Secure Embedded Data Schemes for User Adaptive Multimedia Presentation

Neelu Sinha
Department of Computer Science, Mathematics and Physics, Fairleigh Dickinson University, Madison, NJ 07940
Email: sinha@fdu.edu
Keywords: Adaptive systems, Embedded Metadata, Robust Metadata, Security, Adaptive presentation of Multimedia Content, Digital Rights Management

Contents

Abstract

In this *Digital/Internet Age*, digital multimedia holds an unlimited potential, and virtually all forms of media content, including books, video games, music and software are now available for digital distribution. Digital multimedia libraries, comprising a large amount of such digital media (the so-called Digital Intellectual Property), in the form of images, video, audio and graphics are rapidly growing. Also, due to the unprecedented growth of the World Wide Web, vast amounts of multimedia data is readily available leading to an explosion of multimedia and hypermedia database creation and sharing. Digital information embedding techniques for various types of media, for a variety of applications including digital libraries, museum cataloging, medical and healthcare industries, digital preservation systems, educational systems and personalization systems, are of significant interest in two areas. Firstly, these are useful for the realization of efficient database indexing schemes and customization, which in turn lead to efficient tools for the organization, retrieval, adaptive presentation, and distribution of digital media content. Secondly, these are useful in developing tools to protect, detect and verify ownership and/or usage rights for the Digital Intellectual Property and also the tracking of these in the distribution medium. In either of these applications, information embedding schemes which allow for a detailed level of source description, and which are robust to some of the distortions encountered in the distribution medium (e.g., JPEG compression for images), are particularly attractive. This paper presents an overview of multimedia presentation adaptation through the use of robust information descriptors and a novel information embedding technique for digital images that allows for significantly higher information throughput and increased robustness compared to many of the existing techniques.

1. Introduction

In this ***Digital Age/Internet Era***, all forms of digital multimedia content (images, video, audio, etc.) are available for electronic distribution and commerce. In fact over the last decade we have seen a significant shift in the way this content is being produced, distributed and consumed. This data can vary from a simple collection of a single media to a complex collection with many different formats and mixed media types. Also, it may be used by a variety of users each with their own personal usage requirements. Such personalization or adaptation of data during presentation to web user can often provide a competitive edge to the corresponding distributor/web-site. Although a certain level of personalization can be achieved with the help of user data – identity, preferences, profile, etc. – a significantly higher and satisfying level of personalization is possible if additional program associated data is stored with the multimedia content. Such program associated data, which we'll also refer to as *information descriptors*, can for example include a description of all rendering possibilities and/or transformation rules for the associated host data. Such information descriptors can also lead to efficient database indexing schemes, which in turn could lead to efficient tools for the organization, retrieval, presentation, and distribution of the digital media content. For example, imagine a user trying to find a book in a library without a catalogue, or, searching for a particular song in a vast database of songs without an index, or, trying to find a suitable TV channel without a TV guide, etc. In effect, a description about the what, how, when, who, why, where etc. of a particular resource (such as an image, audio or video data, etc.) is always needed.

The most straightforward approach towards including information descriptors is in the form of tagged data fields along with the host multimedia content. This has been referred to as "metadata" or "data about data" or "information about information". Significant interest in such content descriptors has led to various metadata structure standards. These have been proposed in order to improve searching, content management & sharing, etc. The earliest significant metadata standard was developed by the library community to enable the sharing of library catalog cards [WIPO-DRM]. This was primarily designed to support the "discovery" activity whereby the user may enquire in several different ways: through the name of the item (what editions of this book are in the library?), or through its author (what books by this author are in the library?) or through its subject matter (what books on this subject are in the library?). The tradition of library "discovery" metadata has been carried forward into online practice through "Dublin Core," a standard developed primarily by librarians that was initially conceived as a standard for cross-media and cross-sector discovery on the Internet. The Dublin Core Metadata Initiative (DCMI) [DUB-META] is an organization dedicated to promoting the widespread adoption of interoperable metadata standards and developing specialized metadata vocabularies for describing resources that enable more intelligent information discovery systems. Another metadata initiative is the proposed W3C Resource Description Framework [W3C-RDF]. This is a framework for describing and interchanging metadata and integrates a variety of applications from library catalogs to personal collections of photos, music, etc. Metadata is particularly important for digital images, where users depend on the information added to the image for accurate searching, retrieval and viewing purposes. Also, in order to respect intellectual property rights the end user may need to be validated before allowing image modification capabilities etc.

In many of these applications a level of robustness for the information descriptor is also desirable. For example, some encrypted or hidden information may be stored in the metadata to restrict certain rights (for example: view only, no modifications allowed, etc.) related to copyright protection. Kuo et al., (2004) succinctly describe the explosion of multimedia content as a transition from the ***physical*** to the ***digital*** and further assert that this transition is both a blessing and a curse. The digital medium offers more flexibility in producing, transporting and consuming media content, and at the same time, it also minimizes the efforts needed for unauthorized transport and consumption of this data. Also, since this data represents intellectual property in the digital domain, there is a need to manage, store, and distribute such data using schemes which can protect, detect and verify ownership and/or usage rights of the end user. This is further affirmed by the US Digital Millennium Copyright Act (DMCA), enacted into law in Oct 1998, which was crafted to protect property rights in the digital world while facilitating the robust development of electronic commerce, communications, research and development, and education in the digital age [DMCA-USCOPY].

The collective set of methods and techniques used to impose rules on how content is produced, distributed and consumed is referred to as Digital Rights Management (DRM) (Kuo et al., (2004) and WIPO-DRM). Developers of such schemes try to satisfy two different, and often contradictory, requirements. On one hand, the scheme should be sophisticated enough to protect the objective of intellectual property owners in a distribution environment that may attempt to circumvent those objectives (e.g., unauthorized and often illegal copying and/or distribution of music albums in mp3 format). On the other hand, the scheme should be non-onerous, and, to an extent, open enough to encourage authorized distribution leading to larger commercial and/or other rewards for the intellectual property owners. For example, despite its pitfalls, the Internet often presents a lucrative, cheaper, faster, and ubiquitous connection between owners and potential consumers of digital multimedia content.

There are two possible approaches to achieving the desired robustness for the information descriptors. The first approach is encryption whereby the metadata is protected with the help of suitable public or private key protection. A second approach towards robustness is to incorporate the information descriptors in the form of embedded data which is not easily separable from the host multimedia content. Such embedding has found an important place in solving many DRM problems. Furthermore, data embedding has found other novel applications:

- In a Digital Library (DL) application - as a vehicle for incorporating indexing information which is robust and which may be used to enhance the user experience during the access and/or rendering stage.
- In a Museum Cataloging application – for managing and documenting museum collections (physical and digital collections) by cultural institutions and commercial organizations engaged in creating libraries of digital images for example.
- In Digital Preservation application – for long-term preservation and migration of digital resources in the face of changing technology.
- In real-time Annotation Systems – where experts can attach their own opinions, subjective notes etc. to digital media to provide significant value-add, which can be useful for example in surveillance systems used by the defense community.
- In Educational systems – to generate personalized multimedia presentations or tutoring systems which can adapt to various learning levels of the students.
- In Personalization Systems – which provide alternatives to the traditional "one-size-fits-all" systems based on the user's needs, prior knowledge, abilities, learning styles, etc.
- In Science and Technology fields – to track and protect large volumes of sensitive data.
- In Entertainment as well as Edutainment industries – to offer personalized rendering of the content based on personal preference and tastes of the users.
- In Medical Systems and e-health – to access medical data repositories and image databases.

