Evaluation of Perceptual Quality for Watermarked Images Based on Combination of Fuzzy Similarity Measures Using Neural Network

Methaq Gaata[1,2], Sattar Sadkhn[2], Saad Hasson[2]

**[1] LIRMM UMR CNRS 5506, University Montpellier 2
Montpellier, 34392, France**

**[2] Computer Science Department, Babylon University
Hilla Babylon 00964, Iraq**

Abstract

In this paper, focus is placed on the design a new evaluator to assess the quality of the image watermarking techniques. The main idea is the introduction of an image quality evaluator dependent on a combination of five fuzzy logic-based similarity measures and neural network. In the first stage, fuzzy similarity measures are computed as features of each pair of original and watermarked images and these features are used as input to neural network. In the second stage, these features are combined by using neural network to predict a subjective image quality, known as the Mean Opinion Score (MOS) automatically. The performance of the suggested evaluator is assessed in terms of good correlation with the MOS using the image watermarked database (IVC image database). Experimental results, using 210 tested images, show that the evaluation outputs correlate highly with the MOS scores.

Keywords: *Image Quality Evaluation, Fuzzy Similarity Measures, Neural network.*

1. Introduction

Image Quality Assessment (IQA) has always been an integral part of image and video processing. Many different approaches for IQA with different complexity were developed in the last decade. The main goal of image quality assessment metrics is to provide an automatic and efficient way to predict visual quality. It is essential that these measures show a good relationship with the perception by the human visual system (HVS).

There are two approaches to measure the image quality: either by subjective methods or by objective methods. Subjective image quality assessment methods measure the overall perceived quality. They are carried out by human subjects. The most commonly used measure is the Mean Opinion Score (MOS). The MOS which is a popular method for assessing image quality involves asking people to quantify their subjective impressions according to a predefined quality scale and is often used as a reference during comparison of image quality assessment methods [1]. On the other hand, objective methods usually use mathematical models for simulating of the results of subjective procedures. They measure the quality index by comparing the original and watermarked images. Most of the existing image quality metrics, such as MSE, PSNR, SSIM [2], and so on, belong to this family. One of the properties required for an image quality metric is that it should show objective scores well correlated with subjective quality scores produced by human observers during quality assessment tests.

In the literature, some image quality assessment methods have been developed based on fuzzy similarity measures to simulate the human visual system and measure the image quality in the same way as humans do. Vlachos et al. [3] propose two novel indices for image quality assessment based on the notion of discrimination information between two fuzzy sets. In [4], Nachtegael et al. show how similarity measures for fuzzy sets have been modified in order to be applied in image processing. And, they also discuss a new application of these measures in the context of color image retrieval, indicating the potential of this class of similarity measures. Zhai et al. [5] introduce subjective assessment method for compressed remote sensing images. This method is applied in assessment of reconstructed remote sensing images which are compressed by using JPEG2000 standard.

In our work, a new image quality evaluator is proposed based on a combination of the fuzzy similarity measures and neural networks. The fuzzy similarity measures are employed as input to neural network.

This paper is organized in the following manner. The fuzzy image processing is explained in Section 2. The image quality method is proposed in Section 3. Experimental results are shown in Section 4. Section 5 concludes this paper.

2. Fuzzy Image Processing

Fuzzy image processing is not a unique theory. It is a collection of different fuzzy approaches to image processing. Nevertheless, the following definition can be regarded as an attempt to determine the boundaries: Fuzzy image processing is the collection of all approaches that understand, represent and process the images, their segments and features as fuzzy sets. The representation and processing depend on the selected fuzzy technique and on the problem to be solved [6].

Fuzzy image processing has three main stages: image fuzzification, modification of membership values, and, if necessary, image defuzzification. The fuzzification and defuzzification steps are due to the fact that we do not possess fuzzy hardware. Therefore, the coding of image data (fuzzification) and decoding of the results (defuzzification) are steps that make it possible to process images with fuzzy techniques. The main power of fuzzy image processing is in the middle step (modification of membership values). After the image data are transformed from gray-level plane to the membership plane (fuzzification), appropriate fuzzy techniques modify the membership values. This can be a fuzzy clustering, a fuzzy rule-based approach, and a fuzzy integration approach and so on [7].

Fuzzy image techniques can be applied in several domains of image processing. In this paper, we show how notions of fuzzy set theory are used in establishing measures for image quality evaluation. Objective quality measures or measures of comparison are of great importance in the field of image processing. Such measures are necessary for the evaluation and the comparison of different algorithms that are designed to solve a similar problem, and consequently they serve as a basis on which algorithm is preferred to the other.

