Visual Quality of Watermarking on Mobile Devices

Konstantin Krasavin, Jussi Parkkinen and Timo Jaaskelainen, Department of Computer Science, University of Joensuu, Finland

Abstract

In this study we evaluate visual quality of watermarked images, displayed on mobile devices. We consider a watermarking technique in spatial domain and embed a watermark into an image. We tested 3 different types of images and for each image 5 values for watermarking magnitude were tested. We tried to find out is there a difference in human evaluation of visual quality of watermarked image, depending on image properties and properties of a display. In visual assessment test we used a mobile phone, PDA, and CRT. The visual quality of watermarked image was evaluated by a set of human observers. The results were studied using statistical testing, which showed how does effect image type on perceptual quality of a watermark image.

Introduction

Recently images have become more and more common in the person – to – person communication. The newest developments in mobile messaging allow people to take and exchange pictures using a mobile phone. By Multimedia Message Services (MMS) users can send messages with sound, image, and even with short video.[1] It rises up a problem of protection of multimedia content of a message. We study methods of digital watermarking for mms-messages. Digital watermarking is a technique, where an identifier signal is embedded into an information carrying signal.[2-4] The watermark is embedded in such a way, that it does not disturb the information in normal conditions. The embedded watermark can be extracted for source identification. Watermarking is based on the feature of a human vision system which has different sensitivity to different frequency bands. The sensitivity of the human visual system to specific frequencies depends on screen properties, image size, and viewing distance. Recently an open standard operating system for mobile devices was developed, and it made possible to develop a watermarking applications for mobile devices. We introduce a watermarking technique for mobile devices. We consider specific features of mobile devices and based on visual assessment test give reasonable parameters for digital watermarking. The results of human evaluation were studied using statistical testing.

Digital Watermarking for Mobile Devices

A set of requirements for watermarking includes a perceptual invisibility of an embedded watermark, a proper quality of watermarked image, and watermark robustness. Watermark robustness acts as a trade off between the quality of the watermarked image and the quality of the extracted watermark.

Some methods of watermarking are based on the feature of a human vision system which has different sensitivity to different frequency bands.[5] The sensitivity of the human vision system to specific frequencies depends on screen properties, image size, viewing distance and light conditions. In order to design an efficient watermarking technique specific features of mobile devices have to be taken into account. A technique for mms-messages watermarking was presented in Ref. [6].

Watermarking developed in two domains: spatial and frequency domain.[3] Spatial domain watermarking usually is faster and consumes less memory. Frequency domain watermarking provides better support for human vision system, but it is slower and needs more memory because of domain transformations. Many watermarking methods use discrete cosine or wavelet transformed domain.[7]

In this study we consider a technique for watermarking in spatial domain presented by Bruyndonckx.[8] The proposed method is based on a spatial decomposition of the image in blocks and the pixel classification into homogeneous luminance zones. The choice of the block which will contain one bit of watermark is based on a secret key. The pixels in the blocks are classified in three groups of homogeneous contrast: The hard contrast, where two zones separated by a luminance step can be defined. The progressive contrast, where two uniform zones are separated by a progressive luminance variation. The noise contrast, with a luminance distributed like a random noise. The embedding of a bit in the block is performed through the relationship between luminance mean values:

If b=0: $m_{1B}^*-m_{1A}^* = \beta$

$$m_{2B}^*-m_{2A}^* = \beta \qquad (1)$$

If b=1: $m_{1A}^*-m_{1B}^* = \beta$

$$m_{2A}^*-m_{2B}^* = \beta$$

where m_{1B}^*, m_{1A}^*, m_{2A}^* and m_{2B}^* are the mean values required by b and β is the strength of watermark. The robustness of the embedded watermark is defined by a required difference β between the mean values. For our experiments we chose β = [5,7,10,12,15].