In this paper we also describe a multi-modal scheme of data embedding in digital images using embedded adaptive metadata which can be used in various applications including e-health, e-learning, e-commerce, etc. Recently we introduced a novel data embedding scheme for digital images (Sinha 2000) that is based on Adaptive segmentation and Space-Frequency representation (WASSFR). In this paper we develop the techniques further and describe newly incorporated tools which allow us to increase the data throughput of the algorithm substantially. We also describe design details and robustness analysis for the algorithm. As noted above, data embedding, often, is only one of the needed components in a DRM system. Other components are related to information massaging, such as, Encryption, Channel Coding and Synchronization. In this paper we also provide details on some of these additional components of DRM system. It should be noted that the techniques and concepts described here encompasses several diverse fields including media representation and compression, information theory, coding theory, cryptography, signal processing and human sensory system modeling.

This paper is organized as follows. In this first section we present an introduction to data embedding and its applications. Section 2 looks at the information descriptors for adaptive presentation of multimedia content related to several different applications. In Section 3, the metadata creation and security considerations are discussed. Metadata embedding is discussed in Section 4. In Section 5 we present a novel multi-mode data embedding technique using adaptive segmentation and space-frequency representation. The details of our design and the selection of various parameters

based on different experiments are discussed in Section 6. Section 7 discusses the performance requirements and design validation methodology. The distortion metrics used in the performance evaluation is also discussed in this section. Section 8 is devoted to discussing the results of various robustness studies conducted for a variety of distortions including the JPEG compression and rotation. This section also compares the information throughput of our scheme with other existing schemes. In section 9 we discuss the other components of a complete DRM system, such as the synchronization mechanism, encryption and the channel coding scheme. We also discuss the selection of the information pattern and embedding location in this section. Finally, in Section 10, we draw conclusions and offer some discussions related to future work.

2. Adaptive Presentation of Multimedia Content Using Data Associated Information Descriptors

Consumption of digital multimedia content, audio, video, images, is an important if not predominant component of a typical web user's daily experience. It is often desirable to customize the multimedia presentation depending upon a user's identity, profile, preferences or some combination thereof. Below we discuss several specific scenarios where such customization is desirable. A certain level of customization may be achieved with the help of user data (profile, preferences, etc.) collected at the time of browsing (either through the use of cookies or express questionnaire). However, as will be evident from the examples below, a much higher level of customization is possible if certain information descriptors are included with the multimedia data as part of the database. The exact mechanisms for incorporating such information descriptors will be described in later sections; here we focus on the utility of such descriptors for enhanced customization.

A typical web user has access to a variety of high fidelity audio content either in the streaming or downloaded form. Several commercial sites are attempting to exploit the popularity of audio (such as music or movie sound tracks). Therefore, there is increasing competition among these sites to use a differentiator. In audio presentation certain types of customization can lead to significantly enhanced user experience. The first such customization pertains to *spatialization* or surrounds presentation of audio. This implies creation of an immersive sound field either with the help of multiple speaker configurations (e.g., 5.1) or through a virtual surround field simulated using a two speaker configuration. Subjective tests have indicated that the level of spatialization preferred by a particular listener may vary significantly. While younger listeners respond strongly to highly immersive sound, same may not be the case for other listeners. This suggests the need of adapting the audio spatialization based on user profile or preference. To accomplish varying level of spatialization, a set of parametric sound field information needs to be extracted from a multi-channel audio source and stored with the down-mixed stereo audio as part of the databases. This parametric information thus constitutes the information descriptor in this case to facilitate adaptive multi-channel rendering. In additional to spatialization, adaptive post-processing to add color to the audio (such as warmth, detail, depth, bass, etc.), based on user preference, is also an attractive option.

Enhanced customization is also used in the area of 3-D stereo imaging, which is an application which has a notion of volume, depth, space and examples include realistic gaming where things come alive from your screen, creating 3-D stereo images for example using images from the Mars Exploration Rover Mission, assisting medical personnel in the operating room in surgical procedures, etc. 3-D Stereo rendering works by viewing 2 slightly different images from the two different eyes of a viewer. The scene is thus viewed from slightly different points of view and thus the two viewed images have two different perspectives in the scene, resulting in a 3-D effect. One way to achieve this

is to use a mechanical device with two lenses which focus each eye on a single image of the pair, thereby separating the two images. Another way is to use a red/blue or red/green filter which separates the two images which are then merged into a single image before rendering. Thus, viewing of the image needs to be adapted depending upon the viewing device (user preference) and in order to accomplish this kind of viewing some parameters need to be stored along with the different images. This parametric information thus constitutes the information descriptor in this case to facilitate 3-D stereo rendering.

Data embedding (with customization) is commonly employed in medical and health industries. For example, medical imaging systems use standards such as DICOM (Digital Imaging and Communications in Medicine) [DICOM] which separate image data from its caption (the name of the patient and physician, date, etc.). Embedding the patient name into the image itself, as a security measure, reduces the chances that the association between the patient name and image is destroyed. Of course, the benefits of this security measure over the effect of the diagnosis (due to image degradation) then become an issue. Another technique related to the healthcare industry which can be used to protect intellectual property for example in genetics involves hiding information in the DNA sequences.

Adaptive metadata is prevalent is the Intellectual Property Digital Library hosted by the World Intellectual Property Organization (WIPO) (which is an international organization dedicated to promoting the use and protection of works of the human spirit), where users can search for intellectual property data in the patent and trademark databases. Again, this searching is customized based on the user preferences, access rights, and this application benefits from the use of adaptive metadata whereby parametric information is embedded in the data itself.

3. Metadata Techniques for Incorporating Information Descriptors into Multimedia

We now discuss specific techniques for incorporating content associated information descriptors. In this section we look into the so called *metadata* techniques where the information is expressly tagged to the host data. We take a look at existing standards for metadata. We also consider the security aspects of metadata including encryption schemes. A closer look at security aspects also motivates the need for *embedded* information descriptors which will be the topic of next section.

3.1. Creating the metadata

Over the last several years many metadata frameworks [TASI] have emerged, ranging from the simplicity of the Dublin Core Metadata Initiative (DCMI) [DUB-META] to the complexity of Machine Readable Cataloguing (MARC) [MARC] and include infrastructural developments such as W3C's Resource Description Framework (RDF) [W3C-RDF]. The DCMI starts out with the basic descriptive information and expects users to customize and extend the standard to include information relevant to their needs. For example, the Visual Resource Association Core Categories, Version 3.0 [VRAC-V3], which has been designed for fine art images, have created mappings to translate their metadata into the Dublin Core. In order for different systems to work together or be interoperable [CETIS], metadata should not be randomly assigned to any resource, instead the metadata description should be based on one of the established frameworks and then mappings to the major schema should be created. Metadata creation is also affected by personalization or customization, for example, if we want to individualize the information based on the user’s preference, learning abilities, copyright

usage authority etc. which is a major metadata-related research issue in itself (Hunter 2003). This personalization or adaptation can be implemented either by explicit user preferences or by a learning system which tracks preferences and usage patterns. After the metadata is created within one of the frameworks it then needs to be associated with the image.

