



Department of Information Technology  
Lappeenranta University of Technology

# **A Technique for Digital Color Image Watermarking Using ICA**

The Department Council of the  
Department of Information  
Technology confirmed topic of this  
Master's Thesis on 22.03.2006

**Examiner:** Professor Arto Kaarna  
**Supervisor:** Vladimir Botchko, Ph.D.

**Author:** Venyamin Chebotarev

**Address:** Kirstinkatu 13 C82  
00510 Helsinki,  
Finland

**Phone:** +358 50 4080186

**Lappeenranta  
2006**

## **ABSTRACT**

Lappeenranta University of Technology  
Department of Information Technology

Venyamin Chebotarev

### **A Technique for Digital Color Image Watermarking Using ICA**

Thesis for the Degree of Master of Science in Technology

Examiner: Professor Arto Kaarna

Supervisor: Vladimir Botchko, Ph. D.

44 Pages, 17 figures, 4 tables, 1 appendices.

Keywords: *watermarking, PCA, ICA, spectral image*

With the increase of use of digital media the need for the methods of multimedia protection becomes extremely important. The number of the solutions to the problem from encryption to watermarking is large and is growing every year. In this work digital image watermarking is considered, specifically a novel method of digital watermarking of color and spectral images.

An overview of existing methods watermarking of color and grayscale images is given in the paper. Methods using independent component analysis (ICA) for detection and the ones using discrete wavelet transform (DWT) and discrete cosine transform (DCT) are considered in more detail.

A novel method of watermarking proposed in this paper allows embedding of a color or spectral watermark image into color or spectral image consequently and successful extraction of the watermark out of the resultant watermarked image.

A number of experiments have been performed on the quality of extraction depending on the parameters of the embedding procedure. Another set of experiments included the test of the robustness of the algorithm proposed. Three techniques have been chosen for that purpose: median filter, low-pass filter (LPF) and discrete cosine transform (DCT), which are a part of a widely known StirMark - Image Watermarking Robustness Test.

The study shows that the proposed watermarking technique is fragile, i.e. watermark is altered by simple image processing operations. Moreover, we have found that the contents of the image to be watermarked do not affect the quality of the extraction. Mixing coefficients, that determine the amount of the key and watermark image in the result, should not exceed 1% of the original. The algorithm proposed has proven to be successful in the task of watermark embedding and extraction.

# TABLE OF CONTENTS

ABSTRACT .....	II
TIIVISTELMÄ .....	ERROR! BOOKMARK NOT DEFINED.
TABLE OF CONTENTS .....	III
LIST OF ABBREVIATIONS .....	V
ACKNOWLEDGMENTS .....	VI
1. INTRODUCTION .....	1
2. BACKGROUND .....	4
Applications .....	4
Requirements .....	6
3. TECHNIQUES USED IN WATERMARKING .....	8
PCA (Principle Component Analysis) .....	8
ICA (Independent Component Analysis) .....	9
Wavelet transform .....	14
4. ATTACKS .....	16
LPF .....	18
Median Filtering .....	18
DCT .....	18
5. WATERMARKING FOR COLOR AND GRAYSCALE IMAGES .....	20
A Method Using ICA for Detection .....	20
A Scheme Based on DWT and DCT .....	21

<b>6. SPECTRAL IMAGES.....</b>	<b>23</b>
<b>7. A TECHNIQUE FOR DIGITAL COLOR IMAGE WATERMARKING USING ICA .....</b>	<b>25</b>
<b>Embedding .....</b>	<b>25</b>
<b>Extraction .....</b>	<b>27</b>
<b>8. EXPERIMENT .....</b>	<b>34</b>
<b>Attack influence .....</b>	<b>40</b>
<b>9. CONCLUSIONS .....</b>	<b>42</b>
<b>REFERENCES .....</b>	<b>44</b>
<b>APPENDIX 1 SPECTRAL IMAGE DATABASE .....</b>	<b>1</b>

## LIST OF ABBREVIATIONS

BSS - blind source separation

CC – correlation coefficient

CIE - Commission Internationale de l'Eclairage

CIELAB - CIE  $L^*a^*b^*$  color space

CIELUV – CIE  $L^*u^*v^*$  color space

DCT - discrete cosine transform

DWT - discrete wavelet transform

FFT - fast Fourier transform

HVS – human visual system

ICA - independent component analysis

ICs - independent components

IR - infrared

LPF - low-pass filter

PCA - principal component analysis

RGB – Red, Green, Blue

UV - ultraviolet

WD - wavelet decomposition

## **ACKNOWLEDGMENTS**

This thesis has been done in the Laboratory of Information Processing in the Department of Information Technology of the Lappeenranta University of Technology, Finland.

First of all, I would like to express my deep gratitude to my supervisors, Vladimir Botchko and Arto Kaarna, for their guidance and valuable help during my work.

I thank the organizers of IMPIT program for the opportunity to study in the Lappeenranta University of Technology.

I wish to thank all of my friends in IMPIT program for so many memories, wonderful atmosphere and help in every day problems.

My deepest affection belongs to my parents and my brothers for their support and help, without them this study would not have been possible. And finally, I would like to thank Diana Kalenova for her support, encouragement and all of the advice both during studies and thesis writing.

# 1. Introduction

Digital media development is undergoing dramatic changes nowadays. Digitized information is used in every possible area of our life, from medicine to home finance. Images, texts, audio files, video files contain information normally stored on conventional media. These are now easily shared using the possibilities of the Internet. However, all of these advances lead to a serious problem of security and copyright protection. One of the most widely used copyright protection methods is digital watermarking. Digital watermarking can be defined as imperceptible insertion of information into multimedia data. A more formal definition can be given as a communication method in which information  $h$  is embedded directly and imperceptibly into digital data  $x$  (e.g., image, video, or audio signals), also called original data or host data, to form watermarked data  $y$  [17]. We will limit the scope of this thesis to digital image watermarking, in particular color and spectral images.

Digital watermarking can be generally applied in three main areas [18, 19]: conveying ownership information, verifying that object content has not changed, and conveying object-specific information to a community of recipients (collaborative watermarking application) [17]. Applications that convey ownership information are generally used by organizations that own copyrights or licenses such as news agencies, photo banks, museums. Applications that verify object content use the watermark to determine if the image was somehow modified and help localize the change which can be used e.g. in digital libraries. Collaborative watermarking applications use watermarks to deliver object-specific information to a community of recipients, e.g. automating of accounting of royalties for broadcasting audio content, protection of digital video content [17]. These different areas of application gave rise to the need in different categories of watermarks.

There are several algorithms of watermarking; most of them are intended for dealing with grayscale watermarks. Two types of watermarks exist, visible watermarks and invisible ones. In visible mode, as the name implies, you can

easily see the watermark, which is usually some visual message or a company logo. Visible watermarks should be protected against removal by unauthorized parties. At the same time the watermarks have to resist falsification, meaning that, since it is relatively easy to insert this type of watermark into an image, we have to provide a method to insure that the watermark was inserted by the owner of the image. In invisible mode the watermarked image is perceptually identical to the original one. The watermark in such an image, depending on the embedding procedure, can be extracted using a special algorithm along with additional data (key image) [16].

Another classification is by the ability to corrupt the watermark by some image processing procedure. In this respect three types exist: fragile, semi-fragile, and robust. A fragile watermark is placed in every pixel of the image so, that the slightest change to the image removes the watermark. Watermark here is extracted from every pixel and if that is not possible the image was potentially modified.

In semi-fragile techniques the watermark remains the same as long as the image content does not change while image representation can change. By image representation we mean different standards of image compression (e.g. from GIF to JPEG) [21]. Robust watermark is recoverable despite intentional or unintentional modifications [16]. This kind of watermarking can be performed both in spatial and frequency domains. Several requirements are imposed upon robust watermarking techniques: perceptual transparency, data capacity, robust to unintentional image processing operations, tamper resistance, computational complexity. Robust watermarking techniques can further be classified into private and public. Private methods, also called informed, require the original image for the detection of the watermark while public (oblivious or blind) do not require one [21].

In this work, a novel technique for digital watermarking of color and spectral images using ICA is proposed. A number of methods using grayscale images exist at the moment; however, a color watermark has previously never been used. The aim of the work is to create an algorithm that would allow extracting previously



embedded color digital watermark into color and spectral images. The idea is to create a robust algorithm, which is applicable when watermarked image was subjected to some modifications (filtering, transformation, cropping, etc.). A set of experiments is dedicated to testing robustness against various types of attacks.

In the method proposed in this paper the following techniques are applied: Principal Component Analysis (PCA), Wavelet Decomposition (WD) and ICA. PCA allows us to use a few layers with the largest eigenvalues of eigenvectors. The combination of PCA and WD allows avoiding overlearning when applying ICA for final separation of the original image and the watermark.

## **2. Background**

Digital signals, which can be used as sign, or marks, embedded into digital data are called watermarks. Any types of data are applicable for such purpose (e.g. images, letters, signs, numbers, etc.) [13].

There are several categories for classification of watermark data. Watermarks can be “perceptible” and “imperceptible”. In case of pictures these are called “visible” and “invisible”. A visible watermark is used to denote the owner of a picture. In most cases it contains a logo or any kind of visual message indicating the ownership. Visual inspection is enough for detection of a visible watermark. The invisible watermark may contain the same data like the visible one, but invisible watermarked image does not have any differences compared with original one for the human eye. This type of watermarks can be determined using a number of special algorithms for detection or extraction. Next we will concentrate primarily on invisible watermarks.

### ***Applications***

Digital image watermarking can be applied in a number of areas. Here we will consider several main areas that lead to the development of some specific watermarking methods and tools.

#### **2.1 Fingerprinting:**

A watermark can be used to trace the source of breach of the copyright agreement. Different watermarking keys are embedded into the copies provided to different customers. A unique watermark guarantees the possibility of finding the customer providing the copies of the product illegally to a third party.

#### **2.2 Indexing:**

Watermarking is widely applied in a number of multimedia applications not only for the purpose of protection of the information but also to insert some kind of indexing or comments into video or other multimedia content. As the use of the Internet and in particular search engines increases indexing would simplify access to the data, and watermarking is useful in this respect.

### **2.3 Copyright Protection & Owner identification:**

This application is connected with Fingerprinting, but this is a more generalized area. Watermarking primary use is protection of data copyright information. A watermark representing the information about the owner is inserted into the data to be able to prove the ownership right.

### **2.4 Broadcast monitoring:**

Watermarking can be used to help automate the identification of broadcast programs. A watermark is inserted into the data broadcasted over the network, be it radio or TV. It assures that advertisers receive the airtime they have paid for or the property rights for music or video data.

### **2.5 Copy protection:**

The watermark can also be used to prevent unauthorized copying of the information, which can be detected by the recording or copying devices. The embedded key is a bit stream giving the copy-write information.

### **2.6 Data Authentication:**

Data authentication application seems similar to the one described in section 2.3. Watermarking allows detection of changes or processing introduced into the image. For this purpose a fragile watermark is used, that would not be extracted if the data is corrupted.

### **2.7 Data Hiding (Covert Communications):**

This area of application is probably one of the oldest. In this case the watermark is the carrier of an important message, which is implanted and somehow coded to be able to transmit it in such a way that it would not be detected by the unauthorized person.

