Mobile Image Communication Using JPEG2000

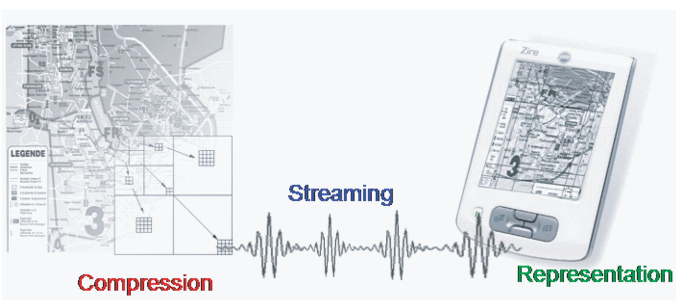
René Rosenbaum, University of Rostock, Albert-Einstein-Str.21, 18057 Rostock, Germany; E-mail: rrosen@informatik.uni-rostock.de
Heidrun Schumann, University of Rostock, Albert-Einstein-Str.21, 18057 Rostock, Germany; E-mail: schumann@informatik.uni-rostock.de

ABSTRACT
Imagery must be adequately processed and adapted to be used in mobile environments. This contribution shows that this can be reached by a rigorous combination of all stages of the image communication pipeline to a single task-aligned system. To achieve this, the modern imaging standard JPEG2000 serves as a foundation for image compression and streaming. Founded on the concept of Regions of Interest and Levels of Detail to describe current demands at client side, powerful generic strategies for the content handling are proposed. The pre-defined task is represented by one of 3 common problems in mobile environments – image browsing, viewer guidance, and content exchange. Beside new ideas to accomplish these tasks on application level, in focus is the strong decrease in the demands for device capabilities by an appropriate image handling. The achieved results show that many resources are saved by an appropriate image communication using JPEG2000.

1. INTRODUCTION
The enthusiasm for mobile computing is still unbroken. With the ability to access information every time and everywhere, the mobile freedom allows for completely new commercial applications. Due to its form factors and the application environment, however, mobile devices are still restricted by *low computing power* and *bandwidth*. As imagery is one of the most important but also most demanding information carriers it must be adequately processed and adapted to be used in such environments (Rosenbaum et al., 2006). How to achieve this by efficient image communication strategies designed to solve a task at hand is shown in this publication. Thereby, the focus is on the achieved results rather than the technological background. However, the required basic technology is briefly explained.

A meaningful image communication strategy consists of the stages *compression, streaming*, and *visual representation* (cf. Figure 1). Although numerous different techniques for each stage have been proposed, an arbitrary combination does usually not lead to an appropriate result. A high performing system can only be guarantied if all stages of the whole communication pipeline are tightly coupled (Rosenbaum & Schumann, 2005). Founded on an appropriate visual representation at the mobile client, this basically means that *only data contributing to the current representation is processed and transmitted*. Although, similar ideas have already been proposed, they are either based on proprietary solutions (Owen et al., 2001) or discuss a specific task only (Deshpande & Zeng, 2001; Ortiz et al., 2004). The ideas proposed in this publication are fully compliant to the international standard JPEG2000 (ISO, 2002), which allows an easy migration of the introduced strategies into existing systems, and due to the generic description of the current demands by Regions of Interest (RoI) and Levels of Detail (LoD) for a broad applicability. To avoid the redundant data transfer in case these demands change during interaction, the principle of a progressive refinement is adopted.

Figure 1. The image communication pipeline



Although the applicability of the proposed JPEG2000-based compression and streaming technology extends far beyond the discussed tasks, this publication focuses on 3 common problems in mobile environments: *Image browsing, Viewer guidance,* and *Content exchange*. Beside the proposal of new ideas for appropriate visual representations and the underlying data handling, of main interest within this publication is their coupling to an appropriate communication system.

This contribution gives an overview about the proposed ideas and achieved results for each single communication stage. Section 2 reviews the idea of RoIs and LoDs and introduces different options for their flexible access within an JPEG2000-compressed image. Section 3 is concerned with a demand-driven streaming of the data based on dynamic RoIs. Section 4 is dedicated to introduce new ideas for the visual representation of the image data depending on the previously mentioned tasks and shows the results of an appropriate image communication. The publication closes with conclusions in Section 5.