Metadata is normally understood to mean structured data about digital (and non-digital) resources that can be used to help support a wide range of operations such as resource description and discovery, the management of information resources (including rights management) and their long-term preservation [UKOLN]. It helps users both to discover the existence of information objects and to understand the nature of what they have found [TASI]. Metadata information can range from simple (author, date of creation, etc.) to complex (learning preferences of the user, copyright information, etc.). It consists of a set of attributes or elements which can be embedded in the resource itself or be contained in a record separate from the resource. In the context of digital resources, there exists a wide variety of metadata formats. One approach by Gilliland-Swetland (2000) states that there are five basic types of metadata:

- Administrative - Metadata used in managing and administering information resources, e.g. copyright, acquisition information
- Descriptive - Metadata used to describe or identify information resources, e.g. controlled vocabularies, user annotations
- Preservation - Metadata related to the preservation management of information resources, e.g. physical condition of resources, preservation actions
- Technical - Metadata related to how a system functions or metadata behave, e.g. digitization information such as formats, compression
- Use - Metadata related to the level and type of use of information resources, e.g. use and user tracking

An addition to these types could be educational metadata, which is increasingly being used by developers and authors to describe the pedagogic content and contexts of digital learning and teaching resources.

3.2. Metadata and Digital Images

Metadata is particularly important for digital images, where users depend on the information added to the image for accurate searching, retrieval and viewing purposes. Without this information the user may find it extremely difficult to search and retrieve the image. The metadata for images may include for example the creator of the original image, capture information, viewing instructions, file format, image resolution, copyright information, etc. Thus, this information not only helps users locate an image but also assist the user in understanding and evaluating the image. As mentioned above, there are different types of basic metadata which incorporate various aspects such as content, context, structure etc. For digital images a vast amount of information can be recorded in the metadata, however, it is important to balance the requirements with the cost of creating and maintaining it.

3.3. Security Consideration for Metadata and Encryption

Metadata contains information whose authenticity and integrity are important considerations. Even though there are preset syntax rules to generate metadata under any of the frameworks described above, it is not difficult to encode metadata which may be deliberately misleading or inaccurate. Thus, systems and users that process metadata need to consider issues related to its accuracy and validity as part of their design. Also, it may be important to provide some mechanisms for data integrity in order to protect the information from unauthorized modifications. Some of these mechanisms may include

signing of metadata with digital signatures, automating some important aspects of metadata creation, using a secret key or a combination of a public and a private key with an encryption mechanism etc. Encryption methods generally have limited robustness (Katzenbeisser et al. (2000)) to protect the embedded information against (intentional or unintentional) modifications which may occur during transmission or storage such as format conversions for images, compression of files, etc. For applications involving data embedding for copyright purposes, it is important that the embedded data be robust against manipulations that may attempt to remove this data altogether. Watermarking offers resilience against such attempts to remove the embedded information and is commonly used for proof of ownership applications. Other applications of watermarking include data monitoring, traitor tracking, fingerprinting for distribution tracking, etc.

4. Embedding Metadata in the Multimedia Content

It is easy to understand the importance of providing information to images in a digital collection and that this information needs to "travel" with the image at all times, which enables the viewing, evaluation etc. of the image. The conventional metadata is usually tagged to the original data in a separate tag field, which enables the "traveling" of the metadata with the image. Some of the common tagging formats include the International Press Telecommunication Council's [IPTC] industry standards for the interchange of news data, the Exchangeable Image File format [EXIF] for digital still cameras to embed a range of technical camera data (such as image height, width, date, time, orientation, camera make, resolution) within an image, Adobes's new XMP format [XMP], JPEG2000 format [JPEG2000], etc. However, support for image tagging is not widespread and only a limited number of applications actually use such tagged metadata. Majority of the applications simply ignore such tagged information and do not even preserve it across applications.

On the other hand, an embedded metadata (as opposed to "tagged") is a piece of information that is hidden directly in the media content, like a watermark, in such a way that it is imperceptible to a human observer, but easily detected by a computer (Parhi et al., (1999)). The principal advantage of using embedded metadata as an alternative to tagged metadata is that the content is inseparable from the metadata. Also since the metadata is embedded within the data itself, based on a particular application, it attempts to satisfy certain robustness properties whereby the embedded information can survive manipulations (both intentional and unintentional) and format changes & conversions. For example, an unintentional loss of conventional metadata may occur when changing a GIF image into a JPEG image using third party image converter/editing software that may not carry the metadata through. Alternatively, metadata may be intentionally manipulated by the user and therefore has little use as a descriptor of restrictive usage rights (e.g., copy protection). The embedded metadata, on the other hand will not only be able to resist any such manipulations and format change, but also allow for recovery of the lost metadata. This is of significant interest for many applications. These may broadly be categorized in two classes. The first class consists of applications related to the management of Intellectual Property (IP) rights, i.e., DRM. In the second class are applications which utilize the embedded metadata as a Robust Information Descriptors (RID) to provide in-place information associated with the multimedia data. The RID is useful, e.g., to enhance the experience for the end users through efficient content based indexing and retrieval of the data (for example in a DL application). Other novel applications of RID which have recently emerged are in adaptive systems which enhance user experience by customizing playback or rendering of the information. It is of course possible that a particular application may encompass ingredients of both the above classes.

As described above embedded data or watermarks are suitable for several applications including: multi-media database indexing, organization, retrieval, presentation, & distribution of digital media content, ownership identification, copyright protection, fingerprinting, traitor tracking, copy protection, authentication, broadcast monitoring, publication monitoring, secret communication, etc. In any of these applications a watermarking system has to comply with several requirements which are in general application dependent (Katzenbeisser et al. (2000)). Therefore, depending upon the type and nature of the application, the requirements and design considerations for a watermarking system will vary. However, in any of these applications a few key requirements need to be met:

- The watermark should contain sufficient information. Therefore watermarking techniques which provide higher information throughput are generally preferable.
- The watermark should be transparent. In other words the watermark should not alter the perceived quality of multimedia content beyond an acceptable limit. In many applications the acceptable level is the so called transparency level.
- The watermark should exhibit high level of robustness to intentional or unintentional attack.

5. Novel Multi-Mode Information Embedding

The proposed scheme consists of three main steps.

(1) *Adaptive Segmentation:* based on *quadtree* decomposition of image and recursive merging of similar nodes.

(2) *Mode Selection:* determination of a suitable space-frequency representation for each merged segment. We allow for one of the following 3 possibilities: Discrete Cosine Transform (*DCT*), Discrete Wavelet Transform (*DWT*), and, Space domain representation. Motivation for these modes is discussed below.

(3) *Masked Noise Power Determination:* This indicates the strength of the watermark that may be embedded into a particular segment of the image.

In steps (1) and (2) above we make use of a novel *entropy based smoothness* measure (*ESM*). Given an image segment $I(x,y)$, its *ESM*, S_I, is estimated in 2 steps, first a density function $W(x,y)$ is computed and then S_I is estimated as the entropy of the density function as summarized below.

$$W(x,y) = \left|I(x,y)\right| \Big/ \sum_{R_i} \left|I(x,y)\right| \qquad (1)$$

$$S_I = W(x,y)\log_2\left[1 \Big/ W(x,y)\right] \qquad (2)$$

5.1. Adaptive Segmentation

Segmentation of an image involves the separation of the image into regions with similar attributes. The primary attribute that we are concerned with is the susceptibility to distortion in space and frequency domain. For example a uniform intensity or textured region in the image is least sensitive to carefully introduced distortions in frequency domain but is more readily affected by any noise introduced in the space domain. An edge on the other hand is more likely to remain undistorted if the injected noise is carefully controlled in the space domain. Therefore, the primary goal of the segmentation scheme is to aid in the determination of the appropriate space-frequency representation for a particular segment of the image.

