

Computational Image Quality Metrics for Watermarking Applications

N.V.Rao¹ and Prof. S.N.N.Pandit²

¹CVR College of Engineering, Department of CSE, Ibrahimpatan, R.R.District, A.P., India
Email: nvrao@ieee.org

²Honorary Director, Center for Quantitative Methods, Osmania University, Hyderabad A.P., India

Abstract—Measurement of image Quality has been a challenging problem in many image-processing fields. Many procedures have been proposed to define metrics for image quality comparison in the context of image compression and watermarking. Subjective and Objective measures are considered to be the prime groups in the classification of quality measures as reported in the literature. Subjective evaluation is not an easy task as it involves environmental conditions. Mean square error (MSE), peak signal to noise ratio (PSNR), correlation coefficients, structural similarity index (SSIM), universal quality index (UQI) are some of the objective measures proposed for evaluation of Image quality. In this paper some widely used measures have been reviewed in general and considered especially through case studies in image compression and digital watermarking.

Index Terms—Image quality assessment, MSE, PSNR, SSIM, UQI, Objective evaluation and Digital watermarking.

I. INTRODUCTION

In general, with human visual system (HVS), the quality of similar images can be assessed based on how they look, and a comparison can also be made. Many procedures have been proposed to define metrics for image quality comparison. Wilsow et al. [6] proposed that image quality metrics can be categorized into three types through

1. Human perception
2. Objective measures based on theoretical models
3. Subjective measures based on mathematically defined models of the HVS.

The first category of metrics employs a selected group of human viewers to judge the quality of selected images. The second category employs matrix representation of images, and mathematically based proven models that use transforms. The third type is based on human visual system models.

1.1 A brief survey of Image quality measures

Image and video data is to be compressed by many times in order to accomplish the transmission at a suitable rate. In fact data compression becomes an enabling technology to bridge the gap between the required high volume of video data and the limited hardware capability. Similarly Watermarking introduces distortion in the digital data. But it is very difficult to accurately measure the amount and visibility distortions caused by watermarks in image or video contents. Eskicioglu et al. in their survey papers presented many approaches to measure image quality in compressed images [7] and watermarked images [8]. They suggested two types, subjective and objective measures. The subjective evaluation is based on human observations or perception. It can not be taken as standard because human observation is based on some critical factors such as environment, motivation and mood. In spite of their complicated algorithms, the human visual system (HVS) based objective measures do not appear to be superior to the simple pixel based measures like the mean square error (MSE), peak signal to noise ratio (PSNR) or root mean square error (RMSE). The objective category includes bivariate measures such as MSE, L_2 -norm and other measures similar to HVS. An ideal image quality measure should be able to describe the amount of distortion, the type of distortion and the distribution of the error [8].

Wang et al. [10] presents a new numerical measure for gray scale images, namely, Universal Image Quality Index (UQI) whose range is [-1 1]. The best value results when both images are the same. The index is computed using a sliding window of size 8x8, leading to a quality map of the image. The overall quality index is the average of all UQI values in the quality map. To make it stable [11], the measure has been generalized to the Structural Similarity Index (SSIM). The overall image quality, Mean of Structural Similarity Index of image (MSSIM) is obtained by computing the average of SSIM values over the windows.

In this work, widely used objective measures are considered for measurement of similarity of the original

and watermarked images and original and extracted watermarks. Similarly image compression with 20%, 30%, 50% are also considered for quality measurement. The present section describes these measures and their formulations. Values of PSNR, MSE are computed and the SSIM index between two images has been computed using MATLAB code given by Wang et al. [11].

In the context of watermarking, the quality of the watermarked image must be very high. The embedded watermark in the host should be perceptually invisible. In general, a PSNR (peak signal-to-noise ratio) larger than or equal to 30 dB in a reconstructed image is considered acceptable [12]. However acceptable value depends on application requirement.

