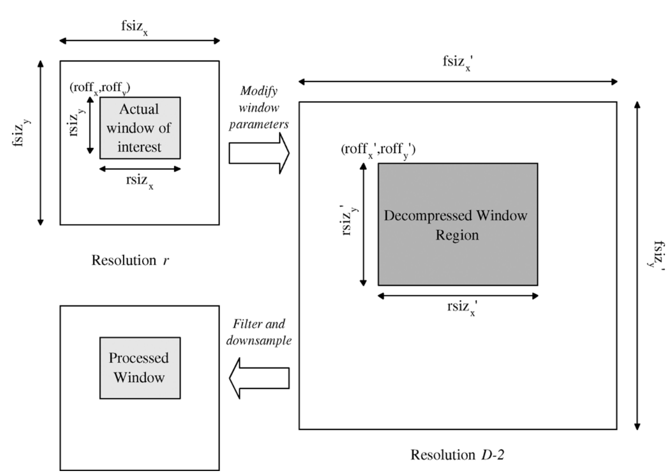
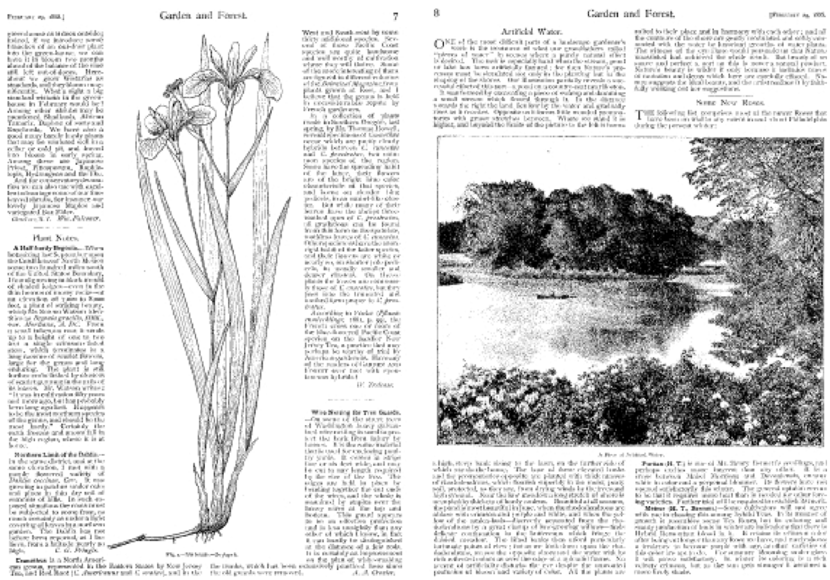


Fig. 8. Modifying the parameters of the window of interest.

Results for the JPIP scheme using a white-0, black-1 assignment policy were obtained as follows. Bilevel images were compressed with 5 levels of transform. For resolutions $0 \leq r < 3$,


Garden and Forest.

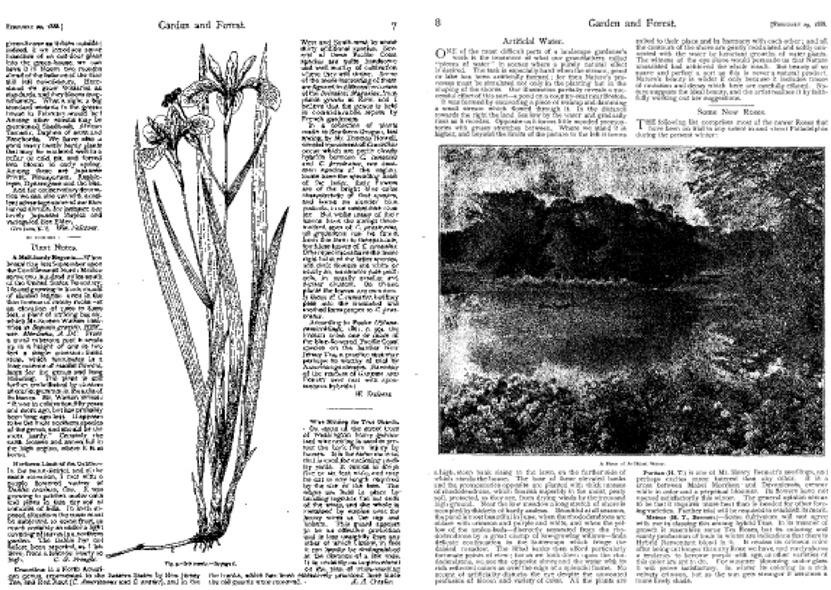
Artificial Water.

