Connection between the need for compression, bandwidth and quality for multimedia transmissions in UMTS

By

Helge Ivar Lislevand
Håkon Oftedal Njaa

Masters Thesis in
Information and Communication Technology

Agder University College
Faculty of Engineering and Science

Grimstad, May 2004

Abstract

The 1990s made mobile communication common among the majority in many countries, with user penetration as high as in the 80th percentile. The mobile telecommunication services has evolved from the analog services provided in the 1980s, to the current digital systems with limited packet data services, and further to the 3rd generation (3G) systems that are currently being rolled out across the world. These systems have been developed with extended data rate support in mind. They are specified to support transport of various data content, with different needs for bandwidth and delay constraints. Amongst these are voice-transportation, web browsing, video-streaming and video-telephony. The bandwidth limitations will of course remain, although with greater bandwidth one will be able to perform tasks that are difficult with today's systems, like streaming video and video conferencing.

The need for compression will in the future be as high as ever. The increase in bandwidth usually leads to new fields of use i.e. new applications with higher demands for bandwidth. Examples of this are new phones with built in video cameras, and phones with high quality digital cameras. There has already been introduced a mobile phone with a built in 2 megapixel camera, and there is no reason to think that the development in this field will stop very soon. Online gaming on mobile devices is no longer science fiction; it will soon be a reality all over the world. These examples show that the need for efficient, high quality compression methods and algorithms are, and will remain high, in the years to come. The introduction of wavelet compression and the use of wavelets in the Joint Photographic Experts Group 2000 (JPEG 2000) standard is one example of new more efficient compression with high quality.

The UMTS standard is supposed to give the user data rates of up to 2Mbps. This is not a very likely transfer rate for most users. The transfer rates will most likely range from 12,2-384Kbps. Interference, movement and fading makes it very unlikely for a user to achieve higher transfer rates.

This thesis looks at image and video compression needs in Universal Mobile Telecommunication System (UMTS) in relation to bandwidth limitations and quality. We try to define what parameters that is most important for the quality of compressed pictures, and present what quality measurement techniques that are most widely used. We also present a quality measurement test based on wavelet-compressed images, where we show the perceptual difference of when a viewer is presented with an image with and without a reference. The subjective quality test is compared to objective measurements, and a conclusion is then presented.

The lack of suitable network simulation software has made it difficult, or to be precise impossible to present a fully working test system. We could have implemented a hardware based simulator, but decided that a task like that might be too time consuming. This being a fact we have made more effort in investigation of the compression and quality part of the thesis description. The focus of the thesis is at image, and image quality. There is not a lot of focus on video and video transmissions. Many of the parameters we discuss in the thesis will though be highly relevant when talking about transmissions of video.

Preface

This report is a result of a master thesis project in Information and Communication Technology at Agder University College (AUC). The work has been carried out between January and May 2004. The thesis has been an internal project for AUC, and no external parties have been involved in the thesis description or as supervisors.

We would like to thank our supervisor Assistant Professor Ola Torkild Aas at Agder University College for all advice and helpful counseling during the thesis period. We would also like to thank Associate Professor Per Henrik Hogstad for letting us attend his lectures on wavelets.

Grimstad May 2004

____________________ ____________________

Helge Ivar Lislevand Håkon Oftedal Njaa

Table of contents

List of figures

List of tables

1 Introduction

1.1 Background

Data compression is extensively used in current computer networks. This is due to bandwidth limitations and the need for effective transportation. Data is also compressed to minimize storage size. All kinds of files are being compressed before transmission or storage, but primarily large files, like images, sound or video.

The 3G mobile telephony systems are more adapted for multimedia files, compared to the current 2.5G systems, with increased functionality in the equipment and extended bandwidth.

The bandwidth is still limited, especially for bandwidth demanding applications like video-streaming and video-telephony.

Video-telephony has been one of the major selling points for the telephony industry, and video compression standards have been made to adapt to the bandwidth limitations of telephony systems. The use of wavelets is the most resent attempt to get satisfactory fidelity in video-telephony.

We have looked into the use of wavelet transformation as compression method of images in reference to bandwidth limitations and quality. We wanted to find out if wavelet transformation could be used to achieve a satisfactory quality and still meet the bandwidth limitations.

Bandwidth limitation is of course not a big problem when transferring single pictures, since time is not an issue. The time used for downloading an single picture is limited only by the patience of the user. There might be a problem with real time applications, like video streaming or video-telephony that have strict delay and jitter constraints. The quality of the service will be perceived as low if one gets high jitter or delay.

1.2 Image Compression

There are several standards for storing and transferring images, both compressed and uncompressed. The information capabilities for each file format differ. Common values for bit depth range from 1 to 64 bits per pixel. Some only store 1-bit black and white images; others store grayscale or color images. The different color image file formats also have different capabilities. Some show only 8-bit color dept, giving a total of 256 different colors, while others use up to 16 bits per color channel, and some again have additional information channels, like alpha planes and transparency planes.

Many images are transferred over the Internet, and the most used image file formats are JPEG and Graphics Interchange Format (GIF). The JPEG file format has been developed by International Organization for Standardization (ISO) and International Telecommunication Union Telecommunication Standardization Sector (ITU-T), and is based on Discrete Cosine Transform (DCT) and Huffman coding. CompuServe Incorporated developed the GIF standard. It uses LZW algorithm for compression.

We have focused on the JPEG 2000 standard, which is a relatively new standard that was developed by the same organizations as JPEG. It uses wavelets and arithmetic coding to compress images. It gives higher quality for the same file size compared to the older JPEG standard. JPEG 2000 also has some extended capabilities. It gives the possibility for downloading images with increasing resolution or quality. Another function is Region Of Interest (ROI). One can define a ROI, and this region will download first with higher quality than the rest of the image.

The software we have been using to compress pictures for testing purposes is a program from LuraTech [9] called LuraWave. It is a small user-friendly program, and it is freeware.

1.3 UMTS

The UMTS standard springs from the Future Land Mobile Telephony System (FLMPS). It was meant to be a global standard for multimedia communications. This evolved to the International Mobile Telecommunications-2000 (IMT-2000) under which several proposed standards were evaluated, but no single standard where agreed upon. This led to a split into two standards and two sub-comities for the IMT-2000 organization, namely the 3rd Generation Partnership project (3GPP) for UMTS and 3rd Generation Partnership Project 2 (3GPP2) for the cdma2000 standard. The UMTS were supported by The European Telecommunication Standards Institute (ETSI) and the Japanese Association of Radio Industries and Businesses (ARIB), while the cdma2000 where mostly an American initiative.

UMTS is the type of 3G systems that will be introduced in Norway, and has already been introduced in some European countries. UMTS gives increased bandwidth and support for multimedia applications compared to the current second generation (2G) and extended 2G (2.5G) systems in operation

UMTS is first and foremost a new radio-interface between the User Equipment (UE) and the base station that is called Node-B in UMTS. It uses a Wideband Code Division Multiple Access (WCDMA) system. This spread spectrum technique spreads the signal over a much larger frequency spectrum than the minimum limit by using a spreading code. Spread spectrum technique was developed to withstand jamming, i.e. deliberate attempts to block communication, and therefore gives a better protection against natural influences, like noise and shadowing. The transmission is spread over a frequency interval as opposed to sending at a single frequency, and this makes it more difficult to eaves drop, since the transmission appears to be noise if one cannot produce the right descrambling codes.

Depending on what services are used, circuit switched or packet switched transportation will be used. For circuit switched communication, the maximum bandwidth will be 64 kbps and for packet switched communications, the maximum will be 384 kbps while moving. UMTS will give up to 2 Mbps in pico-cell environments and close to the B-Node in larger cells. Variations in the transmission channel, such as fading, interference and movement gives limits for the high bandwidth. The low transmit power of the UE also gives limits for how far from the base station one can achieve high bandwidth.

1.4 Quality

Quality measurement is quite a science in itself, and there are loads of techniques described in books and on the internet, that claims to be the best way of conducting a correct quality measurement.

The Human Visual System (HVS) is a very complex subject. The HVS is still not fully explored or understood, and it is a subject witch is worked on quite extensively. We have included a subchapter in this thesis about the HVS. The chapter gives a short introduction on the subject, and is by far a complete lesson.

In this thesis, we have used both objective and subjective quality measurements. The subjective measurements are carried out using a web interface, where the user rates the quality of several compressed pictures according to a five rate scaling system of the degree of impairment. The scaling system is defined by International Telecommunications Union – Recommendations (ITU-R) and is numbered ITU-R 500-3. When the user had gone through the first set of pictures, he is redirected to a new page where he is to grade the same pictures one more time. This time though, he gets an original picture to which he can compare the compressed picture. The results are written to a text file for later use.

The objective quality measurements are done by LuraWave, which in addition to compress pictures to JPEG 2000 lets you compare a compressed image with its original and computes several quality parameters.

We try to do a comparison of the two measurement methods and their results in our discussion chapter. This is an interesting part of the thesis since the subjective and objective results do not always match.

1.5 The connection between bandwidth, quality and compression

To say something about the connection between bandwidth, quality and compression, one has to look at the whole chain from input of a picture in one end, to the reconstruction of the compressed picture in the other end. In this thesis we have looked into different compression algorithms and their use. Our main focus has though been on wavelet compression and JPEG 2000. We have also looked at the UMTS channel and its capacity, bandwidth limitations and coding schemes. In our "Image quality" chapter we have described different quality measurement methods, and different impairments that may occur in a compressed picture. The chapter also includes an introduction to the Human Visual System. In our conclusion chapter, we have tried to reach some conclusions regarding the above mentioned connection.

1.6 Methodology

Most of the work on this thesis is based on reports and standardization papers found on the internet, and with the standardization organizations. We have tried to extract the essence of previously published papers and standards, as a basis for this thesis. Our own work is hopefully useful as a supplementary reading to previously published material.

2 Image Compression

2.1 General

Image and video data compression refers to a process in which the amount of data used to represent image and video is reduced to meet a bit rate requirement. The amount of data should be below or at equal to the maximum available bit rate. To compress data we use an encoder, and to decompress data, we use a decoder. Ideally, an image compression technique removes redundant and/or irrelevant information, and efficiently encodes what remains. Practically, it is often necessary to throw away both non-redundant and relevant information to achieve the required compression. In either case, the trick is finding methods that allow important information to be efficiently extracted and represented. [2] Symmetric compression means that the encoder and the decoder uses the same fundamental algorithm, but opposite way. Asymmetric compression means that the encoder and decoder can work with different algorithms. Most compression methods are physical. They only consider the incoming bits. The algorithms do not consider the meaning of the received data bits.

The reason one needs to compress image data is the limited bandwidth one has to his disposal. The bandwidth in the UMTS system ranges from 12,2 to 384 kbps. The specification operates with speeds up to 2Mbps. This is not a realistic transfer rate for most users. When a user is in the immediate vicinity of a Node-B such transfer rates is possible, but for normal use i.e. when a user is moving, at the best a user should be able to get a transfer rate of 384 Kbps.

2.2 Compression techniques

Historically there are two main categories for witch compression has been used. These are compression for storage and compression before transmission. The latter is our concern in this thesis. We will be focusing on the need for compression in mobile communication systems.

When sending an uncompressed picture from one cellular phone to another, the amount of data that we want to send can be very large. The bandwidth we have to our disposal is though limited. Indeed when a video transmission is what we want, the need for compression can be more than 500:1.

2.2.1 Lossless compression

When dealing with lossless compression, the image after compression and decompression should be identical to the original, and only the statistical redundancy is exploited to achieve compression. Lossless compression can though only achieve a modest amount of compression. To achieve lossless compression one utilizes the possibility to encode the data more efficiently. Examples of compression algorithms are Huffman and Lempel-Ziv-Welch coding.

2.2.2 Huffman coding

This subchapter contains an overview of the properties of the Huffman coding, and an example of use of Huffman coding

Properties:
Variable length coding is highly effective on symbols with high frequency.
The code must be unambiguous i.e. no code word is the start of another word. Also called prefix property. This property results in that words can be sent continuously, without spacing.

Example of Huffman coding:
Suppose you have six characters or symbols that you want to encode. Each of the symbols appears with a certain frequency.

Let us say these are the frequencies:

Table 1 Symbol frequencies

Symbol	A	B	C	D	E	F
Frequency	0.20	0.09	0.15	0.11	0.40	0.05

You first group together the two symbols with the lowest frequency, in this case the symbols B and F. These two symbols make a new symbol BF with a combined frequency of B and F.

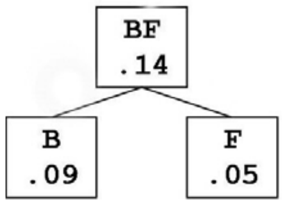


Figure 1 picture of Huffman start

You should now continue to pair the frequencies together until you have used all symbols and made a tree like this

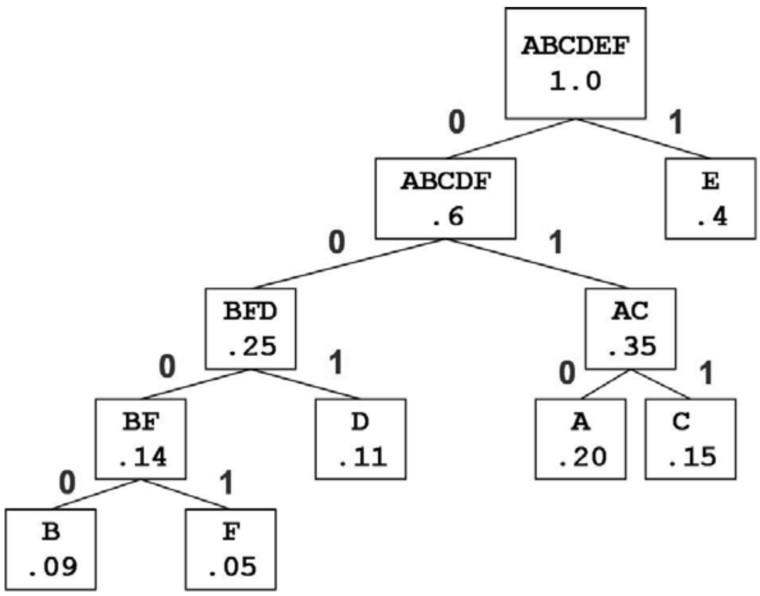


Figure 2 picture of complete Huffman [15]

Now assign zeroes and ones to the branches.

Table 2 The final codes are as follows (read from top to bottom):

Symbol	A	B	C	D	E	F
Code	010	0000	011	001	1	0001

As the table shows this is a highly effective way to encode the symbols instead of coding them with eight bits, or even with four bits.

2.2.3 Lossy compression

With lossy compression, the reconstructed image contains degradations with respect to the original image. Both the statistical and the perceptual (psychovisual) irrelevancy of image data are exploited. The leads to a much higher compression ratio compared to lossless compression. In this mode, image quality can be traded for compression ratio. [2] Lossy compression of video data can lead to ratios of between 10:1 and 50:1 without visibly degrading image quality. If wavelet compression is used a considerable higher compression rate can be achieved. Examples of lossy compression are MPEG, JPEG and JPEG 2000.

2.2.4 Typical image coder

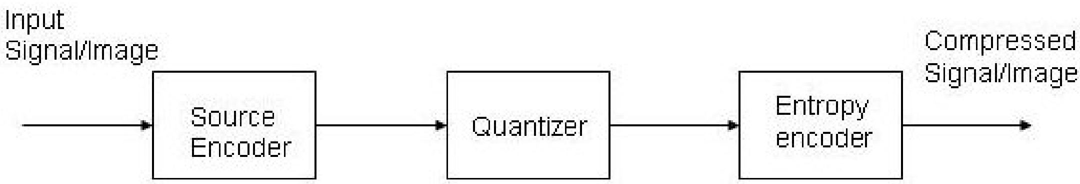


Figure 3 The figure shows a typical lossy image compression system.

A typical image coder consists of a source encoder, a quantizer and an entropy encoder.

The source encoder consists of a variety of linear transforms such as DCT, Discrete Fourier Transform (DFT), Discrete Wavelets Transform (DWT), and others.

The quantizer reduces the number of bits needed to store the transformed coefficients by reducing the precision of those values. Since this is a many to one mapping, it is a lossy process and is the main source of compression in an encoder. Quantization can be performed on each individual coefficient, which is known as Scalar Quantization. Another form of quantization is Vector Quantization, and this is performed by grouping coefficients together.

The entropy encoder further compresses the quantized values losslessly to give better overall compression. The encoder determines the probability for each quantized value and produces an appropriate code based on these probabilities. The reason for doing this is that the output stream should be smaller than the input stream. The most commonly used encoders are the Huffman and the arithmetic encoders. For applications requiring fast execution, simple run-length encoding (RLE) has proved to be very effective.
[3]

2.3 DCT

Discrete Cosine transforms are used in JPEG compression for still images and MPEG-2 compression for video. When DCT is used, the picture is split into 8x8 pixel blocks. There are different blocks for color and for brightness. Each pixel has a level value that tells something about brightness or color. Usually there are three sets of pixels, one for brightness and two for color. A DCT encoder transforms the picture blocs and converts the signals into frequency components. This results in an 8x8 matrix. The top left corner of the matrix illustrates the DC component. The rest of the values in the matrix describe the amplitudes of each frequency in the Fourier series, low frequency components top left and high frequency components bottom right. A quantizer then divides every coefficient by a number N. The quantization reduces the number of bits that are needed to store or transmit the information. This will lead to that many of the coefficients gets a value of zero. The picture is now compressed. The matrix is finally scanned in a zigzag sequence and entropy encoded. This leads to additional compression. DCT compression will often lead to distortion, and visible block effects are usual in the pictures. Examples on entropy encoding are Run Length Coding (RLE), and Variable Length Coding (VLC). Most commonly used encoders are the Huffman and arithmetic coders [16]

2.4 JPEG

JPEG is one of the most widely used compression algorithms for image compression. JPEG is an acronym for Joint Photographic Experts Group. The word joint in the group's name refers to the fact that it is collaboration between ISO and ITU-T.

JPEG is a standard that specifies both lossless and lossy compression, but the lossy version is the most widely used. The basis for JPEG image compression is the Discrete Cosine Transform. 8x8 blocks are coded relative to the first pixel in the block. The next block is then differentially coded with reference to the previous. The blocks are taken from the image in a zigzag order. The coding of the blocks causes visible block effects for high compression ratios.

3 Wavelet Transformation

3.1 Introduction

Wavelet Transformation is a rather new mathematical concept. It has been developed for use in several special fields simultaneously, amongst other signal processing, seismology, optics, and nuclear physics. It is only since the latter half of the previous century it has been viewed as a research-field in its own.

The wavelet transform is closely related to Fourier Transformation, especially Short Term Fourier Transformation, which is a windowed version of the Fourier Transformation. Normal Fourier Transformation only gives frequency information, but no time information. This is because it is based on sine and cosine functions that tend towards plus and minus infinity. It is therefore most suitable for use with stationary signals. Transient signals cannot be reproduced with the Inverse Fourier Transformation. A proposed remedy for this has been to divide the signal in small enough time intervals to view the signal as stationary, the Short Term Fourier Transform. This gives a time-frequency representation of a signal. However, due to the Heisenberg uncertainty principle, one can not exactly pinpoint neither frequency nor at what time one frequency occur. Increased accuracy of one comes at the expense of decreased accuracy of the other.