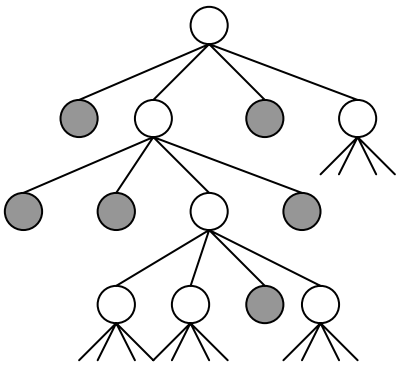

Fig. 1 *Quadtree* after the merge

The first step in this segmentation scheme is a *quadtree* (Pratt 1991) decomposition of the image. The *quadtree* decomposition was chosen for the ease of software implementation and it also aids in robust information extraction. Given an image of size *N x N*, the smallest block size for the *quadtree* leaf node, *B*, is a design parameter. Once the image has undergone the *quadtree* decomposition the *ESM* measure for each node is utilized in a recursive merge algorithm. The merge process begins at the leaf nodes and 4 child nodes of any single parent node are merged together if their entropy differs by less than a pre-determined threshold, and the entropy of the combined segment does not change significantly (see Section 6.1 for suitable value for this *merge* threshold). Any such parent node (now a leaf node) is itself merged with its siblings if all of these are leaf nodes and the merge criterion is also satisfied at this level. After working through all the levels, the merge algorithm results in an *unbalanced quadtree* (an example is shown in Fig. 1), each leaf node of which corresponds to a homogeneous region in the image.

5.2. Multi-Mode Transforms

As indicated above and in contrast to some of the previously proposed schemes, a key aspect of the proposed multi-mode scheme is the selection of an appropriate representation for each of the homogeneous leaf nodes. The goal in matching a representation to a region is to allow for the highest possible strength of the watermark. If one works entirely in the frequency domain, the watermarking ability of such a scheme is somewhat limited because spatial characteristics of the image may be affected in an uncontrolled fashion making distortion around the edges in the image more readily noticeable.

Frequency domain representation such as *DCT* is most suitable for analysis of image segments which are relatively stationary in nature (i.e., uniform in characteristics) or are highly textured. On the other hand if a segment contains edges or certain other non-stationary characteristics, a representation with higher space resolution such as *DWT* allows for a higher level of perceptually invisible distortion. Finally for certain segments with a pronounced edge it was found that leaving it un-transformed (i.e., in space domain) was the best choice.

We found that the smoothness measure *ESM* serves as a convenient tool for mode selection as well. If the *ESM* for a leaf node is less than a threshold T_1 the data is left in space domain. If ESM is larger than a threshold T_2 ($T_2 > T_1$) data is transformed into frequency domain using *DCT*. For *ESM* values in between T_1 and T_2, *DWT* is employed. Selection of T_1 and T_2 is discussed in the next section. The form of two dimensional *DCT* is amply described in (Pratt 1991). For the two dimensional DWT we chose to use Daubechies 4^{th} order wavelet with *symmetric extension* to handle the boundary conditions (Daubechies 1992).

5.3. Information Embedding Using Perceptually Shaped Noise

Information bits are embedded by adding perceptually shaped noise to each transformed leaf node of the image *quadtree*. The information rate, *R*, expressed in bits/(*BxB* block) is a free parameter for the scheme. For example R = 2 indicates that 2 bits are embedded into a leaf of size *BxB* and 8 bits are

embedded in a block of size *2B x 2B* (and so forth). The noise matrix is initially chosen from a dictionary of appropriate size containing mutually uncorrelated noise matrices. Perceptual shaping of noise is achieved by weighting it with the signal. Specifically, the form of this shaped noise addition (for different transform modes) is described below.

Shaped Noise for Space Domain

$$I'_B = I_B + \alpha \overline{N}_B \left(I_B \Big/ \| I_B \| \right) \tag{3}$$

Shaped Noise for Frequency Domain

$$IDCT'_B = IDCT_B + \alpha DCT(\overline{N}_B) \left(IDCT_B \Big/ \| IDCT_B \| \right) \tag{4}$$

where $\overline{N}_B$ is zero mean uniform noise.

The parameter α represents Signal to Mask Ratio (*SMR*) for the segment. Strictly speaking α is a function of segment characteristics. However, it was hypothesized that if the thresholds T_1 and T_2 for the multi-mode transform are properly chosen, a fixed value α (for each transform type) may be used across a wide range of images, simplifying the task of perceptual modeling. The verification of this hypothesis and determination of a suitable α was done experimentally as discussed in the next section.

5.4. Information Formatting and Error Robustness

The components described so far provide a vehicle for embedding raw information bits into the image segments. For embedding meaningful information such as database *id*, image *id,* owner *id* and buyer fingerprint, etc., we first format it into 128 bits packets. This 128 packet itself is organized as two sub-packets each starting with a 12 bits sync pattern (further explained in Section 9.2) and terminated by a 16 bit Cyclic Redundancy Check (*CRC*) which is explained in more detail in Section 8.1. Fig. 2 displays the format of a single sub-packet. Depending upon the application the same packet may be repeated multiple times within an image.

SYNC	SEQ	PAYLOAD	**CRC**
12	2	34	16

Fig. 2 Sub-Packet Format

For the purpose of efficient database indexing the payload may have one or more of the following proposed formats as shown in Table 1 below:

Database Index	User defined format
Image Index	User defined format
Category	Classification information.
Image Type	Photograph, Web graphic, Artwork, etc.
Format Type	Print, Web, etc.
Original Creator	Description of the creator including any signatures
Copyright Owner	Intellectual Property owner
License Owner	Fingerprint of the licensee

License Type	Distribution and other usage rights of the licensee: e.g., modification, sub-licensing etc.
Image Parameter	Width, Height, Date of creation, Aspect ratio, etc.
Image Viewing Guidelines	Best view guidelines for the view/playback software and/or device.
Image Processing Guidelines	Guidelines for compression, contrast modification, etc.

Table 1: Database Indexing Format

6. Design Details: Parameter Selection

This section discusses the selection of the various design parameters. The first parameter is the smallest block size for the *quadtree* leaf node, *B*, in our segmentation scheme. Segmentation subdivides an image into regions with similar attributes and the level to which this subdivision is performed is usually application dependant. We experimented with several typical values including block sizes of 8, 16, 32, etc. on a variety of images and found that choosing B to be 16 gave us enough level of detail as well kept the nontrivial computation requirement to a manageable level. Another parameter of interest based on the information bits to be embedded is the information rate, *R*, which is expressed in bits/(*BxB* block). This is the rate which eventually determines the information throughput of a particular watermarking scheme and the goal is to have a scheme with a high data payload. A higher value of R leads to higher information throughput but may result in reduced robustness. Philosophically our approach is to attempt to maintain as high data throughput as possible while maintaining a desired level of robustness. Therefore, final selection of R can only be made after carefully studying the robustness properties of the underlying watermarking algorithm. We found that proposed algorithm is able to sustain a data rate of 2-3 bits per *BxB* block for a large database of images.

Below we discuss several experiments and data related to the selection of other free parameters in the algorithm.

