Remote display of large raster images using JPEG2000 and the rectangular FishEye-View

René Rosenbaum*
Institute of Computer Graphics
University of Rostock
18059 Rostock
Germany
rrosen@informatik.uni-rostock.de

David S. Taubman
School of Electr. Engineering&Telecommunication
University of New South Wales
Sydney 2052, N.S.W.
Australia
d.taubman@unsw.edu.au

ABSTRACT

In mobile environments, the size of large raster images often exceeds the display area on small output devices. The rectangular FishEye-View as a Focus&Context technique can be used to provide an interactive view of the image without zooming and panning.

This paper presents the application of JPEG2000 for the demand-driven transmission of image data optimized for this display technique. The proposed method is fully compliant to the JPEG2000 standard and enhances its static concept for regions of interest to be fully dynamic. Due to limited processing power at client side almost all calculations belonging to the dynamic RoI-coding are executed at server side.

Keywords

JPEG2000, regions of interest (RoI), rectangular FishEye-View, image coding, image transmission.

1. INTRODUCTION

In mobile environments, the size of large images often exceeds the display area of the user's output device. The method commonly used to present such images is to allow interactive pan and zoom by the viewer. An important drawback of this conventional paradigm is that the display cannot simultaneously provide information about the user's region of interest, as well as surrounding areas. While the user can pan to new locations, contextual information, identifying the position of the view within the image, is lost. These difficulties become particularly apparent when working with the small displays typically available to mobile devices.

To overcome these drawbacks, display techniques from information visualization can also be applied to image data. One of these methods is the FishEye-View introduced by Furnas [Fur81] to display large structures efficiently. This general approach can also be applied to raster images [Rau99]. It belongs to the group of Focus&Context techniques, which display the user's current main area of interest in an undistorted focus region. This focus region is embedded in a spatially distorted representation of the remaining image areas. In this way, the user receives both the information which is of principle interest and useful context information for orientation and further navigation. This strategy substantially facilitates interactive navigation within large images. To distinguish between the importance of different image regions, the concepts of regions of interest (RoI) and level of detail (LoD), as proposed in [RS99], are used.

In mobile environments the image data must be often transmitted in compressed form before it can be displayed. Progressive refinement of the compressed representation is a particularly useful feature, since it provides an interactive user at any given point in the transmission with a preview of the whole image information. As explained in Section 2, the proposed rectangular FishEye-View is coupled with a transmission system which supports compression and

Journal of WSCG, Vol.11, No.1., ISSN 1213-6972
WSCG'2003, February 3-7, 2003, Plzen, Czech Republic.

*Thanks to German Academic Exchange Service and University of New South Wales for supporting the research.

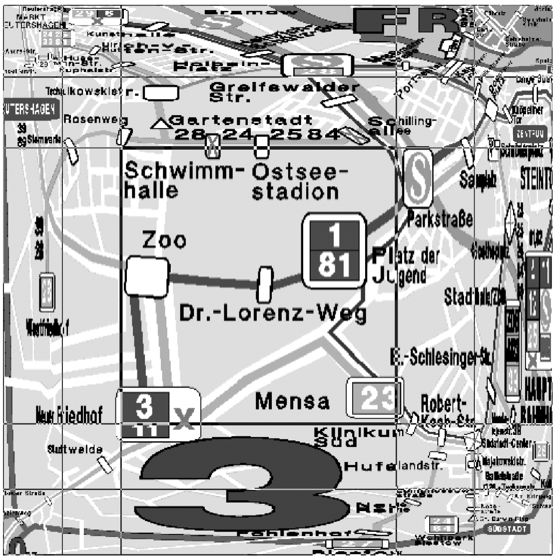


Figure 1. The rectangular FishEye-View.

an adapted progressive refinement. Section 3 serves to familiarize the reader with relevant aspects of the JPEG2000 compression standard, adopted for our work. Section 4 introduces a new approach for dynamic region of interest coding in JPEG2000, and describes how it can be used especially for the rectangular FishEye-View. Our proposal is completed by presenting results and a conclusion in Sections 5 and 6.