2. FLEXIBLE COMPRESSION WITH JPEG2000
Image compression is the foundation for an effective and efficient data handling. Although there are numerous approaches, the JPEG2000 standard represents the current state-of-art and combines a superb compression performance with numerous features. The main important properties regarding a flexible image communication are the modular structure of the data and the ability to support dynamic RoIs.

An RoI is basically a closed pixel area within the image. To determine the current interest, to every RoI is assigned an LoD. Although, there are many approaches to support RoIs within JPEG2000, they usually lack of flexibility to change RoI position or LoD during a running transmission. To achieve this, the concept of dynamic RoIs has been introduced and modeled for JPEG2000 (Rosenbaum & Schumann, 2002). It is based on the encapsulation of the encoded data in independent data containers and supports beside the general options to de- and refine RoIs dynamically, many LoD levels (e.g. gradual resolution, quality, and color). To increase the applicability and performance, different additional concepts, e.g. Limited Spatial Access (LSA) on the data-streams, may be applied (Rosenbaum & Schumann, 2004).

Although JPEG2000 is more complex than its predecessor JPEG, it has been revealed that the strong reduction in the amount of data by the RoI feature heavily decreases this drawback. This can be ensured by producing highly flexible data-streams during encoding.

Result: The result of the compression stage is an appropriately partitioned JPEG2000 image consisting of *modular data containers* able to support the independent reconstruction of single RoIs at the desired LoD. The RoI feature can heavily *decrease the need for computing power and bandwidth*.

3. IMAGE STREAMING WITH DYNAMIC ROIS
Progressive data streaming is an indivisible part of modern image communication. It handles the transfer of the data to the mobile device and basically consists of 3

stages*: Calculation, Sequencing,* and *Signalization* of the compressed data. The following sections discuss the different options for each of these points.

3.1 Calculation

Contrary to pixel domain, the calculation of data parts belonging to a RoI is more complex. As during compression the image is transferred into a scalable representation (ISO, 2002), the content of a pixel region is usually included in multiple data containers. Different approaches to accomplish their calculation have been proposed (Deshpande & Zeng, 2001; Ortiz et al., 2004). However, usually only the spatial dimension of the RoI and less its reconstruction at a reduced LoD is considered. Due to the known structure of the encoded image, this criterion can be fulfilled by a successive selection of RoI containers for each LoD dimension up to the desired LoD (Rosenbaum & Schumann, 2002; Rosenbaum, 2006). By reducing the number of transmitted containers, further bandwidth is saved.

3.2 Sequencing

The goal of sequencing is an interest ordered data transfer by changing the position of data containers within the transmission sequence. This can be achieved by either *global* or *local prioritization*. The global strategy is based on the prioritization between RoIs and thus, allows for a much more flexible adaptability of the streaming process. The basic idea is the assignment of a prioritization value to each RoI (Rauschenbach, 1999). RoIs with values *n* have an advantage of (*n-m*) transmission steps before data of a RoI with prioritization value *m* (*n*>*m*) is included in the sequence. If data from multiple RoIs is to be transmitted, it is interleaved.

The local R-D optimized strategy prioritizes containers belonging to a single RoI. As each element contributes to a pixel region, there might be elements which are truncated at the borders of the RoI. Thus, they do not contribute in the same extent to the refinement of the RoI as fully covered elements. Dependent on the respective overlap, to these elements is internally assigned a smaller priority (Taubman & Rosenbaum, 2003) to achieve a faster quality refinement at client side (cf. Figure 2).

Figure 2. Visual comparison of the reconstruction of a RoI at 6.4kB from the original (left) and R-D optimized sequencing (right).

