PROGRESSIVE IMAGE TRANSMISSION USING LEVELS OF DETAIL AND REGIONS OF INTEREST

UWE RAUSCHENBACH

University of Rostock, Computer Science Department, D-18051 Rostock
email: urausche@informatik.uni-rostock.de
Germany

Abstract: This paper presents a method for progressive image transmission, which consistently handles the problem of levels of detail and regions of interest. We propose a combined method, that allows the specification of the desired degree of detail for overlapping regions of interest in an image, transmits the necessary image data using a modified zerotree algorithm and supports the refinement of image parts by transmitting only differential data. As an application for small displays, we present the technique of rectangular fisheye views, which takes advantage of the properties of the wavelet representation.

Keywords: Progressive Image Transmission, Level of Detail, Region of Interest, Fisheye View

1 INTRODUCTION

With the emergence of the World Wide Web, images have become an important means of communicating information in the formerly text-only Internet. Since a substantial fraction of the users connects to the Internet via modems, efficient image transmission methods are crucial for short response times. The desire to let mobile users participate in the Internet leads to the need to cope with even narrower bandwidth and smaller client displays.

Progressive image transmission can help reducing the latency when transmitting raster images over low bandwidth links. Often, a rough approximation (*preview*) of an image is sufficient for the user to decide whether or not it should be transmitted in greater detail. During image *refinement*, the requested greater *level of detail* can quite often be limited to certain regions of the image (*regions of interest*). We distinguish between the refinement methods *detail on demand*, where the user requests greater detail, and *progressive refinement*, where the system transmits and displays more detail automatically. Both methods can be combined, too. In order to save bandwidth, it is essential that only differential data are transmitted.

The contribution of this paper is to provide a consistent view on the problem of levels of detail and regions of interest. We propose a combined method, which allows the specification of the desired degree of detail for overlapping regions of an image, supports the transmission of the necessary image data and permits the refinement of the image by transmitting only differential data.

2 RELATED WORK

Progressive image transmission has been studied by many authors, and many different methods have been developed. In works of the 70s and 80s (see, e.g., [2], [4], [12]), methods for the progressive transmission of multiresolution images were used to support graphics terminals connected to mainframes via slow links, and already in 1983, Lohscheller [6] used progressive DCT in an image retrieval system. With the emergence of multimedia, focus has been shifted towards image compression methods. As the WWW gained popularity in the 90s, interest in the progressive transmission of images grew again.

Many of today's image compression methods have „built in" a more or less sophisticated progressive transmission mechanism, including simple row interlacing (GIF), scan-based (partially-embedded) progressive encoding (JPEG), and fully-embedded entropy coding (e.g., Zerotrees [9] or SPIHT [10]). The recently proposed Internet Imaging Protocol [5] supports structured access to the FlashPix image data representation, which consists of independent resolution layers, but does not offer differential transmission. Some existing wavelet coders support variable quantization of different regions (Summus WI format [13]) or refinement of the whole image (LuRaWave [7]). A number of adaptive ([3], [8], [14]) and progressive ([1], [5], [6]) image transmission systems have been proposed recently in the literature, which address different parts of the *level-of-detail / region-of-interest* problem.

3 LEVELS OF DETAIL AND REGIONS OF INTEREST

We define a *local level of detail* (local LoD) to be a tuple consisting of *x resolution, y resolution, color,* and *precision.* These dimensions define a vector space (see Fig. 1), in which each level of detail can be interpreted as an individual vector. *Refinement* of an already transmitted local LoD can be realized by retrieving the data corresponding to the difference vector between the local LoD transmitted and the new one requested. A *region of interest* (RoI) is defined as a connected set of pixels in the image (called the *footprint* of the RoI), with a local LoD associated. A set of RoIs (whose footprints may overlap) with one marked *reference RoI* is called a *global LoD*. To permit the automatic scheduling of the transmission of multiple RoIs in a global LoD, each RoI can be prioritized with a *constraint vector*. This vector specifies constraints for the components of the RoI's local LoD relative to that of the reference RoI. These constraints, which are maintained during transmission, could be for example that „the resolution of a certain RoI should always be twice the resolution of the reference RoI".

