ISSN 1392 – 124X INFORMATION TECHNOLOGY AND CONTROL, 2006, Vol.35, No.2

SELF ANNOTATED RASTER IMAGE

Geruta Kazakevičiūtė
Dept. of Graphical Systems, VGTU
Eugenijus Januškevičius
Lab. of Computer- Aided Architectural Design, Dept. of Architecture, VGTU
René Rosenbaum
Institute of Computer Graphisc, Dept. of CS, University of Rostock
Heidrun Schumann
Institute of Computer Graphisc, Dept. of CS, University of Rostock

Abstract. Thenewembeddingmethodofannotationsintorasterimagesispresentedinthearticle.Themethod combines the principles of watermarking and steganography. Using the method proposed it is possible to attach the information to the raster image and bind it to a specified object or region of the image. The JPEG2000 raster images coding standard is used as base to realize the embedding and extracting scheme of image data. The proposed scheme targeted to mobile environments and evaluate their limits. The article reviewes the existing methods of embedding for metadata and annotations; the principles of information embedding and extracting for JPEG2000 images. The main requirements for annotation scheme are defined. Embedding principles and logical structures are analyzed. These requirements are practically adopted for the new method of annotations embedding into the raster image, evaluating the JPEG2000 requirements. The scheme is tested practically and the results of these experiments are presented.

Keywords: information embedding, JPEG- 2000, annotation, hotspot, region of interest.

Introduction

Most of the information a human being receives is via vision. Based on this, there is also a huge interest to process visual information in computers.

Unfortunately, understanding visual image content is a currently unresolved research topic in computer science. To overcome this limitation, visual information should be accompanied by an additional semantic description - metadata. The problem here is that many established image standards do not allow to embed arbitrarily structured metadata and preserve connection between the image and annotation without any additional effort. Another problem, especially in remote environments, is the need for additional bandwidth to transmit the annotation.

The approach of embedding data directly into a host image with the purpose of an additional description (see Fig. 1) is relatively new. It is related to the field of information hiding, and forms, beside Steganography and Watermarking, the third basic part of embedding strategies. As shown in Table 1, the general idea is equal for all strategies, but due to the different focus of each strategy, the properties and thus the implementations are rather different.

In this paper, we propose a new approach to embed arbitrary annotations directly into a raster image without need for further bandwidth or provided structures to embed such metadata. Due to its compression performance, flexibility and numerous features, the JPEG2000 image coding standard is applied as basis for the proposed compliant embedding and extraction scheme. To minimize the influence of additional information to the content, the image is altered directly in the wavelet domain, excluding the need to re-transform images often stored in compressed representation.

The introduced method allows to assign the embedded information to a spatial region. Different objects

Table1. Requirementsforinformationhidingstrategies

	Steganography	**Watermarking**	**Media description**
Protects	Message	Cover	Cover and annotation
Aim	Hidden communication	Copyright protection	Information preservation
Goal	Undetectability	High robustness	Information transmission
Message type	Presence of message is hidden	Presence of message known, message undetectable	Presence of message is known, message detectable
Message length	Long	Short	Long
Message localization	Low	Medium	High

Figure 1. Interactive digital image. The information appears when you point an object

or region, shown within the image can be described independently. This is rather useful, since the introduced approach additionally is robust against cropping. Together this allows to exchange certain image regions between different images without loss of the annotation of the belonging content, which is quite convenient if the image is reused multiple times.

The article is organized as follows: After the introduction, Section 1 deals with existing methods for the embedding of annotations and metadata. Since embedding and extraction of the media description are accomplished directly in JPEG2000-domain, existing methods of embedding information into JPEG2000 imaginery are reviewed. Section 2 derives and defines requirements an appropriate embedding scheme should fulfill. Together with the properties of JPEG2000 these statements are applied in Section 3 to propose a new embedding method for media annotations in JPEG2000 domain. To show the applicability of this approach, results are reported in Section 4. Conclusions and directions for further work close our contribution in Section 5.

1. Existing methods to handle media descriptions in raster images

The value of information often depends on how easy and quickly it can be found, accessed, retrieved, filtered and managed. To give the reader an understanding of our contribution, this section reviews existing methods for describing and embedding additional information in raster images. Here it is reasonable to divide in statements regarding approaches used to organized and structure the description and approaches to embed such a description into the host image.