3.3 Signalization

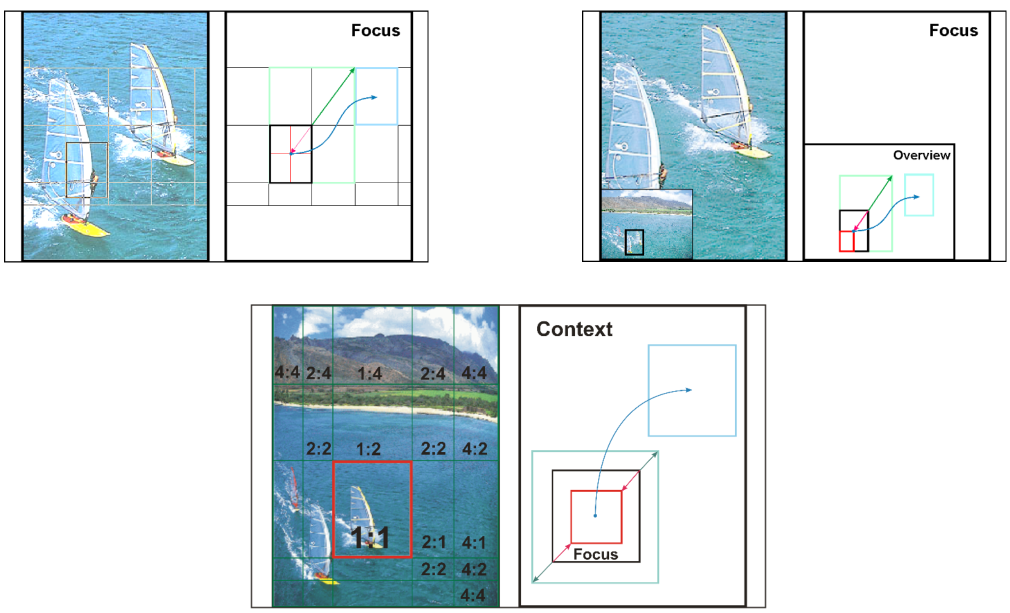
The goal of an appropriate signalization strategy is the proper identification of the transferred data at client side. Due to the reason within *external signalization* to each data chunk is assigned a unique ID, the required bandwidth might significantly increase (cf. to ISO, 2004). Contrary *inherent signalization* identifies each received data element from a predefined progression order. Although inherent in the JPEG2000 codec, the principle must be enhanced to deal with arbitrary container orders as outcome of the sequencing stage. Rosenbaum & Schumann (2002) accomplish this by the un-complex inclusion of new containers within the resulting data-stream. As this also increases the amount of data a strategy for *synchronized signalization* has been proposed by Rosenbaum (2006). Due to the reason server and client running synchronously during transmission, no additional data must be included into the transferred data and much bandwidth can be saved.

Result: A successful streaming of the image delivers the data containers provided by the compression stage in the currently most *appropriate order* to the mobile client. Thereby, the client is able to recognize each container and to process the included information. No more containers than required to reconstruct each RoI at its currently assigned LoD are transferred leading to the effect that much *less processing power and bandwidth* are required.

4. VISUAL REPRESENTATION

Before shown to the user, the received data is usually processed and converted to an appropriate visual representation. This process strongly depends on the respective task and the users needs. The creation of representations used by the 3 interactive tasks: *Mobile image browsing, Viewer guidance*, and *Content exchange*, however,

Figure 3. The image represented by different browsing techniques: Grid-based Zoom&Pan (top/left), largeDetail-View (top/right), and the rectangular FishEye-View with selected RoI scaling values (bottom). Each technique provides different means for interaction (blue: panning, red: zoom in, green: zoom out).



requires strong computing power. In the following, it is shown that the beneficial features of JPEG2000 can be exploited to significantly reduce these needs.

4.1 Mobile Image Browsing

The small screen space of mobile devices imposes a significant usability problem if large imagery is to be displayed. The aim of image browsing is to ease the comprehension of the content by an appropriate representation and means for interaction. Although there are different approaches to achieve this, especially for the strongly limited mobile clients new developments are required. In the next sections, a sophisticated browsing technique for each of the approaches – *Zoom&Pan (ZP)*, *Detail&Overview (DO)*, and *Focus&Context (FC)* – is reviewed and its resource requirements within an image communication system using the introduced streaming technology are discussed. The general approach for any representation is to partition the image in RoIs and to assign well-defined LoD and prioritization values. RoIs which are interesting are assigned a high LoD, RoIs which are not shown the lowest LoD. The prioritization strongly depends on the respective technique and expected behavior. A general taxonomy for image browsing techniques in conjunction with guidelines for their appropriate streaming can be found in (Rosenbaum & Schumann, 2005).

