Interactive Progressive Encoding System For Transmission of Complex Images

Borko Furht[1], Yingli Wang[1], and Joe Celli[2]
[1]NSF Multimedia Laboratory
Florida Atlantic University, Boca Raton, Florida 33431
[2]IBM, Boca Raton, Florida

ABSTRACT

In this paper, we describe an interactive progressive JPEG-based encoding technique, which is suitable for an efficient network transmission of complex images. This system is intended for those applications that require a fast transmission of complex, high-quality, high-resolution images over networks with a limited communication bandwidth. These applications include transmission of medical images, space and earth exploration applications, as well as Internet applications. The proposed technique can be also applied to image archive and browsing systems.

The Interactive Progressive Encoding System (IPES) operates as follows. Users submit requests for imagery to the image database via a graphical user interface. Images in the database are stored in compressed form using standard JPEG compression technique using quantization tables which provide extremely high quality and low compression ratios. Upon initial request, a DCT image, which is a version of the image based only on DC coefficients, is transmitted and reconstructed at the user site. The transmission of a DCT image, even for very complex images, will take a relatively short time. The user can then isolate specific regions of interests within the image and request additional levels of details or additional scans. In each scan, additional DCT coefficients are transmitted which will improve the quality of the selected region of the image. If all levels of detail are sent, the transmitted image is visually indistinguishable from the original.

We developed a prototype of the IPES, which uses the spectral selection progressive encoding. An image database has been created, consisting of JPEG-compressed images, and the system has been tested and evaluated. The obtained results, presented and analyzed in the paper, show that the proposed interactive progressive encoding system is very efficient for transmission of complex, high-quality images over low-bandwidth communication channels.

Keywords: interactive progressive encoding, JPEG image compression, DCT coefficients, spectral selection

1. INTRODUCTION

The sequential JPEG image compression standard provides relatively high compression ratios (1:10 to 1:15), while maintaining high quality of the image [1,2,3,4]. However, the main drawback of the JPEG technique for certain applications that include image transmission, is that it may take long time to receive and display the image. Therefore, the progressive JPEG technique is developed that provides a first sketch of the image in a relatively short period of time. Then, the subsequent scans of the image can be requested by the user, which will improve the quality of the image. However, even the progressive JPEG will require long transmission times for high-resolution, complex images or for cases when the available communication bandwidth at the receiver's side is low.

We extended the conventional progressive JPEG to the interactive progressive JPEG encoding and transmission system, which is capable to reduce the transmission time for those applications, which do not require the complete image to be transmitted with the same quality. In these applications, the user can interactively select at any time regions of interests as well as the acceptable quality of these regions. Only the coefficients that correspond to the selected regions will be transmitted from the server to the user, which can significantly reduce the transmission time.

We developed the prototype interactive progressive encoding system, based on spectral selection progressive JPEG, and applied it to the JPEG-compressed image database. The experimental results are presented in the paper.

2. THE PROGRESSIVE JPEG TECHNIQUE

In a sequential JPEG-based encoding system, an image is encoded in a single left-to-right, top-to-bottom scan. The JPEG encoder consists of a Forward Discrete Cosine Transform (FDCT) block, quantizer, and entropy encoder. In the JPEG sequential decoding, all steps from the encoding process are inversed and implemented in reverse order, as shown in block diagram in Figure 1.

Figure 1. Block diagram of the sequential JPEG codec.

The image is first divided blocks consisting of 8x8 pixels, and then the JPEG algorithm is implemented on each block. The FDCT transforms the image pixels into the frequency domain using the following equations:

$$F(u,v) = \frac{C(u)}{2} \times \frac{C(v)}{2} \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos\frac{(2x+1)up}{16} \cos\frac{(2y+1)vp}{16}$$

where

$$C(u) = \frac{1}{\sqrt{2}} \quad for\ u = 0 \qquad C(u) = 1 \ \ for\ u > 0$$

$$C(v) = \frac{1}{\sqrt{2}} \quad for\ v = 0 \qquad C(v) = 1 \ \ for\ v > 0$$

The transformed 64-point discrete signal is a function of two spatial dimensions x and y, and its components are called spatial frequencies or DCT coefficients. The F(0,0) is the DC coefficient, and the remaining 63 coefficients are the AC coefficients.

For a typical 8x8 image block, most of spatial frequencies have zero or near-zero values, and need not to be encoded. In the next step, all 64 DCT coefficients are quantized using a 64-element quantization table, specified by the application. The quantization reduces the amplitude of the coefficients that contribute little or nothing to the quality of the image. The objective of the quantization is to increase the number of zero-value coefficients and to discard information that is not visually significant. The quantization is performed using the following equation:

$$F_q(u,v) = Round\left[\frac{F(u,v)}{Q(u,v)}\right]$$

where Q(u,v) are quantization coefficients specified by the quantization table. Each element Q(u,v) is an integer from 1to 255, which specifies the step size of the quantizer for its corresponding DCT coefficient.