6.1. Entropy Smoothness Measure & Merge Threshold

Fig. 3 displays the swatches for the *ESM* for various types of image segments. It is readily apparent that for smooth or textured segments the *ESM* value is quite high and becomes lower in the presence of well defined edges. Based on the distribution of *ESM* values we chose an *ESM* difference of 0.01 as the *threshold of merge*. The recursive merge algorithm utilizes the threshold of merge to merge children nodes into their parent node. The recursive merge procedure, which is based on merging of children nodes with the ESM difference less than the threshold of merge, is applied following the first step in the segmentation process after the image has undergone the *quadtree* decomposition.

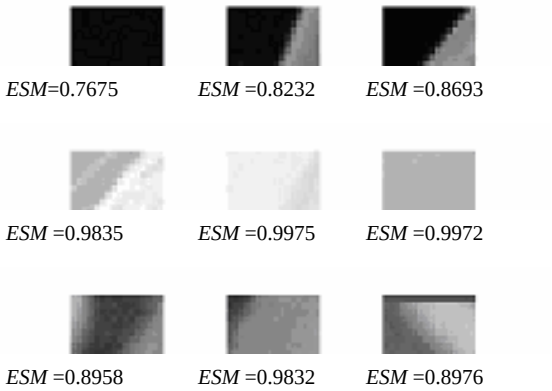

Fig. 3 *ESM* for different image segments

6.2. ESM Thresholds for Mode Selection

Another set of experiments involved finding the two *ESM* thresholds, T_1 & T_2, at which various transforms are applied. For this we subjectively determine *a*, for which image distortion becomes visible ("*detection threshold*"), as a function of the *ESM* threshold. Fig. 4 illustrates the data for *DCT* and the *DWT*. The goal is to choose T_2 such that as a high value of *a* as possible may be used across all the images. Based on this, we chose $T_2 = 0.85$ and $a(DCT) = 0.175$. Similar studies for the wavelet transform yield $T_1 = 0.55$ and $a(DWT) = 0.1$. Finally, a good value for *a*(Space Domain) was found to be 0.49.

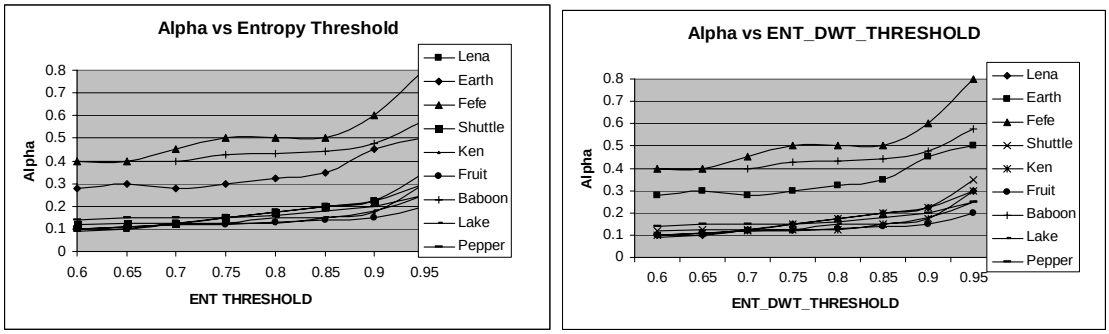

Fig 4. α v/s Entropy Threshold Trend
DCT (left) and *DWT* (right)

Figs. 5a and 5b below illustrate the end result of the complete scheme utilizing the parameter selection described before. Comparing the watermarked image based on this multi-mode scheme to original image on a high resolutions monitor reveals no perceptible distortion. The imperceptibility was verified across a large image database containing multiple *512 x 512* gray scale images.

Fig. 5a ***Boat*** Image - before and after multi-mode information embedding
(with 1568 bits of embedded information)


Fig. 5b ***Elaine*** Image - before and after multi-mode information embedding (with 1568 bits of embedded information)

7. Performance Requirements and Design Validation Methodology

As briefly discussed in Section 4**Error! Reference source not found.**, a watermarking scheme typically has to satisfy several different and often conflicting requirements. These often result in various design issues which affect the performance of the overall scheme. Below we take a closer look at the requirements to be considered in the algorithmic design process and how these translate into objective performance evaluation criteria in the form of distortion and other metrics.

7.1. Design Requirements

The most common requirement is related to watermark imperceptibility. The modifications which are caused by the watermark embedding need to be below a perceptible threshold, which implies the usage of a perceptibility criterion in the design of the watermark. Of course the best methodology for ensuring imperceptibility of distortions is through subjective evaluation. However, in a situation where a large number of algorithmic parameters need to be chosen and the algorithm validated across a large database, a quantified measure for the distortions introduced is highly desirable. The imperceptibility requirement limits the amount of modification that may be made to an image. Research suggests that this limit (in the level of modification) is very sensitive to perceptual shaping of the distortion based on the characteristics of Human Visual System (HVS). It is therefore important to design an algorithm which exploits features of the HVS (Jayant et al. (1993)) and uses a reliable assessment/evaluation. Another important consideration is that the visibility of the watermark may increase if the image undergoes attacks (an image is scaled for example). This is further discussed in Section 8 where we look into the robustness of the proposed scheme.

A second important requirement in watermarking is the robustness of the watermarked data against modifications and/or malicious attacks. These requirements are application dependent and there may be applications where robustness requirements are less important than others. In the design of proposed algorithm we considered the robustness requirement as an important requirement and this had significant influence on the choice of algorithmic structure and various parameters. In fact, robustness was the main driving factor behind the consideration to use a multi-mode transform

algorithm. It has been shown (Katzenbeisser et al. (2000)) that a transform domain technique (rather than a spatial domain one) is more efficient to use if we need a scheme which is resilient to JPEG compression, for instance. On the other hand, if we need a technique to accommodate geometrical transformations (such as scaling, rotation, etc.) then a spatial domain approach is more suitable. Our experiments suggest that by using the multi-mode approach, which matches the appropriate transform to each segment as required, it is possible to ensure robust performance for both these classes of distortions.

In addition to the above two requirements (i.e. imperceptibility and robustness), but often working in opposition to ensuring these, is the overriding goal to achieve the highest possible throughput for the embedded information. Since an embedding scheme will be of little use if it can not accommodate all the information components desired by a particular application, this (throughput) then becomes the third main requirement for a watermarking scheme.

Finally, in some applications, we need to ensure the secrecy of the embedded information, and in this case the issue of watermark security gains importance. In applications such as image database indexing, security is not a critical issue. In cases where secrecy is a requirement, a secret key can be used for the watermark embedding and extraction process.

An additional comment regarding the use of original data in the watermark recovery process is also pertinent from a requirement point of view. In certain applications availability of original data during the recovery stage can not be assumed. For example, if an application involves data monitoring etc. access to the original data may not be possible; or in some applications such as video watermarking it may be impractical to use the original data because of the large amounts of data. In the first phase of our experiments in developing this scheme we assumed that original data is available during recovery. We recently started an investigation to get around this potential deficiency in the scheme, and initial results have been promising.