4.1.1 Grid-Based Zoom&Pan (GbZP)

The main disadvantage of many ZP-based techniques is that no context to the shown image part is provided. This is overcome within GbZP by a navigation grid showing information to the currently displayed region and the whole image (cf. Figure 3). This grid also serves as a means for interaction - another drawback of ZP-based techniques. The traditional slider paradigm and menu-based zooming is replaced by intuitive grid interaction. Each grid cell represents a different image region and corresponds in its dimensions to the displayed region. The grid outline represents the whole image. The information seeking mantra of Shneiderman (1996) -*Overview first, zoom and filter, then details-on-demand*- can be accomplished by 3 fast interactions only. More details to GbZP and the design of the belonging image communication system can be found in (Rosenbaum & Schumann, 2005).

Results: As only a certain image region is shown, creating a GbZP representation is *un-complex*. Furthermore, only data for the currently shown region is transferred. Consequentially, bandwidth requirements strongly correlate with the ratio between this region and the whole image. The *save of bandwidth* for a representation requiring 10% of the space consumed by the full-detail image is *approximately 90%*.

4.1.2 The largeDetail-View (lDV)

To display relevant micro and macro information at the same time, the largeDetail-View provides in additional overview to the whole image (cf. Figure 3). Although, the DO principle is well-known in the area of Information Visualization, it can also be applied for image browsing (Karstens et al., 2004; Rosenbaum, 2006). Within the overview, a viewfinder shows the currently displayed region with regard to the whole image. To concentrate the means for interaction to a particular structure, all browsing tasks are accomplished by a modification of the viewfinder. Though, the effective exploration requires many cognitive switches between the two views, the permanently shown overview eases the orientation within the image significantly. To explore contents covered by the overview, it may be displaced or hidden. An image communication system tailored to the lDV is described by Rosenbaum (2006).

Results: Due to the additional scaling operations needed to create the additional overview, the lDV requires *slightly more computing power* than ZP-based techniques. The same applies for the transmitted data. Due to the very low detailed overview, however, the *additional amount of data is little* and might be even negligible if content covered by the overview is omitted during transmission.

4.1.3 The rectangular FishEye-View (rFEV)

The rectangular FishEye-View has been introduced by Rauschenbach (1999). It is founded on the lens metaphor, and as shown by Karstens et al. (2004) allows for an intuitive user interaction . The consistent embedding of the focus is achieved by a complex belt partition of the background (cf. Figure 3). Belts close to the focus are less scaled than belts at the image borders. Thereby, a single belt is formed by many RoIs with different scaling values. Such a configuration is high-demanding for an appropriate image communication, but can still be handled by the introduced technology. More details to the design of the belonging communication system can be found (Taubman & Rosenbaum, 2003).

Results: The rFEV is very intuitive, but also *much more complex* during the creation of the belonging visual representation than other techniques. Much scaling operations are required to smoothly embed the focus. Regarding the requirements for bandwidth *more resources are required* as for GbZP or lDV. However, there is still much data which can be omitted during data transfer. If the ratio between the image dimensions and the available screen space is 4:1, approximately 70% of the data is negligible.

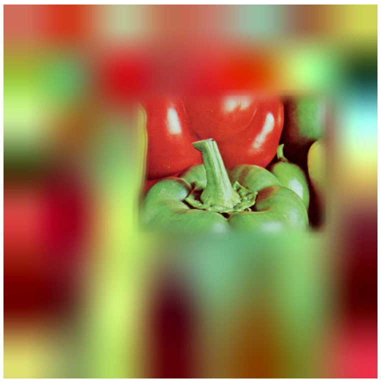
4.2 Viewer Guidance for Raster Imagery

To quickly extract relevant information from large images is difficult. The aim of viewer guidance is to intuitively direct the attention of the viewer to this information. Contrary to image browsing the image is represented in its full dimensions, but usually enriched with additional contents as frames or annotations. Due to the fact, these structures are usually permanently combined with the content, however, this has the significant drawback that the highlighted region may not be changed. To overcome this, two paradigms from Information Visualization - *Depth of Field* and *Tool glasses* - have been adopted to raster imagery.

