Klima M, Pazderak J, Bernas M, Pata P, Hozman J, Roubik K. Objective and subjective image quality evaluation for security technology. In Proceedings IEEE 35th Annual 2001 International Carnahan Conference on Security Technology (Cat. No. 01CH37186) 2001 Oct 16 (pp. 108-114). IEEE.

Objective and Subjective Image Quality Evaluation for Security Technology

Milos Klima, Jiri Pazderak, Martin Bernas, Petr Pata, Jiri Hozman, Karel Roubik
Department of Radioelectronics, Faculty of Electrical Engineering,
Czech Technical University,
Technicka 2, 166 27 Praha 6, Czech Republic
klima@feld.cvut.cz

Abstract

The paper is devoted to the impacts of image compression algoritms on security image data. It compares three fundamentally different evaluation techniques of image – objective criteria, subjective criteria and identification. We have selected two typical security image data (a car plate and a face) with different initial quality and we applied three different compression techniques - two professional (JPEG and LuRaWave - LWF) and one implemented (Karhunen-Loeve transform KLT) [1]. A set of compressed images differing in compression rate was derived from each original image data. Finally the MSE as an objective criterion, the subjective image quality according to the ITU-R Rec. 500 and the identification measure were evaluated and compared.

Introduction

Nowadays a number of different image compression techniques is applied for image communication and/or image storage purposes. These compression techniques have a significant impact on the image perception and consequently the image interpretation. These impacts are not so important in relatively wideband systems such as a digital TV but they are extremely important when narrowband systems are used. The artefacts and distortions introduced by different compression techniques are of different origin and different effect. In this paper we have tried to demonstrate some typical cases and their implications in the security field.

The image quality has a fundamental importance for identification and classification of objects (persons, cars etc.), for analysis of motion and operations, in cases of normal and emergency operations and so forth. Generally we can define a number of image perception-related parameters – resolution in details (image acuity), resolution in levels (gradation), resolution in colors and resolution in motion. All these parameters are involved when compression-induced distortions or artefacts are present.

Lossy image coding standards

Very efficient and rough image data compression is required in two basic cases: vast image databases with a limited storage space and a transmission of image through an extremely narrow communication channel. There is a wide variety of lossless image compression techniques (esp. image computer formats) but they can be used to some extent of compression given by statistical properties of a particular image (the source entropy is a limit). In order to reach higher compression we have to apply some lossy compression formats introducing an additional distortion (impairment, error etc.) after reconstruction. The most frequently applied measure of image quality is calculation of some objective parameter such as MSE, SNR or similar. These parameters are easily calculated but they have a limited relation to the subjective impression (perception) or even to the possibility of identification or classification of objects in the image. The field of subjective image quality measure is widely studied and there is an enormous number of references. We will deal with still B&W images and extreme compression rates in the further discussion.

Basically two main techniques are recently used in lossy compression standards. The discrete cosine transform DCT followed by the data reduction employing a quantisation table is the first one. It is a part of JPEG standard. The second technique is the wavelet transform WT (applied e.g. in the FBI fingerprint database). The WT is a part of the format LuRaWave or JPEG2000. Two of these formats (JPEG and LuRaWave) are implemented in the IRFAN software package we have used.

The third method Karhunen - Loeve expansion is based on statistical properties of image data. By the distribution R we split image matrix into set of M submatrix - elements of vector space $\overset{N}{V}{}_{N}$. Let every such submatrix is one realisation of $N \times N$ dimensional random process. We can write for the correlation between realisations

$$\Xi_{in}^{jm} = E\{ {}^{\rho}[x_j^i| - \bar{[}x_j^i|)(|x_n^m]^{\rho} - |x_j^i])\}_{\rho=1}^{M}, \quad i,j,m,n = 1,\ldots,N. \tag{1}$$

From point of mean square error, the best orthonormal basis $|\{\Phi_i^j\}_s^r]$ of the space $\overset{N}{V}{}_{N}$ can be build from the eigenvectors

$$\sum_{i,j=1}^{N} \Xi_{nj}^{im} |\{\Phi_i^j\}_s^r] = \beta_s^r |\{\Phi_n^m\}_s^r], \quad m,n,r,s = 1,\ldots,N, \tag{2}$$

where β_s^r are eigenvalues of the covariance matrixes Ξ_{nj}^{im}.