“One can investigate what spectral components exist at any given interval of time. This is a problem of resolution, and it is the main reason why researchers have switched to WT from STFT” [19]

“The wavelet transform provides a progressive or “pyramidal” encoding of the image at various scales, which is more flexible than conventional windowed approaches like the FFT” (Fast Fourier Transform) [20]

”Wavelet analysis is capable of revealing aspects of data that other signal analysis techniques miss, aspects like trends, breakdown points, discontinuities in higher derivatives, and self-similarity.” [21]

“Wavelets are mathematical functions that can cut up data into different frequency components, and then study each component with a resolution matched to its scale.” [27]

“Wavelets have advantages over traditional Fourier methods in analyzing physical situation where the signal is transient or contains discontinuities and sharp spikes” [27]

The wavelet transform splits the signal up by using a high pass filter (HPF) and a low pass filter (LPF). The signal that comes out of the low pass filter is then again filtered through a LPH and HPF. This can be done until one only has two components left. (See figure below)

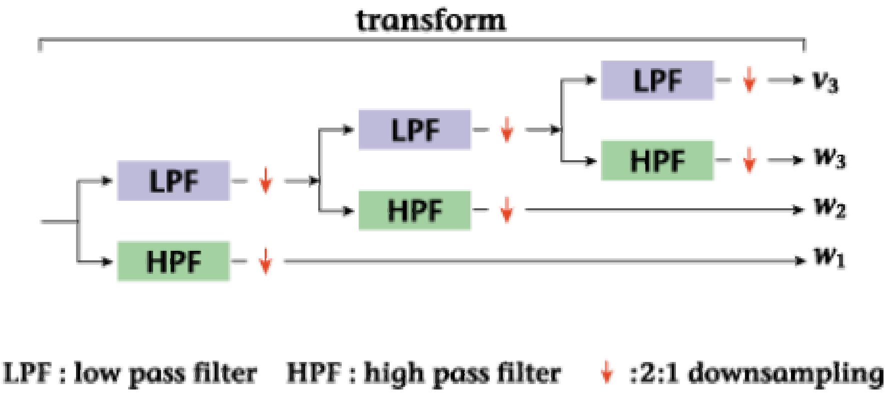


Figure 4 Wavelets transform [22]

This approach gives a higher precision in time for higher frequencies and higher spectral precision for low frequencies. For handling digital images, this is a better solution than Fast Fourier Transformation since it is closer related to human perception.

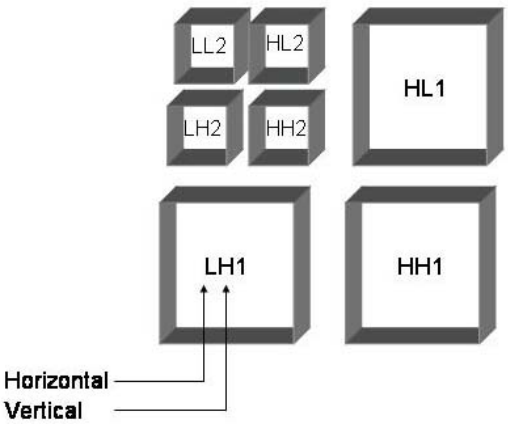


Figure 5 Wavelet sub-bands

The transformation is run both horizontal and vertical, and is split into different layers. The LL2 in the illustration above contains the low frequency information and resembles most the original picture, while HH1 contains the high frequency information. For a low-resolution image, one only uses the LL2. For images with low SNR, one uses the highest bit-planes from all the layers.

For image compression the components below a certain threshold value is set to zero. The threshold value is determined by the desired compression ratio. The higher desired compression ratio the higher the threshold value is set.

$$f(x) \rightarrow W_{\psi}[f](a,b) \xrightarrow{compression} remove\ low\ W\ values$$

The wavelet transform in it self does not lose any information. It is the setting of low values to zero that removes data.

3.2 The wavelet transform

(Contents partly form [27])
"A wavelet is a waveform of effectively limited duration that has an average value of zero." [21]

The definition of continuous wavelet transforms (CWT)

$$W(a,b)=\int_{-\infty}^{+\infty}\overline{\psi_{(a,b)}(x)}\ f(x)dx$$

Where

$$\psi_{(a,b)}(x)=|a|^{-\frac{1}{2}}\ \Psi(\frac{x-b}{a})$$

Ψ is called the mother wavelet, because it is from this function the other wavelets are extracted. The "a", which is called dilatation or scaling index, increases or decreases the spread of the wavelet as it decreases or increases. It is related to frequency in that small "a"-values is good for detecting high frequency components, and high "a"-values are good for detecting low frequencies. The "b", which is called translation, moves the wavelet along the time axis. The absolute value of "a" to the power of minus 0.5 is for normalization. The bar over the wavelet indicates the complex conjugate. The wavelet theory has incorporated the complex wavelet, but the need to use one has not yet been discovered.

For the DWT both the dilatation and translation are discrete, and increased or decreased by a factor of two. This is because a computer is fastest in calculating with numbers that are a power of two due to its binary nature.

$a=2^{-m} \qquad b=n2^{-m} \qquad m=\cdots,-2,-1,0,1,2,\cdots$

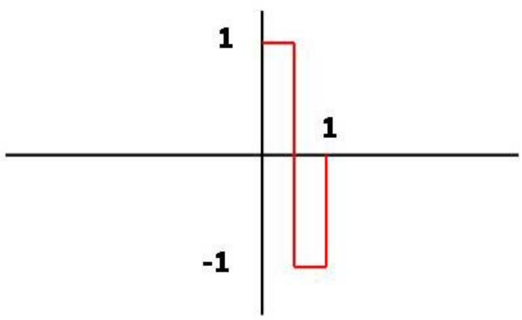
These are the most usual choices of "a" and "b", and gives the expression for the wavelet transform. "a" gives binary dilation and "b" gives dyadic dilation.

$\psi_{(n,m)}(x)=2^{\frac{m}{2}}\psi(2^{\frac{m}{2}}x-n)$ Dyadic wavelets

Some frequently used wavelets

Haar

$$\Psi(x)=\psi_{1,0}(x)=\begin{cases}1 & 0<x\le\frac{1}{2}\\ -1 & \frac{1}{2}<x\le 1\\ 0 & otherwise\end{cases}$$

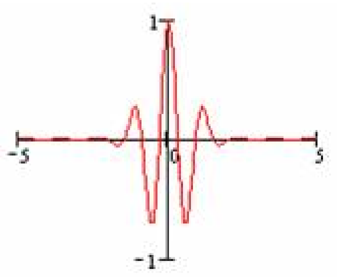


$$\psi_{a,b}(x)=|a|^{-\frac{1}{2}}\Psi\left(\frac{x-b}{a}\right)$$

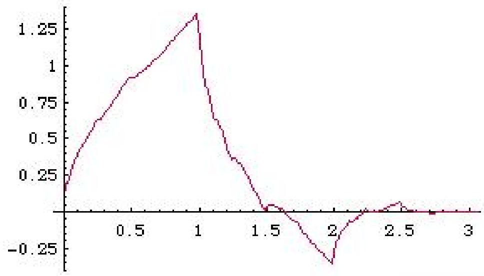
Morlet

$$\Psi(x) = e^{-x^2} \cos(\pi \sqrt{\frac{2}{\ln 2}} x)$$

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}} \Psi\left(\frac{x-b}{a}\right)$$



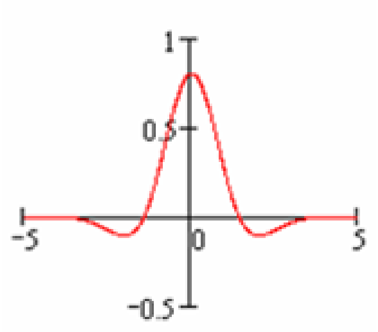
Daubechies



Mexican Hat

$$\Psi(x) = \frac{1}{\sqrt{2\pi}} \left[2 - \left(\frac{x}{\sigma}\right)^2 \right] e^{-\frac{x^2}{2\sigma^2}}$$

$$\psi_{a,b}(x) = |a|^{-\frac{1}{2}} \Psi\left(\frac{x-b}{a}\right)$$



$\sigma = 1$

3.3 JPEG 2000

JPEG 2000 is a newer standard for image compression. It is based on the wavelet transform, and specifies both lossy and lossless compression of images. It has been developed within the same organization as the JPEG standard.

JPEG 2000 has extended capabilities compared to JPEG. If one, for example want both thumbnail and high-resolution image, one can extract both from the same file by adjusting the decoding. It is also possible to download an image with increasing quality or resolution. Then one would be able to start a download with a low quality or low-resolution image and get increasing quality or resolution as the download continues, or one can abort the download when a satisfying quality is obtained. This can be used to differentiate the quality or resolution of images for guests and paying users on a web site. To do this with JPEG, one has to have two files with different quality or resolution.

JPEG 2000 also has support for ROI. The region of interest can be coded to be the first to show when downloading, or be coded with a higher quality than the rest of the image. This can be interesting for medical use, for instance an area with an anomaly in a CAT scan image can be coded as a ROI with high quality, while the rest of the image is coded with lower quality to save storage size and transmission-time.

Lossless JPEG 2000 is based on the reversible Le Galle (5, 3) taps filter, while the non-reversible Daubechies (9, 7) taps are used for lossy compression. This gives higher compression but introduces information loss. The wavelet transformation is done with discrete wavelet filters, and is done as a convolution or lifting operation. The convolution operation is done as a dot product two filter masks and the extended 1-D signal, while the filtering operation alternately updates odd sample values of the signal with a weighted sum of even sample values, and even sample values are updated with a weighted sum of odd sample values. The signal is extended periodically beyond the tile borders the same length as the filter length so that signal samples exists and corresponds to the filter coefficients.

"[JPEG 2000] is not only intended to provide rate-distortion and subjective image quality performance superior to existing standards, but also to provide features and functionalities that current standards can either not address efficiently or in many cases cannot address at all. Lossless and lossy compression, embedded lossy to lossless coding, progressive transmission by pixel accuracy and by resolution, robustness to the presence of bit-errors and region-of-interest coding, are some representative features." [18]

3.3.1 Features

(taken from [18])

- **Superior low bit-rate performance**: This standard should offer performance superior to the current standards at low bit-rates (e.g. below 0.25 bpp for highly detailed gray-scale images). This significantly improved low bit-rate performance should be achieved without sacrificing performance on the rest of the rate-distortion spectrum. Network image transmission and remote sensing are some of the applications that need this feature.
- **Lossless and lossy compression**: It is desired to provide lossless compression naturally in the course of progressive decoding. Examples of applications that can use this feature include medical images, where loss is not always tolerated, image archival applications, where the highest quality is vital for preservation but not necessary for display, network applications that supply devices with different capabilities and resources, and pre-press imagery. It is also desired that the standard should have the property of creating embedded bit stream and allow progressive lossy to lossless build-up.
- **Progressive transmission by pixel accuracy and resolution**: Progressive transmission that allows images to be reconstructed with increasing pixel accuracy or spatial resolution is essential for many applications. This feature allows the reconstruction of images with different resolutions and pixel accuracy, as needed or desired, for different target devices. World Wide Web, image archival and printers are some application examples.
- **Region-of-Interest Coding**: Often there are parts of an image that are more important than others. Published in IEEE Transactions on Consumer Electronics, Vol. 46, No. 4, pp. 1103-1127, November 2000 This feature allows users to define certain ROI's in the image to be coded and transmitted with better quality and less distortion than the rest of the image.
- **Random codestream access and processing**: This feature allows user defined ROI's in the image to be randomly accessed and/or decompressed with less distortion than the rest of the image. Also, random codestream processing could

allow operations such as rotation, translation, filtering, feature extraction and scaling.

- **Robustness to bit-errors:** It is desirable to consider robustness to bit-errors while designing the codestream. One application where this is important is transmission over wireless communication channels. Portions of the codestream may be more important than others in determining decoded image quality. Proper design of the codestream can aid subsequent error correction systems in alleviating catastrophic decoding failures.
- **Open architecture:** It is desirable to allow open architecture to optimize the system for different image types and applications. With this feature, a decoder is only required to implement the core tool set and a parser that understands the codestream. If necessary, unknown tools could be requested by the decoder and sent from the source.
- **Content-based description**: Image archival, indexing and searching is an important area in image processing. Standards like MPEG-7 ("Multimedia Content Description Interface") are addressing this problem currently. Content-based description of images might be available as part of the compression system (for example as metadata information).
- **Side channel spatial information (transparency)**: Side channel spatial information, such as alpha planes and transparency planes are useful for transmitting information for processing the image for display, printing or editing. An example of this is the transparency plane used in World Wide Web applications.
- **Protective image security**: Protection of a digital image can be achieved by means of watermarking, labeling, stamping and encryption. Labeling is already implemented in SPIFF and must be easy to be transferred back and forth to JPEG 2000 image files.
- **Continuous-tone and bi-level compression**: It is desired to have a coding standard that is capable of compressing both continuous-tone and bi-level images. If feasible, this standard should strive to achieve this with similar system resources. The system should compress and decompress images with various dynamic ranges (e.g. 1 bit to 16 bit) for each color component. Examples of applications that can use this feature include compound documents with images and text, medical images with annotation overlays, and graphic and computer generated images with binary and near to binary regions, alpha and transparency planes, and facsimile

3.4 The JPEG 2000 Architecture

Below is a block diagram of the JPEG 2000 encoder (a) and decoder (b). The JPEG 2000 algorithm first performs the discrete wavelet transform on the image. "The transform coefficients are then quantized and entropy coded, before forming the output code-stream (bit-stream)."[18]

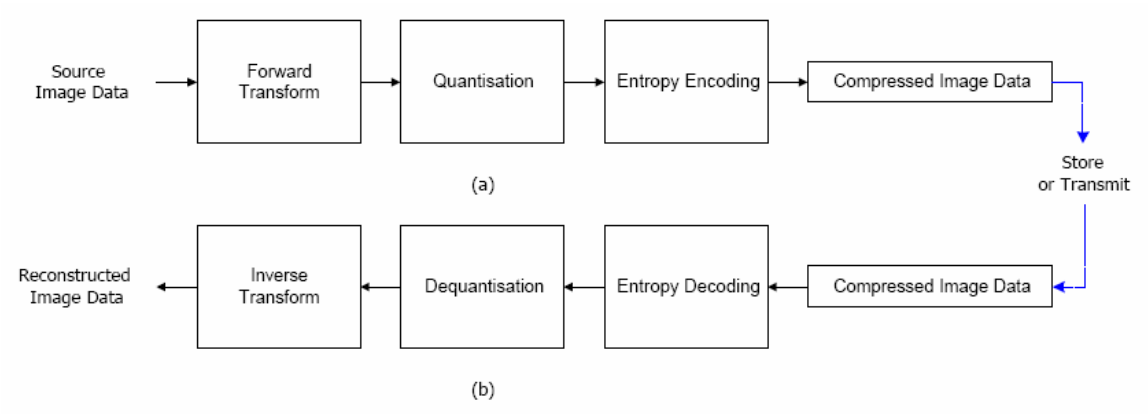


Figure 6 Block diagram of JPEG 2000 a) encoder and b) decoder [18]

”Before proceeding with the details of each block of encoder in figure 4, it should be mentioned that the standard works on image tiles. The term ‘tiling’ refers to the partition of the original (source) image into rectangular non-overlapping blocks (tiles), which are compressed independently, as though they were entirely distinct images. Prior to computation of the forward discrete wavelet transform on each image tile, all samples of the image tile component are DC level shifted by subtracting the same quantity (i.e. the component depth). DC level shifting is performed on samples of components that are unsigned only. If color transformation is used, it is performed prior to computation of the forward component transform” [18]

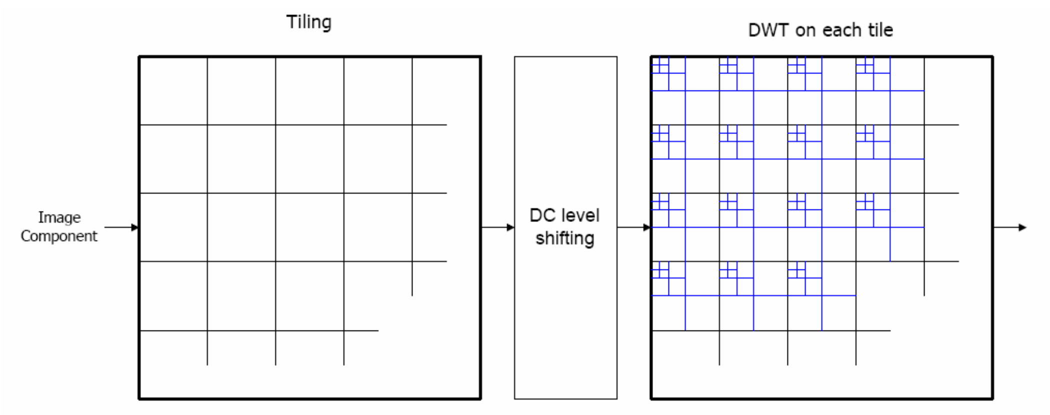


Figure 7 Tiling, DC level shifting and DWT of each image tile component. [18]

The tile size is the same for the entire picture, except at the far right and at the bottom if the picture is a size that is not an exact multiple of the tile size. The tile size can be adjusted during encoding, even to the extent that the picture is considered one tile. Larger tiles gives a better image quality, but at the cost of increased computational complexity and memory requirements.

The dividing and coding of the original image [25]

- The image is decomposed into components.
- The image and its components are decomposed into rectangular tiles. The tile-component is the basic unit of the original or reconstructed image.

- Performing the wavelet transform on a tile-component creates decomposition levels. These decomposition levels can create components with different resolutions.
- These decomposition levels are made up of sub-bands of coefficients that describe the frequency characteristics of local areas (rather than across the entire tile-component) of the tile-component.
- The sub-bands of coefficients are quantized and collected into rectangular arrays of code-blocks.
- The bit-planes of the coefficients in a code-block are entropy coded in three coding passes.
- Some of the coefficients can be coded first to provide a region of interest.

Reconstruction of the coded image[25]

- The image is decomposed into components.
- The image and its components are decomposed into rectangular tiles. The tile-component is the basic unit of the original or reconstructed image.
- Performing the wavelet transform on a tile-component creates decomposition levels. These decomposition levels can create components with different resolutions.
- These decomposition levels are made up of sub-bands of coefficients that describe the frequency characteristics of local areas (rather than across the entire tile-component) of the tile-component.
- The sub-bands of coefficients are quantized and collected into rectangular arrays of code-blocks.
- The bit-planes of the coefficients in a code-block are entropy coded in three coding passes.
- Some of the coefficients can be coded first to provide a region of interest.

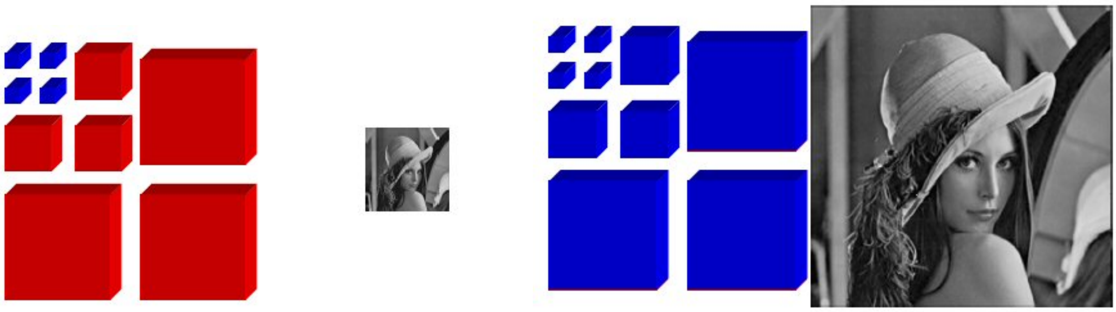


Figure 8 Image decoded with increasing resolution [28]

Figure 9 Image decoded with increasing quality [28]

JPEG 2000 has support for multiple information components. The illustration shows the encoding of a three-component RGB image.

