The Rectangular Fish Eye View as an Efficient Method for the Transmission and Display of Large Images

Uwe Rauschenbach
University of Rostock, Computer Science Department, D–18051 Rostock, Germany
urausche@informatik.uni-rostock.de

Abstract

With the advances of mobile computing technology, the users of mobile hardware expect the same service as users of stationary computers do. Access to large images over the Internet poses two problems to users of wireless data communication devices and mobile computers: low transmission bandwidth and small screen space. In this paper, an integrated image transmission and display method is proposed, which efficiently uses both transmission bandwidth and screen real estate. In order to achieve that, we developed a level–of–detail and region–of–interest transmission scheme for raster images based on embedded zerotree wavelet coding. A special fish eye view technique is built on that transmission scheme, which can benefit directly from the properties of a modified wavelet decomposition.

1. Introduction

Still images play an important role in today's networked multimedia systems, e.g., the World Wide Web. Image transmission accounts for a substantial fraction of the total bandwidth used. Advances in mobile technology are enabling Internet access for the users of mobile computing hardware and mobile data communication devices. These users want to access the same information as the users of stationary computers and wired networks. However, mobile computing environments suffer from limitations imposed by small screens and the low bandwidth of the wireless modems, typically 9600 bits per second for GSM.

In order to decrease the bandwidth demands, lossy image compression methods must be used. The compression scheme should support progressive transmission in order to allow the user to cancel unwanted transmissions at an early stage and to shorten response times. Wavelet–based compression is a good opportunity, since there exist embedded coding schemes like [4, 6] which can achieve this.

Especially for large images, the user often wants a certain resolution or quality only for some *regions of interest (RoIs)* in the image. The desired resolution or quality can be specified in terms of a *level of detail (LoD)* [3]. By combining LoD/RoI support with an embedded compression scheme, bandwidth demands can be decreased further. To put full control over the transmission process to the user, an interactive component for specifying the LoDs and RoIs is necessary. When the user requests refinement of a region of interest, only *differential data* should be transmitted.

If the screen space is not sufficient to display an image, today's systems usually exploit zooming, panning and scrolling. An alternative which provides additional context to the area of interest are fish eye views. They can be seen as an extension to zooming, where different zoom factors are assigned to different regions.

In this paper, we present an integrated image transmission and display method, which saves both transmission bandwidth and screen real estate. We propose the rectangular fish eye view, a focus–and–context display technique for raster images. This technique benefits from the properties of a modified wavelet decomposition scheme and is integrated with a new framework for the embedded image transmission under consideration of levels of detail and regions of interest.

Several authors [1, 2] have presented spatially variable extensions to progressive image coding. These works use a notion of points of interest and an importance function which computes a weight for each wavelet coefficient in the neighborhood of such a point. In contrast to our proposed method, these schemes require to store transmission state information per coefficient instead per RoI and are not coupled with a display technique.

2. Rectangular Fish Eye View: Technique and Requirements

The rectangular fish eye view is a focus–and–context display technique for raster images similar to the rubber sheets [5] proposed by Sarkar et al. for vector graphics. Figures 3 to 6 show an example, which we will discuss in detail later. A *focus region*, displaying a part of the image at full resolution, is surrounded by *context rings*, "squeezed" by ascending powers of two. Each ring is formed by eight adjacent rectangular regions, which are downscaled in x and in y direction by (possibly different) powers of two. Figure 7

This work was supported by The German Science Foundation under contract no. Schu-887/3-3.

depicts the downscaling grid which has been used to generate the view. Because only a part of the image is shown at full resolution at a time, the system must allow the user to pan the focus region interactively, causing a new part of the image to be displayed in detail. Furthermore, interactions for customizing the view by resizing the focus region and changing the number and size of the context rings should also be provided.

The proposed fish eye technique allows the display of large images in a screen–space saving manner, providing detailed information in the image region the user currently pays attention to, and coarse context information in the remaining parts of the image. Since the downscaling factors have been chosen to be powers of two, this display technique can be integrated with a wavelet–based image transmission method in order to save bandwidth. That's why the technique is efficiently applicable to remotely stored large images.