### **2.8 Medical Safety:**

This particular area of application implies that patient's name and some other data concerning the patient is embedded into the medical images, that would simplify medial information search and improve security.

This list of applications does not contain all of the possible areas of application of watermarking, which are way too numerous, however, it contains most of the ones we have found useful or interesting.

## ***Requirements***

Certain requirements have to be imposed upon watermarking. These are common for any types of the algorithm independent of the details of the implementation. These are primarily:

### **Perceptual Transparency:**

This requirement means that the embedded watermark should not be visible for the human eye under typical viewing conditions. The degradation of the perceived quality should not affect the appearance of the image. Of course there are watermarks that are meant to be seen – the “visible” watermarks. However, the requirement we have stated applies to “invisible” watermark algorithms.

### **Robustness:**

The watermark should be robust to intentional or unintentional modifications in the images. The ideal requirement is that the amount of the distortion introduced into the image that affects the watermark is such that the commercial value of the image diminishes significantly.

### **Capacity:**

This is not an absolutely obligatory requirement, it means the possibility of insertion of several independently detectable watermarks in an image.

### **Payload of watermark:**

Payload represents the amount of information that can actually be stored in a particular data stream. This requirement is highly dependent on the host medium,

the intended application as well the quality aimed for. By the addition of for example the year of copyright, the permissions granted on the work and the rating for it, the total number of bits to be embedded should be from 60 to 70 bits. In the design of watermarking scheme, some attention must be given to the fact that this is probably the minimal amount of information that one has to be embedded, therefore, host data should be able to support it while keeping good perceptual transparency.

**Security:**

Watermarking security lies on Kerckhoff assumption that one should assume that the method used to encrypt the data is known to the unauthorized party. This means that watermarking security can be interpreted as encryption security leading directly to the principle that it must lie mainly in the choice of the embedded key.

**Specificity:**

Watermark should be universal, i.e. applicable to any kind of data like images as well as audio and video media. However, this requirement has one serious limitation that the size of the watermark should be appropriate for the insertion into the particular type of data.

Two ways of embedding watermark, that make it invisible are used:

**In the spatial domain.**

The first watermarking scheme that was introduced in this area works directly in the spatial domain. Some image analysis operations (e.g. Edge detection) provide the perceptual information about the image, which is then used to embed a watermarking key, directly in the intensity values of predetermined regions of the image. This method provides a simple and effective way for embedding an invisible watermark into an original image but does not show robustness to common image processing algorithms.

**In the transform domain.**

Another way to produce high quality watermarked image is by first transforming the original image into the frequency domain by the use of Fast Fourier (FFT),

Discrete Cosine (DCT) or Discrete Wavelet (DWT) transforms, or even a combination of these. With this technique, the marks are not added to the intensities of the image but to the values of its transform coefficients. Then inverse transformation with the marked coefficients forms the watermarked image. The use of frequency based transforms allows the direct understanding of the content of the image; therefore, characteristics of the human visual system can be taken into account more easily when it is time to decide the intensity and position of the watermarks to be applied to a given image.

### 3. Techniques Used In Watermarking

#### ***PCA (Principle Component Analysis)***

“In statistics, **principal components analysis (PCA)** is a technique that can be used to simplify a dataset; more formally it is a linear transformation that chooses a new coordinate system for the data set such that the greatest variance by any projection of the data set comes to lie on the first axis (then called the first principal component), the second greatest variance on the second axis, and so on [24].”

PCA is widely used in machine learning and pattern recognition tasks, such as identifying patterns in data, or expressing data as to highlight the patterns in the dataset.

Assuming a zero mean of a two-dimensional dataset  $(x,y)$  PCA is computed as follows:

1. Compute the covariance of  $x$  and  $y$

$$C(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n - 1)} \quad (1)$$

where  $C$ , denoted by  $c_{ij}$ , is a covariance between random variable components  $(x_i, y_m)$  and  $(x_j, y_n)$ ,  $\bar{x}$  and  $\bar{y}$  are means of  $x$  and  $y$ , and  $n$  is the dimensionality of both.

## 2. Calculate the eigenvectors and eigenvalues of the covariance matrix

Since the covariance matrix is square, we can calculate the eigenvectors and eigenvalues for this matrix. Eigenvectors are a special set of vectors associated with a linear system of equations (i.e., a matrix equation) that are sometimes also known as characteristic vectors. Having a square matrix  $A$ , the following equation defines eigenvalues and eigenvectors:

$$Ax = \lambda \hat{x} \quad (2)$$

where  $\lambda$  is called and eigenvalue of the matrix  $A$  and  $\hat{x}$  is the eigenvector corresponding to the eigenvalue  $\lambda$ .

## 3. Choosing components and forming a feature vector

The eigenvector with the largest eigenvalue is the first principle component of the data set. Once eigenvectors are found from the covariance matrix, the next step is to order them by eigenvalue, highest to lowest. This gives you the components in order of significance. Thus, a principal component  $w_1$  of a dataset  $x$  can be defined as follows:

$$w_1 = \arg \max_{\|w\|=1} E\{(w^T x)^2\} \quad (3)$$

With the first  $k - 1$  components, the  $k$ -th component can be found by subtracting the first  $k - 1$  principal components from  $x$ , where  $E\{\}$  defines the expectation value.

## **ICA (Independent Component Analysis)**

”Independent Component Analysis (ICA) is a method for finding underlying factors or components from multivariate (multidimensional) statistic data. It looks for the components that are both statistically independent and non-gaussian” [9].

Given that a set of observations of random variables  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  was generated as a linear mixture of independent components  $\mathbf{s} = \{s_1, s_2, \dots, s_m\}$ :

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = A \begin{pmatrix} s_1 \\ s_2 \\ \dots \\ s_m \end{pmatrix} \quad (4)$$

where  $A$  is unknown matrix, called mixing matrix, ICA consists of estimating both the matrix  $A$  and a random vector  $\mathbf{s}$  based on the observation vector  $\mathbf{x}$ . Generally the number of observations  $n$  can be greater or equal to the number of independent components  $m$ ,  $S$  are independent sources

The model described above defines a linear ICA. Despite its seeming simplicity this model can be applied in a number of areas, in particular at: Blind Source Separation (BSS) and finding basis vectors from data samples for feature extraction. BSS is widely applied in telecommunications, speech recognition and watermarking [8].

Feature extraction in turn is one of the most widely known image or signal processing tasks. ICA assists in this matter in finding orthogonal basis form of data samples, which in itself is aimed to a new linear transformation.

The main advantage of ICA over similar techniques like PCA is that it searches for independent rather than uncorrelated components and works with non-Gaussian random variables. Statistical independence and uncorrelatedness is the same for Gaussian variables, but for non-gaussian variables independence is stronger. Thus, the first stage of ICA is whitening, that leads to decorrelation of random variables using their covariance matrix. Whitening restricts the separating matrix  $W$  to a class of orthogonal matrices, where  $W$  is basically the inverse of  $A$  given above. [7].

There are several methods for finding the independent components, but two of them are basic. First method considers independence as a nonlinear non-correlatedness: If  $s_1$  and  $s_2$  are independent, then any nonlinear transformation  $g(s_1)$  and  $h(s_2)$  are uncorrelated. The nonlinear functions  $g()$  and  $h()$  can be chosen using the principles of the estimation theory and information theory. The



method taken from the estimation theory is maximum likelihood, while information theory provides the measures of independence, such as mutual information. These ideas are employed in the Fast ICA algorithm.

Second approach attempts to look for maximum non-gaussianity. The idea is that according to central limit theorem, the mixture of linear of non-Gaussian random variables is Gaussian. Independent components produce mixture that is non-Gaussian and independent. To measure non-gaussianity kurtosis can be used [9].

ICA cannot be applied to the task of Gaussian random variable separation, due to the fact that the whitened orthogonal mixing matrix does not affect the joint distribution of the observed variables. Thus, original and mixed distributions are identical and there is no way to infer the mixing matrix from the mixture. In order to separate independent Gaussian variables PCA has to be used [9].

#### **Problems Related to ICA:**

1. The mean value of signals is lost due to data centering.
2. The variance of independent components cannot be found. The reason is that, following Eq. 4 both  $\mathbf{s}$  and  $\mathbf{A}$  are unknown. Thus, any scalar multiplied by one of the sources can always be canceled by dividing the corresponding column  $\mathbf{a}_i$  of  $\mathbf{A}$  by the same scalar:

$$\mathbf{x} = \sum_i \left( \frac{1}{\alpha_i} \mathbf{a}_i \right) (s_i \alpha_i) \quad (5)$$

To overcome this lack, variance of observed and independent random variables is restricted to unit, but ambiguity of the sign is still left.

3. The resultant signals can be inverse, and image can have negative contrast.
4. The order of independent components cannot be determined. If the order of elements of vector  $s$  changes, the permutation of columns in matrix  $A$  will compensate this [9].

#### **Fast Fixed-Point ICA Algorithm (FastICA)**

The FastICA algorithm is one of the algorithms that separate independent components based on their non-gaussianity and independency of their linear

transformation. Measures such as kurtosis, and negentropy define gaussianity, and express the likelihood of independent components. The basic form of the FastICA algorithm is as follows [23]:

1. Choose an initial (e.g. random) weight vector  $\mathbf{w}$ .
2. Let  $\mathbf{w}^+ = E\{\mathbf{x}g(\mathbf{w}^T \mathbf{x})\} - E\{g'(\mathbf{w}^T \mathbf{x})\}\mathbf{w}$
3. Let  $\mathbf{w} = \mathbf{w}^+ / \|\mathbf{w}^+\|$
4. If not converged, go back to 2.

The fixed-point algorithm is based upon the gradient descent optimization technique in order to decrease the convergence time and increase reliability. To satisfy the constraints, the fixed-point algorithm utilizes the projection method. This leads to orthogonal projection of the optimized variable onto the constraint set. The difference between gradient descent and fixed-point numerical algorithms implies convergence rule. The gradient descent algorithm ends if absolute difference between new and old values of optimized variable is less or equal to a predefined value. The fixed-point algorithm stops when the gradient direction at old and new iteration is equivalent up to some small predefined angle. Hence, their dot-product is equal to unit. [9].

There are several advantages of FastICA compared with other ICA methods [23]:

- The convergence of the algorithm is cubic (or at least quadratic)
- Easy to use (no step-size parameters to use)
- Finds independent components of any non-Gaussian distribution
- Algorithm can be optimized by choosing a suitable  $g$
- The independent components can be estimated one by one
- Parallel, distributed, computationally simple, requires little memory space.

### **Overlearning**

Overlearning is an unwanted factor that makes finding the independent components (ICs) impossible. Overlearning in ICA produces spurious estimates of the ICs that have a single spike and are zero everywhere else. To overcome the problem of overlearning many approaches have been introduced. Most of them

rely upon some sort of a threshold. Hyvarinen in [10] experimentally showed that PCA as a method of the dimensionality reduction applied at the preprocessing stage of the ICA, specifically during the whitening of the input data, significantly improves the quality of ICs estimation. In our case overlearning arises in the case of ICA watermark extraction from the spectral images. An extension to the ICA algorithm for spectral images that deals with this problem was proposed in [8]. Input represents mixtures of the image, watermark and key image. To make mixtures and sources more independent high-path filtering is used. It is made based on wavelet transform. HPF is first applied to get a separate matrix and then the separating matrix is used to extract the non-filtered data.