After quantization, the 63 AC ceoefficients are ordered into the “zig-zag” sequence in order to place low-frequency coefficients, which are more likely to be non-zero, before high-frequency coefficients. The DC coefficients, which represent the average values of 64 image samples, are coded using predictive coding.

In the last step, the entropy encoder – typically Huffman encoder – is applied in order to convert the DCT coefficients after quantization into a compact binary sequence.

The progressive JPEG encoder produces a sequence of scans, each scan coding a subset of DCT coefficients. Therefore, the progressive JPEG encoder must have an additional buffer at the output of the quantizer and before the entropy encoder. The size of the buffer should be large enough to store all DCT coefficients of the image, each of each is 3 bits larger than the original image samples.

Progressive JPEG compression can be achieved using three algorithms: (a) Progressive spectral selection algorithm, (b) Progressive successive approximation algorithm, and (3) Combined progressive algorithm [1,2].

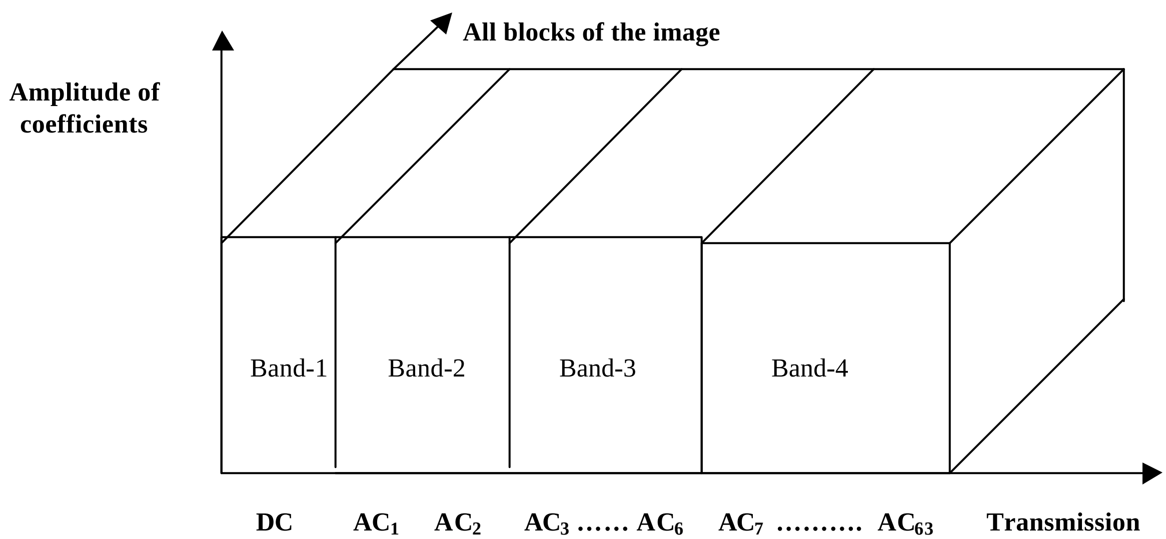

Figure 2. Band transmission in progressive JPEG based on spectral selection.
In each scan, DCT coefficients from the selected band from all 8x8 blocks are transmitted.

In our case, the IPES has been implemented using the spectral selection algorithm. In this algorithm, the DCT coefficients are grouped into several spectral bands. Typically, low-frequency bands are sent first, and then high-frequency bands. The IPES uses four spectral bands:

Band 1: DC coefficient only
Band 2: AC_1 and AC_2 coefficients
Band 3: AC_3, AC_4, AC_5, and AC_6 coefficients
Band 4: AC_7,…AC_{63} coefficients.

Figure 2 illustrates transmission of these four bands in the progressive JPEG system based on spectral selection. Note that in each scan, the DCT coefficients from the selected band of all 8x8 blocks are transmitted.

Interactive Progressive JPEG

In an interactive progressive JPEG system, the first scan, consisting of the DC coefficients only, is the same as in the conventional progressive JPEG system. However, the user is involved in selecting one or a few regions of the image, which will be transmitted in the subsequent scans. In another words, only the DCT coefficients from the blocks that belong to the selected region will be transmitted in the subsequent scans, as illustrated in Figure 3.