2. THE RECTANGULAR FISHEYE-VIEW

Focus&Context techniques are based on the reasonable assumption that a user's degree of interest decreases with distance from the focus region. Accordingly, the rectangular FishEye-View applies varying degrees of distortion to different parts of the image, starting with less distortion near the focus region, and progressing to strong distortion near the borders of the available display. These distortions amount to selective scaling (compaction) of rectangular regions from the original image, as illustrated in Figures 1 and 2. Specifically, the original image data is classified into separate RoI's each of which is assigned a pair of scaling factors, denoted *H:V*. Therefore, each original image region is scaled horizontally by *1/H* and vertically by *1/V*. As may be seen from the figures, the choice of scaling factors is not completely arbitrary, since the overall transformation from original image to display must preserve continuity at the region boundaries.

In addition to scaling parameters, each rectangular RoI may also be assigned a separate priority value, which is used to determine the order in which compressed data will be transmitted from server to client. In order to maximize the perceived impact of progressive refinement, the underlying transmission system delivers higher priority image data (e.g., the focus region) earlier than lower priority data (context data). As the transmitted data arrives, it can be incrementally decompressed and rendered to the user's display. The exact nature of priority values, and the way in which they are used to control the sequence of transmitted data, is described later in Section 4.

Since some of the regions will be scaled at client side, prior to display in the rectangular FishEye-View, not all of the original image information associated with those regions can be used. As an example, the upper left hand region shown in Figure 2 has scaling factors *4:4*, meaning that only one quarter of the original resolution can actually be used, in each direction. As we shall see, the multiresolution properties of the Discrete Wavelet Transform (DWT) employed by JPEG2000 may be exoited to transmit only the information which is actually relevant to the distorted FishEye-View.

3. JPEG2000

In this section, we describe some of the most relevant features of the JPEG2000 image compression standard [J2kBA]. While JPEG2000 does generally provide superior compression efficiency to that of JPEG, its most significant advantages rest with its support for progressive refinement of the image information in both resolution and quality, along with its support for region of interest access to the image.

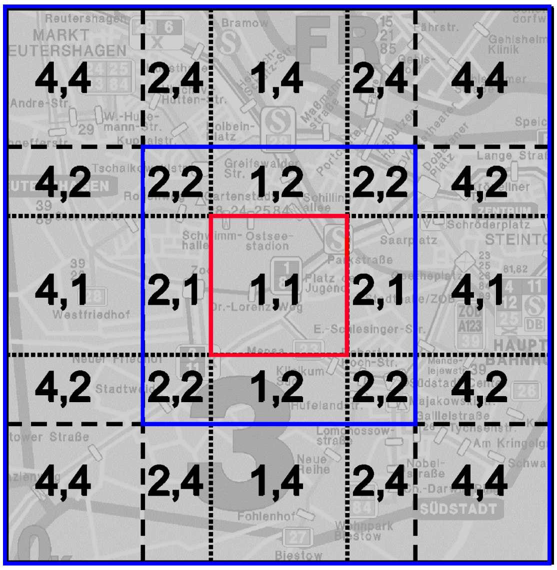


Figure 2. Applied scaling to the different regions of interest in *X* and *Y* direction.

3.1 Multiresolution Decomposition and Compression

The JPEG2000 standard employs a dyadic multiresolution transform, known as the Discrete Wavelet Transform (DWT) to analyze the image into a collection of subband images, each of which contains information from different spatial frequency bands. Each successive resolution level in the image representation consists of three "detail" subbands, which augment the information available from the previous resolution level, having half the number of samples in each direction. One detail subband contains the information required to double the image resolution in the horizontal direction; another contains the information required to double the resolution vertically; while the third completes the information required to double the resolution in both directions. Figure 3 illustrates the detail bands associated with two resolution levels, together with a separate lowest resolution image (top left in the figure).

The DWT coefficients in each subband are further sub-divided into code-blocks, each of which is independently encoded. The encoded code-block bitstreams are then grouped into containers called *precincts*. Each precinct contains a number of corresponding code-blocks from every subband of a single resolution level (Fig.3). Precincts form the base for the creation of the final compressed bitstream, which is explained in the next subsection.