3. Proposed Evaluator

This section describes the main steps of the proposed image quality assessment method for image watermarking techniques. The steps of the proposed evaluator are showed in Fig. 1.

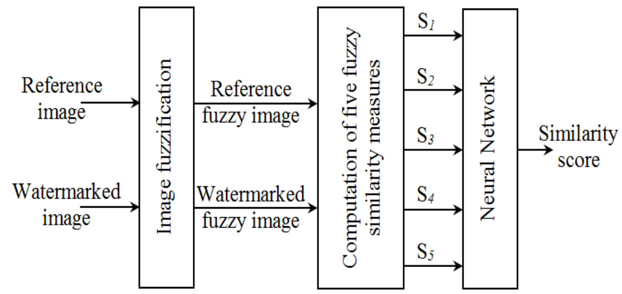

Fig. 1 Flowchart of Proposed method.

3.1 Image fuzzification

Firstly, fuzzification is used to transform original image and watermarked image (crisp data set) into a fuzzy image (fuzzy data set). For fuzzification, we use a fuzzy Gaussian function. Thus in fuzzy set theory, an image X of size $M \times N$ having L gray levels of pixels can be considered as an array of fuzzy singletons, each associated with a membership value. Thus a fuzzy image can be represented as:

$$X = \bigcup_i \bigcup_j x_{ij} \mid \mu(x_{ij}) \tag{1}$$

The membership function of the gray value essentially reflects the membership or belongingness of the pixel to a certain class. The equation for the fuzzy Gaussian function is:

$$\mu(x) = exp(-\frac{(c-x)^2}{2\sigma^2}), \tag{2}$$

where c and σ are the center and width of the fuzzy set, respectively. The default value for center is calculated as the mean of the input data [7]. Fig.2 shows the Gaussian function with different shapes depending on the values of parameters c and σ .

3.2. Features extraction stage

In this evaluator, fuzzy similarity measures are used to calculate the degree of similarity between original and watermarked images. After the original and watermarked images are transformed from gray-level plane to the membership plane (fuzzification), the five of fuzzy similarity measures are computed depending on the membership plane of original and watermarked images.

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012
ISSN (Online): 1694-0814
www.IJCSI.org

These measures (features) are used as inputs to the neural network.

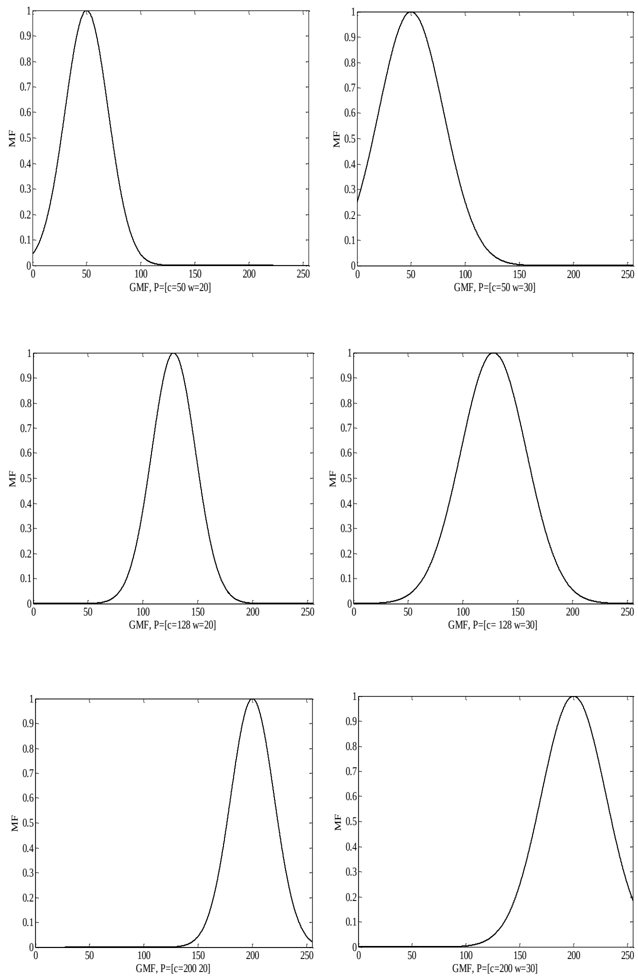

Fig. 2. Different form of Gaussian fuzzy membership functions.

This section describes briefly the five fuzzy logic-based similarity measures used as input to the neural network.