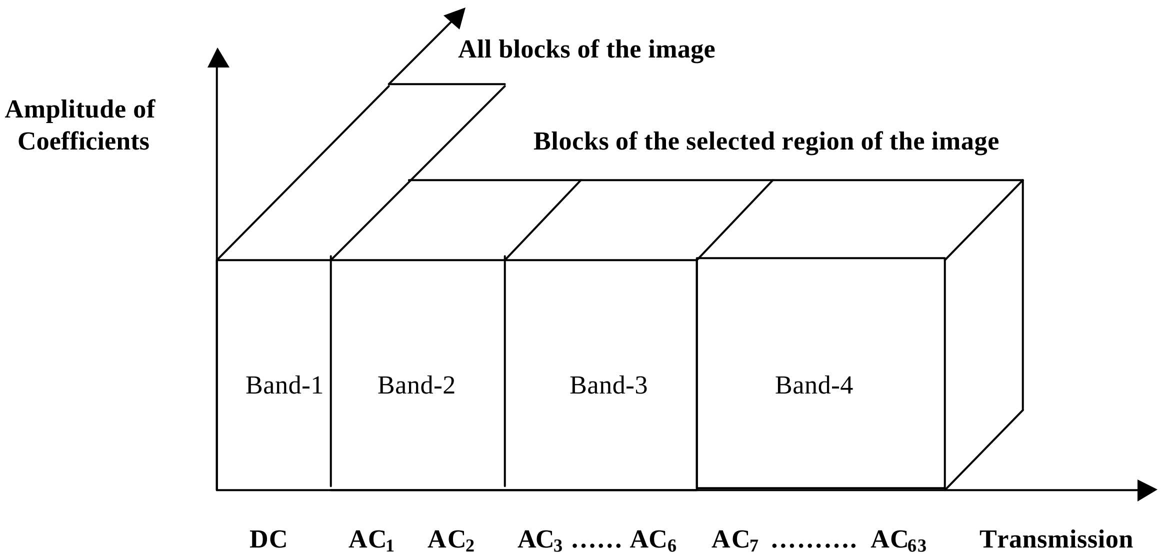

Figure 3. Band transmission in interactive progressive JPEG based on spectral selection. In subsequent scans, only DCT coefficients from the blocks that belong to the selected region will be transmitted within the band.

3. THE IPES PROTOTYPE SYSTEM AND EXPERIMENTAL RESULTS

We developed the prototype interactive progressive encoding and transmission system, which is using spectral selection technique. The system operates in two modes: (a) the encoding mode, and (b) the retrieval and decoding mode. In the encoding mode, typically applied on the server, the IPES compresses a given image and stores it in the image database. The user can select two quality factors, 1 and 2 (two quantization tables), as shown in Figure 4. The quality factor equal to 1 provides better image quality and lower compression ratio than the quality 2. Figure 4 shows the original image, which is compressed at the server using the quality factor = 1.

In the retrieval and decoding mode, the user selects an image from the database by scrolling through a small window (see Figure 4). When an image is selected, the user then presses the “SCAN” button, which will send the corresponding request to the image database at the server. The first band of the selected image (DC coefficients only) will be transmitted from the server, and the DC image will be created and presented to the user (in the small window on the lower left-hand side of the screen, see Figure 5a). The user may now select another image from the database, if the previously selected image is a wrong one, or he/she can continue with receiving this image.

By pressing the "SCAN" button again, the DC image will be presented at the large window on the screen. Now the user can request additional DCT coefficients for the whole image, or just for a selected region of the image. In Figure 5a, only a small region of the image is selected for subsequent transmission of additional coefficients. After the user selects the desired region of the image and presses the "SCAN" button, this request is sent to the server, and the system selects the next band of DCT coefficients for transmission. This time, only the coefficients from the selected region will be transmitted. The new image, shown in Figure 5b, is presented to the user with the improved quality of the selected region.

The user may now stop with the image transmission, if the quality of the region is acceptable, or may select another region from the same image, or may select another image from the database. The user can also request another scan for the same region in order to improve the reconstructed image quality. This is illustrated in Figures 5c and 5d, where the user requested two additional scans for the selected region. Note that after each scan, the number of transmitted bits is registered on the top bar on the screen. Finally, in Figure 5d, the reconstructed region of the image is practically indistinguishable from the original, while the rest of the image represents the DC image, consisting of DC coefficients only.