Numbers of inputs is  $m$ , each input is denoted by  $x_i$ , where  $i=1..m$

- The lost of mean and standard deviation values is replaced by scaling each independent component so that minimum is equal to zero and maximum is equal to unit.
- The permutation of independent components is not important in our case.
- In case of negative (positive) contrast each output signal (image) is reproduced as it is and its inverse.

### **ICA algorithm for the spectral images**

- 1) Input: the  $m$ -dimensional mixture vector  $x$ , the number of principal components  $q$ , the number of wavelet decomposition  $l$ .
- 2) Do:
  - a. For
    - i.  $i = 1, \dots, m$
    - ii. Reduce dimensionality of  $x_i$  using PCA to get  $x_{iq}$ .
    - iii. Decompose  $x_{iq}$  using wavelet transform to get  $x_{iqw}$ .
  - b. Apply the FastICA algorithm to the vector  $x_{kw}$  with elements  $x_{iqw}$  to find a separating matrix  $W^T$ .
  - c. Apply a separating matrix to the vector  $z$  which consist of whitened elements  $x_{iq}$  to get  $y = W^T z$
  - d. For
    - i.  $i = 1, \dots, m$

- ii. Reconstruct spectral images using  $y_i$  and eigenvectors obtained by an eigen decomposition of the covariance matrix of  $x_i$
- 3) Reproduce: the spectral image in a suitable color system (e.g. RGB CIE 1931)

### ***Wavelet transform***

Signal and image processing widely uses Fourier theory, according to which any signal can be represented as a sum of sines and cosines, which is often called Fourier expansion. The most serious disadvantage of such a technique is that it has only frequency resolution and no time resolution. As a solution to this problem a number of techniques have been suggested. The wavelet analysis (often called wavelet transform) is one of the most recent and most successful of the attempts of overcoming the main disadvantage of the Fourier analysis. Wavelet analysis uses a fully modulated scalable window, the window is shifted along the signal and a full spectrum is calculated for every window position. Then this process is repeated over and over again with a shorter or longer window for every cycle. Mathematically defined wavelet transform can be defined as a sum over the time of the signal multiplied by scaled, shifted versions of the wavelet function  $\Psi$  [15]:

$$f^w(a, b) = \int_{-\infty}^{\infty} f(t)\Psi(a, b, t)dt \quad (6)$$

where  $f^w$  are wavelet coefficients,  $a$  and  $b$  are scaling and position coefficients,  $f(t)$  input signal as a function of time. By scaling we mean stretching (or compressing) of the wavelet function [15].

A wavelet is generally a waveform of effectively limited duration that has an average value of zero. Wavelets unlike sinusoids have a limited duration and are irregular and asymmetric. Due to this specific character of the wavelets, they are better suited for analyzing sharp changes [15].

Any signal processing performed on a computer using real-world data must be performed on a discrete signal. Unlike the discrete wavelet transform, the continuous wavelet transform can operate at every scale, from that of the original signal up to some maximum scale that is determined by the end-user. Choosing a subset of scales and positions results in the discrete wavelet transform, with the scales chosen based on powers of two – dyadic scales [15].

Thus the wavelet transform of the function  $f(t)$  becomes [12]:

$$f^w(a, b) = |a|^{-\frac{1}{2}} \int f(t) \Psi\left(\frac{t-b}{a}\right) dt \quad (7)$$

where  $\Psi$  is the wavelet function, which is often called a mother wavelet, defined as [12]:

The original function  $f(t)$  can be restored by the inverse transform [12]:

$$f(t) = \frac{1}{C_\Psi} \iint f^w(a, b) \Psi(a, b, t) \frac{1}{a^2} da db \quad (8)$$

where the constant  $C_\Psi$  depends on  $\Psi$  and is defined as follows:

$$C_\Psi = \int \frac{|\hat{\Psi}(\omega)|^2}{\omega} d\omega \quad (9)$$

A number of different mother wavelets exist. In our study we have used a Matlab implementation of a biorthogonal based wavelet.

## 4. Attacks

An attack, as we have previously defined them, is intentional or unintentional modifications in the watermarked image, that lead to partial or even total destruction of the embedded information. One of the main requirements imposed upon modern watermarking schemes is robustness to such intrusions. There are several types of attacks, that consequently to different watermarking algorithms [11].

### **Active attacks:**

As the name implies such kinds of attacks are aimed at deliberately removing or making the watermark undetectable. Such kinds of attacks are most difficult to protect from and they are often appear in copyright protection, fingerprinting or copy control [11].

### **Passive attacks:**

Passive attacks unlike the active ones are not intended for changing the content of the data under study, but just determining whether the watermark is present. This of course is applicable to imperceptible watermarks, and the target of such an attack is covert communications, where even this knowledge is not acceptable [11].

### **Collusion attacks:**

This particular type of attack has the same goal as the active attacks: to remove the watermark, but the method used for it is different. The idea is having multiple copies of the same dataset, containing different watermarks to construct out of all of these a new copy without the watermark. This is a serious problem especially for the fingerprinting applications (e.g. film industry, etc.). However, such a method presents a certain difficulty for the attacker, due to the fact that this attack requires a significant number of copies which are not always possible to receive [11].

### **Forgery attacks:**

In this particular type of attack the attacker embeds a new valid watermark instead of removing one. This allows making modifications in the data and re-embedding the watermark, so that the forged data would seem genuine [11].

These are just four of the most famous previously used types of intentional attacks. All sorts of signal processing operations involved in transmission compression, cropping, filtering, etc of data can be attributed to the attacks, since these modify the images in such a way that the watermark becomes undetectable.

Some particular examples of the methods used for the attacks are given below [11]:

- Geometrical distortions
- Addition of a constant offset to the pixel values
- Addition of Gaussian or non-Gaussian noise
- Low pass or high pass filtering
- Median filtering
- Lossy compression
- Local exchange of pixels
- Quantization and requantization
- Digital to analog and analog to digital conversion
- Watermarking of watermarked image
- Dithering distortion
- Color reduction
- Printing the image and rescanning / photocopying
- Collusion
- Forgery
- Counterfeiting

For testing attack resistance of the proposed technique we will use three types of modifications: low pass filtering (LPF), median filtering and discrete cosine transform (DCT). The reason for selecting these attacks is that they are a part of StirMark - Image Watermarking Robustness Test [3] a widely used standard for testing the robustness of watermarking techniques. The test includes among others

tests on signal enhancement and compression. LPF and median filter belong to signal enhancement techniques and DCT is part of compression.

### ***LPF***

A low-pass filter is a filter that passes low frequencies well, but reduces frequencies higher than the cutoff frequency. The actual amount of attenuation for each frequency varies from filter to filter. The effect of a low-pass filter is similar to that of a moving average, smoothing the signal, thus leaving the long-term trend [15].

In this work we are using a rotationally symmetric Gaussian low-pass filter of size 3x3 and standard deviation 0.5 according to the following formulae [15]:

$$h_g(n_1, n_2) = e^{-\frac{(n_1^2 + n_2^2)}{2\sigma^2}}$$

$$h(n_1, n_2) = \frac{h_g(n_1, n_2)}{\sum_{n_1} \sum_{n_2} h_g} \quad (10)$$

where  $n_1$  and  $n_2$  are input vectors and  $\sigma$  is the standard deviation.

### ***Median Filtering***

Median filter is widely used to remove noise in the images, especially Gaussian salt-n-pepper noise. What this filter does is that it computes a median over a certain window. In our work we are using a 3x3 two-dimensional sliding window computed over every spectral band in an image. The result is a smoothed image [15].

### ***DCT***

Discrete cosine transform (DCT) is closely related to the discrete Fourier transform. DCT is widely used for the purpose of image compression and is a base of the JPEG standard. It is a separable linear transformation; that is, the two-



dimensional transform is equivalent to a one-dimensional DCT performed along a single dimension followed by a one-dimensional DCT in the other dimension. The definition of the two-dimensional DCT for an input image  $A$  and output image  $B$  is as follows [15]:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N}, \quad 0 \leq p \leq M-1, \quad 0 \leq q \leq N-1$$

$$\alpha_p = \begin{cases} 1/\sqrt{M}, & p = 0 \\ \sqrt{2/M}, & 1 \leq p \leq M-1 \end{cases} \quad \alpha_q = \begin{cases} 1/\sqrt{N}, & q = 0 \\ \sqrt{2/N}, & 1 \leq q \leq N-1 \end{cases}$$
(11)

where  $M$  and  $N$  are the row and column sizes of  $A$ , respectively. If you apply the DCT to real data, the result is also real. The DCT tends to concentrate information, making it useful for image compression applications [15].

## 5. Watermarking for Color and Grayscale Images

### *A Method Using ICA for Detection*

A method for digital image watermarking using ICA was proposed in [22]. It was proven to be effective in detection and extraction of image watermarks. The method was applied to grayscale images.

Three basic steps constitute the watermarking system proposed in [22]:

- 1) Design of the watermark data to be embedded into the original image
- 2) Design of the embedding method to compose watermarked data
- 3) Design of the corresponding extraction method

There are two methods of embedding watermark data to the original one, applicable in the current approach. In first method the watermark ( $M$ ) and the key ( $K$ ) data are inserted in the original image ( $I$ ) with small weighting coefficients [22]

$$X = I + aK + bM \quad (12)$$

The second method uses a small filter coefficient and convolution operation and is defined as [22]:

$$X = I + cK + d*M \quad (13)$$

For watermark recovery authors propose PCA whitening for watermark detection followed by FastICA for the watermark extraction. As an input to these algorithms three images are used: original image and two more generated by adding key image and original image [22]:

$$\begin{aligned} X_1 &= X \\ X_2 &= X + eK \\ X_3 &= X + fI \end{aligned} \quad (14)$$

where  $e$  and  $f$  are arbitrary real numbers. These three images are rearranged into three rows in one matrix, which is used for de-watermarking process.

## ***A Scheme Based on DWT and DCT***

A scheme based on DWT and DCT is considered in [13]. The algorithm uses the properties of the HVS to improve the watermarking performance. The embedding scheme of the method combines DCT and DWT. Such an approach allows good unification between robustness and invisibility.

Watermark embedding is done in [13] in the following manner: first the original image is decomposed into an approximate image  $LL_H$  and sub-bands containing small details  $LH_n, HL_n, HH_n$  ( $n=1,2,\dots,N$ ) by  $N$  level discrete wavelet transform (DWT). Then the approximate image  $LL_N$  is transformed by discrete cosine transform (DCT) to receive the DCT coefficient matrix  $F(u,v)$ .