For efficient bandwidth use, it must be possible to transmit the image data at different resolutions in different regions as needed for display; a grid of regions of interest is needed. Since only data required for display should be transmitted, the underlying image representation must support different downscaling factors in the x and the y direction. Progressive refinement during transmission is strongly desired in order to give the user an early impression whether or not the focus region is positioned correctly. Transmission starts with an initial grid. Each interaction specifies a new grid, which must be shared between transmitter and receiver and maintained at both sides. Creating a new grid must cause only differential data to be transmitted, since redundant data transmissions are intolerable in low bandwidth environments. Prioritization of RoIs is required to allow preferential transmission of the focus region.

In order to meet these requirements, we compose the downscaling grid from rectangular *regions of interest*, and the downscaling factors are modeled as *levels of detail*. A new wavelet decomposition scheme supports different resolutions in x and y direction. Progressive transmission is realized by a variant of the embedded zerotree wavelet method [6], which has been adapted to cater for the needs of LoDs and RoIs. Redundancy–free transmission is ensured by using a scheme which creates a set of partitioning intersections from the original overlapping RoI grids. RoI scheduler components at sender and receiver prioritize, control and synchronize the transmission. They can be controlled using commands. The remaining sections deal with some of these concepts and components in greater detail.

3. Levels of Detail and Regions of Interest

A level of detail (LoD) determines how detailed a part of the image is transmitted. Each LoD can be represented as an ordered set of vectors in a vector space with the dimensions *x resolution*, *y resolution*, *precision* and *color*. The order in the set is described by constraints and determines the sequence of transmission of the data corresponding to the individual elements of the LoD. This mechanism ensures that the data are transmitted in the right order, for instance, that the low–low subband of the wavelet representation is being traversed before the higher subbands, or that higher bitplanes are visited before lower bitplanes, which are important preconditions for the encoding/decoding algorithm to work.

A region of interest (RoI) assigns a level of detail to a rectangular set of pixels in the image, called the *footprint*. For each RoI, three kinds of LoDs are maintained: the *transmission target*, at which the transmission of data for the RoI is terminated, the *transmission state*, which represents the amount of data already transmitted for that RoI, and the *delta state*, which is the set difference between the transmission target and the transmission state. Furthermore, a RoI is assigned *scheduling information*. A more detailed and formal description of the LoD/RoI model can be found in [3].

4. A new wavelet decomposition scheme

The fish eye technique requires the resolutions in x and y direction to differ by more than a factor of two in some grid rectangles (e.g., the top middle rectangle in figure 7 is scaled down in y direction to one–fourth and not scaled in x direction; thus, the resolutions differ by a factor of 4). Unfortunately, the classic wavelet decomposition scheme supports differing resolutions only up to a factor of two. Using this scheme, it would be necessary for some regions to transmit image data at a higher resolution than needed for display and to perform additional downscaling at the receiver side. Since it is important to transmit only data which are required, we introduce a modified wavelet decomposition scheme as shown in figure 1. Additionally to filtering the LL subband as in the dyadic decomposition method, we apply the wavelet filter in x direction to the LH subband and in y direction to the HL subband (marked by the gray boxes in figure 1). By doing this, these subbands are decomposed further in one direction, and we get a subband structure which supports the non–redundant transmission of image parts where the downscaling factors for x and y direction can differ by a factor of 2^n with $n \geq 1$, where n is the number of one–directional decomposition steps.