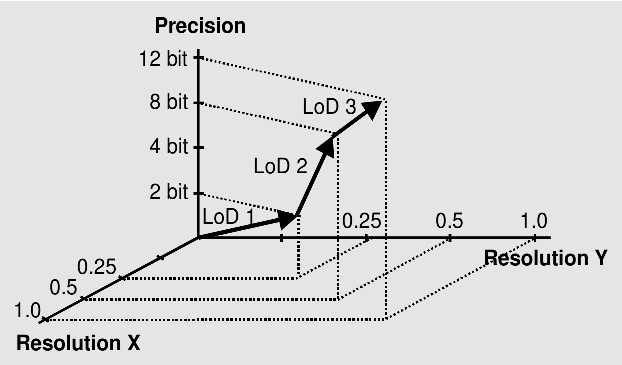

FIGURE 1: AN EXAMPLE LOD VECTOR SPACE

For computing the difference between an old (already transmitted) global LoD and a new one, the set difference of the footprints of the old and the new RoIs is computed, and for these pixel sets the difference vectors of the associated local LoDs are calculated. Since negative difference vector components would mean to discard already transmitted data, any such component is set to zero. Only these difference vectors have to be transmitted.

4 REALIZATION

We are extending the Embedded Zerotree Wavelet developed by Shapiro [9] to implement the concepts outlined above. Wavelets support refinement in the dimensions *resolution* and *precision*, each in steps of powers of two. In every level of detail, the resolutions in x and y direction can differ by a factor of two if desired. The dimension *color* can only have two values - grayscale or color. Shapiro's zerotree method efficiently codes wavelet coefficients by representing a tree of coefficients with the value zero and the same spatial orientation as one *zerotree symbol*. The coefficient array is scanned bitplane by bitplane and in the direction from low frequency subbands to high frequency subbands such that in each tree of coefficients no children are scanned before their parents. Because this traversal order maintains the spatial coherence of the coefficient array, it is very well suited for incorporating regions of interest. The scanning order „bitplane by bitplane and parents before children" can be exploited for the differential transmission of local LoDs by constraining it to the intervals corresponding to the LoD to be transmitted.

By adding RoI schedulers to both encoder and decoder, we have adapted the zerotree method to suit the needs of regions of interest. The schedulers know the sequence of global LoDs as side information, maintain for each RoI the local LoD already transmitted and decide depending on their priority, which bitplane and subbands to transmit/receive next for which regions of interest. Since this mechanism can select the same bitplane multiple times for transmission, we must ensure that each bit in the coefficient array is transmitted at most once and that parent coefficients are always transmitted before their children.

To do this, we compute for each footprint a *multiresolution hierarchy* corresponding to the wavelet decomposition (see Fig. 2b). For each polygon in the hierarchy, it must hold that it is completely inside its parent polygon, the latter scaled by a factor of two. This can be achieved by using the following algorithm for recursively building the polygon hierarchy from the original footprint P_0:

1. Initialize the grid step with 2. Do for each decomposition level $0<i\leq l$:
2. Initialize polygon P_i with a copy of P_{i-1}. Compute the outward normals for each vertex (cf. RoI B in Fig. 2a).
3. Using a modified Bresenham algorithm, move each vertex of P_i in the direction of its outward normal to the next point whose coordinates are multiples of the grid step. Repeat this until the adjacent edges of the moved vertex do not intersect any of the adjacent edges of the original vertex.
4. Double the grid step. If P_i is a self-intersecting polygon, replace it by its convex hull. Go to 2.
5. For each level i and each subband s, compute $P_{i,s}$ by downscaling P_i by the appropriate grid step and translating it relative to the subband's origin.

Let be b the bitplane, i the decomposition level and s the subband to be transmitted for a set of RoIs R selected by the scheduler. The footprints $P_{i,s}$ of the RoIs $r \in R$ are inserted into the *set of positive footprints* F^+, if r has not yet been transmitted at b, i, s; or into the *negative footprint set* F^-, if r has already been transmitted at b, i, s. The scaled footprints $P_{i,s}$ of all RoIs $r \notin R$, which have already been transmitted at b, i, s, are also inserted into F^-. Now, the subband s is traversed by using a modified *scanline fill algorithm*, which computes a set $I_{i,s}$ of spans of coefficients being inside at least one footprint from F^+, but outside all footprints from F^-. The bits at bitplane b of all coefficients on these spans are zerotree coded and transmitted. Since self-intersecting polygons present problems with scanline fill algorithms, we had to replace them by their convex hull in step 4. At the decoder side, inverse wavelet transform and display update after receiving new data can be restricted to the bounding rectangles of the RoIs for which new data have been received.