Some New Roses.

Fig. 9. Improved low-resolution images obtained using the JPIP scheme.

TABLE II
SAVINGS IN DATA TRANSFERRED FOR THE JPIP SCHEME

Image	Amount of data transferred for full resolution image (KB)	Amount of data transferred for the proposed scheme (KB)	Percentage savings
garden2 (5088x7216)	294.17	59.08	79.92
garden3 (5088x7216)	1115.3	119.90	89.25
000012 (7344x5388)	169.90	22.25	86.90
000014 (5728x7500)	737.22	68.35	90.73
000015 (11056x7492)	1548.5	135.67	91.24


Garden and Forest.

Artificial Water.

Some New Roses.

Fig. 10. Images obtained using the JPIP scheme with a white-0, black-1 assignment.

data corresponding to resolution $r+1$ was requested and scaled to gray. The data savings obtained (with respect to transferring the image at full resolution) are presented in Table III and image results are shown in Fig. 10. By comparing Figs. 9 and 10, it may be observed that the JPIP scheme with a white-0, black-1 assignment policy produces images of comparable quality. It must be noted that the white-0, black-1 policy is advantageous only at lower resolutions (specifically, for $r < D - 3$). This is illustrated in Fig. 11, which compares data transfer results for the two versions of the JPIP scheme. For reference, the graph also shows data transfer results for the two assignment policies with zero levels of wavelet transform, which corresponds to transferring the image at full resolution to the client. The data transfer results in the graph are averages over five test images.