0-7803-6636-0/01/$10.00 ©2001 IEEE

Image compression standard as an imaging system

Obviously we can treat the compression algorithm as any other imaging system. In a standard imaging approach we describe an imaging system by the Point Spread Function PSF or the Modulation Transfer Function MTF. We can define the response of imaging system to a point light source (2D Dirac impulse) as the impulse response or the Point Spread Function (PSF). The PSF is frequently used to characterise e.g. the optical blur. The relationship of the imaged object and the original is given by the convolution of the original object with PSF.

$$f_{image}(x,y) = f_{object}(x,y) * h(x,y)$$
$$f_{image}(u,v) = \tag{3}$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f_{object}(x,y).h(u-x, v-y)\,dx\,dy$$

In the frequency domain, based on the rules of Fourier transform, the convolution becomes a product.

$$F\{f_{image}(x,y)\} = F\{f_{object}(x,y)\} \times F\{h(x,y)\} \tag{4}$$

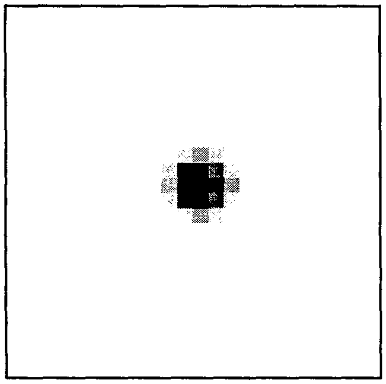

Fig. 1a: PSF – single pixel (15% coeff.) zoomed

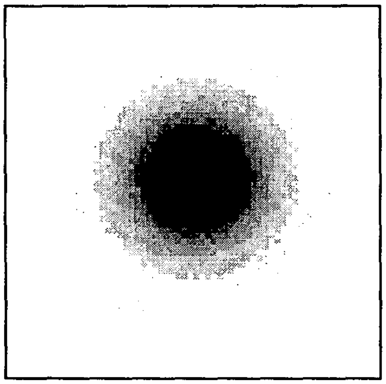

Fig. 1b: PSF – single pixel (3.2% coeff.) full scale

The Fourier transform of the PSF is known as the Optical Transfer Function (OTF).

$$F\{h(x,y)\} = OTF$$
$$OTF(u,v) = MTF(u,v).e^{j.PTF(u,v)} \tag{5}$$

The OTF is generally complex and thus it has a module (absolute value) and argument. The module OTF is called the Modulation Transfer Function MTF or Contrast Transfer Function CTF related to contrast. The argument is called the Phase Transfer Function related to frequency shifts. There is an important assumption that the imaging system is linear and both arbitrary input and output can be given by harmonic series with different spatial frequencies.

We can also define so called "contrast" C that helps to describe the contrast transfer efficiency by MTF with respect to spatial frequencies where B is brightness.

$$C = \frac{B_{max} - B_{min}}{B_{max} + B_{min}} \tag{6}$$

All parts of the imaging systems, i.e. atmosphere, objective, image sensor, image processing, image display and finally the observer's eye can be described by MTF. The MTF of the whole imaging system based on the above-mentioned equations is than given by the product of all particular MTFs.

$$MTF_{sys} = MTF_{atm}.MTF_{obj}.MTF_{sensor}.$$
$$.MTF_{DSP}.MTF_{disp}.MTF_{eye} \tag{7}$$

For the purposes of simulation, however, the compression standard (compression coder – decoder or CODEC) can usually, and conveniently, be treated as a "black box" with a possibility to provide the output images. As an example we have selected one of most the important compression standards - JPEG.