Fig. 2 gives an example illustrating the concepts presented above. In the well-known „boats" image, two RoIs have been transmitted: RoI A with the whole image as footprint at a local LoD of 1/16 the resolution and full precision, and RoI B with a footprint restricted to the foreground boat at full resolution and full precision. A new RoI C has been specified by the user. In figure 2b, the resulting footprint multiresolution hierarchy and its transmission status is shown. Note that the footprint of RoI B has been replaced by its convex hull in the lower resolution subbands because of the self-intersection marked as „scaling problem“ in Fig. 2a, where also the outward normals can be seen. A dark shade of gray means that a coefficient has been transmitted at full precision, light gray means the coefficient has not been transmitted at full precision but the desired local LoD has been reached, and white means that a bit of the coefficient will be transmitted in the current pass. The coefficients of the low-low subband have been fully transmitted previously in RoI A. That's why no data have to be transmitted in the lowest subband for refining RoI C, and $I_{5,0}$ is empty. For all the other subbands, $I_{i,s}$ is non-empty. The corresponding areas are shaded white in Fig. 2b. Note that no data need to be transmitted for coefficients being in the footprints of both RoI C and RoI B.

5 APPLICATIONS

The specification and image transmission method presented above can be applied to reduce the transmission time of images in low bandwidth environments. Before storing an image on the server, its author can prepare RoIs for important parts which will be refined first when the image is transmitted. At the client side, the user may select arbitrary regions of the image for refinement which may overlap the RoIs specified by the author without having to retransmit already received data. Furthermore, an image browser can be implemented which supports zooming and panning through remotely stored large images.

As an image browsing application for mobile computers, we developed a presentation technique which we call *rectangular fish eye view*. This screen real-estate saving technique combines focus and context in one presentation and can be seen as a special case of the *rubber sheets* method by Sarkar et al. [11]. The viewport is split into rectangular parts which show non-overlapping image regions subsampled in x and/or y direction by different powers of two (see figure 3a). In the center rectangle, the *focal region* of the image is displayed at full resolution. The grid in Fig. 3a is generated by our method automatically if five RoIs are specified:

- The whole image with (4,4) subsampling.
- A vertical and a horizontal stripe (delimited by dashed lines in Fig. 3a) with (2,∞) respectively (∞,2) subsampling. The missing value is computed from the overlapped RoI during vector difference calculation.
- A vertical and a horizontal stripe (delimited by dotted lines in Fig. 3a) with (1,∞) resp. (∞,1) subsampling.

This specification is sufficient to automatically create full resolution in the focal region. However, specifying a sixth RoI for the focal region with high priority assigned makes sure that the focal RoI is transmitted with precedence. The user can pan the focal region, causing the system to load the difference between the low resolution already transmitted and the full resolution requested. A wavelet representation of the image naturally supports this browsing method, because the subbands it contains differ in frequency by a factor of two. The inverse wavelet transform only needs to recover the focal region to full detail, in all other image parts it can terminate earlier. Figure 3b shows the result. The original 1200x1544 pixel image has been reduced to 600x772 pixels with a focal region of 300x385 pixels.

6 CONCLUSION

We have presented a method that extends the Embedded Zerotree Wavelet progressive image transmission method by levels of detail and regions of interest. Our method supports the specification of RoIs with arbitrarily shaped, non-self-intersecting, possibly overlapping footprints and associated local LoDs. RoIs can be prioritized. A scheduler and a modified scanline fill algorithm are used for controlling the transmission process. For the refinement of image parts, only differential data are transmitted. Our method can speed up image transmission in low bandwidth environments. Some applications have been sketched. The focus-and-context technique of rectangular fish eye views, which saves transmission bandwidth as well as screen real estate, has been discussed.