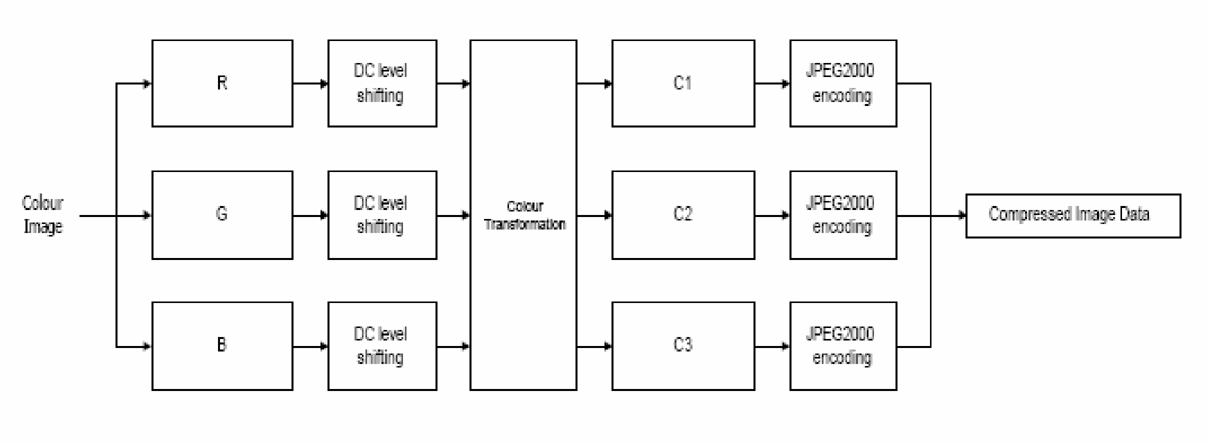


Figure 10 Structure of the JPEG 2000 multiple-component [18]

3.4.1 Region of Interest

To emphasize important aspects of an image one could code the area as a ROI. An ROI is a part of an image that is coded earlier in the code-stream than the rest of the image. The ROI will be emphasized in the encoding procedure and the ROI will be downloaded first with higher priority, and be shown with high bitrate while the rest of the image loads.

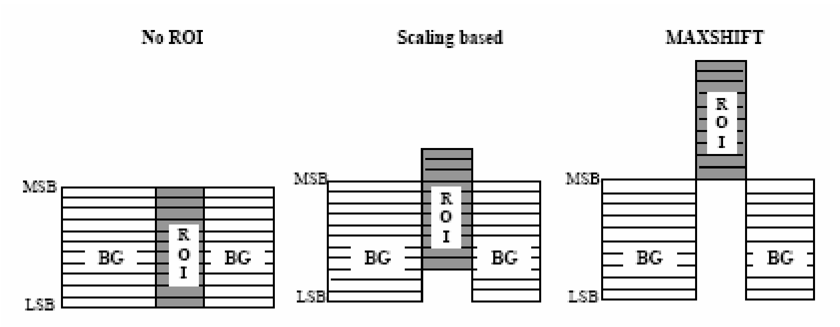


Figure 11 Scaling of ROI coefficients [18]

The ROI is shifted so the bits are associated with higher bit-planes than the other bits in the picture. The ROI-shift can be done as general scaling or max-shift. The higher bit-planes are placed earlier in the bit-streams. By using the max-shift method, the minimum coefficient value of the ROI is shifted to be lager than the maximum background coefficient value, while by using the general scaling method the enhancement of the ROI can be set.

3.4.2 Error Resilience

The Variable Length Encoding (VLE) used for JPEG 2000 is not ideal for error prone channels, but several steps have been made to overcome this problem. The entropy coding is split into code blocks. The coding and decoding of a code block is independent of the others and an error will only affect one code block. Termination of the arithmetic coder is allowed after every coding pass allows the arithmetic decoder to continue decoding after errors.

"To improve the performance of transmitting compressed images over error prone channels, error resilient bit stream syntax and tools are included in the standard.

The error resilience tools deal with channel errors using the following approaches: data partitioning and resynchronization, error detection and concealment, and Quality of Service (QoS) transmission based on priority." [18]

The JPEG 2000 standard also uses a small packet format with per packet resynchronization markers.

4 UMTS

4.1 Introduction to UMTS

Mobile Communication has become more and more popular since the introduction of the first generation systems in the 1980s and second-generation systems in the 1990s. The first generation (1G) systems where analog, voice-only systems with limited security mechanisms. With 2G systems, the step was taken into digital communication, with increased security, but limited data transfer capability. GSM (Global System for Mobile communication) is a 2G system that is widely used in Europe. It can support circuit switched data transfer rates from 9.6-14.4 kbps. The data transfer capability has been extended with the introduction of General Packet Radio Service (GPRS), and these systems are referred to as 2.5G. As the name implies the data-transfer is now done in a packet switched manner instead of circuit switched. These systems have made peer-to-peer communication wireless and the 2G systems provide popular services, i.e. text-messages.

UMTS is a third generation system, with increased security and data transfer capabilities. It has been specified by the 3rd Generation Partnership Project (3GPP), and was intended to be a world wide standard for wideband mobile communication, with support for satellite communications when no base station was available. This was not the result.

Important issues for specifying the 3G systems where extended support for data transmission. Different types of data have different demands to the transport channel. Speech, web browsing and video streaming has for instance very different demands on bandwidth, delay and jitter. By specifying Quality classes the system can take the different demands into account and give different priorities to the classes.

Table 3 Quality Classes in UMTS [23]

Traffic class	Conversational class Real Time	Streaming class Real Time	Interactive class Best Effort	Background class Best Effort
Fundamental characteristics	- Preserve time relation (variation) between information entities of the stream - Conversational pattern (stringent and low delay)	- Preserve time relation (variation) between information entities of the stream	- Request response pattern -Preserve data integrity	-Destination is not expecting the data within a certain time - Preserve data integrity
Example of the application	Voice, video-telephony, video games	streaming video	web browsing, network games	telemetry, emails

4.2 The UMTS channel

In most 2G systems, the user is assigned communication channels as a timeslot at a specific frequency. These are Time Division Multiple Access/Frequency Division Multiple Access (TDMA/FDMA) systems, which uses different frequencies in different radio coverage cells. The frequencies are assigned by national government agencies, and are limited natural resources. To use the assigned frequencies efficiently, extensive radio planning is needed, where topography, buildings and user density is taken into account.

UMTS uses a spread spectrum technique called Wideband Code Division Multiple Access and differentiates between uplink and downlink by using separate frequency bands (Frequency Division Duplex – FDD) or by using different timeslots for up- and downlink (Time Division Duplex - TDD). All the user channels are separated by different spreading codes and the frequency reuse factor is one, i.e. all the users are using the same frequency and no frequency planning is needed.

WCDMA is a spread spectrum technique that gives a more robust channel. Spread spectrum technique was developed to withstand jamming, i.e. deliberate attempts to block communication, and therefore gives a better protection against natural influences, like noise and shadowing. The transmission is spread over a frequency interval as opposed to sending at a single frequency, and this makes it more difficult to eaves drop, since the transmission appears to be noise if one cannot produce the right descrambling codes.

It is important to have fast power control on both user equipment and base station (Node- B). The available bandwidth is dependent on Signal-to-Interference Ratio (SIR) and it is important that none of the users transmit with a signal that is to strong, because it will increase the interference for the other users and limit their bandwidth. If all users were to transmit with the same signal strength, the user closest to the Node-B would overpower all the other users' signals. This is called the near-far problem and is handled by close monitoring of SIR and fast power-control of UEs.

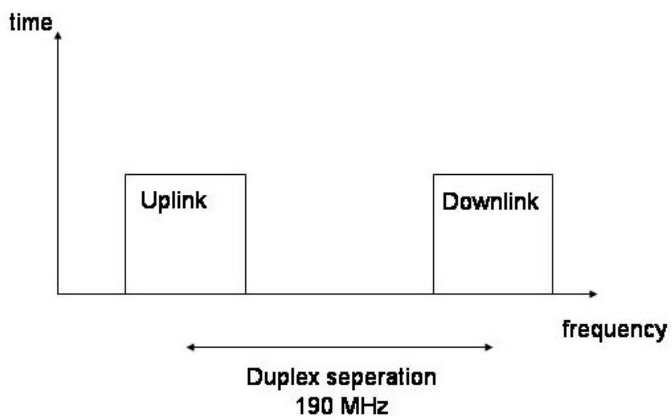


Figure 12 FDD mode [23]

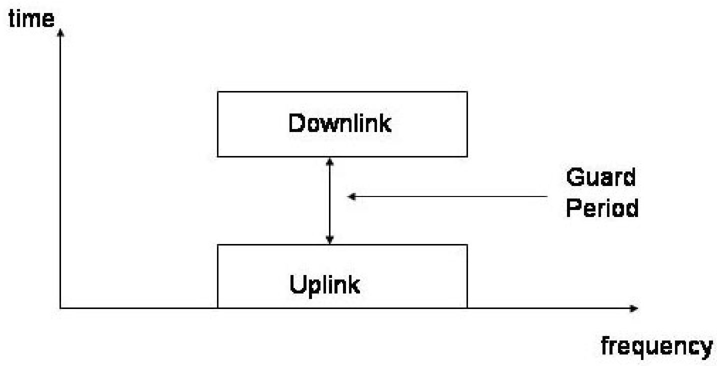


Figure 13 TDD mode [23]

The high bandwidth coverage is limited by several factors. The low transmit power capabilities of the UE gives lower uplink bandwidth with increasing distance to the B-Node, because of worse signal noise ratio. UE size and battery life also prohibits the use of multiple antennas and receivers in the handset, which can extend the range of high bandwidth reception. The B-Node has more complex receivers and transmitters; because power consumption is not a problem here.

The bandwidth is also affected by the UE speed. “At the cell edges the best performance is obtained for high mobile speeds, while close to the base station low mobile speeds perform better.” [23] One of the reasons for this is higher time diversity for higher speeds, which gives coding gain.

4.2.1 The connection between bandwidth and quality in UMTS

The bandwidth in UMTS is larger than the 2G systems. One can expect up to 384kbps transfer rate for most of the cell coverage areas, while the higher bit rates will have a more limited coverage.

Table 4 UMTS bit rate [24]

Class	Rate	Example
High-speed movement	144 kb/s	In vehicles
Pedestrian	384 kb/s	On campus
Stationary	2 Mb/s	In buildings

When sending a single picture from one mobile phone to another, the bandwidth is usually not a problem. If this should take as much as five seconds, it is no problem. Bandwidth limitation is more of a problem when we want to send a sequence of pictures like a video sequence. The possibility for a jerky transmission is definitely high.

The motion JPEG-2000 does not use interframe coding, which is an advantage for high bit error channels like mobile channels, because it makes it more resistant against packet loss, or frame loss.

The bandwidth limit gives a maximum transfer rate of 48 kB per second. The maximum image rate supported by Motion JPEG-2000 is 30 images per second. At maximum image rate a single picture can only be 1,6 kB, not taking overhead and retransmissions into account.

The images we have used in the subjective test have a bit rate of 0.8 bits per pixel. This value has been uses in the calculations:

$$\frac{BW}{bit/pixel} = \frac{384kbps}{0.8bits/pixel} = 480\,000\,pixel/s$$

$$\frac{480000\,pixel/s}{30img/s} = 16000\,pixel/img$$

The bandwidth does not allow large pictures to be transferred at 0,8bpp.

We have a video sequence we want to transmit. The conditions are good, and we get a transmission rate of 384 kbps. According to the new proposed motion JPEG2000 standard, the frames will not be interframe coded i.e. the frames will be coded individually. Let's say that the pixels are coded at a rate of 0.8 bpp. Each frame is 240*180 pixels. We then get 240*180*30*0.8 bits to transfer per second. This amounts to 1036800 bps or about 3 times as much as available bandwidth. According to this we have to code the pixels at around 0.3 bpp, which makes the quality very poor.

Although a very low bit per pixel rate makes the image quality poor, a sequence of pictures is less vulnerable than a single picture is. The shifting of pictures makes is less likely that the eye will notice the impairments in the single frames, and therefore a much lower quality is acceptable in a video sequence, than it would be in a single still picture. Therfore the value 0,8 that came from still images could probably be lowerd to some extent. It is also possible to reduce the frame rate somewhat and use a lower resolution. This will decrease the quality, but the number of pictures per second is probably the one that would degrade the quality perception the most.

5 Image Quality

5.1 Introduction

The main concern in video and image compression is the quality we get on our compressed media. This chapter deals with quality, and some of the different parameters that is important to measure picture quality. So what is meant by image quality, and what parameters are used to measure such a quality? These are some of the questions we have tried to answer in this chapter. We have also tried to explain the two main quality assessment methods, objective and subjective quality measurement.

5.1.1 Factors affecting image quality perception

When going through the factors that affect quality perception, it should be noted that the parameters described here are first and foremost affects the subjective measurements. Often impairments will not be noticed by a viewer. Usually when we compress we will subtract information that is not noticeable to the human eye. Of course, the visible impairment will become noticeable at some point, as the compression ratio increases.
Some of the factors that affect the image quality are:

- Contrast ratio, which defines the dynamic range of the display system.
- Resolution witch is the system's ability to display fine detail.
- Quality of the original pictures.
- The digitization and conversion parameters selected
- Processing and compression algorithms.
- Compression ratio.
- Viewing environment.

5.1.2 Terms

Terms partly from [5]
This list describes some terms and definitions that are important when we discuss image quality.

Image quality: The image quality is measured either subjectively or objectively. The parameters that decide if a picture is to be considered high quality will differ when either of them is used. Generally, it is accepted that key factors to image quality are: Signal to noise ratio, spatial resolution and contrast resolution.

Pixel: A picture element that describes the brightness or color of a discrete point in an image.

Block: Group of pixels. For example a 8x8 block used in DCT,.

Resolution: Parameter that specifies the ability to distinguish video detail in the spatial or temporal dimension.

Chrominance: consists of two components, hue and saturation.

Luminance: refers to the brightness information of the picture /video signal. Luminance is a measure of the amount of light energy that is received by an observer. It is measured in lumens (lm).

Hue: The hue of a color is characterized by the dominant wavelength in the composition.

Saturation: Saturation is a measure of the purity of a color. A pure color has a saturation of 100% and white has a saturation of 0.

Spatial resolution: Spatial resolution is a term that refers to the number of pixels utilized in construction of a digital image. Images having higher spatial resolution are composed with a greater number of pixels than those of lower spatial resolution.

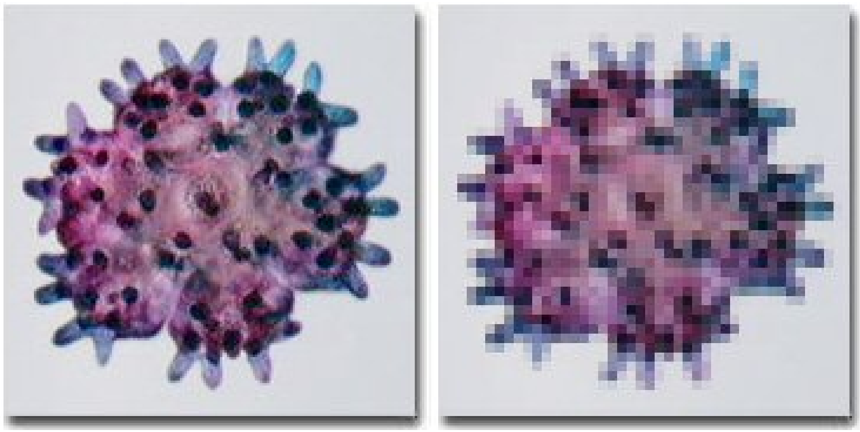


Figure 14 Spatial resolution [8]

The picture on the left has a resolution of 175x175 and the right picture has a resolution of 44x44. Spatial resolution describes how much detail in a photographic image is visible to the human eye. High-resolution images are sharp and more details are visible.

Contrast resolution: Described as the minimum perceptible difference in the luminance between two adjacent pixels a display. So contrast resolution measures the ability to distinguish between differences in image intensity. The contrast resolution of a digital image is given by the number of possible pixel values, and is defined as the number of bits per pixel value

Spatial perceptual information (SI): A measure that generally indicates the amount of spatial detail in a picture. The SI is usually higher for more spatially complex scenes.

Temporal perceptual information (TI): A measure that generally indicates the amount of temporal changes of a video sequence. It is usually higher for high motion sequences.

5.1.3 Impairment terms

The following terms describe different impairments that may occur in pictures.

Block distortion (or tiling): Distortion of the image characterized by the appearance of an underlying block encoding structure.

Blurring: Global distortion of the entire image, characterized by reduced sharpness of edges and spatial detail.

Color errors: Distortion of all or a portion of the final image characterized by the appearance of unnatural or unexpected hues or saturation levels.

Edge busyness: Distortion concentrated at the edges of objects, and further characterized by temporal and spatial characteristics.

Error blocks: A form of block distortion where one or more blocks in the image bear no resemblance to the current or previous scene and often contrast greatly with adjacent blocks.

Jerkiness: Motion that was originally smooth and continuous perceived as a series of distinct “snapshots”.

Mosquito noise: Form of edge busyness distortion sometimes associated with movement, characterized by moving artifacts and/or blotchy noise patterns superimposed over the objects (resembling a mosquito flying around a person's head and shoulders).

Motion-related artifacts: Distortion of motion video potentially observable by the viewer. In some instances, the distortion becomes more observable with increased motion. The distortion may appear as smearing, block distortion, jerkiness, or other impairments.

Motion response degradation: The deterioration of motion video such that the video imagery has suffered a loss of spatio-temporal resolution.

Object persistence: Distortion where the object(s) that appeared in a previous video frame (and should no longer appear) remain(s) in current and subsequent video frames as outline or faded image.

Object retention: Distortion where a fragment of an object that appeared in a previous video frame (and should no longer appear) remains in the current and subsequent video frames.

Quantization noise: A “snow” or “salt and pepper” effect similar to a random noise process but not uniform over the image.

Smearing: This is a local distortion over a sub region of the received image. Smearing is characterized by reduced sharpness of the edges and spatial detail. For example, the portrayal of fast moving object may exhibit smearing.

Spatial edge noise: A form of edge busyness characterized by spatially varying distortion in proximity to the edges of objects.

Temporal edge noise: A form of edge busyness characterized by time-varying sharpness (shimmering) to edges of objects.

5.1.4 Subjective quality measurement

The subjective visual quality measurement plays an important role in visual communications. It is natural that human viewers should judge the visual quality of reconstructed images or video frames since they are the ultimate receivers of the data.

In subjective quality measurement, pictures or video sequences with different quality is presented to the observers who are asked to evaluate their visual quality. The observers are asked to rate the quality by some measure of picture quality. Alternatively, the observers are asked to provide some measure of impairment to the pictures. [1]

The table below shows a five rate scaling system of the degree of impairment. This is numbered ITU-R 500-3

Table 5 Subjective evaluation-scale for images

1	2	3	4	5
Impairment is not noticeable	Impairment is just noticeable	Impairment is noticeable, but it is not objectionable	Impairment is objectionable	Impairment is extremely objectionable

There are several other different subjective quality measurement methods, but they will not be described in this thesis.

5.1.5 Objective quality measurement

Objective quality measurements are conducted by using electrical instrumentation. All types of objective video or image quality assessments are done by measuring differences between the original and received data. For instance, one can measure the signal to noise ratio on the compressed media and compare it to the original.

The Human Visual System is very complicated and not yet fully explored. To make the most of image compression knowledge about the HVS is a key factor. The perception of still images and video sequences is only partly understood, and it is therefore difficult to design objective quality measurement methods that can be compared to subjective methods.