The first similarity measure S_1 is:

$$S_1(A,B) = 1 - \left(\frac{1}{MN} \sum_{x,y=1}^{NM} | \mu_A(x,y) - \mu_B(x,y) |^r \right)^{\frac{1}{r}} \tag{3}$$

where r ≥ 1, $\mu_A(x,y)$ and $\mu_B(x,y)$ are the fuzzy membership values of the pixel of the two images *A* and *B* respectively.

The measures S_2 and S_3 are based on the sigma count as follow:

$$S_2(A,B) = \frac{\sum_{x,y=1}^{NM} min(\mu_A(x,y), \mu_B(x,y))}{\sum_{x,y=1}^{NM} max(\mu_A(x,y), \mu_B(x,y))} \tag{4}$$

$$S_3(A,B) = \frac{\sum_{x,y=1}^{NM} min(1-\mu_A(x,y), \mu_B(x,y))}{\sum_{x,y=1}^{NM} max(1-\mu_A(x,y), \mu_B(x,y))} \tag{5}$$

For the similarity measures S_4 and S_5 it is less obvious to give an intuitive interpretation:

$$S_4(A,B) = 1 - \frac{\sum_{x,y=1}^{NM} | \mu_A(x,y) - \mu_B(x,y) |}{\sum_{x,y=1}^{MN} | \mu_A(x,y) + \mu_B(x,y) |} \tag{6}$$

$$S_5(A,B) = \frac{1}{MN} \sum_{x,y=1}^{NM} \frac{min(\mu_A(x,y), \mu_B(x,y))}{max(\mu_A(x,y), \mu_B(x,y))} \tag{7}$$

3.3. Neural network approach

Artificial Neural Network (ANN) is a powerful data modeling tool that is able to capture and represent complex input/output relationships. The motivation for the development of neural network technology stemmed from the desire to develop an artificial system that could perform "intelligent" tasks similar to those performed by the human brain.

In our work, we propose a method to predict the MOS of human observers using an ANN. Here the ANN is designed to predict the image quality using five features extracted from the reference and watermarked images. Once the features are extracted, we combine these features by an ANN to obtain an index quality.

The first task is to determine the size of network which can approximate the image quality with good results. Therefore, many different sizes of networks were tested during this experiment. Several networks were first trained and then tested with the same test set. The difference among the networks tested was the number of neurons in the hidden layers. It was found that small network does not perform very well, i.e., the network approximation had too large error to the target. As the network size is increased, by adding more hidden neurons,

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012
ISSN (Online): 1694-0814
www.IJCSI.org

the error becomes smaller; however, the training time also increased. Large networks were found to be too slow to train and without any drastic changes on the approximation error. How the error depending on the size of the network are shown in Fig.3. The mean and max error between the MOS and the network prediction based on different sizes is obtained and plotted.

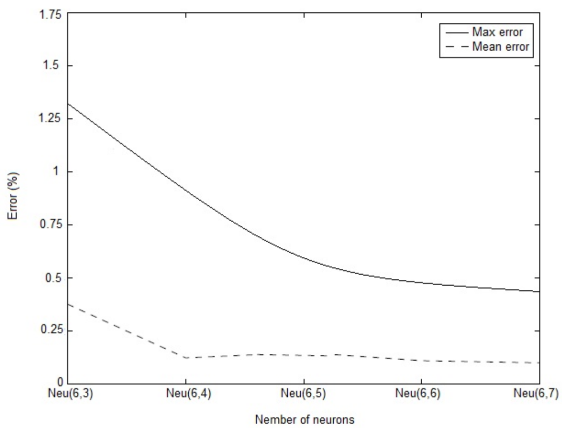

Fig. 3. Error depending on number of neurons.

As Fig.3 shows, the error goes down with the amount of neurons. However, the mean error falls slowly by increasing the amount of neurons. Furthermore, the max error does not improve with much when the amount of neurons gets large. A network with the following size was found as the best solution: 5 inputs, 6 neurons the first hidden layer, 5 neurons the second hidden layer, and 1 output. The inputs correspond to the extracted features and the output is the predicted subjective score (MOS). The ANN model is presented in Fig. 4.

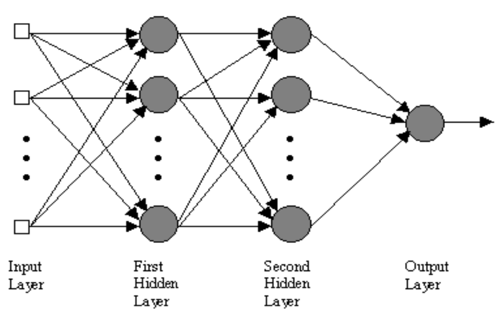

Fig. 4. Artificial neural network model.

