A FLEXIBLE POLYGON REPRESENTATION OF MULTIPLE OVERLAPPING REGIONS OF INTEREST FOR WAVELET-BASED IMAGE CODING

Uwe Rauschenbach[*]

SIEMENS AG
Corporate Technology CT IC 2
Otto-Hahn-Ring 6
D–81730 München
Germany
uwe.rauschenbach@mchp.siemens.de

René Rosenbaum, Heidrun Schumann[†]

University of Rostock
Computer Science Department
Institute of Computer Graphics
D–18051 Rostock
Germany
{rrosen,schumann}@informatik.uni-rostock.de

ABSTRACT

Image transmission over low-bandwidth channels can be speeded up if the image coding mechanism supports regions of interest (RoIs). By such a scheme, image parts not belonging to RoIs can be encoded at a lower bitrate. This paper describes a flexible dynamic RoI scheme which supports the definition of arbitrarily-shaped RoIs before the start of the transmission process and the refinement of a partially transmitted image by the definition of new RoIs at any time during image transmission. The overhead to represent RoIs is kept as small as possible by exploiting a polygon-based RoI representation. Span arithmetic on a multiresolution grid is used to support the redundancy-free transmission of overlapping RoIs.

1. INTRODUCTION AND RELATED WORK

Wireless networks like GSM allow ubiquitous access to image content. Since the bandwidth offered by such networks is very limited, flexible image coding techniques have to be used which support the encoding of regions of interest (RoIs). By such a scheme, image parts not belonging to RoIs can be encoded at a lower bitrate. If the image is encoded on demand at transmission time, the receiving user can control the transmission by specifying RoIs according to his special interest.

Thus, a flexible RoI scheme should support the definition of arbitrarily-shaped RoIs prior to the start of the transmission process and the refinement of a partially transmitted image by the definition of new RoIs at any time during or after partial image transmission. The overhead to represent RoIs must be kept as small as possible, and no redundant data should be transmitted if the user at the receiving end specifies a new RoI to refine an image.

In many existing wavelet-based image coding schemes, RoIs are either restricted to be non-overlapping and of simple shape (rectangular or circular) which is represented geometrically [3, 11, 12, 13], or more complex shapes and possibly overlapping RoIs are stored using bitmaps which is memory-consuming [1, 2, 5, 7, 10]. The new image coding standard JPEG2000 [6] supports implicit RoI coding without overhead by scaling those wavelet coefficients which contribute to the RoI. Only some of the methods with simple RoI geometry [1, 2, 5] support the *dynamic* definition of new RoIs in the course of a running image transmission. JPEG 2000 and the schemes described in [3, 7, 10, 12, 11] allow only the definition of *static* RoIs prior to encoding.

In this paper, we address these problems by proposing a new method. Polygonal shapes are used to represent RoIs geometrically, and local levels of detail combined with span arithmetic on a multiresolution grid support the redundancy-free transmission of overlapping RoIs in a wavelet-based image coding system. The paper is structured as follows: After discussing the basic idea of the method, we will describe the algorithm for the span computation. We briefly present an example application. Finally, we conclude with directions for future work.

2. BASIC IDEA OF THE NEW METHOD

As the foundation of our method, we use an EZW coder similar to the one proposed by Shapiro [9]. This encoder generates an embedded data stream by traversing the wavelet coefficient array in two passes per bitplane. Within each bitplane, the coefficients of each subband are visited in scanline order. In the field of computer graphics, span-based algorithms are used for visibility computations during scanline order rendering of 3D scenes (cf. [4]). We modify this idea to support RoIs in wavelet-based image coding.

2.1. RoIs and local LoDs

Each RoI is represented as a polygonal area, called the *footprint*, and a *local level of detail (LoD)*. The LoD stores the *transmission target* and the *transmission state* which are both sets of flags representing all possible combinations of color channel, bitplane and coefficient subband. By means of the transmission target, the user can specify which bits of which coefficients shall be transmitted. Analogously, the system stores information about the already transmitted data in the transmission state. Such, transmission state information can be stored per RoI instead per wavelet coefficient to save memory. The transmission of a RoI is finished if transmission target and transmission state are equal. A priority can be assigned to each RoI which is used as a bitplane offset. This offset controls how many coefficient bitplanes must be transmitted for the according RoI before the highest coefficient bitplane of other RoIs with a priority of zero can be transmitted.