7.2. Performance Evaluation and Distortion Metrics

The design and refinement of a watermarking system is often an iterative process based upon proper evaluation and benchmarking of the technique. As noted above, this requires the evaluation of the robustness and distortion in conjunction with the information throughput. Furthermore, quantitative measures are desirable to study the performance of the algorithm *w.r.t.* the key requirements. The design/study of a watermarking algorithm from a robustness perspective may make use of the following measures and considerations:

- The integrity of recovered information is conveniently measured as a Bit Error Rate (*BER*); i.e., fraction of information bits which were found to be in error under a particular situation. The *BER* is typically an averaged value across a large image database.
- The watermark embedding strength is another important measure and there is a trade-off between the strength and the perceptibility of the watermark. Increased robustness usually requires a stronger embedding, and this in turn may increase the perceptibility of the watermark.
- Both the size and the nature of the embedded data affect the robustness of a watermark. Keeping this in perspective, an image database which involves a range of image sizes (from small pictures to large ones) and types (from scanned natural images to computer generated ones) should be used during the design and validation.
- The amount of secret information (such as the seed or key used in generating the watermark) does not directly impact the robustness, however, it plays an important role in

the security of the system. Thus, the system design should follow basic cryptographic principles in generating a key.

Given the trade-off between robustness and perceptibility of the watermark, we need to consider the perceptibility in the evaluation process. We can do this either in a subjective way (best measure but sometimes impractical) or by using a quality metric. A properly designed subjective study should make use of individuals with different experiences ranging from researchers to professional photographers. Nevertheless, even the best subjective studies sometimes have a repeatability problem, i.e., these subjective tests can give varying results. Also, subjective tests are more practical for final quality evaluation rather than during initial research or in a design environment. Thus, we decided to use a quantitative distortion metric in the design and evaluation of the proposed scheme. The most popular distortion measures in the field of image and video coding are the *signal-to-noise ratio* (*SNR*) and the *peak signal-to-noise ratio* (*PSNR*) which are usually measured in decibels (Katzenbeisser et al. (2000)). It is widely accepted that the *PSNR* (in contrast to *SNR*) provides a better correlation to the sensitivities of the human visual system. In the present work we employed a variation of the *PSNR*, termed Composite PSNR (*CPSNR*) measure, which may be viewed as a normalized PSNR measure, whereby the normalization is performed to take into account the fraction of pixels which may have been modified by the watermarking algorithm. This normalization ensures that *CPSNR* can be calibrated for comparison across different images and different selection of parameters even though each case may result in a different subset of image being modified. Formulae to compute *SNR*, *PSNR*, *and*, *CPSNR* are as below:

First the Mean Squared Error (*MSE*) is computed as:

$$MSE = \frac{\sum [I(i,j) - I_w(i,j)]^2}{N^2} \qquad (5)$$

Next, the Root Mean Squared Error (*RMSE*), which is simply the square root of the *MSE* is computed and then the PSNR is computed assuming a pixel's luminance range between black (0) and white (255) as:

$$PSNR = 20\log_{10}\left(\frac{255}{RMSE}\right) \qquad (6)$$

We take this further and use a normalization based on the subset of image that has gone through some modifications made to the original image to compute the *CPSNR*. First we compute the Composite Mean Squared Error (*CMSE*) as:

$$CMSE = \frac{\sum [I(i,j) - I_w(i,j)]^2}{frac * N^2} \qquad (7)$$

where $frac = N_p / N^2$ and N_p is the total number of pixels actually modified.

The Composite *RMSE* (or *CRMSE*) which is the square root of the *CMSE* is computed next and finally the Composite *PSNR* is computed as:

$$CPSNR = 20\log_{10}\left(\frac{255}{CRMSE}\right) \qquad (8)$$

8. Robustness to Various Distortions

Distortion to the watermarked image or "attacks" can be coarsely categorized into four classes (Kutter 1999) and Voloshynovskiy et al. (2001)): Removal, Geometric, Cryptographic, and, Protocol attacks. In this section, we subject our watermarking algorithm to a wide variety of distortions from some of these attack classes and measure the robustness of the scheme. Also, as noted in Section 7.2 above, both the size and the nature of the embedded data affect the robustness of a watermark. Thus, for our robustness studies, we chose a vast image database from [USC-SIPI] which covers a range of image sizes and types. We report the results of a detailed study of the robustness of our scheme to several distortions including the JPEG compression and rotation geometric transformations. For each of the class of distortion, the watermarked images were modified with a varying magnitude of distortion.

8.1. The Cyclic Redundancy Checksum (CRC)

The robustness was measured by counting the number of total packets for which the *CRC* (Cyclic Redundancy Checksum) is correctly verified. The *CRC* [NIST-CRC] is a powerful technique for obtaining data reliability and is commonly used for detecting transmission errors. We used a 16 bits CCITT recommended *CRC* in this work. The CCITT, now known as the ITU-T (Telecommunication Standardization Sector of the International Telecommunications Union), located in Geneva, Switzerland, is the primary international body for fostering cooperative standards for telecommunications equipment and systems. Our 128-bit data packet is divided by a fixed divisor and the remainder number is appended to the message as the 16-bit *CRC* which is then sent to the watermark decoder. When the watermarked data is received, the remainder is recalculated and compared to the transmitted remainder. If the numbers do not match, an error is detected. It should be noted that meaningful decoding of the information is possible as long as at least 2 packets are error free. The number of correct packets as a function of the severity of distortion was measured for each of the images.

8.2. Robustness to JPEG Compression and Rotation

JPEG compression is currently one of the most widely used compression algorithm for still images. When images are prepared, for example, for web publishing, they are usually resized and compressed to meet various layout and bandwidth requirements. To illustrate the robustness of our scheme to JPEG compression, we used a wide range of compression percentages (also referred to as the JPEG quality factors). We also considered the effects of geometric transformations, in particular, the rotation, in our robustness studies. The results are presented below in Fig. 6 which illustrates the robustness of the scheme to JPEG compression and rotation geometric transformations.

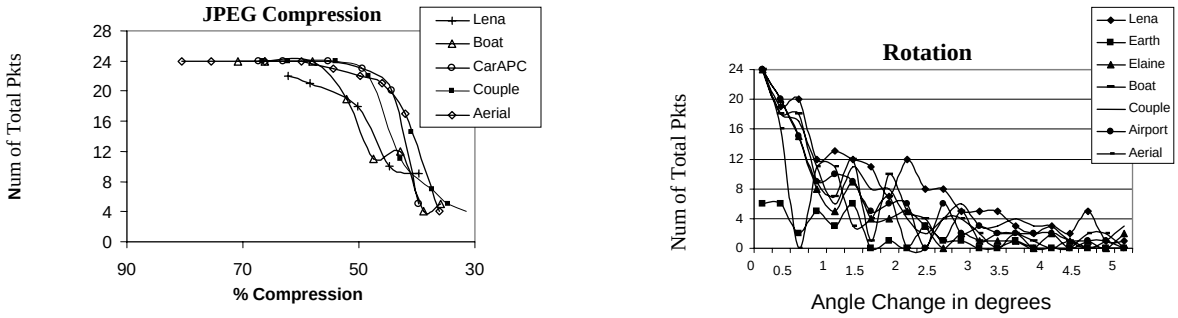

Fig. 6 Robustness to JPEG Compression and Rotation

8.3. Comparison to Other Schemes

In the evaluation and comparison of image watermarking schemes it is important to use an image database that encompasses a broad range of image sizes and types. The Signal and Image Processing Institute at the University of Southern California [USC-SIPI] maintains an extensive image database to support research in image processing and analysis. The images are of various sizes such as 256x256 pixels, 512x512 pixels, or 1024x1024 pixels. All images are 8 bits/pixel for black and white images. For our robustness studies we used a variety (both size and type) of images from this database ranging from small pictures to large pictures and computer generated ones to scanned natural ones. We believe, based on the review of other published watermarking schemes, that utilization of such a standardized database is not the norm for these schemes. Therefore, validation of the proposed watermarking scheme using images from this standard database provides a unique affirmation in comparison to other schemes.