1.1. Description of raster image data (structure of the embedded data)

The ease of information usage is determined by its structure and description. The most widely used paradigm for information description in digital media is the use of supplementary data - metadata. It is usually organized in pairs of attribute and value, and to some extent is self-describing. The application of metadata is closely related with the media storage format, as usually the storage format defines the options for metadata storage and retrieval. Basic implementations of this paradigm can be found in applications where the available information for the description is supposed to be fixed and non-expandable. Typical examples are the *EXIF* information, used in digital cameras, or *Dublin Core Metadata Element Set* (DCES) [1], used in HTML or in many existing image file formats. Here certain attribute sets and their belonging values are used to describe the image. Since this attempt is by far not sufficient to describe arbitrary media contents more sophisticated strategies have been proposed. Here information is described and organized by using extendable, hierarchical and freely-definable structures. Formats like DIG35 [2] or MPEG7 [3] are milestones to create and handle images descriptions in a rather flexible way. They are based on the well known XML-description scheme [4] and believed to be the best paradigms to handle structurized data.

1.2. Embedding of media descriptions

Embedding information into an image file format can be done in a variety of ways. The most simple is the use of features provided by the file format. Common examples are JPEG (EXIF) [5] or the PNG image file formats. Here fixed fields are reserved for attribute values and might be filled with the current description, whereby quantity and kind of attributes depend directly on the file format used.

Another and more flexible approach to embed arbitrary metadata is used in modern formats like JPEG2000-image file format [6] or in parts in the TIFF or TGA format [7]. Here the concept of containers or boxes to store the information is used. The number of boxes is not limited and they can be filled with arbitrary content, e.g user-defined attribute pairs. Since this attempt is by far more flexible than the described static approach, more sophisticated schemes to describe information, e.g. MPEG7, can be applied.

However, all those approaches simply attach the description to the image. Since the attached information is not a standart part of the image data and there is no uniform support for certain attributes in each image file format, a big drawback of this approach is that most of the description is lost if the file format is changed. Furthermore, the simply attached information creates additional payload when the image is transmitted.

Instead of attaching the media description, we propose to embed it directly into the image, to overcome this limitations. Since almost all techniques do not distinct in the kind of the embedded annotations as long as it is in binary format, we refer to the media description further on simply as data. The general idea is not new but is mostly used in application fields as digital Steganography or Watermarking. The advantages of this approach are that the embedded data become an inseparable part of the media and take almost no additional storage space. Many algorithms applied in Steganography might also be suitable for embedding of metadata, since the general focus to emphasize capacity is one of main requirements for media description. Nevertheless, these approaches might also be adapted to further increase capacity by omitting the requirement of undetectability. Many of these approaches are designed for the spatial domain as it provides more capacity and controllability for embedding and extraction of the additional information. Here, the most common approach to embed data is to substitute certain bits of the available pixel representation by bits of the additional data [8]. This technique is known as LSB-approach in Steganography. As more significant a bit is as more a manipulation will be noticeable, as less significant a bit is as higher the probability that it will be altered during unintended manipulations. An interesting steganographic approach of embedding descriptive data into an host image using aditional data stream was proposed by S. Areepongsa [9]. Although the focus of their paper is on MPEG7-data, it is conceivable to apply the approach to other kinds of description schemes too.

Since it can be shown that capacity is inversely proportional to robustness [10], these approaches are highly vulnerable if the image is manipulated after embedding. This is a significant drawback, because nowadays, image data are mostly used in compressed form to reduce storage space and transmission costs, which often includes a lossy transformation of the image content. Due to their special alignment, watermarking techniques are more robust against such manipulations than steganographic approaches [10]. Most of them have been designed for a certain transformation domain and embed the additioanl data (watermark) directly into the transformed image representation to be resistant against modifications imposed by the transformation itself. Dependent on the used transformation domain, e.g. Discrete Cosines Transform (DCT) [11] or Discrete Wavelet Transform (DWT) [12, 13], or specific file formats using transformation domains, e.g. GIF [14], JPEG [15] or JPEG2000 [16, 17], many different techniques have been proposed. Nevertheless, a general problem with all of these approaches is, the capacity is in general very low and the whole image is used for embedding to reduce the risk the embedded data can be simply removed by cropping parts of the image.

1.3. Embedding data into JPEG2000 compressed imaginery

To create a reasonable technique for media embedding, relevant properties and approaches from Steganography and Watermarking approaches must be combined and enhanced. Due to a variety of available techniques, we focus on DWT-based approaches, since many of them deliver very good results and can easily be applied to the JPEG2000 image coding standard.