5.2 The affect screen size has on quality

The screen of a 3G phone will range from 1 ½ x 2 inches to approximately 2 x 4 inches. This fact has a dramatic impact on a person’s perception of either images or video. On a screen of this size, details in a picture will be almost impossible to detect for the viewer. A landscape picture will for instance show you the general outline, but details like small rocks bushes etc. will not be noticed due to the small screen. This fact leads to the possibility of removing those details from an image, since a viewer does not notice them anyway. One can expect the quality of the screens to become better in the years to come, expansion of the screens can also be expected though only to a certain level. It is obviously not practical to have too big screens on a mobile device like a UMTS phone.

5.3 Redundancy

(Partly collected from [1])

What we want to take advantage of when we compress data is redundancy. Redundancy means information we don't need to reconstruct the original data adequately. This subchapter deals with different types of redundancy.

5.3.1 Statistical redundancy

Statistical redundancy is divided into two types. These are interpixel and coding redundancy.

Interpixel redundancy:

Interpixel redundancy means that pixels of an image frame and pixels of a group of successive image or video frames are statistically dependent. On the contrary, they correlate to various degrees. Interpixel redundancy can be divided into to categories, spatial redundancy and temporal redundancy.

Temporal redundancy:

Temporal redundancy (also called interframe redundancy) is concerned with the statistical correlation between pixels, from successive frames in a temporal image or video sequence.

Spatial redundancy:

Spatial redundancy (also called intraframe redundancy) represents the statistical correlation between pixels within an image frame.

Coding redundancy:

By coding-redundancy, we mean the statistical redundancy associated with coding techniques.

5.3.2 Psycho-visual redundancy

While interpixel redundancy inherently rests in image and video data, psycho visual redundancy originates from the characteristics of the Human Visual System. The HVS does not precept visual information in the same way as a camera does. This is what we want to take advantage of in psycho visual redundancy techniques. Some information may be more important than other information. This implies that if we apply fewer data to represent less important visual information, perception will be unaffected. This means that some information is what we call psycho visual redundant.

Several characteristics used in compression utilize the HVS. The most important are luminance masking, texture masking, frequency masking, temporal masking and color masking.

5.4 HVS-The Human Visual System

In many image-processing applications, the objective is to help a human observer perceive the visual information in an image. Therefore, it is important to understand the Human Visual System.

The HVS consists mainly of the eye (image sensor or camera), optic nerve (transmission path), and brain (image information processing unit or computer). It is one of the most sophisticated image processing and analysis systems around.

As said earlier the HVS is very complex and not fully explored yet. Its understanding would also help in the design of efficient, accurate and effective computer/machine vision systems. This been said, many aspects have been clarified already, and this chapter will give a short walkthrough of some of the features of the HVS
[7]

5.4.1 Color perception

The human eye has two types of photoreceptors: rods and cones (light sensors).

Humans perceive color using cone-shaped cells at the back of the retina (from dictionary.com *retina* is described as follows: A delicate, multilayered, light-sensitive membrane lining the inner eyeball and connected by the optic nerve to the brain.) There are 6-7 million cones, located in the central portion of the retina. The cones are responsible for the photopic vision, that is the bright light vision and color perception. The cones can resolve fine details. Most color perception is concentrated in the fovea, (from dictionary.com *fovea* is described as follows: small depression in the retina containing cones and where vision is most acute) a small area about 1,5mm across where all but peripheral vision is sensed. The photoreceptors around fovea responsible for spatial vision (still images).

The rods are distributed over the entire retina, and there are between 75 and 150 million of them. They are responsible for the scotopic vision (dim light vision). The rods are not color sensitive and give a general overall picture. They do not contribute to the detailed vision. The blind spot is a point on retina where optic nerve emerges, and it does not contain any photoreceptors. The photoreceptors around the periphery are responsible for detecting motion.

The three types of color cells in the eye roughly correspond to their sensitivity to red, green, and blue colors, centered at wavelengths of 588 nanometers (nm), 531 nm, and 419 nm. These are not narrow-band sensors. The sensitivity of the red and green cones extends from about 400 to 750 nm. The numbers of green, red and blue cones are not equivalent. The average person has only 1 blue cone for every 20 green and every 40 red cones. This is why image detail in the blue is not perceived as well as in the green and red. As said earlier the eye contains roughly 6-7 million cones. Because of this, image capture can be tuned to bias capture of green details above blue, and most sensors use an array with twice as many green as red or blue sensor elements. As the absorption spectrum of red and green cones is quite similar, the eye must perform a difference operation between red and green signals to distinguish between these colors. Perception of light intensity is therefore significantly stronger than perception of color. Color perception is highly adaptive to both intensity and lighting color. Within a few seconds or minutes of transitioning from one to another light, our mind adapts to the new illumination and our perception of color changes appropriately. When wearing heavily tinted yellow lenses on a snowy day for instance, the snow initially appears yellow, but within a few minutes is perceived as white again. Both of these light adaptation capabilities are challenges for all photographic methods. [6] [7]

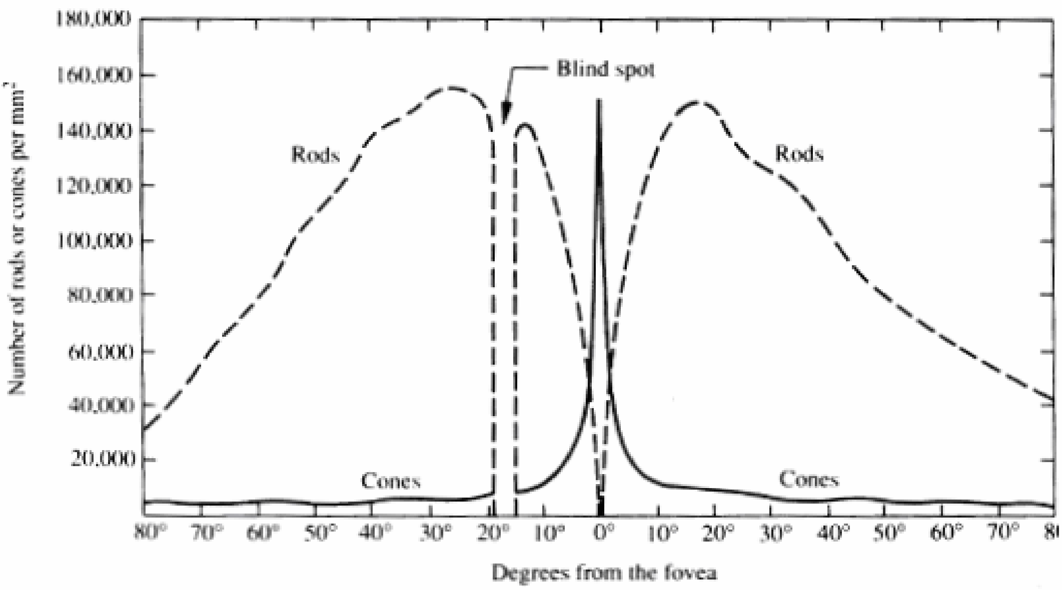


Figure 15 Picture from [7]

5.4.2 Resolution perception

The ability to resolve detail is measured in lines per unit of radial angle. This is equivalent to lines per unit of distance if the viewing distance is fixed. At 50 to 60 lines per degree, humans lose the ability to see individual lines, and a gray single color is achieved. This is a result of the density of cone cells in the fovea, generally accepted as 120 cells per degree in adult humans. One needs two cells to perceive a line and the white space next to a line. At a viewing distance of 30 cm (or about a foot), the average person can distinguish line spacing of 0.1 mm, or about 250 lines per inch in black and white. Color variations at uniform brightness are perceived with substantially less resolution.

Interestingly, human contrast perception is stronger at five lines per degree than at lower frequencies. In other words, people see 25 lines per inch more strongly than 2 lines per inch, when viewed at a distance of 1 foot. This is an anatomically driven result of the way the eye perceives edges and contrast. [6]

Luminance masking:

> Luminance masking (also referred to as luminance dependence and contrast masking) concerns the brightness of the HVS. To illustrate the effect of luminance masking, here is a sequence of pictures:

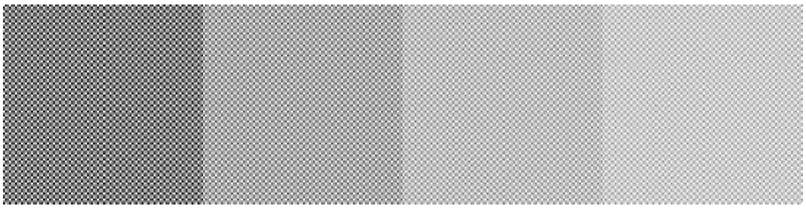


Figure 16 Luminance masking

Each of these images have the same amplitudes but with different mean, the lowest on the left. As can be seen, the pattern is more noticeable towards the left. When the average brightness is higher, the same amount of regional change amounts to a lower contrast as compared to a lower average brightness. Thus, the same variation in a bright region would be less visible than in a darker region [4]

Texture masking:

(Also called detail dependence, spatial masking and activity masking) The number of quantization levels, which has a considerable effect on bit rate, should be adapted to the intensity variations in the different parts of a picture/frame.

Frequency masking:

(Also called frequency dependence) While the two already mentioned characteristics are picture dependent, frequency masking is picture independent. High frequency random noise is added to a picture before quantization. This improves the quality because the low frequency quantization error is shifted to high frequency noise, and the HVS is less sensitive to high frequency content. This means that we can drop some high frequency coefficients with small magnitudes to achieve data compression without noticeably affecting the perception of the HVS.

Temporal masking:

Also, like frequency masking, picture independent. The HVS needs time to adapt itself to a scene when the scene changes abruptly. During this change, the HVS is not sensible to details. The masking takes place both before and after the abrupt change.

Color masking:

Any visible light corresponds to an electromagnetic spectral distribution, so a color as a sensation of visible light, is energy with intensity as well as a set of wavelengths associated with the electromagnetic spectrum. Intensity is an attribute of visible light. The composition of wavelengths has another attribute: chrominance. Chrominance consists of two attributes, hue and saturation.

Hue:

The hue is characterized by the dominant wavelength of the color.

Saturation:

The saturation is a measurement of the purity of a color. A pure color will have a saturation of 100%. White light will have a saturation of 0.

6 Results

6.1 General

This chapter gives a walkthrough on our work during the months we have been working on the thesis. In this chapter, we have tried to give a description of the different parts of the thesis, but have gone into the subjective and objective test results more detailed than the other subjects we have been working on.

6.2 Channel simulator

At the start of the project period, it was our intention to make a complete system that we could use for our simulations. As the thesis-description says: *"The lack of commercially available equipment for 3G systems makes it important to develop a model for a system based on available software. This could become a critical part of the thesis, because the availability and types of software is unknown."* This part of our project indeed turned out to be a critical part, and after two months of trying to get our hands on suitable software we had a project meeting with our supervisor and came to the agreement that the channel simulation part of the project had to be terminated. Before we came to this conclusion, we had been in contact with Sony-Ericsson, Motorola, Nokia and several other companies trying to get the necessary software for channel simulation. The big companies all had programs for simulation but said they considered them company secrets. They were all helpful though in trying to help us with finding available software on the Internet. We found several programs, but the problem was that these programs were meant to simulate a network as a network itself, and not to simulate traffic through the network.

Examples on programs we considered are Glomosim [10] and NS-2 [11]. We also tried to make a simulator using Simulink [12] from MathWorks. The latter turned out to be an enormous task witch we abandoned after just a few days. We then turned to Professor Matthias Pätzold at AUC who told us that only a few alternatives were commercially available and they were hardware based. They would also come with a formidable cost. We could of course make a channel simulator ourselves, implementing it on a signal processor, but he felt that this task alone would be to comprehensive for a master thesis.

Since the channel simulator part of the thesis had to be cancelled, we decided to concentrate on each end of the thesis description i.e. the compression and quality measurement parts. The part of the thesis that deals with channels and bandwidth has therefore been given a much lower priority, as the observant reader will have noticed.

6.3 Wavelet compression software

For wavelet compression we have used a program from Luratech [9] called LuraWave. Although it is a small program, it has several useful features. LuraWave compresses the images to JP2 (JPEG 2000). The program also lets you determine the quality of the compressed pictures.

The program compares the original with the compressed picture and does the following calculations

MAE	Maximum Absolute Error	$\max\left\vert f(x,y) - \tilde{f}(x,y)\right\vert$
MSE	Mean Squared Error	$\frac{1}{N \cdot M}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left[f(x_i,y_j) - \tilde{f}(x_i,y_j)\right]^2$
RMS	Root Mean Squared Error	$\sqrt{MSE}$
SNR	Signal to Noise Ratio	$10\log\left(\frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} f(x_i,y_j)^2}{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left[f(x_i,y_j) - \tilde{f}(x_i,y_j)\right]^2}\right)$
PSNR	Peak Signal to Noise Ratio	$20\log\left(\frac{255}{RMS}\right)$

Equations from LuraWave

We did an extensive search on the net for suitable compression software, but most of the programs we found were commercial i.e. if someone wanted to use them, they would have to pay. Finally we came across LuraWave witch suited our needs perfectly, since it also in addition to the ability to compress pictures, did the earlier mentioned calculations.

In addition to LuraWave, we wanted to test how other JPEG 2000 compression software performed. We decided to compress the same pictures as we had used earlier and compress them at the same ratio. We decided to use Paint Shop Pro 8, who has a built in JPEG 2000compression tool, and we found a JPEG 2000 plug-in for Photoshop 7. The plug-in is called “fnord” and can be downloaded from [26]

6.3.1 Compression of test pictures

The test pictures we used were compressed at a ratio of 30:1. When we compared them with the original pictures the degradation was very much noticeable on all of them, but we expected that when standing by themselves, they would vary from being graded as ok to poor in quality.

6.4 Web pages

For the subjective testing purposes, we decided to make a set of web pages. The web pages were constructed using ordinary HTML, PHP and JavaScript.
The start page looks like this:

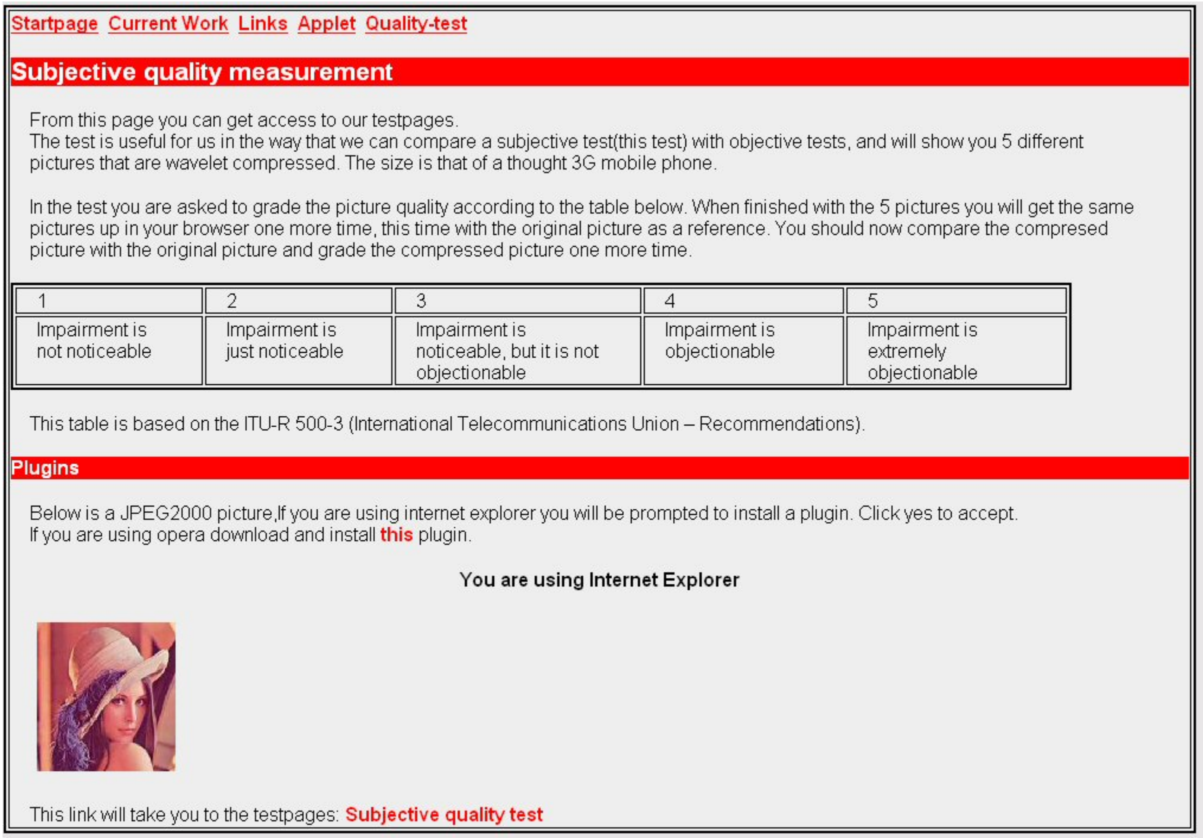


Figure 17 Webpage

When the user enters the webpage, the server determines if he is using Internet explorer. If the test returns true he is prompted to install a JPEG 2000 plug-in, which allows him to watch JPEG 2000 in his browser. If the browser is identified as anything else, the plug-in will not be automatically installed. If the user is using Opera, he can download and install a plug-in that is included on the web page. When the user has installed the required plug-in, he is able to display the "Lena" picture in the left bottom corner of the page. The plug-ins that we have used are supplied by jp2IE [13] for Internet Explorer and Elysium [14] for Opera. The page also tells the user how he should rate the picture in the test.

This is the first part of the test, consisting of a compressed picture without a reference picture.

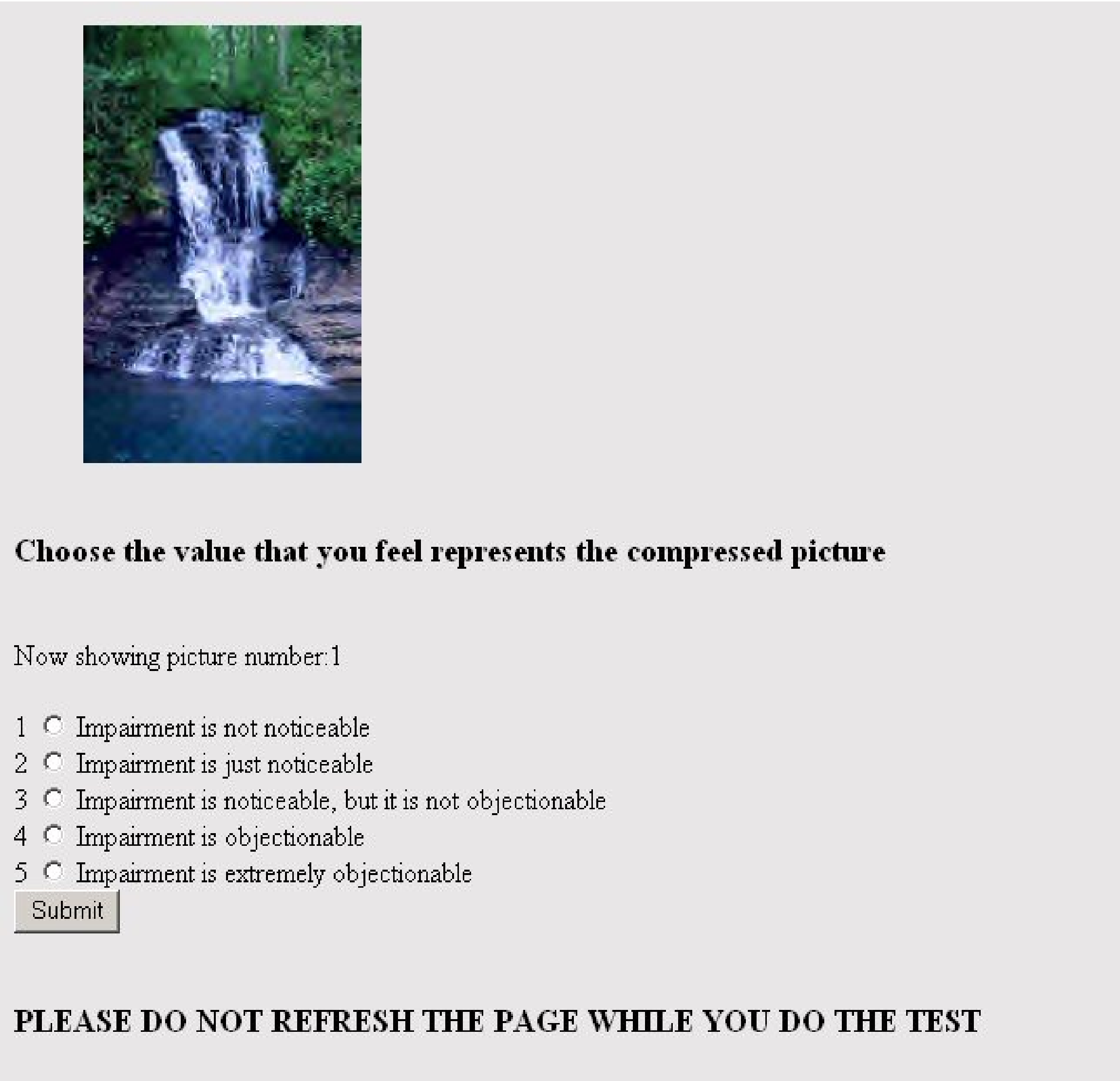


Figure 18 Test page without reference picture

All the pictures have been compressed using LuraWave. The pictures are originally BMP pictures and have been compressed with a compression rate of 30:1. The pictures are stored as JP2.