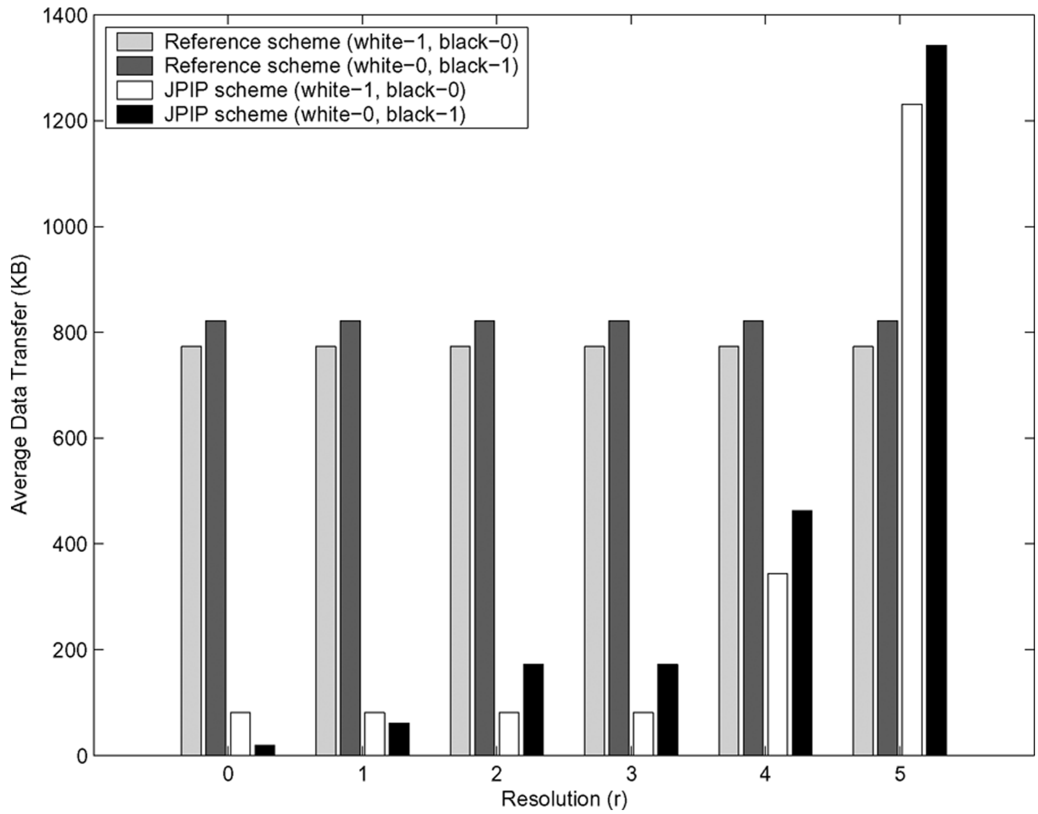


Fig. 11. Comparison of the two versions of the JPIP scheme.

TABLE III
DATA TRANSFER RESULTS AT VARIOUS RESOLUTIONS FOR THE JPIP SCHEME, USING WHITE-0, BLACK-1. ($D = 5$ LEVELS)

Image	Amount of data transferred for full resolution image (KB)	$r = 2$		$r = 1$		$r = 0$	
		Data transfer	Percentage savings	Data transfer	Percentage savings	Data transfer	Percentage savings
garden2 (5088x7216)	352.48	140.93	60.02	53.38	84.85	16.97	95.18
garden3 (5088x7216)	1138.22	202.54	82.21	63.03	94.46	17.4	98.47
000012 (7344x5388)	216.56	50.22	76.81	19.89	90.81	7.37	96.60
000014 (5728x7500)	772.76	147.81	80.87	54.04	93.01	17.68	97.71
000015 (11056x7492)	1628.63	318.44	80.45	113.51	93.03	36.4	97.76

It may be observed from the graph that when the image at full resolution ($r = 5$) is being viewed, both JPIP schemes perform poorly, since use of the wavelet transform results in an increase in file size. However, at all lower resolutions, the JPIP schemes provide significant data savings. Up to resolution $D - 3$ (corresponding to $r = 2$ in the graph), both versions of the JPIP scheme fetch identical resolutions from the server in order to scale-to-gray. However, since the use of white-0, black-1 results in a drop in compression performance, this scheme requires slightly more data. For all lower resolutions, $r < D-3$, the JPIP scheme with a white-0, black-1 assignment provides increased data savings, since it only accesses one resolution higher than the resolution of interest.

IV. CONCLUSION

This paper introduces two efficient schemes in order to improve resolution scalability for bilevel imagery in JPEG2000. The first method suggests the use of a particular black/white assignment policy in order to improve the quality of the low-resolution image, and works well for certain commonly occurring types of bilevel imagery. The second approach employs the JPIP protocol, and produces images that are comparable to those that result from downsampling the full resolution image, but requires only 15% of the data. The JPIP scheme may be further modified by using the alternate black/white assignment policy, which provides increased data savings at low resolutions (for $r < D-3$). The proposed schemes can be implemented in a fully JPEG2000 Part 1 compliant fashion.

APPENDIX

In Section III-A, we noted that replacing the floor operator in (1) and (2) with a ceiling operator results in improved image quality at low resolutions, for certain types of bilevel imagery. However, since this would make the encoder noncompliant with the standard, we make use of an alternate black/white assignment policy. In this Appendix, we show that the two strategies are equivalent.