The ultimate receivers of the data are users and hence visual quality should be judged by human view. In general, image quality is evaluated subjectively by rating the image quality, and also by measuring the image impairment. A fine scale rating system of the degree of impairment: Noticeable/just noticeable/definitely noticeable but not objectionable/ objectionable/ extremely objectionable is now being used in industries as ITU-recommendations [13]. However, subjective quality assessments are costly and time consuming because human eyes are easily fatigued. Evaluation depends on user and hence may not be standardized.

II. OBJECTIVE QUALITY MEASURES

Signal to noise (SNR) measures are estimates of the quality of a reconstructed or modified image compared with the original. They are easier to compute and reasonable estimates as a single number that reflect the quality of the reconstructed image.

2.1 Mean Square Error (MSE)

$$MSE = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N (f(x, y) - f^1(x, y))^2$$

where f and f^1 are the host and stego (watermarked) images. $M \times N$, the size of the images. Small values of MSE indicates acceptable degradation.

RMSE is the root mean squared error

$$RMSE = \sqrt{MSE}$$

2.2 Signal to noise Ratio (SNR)

$$SNR = 10 \log_{10} \left(\frac{MSE2}{MSE1} \right) dB$$

Where MSE1 is the mean square error of original image and MSE2 is the mean square error of processed image.

2.3 Peak Signal to Noise Ratio (PSNR)

Peak signal to reconstructed image measure is computed by

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) dB$$

PSNR penalizes the visibility of noise in an image. In multimedia applications, any image with more than 30 dB is accepted in general.

2.4 SVD based Gray-scale Image Quality Measure

In [4,9], a graphical measure (which is a bivariate measure) that computes the distance between the singular values of the original image and the singular values of the distorted image block, has been defined as

$$D_i = \sqrt{\sum_{i=1}^n (S_i - S_i^1)^2}$$

Where S_i and S_i^1 are the singular values of the original and distorted block respectively. This numerical measure is a Minkowski metric $(\sum |S_k - \bar{S}_k|^\beta)^{1/\beta}$ with $\beta = 2$.

If the image is of size $K \times K$, we have $K/n \times K/n$ blocks.

The set of distortions when displayed in a graph represents a distortion map. A numerical measure is derived from the graphical measure. It computes the global error expressed as a single numerical value depending on distortion type

$$M-SVD = \frac{\sum_{i=1}^{(K/n) \times (K/n)} |D_i - D_{mid}|}{(K/n)(K/n)}$$

Where D_{mid} represents the mid point of the sorted D_i 's, $K \times K$ is the image size, and n is the block size.

2.5 Difference Image Visibility

An important technique for displaying errors is to construct an error image which shows the pixel-by-pixel errors. The simplest computation of this image is to create an image by taking the difference between the reconstructed and original pixels. In order to create an image, the difference is multiplied by a constant and the entire image is converted to gray level by adding a constant for visibility as shown below. (to make it non-negative)

$$E(x, y) = 2 [f(x, y) - f^1(x, y)] + 128$$

More details are given in [1,2]

2.6 Frobenius norm

One of the good ways of measuring error in images is that the error measure agrees with human perception. Frobenius norm, which is related to L2 - norm for

functions and also to the concept of energy, is used in image processing.

Consider a digital image represented as the MxN matrix I and its elements denoted as I_{ij} , $i = 1,2,3,\dots, M$, $j = 1,2,3, \dots, N$. Frobenius norm of the matrix I is defined as

$$\|I\|_{Fro} = \sqrt{\sum_{i=1}^M \sum_{j=1}^N |I_{ij}|^2}$$

The relative error (Frobenius Error) is given by

$$E_{rel} = \frac{\|I - I_c\|_{Fro}}{\|I\|_{Fro}}$$

Where I_c represents a modified image. Frobenius norm and singular values of an image are related through

$$\sqrt{\sum_{i=1}^M \sum_{j=1}^N |I_{ij}|^2} = \sqrt{\sum_{i=1}^N \sigma_i^2}$$

σ_i , $i = 1, 2, \dots, N$ denotes the singular values of I .

One way of measuring error in images is to compute a measure that *agrees with our eyes* or in other words, in a manner *what you see is what you get*. It is related to L2 – norm for functions, which in turn, is related to the concept of energy. For this reason, Frobenius norm is used in image processing.