ACKNOWLEDGEMENT

The author wishes to thank Prof. Heidrun Schumann for encouragement and Tilo Strutz for fruitful discussions about wavelets. Special thanks to the Warnow Regional Public Transport Board and the Nitschke Advertising Studios for permitting the use of their map image data. The work presented is supported by the German Science Foundation under contract no. Schu-887/3-1.

REFERENCES

[1] M.J. Dürst and T.L. Kunii, Progressive transmission increasing both spatial and gray scale resolution, *Proc. Int. Conf. on Multimedia Information Systems '91* (Singapore: MacGraw-Hill, 1991), 175-186.

[2] H.M. Dreizen, Content-driven progressive transmission of gray-scale images, *IEEE Trans. on Comm.,* 35(3), 1987, 289-296.

[3] A. Fox and E.A. Brewer, Reducing WWW Latency and Bandwidth Requirements by Real-Time Distillation, *Proc. 5th Int. WWW Conference,* Paris, France, 1996.

[4] K. Knowlton, Progressive Transmission of Gray-Scale and Binary Pictures by Simple, Efficient, and Lossless Encoding Schemes, *Proc. of the IEEE*, 68(7), 1980, 885-896.

[5] H.J. Lee, High-resolution graphics with FlashPix and the PsiFi Internet Imaging Protocol, *Web Techniques*, 12/96, Online Magazine, http://www.webreview.com/97/06/06/feature/

[6] H. Lohscheller, A Subjectively Adapted Image Communication System, *IEEE Trans.Comm.,* 32(12), 1984, 1316 ff.

[7] *LuRaWave image compression*, http://www.luratech.com.

[8] U. Rauschenbach and H. Schumann: Adaptive Image Transmission, *Proc. 5th Int. Conf. in Central Europe on Computer Graphics and Visualization 97*, Plzen, Czech Republic, 1997.

[9] J.M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, *IEEE Trans. Signal Processing*, 41(12), 1993, 3445-3462.

[10] 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.

[11] 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.

[12] S.L. Tanimoto, Image Transmission with gross information first, *Computer Graphics and Image Processing,* 9(1), 1979, 72-76.

[13] *Summus wavelet image format,* http://www.summus.com.

[14] J. Zhang, Medium Support in Mobile Computing, *Proc. Conf. on Distributed Multimedia Systems,* Vancouver, Canada, 1997.