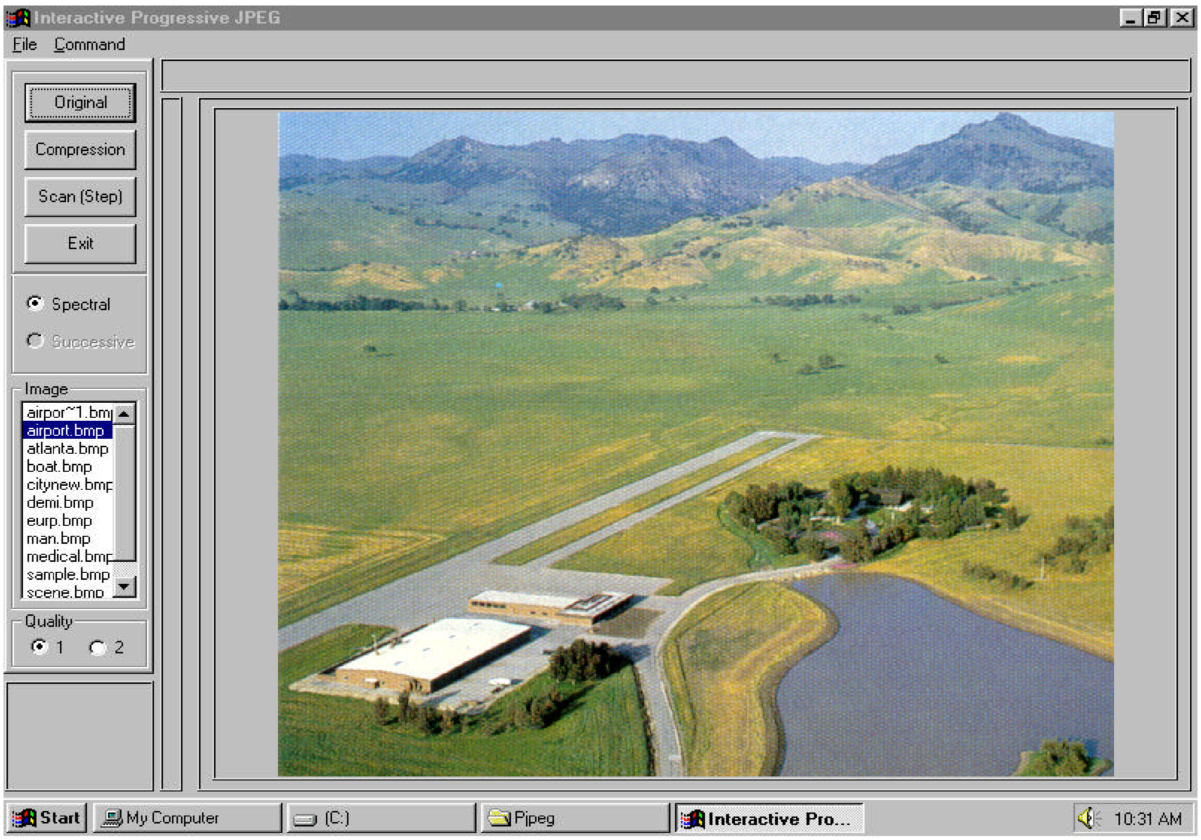

Figure 4. The IPES user interface. The original image "Airport" is selected from the database and progressive JPEG compression is applied.

Experimental Results

To illustrate the efficiency of the IPES system, the image "Airport," shown in Figures 4 and 5, is used. The original color image consists of 480x560 pixels and requires 806.4 KB (or 6.45 Mbits) for its storage. First, the image is compressed with the quality factor=1 and stored in the database. Its size in compressed form is 350 KB (or 2.8 Mbits); the compression ratio has been 2.3.

The following three experiments were performed. In the first experiment, all four scans for the whole image were requested, and the complete image was reconstructed. The image was indistinguishable from the original. The total number of bits transferred in four scans from the server to the client was 2.54 Mbits.

In the second experiment, the region shown in Figure 5a ("building") was selected for the scans 2 to 4. The final reconstructed image is shown in Figure 5d. The total number of bits transferred was 495 Kbits.

Finally, in the third experiment we selected two regions for the scans 2 to 4 – "building" and "ramp," as shown in Figure 6. The total number of bits transferred for both regions was 698 Kbits. The complete results of these three experiments are presented in Table 1 and Figure 7. Table 1 reports the number of bits transmitted in each scan, while Figure 7 shows the cumulative number of transmitted bits for these three cases.