2.7 Structural Similarity Index (SSIM)

Wang and Bovik [10] proposed universal Image Quality index (UQI) and improved in [11] as Structural SIMilarity (SSIM) index. In UQI, the dynamic range of quality is [-1 1] and the best value is achieved when $y_i = x_i$, $i = 1,2,\dots,n$. This index models any distortion as a combination of three different factors i) loss of correlation, ii) mean distortion and iii) variance distortion. A sliding window of 8x8 size is used for computing the index that results in quality map of the image. The overall index is the average of all UQI values in the quality map. To avoid some mathematical instability, SSIM is proposed in [11] as a refinement. MSSIM is obtained by computing the average SSIM values over all windows.

The universal image quality index is defined as

$$Q = \frac{4\sigma_{xy} \bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)[\bar{x}^2 + \bar{y}^2]}$$

Where $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$,

$$\sigma_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2, \sigma_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2$$

$$\sigma_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Where x_i, y_i , $i = 1,2,3,\dots, n$ represents the original and distorted signals respectively.

2.8 SSIM index

SSIM index between X, Y is defined as

$$SSIM(X, Y) = \frac{(2\bar{x}\bar{y} + C_1)(2\sigma_{xy} + C_2)}{(\bar{x}^2 + \bar{y}^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

where C_1, C_2 are small constants given by $C_1 = (K_1R)^2$ and $C_2 = (K_2R)^2$ respectively, R is the dynamic range of pixel values i.e $R = 255$.

K_1 and $K_2 \ll 1$ are two scalar constants. ($K_1 = 0.01$, $K_2 = 0.03$ for the experiments presented in [11])

2.9 Normalised Correlation Coefficient (NCC)

The presence of watermark is evaluated quantitatively by measuring the similarity between the original and extracted watermarks. [3]. A commonly used similarity measure is given by the Normalized Correlation Coefficient, which is represented with ρ .

$$\rho(W, W^1) = \frac{\sum_{x=1}^m \sum_{y=1}^n W(x, y)W^1(x, y)}{\sqrt{\sum_{x=1}^m \sum_{y=1}^n |W(x, y)|^2} \sqrt{\sum_{x=1}^m \sum_{y=1}^n |W^1(x, y)|^2}}$$

Where W and W^1 represents the original and extracted watermark sequence of $m \times n$ size respectively. Watermark is detected if $\rho(W, W^1) > T$, where T is a specified threshold, chosen to minimize false alarm (detection indicates positively even if no watermark is embedded). But it is clear that the extracted watermark is also a visually recognizable pattern, which can be used to subjectively evaluate the performance of the watermarking scheme. In addition to visual similarity, a high correlation between W and W^1 indicates that they are the same or statistically similar.

2.10 Pearson's Correlation Coefficient (PCC)

In objective evaluation, statistical measures like Pearson's correlation coefficient can be used [5]. Pearson product moment correlation coefficient is a dimensionless index that ranges from -1.0 to 1.0 and reflects the extent of a linear relationship between two data sets to make it clear further the square of the correlation coefficient gives the percentage of the variation explained in one of the variable when the variation in the other is taken into account to predict its value. In the case of SVD based watermarking methods, these moment correlation coefficients are computed between the original vector of singular values and

extracted vector of singular values. This is done for every quadrant in case of DCT or DWT based (multi-band) embedding.

2.11 Pearson's Correlation Coefficient: A general form

The correlation coefficient r as given [13] (also called Pearson's product moment correlation after Karl Pearson) is calculated by

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

Where $x_i, y_i, i = 1, 2, \dots, n$, represent original and distorted signals

The correlation coefficient may take any value between -1.0 and +1.0. A linear relationship between x and y is assumed.

The above formula can be replaced by the following equivalent, which avoids to use the means and is therefore much faster to compute.

$$r = \frac{\sum_{i=1}^n x_i y_i - \frac{1}{n} \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{\sqrt{\sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2} \sqrt{\sum_{i=1}^n y_i^2 - \frac{1}{n} (\sum_{i=1}^n y_i)^2}}$$

The correlation coefficient stands in close relationship to linear regression [13]. The square of r is called the *goodness of fit* and denotes the portion of total variance in a regression model.