The watermark  $M(i)$  with ( $i=1,2,\dots,I$ ),  $I$  being the length of the watermark sequence chosen to have a normal distribution with zero mean and unit variance. The watermark is embedded into the low-frequency DCT coefficients  $F(u,v)$  using the following formula [13]:

$$F'_{k+i} = F_{k+i} + a \times M(i) \quad i = 1,2,\dots,I \quad (15)$$

where  $F_k$  is the  $K^{\text{th}}$  coefficient after the matrix  $F(u,v)$  is zigzag scanned and  $k$  is the jumping-off point of watermark embedding and  $a$  is the embedding intensity factor. After that thus modulated DCT coefficients are inverse zigzag scanned into the 2D DCT coefficient matrix  $F'(u,v)$ , which then is transformed into the approximate image  $LL'_N$  by 2D inverse DCT. And finally using this approximate image and the detail sub-bands an inverse 2D DWT transform is performed.

Watermark detection is done using ICA, the image that needs to be detected is transformed by  $N$  level DWT and the approximate image  $LL''_N$  received as a result is then transformed by 2D DCT and zigzag scanned into 1D sequence. The DCT coefficients are chosen to correspond to the coefficients used in the embedding stage. In order to justify the watermark sequence a similarity measure of the original  $M$  and the extracted  $M^{\wedge}$  watermarks is used [13]:

$$p = \sum_i M(i) \cdot M'(i) / \text{sqrt}(\sum_i (M'(i))^2) \quad (16)$$

if  $p > T_o$  where  $T_o$  is a threshold set to be five times of the variance so it is normally set to 5.

## 6. Spectral Images

Spectral or as they are also called multispectral images have become an area of particular interest recently. A spectral image can be defined as an image captured at different optical wavelengths [14, 2]. Conventional images, that are used in everyday life are either color or gray-scale (a particular case might be black-and-white images). Where grayscale or black-and-white images can be presented as one-dimensional matrices, where each element of the matrix is a gray level or intensity at a certain point of the image [4]. Color images, e.g. RGB, CIELAB or CIELUV color images, consist of three planes, representing red, green and blue color intensity values respectively for RGB images or luminance and two chrominance color components [6, 4, 25]. Thus the color image is a three dimensional matrix, with the third being the spectral dimension. Spectral images, being a particular class of images are also presented as three dimensional matrices, the only difference is that the third dimension can contain more than 3 components [2] (see Fig.1). The spectral images used in this thesis are sampled in the visible range of wavelength (from 380 to 780 nm), however, an infrared (IR) or ultraviolet (UV) components are also possible to use [5].



**Figure 1.** Cubic representation of spectral image

The growing popularity of spectral images can be explained by several factors. Conventional three-color systems are insufficient to specify the appearance of objects under varying illuminants and viewing conditions. Such phenomena as metamerism, when two colors appear to match (mismatch) under one source of light and mismatch (match) under another [5] are solved using spectral images. Thus, the spectral reflectance of the object, defined as the function of wavelength, is necessary [2]. Spectral reflectance can be measured either evenly or at unequal intervals. Number of bands in spectral images is different, however, experiments show that 7-8 bands accurately represent original reflectance spectra [1, 2].

Areas of application of spectral images vary from remote sensing, environmental monitoring and industrial quality control to Internet museums and on-line stores [14]. The trend is to increase the number of commercial applications.

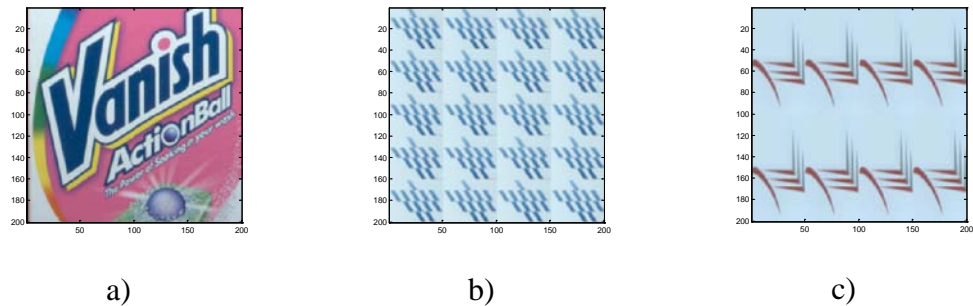
## 7. A Technique for Digital Color Image Watermarking Using ICA

### *Embedding*

For embedding the watermark and key data to the original dataset we use the method proposed in Section 2.2.2. The idea of the method is to create a watermarked image by summing the original, key and watermark images with different coefficients.

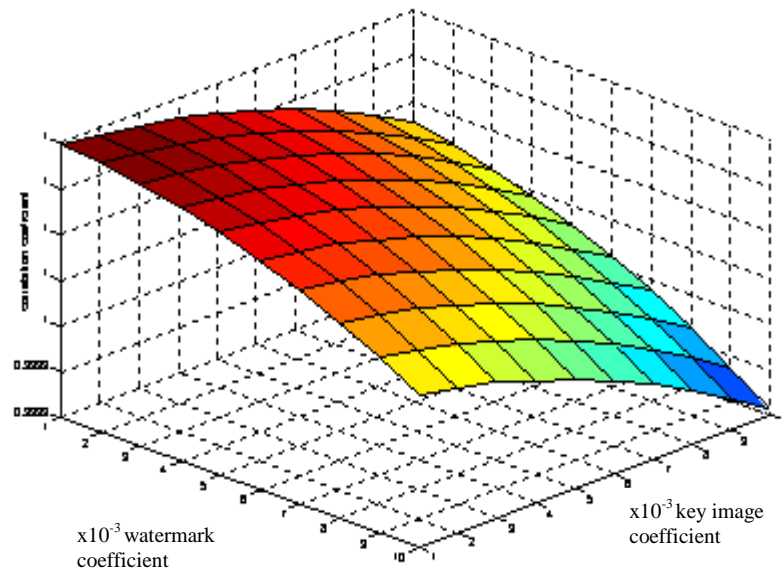
$$X = (1-a-b)I + aK + bM \quad (17)$$

where  $I$  – the original image,  $K$  – key image,  $M$  – watermark image,  $a$  and  $b$  – the mixing coefficients of key and watermark image consequently. We modified the method so that the sum of the coefficients for all three images is equal to 1. Examples of the original ( $I$ ), key ( $K$ ), watermark ( $M$ ) images are presented in Figure 2.



**Figure 2.** Color reproduction of images used in creating watermarked image. a) Original image, b) Key image, c) watermark image.

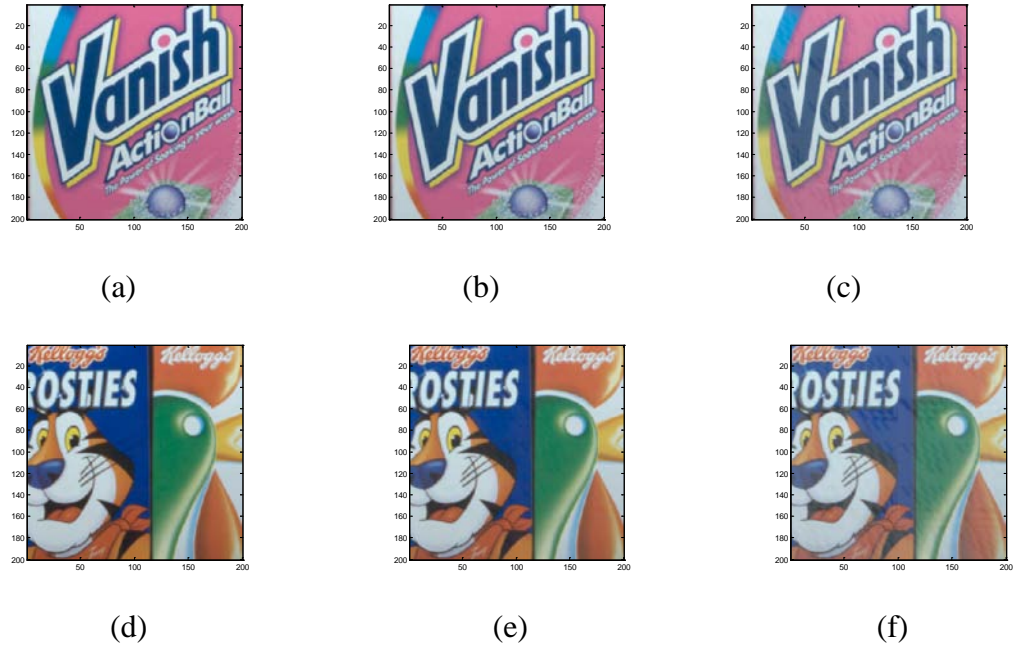
For this method we need to determine coefficients  $a$  and  $b$  both for the watermark and key images. The Figure 3 shows the correlation coefficient distribution between the original and watermarked image with different mixing coefficients ( $a$  and  $b$ ).



**Figure 3..** Correlation coefficient between original and watermarked images.

The range of coefficients was chosen based on visual inspection. The maximal value of coefficients, applicable for the image “vanish”, is 0.01. Figure 4 illustrates the influence of the amount of watermark in the mixture where different images with different coefficients  $a$  and  $b$  are used.





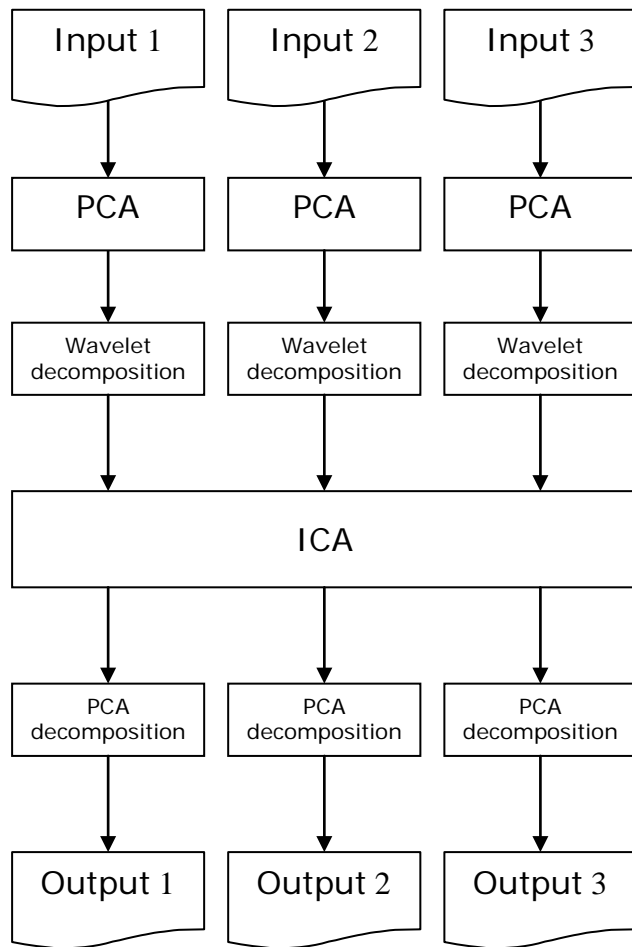
**Figure 4.** Visual inspection results of color reproductions of “vanish” (a - c) and “kellogs” (d - f) images. (a, d) original image, (b, e) watermarked image with coefficients  $a=b=0.01$ , (c, f) watermarked image with coefficients  $a=b=0.5$

Looking at Figure 4, it is possible to say that with the increase of the mixing coefficients  $a$  and  $b$  the watermark becomes more visible in the resultant image and the quality of the appearance degrades.