Another distinguishing aspect of this work is that robustness studies were performed across a wider range of attack classes. Watermarking attacks can be coarsely categorized into four classes of attacks (Voloshynovskiy et al. (2001)): Removal attacks, Geometric attacks, Cryptographic attacks, and, Protocol attacks. Removal attacks aim at the complete removal of the watermark information from the watermarked data and this category includes JPEG compression. In contrast to removal attacks, geometric attacks do not actually remove the watermark, but introduce distortion, as in the case of image rotation, image scaling, etc. Cryptographic attacks try to break the security by using a brute-force search to determine the embedded information and then remove the information or embed a misleading watermark. Of course, these attacks are restrictive because of the computational complexity involved in the brute-force search methods. The protocol attacks try to attack the entire watermarking application concept, for example, the copy attack not only destroys the watermark but also estimates it from the watermarked data and then copies it to other data. It is interesting to note that content based watermarking can improve the resistance to such attacks. In our robustness studies, we studied robustness to both the removal and geometric attack classes. Although impact of removal attacks has been studied widely by workers of the other watermarking studies, studies involving impact of geometric distortions, including rotation, are hard to find for most if not all of these schemes. Our studies in this area (i.e. the robustness of the proposed scheme to geometric distortions), therefore provides further validation for the scheme. Also, the cryptographic and protocol attacks depend upon the particular application and robustness studies for such attacks are currently underway and we will report these in future publications.

Finally, we believe the key distinguishing aspect of the proposed scheme is the high information throughput afforded by the scheme. The multi-mode embedding allows for a natural perceptual shaping of the introduced distortion allowing the scheme to hide significantly higher level of distortion than is possible by previously reported techniques which operate in a fixed domain. This then results in significantly higher information throughput. Among the previously reported schemes which allow for high information throughput, the scheme reported by Miller et al. (Miller 2004) is particularly noteworthy and has been reported to embed about 1380 information bits in a typical image. In comparison, the proposed scheme is capable of embedding at least 1500-2048 information bits in the image without any perceptible distortion. We believe, that the information throughput will move even higher once information compacting schemes similar to Miller 2004 have been incorporated and the perception model used in the scheme is further refined.

9. Components of a Digital Rights Management System

In this section we focus on DRM application of the watermarking scheme. As noted in Section 1, watermarking is only one of the needed components in a DRM System. In addition, for a practical DRM System some other components related to information massaging, such as, Encryption, Channel Coding and Synchronization are also needed. In this section we provide details on some of these additional components.

9.1. Selection of Information Pattern and Embedding Location

To make the watermark robust to attacks such as geometric operations and scaling, and border area elimination, careful selection of the information pattern and embedding location is necessary. The information pattern is derived from a random independent and identically distributed process for increased image quality. For fingerprinting applications, this random sequence may be generated using unique receiver identification as a key. In terms of embedding location the following scheme is used. Since elimination of a suitably chosen border region often results in an acceptable loss in terms of image usability, it is conceivable that border elimination will often be utilized by a malicious attacker.

It is therefore convenient to work outwards from a properly identified centroid of the original image. The information carrying blocks (as identified by the entropy criterion in Section 6.1 above) may then be sequentially arranged, for example, by traversing in a counter-clockwise direction and always picking up the next closest block to the centroid. This represents an outward spiral scan of the image. Within a particular block itself, the information carriers may be arranged in any suitable order. In general the intra-block ordering depends on the characteristics of the block (i.e., time domain, wavelet, or, DCT). Associated with each block is a choice function $C(n; k)$ which selects k out of an initially identified n perceptually significant potential carriers as the actual carriers. In a particular $C(n; k)$ the ratio k/n depends on the desired information rate, the function otherwise may be random and based on a *key*. In the proposed scheme, a particular key is used for watermarking, and a different, user dependent key, selected from a finite predetermined set, is used for fingerprinting.

9.2. Synchronization Mechanism

For the purpose of synchronization a fixed bit sync pattern is added to the information sequence. This pattern may also function as a universal/private key during the watermark detection stage. The pattern is set by the original owners of the media and a pattern is also provided as part of a key to legal customers for the purpose of fingerprinting. The embedded bits therefore are formatted as information packets consisting of a synchronization pattern followed by the payload. For our studies, we used a 12 bit Barker Sequence (Ziemer 1995) to identify the start of the data. The actual payload itself is broken down into a length field followed by the actual information bits. Depending upon the available information bandwidth in an image and the desired throughput, a particular information packet may be repeated several times for increased robustness.

9.3. Encryption

Design considerations for DRM System are based on its applications. In applications such as copyright protection, the secrecy of the embedded information needs to be maintained and security is an issue. In other applications (for e.g. image database indexing), security is not really an issue. Thus, if secrecy is a requirement, then a secret key must be used for the embedding (and extraction) process.

It is possible to combine one or several public keys with a private (secret) key (Hartung 1997)and embed a public/private watermark, whereby the image can be watermarked using a private key, and the public key can be used to verify the watermark. Several public key watermarking algorithms are discussed in (Wong 1998). For applications where secrecy is a requirement, a rudimentary form of encryption is provided by the synchronization pattern as well as the key used in choosing information embedding location. In addition the length information in each information packet is encrypted using a standard 16 bit encryption scheme, while the actual payload is encrypted using a 16-64 bit key (based on the desired information throughput).

9.4. Channel Coding

In the design of the watermark detector, we take into account the basics of communication theory, since; recent developments in the theoretical analysis of digital watermarking schemes have shown a direct link between digital watermarking and communications theory (Martin 2001). Thus, channel coding techniques (such as the soft decision version of the Viterbi algorithm for convolution codes) are applied in watermarking context to improve the performance of the watermark extractor. Channel coding may be used to detect and/or correct potential errors in the detected watermark. Error correcting codes such as convolution codes and Reed-Solomon codes are example of some popular classes of channel codes in practical communication systems. In our scheme the information pattern is coded with the help of a rate 1/3, convolution channel code as described in (Lin 1983). One of the most common problems for the detection algorithm in a watermarking scheme is that of information deletion and/or addition (of spurious information). We attempt to correct for these situations with the help of soft decision channel decoding algorithms. The information is organized in relatively small fixed (and known) length information sub-blocks which are first tagged with an error detection code such as the CRC (as described in Section 8.1) and then coded further with the convolution code.

10. Discussions and Directions for Future Research

An information embedding scheme for digital images using a multi-mode based on adaptive segmentation approach has been presented. Potential applications of this are in multimedia database indexing (for images) and Digital Rights Management and can easily be extended to other multimedia applications. Extraction of the embedded information is performed using a maximum likelihood detection algorithm. The proposed scheme provides the framework of a high throughput information embedding scheme that is robust to a wide class of distortions. Here we briefly discuss some of the areas for ongoing/future research. In the results presented above it is assumed that original image is available at the time of extraction. As noted above, in certain applications it may be desirable to extract the message without access to original image. It is possible to design a practical, though potentially less robust, detector for such applications. In general a grid search algorithm is needed to extract information from a (potentially) distorted image, but exhaustive search may be prohibitively expensive. Therefore, our future work in this direction is based on developing fast heuristic search schemes to facilitate robust information extraction without access to the original image. Another interesting area for research is the development of suitable coding schemes for the embedded information which can then be combined with powerful iterative decoding schemes such as turbo-coding to increase the robustness of the scheme (Proakis (2001)). In the last few years several benchmarking technologies have emerged in order to attempt to better evaluate watermarking techniques. StirMark [STIR-MARK], Checkmark [CHECK-MARK] and Optimark [OPTI-MARK]

are some of these technologies. Our future work direction also involves the possibility of testing our scheme with one or more of these benchmarks.