The new decomposition scheme may introduce artefacts in the focus region during the early stages of progressive transmission, as shown in figure 2A. Horizontal and vertical structures tend to intrude into areas of low activity. This artefact can be reduced using a combination of two measures: First, we can reduce the number of one–directional

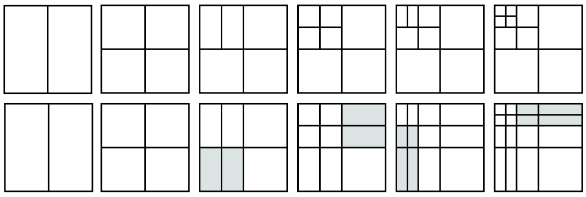

Figure 1. Classic (top) and new (bottom) decomposition

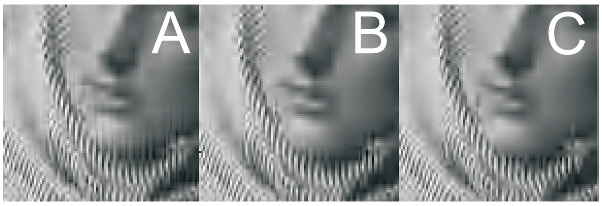

Figure 2. Artefacts of new decomposition scheme at 0.25bpp

decomposition levels to the minimum number needed. For the grid in figure 7, just two those levels are required. The figures 2A (4 levels) and 2B (3 levels) show the effect of this reduction. Second, we can use a Haar filter (cf. figure 2C) instead of a biorthogonal filter for the one–directional steps. A combination of these two measures efficiently suppresses the artefacts for our application. Furthermore, progressive refinement guarantees that they are not visible for a long time.

5. Redundancy–free progressive transmission

We modified the embedded zerotree wavelet algorithm to support the new decomposition scheme and integrated RoI and LoD support into it. The LoD vector space dimensions are mapped to wavelet representation properties as follows: the spatial resolutions are represented by the subbands, the precision by the coefficient bitplanes and the color dimension has the two values Y and $(C_b + C_r)$. During coding and decoding traversal of the wavelet array, we visit only those coefficient bit positions in those subbands which have a corresponding element in the current delta state. RoIs are supported by constraining the traversal to the wavelet coefficients contributing to the RoI's footprint. In order to achieve that, the footprints are transformed to wavelet space, creating a multiresolution representation.

Figure 7 illustrates how RoIs and LoDs can be exploited to generate the downscaling grid of a rectangular fish eye view. According to this grid, we define 25 RoIs with rectangular footprints and specify their transmission targets as follows: The dimensions *x resolution* and *y resolution* of the target LoD describe the downscaling of the image data for the individual regions; they limit the data to be transmitted to the according subbands. The dimensions *precision* and *color* are not needed for the spatial layout of the fish eye view; they are used for progressive refinement, i.e., set to their maximum values in the transmission target.

To ensure redundancy–free transmission even if RoIs have overlapping footprints, the transmission framework creates a set of new RoIs with non–overlapping footprints called *partitioning intersections*. For each element of such a set, a compound transmission target, transmission state and delta state are computed using set operations from the corresponding parameters of the original RoIs overlapping this new RoI. Only the data represented by the compound delta states are transmitted for each of the partitioning intersections.

6. Transmission Control

Both receiver and transmitter use a RoI scheduler component to keep track of the set of regions of interest and to compute the partitioning intersections. Since explicit storage of the intersections would involve a large memory overhead, we prefer an implicit realization: During coding/decoding traversal of the wavelet coefficient array, we use *span arithmetic* known from the scan conversion of 2D shapes to determine the delta states for the different parts of the current scanline. The current bit of the coefficients on a span is transmitted, if the span is inside at least one footprint whose RoI has – according to its transmission state – not yet transmitted data for the current subband and the current bit, but outside all RoIs which already have.

The schedulers decide depending on scheduling parameters and delta states which subband and which bitplane of the wavelet coefficient array to zerotree–code next for which RoI. To control transmission, each RoI is assigned a *priority* number $p_i \in \{\ldots, -1, 0, 1, \ldots\}$ which influences the transmission sequence as follows: Assume that the number of the bit plane to be transmitted next is denoted by $b_i \in \{1, 2, \ldots\}$, where 1 represents the most significant coefficient bit. From all regions of interest, the RoI scheduler selects a RoI r_j to be transmitted next depending on the following criterion: $\forall i : b_j - p_j \leq b_i - p_i$. During transmission, this gives a higher prioritized RoI r_x a "lead" of $p_x - p_y$ bit planes over a RoI r_y. Since the focus region should be treated with preference, we assign $p = 3$ to the focus, $p = 1$ to the context ring adjacent to the focus and $p = 0$ to the remaining rings.