3.2 JPEG2000 Bitstream Formation

Precinct structures play an important role in enabling resolution- and region-based access to JPEG2000 compressed images. The precincts can have different dimensions in each resolution level and are used to lay down rules for the creation of packets, which form the final bitstream.

To allow progressive refinement of the image representation, the encoded data associated with each code-block is spread over a number of *layers*. This is generally done by an optimization procedure, which aims to uniformly increase the image quality by similar perceptual increments in each successive layer.

The final JPEG2000 bitstream is constructed by concatenating a list of so-called *packets*, where each packet represents the compressed data contributions to a single layer, from code-blocks belonging to a single precinct. Thus, each precinct contributes one packet for each layer.

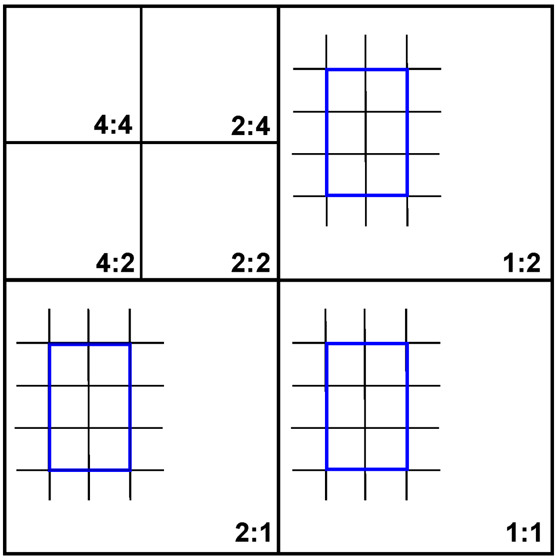


Figure 3. Multiresolution decomposition in JPEG2000 using different scaling values (blue: subband areas belonging to a single precinct).

In the next section, we briefly review existing proposed methods for adapting the image data transmission, based on RoI's.

3.3 Regions of Interest in JPEG2000

In most of the proposed schemes for RoI encoding in JPEG2000, the supported region has to be defined during the image encoding process [J2kBA,J2k2A,WBEB02,GCT00]. This has the drawback that the RoIs cannot easily be defined dynamically, in accordance with a user's interests. One scheme has been proposed in [RSI02] to support dynamic regions of interest. Here, JPEG2000's flexible packet syntax is used to adjust the sequence in which compressed data packets are transmitted, so as to reflect the relevance of different precincts to a region of interest. Even though the order of the compressed data is determined dynamically, the transmitted information always constitutes a compliant JPEG2000 bitstream, which can be reconstructed and rendered progressively using a compliant JPEG2000 decoder. No further RoI handling is required at the client side, and the decompressed data is suitable for display within a rectangular FishEye-View. The next section is concerned with particular considerations for enhancing the efficiency and sequencing of transmitted information to be used with the FishEye-View.

4. DYNAMIC ROIS FOR THE FISHEYE-VIEW

The basic idea of our approach is the prioritized transmission of JPEG2000 packets, so as to improve the efficiency and rate of refinement of relevant regions in the FishEye-View. It is helpful to differentiate between refinement in resolution and refinement in quality. With regard to resolution, we take advantage of the fact that not all of the original image resolution will actually be displayed in non-focal regions -- those with scaling factors $H > 1$ or $V > 1$. This means that some original image subbands can be discarded in such regions, without any loss of information.

We use the term "Level of Detail (LoD)" and refer to the collection of DWT subbands which are relevant to the FishEye-View in any given region as *the final LoD*. These are the subbands, *b*, whose scaling factors $H_b : V_b$ (see Fig.3) satisfy $H_b \geq H$ or $V_b \geq V$, where $H : V$ are the scaling factors for the region. By pruning the transmitted information to avoid data which lies beyond the final LoD, valuable transmission bandwidth can be saved.