4.2.1 Depth of Field

The effect aroused by embedding a highly detailed focus region into a blurred background is called Depth of Field (DoF) (cf. Figure 4/left). As humans do

Figure 4. The depth of field effect (left) and tool glass approach (right) applied to still imagery by blurring the background or removing texture from the current focus



not like to look at blurred objects, it is a very effective means of guiding the attention of the user to the focus. However, the demands of DoF representations often exceed the abilities of a mobile device. Due to this, new ideas for their creation and transmission have been introduced by the authors (2006). They are mainly based on the Discrete Wavelet Transform (DWT) as integral part of the JPEG2000 codec. As the DWT inherently separates detail and approximation of the image, the later can be used for blurring. Thereby, the modular structure of the data-stream allows the removal of the details without to decode the image. This also applies for an additional decrease in the quality of the reconstructed signal to advance the blurring effect. Even such complex image processing operations can be supported by an appropriate image handling (Rosenbaum & Schumann, 2006; Rosenbaum, 2006).

Results: All operations required to create a DoF representation of the content are accomplished in compression domain. Thus, the implementation is more than *5 times faster* compared to traditional pixel based processing. As the blurring of the large background regions is accomplished by removing data containers, the coupling with the proposed streaming technology leads to a *tremendous save of bandwidth*,. In case the focus covers 1/10th of the image, almost *90% of all available data can be skipped* during transmission.

4.2.2 Tool Glasses

The main idea of tool glasses is to reduce the content shown within the focus. Due to the resulting inconsistencies, the embedded focus immediately attracts attention. The reduction of presented information within edge or detail glasses, however, is also a great means for content exploration. In any event, the extraction of details requires much computing power. To extent the application area for such tools to mobile environments, new ideas are required. To achieve this, Rosenbaum & Schumann (2006) again take advantage of properties of the DWT. To all contents belonging to the focus, the approximation of the image is removed and replaced by more suitable data without to decode the image. This leads to a representation as shown in Figure 4/right, and is an proceeding can be fully supported by an appropriate communication system (see also (Rosenbaum, 2006)).

Results: Working in compression domain is *un-complex*. The proposed texture removal strategy is up to *20 times faster* compared to pixel-based techniques, and thus, is rather appropriate for mobile devices. During interactive exploration, the *modification* of the new focus region can be achieved *in real-time* on current mobile hardware. Due to the fact, the large background regions are shown in full detail, the *save of bandwidth is little.*

4.3 Content Exchange Between Imagery

The exchange of content between two or multiple images in consequence of modifications is a frequent task in mobile environments. Especially the often changing data in commercial information systems requires efficient mechanisms for remote content updates as achieved by non-interactive image editing functionality. As moving imagery also helps creative, editorial and business professionals to create powerful communications, the relevance of Motion-JPEG2000 for commercial applications is similar. However, data exchange is easy to accomplish in pixel domain, but requires much resources if the content is only be available in compressed representation. To overcome this, the content might directly be processed in this form. Rosenbaum & Taubman (2003) and Rosenbaum (2006) present a scheme for an exchange of JPEG2000-compressed content. After briefly describing the underlying idea, two of its main applications are explained in more detail.

4.3.1 Content Exchange in JPEG2000 Domain

Content exchange in compression domain requires the access of all encoded data contributing to the reconstruction of a closed image region directly on the data-stream. The crucial challenge here is the multi-resolution representation of the content. This is considered by the introduced Limited Spatial Access (cf. Section 2) leading to an appropriate partition of the image during compression. If this principle is applied to a source and a destination image, corresponding data containers can simply be exchanged without to violate the compliance of the resulting data-stream (cf. Figure 5/left). By taking advantage of the multi-resolution property, it is even possible to reduce the source contents in their spatial dimensions (cf. Figure 5/right). The decoding of the modified data-stream leads to a reconstructed image with contents from both images. If a remote scenario with frequent content changes at server side is considered, this idea can easily be applied for an efficient data streaming. Instead of transferring the whole modified image, only data belonging to the altered regions is transmitted and merged at client side with static contents.