2.2. Handling overlapping RoIs

The dynamic definition of arbitrarily-shaped RoIs often leads to overlapping RoIs in practical applications. In this case, redundant transmission of coefficients belonging to multiple RoIs must be avoided. Overlapping RoIs split the set of wavelet coefficients into *partitioning subsets*, each subset containing the coefficients belonging to a particular set of RoIs.

2.3. Scanline discretization on a multiresolution grid

The original scanline algorithm [4] is used in computer graphics to solve the problem of discretization in raster images. Here, we adopt it to describe the shape of RoIs and to efficiently handle the common problem of overlapping and polygonal RoIs.

The original EZW algorithm traverses the wavelet coefficients within one subband in scanline order. RoIs can be integrated by restricting traversal to the coefficients which contribute to the footprints of the RoIs. Since the basic EZW algorithm exploits inter-band relationships by defining a zerotree structure, we must make sure that no child coefficients in a zerotree are transmitted before their parent coefficients. To ensure this property, we control the coefficient array traversal by discretizing the RoI footprints on a *multiresolution grid.*

The wavelet subband array defines a multiresolution grid on the image. For each decomposition level k, we can scale the RoI footprint by 2^{-k}. Each coefficient in a subband at level k contributes to 2^k x 2^k pixels in the image. We will call such a grid cell corresponding to a wavelet coefficient at a particular level a *superpixel.* Figure 1 illustrates the idea of a multiresolution grid.
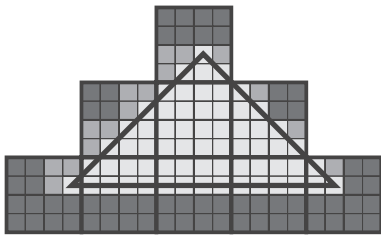

Fig. 1. Three-level multiresolution representation of a triangle.

In the proposed solution, a coefficient contributes to a RoI if the footprint of that RoI intersects the superpixel corresponding to the coefficient. This differs from the original scanline algorithm, which just approximates the shape of an object.

We can now traverse all superpixels covered by a particular RoI scanline by scanline and model the partitioning subsets by using *spans* of superpixels. Such a span is a sequence of consecutive superpixels on a scanline which contribute to the same set of RoIs. A span may be started or terminated only at superpixels intersected by an edge of a RoI.

To avoid redundancy, we want to transmit each bit of each wavelet coefficient at most once. That means, when encoding a coefficient of a RoI, we must check if the current bit of this coefficient has already been transmitted by a RoI encoded earlier. We do this by checking the flags in the transmission state of all RoIs which overlap the superpixel corresponding to the coefficient. This check has to be done once per span only.

3. DESCRIPTION OF THE ALGORITHM

3.1. Transmission steps

The basic algorithm is embedded into an image transmission framework described in greater detail in [8]. The encoding and decoding processes are controlled by *RoI schedulers* running synchronously. They perform a sequence of *transmission steps* which terminates if for all RoIs the transmission target is equal to the transmission state. For each such step, the schedulers select a RoI r_0 as *active* and determine the color channel, bitplane and subband to be encoded resp. decoded next.

Prior to a transmission step, a new RoI with associated transmission target may be defined by the client system. The footprint must be a regular polygon, such that all edges can be classified in a preprocess as *opener* (which means that the area left of the edge does not belong to the polygon but the area right of the edge does), *closer* (which means that the area left of the edge belongs to the polygon but the area right of the edge doesn't) or *horizontal*.

Each transmission step considers the area covered by the active RoI r_0 and all other RoIs which intersect r_0 and have already transmitted the current color channel, bitplane and subband according to their transmission state. The bounding box of this area is traversed scanline by scanline where the scanlines consist of superpixels of the according size. Span computation is performed for each scanline, considering all edges intersecting that scanline. For each span, we check the transmission state of all RoIs to which this span contributes. If there is no RoI with the corresponding flag set, we encode resp. decode the current bit of the wavelet coefficients on the span.

Having finished the transmission step, we set the according flag in the transmission state of the current RoI.