With regard to quality, we take advantage of the fact that the perceived impact of any given layer on visual quality is generally lower in the context regions than the focus region. Equivalently, uniform perceived image quality generally involves a smaller number of layers for precincts which contribute only to the context region than for precincts which contribute to the focus region. One reason for this is simply that the focus region is by definition more important to the interactive user. Another reason is that non-focus regions are scaled. Scaling an image region by factors $H : V$ reduces the impact of quantization on mean squared error in the rendered image by the factor *HV*. Although formal means may be derived for accommodating this later phenomenon through a re-sequencing of the original packets, it is less clear how the former aspect should be accommodated in a rigorous manner. For the purpose of the presented study, a simple prioritization scheme is employed to determine the number of layers from a context region which should be considered perceptually similar to a number of layers from the focus region.

To create a compliant RoI-enhanced bitstream which can be easily decoded at client-side, our approach is based on [RSI02]. The original image is compressed only once, into a JPEG2000 compressed bitstream, which can be used to serve clients with different focus regions. For any given rectangular FishEye-View, specified by the remote client, a server determines the RoIs, their scaling values and their final LoDs. For each RoI and each resolution, the server determines the set of precincts from that region which contribute (either completely, or in part) to the RoI. A precinct contributes to the RoI if its spatial region of support overlaps with that of the RoI and its resolution contains one or more subbands which belong to the RoI's final LoD.

The set of contributing precincts determines the collection of packets from the original compressed image bitstream which are relevant to the FishEye-View. An RoI-scheduler then determines the order in which these packets are to be transmitted. The RoI-enhanced bitstream is created layer-wise, starting with the lowest layer. To ensure that the sequence of constructed layers forms a compliant JPEG2000 bitstream, each layer must contain at least a place holder for every packet in the original image. However, only the contributing precincts are assigned non-empty packets in the current layer. All other precincts are assigned "empty packets", which contribute no compressed data from their code-blocks to the layer being constructed.

In order to exploit the different priority associated with precincts which contribute to different RoIs, the scheduler generally creates a transmitted bitstream which contains more layers than the original bitstream. The first layer created by the scheduler contains all first packets from those precincts which contribute to the highest priority RoI (the focus region), with all other packets empty. Subsequent layers created by the scheduler contain later packets from the precincts which contribute to the highest priority RoI and earlier packets from the precincts which contribute only to lower priority RoIs (context).

For our current work, a simple prioritization scheme is implemented, with each precinct, *i*, assigned an integer priority index, p_i. Suppose that the index of the first layer not yet transmitted from precinct *i* is $l_i \in \{1 \leq l < L\}$, where *L* represents the number of layers in the original JPEG2000 data stream. When the scheduler creates a new layer for transmission, it includes a new packet for precinct *i* only if this is a contributing precinct and $l_j - p_j \geq l_i - p_i$ for all contributing precincts, *j*. This policy gives a region r_a, having higher priority, a competitive edge of $p_a - p_b$ layers over a region r_b, having lower priority.

As described in Section 3.1 every packet contains image information from every subband of a resolution level. Thus, a certain packet may contain more image data than necessary to refine a scaled region. This happens in every RoI where the final LoD contains some but not all subbands from a resolution level. These RoIs have scaling factors $H:V$ with $H \neq V$. To avoid such redundant transmission, it is necessary to generate new compressed data packets, which contain non-empty contributions only from the code-blocks which belong to subbands within the final LoD. The server can perform the necessary *repacketization* of code-block contributions at some expense in computational complexity and signaling overhead within the JPEG2000 packet headers. As we shall see in the next section, however, this repacketization can have a significant beneficial impact on the efficiency with which data is transmitted for the FishEye-View.

The dynamically created packets can be scheduled in the same manner described previously. Note, however, that the RoI-scheduling task becomes more complex, since it must keep track of both original image packets which have been fully transmitted, and also original packets which have been only partially transmitted. Later changes in the user's focus region may require these partial packets to be completed by the transmission of additional complementary packets.