The page has been made almost entirely using PHP but there is a JavaScript that redirects the user to the next page when he has cycled through the first five compressed pictures. The input from the user is written to a file on the server.

This is part two of the quality test. The page consists of a compressed picture and a reference picture

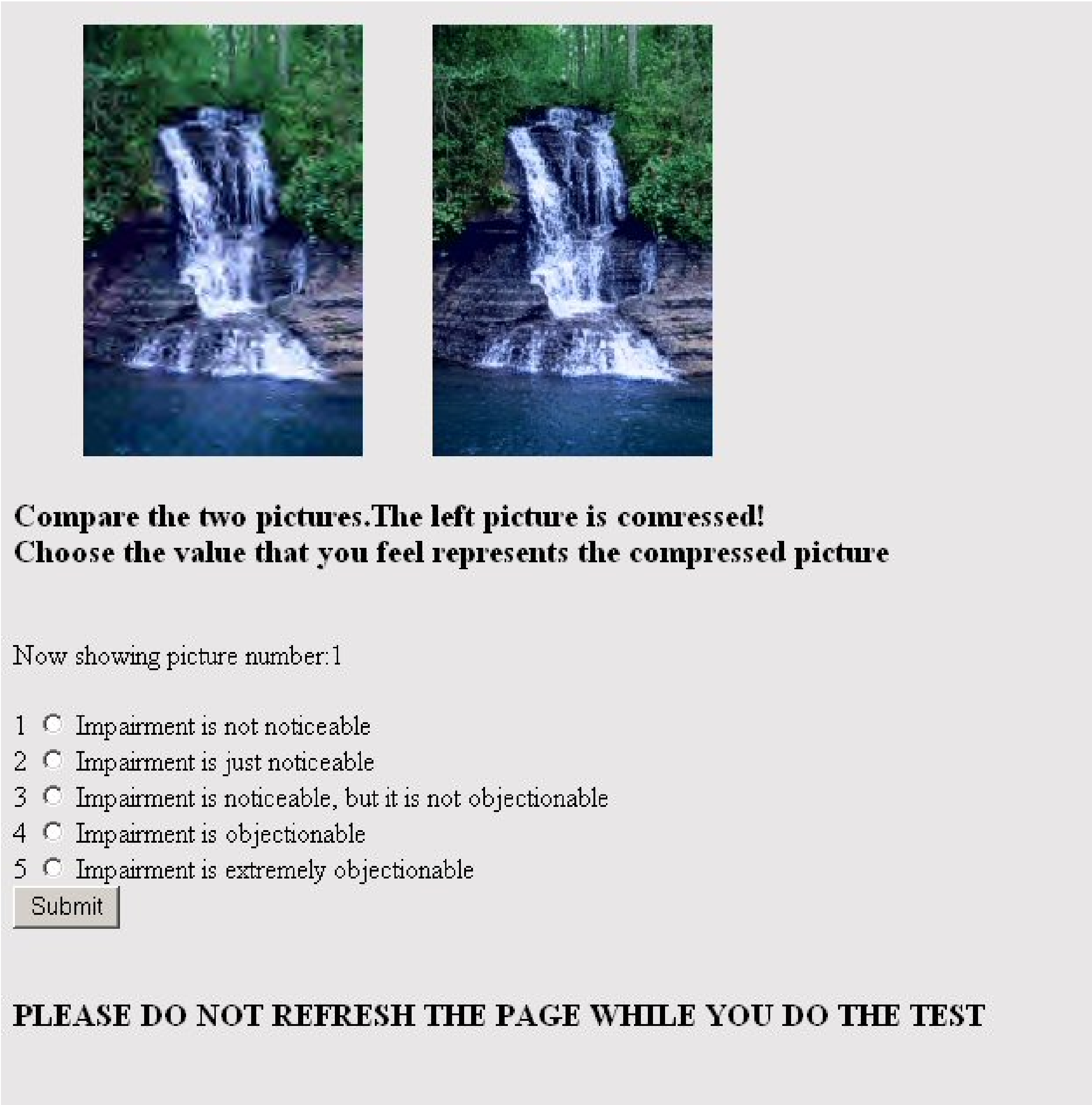


Figure 19 Test page with reference picture

The user gets the same compressed picture as in the first part. This time it is located at the left side of the screen. On the right side of the screen, the original picture is displayed as a reference. The user is asked to compare the two pictures and grade the compressed picture with respect to the original one. When the user finishes the test, he is redirected from the pages.

6.5 Test results

We decided to use a subjective test that was divided into two parts. First, we wanted to see how the users reacted to a compressed picture without a reference picture. When they had gone through the pictures they would be redirected to a new page witch had, in addition to the earlier compressed pictures, the original uncompressed pictures. In theory, the compressed pictures should score lower when compared to the original pictures than they did by themselves, so we wanted to see if our test could support this. In addition to the subjective test results, you will find a subchapter called objective test results were the objective measurements we have been able to do is described.

6.5.1 Subjective test results

This is the results from the subjective test. A number of people have taken the test during the last month of the thesis period.

Table 6 Subjective Test Results

Picture 1	Picture 2	Picture 3	Picture 4	Picture 5	Picture 1 with reference	Picture 2 with reference	Picture 3 with reference	Picture 4 with reference	Picture 5 with reference
	1	3	2	2	1	2	3	4	1
			4	3					
2	4	1	3	1	4	5	2	3	2
2	4	3	2	1	3	5	3	2	2
3	3	4	2	1	3	4	3	2	1
2		2	2	1	4		4	4	3
3	4	2		2	4	5	4		3
2	4	3	2	2	4	5	4	3	3
4	4	3	3	3	3	4	3	2	2
3	4	4	2	2	4	4	3	2	2
5	5	5	3	3	4	4	3	3	3
2	4	3	3	2	3	5	3	4	3
1	3	2	3	1	3	4	3	4	2
2	4	2	2	1	3	5	3	4	2
2	3	2	2	1	4	5	3	3	2
3	3	2	2	1	3	4	3	4	3
3	4	2	2	1	4	5	5	3	2
2	2	2	3	1	4	5	3	3	2
2	3	2	2	1	3	4	3	3	2
2	3	2	2	2	3	5	3	4	3
3	3	2	2	2	3	4	3	4	3
1	4	2	2	1	3	5	3	3	3
3	3	2	2	1	4	4	3	3	2
2,47619	3,42857	2,5	2,36364	1,56522	3,36364	4,42857	3,18182	3,19048	2,31818

The table above shows the single results for each individual image. The bottom row shows the mean score for each picture.

Table 7 Explanation of values

Number	Explanation
1	Impairment is not noticeable
2	Impairment is just noticeable
3	Impairment is noticeable, but it is not objectionable
4	Impairment is objectionable
5	Impairment is extremely objectionable

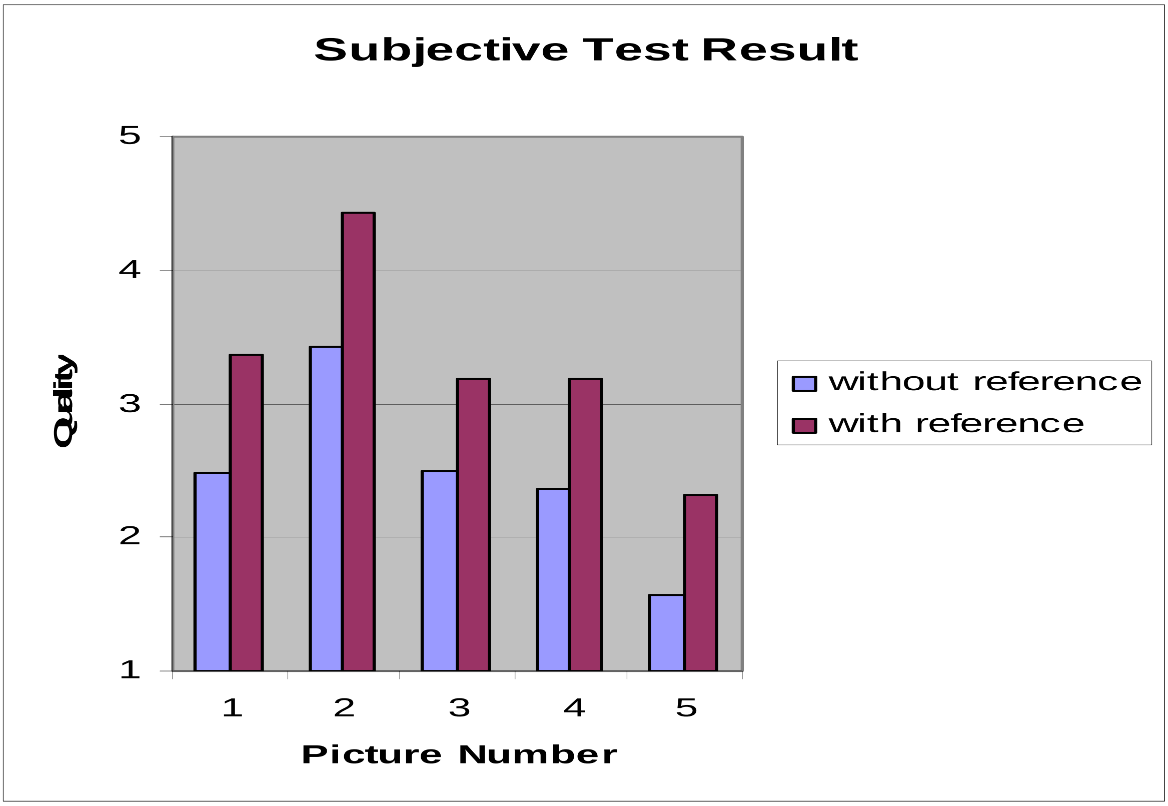


Figure 20 Illustration of subjective test results

It is important to notice that lower values means higher quality, when reading the illustration. As the reader will notice, the score is higher for the pictures when they are shown without a reference picture. This is just what we expected when the test was planned. It is of course easier to notice impairments in a picture when you have a reference.

6.5.2 The test pictures

The pictures used in the test are picked for their varying scenes, from high detailed, to smooth with a sudden transition from dark to white. For instance test picture five has a smooth black background with a sudden transition to the lighter colors of the flower. Wavelets are supposed to be a very good tool to detect edges, and this picture would be perfect for that purpose. Picture one on the other hand has a very high amount of fine details. The picture would therefore be good for testing if the software were able to distinguish the small variations in the picture. Picture four has the blue sky, with some variations in color, and is suitable to find out if the software is splitting the image into small blocks. If that is the case, block effects could be a possible result. Picture number two has details, smooth surfaces and a transition from light to dark color. Finally picture number three has much of the same qualities as picture number one.

6.5.3 Objective results and parameters

This subchapter deals with the objective results we have been able to get using the software from LuraTech. The original and compressed pictures are compared, using LuraWave. The following parameters are listed:

- MAE – Maximum Absolute Error
- MSE – Mean Squared Error
- RMS – Root Mean Squared error
- SNR – Signal to Noise Ratio
- PSNR – Peak Signal to Noise Ratio

Two of the error metrics used to compare the compressed images with original pictures is the Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR). The MSE is the cumulative squared error between the compressed and the original image, and PSNR is a measure of the peak error. The mathematical formulas for the two are:

MSE Mean Squared Error $$\frac{1}{N \cdot M}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left[f(x_i, y_j) - \tilde{f}(x_i, y_j)\right]^2$$

PSNR Peak Signal to Noise Ratio $$20\log\left(\frac{255}{RMS}\right)$$

Where f(x,y) is the original image, $\tilde{f}$(x,y) is the approximated version (which is actually the decompressed image) and M,N are the dimensions of the images. A low value for MSE means lesser error, and as seen from the inverse relation between the MSE and PSNR, this translates to a high value of PSNR. Logically, a higher value of PSNR is good because it means that the ratio of Signal to Noise is higher. Here, the 'signal' is the original image, and the 'noise' is the error in reconstruction. A compression scheme having a lower MSE (and a high PSNR), is a better scheme than one having a high MSE and a low PSNR.

6.5.4 Measurements

On the next pages, we have the collection of pictures used in our test. The test pictures and the measurements are combined in one image for easier viewing of the results. The results are given in the table here:

Table 8 Measurements using LuraWave

Picture num:	1	2	3	4	5
MAE	11.1068	14.2515	7.88418	7.1828	10.4683
RMS	14.5506	19.9976	10.5007	10.0942	14.8048
MSE	211.72	399.905	110.265	101.893	219.183
SNR	16.7252	17.2473	21.8546	20.9593	14.5929
PSNR	24.8732	22.1112	27.7064	28.0494	24.7227