III. EXPERIMENTS AND RESULTS

In this section, we consider two applications and evaluate the objective quality measures and also display the images for subjectivity. The first application is digital watermarking, in which quality assessment is required to assess the quality of the watermarking images. The second application is image compression. In this application, quality metrics are needed to evaluate the quality of reconstructed images and also to determine the appropriate compression factor.

3.1 Digital watermarking

Digital image watermarking is the process of embedding a logo or data called watermark into a host image. One of the main characteristics of a digital image watermarking is the imperceptibility. The quality of the image should not be compromised for the sake of authenticity. The primary quality observation is through human visibility or how the image appears to the human eye. In addition, some objective measures based on statistics can provide the distortion quantitatively,

enabling the limit of acceptability. A case study of digital watermarking employing Satish Chandra's SVD algorithm [3] has been considered. The host, watermark and stego images considered are shown in figure 1.

The algorithm proposed by Chandra [3] treats host and watermark images globally for computation of SVD and for embedding watermark. A brief account of algorithm is given below for completeness.

Watermark Embedding Technique:

Let X represents the cover image of size $M \times N$, let W , the watermark be an array of size $P \times Q$. Singular value decomposition (SVD) of X and W are computed as

$$X = U \Sigma_x V^T$$

$$W = U_w \Sigma_w V_w^T$$

The diagonal elements of Σ_x and Σ_w represents the singular values of X and W respectively and are represented by

$$[\sigma_x = \sigma_{x_1} \quad \sigma_{x_2} \quad \sigma_{x_3} \quad \dots \quad \sigma_{x_N}]$$

$$[\sigma_w = \sigma_{w_1} \quad \sigma_{w_2} \quad \sigma_{w_3} \quad \dots \quad \sigma_{w_Q}]$$

1 The watermark is embedded into singular values of X according to relationship

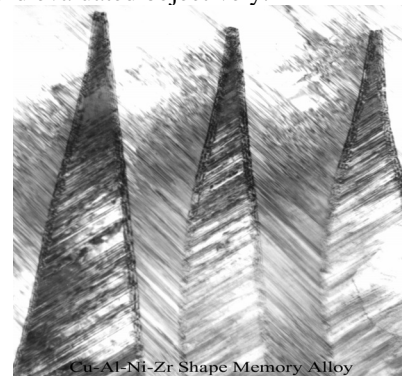
$$\sigma_{y_i} = \sigma_{x_i} + \alpha \sigma_{w_i}$$

Where α is the scaling parameter which determines the embedding strength and is chosen to maintain perceptual fidelity of the watermarked image Y .

Let $\Sigma_y = [\sigma_{y_i}]$, i.e Σ_y represents the diagonal matrix whose elements corresponds to σ_{y_i} . The watermarked image Y is computed as

$$Y = U \Sigma_y V^T$$

'Alpha' (α), the strength of watermark, is varied so as to decide the acceptable level of the watermark. The original and watermarked images are compared for visibility and evaluated objectively.



(a)

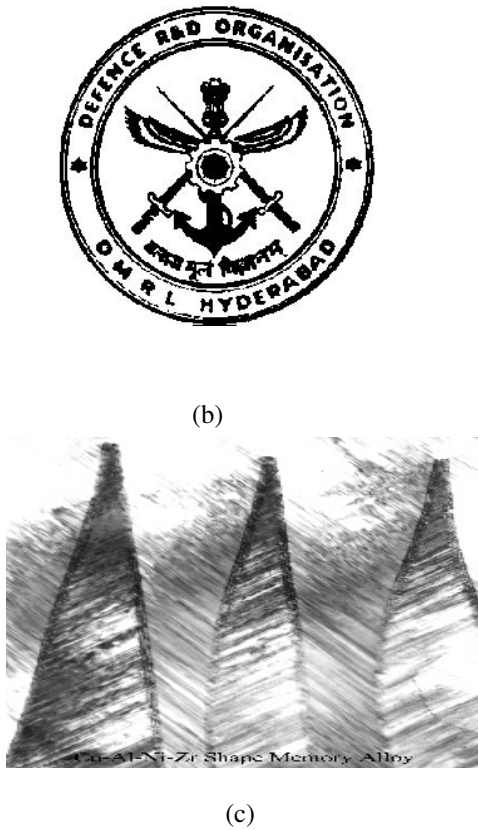


Figure 1. Host, Watermark and Stego Images:
(a) Host (Original) (b) Watermark (c) Stego Images

Table 1:

Table showing the error metrics between original and watermarked images with respect to the controllable parameter alpha.