One possible way is the modification of coefficients resulting from the DWT of the image [12]. Thus, the first lossy transformation can be skipped before embedding. However, in many cases a subsequent quantization step alters these coefficients again, and must also be considered in order to be able to restore the embedded data successfully. To overcome problems imposed by the quantization step, different strategies, either by steering the embedding dependent on the belonging quantization value [13] or by embedding after quantization have been proposed. For the last, there are two options: either modifying high index values, possibly causing noticeable distortions in the final image (REF), or modifying zero index values, produced by the JPEG2000 dead zone quantization [19]. The key point of this approach is that slightly altering zero quantization indexes will not change the restored image greatly. However, the negative issue of the method is that compression performance will decrease.

A combined technique derived from Steganography, is the LSB-approach applied to DWT coefficients. Bit plane modifications usually have the highest possible informational capacity (up to 4 bits/pixel), but they are not robust to even the slightest modifications. As our goal is to maximize capacity, we can adopt the LSB-approach, keeping in mind its disadvantages. Instead of pixels, the quantized coefficients are used in order to avoid further manipulations and to guarantee a full recovery of the embedded data. Two techniques already have been proposed using this approach to embed data into a JPEG2000 bit stream. One is proposed by Po Chyi Su [16] and the other proposed by H. Noda et al. [17]. Although, these approaches are suitable to embed data globally all over the image, it is not clear how the aspired range limitation of the embedding and following detection during extraction can be accomplished.

Both method apply the reasonable assumption that before and after the arithmetical encoding [1], the embedding leads to unpredictable results. Thus, both techniques embed the additional data directly after the quantization step. Nevertheless, to keep compliance they still face a number of problems. The most important problem might be the ability of the encoder to remove certain less significant bit streams to achieve a determined bit rate. As already mentioned, additional data might be embedded at less significant positions to keep modifications invisible. Unfortunately, these positions might be the first to be removed to keep a certain bit rate.

Su's[16] method is based on a compliant configuration of the encoding process, called "lazy coding" [19], and does not consider a possible remove of less significant information. The main aim of lazy coding was to speed up the encoding process by skipping the 4 lowest bit planes during arithmetical encoding. They are passed to the rate control stage without being further processed. Thus, selected bits of each coefficient can be modified and afterwards passed to the rate control stage. In some cases rate controller may exclude some bit planes from further processing. In this case, embedded data may be partially or completely lost. Furthermore, this technique must be applied directly at encoding time.

Noda's scheme [17] is more sophisticated and based on a partial roll-back of the JPEG2000-encoding procedure. Basically, the encoding is undone until the dequantization stage. The idea here is, that the data has already passed the rate controller during the first encoding. Since an aspired bit rate has already been established, no information is removed while passing the rate control again. The additional information can be embedded after the quantization stage and the manipulated image data are again processed by the remaining parts of the JPEG2000 pipeline. To be absolutely sure the embedded data will be kept during the second processing, the target bit rate may be set at a lower value for initial processing and set to the desired value for the second and final run. This technique can be applied during encoding as well as to already encoded JPEG2000 bit streams, which increases the application field of this technique.

After deriving requirements for self-annotated raster images in the next section, it will be shown how the second approach can be enhanced to support local image annotations.

[1] for detailed description of JPEG2000 coding see [6] and [19]

2. Basic requirements for region-based embedding schemes

The requirements are based on limitations imposed by the used hardware. Here, we focus on mobile devices, like cell phones, sub-notebook computers and other pocket-sized information appliances, since they are equipped with the most limited hardware components. Requirements for other meaningful device classes can be easily derived from this device class by relaxing or omitting one or more of the introduced constraints.

The limitation of mobile devices are mostly imposed by their small construction and can be stated as:

Limited display size: Due to the small outline, devices can not be equipped with large displays. Further on, having many pixels and high resolutions are the main energy consumers, which is probably the main enemy in mobile environments.

Limited processing power and storage space: Small and handy devices can not provide as much power and storage space as devices without this design concept. Additionally, strong power architectures consume much more energy. To make the application usefull, the extraction of embedded data must be as simple as possible.

Slow and expensive communications: Wireless communication is problematic with regard to many aspects. Beside the point that the ether is overcrowded with different sources of radio signals and it is hard to establish a reliable link in the most popular ISM band (2.4 GHz), transmissions are slower and much more expensive than in wired environments. The scheme must neither add extra data to the information transfered nor need any aditional tranfers to retrieve the data for annotations.