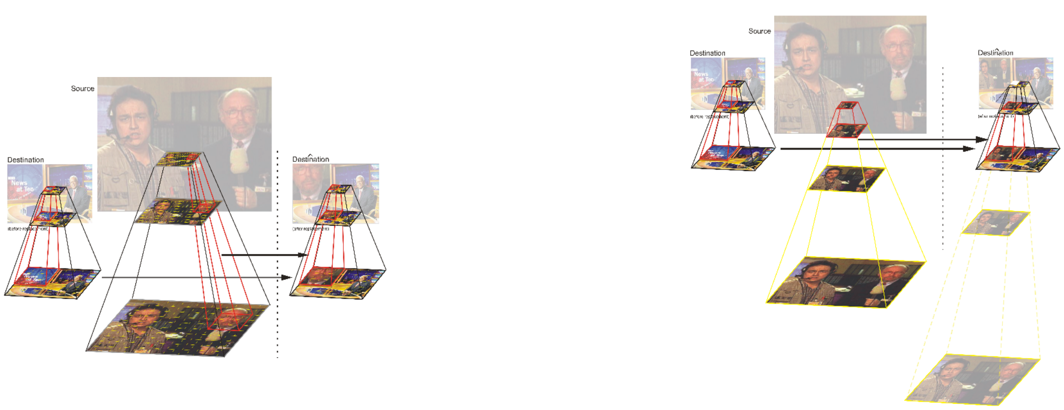
Results: Working in JPEG2000 domain *strongly decreases the complexity* of the exchange task for encoded imagery. As no decoding is required, the proposed strategy is *9 times faster* than pixel based processing. The performance gain even increases if the result has to be in encoded form for transmission or storage. The save of bandwidth strongly correlates with the ratio between the modified region and the whole image. If it covers a 1/10th of the image, 90% of all available data may be neglected.

4.3.2 An Image Editing Framework

Although content exchange in JPEG2000 domain is un-complex, all data containers involved in the exchange procedure must exactly fit to their respective counterparts. This is not always granted and limits the applicability of the approach. To overcome this, an editing framework has been proposed by Rosenbaum (2006). It is shown that the probability of a perfect fit increases with a subsequent un-do of the encoding procedure, and the recursive application of the exchange procedure based on different container types always leads to the desired result.

Results: The evaluation of the proposed framework has revealed that its performance is never slower than a full pixel-based exchange, but *in average much faster*. This means that by a universal applicability, *much less computing power* is required for content exchange at client side. As approximately the same amount of data is transferred in both cases, there is no significant difference in the use of bandwidth compared to the general exchange scheme.

Figure 5. The basic scheme for content exchange in JPEG2000-domain (left) can also be enhanced to inherently reduce the spatial dimensions of the included image region (right)



4.3.3 Motion-JPEG2000

Although Motion-JPEG2000 offers a number of very desirable features for digital video processing, its compression performance is not as high compared to other codecs. This is mostly due to the neglect of interframe redundancies. This is considered by a strategy designed to advance the compression performance of Motion-JPEG2000 video (Rosenbaum, 2006). By detecting and omitting data belonging to regions which are static with regard to two subsequent frames, much data is removed from the original data-stream. The original frame can easily be reconstructed by re-including this data from the previous frame whenever required. Due to the fact that each single frame within a compliant Motion-JPEG2000-video is encoded with JPEG2000, the proposed ideas for static or remote context exchange in JPEG2000 domain can be applied without limitations.

Results: Due to the foundation on the described basic scheme, detection, removal and reconstruction are *little demanding regarding computing power*. All stages are accomplished on encoded data and no decoding is required. Thereby, this procedure is also applicable to video streaming. By transmitting only the dynamic parts of each frame and reconstructing the whole frame at client side usually *much bandwidth is saved*.

5. CONCLUSIONS

By focusing on the results achieved by the design of appropriate communication systems, this publication has shown the potential of image communication using JPEG2000. To reduce the impact of *limited computing power* and *bandwidth* in mobile environments, all stages of the image communication pipeline have been tightly coupled. Based on the generic concepts for a *demand-driven streaming* based on *dynamic RoIs and JPEG2000*, the three major tasks in mobile environments – *Image browsing, Viewer guidance* and *Content exchange* could been implemented very efficiently. The respective communication systems *significantly reduce the need for computing power* (DoF-80%, content exchange-89%, Tool glasses-95%) *and bandwidth* (often 90%). Overall, the achieved results confirm the authors' assertion that such a system can *greatly increase the performance* in mobile environments. The rigorous compliance of the proposed technology allows for an easy migration in existing systems.