3.2. Span computation

We will now describe in detail how the spans are computed. I.e., we now consider one scanline. All polygon edges which intersect the current scanline are stored in a data structure called the *active edge table (AET)*. This structure is updated from scanline to scanline to ensure that it contains all valid edges. Each edge in the AET holds the intersection coordinate X_S of that edge with the current scanline. A fast incremental update of X_S from one scanline to another is supported by separately storing the edge slope, the whole-numbered and the fractional component of X_S as three integers. The AET is sorted in ascending X_S order.

For each edge in the AET, the algorithm checks whether this edge opens or closes a span. Furthermore, two counters are maintained by evaluating the transmission states of the RoIs overlapped by the spans. The counter C_{neg} contains the number of RoIs which did already transmit data for the current channel/bitplane/subband combination, and C_{pos} stores the number of RoIs which didn't. If a span is opened or closed, the algorithm evaluates the counters in order to decide whether the coefficients on the span have to be encoded resp. decoded. This is the case for spans with $C_{neg} = 0$ and $C_{pos} > 0$.

Because of the multiresolution grid, two or more edges may intersect the same grid cell. To handle this case correctly, the edges which intersect a particular grid cell are evaluated using a multi-pass algorithm. The first pass considers only edges classified as *opener*, the second one edges classified as *closer*. For concave polygons, a third pass must handle the special case that two *opener*-edges of the same polygon intersect the current grid cell.

These three passes are sufficient to handle regular concave polygons, i.e. polygons in which the vertices are the only points where two edges of the polygon intersect. This property is checked for all resolution levels each time a new RoI is defined by the user. If the regularity test fails, the polygon is replaced by its convex hull.

3.3. Intersection point computation

The correct computation of the intersection coordinate X_S plays a crucial role in maintaining the parent-child relationship in the zerotrees. It is important to decide whether the intersection of the edge with the upper or the lower superpixel boundary is used. Figure 2 shows the four possible cases.
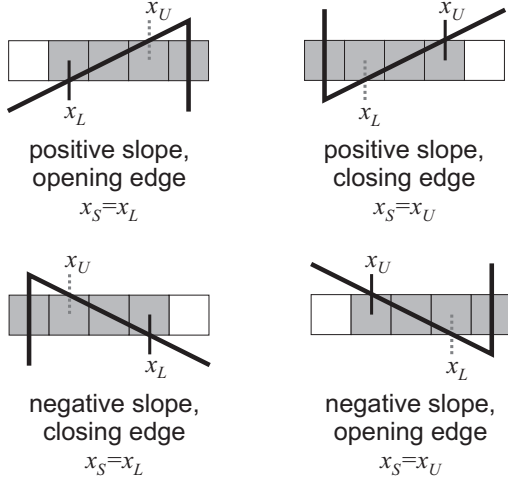

Fig. 2. Computing the intersection x coordinate X_S with a non-vertical edge.

When computing the intersection coordinate at joins of two edges, we assume one of the edges as vertical for simplification and compute two X-values X_i ($i \in \{1, 2\}$). X_1 is computed as depicted in figure 3, assuming the lower edge as vertical. X_2 is computed analogously, assuming the upper edge as vertical and modifying figure 3 as follows: both edges are mirrored about the X-axis and "positive" resp. "negative" slope are swapped. As the resulting value for X_S, we use the minimum of X_1 and X_2 (if both edges are openers) or the maximum of X_1 and X_2 (if both edges are closers).

Horizontal edges require special handling. Since we are using a multiresolution grid, non-horizontal edges may become horizontal at coarser grid resolutions. We refer to these edges as *collapsed edges*. Both horizontal and collapsed edges are handled by temporarily replacing them by zero, one or two auxiliary vertical edges which intersect the current scanline only, in contrast to the classical scanline algorithm, where they are ignored. Figure 4 shows the possible cases.

4. AN APPLICATION

Based on the proposed algorithm, we have realized a system which supports both *image author* and *image viewer* in low-bandwidth environments. The *authoring tool* can be used by the image author to prepare an image for transmission. It allows the flexible definition of arbitrarily-shaped RoIs, offers control over the transmission
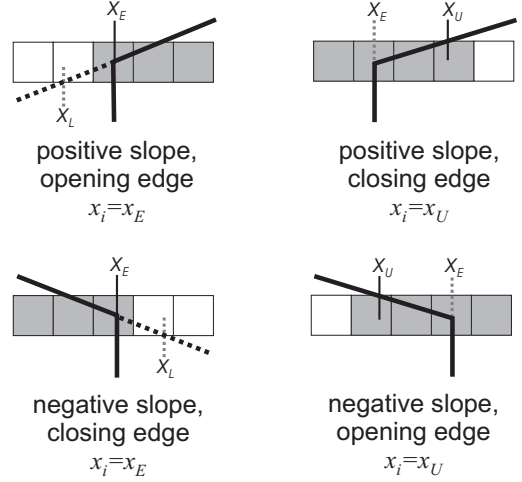