### ***Extraction***

A next logical step after embedding the watermark is watermark detection and extraction. By extraction we mean a scheme that would allow us to extract the watermark in the original form. Depending on the algorithm this might not be possible, for that reason a procedure that allows detecting if the specific watermark is present in the image is introduced, such a procedure is called watermark detection. The choice of the extraction and detection algorithm depends on embedding procedure.

The watermarking scheme we are using allows us extraction of the embedded watermark. The scheme of the extraction technique proposed in this paper is shown in Figure 5.

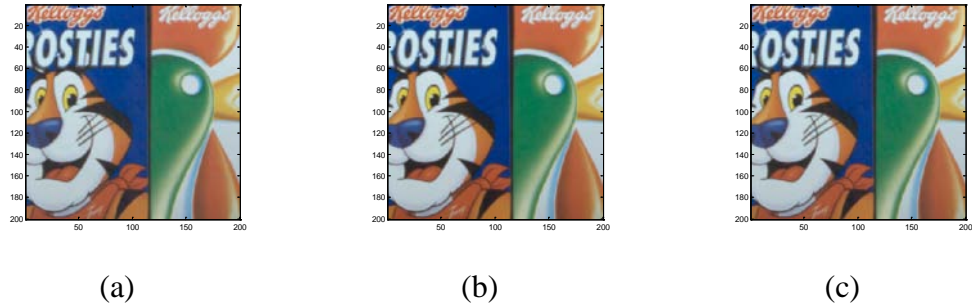


**Figure 5.** Extraction scheme for the watermarking algorithm

The extraction algorithm consists of a number of steps:

- *Inputs:* The algorithm requires a total of three inputs. First one is a watermarked picture available for public access. The rest of the images are additional pictures: two mixtures of the original and key images. The difference between the two is the quantitative proportions of each in the mixture. These two additional images are created similarly to the watermarked image using Eq. 17 with  $a = 0$ . The values of the mixing coefficients ( $b$ ) in this case is determined experimentally. An example of the images input into the algorithm is given in Figure 6.

In the example below, we have chosen as one of the input images a mixture of the original and key images, where the key image constitutes 1% of the total mixture and the original image as the second one.



**Figure 6.** Color reproduction of extraction algorithm input images (*kellogs*) (a) watermarked image with coefficients  $a=b=0.01$ ; (b) original (*kellogs*); (c) additional image with coefficient  $b = 0.01$

The next two steps can be called data preprocessing steps. The purpose of these data modifications is to avoid overlearning in the ICA algorithm, which is one of the most serious ICA problems

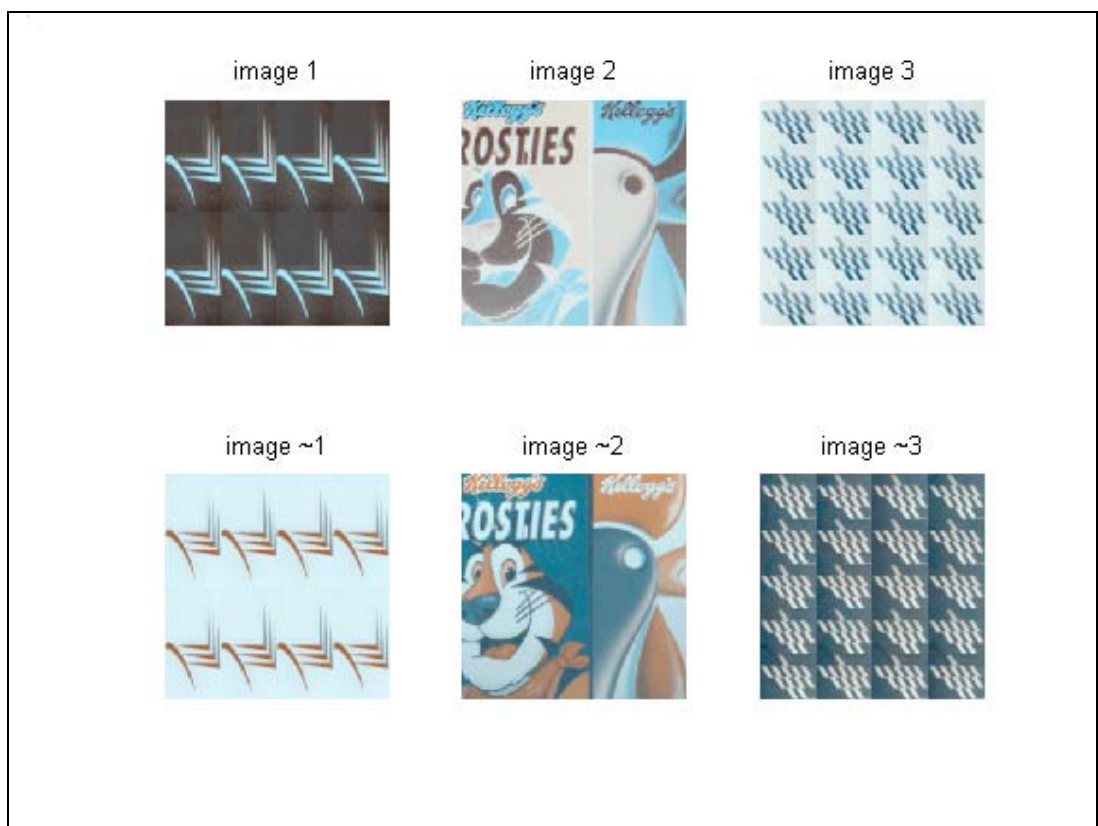
- *PCA*: Principal Component Analysis is used for dimensionality reduction. The principal components of the input images are computed and ones with the highest energy are retained. PCA works in the spectral domain of the images leaving the colors that are dominant in the image.
- *WD*: Another step of data preprocessing which allows reducing correlation by leaving outline data only. After this transformation the basic color separation lines are left in the image.

The outputs of the preprocessing steps are three images with dimensionality reduction performed on them both in the spatial and spectral domains.

- *ICA*: Independent Component Analysis is a basic step of the algorithm, described in background part of the thesis. In this stage incoming preprocessed data is separated into original picture watermark and key images. The output of the ICA algorithm can be both with positive and negative contrast. For that purpose inverse images of the ICA outputs are also obtained.
- *PCA decomposition*: This is inverse step of PCA and used for transformation of a few eigenvectors with a highest eigenvalues into full

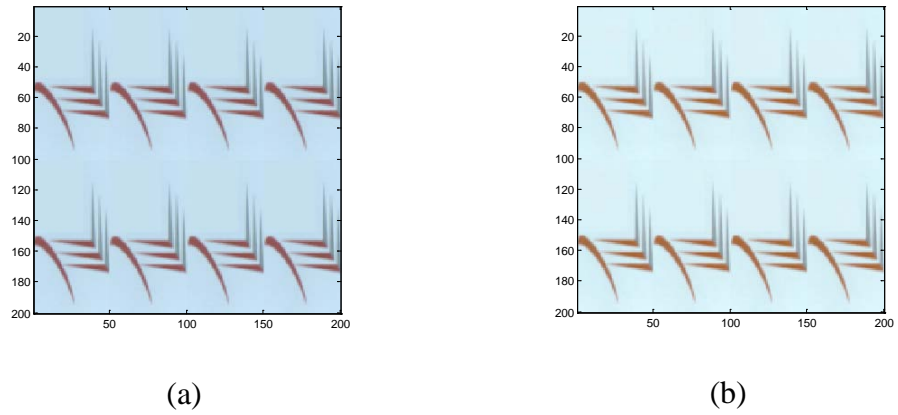
eigenspecter. This step allows us to see images with the same appearance, like they were presented before the preprocessing operations. The information in the spectral domain is restored.

- *Outputs:* The outputs of the whole scheme are original, watermark and key images. An example of the output is given in Figure 7, these images are received after inputting the images presented in Figure 6 into the extraction algorithm.



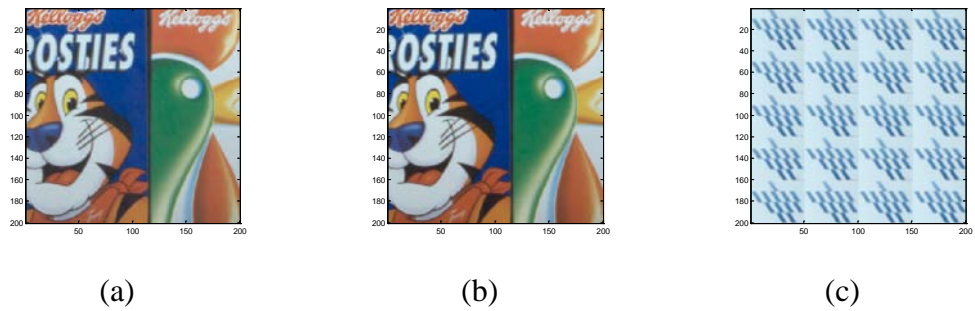
**Figure 7.** Color reproductions of extraction algorithm output images

An important quality factor of the algorithm is the similarity between the extracted watermark and the original watermark images. An example of both of the images is given in Figure 8, visually these images appear similar which means that the extraction algorithm is effective. To test the performance of this algorithm one of the most important parameters of the quality of the extraction – correlation coefficient between extracted and original watermark images, can be computed. In our case the correlation coefficient between these two images is equal to 0.77



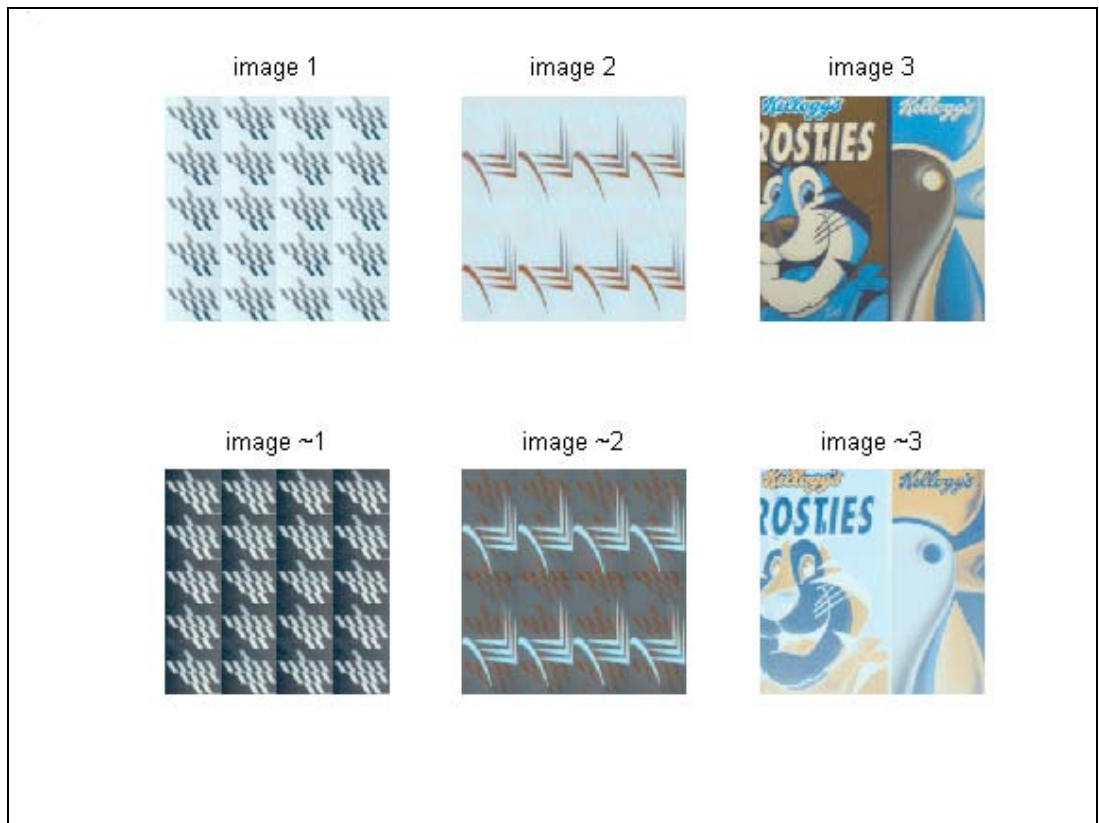
**Figure 8.** Color reproduction of original (a) and extracted (b) watermark images

If we change the third additional image used in the previous example to key image ( $b = 1$ ) and the rest of the inputs leave the same as shown in Figure 9.