To be suited for limited display size, the scheme must be applicable to the partial content of the image, visible on the screen. As images might be rather large compared to the available screen dimensions, it might be useful to transmit only a part of the image, as this is a common feature of modern transmission systems. This also decreases requirements for processing power and storage space needed to handle the whole image. It is mandatory to be able to find and extract all embedded annotations belonging to regions shown on the screen. This also applies if the image is scaled regarding quality, as modern image coding standards, like JPEG2000, allow.

Since there is no limitation regarding the image content to be described, shape and number of image regions, and the length of each description are a priori only limited by the properties of the host image. It is obvious that the sum of all annotations can not exceed its provided capacity. Nevertheless, this is a strong demand assuming an optimal and redundancy-free usage of the provided capacity, which is rather hard to fulfill for arbitrary application fields.

Based on these assumptions, we are able to derive the basic requirements for embedding region-based image descriptions into raster images:

1. All embedded data belonging to regions shown on screen must be retrievable.
2. The extraction does not depend on current image quality.
3. The need to transmit annotations separately must be avoided.
4. Annotations may contain any binary data.
5. Regions to be described may have any shape.
6. The number of embedded annotations is only limited by the provided capacity of the host image.
7. The length of embedded annotations is only limited by the provided capacity of the host image.

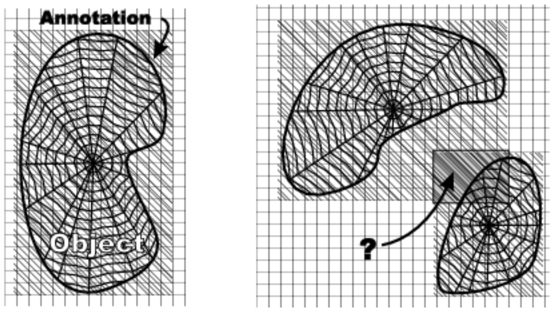
In this article, we are targeting on requirements 1-5. As already mentioned, completely using the provided capacity of the host image is often not possible, especially if dealing with constraints imposed by working with the image in transformation domain.

Based on these requirements and properties of JPEG2000, in the next section, an approach for region-based embedding of additional information is proposed.

3. Application's scheme for information embedding in JPEG2000

We described the basic requirements for the embedding scheme. Here we will show how to achieve these goals by a new approach. Some of the requirements, described in Section 2 are more easy to implement in JPEG2000 image compression standard we are focusing on.

3.1. Spatial and logical structure of the embedded information



(a) Annotation is embedded in the rectangular ROI

(b) Collision of two annotations

Figure 2. The simplest approach of annotation embeding

In our approach, a user needs to get the information about the object he is pointing at. Thus the information should be embedded in a local region of the annotated object - Region of Interest (ROI). This way the embedded data are strongly spatially related to the displayed object.

The most simple approach could be inserting the annotations in the bounding box, surrounding the object or ROI (Fig 2(a)). Two problems arise in this approach:

1. *space for embedding*, in the case we need to embed a lot of information about small object of the image. There are two posible ways to resolve space issue: use of spatially localized tree-like structure of DWT coefficients or expanding an annotation to the unused areas of the image. In this case we must be able to identify the object annotation is bound to.

2. *localization*, when information from one object can be by mistake treated as information of the other object (Fig. 2(b)). To prevent this problem, some tagging of both object and annotation should occur.

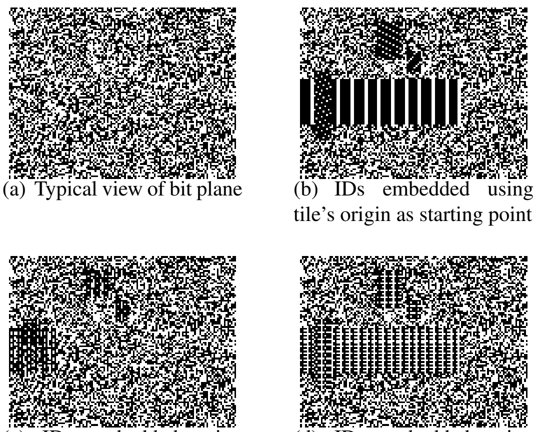
The information for embedding may be split into two parts: *ID*, used to uniquely identify any ROI in the image and the *content*, describing the ROI. As only ID needs to be spatially tied with the image, the amount of information, needed to embed directly over the ROI is several times less than in previous approach. Moreover, as the annotation can be placed anywhere in the image, the distortions for the image are less concentrated and thus less visible, which will result better visual appearance of the image.