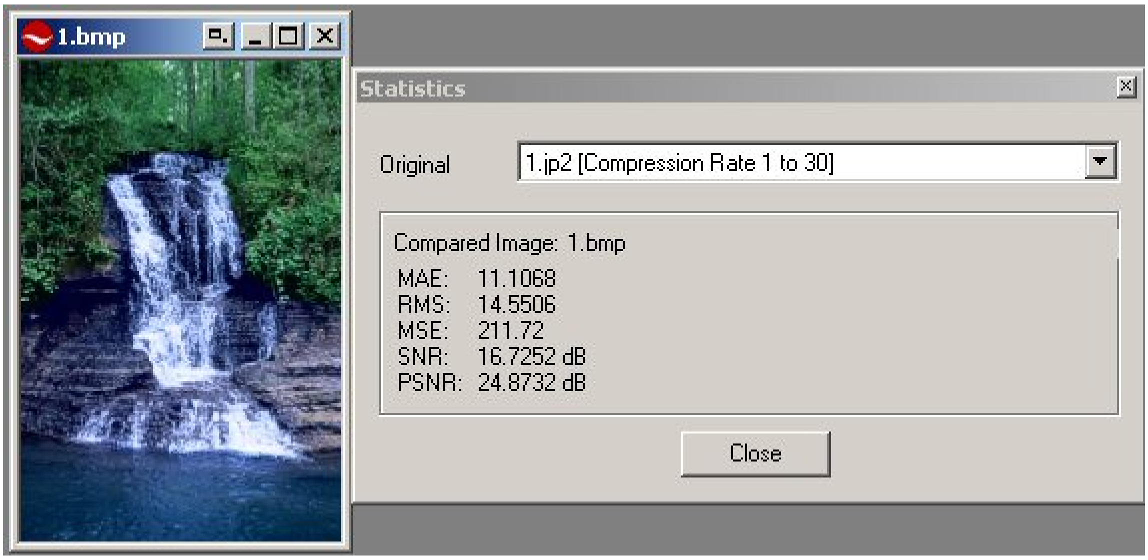


Figure 21 Test picture 1

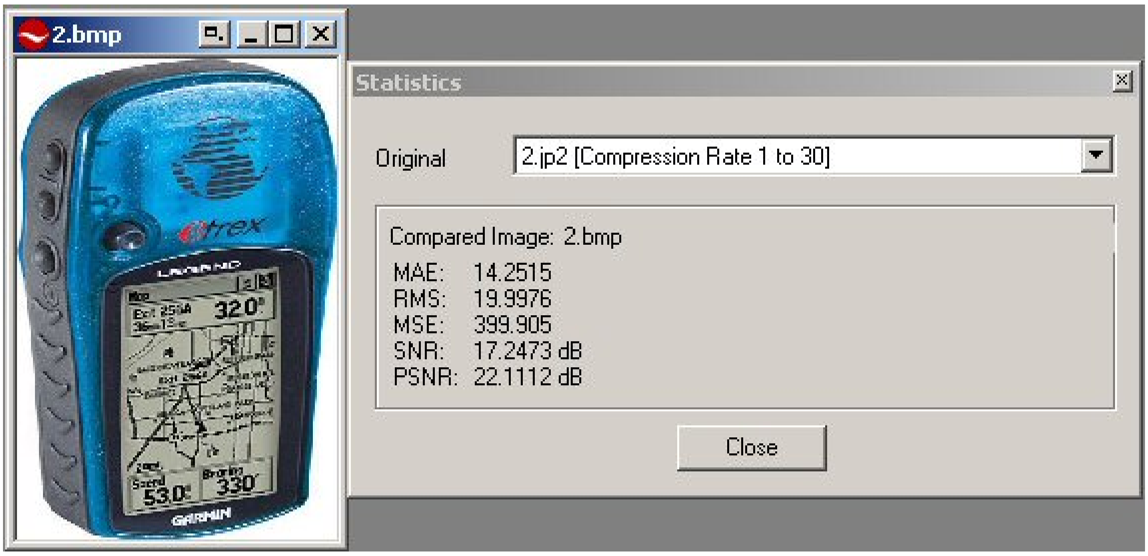


Figure 22 Test picture 2

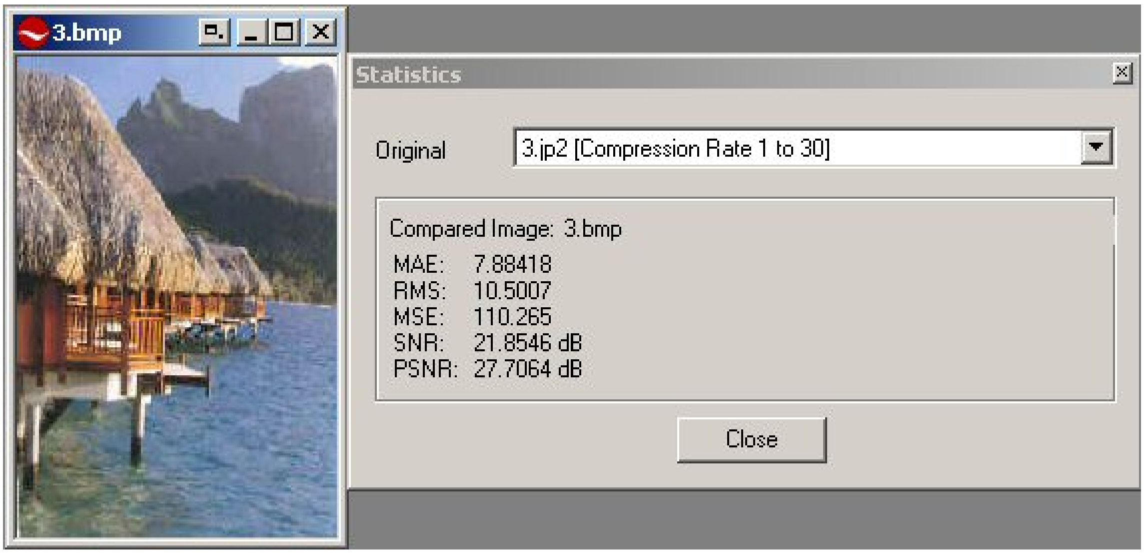


Figure 23 Test picture 3

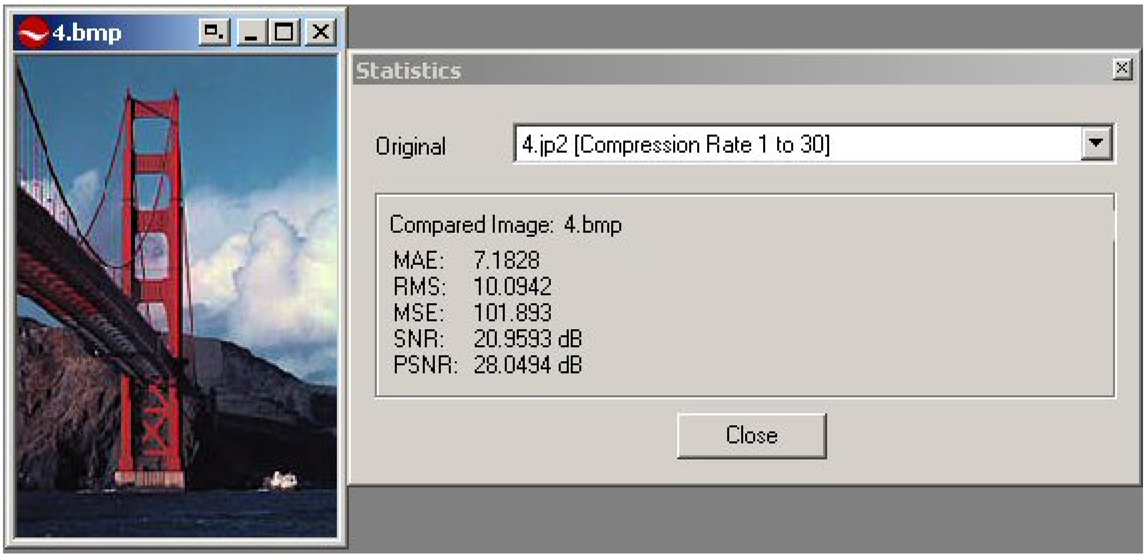


Figure 24 Test picture 4

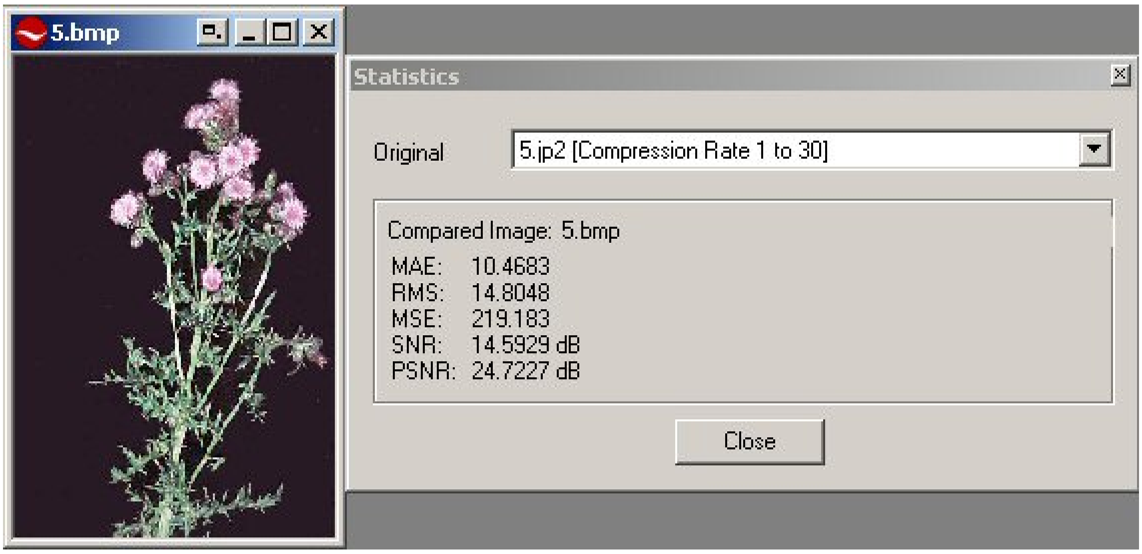


Figure 25 Test picture 5

More on the results can be found in the discussion chapter.

6.6 Comparison of JPEG 2000compression software

This chapter deals with different JPEG 2000 compression tools. We have tested three different software packages; Paint Shop Pro 8, Photoshop 7 and LuraWave. To use Photoshop 7 to compress JPEG 2000 images, you need a plug-in. We have used a plug-in called “fnord” that can be found here. [26]

6.6.1 Pictures

Below is a listing of the different compression tools we have tested. The same input pictures have been used for all of the programs.

BMP LuraWave Paint Shop Pro 8 Photoshop 7 plug-in

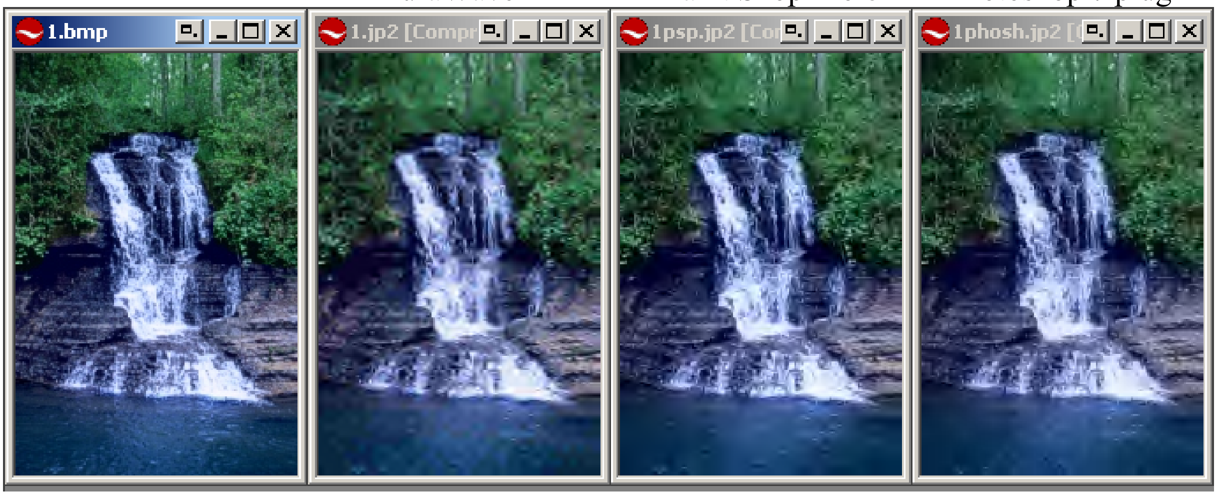


Figure 26 Series 1

Figure 27 Series 2

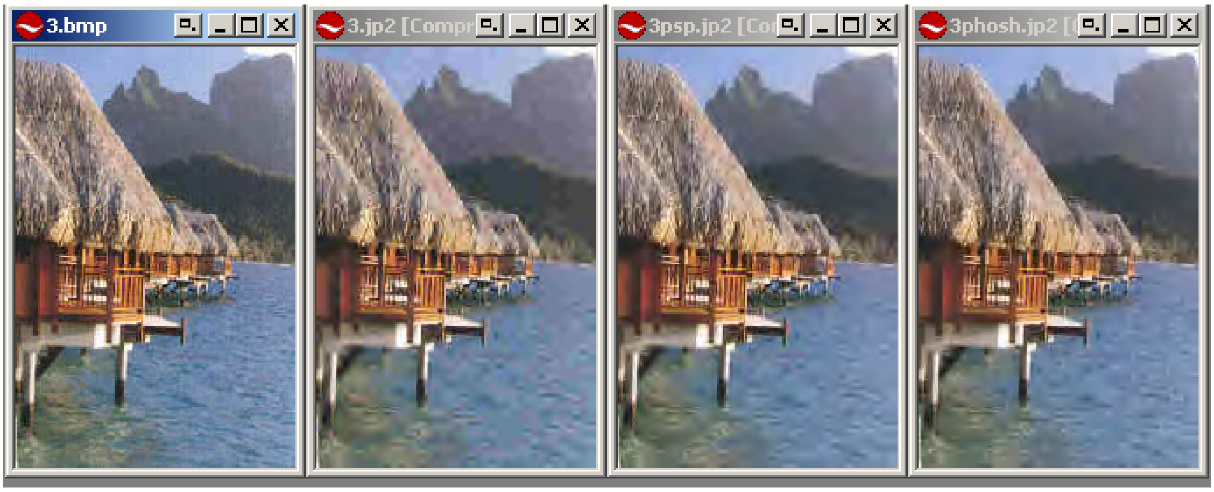


Figure 28 Series 3

Figure 29 Series 4

Figure 30 Series 5

The compressed pictures are clearly of different quality, there is no question about that. It's obvious even when we just consider them subjectively. They do not compress the pictures equally.

6.6.2 Software differences

There is clearly a difference in quality between pictures compressed with different software, even when the pictures are compressed by exactly the same ratio. We have made the same objective measurement with the pictures compressed with Paint Shop Pro and Photoshop as we did with the pictures compressed using LuraWave. The results are listed in the table below.

Table 9 Results using Paint Shop Pro 8

Picture num:	1	2	3	4	5
MAE	9.47735	12.6538	6.93635	6.17382	7.86786
RMS	12.407	18.1333	9.3718	8.87888	13.5457
MSE	153.934	328.815	87.8306	78.8345	183.486
SNR	18.175	18.2134	22.8891	22.114	15.4919
PSNR	26.2575	22.9613	28.6943	29.1636	25.4948

Table 10 Results using Photoshop 7 with plug-in

Picture num:	1	2	3	4	5
MAE	9.82216	12.8233	6.89531	6.23605	8.1673
RMS	12.8178	18.204	9.3683	8.95406	13.8819
MSE	164.297	331.385	87.765	80.1753	192.708
SNR	17.7911	18.1796	22.8713	22.0408	15.0818
PSNR	25.9745	22.9275	28.6976	29.0904	25.2818

Table 11 Results using LuraWave (same table as earlier)

Picture num:	1	2	3	4	5
MAE	11.1068	14.2515	7.88418	7.1828	10.4683
RMS	14.5506	19.9976	10.5007	10.0942	14.8048
MSE	211.72	399.905	110.265	101.893	219.183
SNR	16.7252	17.2473	21.8546	20.9593	14.5929
PSNR	24.8732	22.1112	27.7064	28.0494	24.7227

We have not been able to find out how the different programs work i.e. how they exactly compress their pictures. The information included with the programs is very sparse, but they certainly do not compress pictures equally. More on this in the discussion chapter.

7 Discussion

7.1 Introduction

In this chapter, we try to evaluate the results, and the work that we have been doing in this thesis. This chapter is divided into three parts; channel, compression, quality measurement and different software solutions.

7.2 Channel

As described earlier the channel simulation part of the thesis was cancelled after two months of work. Because of this, our thesis had to be more concentrated on the other parts in the thesis description.

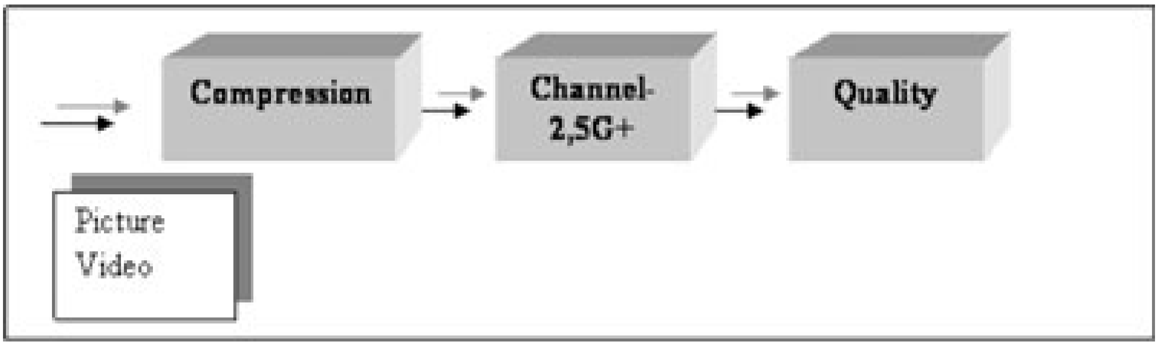


Figure 31 Blocks in the thesis

The fact that the channel simulation part of the thesis was cancelled was really a disappointment for us, since much of the thesis description was built around this. The lack of available channel simulation software was something our supervisor had taken into account when the description was made. We did not though; expect that it would be as difficult as it turned out to be.

After being in contact with Nokia, Sony-Ericsson and Motorola, the results we were left with were some links to open source programs that could be used for network simulation. All the software that the companies had, were considered company-secrets.

The problem with the open source programs is that they are not designed for simulating traffic through a network, but simulation of the whole network. In addition to the big companies, we approached others but none of them had any software that would suit our needs. We also contacted Professor Matthias Pätzold with an enquiry about suitable software. Professor Pätzold pointed out that only a few real-time channel simulators were available, and these were hardware based. He also pointed out that we could try to implement a channel simulator ourselves, on a signal processor, but this would be a highly time consuming task, witch he felt would be too big in addition to our other work.

When these alternatives had been tried, we thought we should try to develop a software model using Simulink from MathWorks [12]. The Simulink solution would also turn out to be very comprehensive task. MathWorks had some demos of WCDMA, but even this small part of a system were enormous, and we soon found out that to expand their model would be way to much work for this thesis.

One could of course argue that we should have gone for a solution either with a signal processor, or with Simulink, but we felt that the most important part of the thesis would have to be the compression and quality parts of the system. If we continued the work on a channel simulator, we would run the risk of sitting in June with no results at all. Our supervisor Ola T. Aas therefore recommended in a meeting that we terminated the part of the thesis that aimed at a working channel simulator, and concentrated on the compression and quality tasks.

7.3 Compression

Since the JPEG 2000 format is the only real well known format for wavelet compressed images, we decided soon in the thesis process that this was our only real choice as file format. The standard is well documented, and several articles have been published about the format. Finding suitable compression software for JPEG 2000 compression has not been as easy a task as one should think. There are really a lot of available software, but most of it is commercial i.e. you have to pay for it. We did find a number of software solutions, and ended up choosing LuraWave from Luratech [9]. LuraWave was used for compression of the pictures that we published on our test pages on the web.

7.3.1 LuraWave

LuraWave from Luratech [9] is a practical program for compression of pictures to JP2. We found several programs, but decided to go for LuraWave since it also gives you useful information about the quality of compressed pictures. LuraWave lets you compare the compressed picture with the original one and computes the following properties: MAE, MSE, RMS, SNR and PSNR. More on these properties can be found later in this chapter.

7.3.2 Other JP2 compression tools

Wavelet toolbox from MathWorks is a toolbox for MatLab. We have tried this, but found it to be a little too troublesome for our use.

LuraTech has a LuraWave plug-in for Photoshop and this plug-in cost about 80$. We have not tested this plug-in.

IWcompressor [17] is a shareware program that can be downloaded for 19$. Not tested

There are also several other programs that can be used for JPEG 2000 compression, and most of them are shareware or software you have to pay for. We have not been able to find any article that describes a thorough test of this type of software, so that could probably be an interesting approach for further work.

7.4 Quality measurement

We have two quality measurements, objective and subjective. The way that the measurements are carried out follows well-known patterns. The principles of objective and subjective quality measurements are described in the image quality chapter.

7.4.1 Subjective test results

The test persons submit the subjective quality measurements through a web solution. The resulting grades are, as described in the chapter about results, saved to a text file, and the mean values of the test are plotted on a bar graph. The bar graph from the subjective test is shown below.

From the bar graph you can tell that the results are quite what we expected. When a user judges a compressed picture without a reference picture, it is graded higher than when a reference picture is present.

This table shows the mean values of the results in the subjective test, as plotted on the bar graph.

Picture 1	Picture 2	Picture 3	Picture 4	Picture 5	Picture 1 with reference	Picture 2 with reference	Picture 3 with reference	Picture 4 with reference	Picture 5 with reference
2,47619	3,42857	2,5	2,36364	1,56522	3,36364	4,42857	3,18182	3,19048	2,31818

This table shows the difference in mean value between the picture with, and without the reference picture.

Picture number	1	2	3	4	5
Difference in mean value	0,88745	1,00000	0,81818	0,82684	0,75296

The table shows that picture number 2 has the biggest difference between the test picture and the test picture with a reference picture, while picture 5 has the least difference. This indicates from the test subject's point of view, that they rate picture 5 as the best, and picture 2 as the poorest of the five test pictures.

7.4.2 Objective test results

The objective measurements have been performed with LuraWave. The program calculates the following parameters:

MAE	Maximum Absolute Error	$\max\left\|f(x,y)-\tilde{f}(x,y)\right\|$
MSE	Mean Squared Error	$\frac{1}{N\cdot M}\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left[f(x_i,y_j)-\tilde{f}(x_i,y_j)\right]^2$
RMS	Root Mean Squared Error	$\sqrt{MSE}$
SNR	Signal to Noise Ratio	$10\log\left(\frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}f(x_i,y_j)^2}{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}\left[f(x_i,y_j)-\tilde{f}(x_i,y_j)\right]^2}\right)$
PSNR	Peak Signal to Noise Ratio	$20\log\left(\frac{255}{RMS}\right)$

The results from the objective measurements are listed in the table below.

Table 12 Results using LuraWave (same table as earlier)

Picture num:	1	2	3	4	5
MAE	11.1068	14.2515	7.88418	7.1828	10.4683
RMS	14.5506	19.9976	10.5007	10.0942	14.8048
MSE	211.72	399.905	110.265	101.893	219.183
SNR	16.7252	17.2473	21.8546	20.9593	14.5929
PSNR	24.8732	22.1112	27.7064	28.0494	24.7227

The table shows that picture 5 has the clearly lowest Signal to Noise Ratio. It does not however have the highest Mean Squared Error.

From the table it is interesting to notice that picture 4 have the best results, and not picture 5 as in the subjective test. This indicates that the HVS (the subjective test) does not, as described earlier, perceive the same factors as an objective test does.

We could of course have used a greater number of pictures for the test, and the scenes could possibly have varied to a greater extent. The pictures we have chosen do though work well for the test, and we have been getting some clear and proper results.

7.5 Different software solutions

In addition to LuraWave, we wanted to test other JP2 compression software solutions. The choice of additional software fell on Paint Shop Pro 8 and Photoshop 7 with a JP2 plug-in. The picture that was compressed for the test pages was also used with these programs. LuraWave was used on all compressed pictures to objectively measure the quality.

7.5.1 Performance of the different software solutions

The original BMP pictures have been compressed with Paint Shop Pro, Photoshop and LuraWave. In LuraWave, we have made comparison between the original pictures and the compressed pictures from the different platforms.