5. RESULTS

To demonstrate the effectiveness of the proposed approach, we have implemented a client/server image transmission architecture with a bi-directional communication. The client simply consists in a GUI, to define and display the rectangular FishEye-View, a JPEG2000-compliant decoder to convert incoming image data, and a communication component to handle data transmission from the server and the delivery of request parameters to the server. The server includes a transcoding element, which dynamically transcodes an original bitstream into one with the RoI-enhanced layering, as described above. The transcoded RoI-enhanced layers are transmitted to the client packet by packet, in a sequence which is compatible with the JPEG2000 data stream syntax. We use the Kakadu JPEG2000 tools [Kak02] to accomplish this dynamic transcoding, including, where necessary, the repacketization operations described in the previous section.

Figures 4 and 5 present numerical results which indicate the impact of the proposed dynamic RoIs on transmission efficiency. The upper curve, marked "original" in each of the figures, corresponds to the original transmission scheme proposed in [RSI02]. In this scheme, all information relevant to the FishEye-View must eventually be transmitted to the client. A second curve in each figure identifies the performance of our proposed transmission scheme, where original packets are not sub-divided to accommodate differences in the horizontal and vertical scaling factors. The third, lowest curve, in each figure corresponds to the proposed scheme, using repacketization to avoid the transmission of subbands which do not belong to the final LoD of any RoI. Both figures correspond to the same image, RoI parameters, and transmission sequence, with Figure 4 identifying the size of each transmitted RoI-enhanced layer, and Figure 5 identifying the cumulative size of the layers. The original image considered here has a

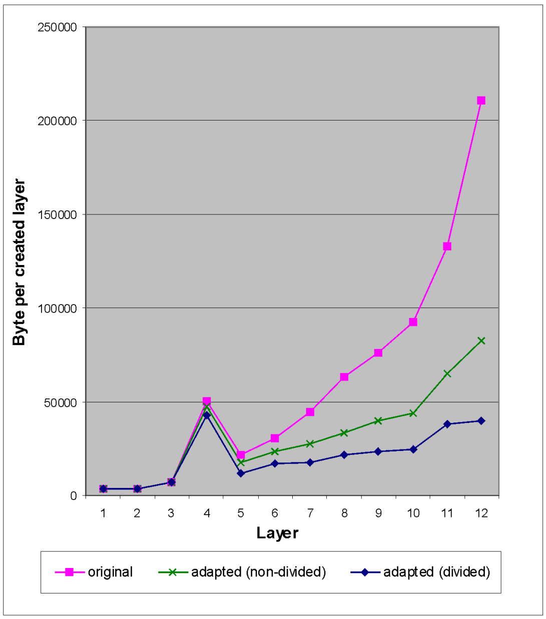


Figure 4. The length of a layer increases with the number of non-empty packets to include.

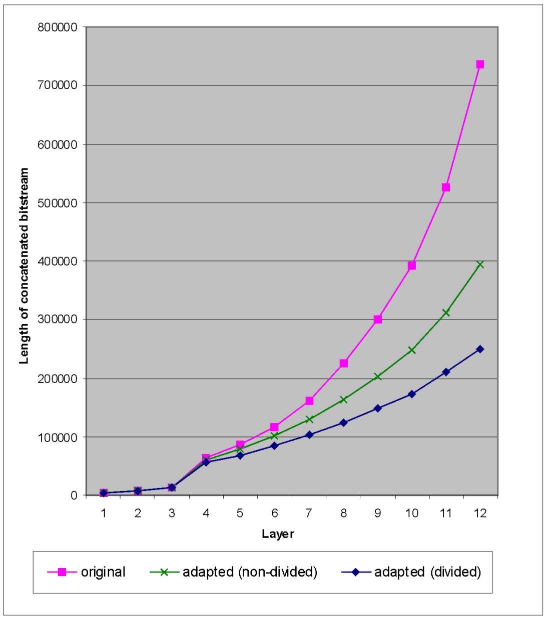


Figure 5. Comparison of the different methods using the length of the final bitstream.