ALPHA	MSE	SNR	PSNR	R_Error	Q_Fact
0.0010	0.0464	58.6578	61.4650	0.001	1.000
0.0060	1.6706	43.1433	45.9020	0.007	1.000
0.0110	5.6152	37.9268	40.6371	0.013	1.000
0.0160	11.8801	34.7203	37.3826	0.019	0.999
0.0210	20.4654	32.4062	35.0206	0.025	0.998
0.0260	31.3710	30.5987	33.1655	0.030	0.998
0.0310	44.5970	29.1182	31.6378	0.036	0.997
0.0360	60.1433	27.8665	30.3389	0.042	0.995

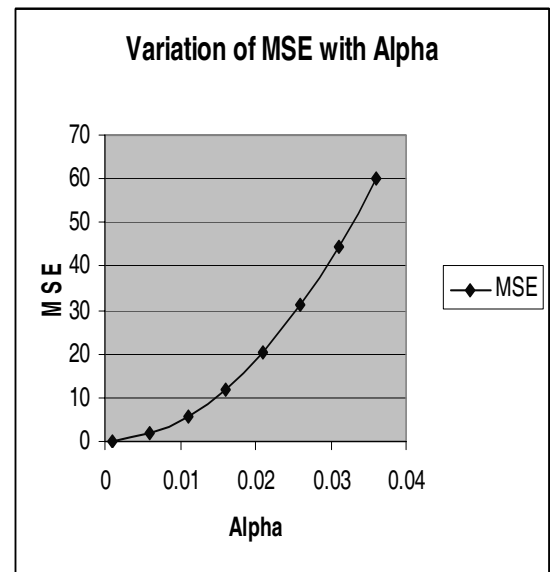


Figure 2.

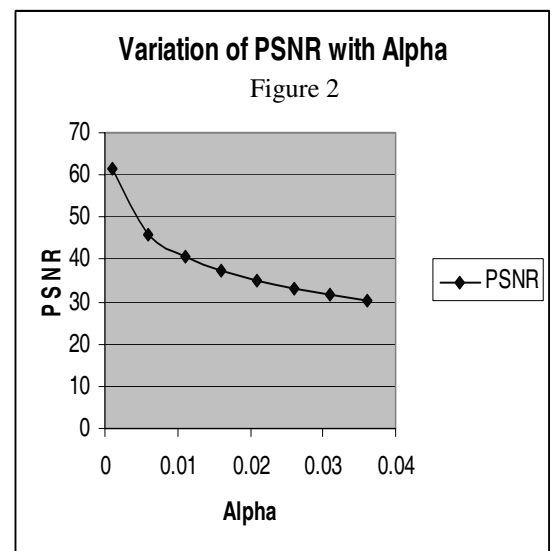


Figure 3

MSE, SNR, PSNR, R-error (Frobenius), Q-factor (SSIM index) are tabulated for different values of alpha in table 1. These five measures are considered in this paper for illustration. PSNR and SSIM index are widely accepted for image quality assessment. Variation of MSE and PSNR with alpha are shown graphically in figures 2-3. In general a PSNR > 40 dB is acceptable for watermarking applications in quality critical applications like Medical Image watermarking.

3.2 Image compression

Quality in image compression depends on closeness of the reconstructed image with the original. On the other hand the data is to be reduced especially the redundant data in coding is to be eliminated, resulting in optimum data with storage reduction. The reconstructed image from compressed image data is expected to be very close to the original without any noticeable difference. In this present paper, compression is achieved by JPEG compression technique.