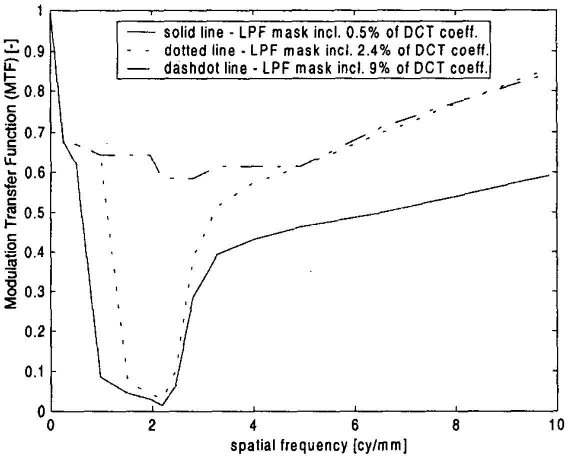

Fig. 2: MTF – examples

Using the above mentioned theory we can consider our "black box" (compression CODEC of JPEG) as an imaging system with PSF and MTF. We used these parameters because of their important relevance to evaluation of the subjective and objective image quality.

The following examples were calculated in the MATLAB in order to demonstrate the above mentioned objective characteristics for the JPEG standard. The PSF and MTF characteristics can be used in linear system only. Therefore the examples are calculated by the DCT (not by the complete JPEG coder). Figs 1a, b show the PSF where selected DCT coefficients are omitted (LP filtering). Some cases of the MTF for different amount of DCT coefficients are drawn in Fig. 2.

Relation between objective and subjective image quality - vision models

The subjective quality of image can be explained more deeply applying the human vision models [1]. Nearly all the published models have a general structure comprising the optical characteristics of the human ocular system, retinal sampling and some functions of the primary visual cortex, implemented as a multi-pathway algorithms conducting image processing in several intervals of spatial frequencies and for several edge orientations. Parameters of these models are derived from various psycho-physiological experiments. Such a model processes two input images: an original, not distorted, image and a tested affected one. Output of the model is commonly called a „Just Noticeable Difference“ (JND) map. Each point in the JND map indicates whether and how much a difference between the corresponding pixels in the original and affected images is visible for a human observer.

A further improvement of the model prediction accuracy can be made considering other psycho-physiological phenomena. A significant advancement of the model can be reached by employing algorithms for an automatic detection of the „regions of interest“ [3]. Impairment present in a region of interest affects the overall subjective image quality much strongly than an impairment of the same intensity present in another region of the given image.

Another important thing to consider during modelling of the human visual system is a difference between the subjective image quality, which can be defined as an impression that the given image evokes during its observing, and the information content of the perceived image that the image, processed by the human visual system, offers to the observer. The information content is undoubtedly of the higher importance in the security technology than the overall subjective image quality. The offered information content is indeed uneasy to express. Nevertheless, two different extremes can be defined according to the way of observing and analysing the image.

The first extreme appears during the „off-line“ examination of an image, when the observer has enough time to analyse the given image, to scan its each position by eye and evaluate it. The perceived image quality in this case can be evaluated using a model described above. The information the observer gains is limited primarily by the examined image quality rather than by any psycho-physiological phenomenon.

The second extreme occurs during the „real-time“ observing of a scene, when the time factor plays an essential role. Likewise, the time factor could necessarily be considered in a full connection with the human visual system characteristics, psychophysiology and psychology. Attraction of the observer's attention, various masking effects and seeking for „regions of interest“ are some examples, which must be involved into the model of the visual system designed for this second case. The effect of the visual attention attraction on the perceived information content can be seen from figures.

Fig. 3: Image of the real scene

Fig. 4: Image of the scene from Fig. 3 recorded by the human visual system. Explanation in text.

An original image of a scene recorded by a photo camera is presented in Fig. 3. The picture in Fig. 4 is an output image from a model of the earliest stages of the human visual system, comprising the ocular optics, retinal sampling and visual information processing in the nearest neuron layers to retina. The image in Fig. 4 is an image identical to that the observer see, while the observer's attention is fixed continuously on a face at the upper - left region. The applied model simulates, among others, rapidly decreasing spatial resolution (better to say „angular resolution“) with the increasing angle from the fixed attention point. The deterioration in the spatial resolution is furthermore more significant in diagonal directions than in the vertical one, while the least deterioration is performed for the horizontal direction. From picture in Fig. 4 one can conclude that during continuous fixation to an unvaried point in the scene the observer cannot determine and retain such peripheral patterns, which are situated angularly far from the visual axis, even though the patterns are quite large, as e.g. the number plate of a car in the Fig. 4. Therefore the observer cannot recall the patterns later, because the patterns were all the observing time under resolution ability of the observer's visual system.