The most important question at this point is the localization of region. Fig. 3(a) shows typical view of a lower bit planes. Fig. 3(b) shows the same bit plane with ID of ROI embedded. There is no fast and reliable method of finding the ROI and determining its ID, as the origin of embedding is unknown. As the processing of ROI should be fast (preferably real time), we need to create a reliably identifiable structure for fast processing.

To find the free-shaped boundaries of the ROI (requirement 5 in Section 2), we can use an approach, called 'masking'. The idea is that we can determine a mask, consisting entirely of ones (or zeros) to indicate the boundaries of ROI. The distribution of bits in lowest bit planes is in most cases random, so using 4- neighbor context we can determine if the pixel still belongs to the ROI (Fig. 4(a)). Using LSB modification, we can change up to 1/3 of lowest bits in a byte before the changes become visible by humans eye. In our case, we can embed the mask in the LSB plane and the ID may be embedded in the LSB+1 bit plane. We will be able to identify ROI when only single index will be restored. Small errors in the LSB plane will not prevent us from restoring actual shape of the ROI as long as bits will stay 4- connected. The problem of this approach is the amount of embedded information is relatively high and it has no tight relations with the image. Using this approach, we disturb, especially in the LSB plane, the natural entropy of the image.

Another approach, which is more friendly to the annotated information, is to fill the LSB bit plane of the ROI using pattern, consisting of multiple frames (Fig. 4(b)). In this case, the entropy of pattern is closer to the natural entropy of the bit plane and the restoration process will result less visible distortions. Fig. 3(b) shows hot spot index embedded using this approach.

The ID of the ROI can be embedded at predetermined bits, taking JPEG- 2000 tile's origin as the base of our coordinate system. As we need to embed 8-bit IDs and possibly error correction codes, we can arrange the information into matrix of 4-by-3 bits, where a matrix of 4-by-2 bits is used for index and the last column is used for error correction bits, and embed it into every cell of 4-by-5 bits. These dimensions are well coordinated with the size of JPEG2000 code block (4 bit height) and ID can be embedded in the JPEG2000 encoding stage



(a) Typical view of bit plane

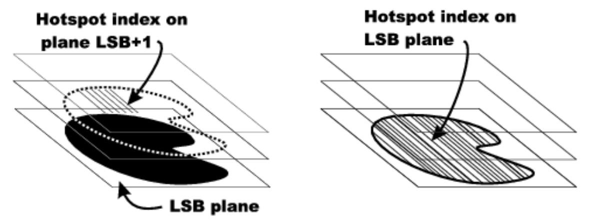
(b) IDs embedded using tile's origin as starting point

(c) IDs embedded using regular grid

(d) IDs embedded using frame structure

Figure 3. Different types of ID embedding

or bit stream, without decompressing the image. Fig. 3(c) shows bit plane with IDs embedded using just described method. The localization of embedded ID can be performed scanning every 5 columns of the tile and searching for the bitmap pattern with correct CRC code. This approach is not robust against cropping nor prevents false positives.



(a) Embedding hot spot index in two bit planes

(b) Embedding hot spot index into single plane

Figure 4. Methods for embedding hot spot index

Another approach is based on industry - adopted technology of semi-automatic object identification, called bar codes [22]. The simplest bar code is a set of stripes of varying width, encoding the information. Bar codes are designed to be reliably recognizable in different viewing conditions and they contain error correcting codes. Bar codes are well - suited for automated recognition purposes as they can be read either by linear scanner or camera. In order to recognize the code a single valid row must be found over the whole length of the code while the code can be severely physically damaged. The size of strips and synchronization of bar codes is determined after it was scanned, using special starting and ending stripes. The main disadvantages of bar codes are: most of them require "quiet zone" before and after the code to read it reliably; they take a lot of place.

To overcome these issues, a new subset of bar codes, called 2D - bar codes were introduced. 2D bar codes are characterized by high informational density and most of them do not need the "quiet zone" around. Most types of 2D bar codes are arranged in such way, that the information can be read at any rotation angle and the software is able to restore the correct orientation of the code. Fig. 5 shows some patterns of linear and 2D bar codes.