The results clearly show that there are differences in the quality of the compressed pictures. We have listed the same results below as in the image quality chapter for easier reading

Table 13 Results using Paint Shop Pro 8

Picture num:	1	2	3	4	5
MAE	9.47735	12.6538	6.93635	6.17382	7.86786
RMS	12.407	18.1333	9.3718	8.87888	13.5457
MSE	153.934	328.815	87.8306	78.8345	183.486
SNR	18.175	18.2134	22.8891	22.114	15.4919
PSNR	26.2575	22.9613	28.6943	29.1636	25.4948

Table 14 Results using Photoshop 7 with plug-in

Picture num:	1	2	3	4	5
MAE	9.82216	12.8233	6.89531	6.23605	8.1673
RMS	12.8178	18.204	9.3683	8.95406	13.8819
MSE	164.297	331.385	87.765	80.1753	192.708
SNR	17.7911	18.1796	22.8713	22.0408	15.0818
PSNR	25.9745	22.9275	28.6976	29.0904	25.2818

Table 15 Results using LuraWave (same table as earlier)

Picture num:	1	2	3	4	5
MAE	11.1068	14.2515	7.88418	7.1828	10.4683
RMS	14.5506	19.9976	10.5007	10.0942	14.8048
MSE	211.72	399.905	110.265	101.893	219.183
SNR	16.7252	17.2473	21.8546	20.9593	14.5929
PSNR	24.8732	22.1112	27.7064	28.0494	24.7227

As you can see, the MSE and SNR are better in the imaging programs, than it is in LuraWave. The two imaging programs perform almost the same, Paint Shop Pro perform slightly better in the measurements while LuraWave's performance is considerably lower.

These pictures show the different compression tools in use. From left to right we have:
BMP LuraWave Paint Shop Pro 8 Photoshop Plug-in

When you look at the pictures subjectively, there are clearly quality differences between them. The LuraWave picture appears to have tiling, and the picture is partially blurred. The other compressed pictures appear to have a considerably higher quality. There is clearly some smearing in parts of the pictures; but in general, they have a higher subjective quality. It is very difficult to distinguish between the pictures of Paint Shop Pro 8 and Photoshop 7. None of them is clearly better than the other, when we try to subjectively separate them.

These assumptions are confirmed when we compare this with the objective measurements. The performances of LuraWave are considerably lower than the others. There is no information on how any of the programs do the actual compression, but in the JPEG 2000 specification it is stated that the developers are free to divide the pictures into blocks as they see fit. One would suspect that LuraWave uses smaller blocks than the others do, hence the clear tiling on the LuraWave pictures.

This been said it is highly scene dependant how the programs perform. In the next example, LuraWave performs slightly lower on the objective measurements, but subjectively it appears to be the best of the three software solutions.

Here one can clearly see that Paint Shop Pro 8 and Photoshop 7 plug-in have had problems with the sudden transition from the black background to the lighter colored parts of the flower.
The same thing is visible in the LuraWave picture, but it seems that LuraWave has been able to cope with the transition more gracefully.

8 Conclusion

8.1 Introduction

In this chapter we try to draw some conclusions from the work that has been done in this thesis. First and foremost we try to draw conclusions from the results we have been getting from our objective and subjective quality tests, but we also try to make some points related to the thesis work in general.

8.1.1 UMTS channel

As mentioned earlier, the UMTS channel part of the thesis was terminated, due to the lack of available suitable software. This turned out to be crucial for the project, and we had to change our focus somewhat from a complete system to the endpoints of the system i.e. compression and quality measurement. All in all we used a lot of time looking for usable software, time that could have been spent on the other parts of the project.

8.1.2 Compression and tools

We have concentrated mainly on JPEG 2000, because it is well documented and wavelet based. We have also compared our preferred software, LuraWave with two alternatives, Paint Shop Pro 8 and Photoshop 7 with a JPEG 2000 plug-in.

When we compared the pictures that were compressed with the different programs, it became clear that there were considerable differences in the quality between them. This leads to the conclusion that they do not compress the pictures in the same way. Although the JPEG 2000 standard describes clearly witch wavelets that should be used, it does not say anything about how a picture should be divided before compression. Especially on the LuraWave pictures there are clearly block effects, and this is probably the result of the way the picture is split up before the compression process.

It’s clear that wavelet compression, in this case JPEG 2000 is much more effective than JPEG. At the same compression rate a JPEG 2000 image will most often have a very much higher quality than a JPEG image.

8.1.3 Tests and results

The testing methods we have used, seems to have worked in a satisfactory way. The subjective test method we have used has been used frequently earlier by several researchers and is well described. The objective test method is a lot easier, since the calculations have all been done with LuraWave. These results have not been compared to other objective measurements done by other types of software. This would of course have made our calculations more foul proof, and this would have been a clear advantage. Our subjective and objective measurements have given us some great results to work with. The results have shown us that there is a need for software who can emulate the HVS in a satisfactory way. This would be a great step forward, and such technology would make the testing a whole lot easier. A lot of effort is being put into this research field, but no satisfactory solution has yet been developed.

The test results showed us that there is clearly a difference between objective and subjective measurements. A human test subject emphasizes different parameters than a test program does. In our test, the picture that had the best objective results did not score highest subjectively. One interesting discovery we made while comparing the three software solutions we tested, was that they do not compress the pictures in the same way. The results from the three programs are clearly different. The quality of the compressed pictures varies greatly, even though there are compressed to the same format and with the same rate. For instance one of the pictures clearly showed tiling. The same picture did not show any signs of tiling when compressed by one of the other programs. This is very interesting, and proves that the programs compress JPEG2000 pictures differently.

8.1.4 The connection between bandwidth, quality and compression in UMTS

There is obviously a clear connection between the bandwidth of a channel and the need for compression of pictures and video that is to be transmitted over this channel. The channel capacity basically decides the need of compression. The quality is very much decided by the rate of compression, so it all comes down to bandwidth in the end. All in all we think that it is possible to get a satisfactory quality in a video transmission with a bit rate of 384kbps. The possibility of lowering frame rate to for example 24 frames pr. second, and code the pixels at a lower rate (for example 0,5 bpp) makes it possible to stream a video sequence with an acceptable quality. For bit rates lower than this though, an adequate quality can not be expected, at least with today’s compression techniques and algorithms.

8.1.5 Further work

One possible way to take this work further might be to try and get a complete system up and running. It could be very interesting to have a fully working channel simulator to send the data through. It could also be very interesting to do a broad comparison of available JPEG2000 compression tools, and find out more exactly what differences they have.

8.1.6 General reflections

Generally this project period has been very instructive. None of us had any special experience with compression or wavelets before we started on the thesis, so it has been a long learning process for the both of us.

9 Bibliography

[1] Image and video compression for multimedia engineering, fundamentals algorithms and standards by Yun Q. Shi and Huifang Sun. ISBN: 0-8493-3491-8
[2] http://164.214.2.51/ntb/2002SICS/Comp_Tutorial_Apr02_pt2.PDF
[3] http://www.acm.org/crossroads/xrds6-3/sahaimgcoding.html
[4] http://ise.stanford.edu/class/psych221/projects/98/dctune/yuke/page2.htm (28.01.04)
[5] http://www.etek.chalmers.se/~e4joni/xjobb/report.pdf
[6] http://www.codesta.com/knowledge/technical/digital_photography/page_01.jsp#color_perception (24.02.04)
[7] http://www.eecs.wsu.edu/~cs445/ (22.02.04)
[8] http://www.mic-d.com/java/digitalimaging/spatialresolution/ (16.04.04) spatial resolution picture
[9] http://www.luratech.com/
[10] http://pcl.cs.ucla.edu/projects/glomosim/
[11] http://www.isi.edu/nsnam/ns/
[12] http://www.mathworks.com/
[13] http://www.jp2tools.de/English/index.htm
[14] http://www.elysium.ltd.uk/JPEG 2000.html
[15] http://www.si.umich.edu/Classes/540/Slides/4Compression.ppt (28.04)
[16] http://fag.grm.hia.no/olata/dat2430eng/
[17] http://www.barrt.ru/parshukov/
[18] IEEE Transactions on Consumer Electronics, Vol. 46, No. 4, pp. 1103-1127, November 2000 THE JPEG 2000 STILL IMAGE CODING SYSTEM:
AN OVERVIEW
[19] http://www.public.iastate.edu/~rpolikar/WAVELETS/WTtutorial.html (09.02.04) (can now be found at http://users.rowan.edu/~polikar/WAVELETS/WTpart1.html)
[20] Image Processing Handbook 3rd edition John C. Russ s140 ISBN 3-540-64747-3
[21] http://www.mathworks.com/access/helpdesk/help/pdf_doc/wavelet/wavelet_ug.pdf (27.01.04)
[22] http://www.nhk.or.jp/strl/publica/bt/en/le0009-1.html (17.02.04)
[23]WDCMA for UMTS-Radio access for third generation mobile communications-Second Edition, H. Holma, A. Toskala, ISBN 0-470-84467-1
[24] http://www.d.kth.se/~d98-fah/hacklin/thesis/latex2html/output/node30.html
[25] The JPEG 2000 Still-Image Compression Standard. Majid Rabbani, Diego Santa Cruz jj2000 (http://jj2000.epfl.ch/jj_publications/papers/011.pdf) (5.3.04)
[26] www.fnordware.com
[27] http://fag.grm.hia.no/fagstoff/perhh/htm/fag/matem/datwwww/wavelet.htm
[28] http://jj2000.epfl.ch/jj_tutorials/ (14.5.04)

10 Acronyms and Abbreviations

2.5G - Generation 2.5 .... VII, 2, 18
2G - 2nd Generation 2, 18, 19
3G - 3rd Generation .... I, VII, 2, 18, 26, 31, I
3GPP - 3rd Generation Partnership Project 1, 18
3rd generation - 3G I
CAT - Computer-Aided Tomography 11
DCT - Discrete Cosine Transform 1, 6, 7, 23
DFT - Discrete Fourier Transform 6
DWT - Discrete Wavelet Transform 6, 10, 14
FDD - Frequency Division Duplex 19, 20
FDMA - Frequency Division Multiple Access 19
FFT - Fast Fourier Transform 8
GPRS - General Packet Radio Service 18
GSM - Global System for Mobile communication 18
HPF - High Pass Filter 8
HVS - Human Visual System 2, 26, 27, 29, 30, 48, 51
IEEE - Institute of Electrical & Electronics Engineers 13
ISO - International Organization for Standardization 1, 7
ITU-R International Telecommunication Union - Recommendations Sector .... 2, 25, VII
ITU-T - International Telecommunication Union - Telecommunication Stadardization Sector 1, 7
JPEG - Joint Photographic Expert GroupI, 1, 3, 6, 7, 11, 12, 13, 14, 16, 17, 18, 32, 33, 40, 41, 45, 49, 51
JPEG 2000 - a wavelet based compression standardI, 1, 3, 6, 11, 12, 13, 14, 16, 17, 18, 32, 33, 40, 41, 45, 50, 51
LPF - Low Pass Filter 8
MPEG - Moving Picture Experts Group 6, 7, 13
RLE - Run-Length Encoding 7
ROI - Region Of Interest 1, 11, 12, 13, 17
SIR - Signal - Interference Ratio 19, 20
STFT - Short Term Fourier Transform 8
TDD - Time Divison Duplex 19, 20
TDMA - Time Division Multiple Access 19
UE - User Equipment 2, 20
UMTS - Universal Mobile Telecommunication System .... I, 1, 2, 3, 18, 19, 20, 26, 50, 53, I
VLC - Variable Length Coding 7
wavelet - from small wave. .... I, VII, 1, 6, 7, 8, 9, 10, 11, 14, 15, 32, 45
WCDMA - Wideband Code Division Multiple Access 2, 19, 44

A. Appendix

Master of Science – Information and Communication technology – AUC

Master thesis

Title:
Connection between the need for compression, bandwidth and quality for multimedia transmissions in UMTS

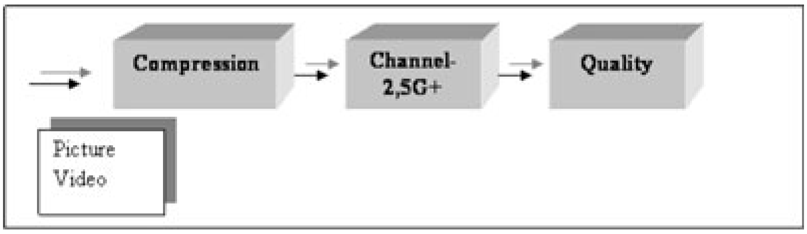
Students:
Helge Ivar Lislevand hilisl99@siving.hia.no
Håkon Oftedal Njaa honjaa02@siving.hia.no

Supervisor:
Ola T Aas Ola2.Aas@uwe.ac.uk

Project description:
Next generation mobile systems, from 3G and beyond, will have a much higher channel capacity and the terminals will be more adapted for multimedia than today's 2,5G system. Transmission of pictures and video will demand efficient compression methods. There is a close connection between the channel capacity (bandwidth) and the need for compression. For a user the quality is of the outmost importance. We believe that it could be interesting to study the connection between the three factors: need for compression, bandwidth and quality for a UMTS system.

The thesis will have the following structure:

- A theory description of compression methods and important compression parameters, with a special focus on wavelets
- Software search to find adequate software for a simulator construction and testing.
- Tests and measurements on simulator and possibly actual equipment.
- Discussion and evaluation of the tests with a conclusion.



This figure shows the system. The input will be a picture and/or a video sequence. By varying the parameters in the compression process and the transmission channel, one will be able to test information with different quality. The development of a model for a complete system is an extensive and time-consuming task. Development of simulation methods adapted for the measurement of interest is an important part of the thesis.
The lack of commercially available equipment for 3G systems makes it important to develop a model for a system based on available software. This could become a critical part of the thesis, because the availability and types of software is unknown. Quality measurements of

digital pictures, and if time allows it, video, will be a decisive part of the thesis. Perception of quality is subjective and demands special measuring methods. We do however believe that we can make objective and subjective tests of a system that can give us interesting information of characteristics and expectations of next generation of mobile systems.

B. JPEG 2000 filters

For lossy compression, Daubechies 9-tap/7-tap filter is used.

Table 16 Daubechies 9/7 analysis filter coefficients [18]

Analysis Filter Coefficients		
i	Low-pass Filter hL(i)	High-pass Filter hH(i)
0	0.6029490182363579	1.115087052456994
±1	0.2668641184428723	-0.5912717631142470
±2	-0.07822326652898785	-0.05754352622849957
±3	-0.01686411844287495	0.09127176311424948
±4	0.02674875741080976	

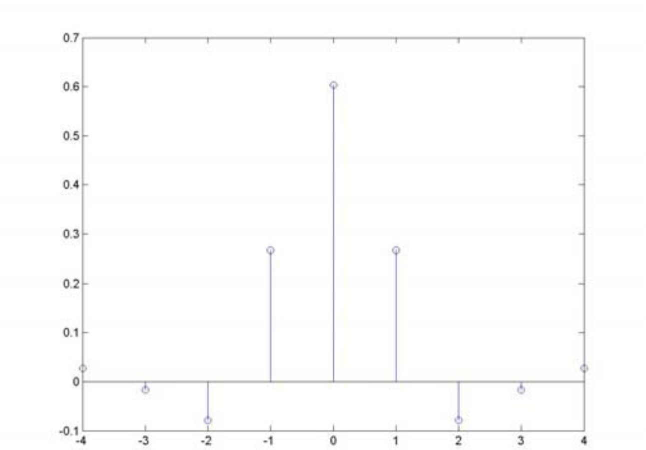


Figure 32 Daubechies 9-7 Low Pass Analysis Filter

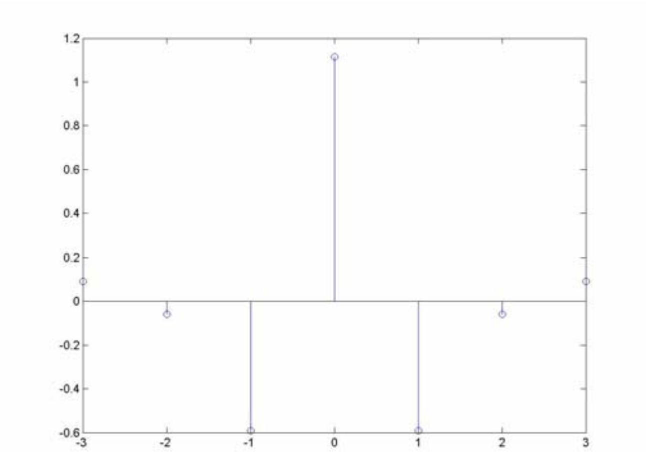


Figure 33 Daubechies 9-7 High Pass Analysis Filter

Table 17 Daubechies 9/7 synthesis filter coefficients [18]

Synthesis Filter Coefficients		
I	Low-pass Filter gL(i)	High-pass Filter gH(i)

0	1.115087052456994	0.6029490182363579
±1	0.5912717631142470	-0.2668641184428723
±2	-0.05754352622849957	-0.07822326652898785
±3	-0.09127176311424948	0.01686411844287495
±4		0.02674875741080976

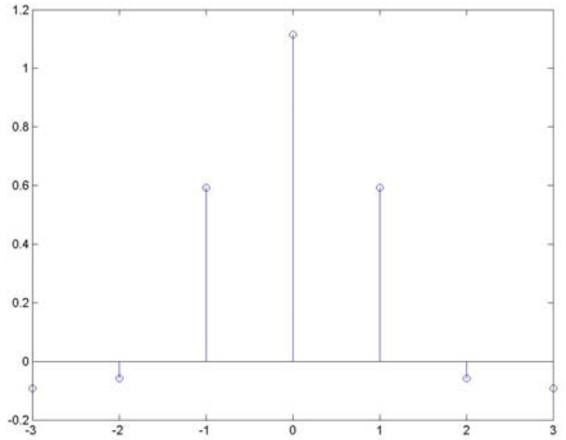


Figure 34 Daubechies 9-7 Low-Pass Synthesis Filter

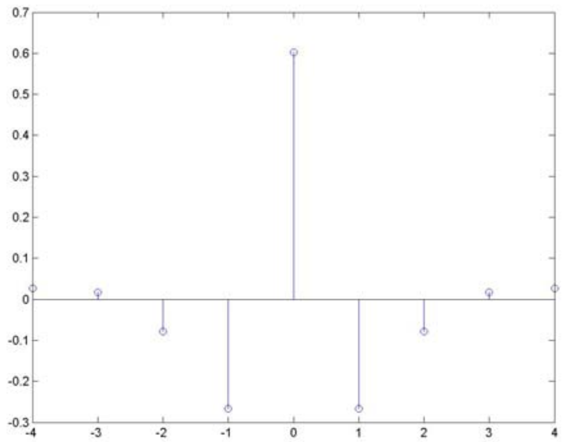


Figure 35 Daubechies 9-7 High Pass Synthesis Filter

Lossless compression is implemented with a 5-tap/7-tap filter

Table 18 5/3 analysis and synthesis filter coefficients [18]

	Analysis Filter Coefficients		Synthesis Filter Coefficients	
i	Low-pass Filter hL(i)	High-pass Filter hH(i)	Low-pass Filter gL(i)	High-pass Filter gH(i)
0	6/8	1	1	6/8
±1	2/8	-1/2	1/2	-2/8
±2	-1/8			-1/8