Subjective vs objective image quality and methodology of subjective tests

Our work was devoted to the evaluation of image quality in both objective and subjective way. We applied the MSE Mean Square Error criterion as an objective measure of the image quality. The subjective quality was evaluated by a set of subjective tests as described further. For the evaluation of the quality of the images we chose the subjective tests based on the Recommendations ITU-R BT.500 [1]. The used method is slightly modified DSIS (double-stimulus impairment scale) method adapted for the evaluation of the image quality.

Two different sets of images (a car plate, a face) were prepared for the tests. The size of car plate number was set to 55×11 pixels and the size of face approx. 60×85 pixels as typical scenes. These images were then compressed using three different types of lossy compression methods (JPEG, LuRaWave, KLT) with a set of compression rates. These images were ordered into two random sequences, each containing the same image but different compression ratio and/or method. In front of each image in the sequence the reference (original) image was placed. Then each individual test consisted of two presentations, the reference image first and then compressed image. The quality was evaluated using 5-degree scale with two valid figures (1, 1.1, 1.2, ...4.9, 5). At the beginning the reference, maximally degraded image and one middle degraded image were shown to the observers with an explanation of important image parameters. The observers were students from the Faculty of Electrical Engineering but not professionals.

The test images were taken by scanning of high quality slides. For the scanning the resolution of 1200 dpi was used and the excision (the area chosen) was subsampled to the output resolution 720x576 pixels. The photographic equipment was based on camera Nikon FE10 with lenses 35-70/3.5 and the dia film Fuji Provia 100 ASA. The digital camera OLYMPUS Camedia C-1400L was used. This camera has an image sensor CCD dia 2/3" with a resolution 1.4 million pixels. The image can be stored as SHQ (Super High Quality) in a raster 1280 x 1024 pixels in the JFIF format. The car plate image was taken by digital camera and the face image by the film camera.

The reference images were processed by three compression methods with different compression rates as described. The monitor Nokia 19" was used for displaying. For the test the room with dark curtains (window hangins) was used. The viewing conditions were similar to the laboratory environment conditions defined in [1]. Because of the maximum available luminance of the screen was 120 cd/m^2, the luminance of inactive screen was set to the 2.4 cd/m^2, luminance of black level on the screen in a completely dark room was set to 1.2 cd/m^2 and luminance of the background was set to 1.8 cd/m^2 according to the standard. The viewing distance was 5 x picture height and the observation angle was 0° - axial.

The group of 11 observers was involved in the subjective image quality evaluation and the group of 6 observers was involved in the recognition test (the identification of the car plate or the face).

Objective and subjective image quality evaluation – the results

This part summarises our results of simulation and experiments. Figs 5, 6 show the original version of both the images – the car plate and the face.

Fig. 5: Original image (a car plate)

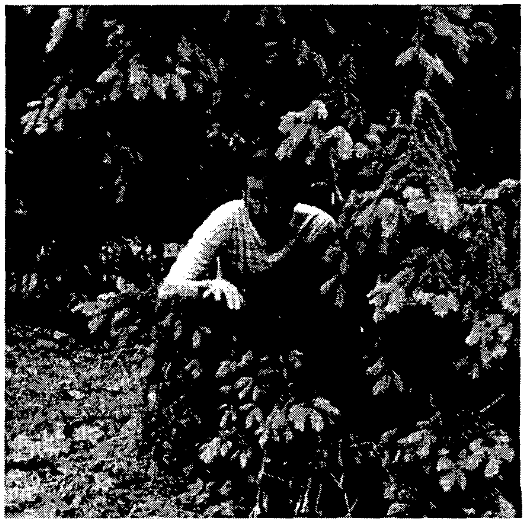

Fig. 6: Original image (a face)

Figs 7 till 9 demonstrate relevant dependencies of the Root Mean Square Error RMSE as an objective measure and the subjective image quality as function of the compression rate. The compression rate was derived from real sizes of image files. Moreover identification limits is shown there. The JPEG compression relates to Figs 7a, b and the worst case of compressed images is shown on Figs 7c, d. In the same way the LWF resp. KLT compression is displayed on Figs 8a, b, c, d resp. 9a, b, c, d.