Sender and receiver must always be kept synchronized. In order to achieve that, control commands to define new RoIs or to change parameters of existing RoIs are embedded into the data stream. If the receiver issues a command to the sender (e.g., to specify a new grid), it is put into a command queue there. Each time after one coefficient bit plane has been encoded for one subband and one RoI, this queue is evaluated; the commands are executed by the sender's RoI scheduler and interleaved with the code stream in order to be executed by the receiver at the same point during the decoding process.

7. Results

7.1. Example

This section will give an example of the proposed method simulating a transmission of the scanned Rostock public transport map over a 7200bps channel. The grid in figure 7 has been used to generate a fish eye view of 512x512 pixels with a focus region of 256x256 pixels from a 1024x1024 pixel image. Figure 3 shows the transmitted image after 27 seconds. The focus region may have been positioned initially by knowing some context information, e.g., the position of the viewer or his travel plans. Although only a small fraction of the image data has been transmitted at this early point (three–fourth of all wavelet coefficients have been completely ignored, from the remaining ones, only some bits have been encoded), the information in the focus region is recognizable. Compared to the 27 seconds transmission time, transmitting the whole image at the LoD of the focus region would have taken 115 seconds.

At this point, the user moves the focus down and to the left (see figure 4). The viewing software immediately changes the layout based on the data already received; the left and the lower part of the new focus region are reconstructed from the available data at a lower level of detail. To update the RoI set, a request is sent to the server specifying the new grid. The server converts this request into commands instructing the encoding RoI scheduler to stop encoding all existing RoIs and to define new ones with the appropriate local LoDs. Given the new grid, only those parts of the wavelet array which have not yet been transmitted are now encoded and sent as differential refinement information. Only 6 seconds after moving the focus (plus latency for sending the request to the server), the new focus region is available at the same transmission state as the old one (see figure 5). Further progressive refinement is carried out automatically, and 88 seconds after starting the transmission, the readability of the whole fish eye (see figure 6) has reached a stage which can not be further improved by transmitting more data (although the artefacts could still be reduced).

7.2. Discussion

The proposed image browsing method has been developed to cope with the bandwidth constraints and screen size limitations of mobile computing environments. Bandwidth is used efficiently by utilizing compression and redundancy–free transmission. The display technique is integrated with the transmission scheme and can benefit from its properties. By using a focus region and context rings, screen space is saved while the important information in the image is maintained. An embedded code stream supports early feedback to correctly position the focus region. Savings in screen space depend on the layout of the downscaling grid but are substantial: The view in figure 6 occupies only one fourth of the space the original image would need. Since only those data are transmitted which are needed for the presentation, the same savings apply to the transmission bandwidth additionally to the savings achieved by the wavelet–based image compression.

8. Acknowledgement

I wish to thank my advisor, Heidrun Schumann, for her support, and Tino Weinkauf for implementation work and for contributing ideas. Special thanks to the Warnow Regional Transport Board for permitting the use of their map image.

9. References

[1] E.C. Chang, C.K. Yap and T.J. Yen, "Realtime Visualization of Large Images over a Thinwire", *Proc. IEEE Visualization '97*, Phoenix, Arizona, Oct. 19–24, 1997.

[2] T. Frajka, P.G. Sherwood and K. Zeger, "Progressive Image Coding with Spatially Variable Resolution", *Proc. ICIP '97*, Santa Barbara, California, October 1997, Vol. 1, 53ff.

[3] U. Rauschenbach and H. Schumann, "Demand–driven Image Transmission with Levels of Detail and Regions of Interest", *Computers and Graphics*, 23 (5), 1999.

[4] A. Said and W.A. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", *IEEE Trans. Circuits and Systems for Video Technology*, 6 (3), 1996, 243–250.

[5] M. Sarkar et al., "Stretching the rubber sheet: A metaphor for visualizing large layouts on small screens", *Proc. ACM Symp. on User Interface Software and Technology*, 1993.