The use of 2D bar codes could allow us to place IDs anywhere in the image, without need to synchronize its boundaries with the tile's coordinate system (a problem in the case we need to annotate narrow objects). The two main types of 2D bar codes are: 1) centered pattern (like Fig. 5(b)) and 2) framed pattern (like Fig. 5(c)). Centered 2D bar codes can accommodate more information, but due to the limited space they are not suitable to use in our approach. Framed 2D bar codes are more suitable in our case, as they can be scaled freely.

To use the principle of 2D barcodes, we need only a slight alterations to our previous approach. As the height of JPEG2000 stripe is 4 pixels and the frame of 2D barcode uses 1 unit from bottom and left sides, we have a stripe of 3 pixels height to embed the information. We transpose the matrix from previous approach and use 2-by-4 bits for ID and use one row for error correction bits. The surrounding frame must contain unique, clearly artificial pattern to help us to detect it in the image.

We decided to use the following application scheme:

- The precinct - the basic element of compliant JPEG2000-based image transmission and spatial organization - is transmitted every time a part of the image is requested. If the embedded annotation will not cross precinct's boundary, we will be able to retrieve the information about any object in the precinct. In the case an object is covered by two or more precincts, to avoid duplication of the same data, the information should be embedded to align with tile rather than precinct boundaries.
- The ID is embedded in the spatial region, coin-

(a) EAN13 composite

(b) Aztec

(c) Datamarix

(d) PDF417

Figure 5. Different types of bar codes

ciding with the ROI. (Only rectangular ROIs were used in testing environment.) To make ID acessible as fast as possible, it should be embedded in the first quality layers of a precinct. A caution must be taken here, as distortions introduced in earliest quality levels here accumulate in later quality levels and are not correctable. This means the space to embed the information will be limited or the quality of the image will degrade.

- Annotations can be either scattered over the whole image or placed over a certain region. Expandable data space, not showing on the image, but defined to hold annotations, may be designed to hold the annotations that do not fit into the designed space.

- Annotation may use any 8-bit value as data, except some predetermined values, like zero symbol(0x0), which should be 'escaped' in special predetermined way.

- The number of annotated objects is determined by the size of JPEG2000 image tile. In the case unsigned 8-bit integer value is used for ID, up to 255 unique IDs can fit in the single tile. More annotations can be embedded either using more and smaller tiles for the image or using 16-bit numbers for the ID, thus having 65535 unique IDs per single tile.

4. Results

To test our scheme, we used prototyping with MatLAB environment and some of its toolboxes: Image Processing toolbox and Wavelet Toolbox.

We used a simplified JPEG2000 processing routine:

1. As JPEG- 2000 is using YUV color space, all images we used were gray scale.

2. We assume the size of tile is equal to the size of the image for relativelly small test images

3. DWT decomposition was performed using 'db4.4' wavelet filter, the same as '9/7' wavelet filter, used for lossy compression in JPEG2000

4. quantization process was simulated rounding wavelet coefficients to the lower integer value

5. rate-control issues on embedded information were out of scope in current work phase.

First of all, we needed to set allowable PSNR value to decide if embedding was successful. Visible artifacts appear in the image when PSNR value is above -32dB. So the desirable PSNR value is -35…-32dB.

The next step was to determine possible extents of bit plane modifications. These modifications are limited by 2 factors:

1. the minimal size of a hot spot, as dimensions of hot spot in LL sub-band are decreasing in geometrical progression when making DWT decomposition and

2. the PSNR value, influenced by distortions.

The minimal size of hot spot is 4 rows and 5 columns and these dimensions did not allow us to make more than 3 levels of DWT decomposition, working with relatively small images (Lena, Mandrill).

The distortions of the image are of 2 origins: modification of LL sub-band and modification of coefficient sub-bands. Wavelet coefficients appear in number and defined bit depth as pixel values, with the important difference that they belong to a certain decomposition level. Due to this, they have a different range of their spatial contribution. Unlike pixels with a range of 1, their contribution to reconstruct a certain region in pixel domain can be stated as 2^d for subbands LH, HL and HH created by decomposition step d, and 2^{d+1} for the approximation LL. This is of very importance if the embedding is coefficient-based and must be restricted to a certain region.

Nevertheless, we decided to embed the annotation's ID into LL sub-band. There are several reasons:

- the LL sub-band is the first sub-band to be transmitted;
- ROIs are tightly spatially localized here;
- if we will ever need to get the total number or area of ROIs, only a little subset of the whole image data will be queried.