Conclusion

The mutual relation of the objective distortion measure, subjective quality and identification is a very complex problem. Our paper has demonstrated all three parameters in a common figures for two typical security scenes. It can be deducted that the critical security procedure – identification – is in a very vague relation to the both objective and/or subjective image quality. The car plate number and face are of different origin and the face identification can succeed under much more severe distortion conditions. Therefore the application and implementation of general image compression algorithms and standards have to be carefully evaluated for a particular image type because they are optimized from the objective or subjective image quality point of view.

Acknowledgements

This work has been supported by the research program No. MSM 212300014 "Research in the area of information technologies and communications" sponsored by the Ministry of Education, Youth and Sports of the Czech Republic.

The authors want to acknowledge Mr. Kloucek for the technical support during the tests and processing of data.

References

[1] Bovik A., Handbook of Image Processing, Academic Press, 2000

[2] Safranek R.J., Comparison of the coding efficiency of perceptual models, Human Vision, Visual Processing and Digital Displays, Proc. SPIE, 2179, 1995

[3]Osberger, W.: Perceptual Vision Models for Picture Quality Assessment and Compression Applications. Doctoral Theses. Queensland University of Technology, 1999.

[4] ITU-R Recommendation BT.500-10 (2000), Methodology for the subjective assessment of the quality of the television pictures.

[5] ITU-R Recommendation BT.1129-2 (1998), Subjective assessment of standard definition digital television (SDTV) systems.

[6] Miano, J.: Compressed Image File Formats: JPEG, PNG, GIF, XBM, BMP (ACM Press). Addison-Wesley Longman, Inc., 1999. 264 p.

Klima M, Pazderak J, Bernas M, Pata P, Hozman J, Roubik K. Objective and subjective image quality evaluation for security technology.
In Proceedings IEEE 35th Annual 2001 International Carnahan Conference on Security Technology (Cat. No. 01CH37186) 2001 Oct 16 (pp. 108-114). IEEE.

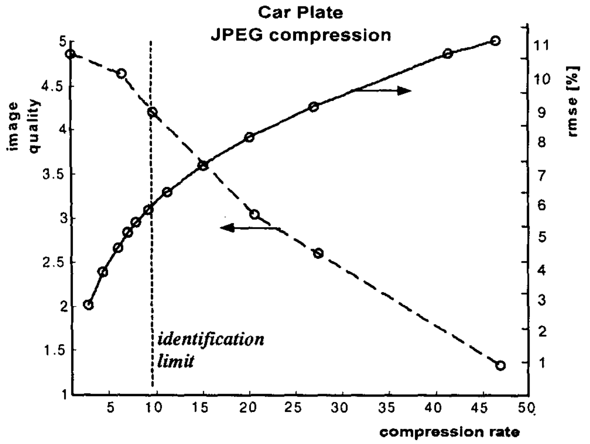

Fig. 7a: RMSE resp. subjective image quality as a function of compression rate - JPEG

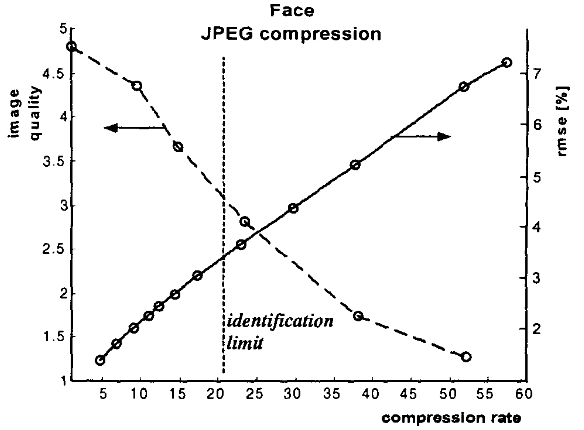

Fig. 7b: RMSE resp. subjective image quality as a function of compression rate - JPEG