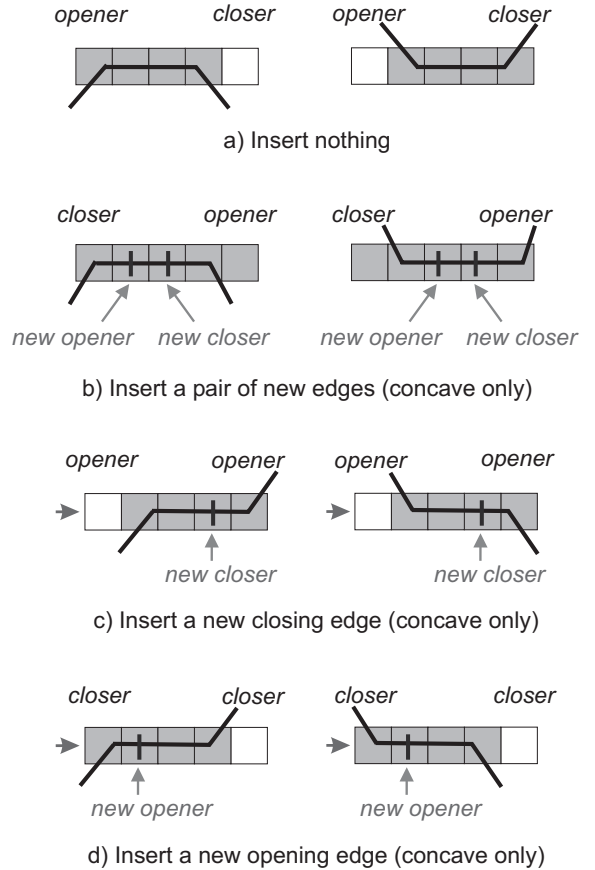**Fig. 3**. Computing the intersection x coordinate X_S at joins.

Fig. 4. Handling a collapsed edge by inserting auxiliary vertical edges for the current scanline.

target and contains a transmission simulation module. Figure 5 (left) shows that the author has defined two overlapping RoIs: the background at low resolution and priority and the fishing boat as the main feature of the image at full resolution and full coefficient
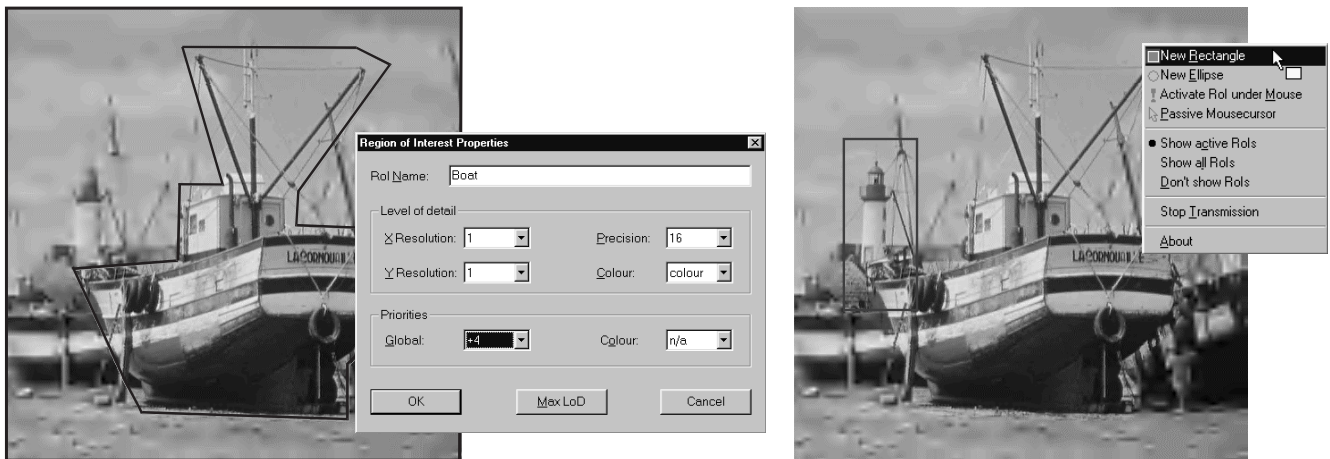
Fig. 5. Support for image author (left) and image viewer (right) using the proposed scheme.

precision. Additionally, the fishing boat has been assigned a higher priority. The simulation shows the degree of detail of the image after 9 seconds transmission time over a simulated GSM connection.