An example of Lena image is considered, compressing it with 20%, 30%, and 50% Factors. The original and resulting compressed figures are shown as figures 4(a)-(d). The compressed image and the original are compared for quality measures. It can be noticed that a PSNR>40 is acceptable from the quality point of view.



(a)



(b)



(c)



(d)

Figure 4:

- (a) Original Image
- (b) 20% Compressed image
- (c) 30% Compressed image
- (d) 50% Compressed image

SUMMARY AND CONCLUSIONS

Both subjective and objective assessments are necessary for image processing applications. Objective measures numerically give an indication of the quality and subjective test is for user's view. In this paper good number of objective measures have been reviewed and demonstrated through two applications. These objective measures definitely help in selecting critical parameters in the applications such as strength of watermark in watermark embedding and compression factor in image and video compression.

REFERENCES

- [1] A.N. Netravali and B.G. Haskell, *Digital Pictures: Representation, Compression and Standards*, Plenum Press, New York, 1995.
- [2] M. Rabbani and P.W. Jones, *Digital Image Compression Techniques*, SPIE , Vol. TT7, Optical Engineering Press, bellvue, Washington, 1991.
- [3] D.V. Satish Chandra, *Digital Image Watermarking using Singular Value Decomposition*, Proceedings of 45th IEEE Midwest Symposium on Circuits and Systems, Vol. 3, pp. 264-267, 2002.
- [4] Aleksandr Shnayderman, Alexander Gusev and Ahmet M. Eskicioglu, *An SVD-Based Gray-Scale Image Quality Measure for Local and Global Assessment*, IEEE Transactions on Image Processing, Vol. 15, No.2, pp. 422-429 February 2006.
- [5] Alexander Sverdlov, Scott Dexter and Ahmet M. Eskicioglu, *Robust DCT-SVD Domain Image Watermarking for Copyright Protection: Embedding Data in All Frequencies*, 13th European Signal Processing Conference (EUSIPCO-2005), Antalya, Turkey, September 4-8, 2005.

- [6] Dale L. Wilson, Andrian J. Baddeley and Robyn A. Ownes, *A New Metric for Gray-Scale Image Comparison*, International Journal of Computer Vision, Vol. 24, No.1, pp. 5-17, 1997.
- [7] A.M. Eskicioglu, *Quality Measurements for Monochrome Compressed Images in the past 25 years*, Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Vol. 4, pp. 1907-1910, Istanbul, Turkey, June 5-9, 2000.
- [8] Aleksandr Shnayderman and Ahmet M. Eskicioglu, *Evaluating the Visual Quality of Watermarked Images*, IS&T/SPIE's 18th Annual Symposium on Electronic Imaging, Security, Steganography and Watermarking of Multimedia Contents, VIII Conference, San Jose, CA, January 15-19, 2006.
- [9] Aleksandr Shnayderman, Alexander Gusev and Ahmet M. Eskicioglu, *A Multidimensional Image Quality Measure Using Singular Value Decomposition*, Proceedings of the SPIE Image Quality and System Performance Conference, San Jose, CA, Vol. 5294, pp. 82-92, January 19-20, 2004.
- [10] Zhou Wang and A. Bovik, *A Universal Image Quality Index*, IEEE Signal Processing Letters, Vol. 9, No. 3, pp. 81-84, March 2002.
- [11] Zhou Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, *Image Quality Assessment: From Error to Structural Similarity*, IEEE Transactions on Image Processing, Vol.13, No.4, April 2004..
- [12] C.- C. Chang, K.- F. Hwang and M. - S. Hwang, *Robust Authentication Scheme for Protecting Copyrights of Images and Graphs*, IEE Proceedings on Vision, Image and Signal Processing, Vol. 149, No.1, pp. 43-50, February 2002.
- [13] CCIR Recommendations 500-3, *Method for the Subjective Assessment of Quality of Television Pictures*, Recommendations and Reports of the CCIR, 1986, XVI th Plenary Assembly, Vol. XI, Part I.
- [14] H. Lohniger, *Teach Me Data Analysis*, Springer -verlag, 1999.