Replacing the floor operator in (1) and (2) with a ceiling operator, we have the following expressions for the high- and low-pass sequences

$$d_{\mathrm{ceil}}[n] = x_1[n] - \left\lceil \frac{1}{2}\big(x_0[n] + x_0[n+1]\big) \right\rceil .$$

For simplicity, we define $a[n] = (1/2)(x_0[n] + x_0[n+1])$. We then have

$$d_{\mathrm{ceil}}[n] = x_1[n] - \lceil a[n] \rceil . \tag{3}$$

Similarly

$$s_{\mathrm{ceil}}[n] = x_0[n] + \left\lceil \frac{1}{2} + \frac{1}{4}\big(d_{\mathrm{ceil}}[n-1] + d_{\mathrm{ceil}}[n]\big) \right\rceil . \tag{4}$$

For bilevel imagery, use of the alternate assignment policy may be seen as replacing $x[n]$ by its "complement," $(1 - x[n])$. Thus, for the alternate assignment, we have

$$\begin{aligned} d_{alt}[n] &= \big(1 - x_1[n]\big) - \left\lfloor \frac{1}{2}\big(1 - x_0[n] + 1 - x_0[n+1]\big) \right\rfloor \\ &= \big(1 - x_1[n]\big) - \left\lfloor 1 - \frac{1}{2}\big(x_0[n] + x_0[n+1]\big) \right\rfloor \\ &= \big(1 - x_1[n]\big) - 1 - \lfloor -a[n] \rfloor \\ &= -x_1[n] + \lceil a[n] \rceil = -d_{\mathrm{ceil}}[n] \end{aligned} \tag{5}$$

and

$$\begin{aligned} s_{alt}[n] &= \big(1 - x_0[n]\big) + \left\lfloor \frac{1}{2} + \frac{1}{4}\big(d_{alt}[n-1] + d_{alt}[n]\big) \right\rfloor \\ &= \big(1 - x_0[n]\big) + \left\lfloor \frac{1}{2} - \frac{1}{4}\big(d_{\mathrm{ceil}}[n-1] + d_{\mathrm{ceil}}[n]\big) \right\rfloor \\ &= \big(1 - x_0[n]\big) + 1 \\ &\quad + \left\lfloor -1 + \frac{1}{2} - \frac{1}{4}\big(d_{\mathrm{ceil}}[n-1] + d_{\mathrm{ceil}}[n]\big) \right\rfloor \\ &= \big(2 - x_0[n]\big) + \left\lfloor -\frac{1}{2} - \frac{1}{4}\big(d_{\mathrm{ceil}}[n-1] + d_{\mathrm{ceil}}[n]\big) \right\rfloor \\ &= \big(2 - x_0[n]\big) - \left\lceil \frac{1}{2} + \frac{1}{4}\big(d_{\mathrm{ceil}}[n-1] + d_{\mathrm{ceil}}[n]\big) \right\rceil \\ &= 2 - s_{\mathrm{ceil}}[n]. \end{aligned} \tag{6}$$

It may be observed from (4) that when the input sequence $x[n]$ is bilevel, $s_{\mathrm{ceil}}[n]$ takes values in $\{0, 1, 2\}$. Thus, we note from (6) that the low-pass sequence $s_{alt}[n]$ is equivalent to "complementing" the sequence $s_{\mathrm{ceil}}[n]$. Hence, when decompressed and displayed, the two representations will be identical.

REFERENCES

[1] [Online]. Available: http://j2karclib.info

[2] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards, and Practice*. Boston, MA: Kluwer, 2002.

[3] D. S. Taubman and M. W. Marcellin, "JPEG2000: Standard for interactive imaging," *Proc. IEEE*, vol. 90, pp. 1336–1357, 2002.

[4] Information Technology–JPEG2000–Image Coding System–Part 1: Core Coding System 2000, ISO/IEC 15444-1.

[5] I. H. Witten, A. Moffat, and T. C. Bell, *Managing Gigabytes*. San Mateo, CA: Morgan Kaufmann, 1999.

[6] W. Sweldens, A. F. Laine and M. Unser, Eds., "The lifting scheme: A new philosophy in biorthogonal wavelet constructions," in *Proc. SPIE Wavelet Applications in Signal and Image Processing III*, 1995, vol. 2569, pp. 68–79.