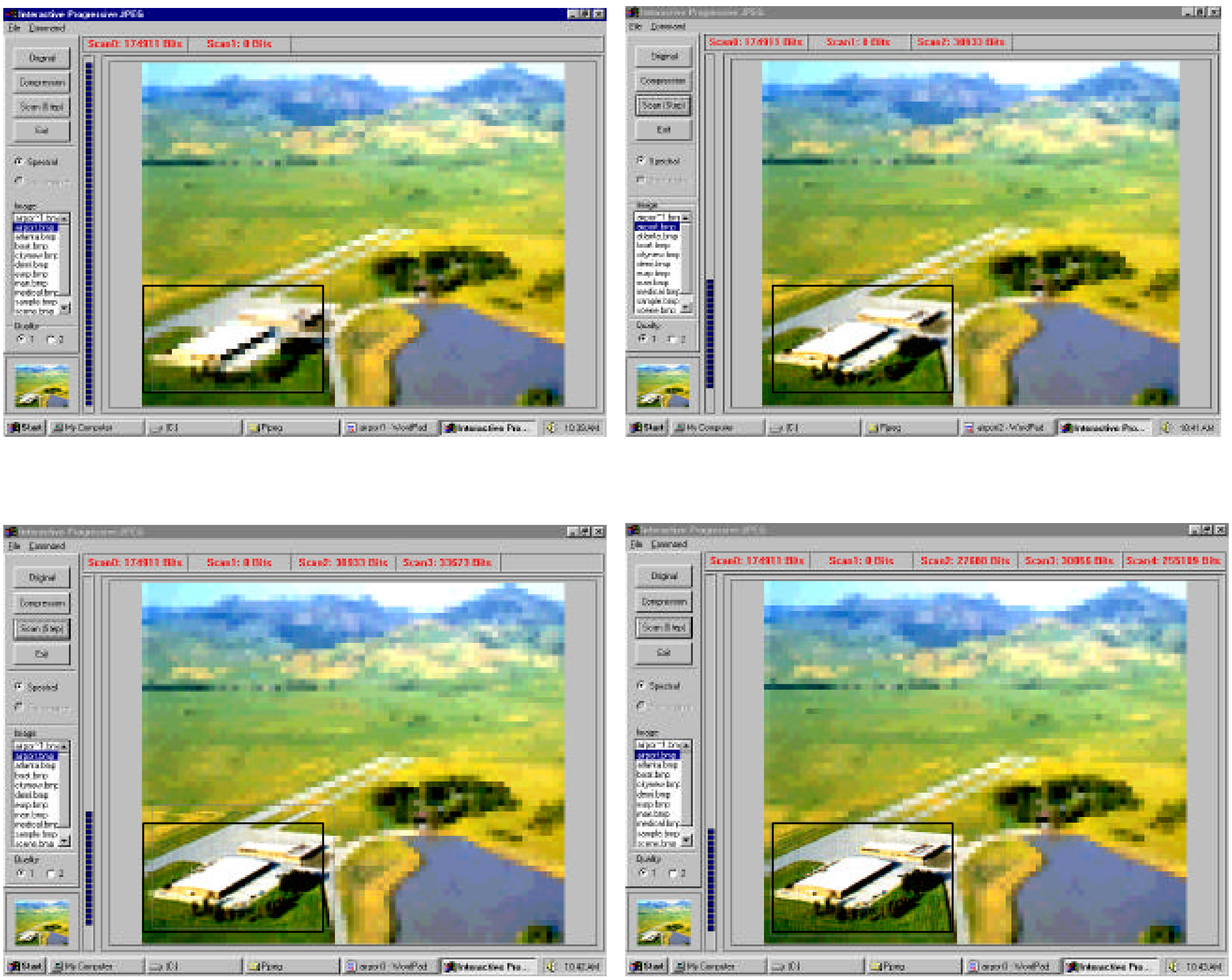

Figure 5. The region referred as to "building" is selected and four scans were implemented:
(a) scan 1 - the DC image, (b) scan 2, (c) scan 3, and (d) scan 4.