Fig. 7c: The worst case of compressed image JPEG

Fig. 7d: The worst case of compressed image JPEG

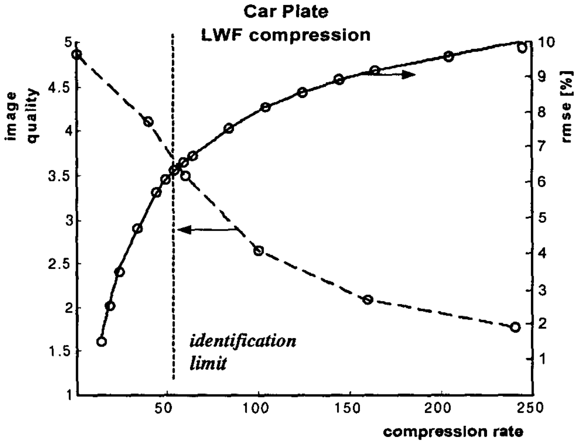

Fig. 8a: RMSE resp. subjective image quality as a function of compression rate - LWF

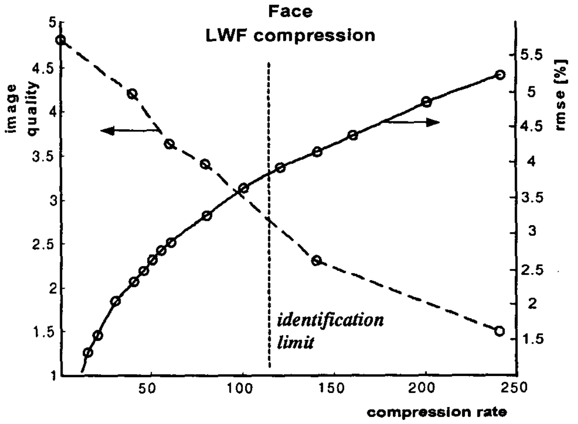

Fig. 8b: RMSE resp. subjective image quality as a function of compression rate - LWF

Fig. 8c: The worst case of compressed image LWF

Fig. 8d: The worst case of compressed image LWF

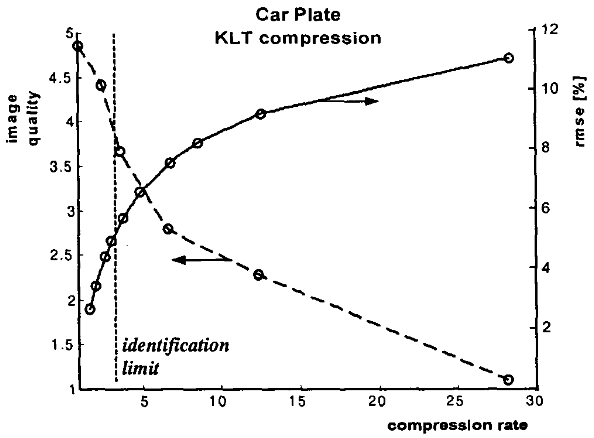

Fig. 9a: RMSE resp. subjective image quality as a function of compression rate - KLT

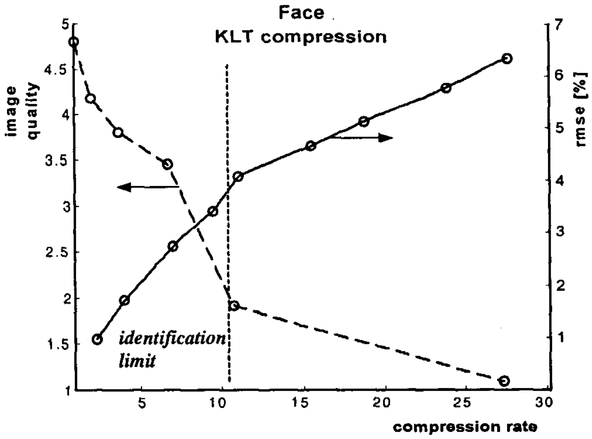

Fig. 9b: RMSE resp. subjective image quality as a function of compression rate - KLT

Fig. 9c: The worst case of compressed image KLT

Fig. 9d: The worst case of compressed image KLT