[7] Information Technology–JPEG2000–Image Coding System–Part 6: Compound Image File Format 2003, ISO/IEC 15444-6.

[8] Information Technology – Coded Representation of Picture and Audio Information – Lossy/Lossless Coding of Bi-Level Images, 1999, ISO/IEC 14492.

[9] D. Taubman and R. Prandolini, "Architecture, philosophy and performance of JPIP: Internet protocol standard for JPEG2000," presented at the Int. Symp. Visual Communications and Image Processing, 2003.

[10] [Online]. Available: http://www.kakadusoftware.com

[11] C. E. Duchon, "Lanczos filtering in one and two dimensions," *J. Appl. Meteorol.*, vol. 18, pp. 1016–1022, Aug. 1979.

Rahul Raguram (M'05) received the B.S. degree in electronics and communication engineering from Visveswaraiah Technological University, Bangalore, India, in 2004, and the M.S. degree in electrical engineering from the University of Arizona, Tucson, in 2007. He is currently pursuing the Ph.D. degree in computer science from the University of North Carolina, Chapel Hill.

His research interests include computer vision, multiple-view geometry, robust estimation techniques, image and video coding, and image processing.

Michael W. Marcellin (S'81–M'87–SM'93–F'02) was born in Bishop, CA, on July 1, 1959. He graduated *summa cum laude* with a B.S. degree in electrical engineering from San Diego State University, San Diego, CA, in 1983, where he was named the most outstanding student in the College of Engineering. He received the M.S. and Ph.D. degrees in electrical engineering from Texas A&M University in 1985 and 1987, respectively.

Since 1988, he has been with the University of Arizona, Tucson, where he holds the title of Regents' Professor of Electrical and Computer Engineering and of Optical Sciences. His research interests include digital communication and data storage systems, data compression, and signal processing. He has authored or coauthored more than 200 publications in these areas. He is a major contributor to JPEG2000, the emerging second-generation standard for image compression. Throughout the standardization process, he chaired the JPEG2000 Verification Model Ad Hoc Group, which was responsible for the software implementation and documentation of the JPEG2000 algorithm. He is coauthor of the book *JPEG2000: Image Compression Fundamentals, Standards and Practice* (Kluwer, 2002). This book is intended to serve as a graduate level textbook on image compression fundamentals, as well as the definitive reference on JPEG2000. He served as a consultant to Digital Cinema Initiatives (DCI), a consortium of Hollywood studios, on the development of the JPEG2000 profiles for digital cinema.

Prof. Marcellin is a member of Tau Beta Pi, Eta Kappa Nu, and Phi Kappa Phi. He is a 1992 recipient of the National Science Foundation Young Investigator Award and a corecipient of the 1993 IEEE Signal Processing Society Senior (Best Paper) Award. He has received teaching awards from NTU (1990, 2001), IEEE/Eta Kappa Nu student sections (1997), and the University of Arizona College of Engineering (2000). In 2003, he was named the San Diego State University Distinguished Engineering Alumnus. He is the recipient of the 2006 University of Arizona Technology Innovation Award. From 2001 to 2006, he was the Litton Industries John M. Leonis Professor of Engineering. He is currently the International Foundation for Telemetering Professor of Electrical and Computer Engineering at the University of Arizona.

Ali Bilgin (S'94–M'03–SM'08) received the B.S. degree in electronics and telecommunications engineering from Istanbul Technical University, Istanbul, Turkey, the M.S. degree in electrical engineering from San Diego State University, San Diego, CA, and the Ph.D. degree in electrical engineering from the University of Arizona, Tucson, AZ.

Dr. Bilgin is currently a Research Assistant Professor with the Department of Electrical and Computer Engineering and the Department of Radiology, University of Arizona. His current research interests are in the areas of signal and image processing and include image and video coding, data compression, and magnetic resonance imaging. He has coauthored over 100 research papers in these areas