Visual Assessment Tests

In digital watermarking of images, a watermark is embedded into an image in such a way, that it is imperceptible. The magnitude of an embedded watermark defines the robustness and quality of watermarked image. We tested a number of watermark magnitudes for a set of images. For the visual assessment test we chose three images and five magnitudes of watermark. The watermarked image was shown on a mobile phone, PDA, and CRT display with 176 × 208, 240 × 320, and 1024 × 768 display resolutions, respectively. The quality of watermarked image is evaluated by 20 human observers.

The visual evaluation of image quality has been studied by a number of researches for many years. The assessment of watermarked image quality is done by means of direct category scaling.[9] The categories are associated with the numerical scores. In quality assessment experiment, subject is asked to classify watermarked images into a number of descriptive categories. A modified subjective mean opinion score (MOS) is used for measuring the quality of a watermarked image (Table 1):

Table 1: MOS Grading Scale

Score	Descriptive impairment
10	Imperceptible
8	Perceptible, but not annoying
6	Slightly annoying
4	Annoying
2	Very annoying

Subjective quality evaluations were carried out for three test images (Fig. 2). The evaluation is done in the way that a direct comparison of original and watermarked image is possible (Fig. 1). The subjects observed the images in normal office lights, without time limitation, and they could adjust the device position and viewing distance freely. Twenty subjects participated in the experiments. The subjects did not have any knowledge about strength of the watermark in each test image.

In one session, the original and one of five watermarked images was displayed on a device. The observers are asked to assign a score to each watermarked image. Each score is in the range of two to ten according to the scale of Table 1.

The session included the evaluation of image quality on a mobile phone, a PDA, and a CRT screen. The evaluation is done between the same kinds of device, independently of each other. Totally 45 images was evaluated by an observer.

For each watermarked image, the scores are averaged among all observers to obtain the MOS grade for a specific image. In Fig. 3 the MOS-scores for a specific image are shown. The normalization of individual quality scores is done using the z-score transform described by Hays.[10] This transform converts each score into a z-score, which indicates the deviation from the mean score. In Fig. 4 the z-scores for an each image are shown.

Test Images

For our experiments we selected three different images that represent a range of characteristics: one image is an image of human face, the second one is a technical and artificial image, and the third one is a high detail image (Fig. 2).

Figure 1. Experiment environment

We chose an image “Lena” as a human face image, map image as an artificial image, and an image “Sailor boats” as a high detailed image. Every image is watermarked with five different magnitudes. For experiments, we used images with sizes according to display resolution of a particular device. On the mobile phone we used images of size of 128 × 128 pixels, on the PDA we used images of size of 256 × 256, and on the CRT screen we used images of size of 512 × 512. In order to have a similar image distortion of watermarked image on different devices, we increased the size of watermarking according to dimensions of a particular image (Table 2). It equate ratio of embedded watermark and image size for a different sized images.

Table 2: Watermark Size

Image size, pixels	Watermark size, bits
128 × 128	256
256 × 256	1024
512 × 512	4096

Statistical Evaluation

We used an *Analysis of Variance* (ANOVA) to test the hypothesis that means over all visual quality grades are equal.[11,12] The ANOVA can be used under the natural assumption that the grade distributions are normally distributed, and the natural assumption was that the distribution is a normal one, and variances are similar.

Figure 2. Test images: Lena, Map, Sailor boats. At the upper row are the original images, and at the bottom row are the watermarked images with magnitude equaling to 15

The Fisher test is used to test the hypothesis that the variances between groups are the same. We tested the null hypothesis H_0: variances between the distributions of visual quality grades are equal, and H_1: that there the variances are different. Based on calculated values, with a significance level 95 % (α = 0.05) we rejected the null hypothesis that the variance are the same for some groups (Table 3).

Table 3: Pairs of Watermarking Magnitude, Where the Variances in Visual Quality are Different. Based on F-test with Significance Level 95%

	Mobile phone	PDA	CRT
Lena		(10,12) (10,15)	(5,15)
Map	(10,15)		(5,15) (7,15)
Sailor boats	(5,10) (7,10) (7,15)		

At the table are shown pairs of magnitudes, which have different variances in visual grades. For mobile device, the evaluation of the visual quality of Lena image was consistent among observers. For PDA and CRT, the variances were different for some pairs of magnitude. The Map image was evaluated on PDA with the same variance of grades. The Sailor boats image has a consistent variance in evaluation on PDA and CRT.