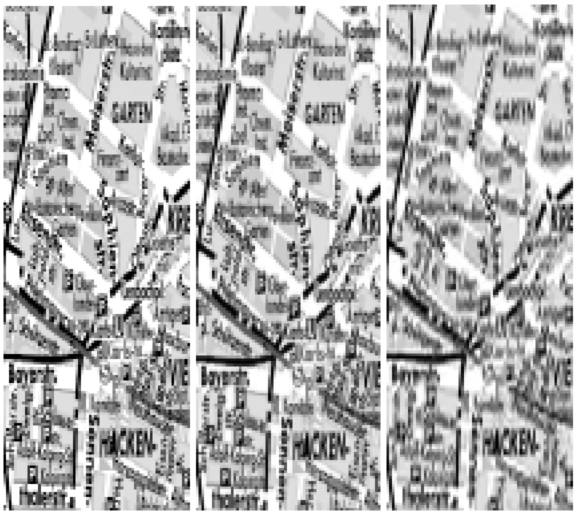
size of 1024×1024, and a compressed representation of length *727* kB, consisting of *9* layers, with *2* levels of DWT decomposition, as in Figure 3. Additionally, we use precincts of size 32×32 in the spatial domain and shrink this size progressively for precincts in lower resolution levels, by factor *2* in each direction. Thus, precincts in every resolution level contribute to a spatial area of approximately the same size. This policy for sizing precincts limits the extent of their spatial contribution and helps to speed up our implementation.

The cumulative bit-rate associated with each successive layer in the original bitstream is $\sqrt{2}$ times that of the previous layer. The focus region for the FishEye-View has been placed in the center of the image, having *1/8* of the full image width and height. A prioritization index, p_i, of *3* is assigned to the precincts which contribute to the focus region. The prioritization index is reduced to *1* in the first context belt (regions scaled by at most *2* in either direction), and *0* in the outer belt (regions scaled by at most *4* in any direction). Accordingly, the final RoI-enhanced bitstream contains *12* layers.

The results presented in Figures 4 and 5 are easily understood. The original scheme produces the largest layers, and hence has the lowest transmission efficiency, because all subbands from all resolutions must be included in both focus and context regions. The performance of the FishEye-adapted transmission scheme is clearly superior, and most notably so when repacketization is used to avoid redundant transmission of subbands which are not relevant to the scaled RoIs. From Figure 4 we see that the ratio between the most and least efficient transmission schemes can be as large as *4.5:1*. This is because RoIs which have large scaling factors generally represent the largest portion of the original image, and these are the very RoIs which stand to benefit most from the proposed FishEye-adapted transmission scheme.

Evidently, the effect of FishEye adaptation is significant only in the later layers. This is because the prioritization scheme delays the scheduling of context information until later layers; FishEye adaptation affects only the packets which belong to the context region, not the unscaled focus region. The strong increment in transmitted length observed at layer 4 corresponds to the first appearance of packets from the second belt; these represent the largest region on the original image.

The cumulative layer sizes presented in Figure 5 are perhaps more indicative of the overall impact of our proposed transmission scheme, revealing a maximum overall gain of approximately *3:1* in transmission efficiency relative to the original approach. With regard to objective and subjective quality, the decoded and displayed FishEye-View image appears to be almost identical (cp. Figure 6) using all three methods compared in Figures 4 and 5. This confirms our assertion that subbands not belonging to a region's final LoD can be safely omitted.



(original) (non-divided) (divided)

Figure 6. The scaled image content appears almost identical for all three methods if displayed within FishEye-View.

Dividing packets into their subband contributions places increased demands on the server's computational resources. The additional state memory required to keep track of these subband contributions, also places demands on the server's resources. Nevertheless, even low powered servers can benefit from our proposed approach to delivering original undivided packets, since fewer packets must be handled than in the original case.

Regarding processing at the client side, both the proposed approach with undivided packets and the proposed approach with full repacketization tend to reduce the client's complexity, since they reduce the amount of data which must be decoded. With the original scheme, a considerable amount of data may be decoded which does not substantially contribute to the rendered FishEye-View. In fact, the lowest client complexity is achieved when the server repacketizes all original packets which contribute only partially to the final LoD.