The tree - like structure of DWT decomposition may be the ideal way of spatial localization of the embedded information. We can use this peculiarity and embed annotations into DWT tree of the image tile, instead of scattering them all over the whole image. In this case, the ROI is 'covered' with the embedded annotations and the tree-like structure becomes highly spatially localized. The shortcoming is that modifications to the image region are more concentrated and can lead to artifacts in the displayed image. The distortions in detail sub-bands are less like to cause artifacts, because the extents of modifications are lower.

Testing of the scheme revealed the week places. First of all, as we embed IDs index into the whole area of ROI, the artifacts of the image become visible and the value of PSNR quickly approaches to -35dB. In order to minimize influence of modifications, only the LSB plane of LL sub-band could be used for embedding. The detail sub-bands, as expected, are less sensitive to additional information. In the case we used scattered annotation embedding, we could use up to 3 lowest bits of the coefficients to embed the information, without noticeable influence to the PSNR value and image's quality. In the case of concentrated, tree-like annotation embedding, to keep the PSNR value at the same level, we had to use only 2 bits of a coefficient in the detail (LH, HL and HH) sub-bands.

The second problem is the size of the ROI. In our approach, the framed ID of the ROI is placed in LL sub-band. As the size of the ROI decreases in geometrical order during DWT decomposition, it is hard to annotate narrow objects using high levels of DWT decomposition, because there is no space to embed ID. If we embed the information in such places, the contours of the object under the ROI are mangled due to the image distortions. Figure 7 shows artifacts that appeared due to the information embedding into digital image.

Figure 6. All embedded hotspot regions are highlited with annotation placed on top. Image 'Goldhill', PSNR=-33.51dB

The other problem that had arised is the identification of the ROI. Using lower levels of DWT, not only makes the LL sub-band smaller, but also makes the identification of ROI much harder. The overall entropy of the bit plane tends to decrease and the ID of the ROI in many cases is embedded in a single instance. When using higher levels of DWT, we can use neighboring regions to check and verify the presence of ID. We have no such an opportunity when using lower levels of DWT.

Evaluated informational capacity of the scheme is about 10% of the image's size and as expected is highly image dependent. The trend is that images from natural origin have more informational capacity than synthetic ones because of higher natural entropy.

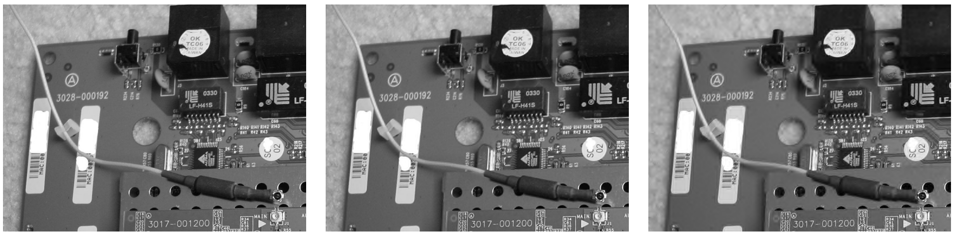
The example of the embedded hotspot regions with annotations are shown in Figure 6.

5. Conclusion and future work

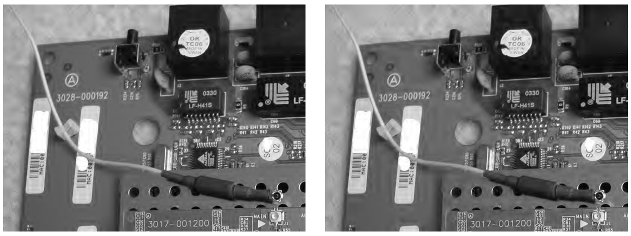
The article presents novel approach to annotations storage for raster images, based on methods of steganography and watermarking.

The proposed application scheme embeds the describing information and connects it to the determined spatial region of the image. It uses JPEG2000 image standard to carry the information.

The proposed scheme is primarily designed to work in mobile environments and estimates the limitations of these environments.



(a) LL bit plane 1, number of DWT - 1, PSNR -42.8061dB (b) LL bit plane 1, number of DWT - 2, PSNR -36.9773 (c) LL bit plane 1, number of DWT - 3, PSNR -34.1435



(d) LL bit plane 1, number of DWT - 4, PSNR -33.8231 (e) LL bit plane 1, number of DWT - 5, PSNR -32.4358

Figure 7. Influence of information embedding to appearance of the image

It is possible to embed the data in the digital image and restore them, although the influence of information embedding is visible to the eye. The performance of the scheme may be evaluated as good. It may be enhanced using full JPEG2000 processing pipeline.