REFERENCES

S. Deshpande & W. Zeng (2001). Scalable streaming of JPEG2000 images using hypertext transfer protocol, *Proceedings of the 9th ACM international conference on Multimedia*

ISO/IEC JTC1/SC29/WG1 N2678 (2002). JPEG2000 image coding system, Part 1, FPD

ISO/IEC JTC1/SC29/WG1 N3052R (2004), JPEG2000 image coding system, Part 9, FDIS

B. Karstens, R. Rosenbaum, & H. Schumann (2004). Information presentation on handheld devices,Idea Group Publishing, *E-Commerce and M-Commerce Technologies*

J.P. Ortiz, V.G. Ruiz, & I. Garcia (2004). Remote Browsing of JPEG 2000 Images on the Web: Evaluation of Existing Techniques and a New Proposal, *Proceedings of VIIP*

M. J. Owen, A. K. Lui, E. H. S. Lo, & M. W. Grigg (2001). The design and implementation of a progressive on-demand image dissemination system for very large images, *Proceedings of the 24th Australasian conference on Computer science*

U. Rauschenbach (1999). The rectangular FishEye View as an efficient method for the transmission and display of large images, *Proceedings of IEEE ICIP*

R. Rosenbaum, & H. Schumann (2002). Flexible, dynamic and compliant region of interest coding in JPEG2000. In *Proceedings of IEEE ICIP*

R. Rosenbaum, & D. S. Taubman (2003). Merging images in the jpeg2000 domain. *Proceedings of VIIP*

R. Rosenbaum, & D. S. Taubman (2003). Remote display of large raster images using jpeg2000 and the rectangular fisheye-view. *Journal of WSCG*

R. Rosenbaum, & H. Schumann (2004). Limited spatial access in jpeg2000 for remote image editing. *Proceedings of VIIP*

R. Rosenbaum, & H. Schumann (2005). Grid-based interaction for effective image browsing on mobile devices, *Proceedings of Electronic Imaging - Multimedia on mobile devices*

R. Rosenbaum, & H. Schumann (2005). Jpeg2000-based image communication for modern browsing techniques. *Proceedings of Electronic Imaging - Image and Video Communications and Processing*.

R. Rosenbaum (2006). Mobile image communication using JPEG2000, *PhD-thesis*, University of Rostock

R. Rosenbaum, & H. Schumann (2006). Jpeg2000-based viewer guidance for mobile image browsing. *Proceedings of IEEE MultiMedia-Modeling*

R. Rosenbaum, Ch. Tominski, & H. Schumann (2006). Graphical Content on Mobile Devices, Idea Group Publishing, *Encyclopaedia of e-commerce, e-government and mobile commerce*

B. Shneiderman (1996). The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations, *Proceedings of the IEEE Symposium on Visual Languages*

D. S. Taubman, & R. Rosenbaum (2003). Rate-distortion optimized interactive browsing of jpeg2000 images. *Proceedings of IEEE ICIP*

Related Content

The Role of Innovative and Digital Technologies in Transforming Egypt Into a Knowledge-Based Economy

Sherif H. Kameland Nagla Rizk (2019). *Handbook of Research on the Evolution of IT and the Rise of E-Society (pp. 386-400).*

www.irma-international.org/chapter/the-role-of-innovative-and-digital-technologies-in-transforming-egypt-into-a-knowledge-based-economy/211624

Understanding the Context of Large-Scale IT Project Failures

Eliot Richand Mark R. Nelson (2012). *International Journal of Information Technologies and Systems Approach (pp. 1-24).*

www.irma-international.org/article/understanding-context-large-scale-project/69778

Modernizing the Academic Library

Jennifer Ashley Wright Joe (2021). *Encyclopedia of Information Science and Technology, Fifth Edition (pp. 1757-1766).*

www.irma-international.org/chapter/modernizing-the-academic-library/260304

Design and Analysis of AWG and FDL-Based Optical Switch in Data Centre Network

Vaibhav Shuklaand Dilip Kumar Choubey (2021). *Encyclopedia of Information Science and Technology, Fifth Edition (pp. 980-1001).*

www.irma-international.org/chapter/design-and-analysis-of-awg-and-fdl-based-optical-switch-in-data-centre-network/260244

An Empirical Evaluation of a Vocal User Interface for Programming by Voice

Amber Wagnerand Jeff Gray (2015). *International Journal of Information Technologies and Systems Approach (pp. 47-63).*

www.irma-international.org/article/an-empirical-evaluation-of-a-vocal-user-interface-for-programming-by-voice/128827