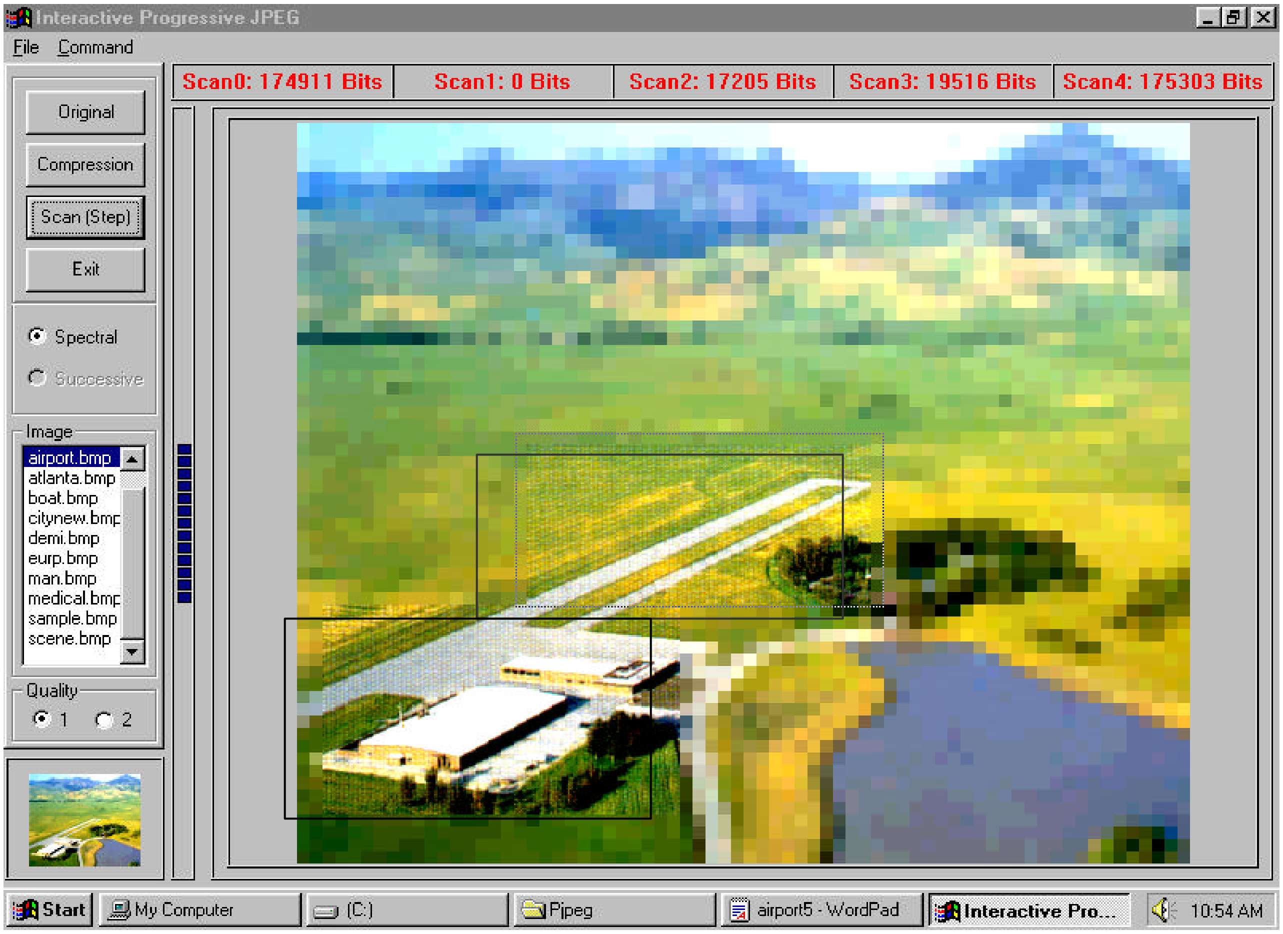

Figure 6. Two regions, referred to as "building" and "ramp," were selected for the subsequent scans 2 to 4. The image shown is reconstructed after the 4th scan.

Table 1. Results of the experiments for the interactive progressive encoding and transmission system.

SCAN	COMPLETE IMAGE [KBITS]	REGION "BUILDING" [KBITS]	TWO REGIONS "BULIDING" & "RUNAWAY" [KBITS]
1	175	175	175
2	190	28	36
3	215	31	51
4	1964	261	436
TOTAL	**2544**	**495**	**698**

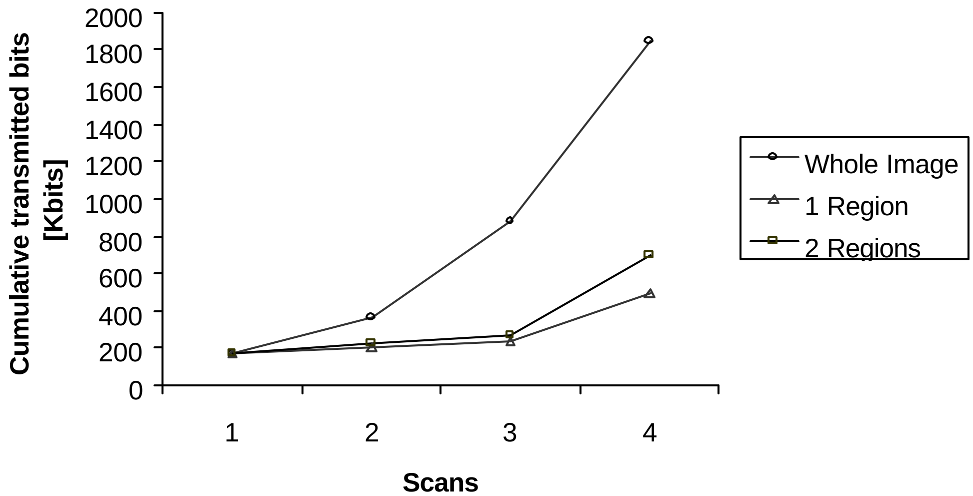

Figure 7. Cumulative number of transmitted bits for three cases: (a) the whole image was transmitted, (b) only one region was transmitted, and (3) two regions were transmitted.

4. APPLICATIONS

Two major applications for the proposed interactive progressive image encoding and transmission system are [5]:

- Retrieval and transmission of complex images over low bandwidth communication channels, and
- Archiving and browsing of visually lossless image databases.

We discuss both applications next.

Retrieval and Transmission of Complex Images Over Low Bandwidth Communication Channels

In these kind of applications, we assume that the user is at the end of a low bandwidth communication link, which could be as low as 2.4 Kbps, and he/she requires to receive an image in a short period of time. In certain cases, the quality of the image is not as important as the transmission time. In addition, in such applications the user may only be interested in a section of a complex image. Possible applications include image transmission over the Internet, real-time retrieval and transmission of medical images, and many others.