The image viewer who downloads the image may, however, be interested in another part of the image (e.g., the lighthouse). To provide RoI functionality to the viewer, a web browser plugin has been developed. This *viewing tool* allows the easy specification of rectangular or elliptic RoIs[1]. Figure 5 (right) shows that the viewer has selected the lighthouse as his main area of interest, which in part overlaps the fishing boat. By means of the redundancy-free transmission, the overlap is handled automatically. Since the new region of interest is small and only differential data are transmitted, refinement can be performed very fast.

5. CONCLUSION AND FUTURE WORK

In this paper, we have proposed a new method for the incorporation of regions of interest into a wavelet-based image coding scheme. Unlike many existing systems, the proposed scheme allows the *dynamic* definition of regular *polygonal RoIs*. Span-based traversal of the wavelet coefficient array ensures *redundancy-free refinement* of already transmitted images and *saves memory and transmission bandwidth* since no bitmap representation of the RoI shape is needed.

RoI support in the new image coding standard JPEG 2000 is mainly targeted at the archiving of medical images. Future work will concentrate on realizing dynamic RoIs based on JPEG 2000 by transcoding an already encoded data stream.

6. REFERENCES

[1] E. Atsumi and N. Farvardin. Lossy/lossless region-of-interest image coding based on set partitioning in hierarchical trees. In *Proc. IEEE ICIP98*, Chicago, Illinois, October 4-7 1998.

[2] E.C. Chang, C.K. Yap, and T.J. Yen. Realtime visualization of large images over a thinwire. In *Proc. IEEE Visualization*, Phoenix, Arizona, Oct. 19 - 24 1997.

[3] A.T. Duchowski. Representing multiple regions of interest with wavelets. In *Proc. SPIE*, volume 3309, page 975ff., 1997.

[4] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes. *Computer Graphics – Principles and Practice, Second Edition*. Addison-Wesley, 1990.

[5] T. Frajka, P.G. Sherwood, and K. Zeger. Progressive image coding with spatially variable resolution. In *Proc. IEEE ICIP97*, volume 1, page 53ff., Santa Barbara, California, October 1997.

[6] JPEG 2000 image coding system. Committee Draft CD15444-1, Version 1.0, ISO/IEC JTC 1/SC 29/WG 1 (ITU-T SG8), 9 December 1999.

[7] D. Nister and C. Christopoulos. Lossless region of interest with a naturally progressive still image coding algorithm. In *Proc. IEEE ICIP98*, volume 3, pages 856–859, Chicago, Illinois, October 4-7 1998.

[8] U. Rauschenbach and H. Schumann. Demand-driven image transmission with levels of detail and regions of interest. *Computers and Graphics*, 23(6):857–866, December 1999.

[9] J.M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, Dec. 1993.

[10] J.M. Shapiro. Apparatus and method for emphasizing a selected region in the compressed representation of an image. U.S. Patent 5,563,960, October 1996.

[11] D. Shin, H.-H. Wu, and J.-C. Liu. A region of interest (ROI) based wavelet compression scheme for medical images. In *Proc. SPIE*, volume 3031, pages 790–798, 1997.

[12] A. Signoroni and R. Leonardi. Progressive ROI coding and diagnostic quality for medical image compression. In *Proc. SPIE*, volume 3309, pages 674–685, 1997.

[13] T. Yu, N. Lin, S.J. Liu, and A.K. Chan. A region-of-interest based transmission protocol for wavelet-compressed medical images. In *Proc. SPIE*, volume 3078, pages 56–64, 1997.

---

[1]Elliptic RoIs are approximated using a polygon. The definition of RoIs with a more complex shape is possible, but would require complex interactions taking too long to control a viewing tool in the course of a running transmission.