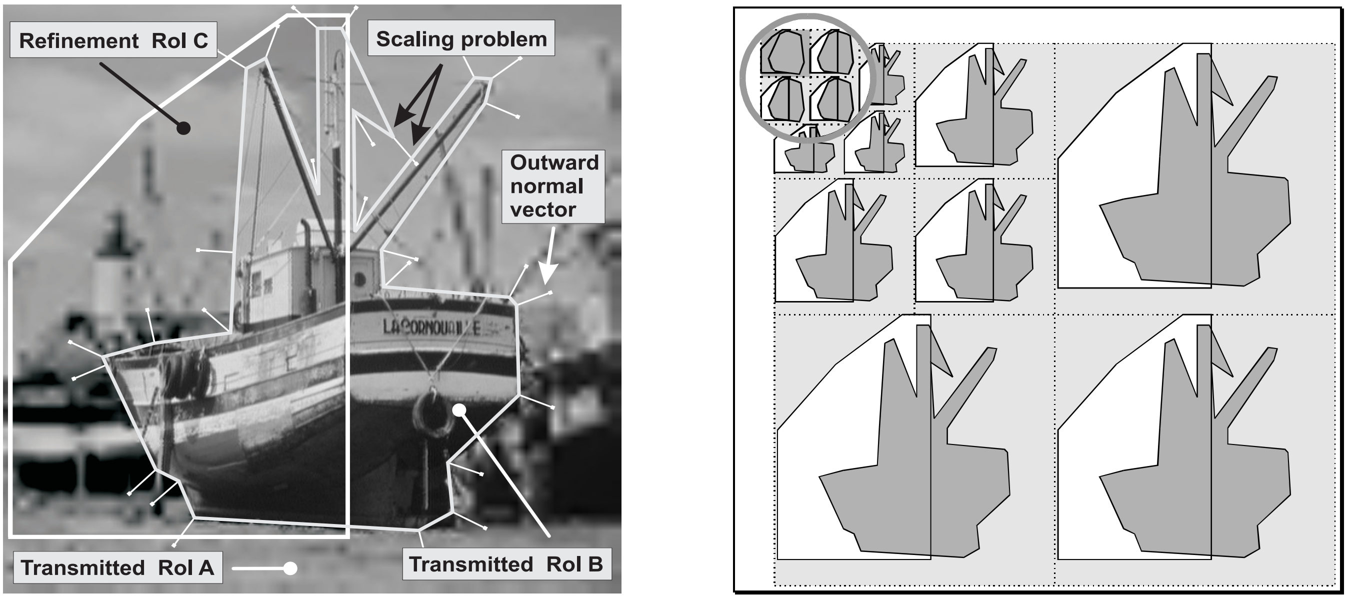

FIGURE 2: A) THREE ROIS IN THE „BOATS“ TEST IMAGE. **B)** FOOTPRINT POLYGON HIERARCHY

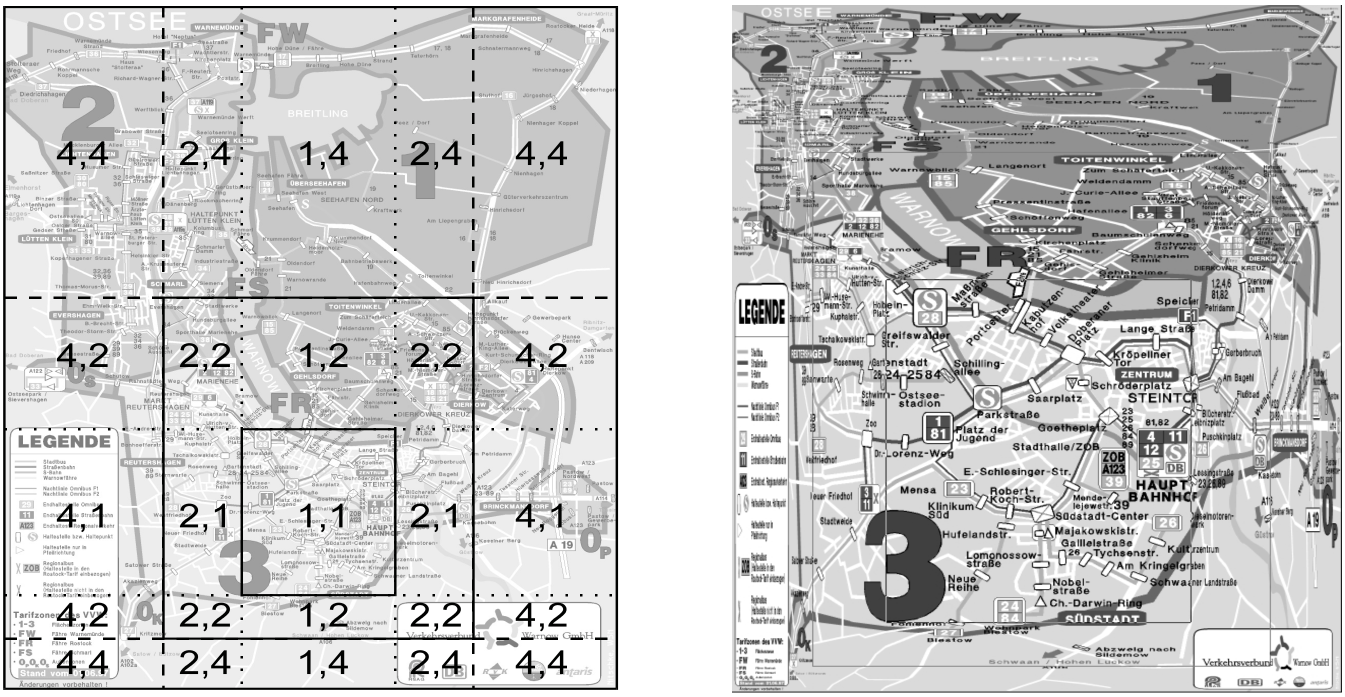

FIGURE 3: RECTANGULAR FISHEYE VIEW OF THE ROSTOCK PUBLIC TRANSPORT MAP. **A)** SUBSAMPLING GRID (X,Y). **B)** RESULT.