Future research work should concentrate on finding a way of embedding ID, minimizing its influence to the final image; evaluating and testing the capacity of the scheme in real environment. A special interest is the behaviuor of the scheme in non wavelet-based image file formats, especially considering popular nowadays DCT based JPEG image file format.

References

[1] Dublin Core Metadata Initiative Dublin Core Metadata Element Set, Version 1.1. *Reference Description, http://www.dublincore.org/documents/dces/.*

[2] Digital Imaging Group DIG35 Specification - Metadata for Digital Images, Version 1.1. *Working Draft, Digital Imaging Group Inc*., 2001.

[3] **J. Martinez, R. Koenen, F. Pereira.** Mpeg-7: the generic multimedia content description standard. *IEEE Multimedia*, 2002.

[4] Extensible Markup Language (XML), 1.0 (Third Edition). *W3C Recommendation*, 04 *February* 2004.

[5] Exchangeable image file format for digital still cameras: Exif Version 2.2. *Standard of Japan Electronics and Information Technology Industries Association, Established in April*, 2002, 154.

[6] ISO/IEC JTC 1/SC 29/WG 1 N2678 JPEG 2000 image coding system, Part1. *Final Publication Draft*, 2002.

[7] Truevision, Inc. Truevision TGA file format specification, Technical Manual Version 2.2. *January*, 1991.

[8] **C. Kurak, J. McHugh.** A Cautionary Note on Image Downgrading. *Proc*. 8*th Annual Computer Security Applications Conference, San Antonio*, 1992.

[9] **S. Areepongsa, N. Kaewkamnerd, Y.F. Syed, K.R. Rao.** Information Hiding in Image Retrieval Systems. *ICCS* 2000.

[10] **F.A.P. Petitcolas, R.J. Anderson, M. Kuhn.** Attacks on Copyright Marking Systems. *Second Workshop on Information hiding, Vol*.1525 *of Lecture Notes in Computer Science, Springer Verlag*, 1998.

[11] **M.D. Swanson, B. Zhu, A.H. Tewfik.** Robust Data Hiding for Images. *IEEE - Digital Image Processing Workshop*, 1996.

[12] **M. Barni, F. Bartolini, V. Cappelini, A. Lippi, A. Piva.** ADWTbased technique for spatio-frequency masking of digital signatures. *Proceedings of Security and Watermarking of Multimedia Contents, San Jose*, 1999.

[13] **J. Onishi, K. Matsui.** A Method of Watermarking with Multiresolution Analysis and PN Sequence. *Systems and Computers in Japan*, 1998.

[14] **R. Machado.** EzStego. *http://www.stego.com*, 1997.

[15] **D. Upham.** JPEG-JSTEG. Modifications of the Independent JPEG Group's JPEG Software for 1-bit steganography in JFIF output files. *ftp://ftp.funet.fi/pub/crypt/steganography/.*

[16] **Po-Chyi Su, C.-C. Jay Kuo, Fellow.** Steganography in JPEG2000 Compressed Images. *IEEE Transactions on Consumer Electronics, Vol.*49, *No.*4, 2003.

[17] **H. Noda, J. Spaulding, M.N. Shirazi, M. Niimi, E. Kawaguchi.** Bit-Plane Decomposition Steganography Combined with JPEG 2000 Compression. *Proc. of Information Hiding*, *5th Intern. Workshop*, 2002, 295-309.

[18] **D. Taubmann.** High performance scaleable image compression with EBCOT. *Proceedings of IEEE International Conference On Image Processing*, 1999.

[19] **D. Taubmann.** JPEG2000 image compression standard. *Proceedings of IEEE International Conference On Image Processing*, 1999.

[20] **M. Droogenbroeck, J. Delvaux.** An entropy based technique for information embedding in images. *Proceedings of* 3*rd* *IEEE Benelux Signal Processing Symposium* (*SPS*-2002), *Leuven, Belgium*, 2002.

[21] **H.W. Tseng, Ch.Ch. Chang.** High Capacity Data Hiding in JPEG-Compressed Images. *Informatica, Vol.*15, *No.*1, 2004, 127 - 142.

[22] *http://www.barcodeart.com/science/science.html bar codes, including 2D bar codes*.

[23] **M. Celik, G. Sharma, A.M. Tekalp.** Reversible data hiding. *Proc. IEEE ICIP, Rochester, NY, Sept*., 2002.

Received January 2006.