Even though the results presented here suggest that significant gains in transmission efficiency can be obtained by a FishEye-adapted transmission scheme, it is worth noting that we are unable to fully eliminate the transmission of redundant information for regions with certain scaling parameters. As shown in Figure 2, the FishEye-View involves some RoIs whose horizontal and vertical scaling parameters can differ by a factor greater than *2*. Unfortunately, the JPEG2000 decomposition scheme (see Fig.3) supports different horizontal and vertical resolutions only up to a factor of *2*, so we are forced to transmit redundant information for some regions. Consider, for example, a region with scaling parameters *4:1*, as shown in Figure 2. Such a region must be reconstructed at *1/4* the original resolution in the horizontal direction, but at full resolution vertically. This suggests that only the subbands marked *4:4*, *4:2* and *2:1* in Figure 3 need actually be sent to the decoder. The total number of transmitted subband samples in this case is *3/8* of the original image samples, while a perfectly efficient scheme would send only *1/4* the original samples. This necessitates a transmission of *1/8* more image samples, but ensures that our approach is still JPEG2000-compliant.

6. CONCLUSIONS

In this paper, we have proposed new methods for transmitting a JPEG2000-compliant bitstream, so as to allow efficient remote browsing of large images, using a rectangular FishEye-View. This method is based on the successive creation of new JPEG2000 layers, using already encoded packets from an original compressed image. After each layer has been dynamically created, it can be transmitted immediately to the client. This can be done without regard for the effect which future client interaction may have on the packets to be included in subsequent layers. The resulting stream of transmitted data can be decoded by a compliant JPEG2000 decoder, at any point in the transmission process, allowing progressive refinement of the displayed FishEye-View.

To avoid the transmission of redundant image data, we assign a final LoD to each region of the FishEye-View, based on the scaling distortions associated with that region. Only image data which contributes to the refinement of this LoD is considered for transmission. This strategy saves significant transmission bandwidth, without damaging the visual properties of the image, rendered within the FishEye-View. The JPEG2000 standard further allows an intelligent server to repacketize the original compressed data on the fly, so as to eliminate specific subbands which do not contribute to any final LoDs. Our experiments show that this can lead to further improvements in transmission efficiency, at the expense of some additional computation and state management memory in the server.

7. REFERENCES

[Fur81] Furnas, G. W.: "The FISHEYE view: a new look at structured files", Bell Laboratories Tech. Report, Murray Hill, New Jersey, 1981.

[Rau99] Rauschenbach, U. The rectangular FishEye View as an efficient method for the transmission and display of large images, in Proceedings of IEEE-ICIP'99, Kobe, Japan, 1999.

[RS99] Rauschenbach, U., Schumann, H. Demand-driven Image Transmission with Levels of Detail and Regions of Interest, Computers and Graphics, Vol23/6, December, 1999.

[J2kBA] JPEG2000 image coding system, part1, Final Publication Draft, ISO/IEC JTC 1/SC29/WG 1 N2678, 19, July 2002.

[J2k2A] JPEG2000 image coding system, part2, Final Publication Draft, ISO/IEC JTC 1/SC29/WG 1 N2679, 19 July 2002.

[WBEB02] Wang, Z., Banerjee, S., Evans, B., Bovik, A. Generalized bitplane-by-bitplane shift method for jpeg2000 RoI coding, in Proc. IEEE-ICIP'02, Rochester, N.Y., September 22-25 2002.

[GCT00] Grosbois, R., Santa Cruz, D., Ebrahimi, T., New approach to jpeg2000 compliant region of interest coding, SPIE 46th annual meeting, Applications of Digital Image Processing, July 2001.

[RSI02] Rosenbaum, R., Schumann, H., Flexible, dynamic and compliant region of interest coding in JPEG2000, in Proc. IEEE-ICIP02, Rochester, N.Y., September 22-25 2002.

[Kak02] KAKADU, a comprehensive, heavily optimized, fully compliant software toolkit for JPEG2000 developers, Version 3.3, http://www.kakadusoftware.com/, 2002.