First we calculated a single factor ANOVA to the distribution of the grades of watermarked Lena image. The same way, the ANOVA procedure was applied to the group of grades for Map image, and Sailor boats image. That was done separately to the grades of mobile phone, PDA and CRT watermarked images.

We tested the null hypothesis H_0: means between the distributions of visual quality grades are equal, and H_1: that there is a difference. Based on calculated values, with a significance level 95 % (α = 0.05) we rejected the null hypothesis that there is no difference in visual quality.

The results of ANOVA tests show that the visual quality of watermarked images is significantly different for different watermarking magnitudes (Table 4).

Table 4: Significant Difference in Evaluation, Based on ANOVA with Significance Level 95%

	Mobile phone	PDA	CRT
Lena	+	+	+
Map	+	+	+
Sailor boats	+	+	+

Although that there is a difference in distribution among visual quality in a group of images, it does not answer to the question is there a significant difference between visual quality of sequenced watermark magnitudes.

The Student test is an equivalent of ANOVA test used to test for a difference in means between two distributions. The two-tailed test is used since there is no reason to assume that there is a difference between the groups.

We tested the null hypothesis H_0: means between the two distributions of grades are equal, and H_1: that there is a difference. Based on calculated values, with a significance level 95 % (α = 0.05) we rejected the null hypothesis that there is no difference in visual quality. We used t-test for all possible pairs of watermark strength values, and found out that some of the pairs of watermark strength do not showed a significant difference (Tab.5). At the table are shown pairs of values for magnitude of embedded watermark, for which is no significant difference is found.

Table 5: Pairs of Watermarking Magnitude, Where there is no Difference in Visual Quality. Based on T-test with Significance Level 95%

	Mobile phone	PDA	CRT
Lena	(7,10) (12,15)	(5,7) (12,15)	(7,10)
Map	(5,7) (10,12) (10,15) (12,15)	(7,10) (7,12) (7,15) (10,12) (10,15) (12,15)	(10,12) (12,15)
Sailor boats	(5,10) (7,10) (7,12)	(5,7) (10,12) (12,15)	(5,7) (10,12)

Results and Discussion

For each of test images, quality scores both for MOS-score and for z-score were obtained independently for each of devices. In Fig. 3 are shown MOS-scores for each of watermarked image displayed on the mobile phone, on the PDA, and on the CRT screen. In Fig. 4 are shown z-scores for each of the images. In each plot, the horizontal axis indicates the watermark magnitude, and the vertical axis indicates a score for a specific image.