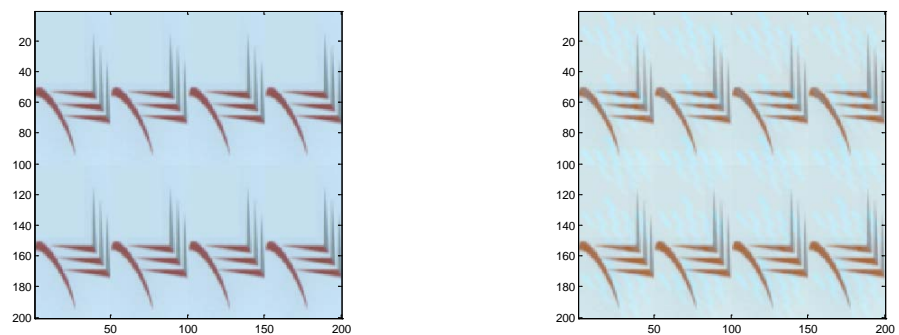
**Figure 9.** Color reproduction of extraction algorithm input images (*kellogs*) (a) watermarked image with coefficients  $a=b=0.01$ ; (b) original (*kellogs*); (c) key image

The output of the extraction algorithm in such a case is presented in Figure 10.



**Figure 10.** Color reproductions of extraction algorithm output images

Again the original and extracted watermarks have to be compared. Both are presented in Figure 11. The quality of the extraction appears to be slightly worse, moreover, the correlation coefficient of the two is equal to 0,75.



**Figure 11.** Color reproduction of original (a) and extracted (b) watermark images

The algorithms both embedding and extraction are tested in experimental part of this work



## 8. Experiment

In the experiment we are using a database of multi-spectral images which have been acquired using an Applied Spectral Imaging ©Spectracube camera. This device is an interferometry based, semi-portable digital camera which is able to capture in a single exposure a full 2D spatial array of spectra (i.e. a spectral image). The database we use consists of a set of high spatial and spectral resolution reflectance images of every day objects. The MATLAB<sup>TM</sup> .mat files contain full spectral resolution reflectance data from 400nm to 700nm at 10nm steps (31 samples). The image matrix for each object is HEIGHTxWIDTHx31. The images have been captured in a VeriVide viewing booth with a black cloth background under CIE illuminant D75. Each image has been captured twice: once with a white tile and once without. The illuminant has been estimated from the white tile and the spectral data divided by this estimate, in order to arrive at reflectance measurements [20]. The images used in this thesis are presented in Appendix 1. For the purpose of the experiments a total of 5 images were selected. An area of 200x200 pixels was cut out of each of the images, so that the most informative and colorful area would be present in the fragment. As the watermark and key images spectral images of LUT and Joensuu University were selected (see Appendix 1).

The experimental part of this thesis is dedicated to testing different parts of the algorithm proposed in this work (see Section 7). The overall aim is to prove that the algorithm is capable of watermark embedding and extraction. Another problem to be experimented upon is the influence of the embedding coefficients  $a$  and  $b$  (see Eq. 17) on the extraction quality. And the final part is dedicated to testing various types of attacks on the algorithm.

The overall plan of the experiments can be given as follows:

1. Extraction possibility with different embedding coefficients;
2. Extraction quality with different mixtures (additional images);
  - a. Dynamic watermarked image and constant additional images with different content of key image in them;



- b. Dynamic watermarked image and dynamic additional images;
3. Extraction quality with attacks influence.

By dynamic we mean that the content of the input image is changing during each iteration depending on current value of  $a$  and  $b$ . While by constant we mean that input image is created once at the beginning of the experiment.

For the purpose of the experiments five pictures were taken. Since the content of the tested image does not make an influence on the results of experiments, displaying of results for one image was decided. The results for image "kellogs" were chosen for display in the experimental part of the thesis.

The main goal of the first set of experiments is the possibility of a choice of the mixing coefficients for watermark and key image embedding into the original image, so that the resulting image would allow us the best watermark extraction, meaning that the extracted watermark image would be of the best quality possible.

Difficulty of such an experiment is that the two coefficients we are looking for are not the only ones influencing the quality of the watermark extracted. Two additional images, as it was shown in Chapter 7 influence the result of the algorithm.

A watermarked image ( $X$ , see Eq. 17) is formed by mixing the key image ( $K$ , see Eq. 17) and the watermark ( $M$ , see Eq. 17) into the original image. For that purpose a combination of mixing coefficients ( $a$  and  $b$ , see Eq. 17) in the interval 0.0001 to 0.005, with a step of 0.0005 is taken, resulting in a total of 100 different combinations. To assess the quality of the watermark extracted correlation coefficient (CC) was used, computing correlation between the original watermark and the one obtained on the output of the algorithm.

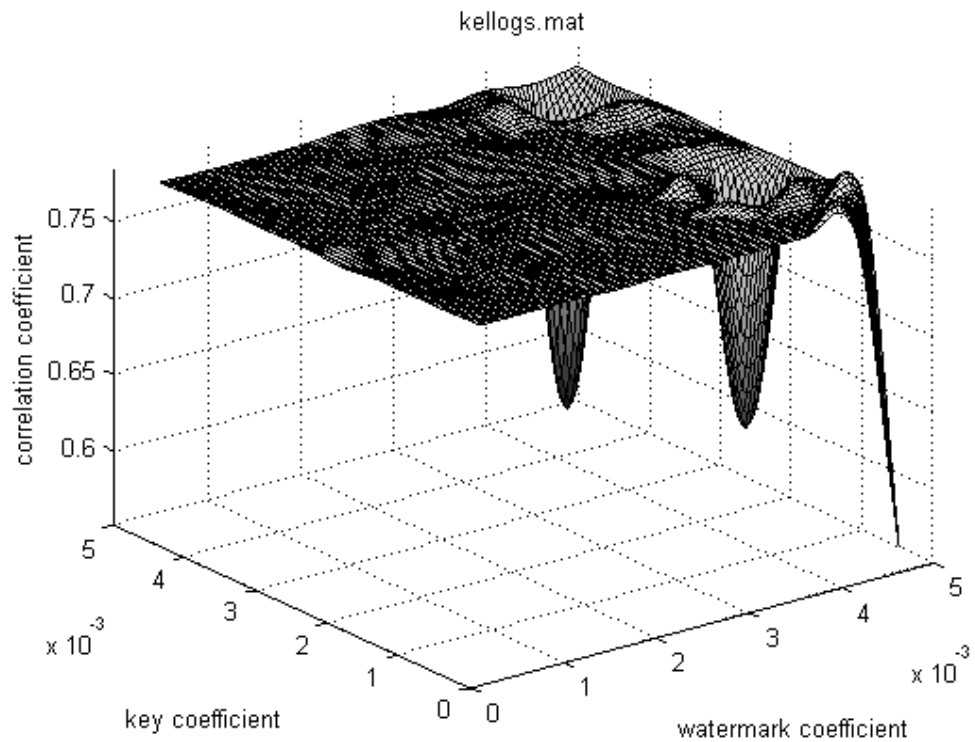
Image  $I$  was chosen as one of the additional images, while the second image was formed via mixing of  $I$  and  $K$  before the experiment.

As the input of the algorithm in the first experiment three images were used:  $X$  formed during the experiment,  $I$  and an additional image with  $K$  mixed in, given that the portion of  $K$  was equal to 0.1% of the total image. The way the inputs in the experiment are formed is given in Table 1.

**Table 1. Input values for the first experiment**

Input 1	$(1-a-b)I+aK+bM$
Input 2	$I$
Input 3	$0.999I+0.001K$

As the result a surface of the CC distribution, shown in Figure 12, was obtained.



**Figure 12.** Correlation coefficient between original and extracted watermark

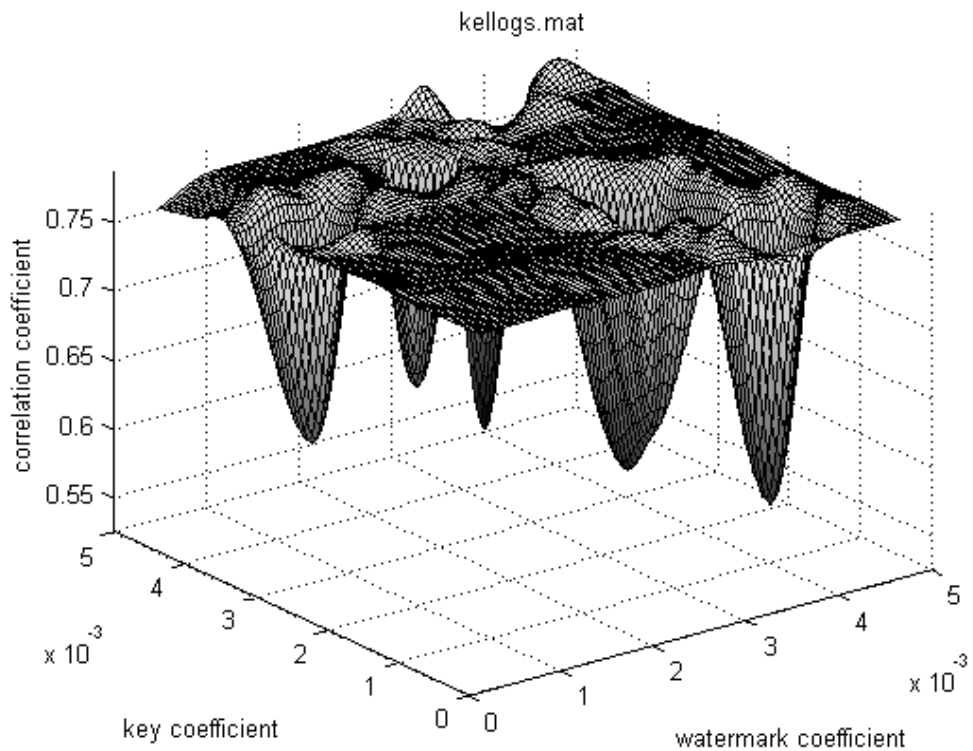
In Figure 12 X and Y axes represent the values of key and watermark coefficients  $a$  and  $b$  while Z – correlation coefficient. The surface shown is smooth, except for several bumps that can be attributed to the fact that ICA is an unstable algorithm that sometimes produces undesirable results.