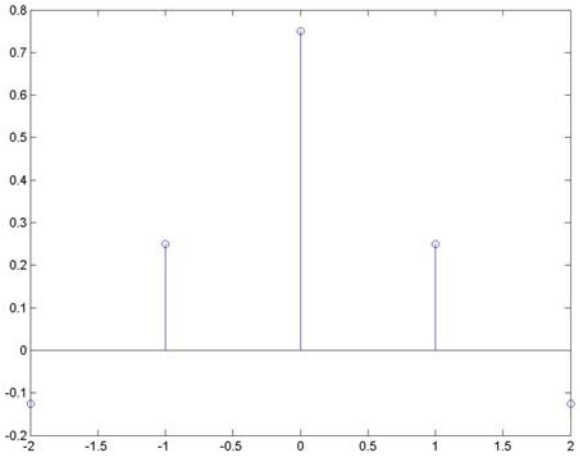


Figure 36 5-3 Low Pass Analysis Filter

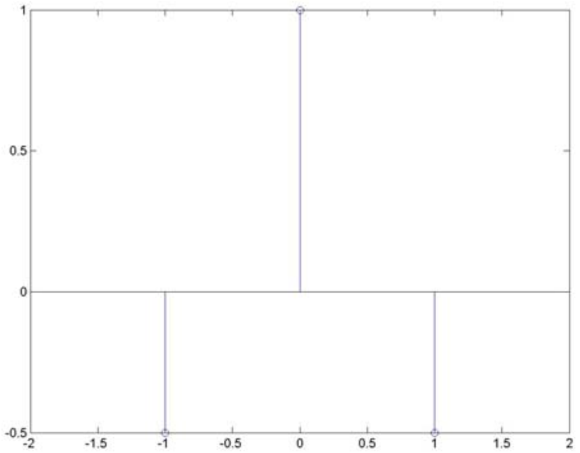


Figure 37 High Pass Analysis Filter

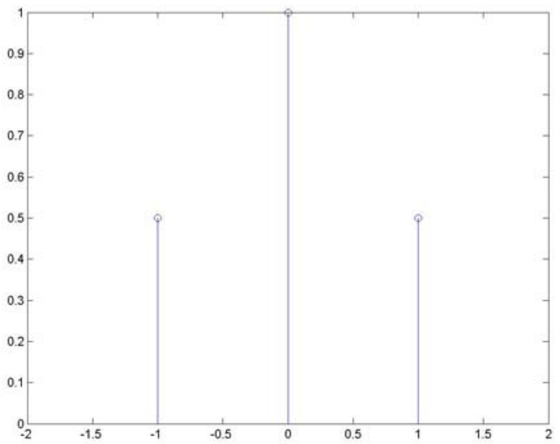


Figure 38 Low Pass Synthesis Filter

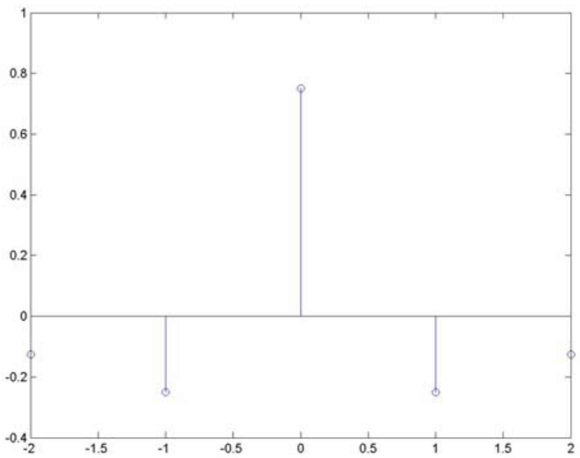
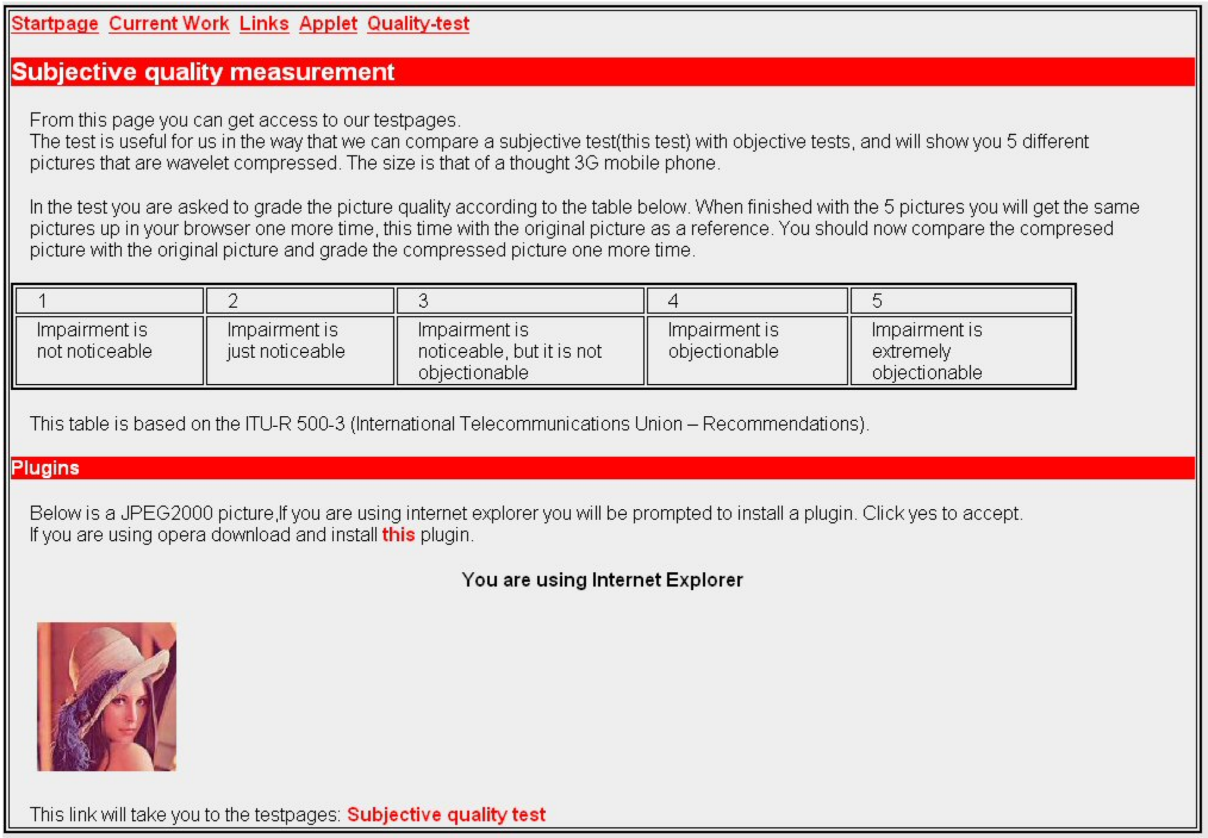


Figure 39 High Pass Synthesis Filter

C. Source code

This is the source code for the WebPages that deals with the subjective test. For easier understanding, a picture of the webpage is included with the corresponding source code.

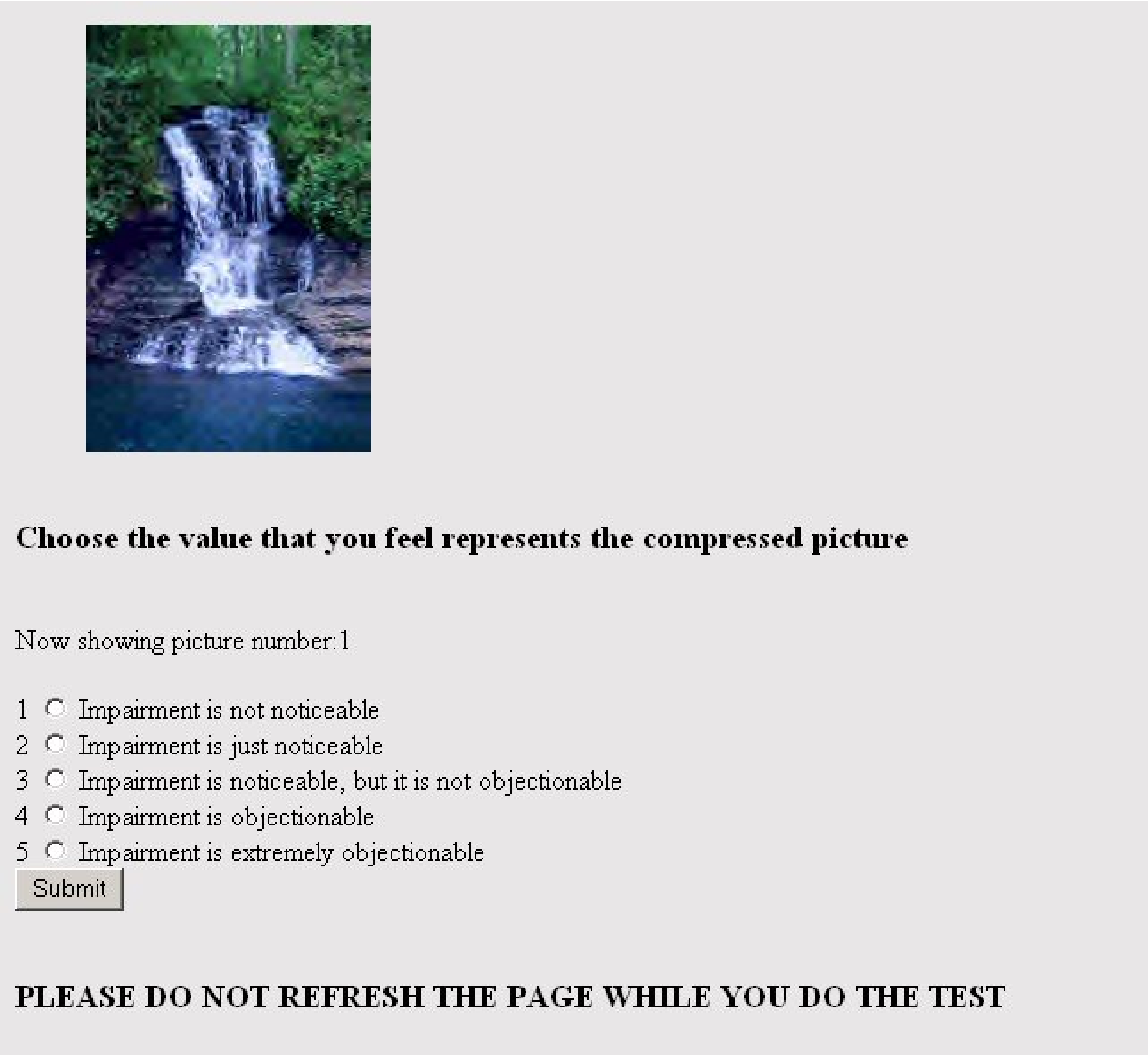


Quality test page

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">

<html>

<head>
<meta name="GENERATOR" content="Microsoft FrontPage 5.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">
<title>Quality measurement</title>
<link rel="stylesheet" type="text/css" href="styles.css">
<style>
<!--
table.MsoTableGrid
	{border:1.0pt solid windowtext;
	font-size:10.0pt;
	font-family:"Times New Roman"}
-->
</style>
</head>

<body>
```

```
<table>
        <tr>
            <td>
                <a class="meny" href= "index.htm">Startpage</a> 
                <a class="meny" href= "current.htm">Current Work</a> 
                <a class="meny" href= "links.htm">Links</a> 
                <a class="meny" href="Applet.html">Applet</a> 
                <a class="meny" href="imagequality.htm">Quality-test</a> 


                <h1>Subjective quality measurement</h1>
                <p>From this page you can get access to our testpages.<br>
                The test is useful for us in the way that we can compare a subjective test(this
                test) with objective tests,
                and will show you 5 different pictures that are wavelet compressed. The size
                is that of a thought 3G mobile phone.</p>
                <p>In the test you are asked to grade the picture quality according to the
                table below. When finished with the 5 pictures you will get the same pictures
                up in your browser one more time, this time with the original picture as a
                reference. You should now compare the compresed
                picture with the original picture and grade the compressed picture one more
                time.
                </p>

                <table>
  <tr>
    <td><p>1</p></td>
    <td><p>2</p></td>
    <td><p>3</p></td>
    <td><p>4</p></td>
    <td><p>5</p></td>
  </tr>
  <tr>
    <td><p>Impairment is not noticeable</p></td>
    <td><p>Impairment is just noticeable</p></td>
    <td><p>Impairment is noticeable, but it is not objectionable</p></td>
    <td><p>Impairment is objectionable</p></td>
    <td><p>Impairment is extremely objectionable</p></td>
  </tr>
 </table>


 <p>This table is based on the ITU-R 500-3 (International Telecommunications Union –
Recommendations).</p>

 <h2>Plugins</h2>
        <p>Below is a JPEG 2000 picture,If you are using internet explorer you will be prompted to
        install a plugin. Click yes to accept.<br>
         If you are using opera download and install <a href= "plugin.rar">this</a> plugin.</p>

  <?php
        if (strpos($_SERVER['HTTP_USER_AGENT'], 'MSIE') !== false) {
        ?>
        <center><b>You are using Internet Explorer</b></center>
        <?php
        } else {
        ?>
        <center><b>You are not using Internet Explorer</b></center>
        <?php
        }
```

```
        ?>

    <p><object classid="clsid:7AAA3403-FE71-11D2-A1A9-D484164D354E"
        CODEBASE="http://www.jp2Tools.de/jp2IE.cab"
        WIDTH="130" HEIGHT="140">
        <param name="SRC" value="images\lena.jp2">
        <param name="Scaling" value="0">
        <param name="ShowToolbar" value="0">
        </object></p>

    <p>This link will take you to the testpages: <a href= "qualtest.php">Subjective quality test</a> 

            </td>

        </tr>
    </table>


</body>

</html>
```



Test page 1

```
<?php session_start(); ?>
<html>

<head>
<meta name="GENERATOR" content="Microsoft FrontPage 5.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">
<title>New Page 1</title>
</head>

<body>
```

```
<?php

	session_register("count");
	++$count;
	if($count==6){

		echo '<script type="text/javascript">
		<!--

		 location.href = "qualtest2.php";
		 //-->
		</script>';

		}
?>
	<p><object classid="clsid:7AAA3403-FE71-11D2-A1A9-D484164D354E"
	CODEBASE="http://www.jp2Tools.de/jp2IE.cab"
	WIDTH="240" HEIGHT="240">
	<param name="SRC" value="images\<?php echo $count; ?>.jp2">
	<param name="Scaling" value="0">
	<param name="ShowToolbar" value="0">
	</object></p>

	<br><H3>Choose the value that you feel represents the compressed picture</H3><br>
	Now showing picture number:<?php echo $count; ?><br>
	<form action=<?php echo $_SERVER['PHP_SELF']?> method="POST">
	<p>
	1 <input type="radio" value="V1"  name="R1"> Impairment is not noticeable<br>
	2 <input type="radio" value="V2"  name="R1"> Impairment is just noticeable<br>
	3 <input type="radio" value="V3"  name="R1"> Impairment is noticeable, but it is not objectionable<br>
	4 <input type="radio" value="V4"  name="R1"> Impairment is objectionable<br>
	5 <input type="radio" value="V5"  name="R1"> Impairment is extremely objectionable<br>
	<input type="submit" value="Submit" name="B1"><br>
	</p>
	</form>
<?php
	if(array_key_exists('B1',$_POST)){
	if(array_key_exists('R1',$_POST)){
	$inn=($_POST['R1']);
	$fp=fopen("test.txt","a");
	$pn=($count-1);
	fwrite($fp,"Picture number:$pn is given the value:$inn\n");
	print"you chose: $inn for picture number:$pn";

	}
	else{echo"Please select a value";
	//reduce();
	}

	}
?>


<br><H3>PLEASE DO NOT REFRESH THE PAGE WHILE YOU DO THE TEST<br></H3>


</body>

</html>
```

Compare the two pictures.The left picture is comressed!
Choose the value that you feel represents the compressed picture

Now showing picture number:1

1 ○ Impairment is not noticeable
2 ○ Impairment is just noticeable
3 ○ Impairment is noticeable, but it is not objectionable
4 ○ Impairment is objectionable
5 ○ Impairment is extremely objectionable

[Submit]

PLEASE DO NOT REFRESH THE PAGE WHILE YOU DO THE TEST

Test page 2

```
<?php session_start(); ?>


<html>

<head>
<meta name="GENERATOR" content="Microsoft FrontPage 5.0">
<meta name="ProgId" content="FrontPage.Editor.Document">
<meta http-equiv="Content-Type" content="text/html; charset=windows-1252">
<title>New Page 1</title>
</head>

<body>
```

```
<?php

            session_register("count1");
            ++$count1;


            if($count1==6){

                        echo '<script type="text/javascript">
                        <!--

                         location.href = "thanks.php";
                         //-->
                        </script>';

                        }
?>
            <object classid="clsid:7AAA3403-FE71-11D2-A1A9-D484164D354E"
            CODEBASE="http://www.jp2Tools.de/jp2IE.cab"
            WIDTH="240" HEIGHT="240">
            <param name="SRC" value="images\<?php echo $count1; ?>.jp2">
            <param name="Scaling" value="0">
            <param name="ShowToolbar" value="0">
            </object>

            <img src="images\<?php echo $count1; ?>.bmp">


            <br><H3>Compare the two pictures.The left picture is comressed!<br> Choose the value that you feel
represents the compressed picture</H3><br>
            Now showing picture number:<?php echo $count1; ?><br>
            <form action=<?php echo $_SERVER['PHP_SELF']?> method="POST">
            <p>
            1 <input type="radio" value="V1"  name="R1"> Impairment is not noticeable<br>
            2 <input type="radio" value="V2"  name="R1"> Impairment is just noticeable<br>
            3 <input type="radio" value="V3"  name="R1"> Impairment is noticeable, but it is not
objectionable<br>
            4 <input type="radio" value="V4"  name="R1"> Impairment is objectionable<br>
            5 <input type="radio" value="V5"  name="R1"> Impairment is extremely objectionable<br>
            <input type="submit" value="Submit" name="B1"><br>
            </p>
            </form>
<?php
            if(array_key_exists('B1',$_POST)){
            if(array_key_exists('R1',$_POST)){
            $inn=($_POST['R1']);
            $fp=fopen("test.txt","a");
            $pn=($count1-1);
            fwrite($fp,"Dual picture number:$pn is given the value:$inn\n");
            print"you chose: $inn for dual picture number:$pn";
            }
            else{echo"Please select a value";
            }
            }
?>
<br><H3>PLEASE DO NOT REFRESH THE PAGE WHILE YOU DO THE TEST</H3>
</body>

</html>
```

The stylesheet used is this:

```
h1 {
color: white;
background-color: red;
font-size: 1.3em;
font-family: sans-serif;
}

h2 {
color: white;
background-color: red;
font-size: 1.05em;
font-family: sans-serif;
}

body {
font-size: 1em;
font-family: sans-serif;
color: #000000;
background-color: transparent;
}

p{
margin-left: 1em;
margin-right: 2em;

}

table{
width: 90%;
height: auto;
margin-bottom: auto;
border: 2px solid black;
padding:0;
vertical-align: top;
text-align: left;
font-size: 1em;
font-family: sans-serif;
color: #000000;
background-color: #EEEEEE;
}

b{
font-family:  sans-serif;
font-size: 1.02em;
}

ul{
	list-style-type:none;
}

tr{
vertical-align: top;
height:1%;
border: 1px solid black;
}

td{
```

```
vertical-align: top;
border: 1px solid black;
}

a{
color: red;
background-color: #EEEEEE;
font-family: sans-serif;

}

a:link, a:visited {
font-weight: bold;
text-decoration:none;
}

a:active, a:hover {
color: white;
background-color: red;
font-weight: bold;
text-decoration:none;
}



a.meny:link, a.meny:visited{
font-weight: bold;
text-decoration:underline;
}

a.meny:active, a.meny:hover{
color: white;
background-color: red;
font-weight: bold;
}
```