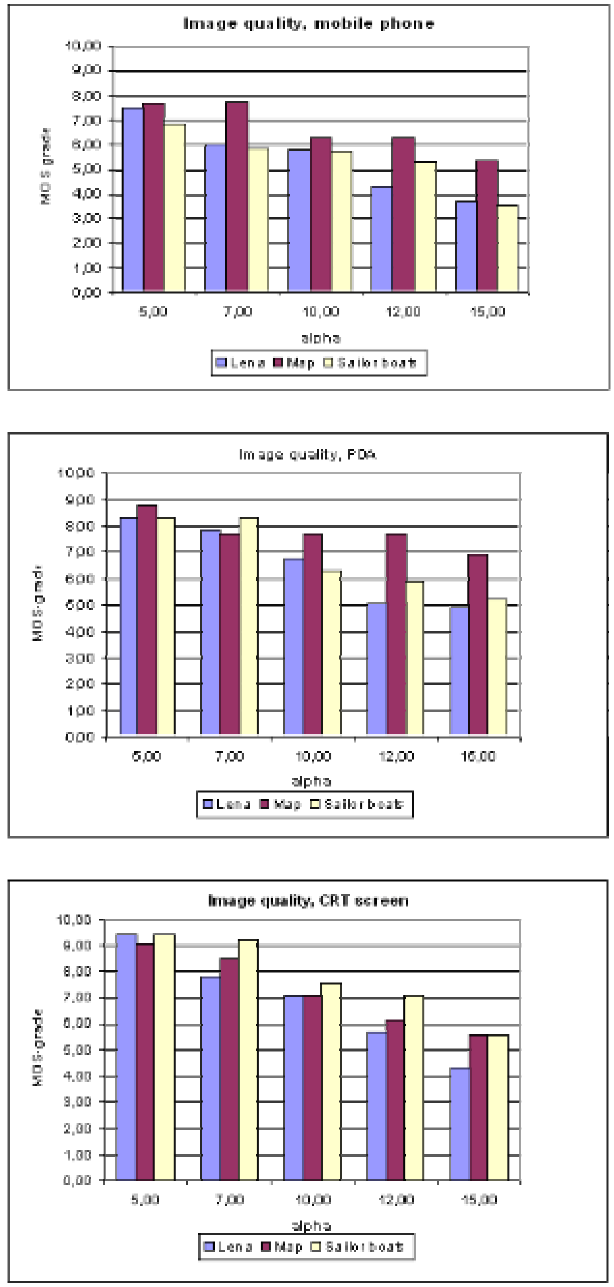

Figure 3. MOS-scores for the mobile phone, PDA and CRT screen.

Averaged results (Fig. 4) show that there is a difference in evaluation of images on different kind of devices. However, the difference in mean values is not significant for some watermark magnitudes, and it rises up a question about significance in visual quality evaluation. As can be seen in Fig. 4, there are difference in evaluation of tested images. In plot of image quality for mobile phone, the slope has clear breakpoints. The reason can be that it is difficult to distinguish images that have similar quality on small sized display of a mobile phone. The plot of image quality for CRT monitor has smoother curves, because larger image and larger screen let us notice the small differences in image quality. The plot for the PDA shows properties of something about the middle of mobile phone and CRT screen. On a mobile phone, the images with strongest watermark have better visual quality than on a CRT.

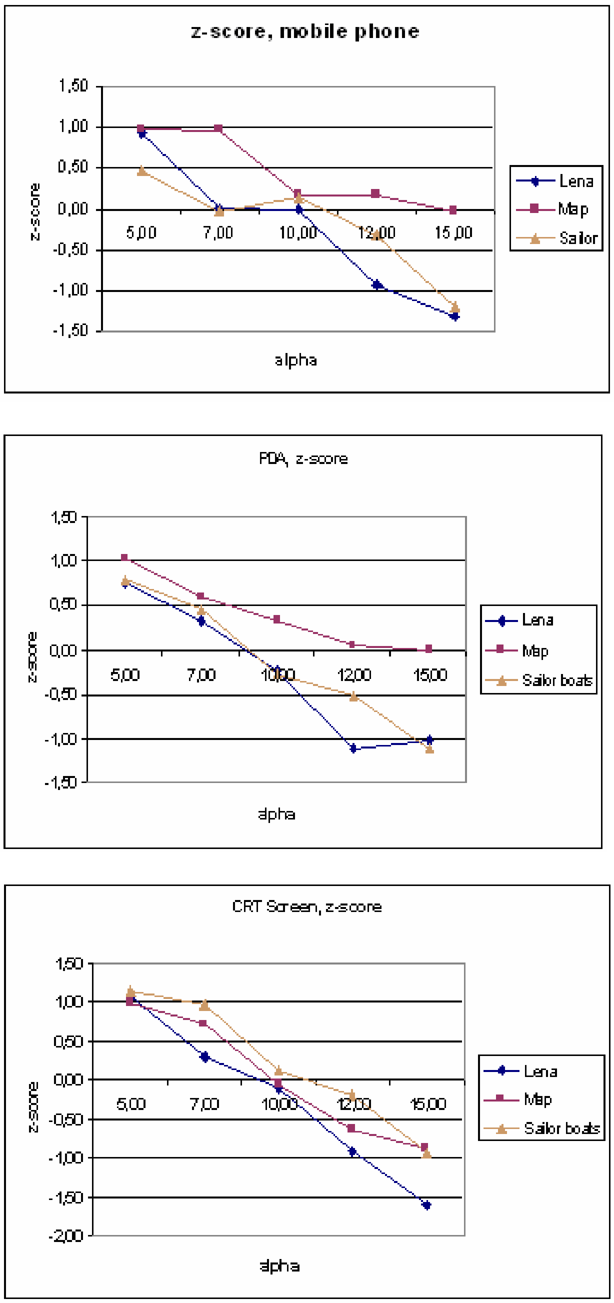

Figure 4. Z-scores for a mobile phone, PDA, and CRT screen.