During the second experiment the percentage of  $K$  mixed in totaled 10%, whilst the rest of the conditions of the experiment remained the same. The experimental setup of this experiment is given in Table 2.

**Table 2. Input values for the second experiment**

Input 1	$(1-a-b)I+aK+bM$
Input 2	$I$
Input 3	$0.9I+0.1K$

The result obtained the experiment is shown in Figure 13.



**Figure 13.** Correlation coefficient between original and extracted watermark

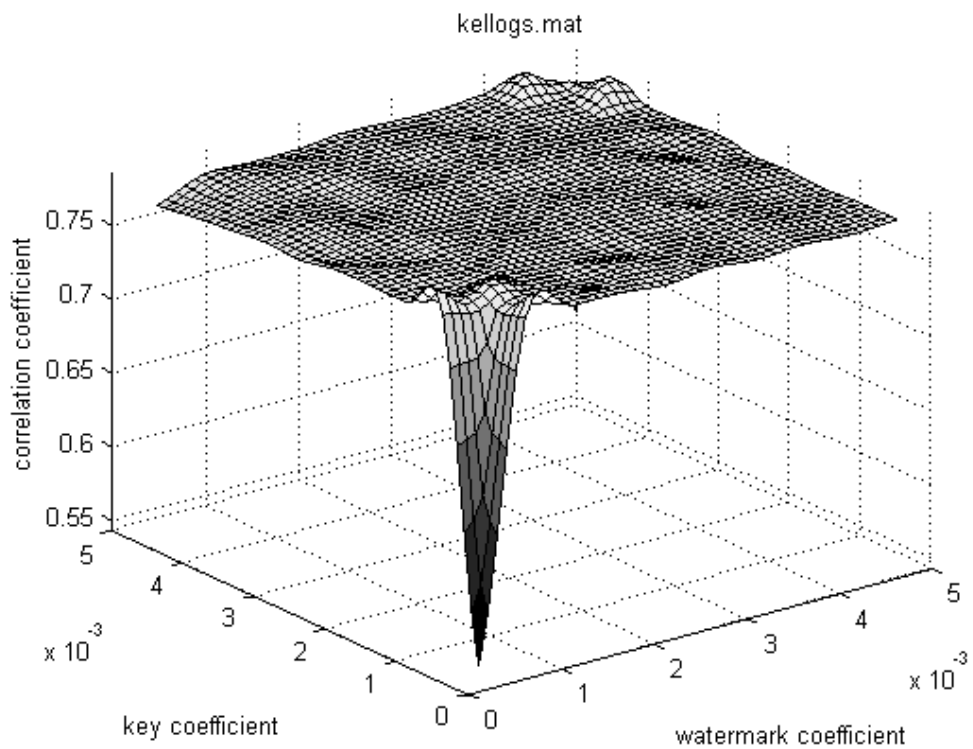
The axes in Figure 13 are similar to those in Figure 12. The surface shown is now more uneven, but the changes in the CC have a randomized character, based on what it can be stated that there is no or a very weak connection between the  $a$  and  $b$  change and the CC.

On the next stage of the first experiment the key image mixed into the original image constituted 10% percent, while the other signals remained the same. The setup of this experiment are shown, in turn, in Table 3.

**Table 3. Input values for the third experiment**

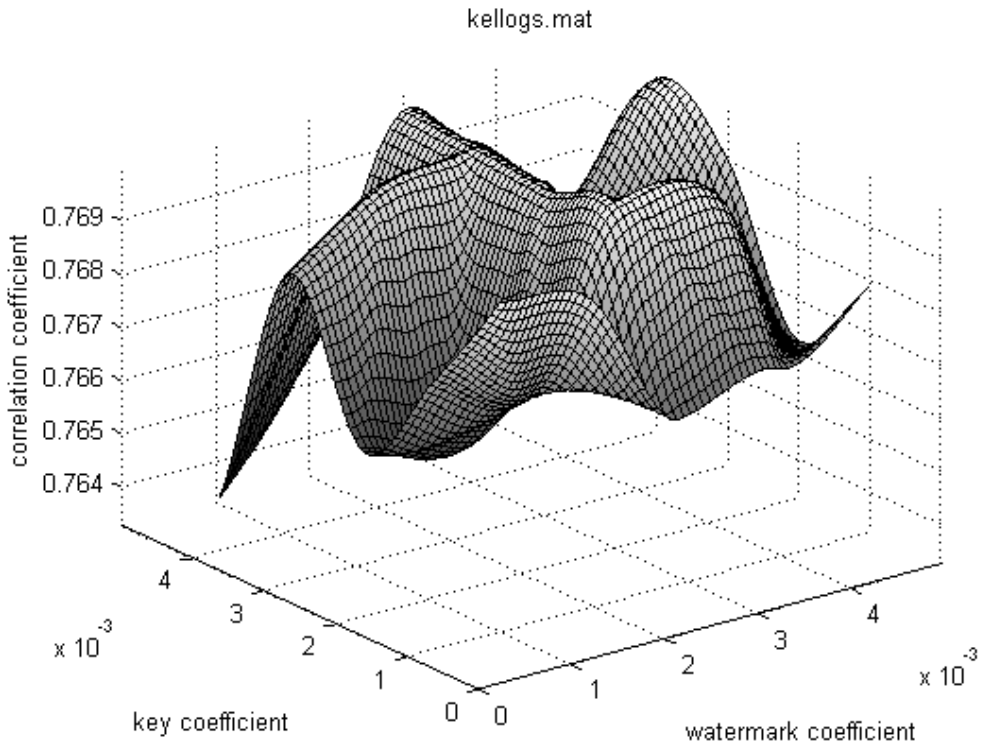
Input 1	$(1-a-b)I+aK+bM$
Input 2	$I$
Input 3	$0.1I+0.9K$

The correlation coefficients received as the result are plotted in the Figure 14.



**Figure 14.** Correlation coefficient between original and extracted watermark

The dynamics of the correlation coefficient subject to key image and watermark coefficients decrease in the experiments described above are given in Figure 15.



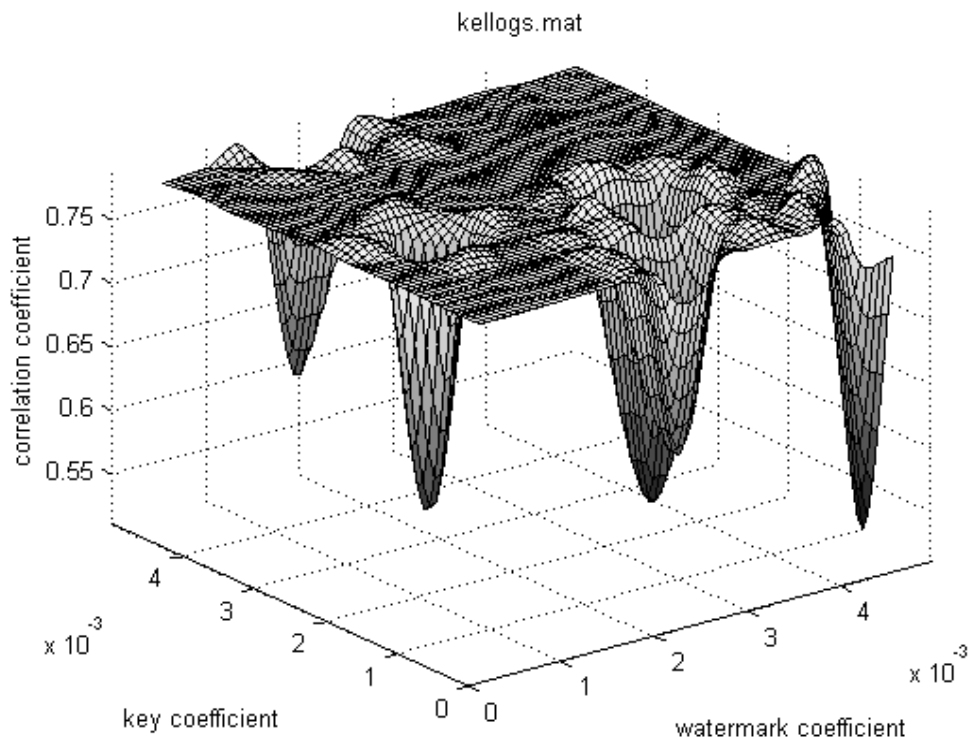
**Figure 15.** Correlation coefficient between original and extracted watermark

In the experiments considered before mixture of the images obtained before the start of the experiment were used as additional images. An additional image is generated along with the watermarked image in the following example, where the quantitative constituent of the additional image is equal to the quantitative constituent of the key image in the watermarked image. The second additional image was created in such a way that the key image would have a coefficient  $b$  of 0.0001.

**Table 4.** Input values for the forth experiment

Input 1	$(1-a-b)I+aK+bM$
Input 2	$0.9999I+0.0001K$
Input 3	$(1-a)I+aK$

Figure 16 illustrates the dependence of the correlation coefficient and the coefficients of the key image in the shared image and the additional image, and also of the quantitative constituent of the watermark in the watermarked image.



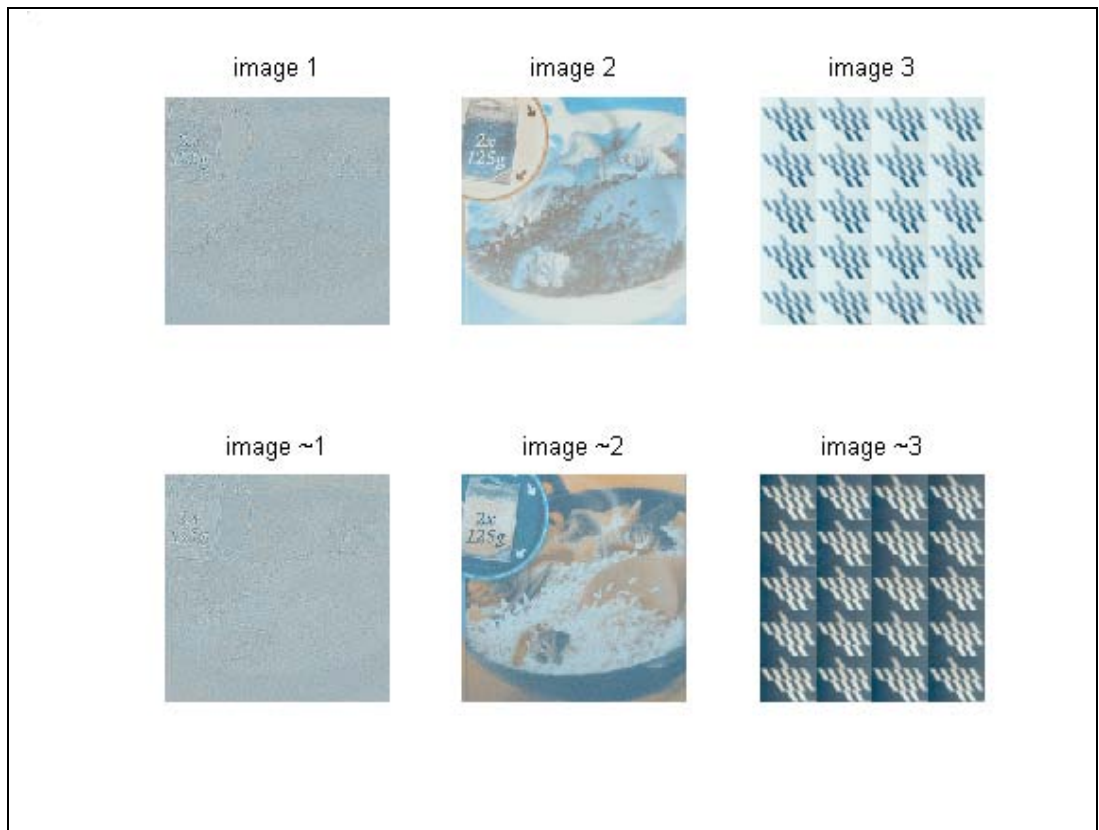
**Figure 16.** Correlation coefficient between original and extracted watermark

Again the figure shows that there is a weak or no dependence of the CC on the  $a$  and  $b$ .