[6] J.M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. Signal Proc.*, 41 (12), 1993, 3445–3462.

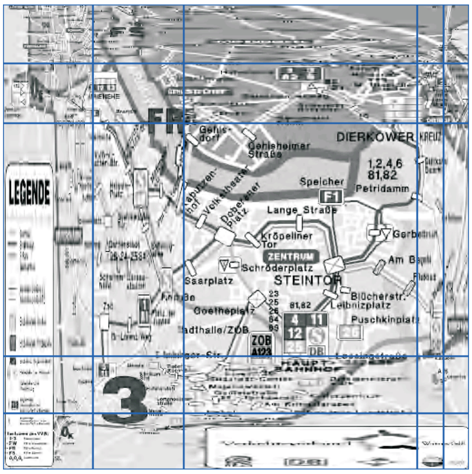

Figure 3. Initial fish eye view, 512 x 512, 24562 bytes

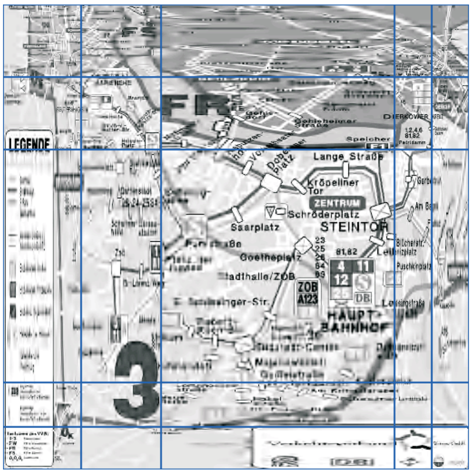

Figure 4. Fish eye view after moving focus, 24562 bytes

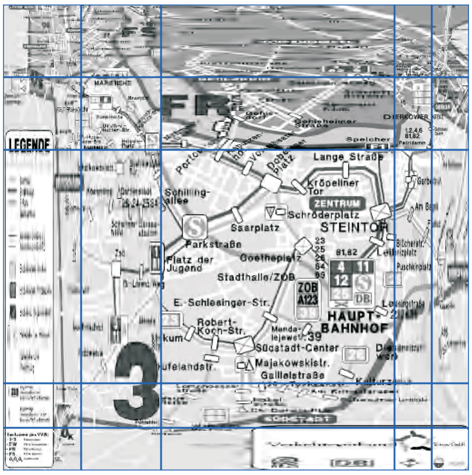

Figure 5. Fish eye view with the new focus, 30091 bytes

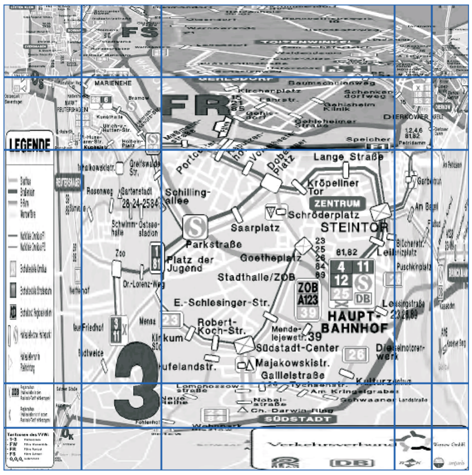

Figure 6. Fish eye view with the new focus, 79617 bytes

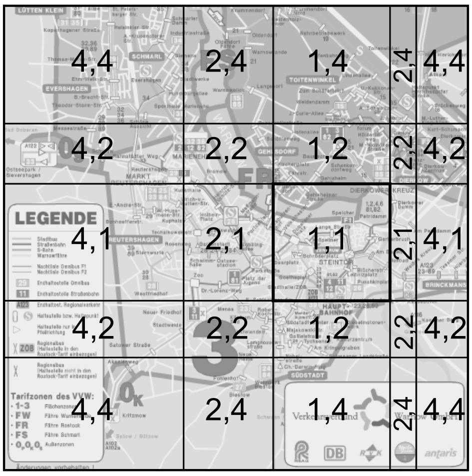

Figure 7. RoI grid of the fish eye view (1024 x 1024)