The statistical tests showed that between some pairs of watermark magnitude is no significant difference for visual quality evaluation. For that reason we could chose the higher magnitude of an embedded watermark, and get more robust watermarking keeping the same visual quality of watermarked image.

The difference between image types can be seen from the results. For natural object, like a human face image, the quality differences can be seen clearly also in low resolution displays of a mobile phone. For artificial (map) image, the quality decrease is not so disturbing.

For a high resolution display (CRT) the quality decreases almost linearly according to the β- value, independently of image type.

It is shown that watermarking can be done for mobile without disturbance to the visual quality of the image. However, the PDA tests show that improvement of display resolution in mobile phones change the situation.

Conclusions

In this paper, we described an experiment of image quality evaluation, in which subjects had to judge the watermarked images using numerical category scaling. We evaluated the quality of watermarked images, displayed on mobile phone, PDA, and CRT screen.

We can conclude that for the high resolution displays, the image quality is important factor in watermarking. We have implemented a method for image watermarking for mobile phones and low strength watermark does not disturb the image visually.

Acknowledgment

The authors would like to thank all participants of the image quality evaluation.

References

1. Miraj E. Mostafa. “MMS – The Modern Wireless Solution for multimedia Messaging”, in The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Lisboa, Portugal,2002.
2. Ingemar J. Cox, Matthew L. Miller, Jeffrey A. Bloom. ”Digital Watermarking”. Academic Press, San Diego, USA, 2002.
3. Ingemar J.Cox, Joe Kilian, Tom Leighton and Talal Shamoon. “Secure Spread Spectrum Watermarking For Multimedia”, in IEEE Trans. on Image Processing,vol.6, no.12,1997, pp.1673-1687.
4. Fabien A. P. Petitcolas, Ross J. Anderson, and Markus G. Kuhn, “Information Hiding—A Survey”, in Proc. of the IEEE,vol.87, no.7, 1999, pp.1062–1078.
5. Olzak L.A. and Thomas J.P. “Handbook of Perception and Human Performance”, Wiley, New York, USA, 1986.
6. Krasavin K., Parkkinen J.,”Watermarking of mms messages”, FINSIG’03, Tampere, Finland, 2003.
7. Yun Q. Shi and Huifang Sun. “Image and Video Compression for Multimedia Engineering”, CRC Press LLC, Boca Raton, Florida, 2000.
8. O. Bruyndonckx, J.J. Quisquater, B.M. Macq. Spatial method for copyright labeling of digital images, in IEEE Workshop on Nonlinear Signal and Image Processing,1995, pp. 456 - 459.

9. A.M. Dijk, J.B.Martens, and A.B.Watson, “Quality assessment of coded images using numerical category scaling”, in Advanced Image and Video Communications and Storage Technologies.vol.2451,1995,pp.90-101.
10. W. Hays, Statistic, fourth edt., Rinehart and Wilson Inc., New York, USA, 1988.
11. Kachigan, Sam K. Statistical Analysis. Radius Press, New York, USA, 1982.
12. Levin, Irwin P, Relating statistics and experimental design, Sage Publications, Thousand Oaks, Canada, 1999.

Author Biography

Konstantin Krasavin received his MSc in mathematics from the University of St.Petersburg (2001) and MSc in computer science from the University of Joensuu (2003).He currently continuing PhD research on Watermarking of spectral images at the Department of Computer Science, University of Joensuu, Finland. His research areas are Image Processing, Digital Watermarking and Pattern Recognition.