As an example, assume a color image with 800x800 pixels at 24 bits per pixel. In uncompressed form, the image requires 1.92 MB for storage. If the image is transmitted over a low bandwidth communication link, say 28.8 Kbps, it will require about 533 seconds or about 9 minutes to receive the complete image on the other end. Even if the image is compressed with a low compression factor of 1:2 or 1:3 (or 2.5 to 3.0 bits per pixel), the transmission time will be between 3 to 5 minutes, which is unacceptable in many applications. In some cases, the user is even not sure which image he/she is looking for. Therefore, it may happen that a wrong image is received, and another request must be sent for another image.

The proposed interactive progressive encoding and transmission system will resolve this problem in the following way. All images in the database are JPEG-encoded. The user interactively selects an image from the database, and this image is then transmitted over the communication link in several scans. In the first scan, only DC coefficients of each 8x8 block are transmitted, and a DC image is created on the other end. Because there are only a very few coefficients, the time for transmission of the DC coefficients is relatively short, and the user can see the first image almost instantly. If the wring image is transmitted, the user can immediately stop the transmission and request another image. If the image is correct, the user can request another level of detail either for the whole image or for any image region. In the latter case, the appropriate set of DC coefficients is extracted from the compressed image file by the server and transmitted to the user. The user can then stop transmission, if he/she is satisfied with the quality of the reconstructed image and its regions, or he/she can continue requesting additional scans until the desired level of fidelity is achieved in all image regions of interest. In this way, the total transmission time of particular regions of an image can be drastically reduced to less than 1 minute. The emphasis in these applications is a reduction of the time to transmit and receive a useful image or its subsections.

Archiving and Browsing Visually Lossless Image Databases

In these applications, we assume that relatively complex images in a database are JPEG-compressed with a very high-quality factor; compression ratios are typically between 1:2 and 1:3 (which is about 2.5 to 3.0 bits per pixels). This allows almost visually lossless image reconstruction on the other end. The communication media include local area networks and therefore the available network bandwidth is relatively high. Typical applications include medical imaginary, space exploration applications, military applications, and others.

In these applications, the user browses through the images in the database in order to quickly identify the correct image. If a standard encoding system is used, it will take a long time to find the image. With the IPES, the user will first receive the DC image, and when the right image is found, he/she may request the remaining scans of the image. The reconstructed quality of the whole image or of the desired regions of the image is almost indistinguishable from the original. The emphasis in these applications is on quality of the image.

In the following example, we illustrate the efficiency of the proposed interactive progressive encoding system in retrieving individual images from a group of images. In Figure 8a, an image is shown that integrates 12 different images. This image is compressed and stored as one image. The size of the image is 624x320 pixels, and in uncompressed form it requires 600 KB (or 4.8 Mbits). When compressed using the quality factor=1, the compressed image requires 204.5 KB – compression ratio has been 2.9.

The IPES was used to extract two images from the group of images. In the first scan, the DC image was obtained. Then, we selected only two regions for subsequent scans. The final image is shown in Figure 8b, in which two images were perfectly reconstructed – " White House" and "Presidents." It required only 286 Kbits to transmit in order to obtain the image.

5. CONCLUSION

We presented the interactive progressive encoding system, which appears to be effective for transmission of complex images over low bandwidth communication channels as well as for archiving and browsing visually lossless image databases. We designed and developed the prototype IPES system, which is an extension of the spectral selection progressive JPEG system. The prototype system uses four scans based on four frequency bands. In our future research we intend to investigate interactive progressive systems that apply more than four scans, say 5 to 10 scans, as well as to develop interactive systems that use the successful approximation technique and the combined technique.

In addition, we are currently researching the interactive progressive video encoding and transmission systems. The basic idea consists of extending the IPES to real-time transmission of high-resolution video, as briefly described next. In a real-time video transmission application, such as live video broadcasting or videoconferencing, at the beginning of transmission the video is compressed using a standard technique (such as MPEG, H.26X, or M-JPEG), and transmitted over the network. For a high-resolution video (such as HDTV), the limited bandwidth will cause that the quality of the reconstructed video will be low (compression ratio must be very high). However, now the user can select a region of video of interest, and can request to receive only this region of the video. This request will be sent to the server, and from that time, the server will transmit only the selected region. Because the size of the selected region will typically be smaller than the original frame size, the quality of the video can be significantly improved. At any time, the user can return to the full frame, and then select another region of interest. There are a number of applications that can benefit from this approach, such as military and space exploration applications.