### ***Attack influence***

Final stage of the experiments involved testing of the robustness of the algorithm. For that purpose a part of the StirMark test [3] was used. Three types of modifications have been chosen for that purpose (see Section 4): low pass filtering (LPF), median filtering and discrete cosine transform (DCT).

All of the techniques were applied to watermarked images and input into the extraction algorithm proposed. An example of the output is shown in Figure 17.



**Figure 17.** Color reproduction of the output of the extraction algorithm applied to a watermarked image after an attack

Looking at Figure 17 it is possible to say that the watermark could not be extracted after applying the attack prior to the extraction. In this case a median filter was used as the attack method. Similar results were obtained using DCT and LPF applied to all 5 images out of the image database. Judging by these results it is possible to say that the algorithm proposed in this paper is fragile, which means that the watermark embedded in the image is altered even by simple image processing operations performed on the watermarked image.

## 9. Conclusions

In this work digital image watermarking is considered. A novel method of digital watermarking of color and spectral images using a color or spectral watermark is introduced in this paper. The method is based on previously developed methods created for gray-scale images. The paper starts with an overview of existing methods watermarking of color and grayscale images, in particular Methods using independent component analysis (ICA) for detection and the ones using discrete wavelet transform (DWT) and discrete cosine transform (DCT).

A novel method of watermarking proposed in this work allows embedding of a color or spectral watermark image, key image into color or spectral image consequently and a successful extraction of the watermark out of the resultant watermarked image. The idea is to form the resultant image via summation of the original, watermark and key images with different weighting coefficients, so that the overall sum of weights does not increase one.

A number of experiments have been performed on the quality of extraction depending on the parameters of the embedding procedure, i.e. weighting coefficients. We have found that mixing coefficients of the key and watermark image should not exceed 1% (0.01) of the original otherwise the watermark becomes visible. Moreover, mixing coefficients change has a very weak influence on the quality of the watermark extraction, so that even very low values of these allows successful watermark extraction. Furthermore, we have found that the contents of the image to be watermarked do not affect the quality of the extraction, except for the case when all three images (original, key and watermark) have the same content.

Another set of experiments included the test of the robustness of the algorithm proposed. Three techniques have been chosen for that purpose: median filter, low-pass filter (LPF) and discrete cosine transform (DCT), which are a part of a widely known StirMark - Image Watermarking Robustness Test. The study shows that the proposed watermarking technique is fragile, i.e. watermark is altered by



simple image processing operations. Creating a more robust algorithm could be one the topics for further investigation.

To conclude it is possible to say that the algorithm of color and spectral image watermarking, proposed in this paper, has proven to be successful in the task of watermark embedding and extraction.

## References

1. Berns R.S., Imai F.H., Burns P.D., Di-Y. Tzeng: Multi-spectral-based Color Reproduction Research at the Munsell Color Science Laboratory. 14-25, Electronic Imaging: Processing, Printing, and Publishing in Color, 18-20 May 1998, Zurich, Switzerland, part of Europt. SPIE Proceedings, Volume 3409, 1998, ISBN 0-8194-2861-2.
2. Bovik A., Handbook of Image and Video Processing, Academic Press, 2000
3. Fabien A.P., Petitcolas R.J., Anderson M., Kuhn G., Attacks on copyright marking systems, in David Aucsmith (Ed.): Information Hiding, Second International Workshop, IH'98, Portland, Oregon, USA, April 15-17, 1998, Proceedings, LNCS 1525, Springer-Verlag, pp. 219-239.
4. Gonzalez R.C., Woods R.E., Digital Image Processing. Prentice Hal, Inc., 2002.
5. Hauta-Kasari M., Computational Techniques for Spectral Image Analysis. Ph.D. thesis, Lappeenranta University of Technology, 1999.
6. Hunt R., Color Research and Application, 19 (1994).
7. Hyvarinen A., and Oja E., Independent component analysis: Algorithms and Applications. Neural Networks, Vol. 13(4-5), pp. 411-430, (2000)
8. Hyvarinen A., Karhunen J., and Oja E., "Independent Component Analysis". Wiley Interscience, 2001. 504 pages. Japanese translation by Tokyo Denki University Press, 2005.
9. Hyvarinen A., Karhunen J., Oja E., Independent Component Analysis, John Wiley, final version of 7 March 2001.
- 10 Hyvarinen A., Sarela J., Vigarior R., Spikes and bumps: Artefacts generated by independent component analysis with insufficient sample size. In proc. Int. Workshop on independent Component Analysis and Signal Separation (ICA 99), pp. 425-429, Aussois, France, 1999.
- 11 Kankanhalli M., Ramakrishnan K., Content Based Watermarking of Image, ACM Multimedia (1998), pp. 61 - 70.
- 12 Krestyannikov E., "Wavelet Filter Parameterization for Compression of Spectral Images" Master's thesis, Lappeenranta University of Technology 2003
- 13 Liu J., Zhang X., Sun J., Lagunas M., A Digital Watermarking Scheme

Based on ICA Detection, in Proc. ICA 2003, pp.215 - 220.

- 14 MacDonald L.W., Luo M.R., edited: Color Image Science (John Wiley & Sons, 2002).
- 15 Matlab website <http://www.mathworks.com> (cited 18.07.06)
- 16 Memon N., Wong P., Protecting Digital Media Content, Communications of the ACM, vol. 41, No. 7, (1998), pp. 35 - 43.
- 17 Mintzer F., Braudaway G., Bell A., Opportunities for Watermarking Standards, Communications of the ACM, vol. 41, No. 7, (1998), pp. 57 - 64.
- 18 Mintzer F., Braudaway G., Yeung M., Effective and ineffective image watermarks, in proceedings of the IEEE ICIP 97 (Santa Barbara, California, October 26-29), 1997, pp. 9-12.
- 19 Mintzer F., Lotsipiech J., Morimoto N., Safeguarding digital library contents and users. DLIB Mag (Dec, 1997)
- 20 Multispectral Image Database  
<http://www2.cmp.uea.ac.uk/~pm/chromagenic/msdb2.html> (cited 16.07.06)
- 21 Rabbani M., Invisible Digital Image Watermarking, PICS 2003, Tutorial, (2003)
- 22 Shen M., Zhang X., Sun L., Beadle P., Chan F., A Method for Digital Image Watermarking Using ICA, in Proc. ICA 2003, pp. 209 - 214.
- 23 Smith L. A Tutorial on Principle Component Analysis (2002)
- 24 Wikipedia, the free encyclopedia  
[http://en.wikipedia.org/wiki/Principal\\_components\\_analysis](http://en.wikipedia.org/wiki/Principal_components_analysis) (cited 16.07.06)
- 25 Wyszecki G., Stiles W.S., Color science: concepts and methods, quantitative data and formulae, 2nd edition, Wiley, New York, 1982.

# Appendix 1

## Spectral Image Database



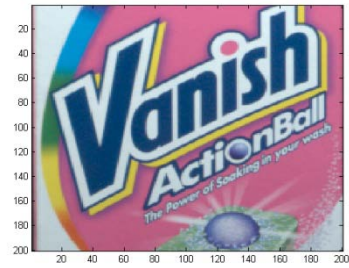
(a1)



(a2)



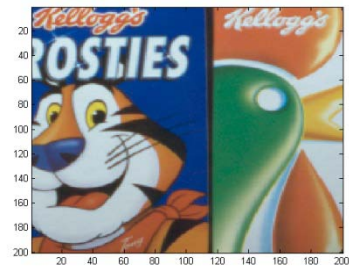
(b1)



(b2)



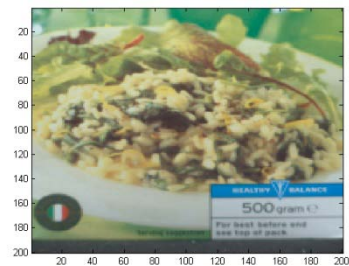
(c1)



(c2)



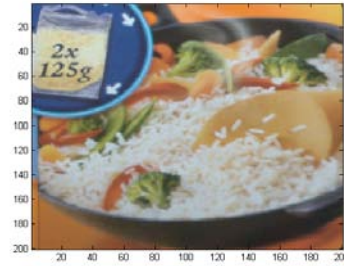
(d1)



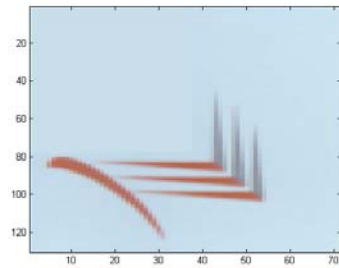
(d2)



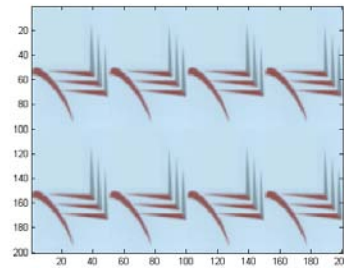
(e1)



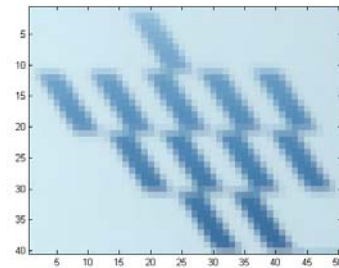
(e2)



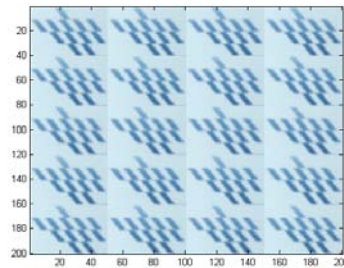
(f1)



(f2)



(g1)



(g2)

Figure 1. Set of color representations of spectral images used in the experiments: (a) sensodyne; (b) risotto; (c) vanish; (d) unclbens; (e) kellogs; [20] (f) lutlogo, (g) ujlogo.