The multi-mode embedding framework is also attractive for information embedding in other multimedia content such as audio and video and we plan to evaluate the scheme for these classes of signals in the future.

References

Barni, M. and Bartolini, F. “Data Hiding for Fighting Piracy” *IEEE Signal Processing Magazine*, Vol. 21, No. 2, March 2004.

Daubechies, I. (1992) “Ten lectures on wavelets” CBMS-NSF conference series in applied mathematics, SIAM Ed.

Gilliland-Swetland, A.J. (2000). *Setting the stage: Defining metadata.* In Baca, M. (Ed.) Introduction to Metadata: Pathways to Digital Information, 2nd ed. (Los Angeles: Getty Information Institute).

Hartung, F. and Girod, B. (1997) “Fast Public-Key Watermarking of Compressed Video” *Proceedings IEEE International Conference on Image Processing*, Vol.1, Santa Barbara, CA, October 1997.

Hunter, J. (2003) “Working towards MetaUtopia - A Survey of Current Metadata Research” *Library Trends, Organizing the Internet*. 52 (2).

Jayant, N., Johnston, J. and Safranek, R. (1993): “Signal Compression based on Models of Human Perception,” *Proc. IEEE*, 81(10), 1993.

Katzenbeisser, S. and Petitcolas, F.A.P. (2000) *Information Hiding Techniques for Steganography and Digital Watermarking* (Boston: Artech House, Inc.)

Kuo, J. C., Kalker, T., and Zhou, W. (2004) “Digital Rights Management” *IEEE Signal Processing Magazine*, Vol. 21, No. 2, March 2004.

Kutter, M. and Petitcolas, F. (1999) “A Fair Benchmark for Image Watermarking Systems”. *Electronic Imaging ’99: Security and Watermarking of Multimedia Contents*, vol. 3657, San Jose, CA, 226-239

Lin, S. and Costello, D.J. (1983) *Error Control Coding* (New Jersey: Prentice Hall).

Martin, J.R.H and Kutter, M. (2001) “Information Retrieval in Digital Watermarking” *IEEE Communications Magazine*, pp. 110-116, August 2001.

Miller, M. L., Doerr, G. J., and Cox, I. J. (2004) "Applying Informed Coding and Embedding to Design a Robust, High capacity Watermark" *IEEE Trans. on Image Processing*, 13, 6, 792-807.

Moulin, P., Kalker, T., Cox, I., Dittmann, J., Duhamel, P., Jain, A., Katzenbeisser, S., and Lagendijk, R. (2004) "Guest Editorial: Supplement on Secure Media – I" *IEEE Transactions on Signal Processing*, Vol. 52, No. 10, October 2004.

Parhi, K.K. and Nishitani, T., Ed. (1999), *Digital Signal processing for Multimedia Systems* (New York: Marcel Dekker).

Pratt, W.K. (1991) *Digital Image Processing* (New York: John Wiley & Sons, Inc.)

Proakis, J. G. (2001) *Digital Communications* (New York: McGraw-Hill Companies, Inc.).

Sinha, N. (2000), "A Novel Watermarking Technique for digital images based on Adaptive Segmentation and Space-Frequency representation" *Proc. IEEE 2000 International Symposium on Information Theory and its Applications, ISITA 2000*, Vol. II, 972-975.

Voloshynovskiy, S., Pereira, S., Pun, T., Eggers, J. and Su, J. (2001): "Attacks on Digital Watermarks: Classification, Estimation-Based Attacks, and Benchmarks," *IEEE Communications Magazine*, pp. 118-126, August 2001.

Wong, P.W. (1998) "A Public Key Watermark for Image Verification and Authentication" *Proceedings of the International Conference on Image Processing*, Vol. 1, Chicago, IL, Oct 1998

Ziemer, R.E. and Tranter W.H. (1995) *Principles of Communications, Systems, Modulations and Noise* (New Jersey: John Wiley & Sons).

Links

[CHECK-MARK] Checkmark Benchmarking at http://watermarking.unige.ch/Checkmark/

[CETIS] CETIS – Centre for Educational Technology Interoperability Standards at http://www.cetis.ac.uk/.

[DICOM] Digital Imaging and Communications in Medicine at http://medical.nema.org/.

[DMCA-USCOPY] The Digital Millennium Copyright Act of 1998, US Copyright Office Summary, 1998, at http://www.copyright.gov/legislation/dmca.pdf

[DUB-META] Dublin Core Metadata Initiative at http://www.dublincore.org/about/.

[EXIF] Exchangeable Image File Format at http://it.jeita.or.jp/document/publica/standard/exif/english/jeida49e.htm.

[IPTC] International Press Telecommunications Council at http://www.iptc.org/pages/index.php.

[JPEG2000] JPEG 2000 format at http://www.jpeg.org/jpeg2000/.

[MARC] MARC standards at http://www.loc.gov/marc/.

[NIST-CRC] National Institute of Standards and Technology – Cyclic Redundancy Check, at http://www.nist.gov/dads/HTML/cyclicRedundancyCheck.html

[OPTI-MARK] Optimark Benchmarking at http://poseidon.csd.auth.gr/optimark/

[STIR-MARK] Stirmark Benchmark at http://www.petitcolas.net/fabien/watermarking/stirmark/

[TASI] Technical Advisory Service for Images at http://www.tasi.ac.uk/index.html.

[UKOLN] UKOLN Metadata at http://www.ukoln.ac.uk/metadata/.

[USC-SIPI] University of Southern California - Signal & Image Processing Institute - The USC-SIPI Image Database at http://sipi.usc.edu/services/database/Database.html

[VRAC-V3] Visual Resource Association Core Version 3 at http://www.vraweb.org/vracore3.htm.

[W3C-RDF] World Wide Web Consortium Resource Description Framework at http://www.w3.org/RDF/

[WIPO-DRM] WORLD INTELLECTUAL PROPERTY ORGANIZATION standing committee on copyright and related rights Tenth Session Geneva, November 3 to 5, 2003 “Current Developments in the field of Digital Rights Management”, SCCR/10/2 Rev. May 2004 at http://www.wipo.int/documents/en/meetings/2003/sccr/doc/sccr_10_2_rev.doc.

[XMP-ADOBE] Adobe Extensible Metadata Platform (XMP): Adding Intelligence to Media at http://www.adobe.com/products/xmp/main.html.

Author Information

Neelu Sinha received her MS and PhD in Electrical Engineering from Iowa State University in Ames, Iowa, USA. She is currently an Assistant Professor in the Department of Computer Science, Mathematics and Physics at Fairleigh Dickinson University in Madison, New Jersey. In the past she has worked for Control Data Corporation in Plymouth, MN and the Bell Laboratories in Whippany, NJ. Her current areas of interest are digital image and audio watermarking, information security, multimedia database indexing and content based retrieval.