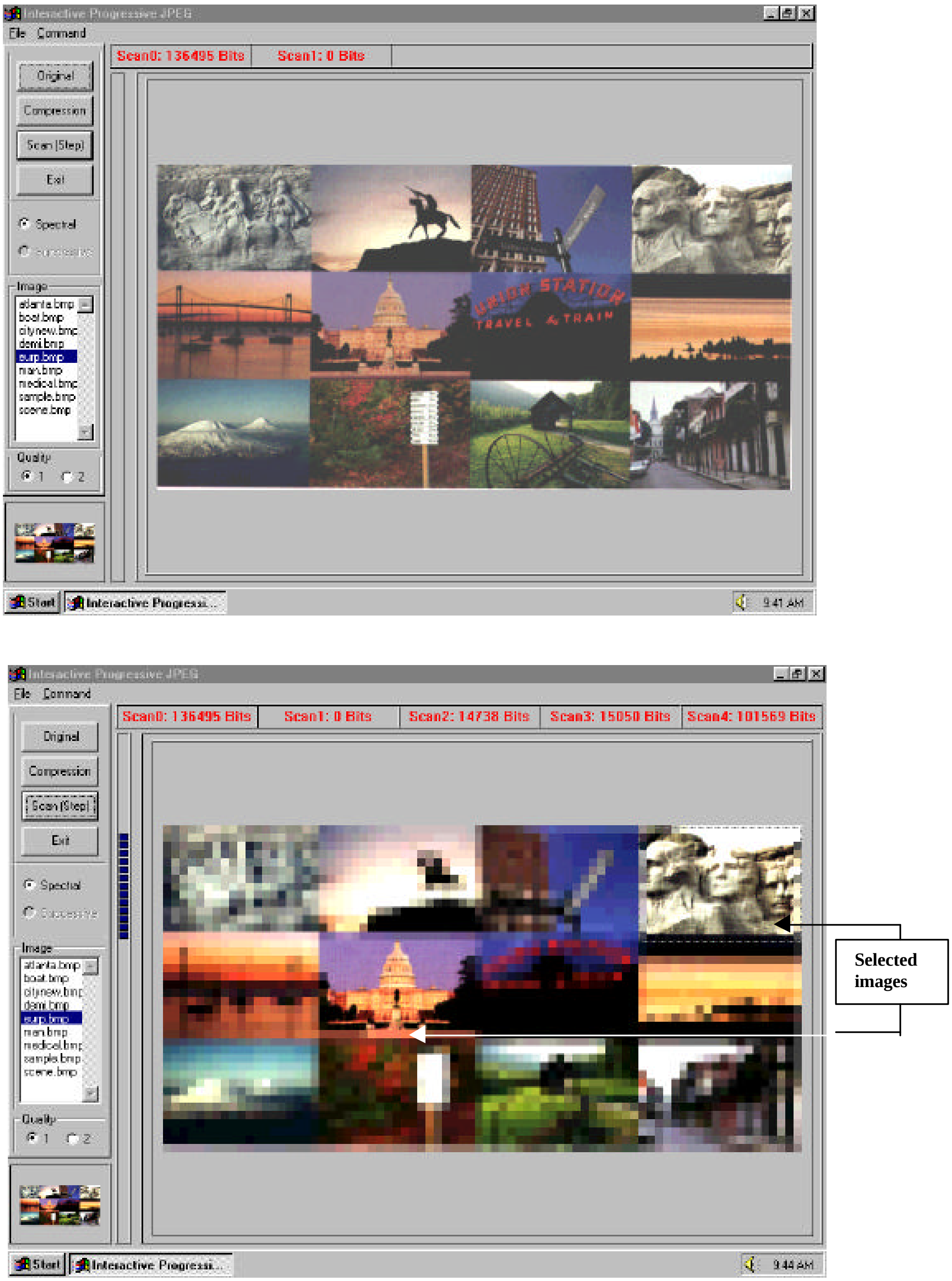

Figure 8. The interactive progressive encoding system applied to extract images from a group of images. (a) The group of 12 images is stored as an image. (b) Tow images were extracted using the IPES system.

REFERENCES

1. W.B. Pennenbaker and J.L. Mitchell, “JPEG Still Image Data Compression Standard”, *Van Nostrand Reinhold*, New York, 1993.
2. B. Furht, “A Survey of Multimedia Compression Techniques and Standards. Part I: JPEG Standard,” *Real-Time Imaging Journal*, Vol. 1, pp. 49-67, 1995.
3. G. Wallace, “The JPEG Still Picture Compression Standard,” *Communications of the ACM*, Vol. 34, pp. 30-44, 1991.
4. R. Steinmetz, “Data Compression in Multimedia Computing – Standards and Systems, Part I and II,” *Multimedia Systems*, Vol. 1, pp. 166-172 and 187-204, 1994.
5. T.S. Wilkinson, “Wavelet-Based Compression for Interactive Progressive Image Transmission”, *Proceedings of Combined Industry, Space and Earth Science Data Compression Workshop*, Snowbird, Utah, pp. 1-10, April 1996.