The input and output values are first normalized to the range [0, 1]. For the output, values 0 and 1 denote worst and best quality respectively. The unibipolar sigmoid function is used as activation function in the hidden and output layers.

The back propagation learning algorithm (supervised learning) is used to train the ANN. In order to train our ANN and test the efficiency of the obtained model, we used the IVC-FourierSB database [8] as image database. Some original images are presented in Fig. 5.

Fig. 5. Some original images of the IVC-FourierSB database.

This database provides the original image and its watermarked version with their MOS values. The database consists of 5 original gray-scale images, 210 watermarked images were generated from watermark addition in 6 different frequency (Fourier) sub-bands of various frequency range, with 7 embedding strengths. The subjective evaluations were prepared using a DSIS (Double Stimulus Impairment Scale) method with five equal regions marked with adjectives "Bad", "Poor", "Fair", "Good" and "Excellent" as it is shown in Table 1.

Table 1. Five-grade scale recommended by ITU [1]

Rating	*Quality*	*Impairment*
5	Excellent	Imperceptible
4	Good	Perceptible, but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

Once the database is chosen, we divide it into training set and testing set. Network was trained on 70% of the images from the available set, and the other 30% were used to test the performance of the network. The desired output of the neural network is the MOS value of the test image.

The ANN training procedure is terminated either when the training error is less than 10-3 or when 10000 iterations have been executed. The training error used is the mean square error which is the average squared error between actual output of the network and desired output for all the training patterns. The training data consists of a set of sample inputs (objective metrics) and the desired outputs (MOS) corresponding to those inputs. The learning rate and momentum term for training of NN were chosen as 0.1~0.15 and 0.8~0.9, respectively. The learning rate and momentum term are the parameters of the well-known back propagation algorithm. According to the some preliminary runs and the suggestion of many previous studies these were set. Normally, the weight values are initialized as random numbers between -1 and 1. Once trained, the network weights are saved and can be used to compute output values for new input samples.

4. Experimental results

After the learning process, we test the efficiency of the proposed way using the test set. To test the ability of the neural network to predict the MOS, its performance is assessed by using Pearson and Spearman correlations with MOS values used as a reference. The network is tested with a new test set which the network has not seen before. Table 2 shows the Pearson and Spearman correlations between the MOS and the network predicted of the MOS.

Table 2: Pearson and Spearman correlations.

Results of the proposed evaluator	
Correlation type	***Obtained value***
Pearson	0.9726
Spearman	0.9682
Outlier	1%

The Pearson and Spearman correlation coefficients obtained for the all test images are respectively equal to 0.9726 and 0.9682.

To better visualize for the obtained results, the polynomial fitting function is used to summarize the visual relationship between the actual MOS and the network predicted of the MOS as shows in Fig. 6.

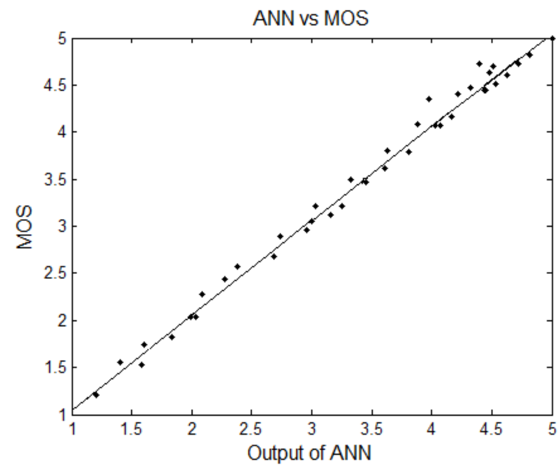

Fig.6. MOS vs. network prediction.

The straight line in Fig. 5 makes it easier to visualize the correlation between the target and the network output. It can clearly be seen that our network approximates the MOS with good results from the graphs. The ANN is able to give accurate prediction of the MOS for new values of the inputs (fuzzy similarity measures) taken from testing samples.

The proposed evaluator is also compared to three full reference measures, namely the UIQI [9] (pixel-based), SSIM [10] (structural-based), C4 [11] (correlation measures-based) and the VIF [12] (mutual information-based) see Table 3.

Table 3: Comparison of image quality assessment methods.

Metrics	***Correlation Coefficients***	
	Pearson	***Spearman***
UIQI	0.835	0.870
SSIM	0.894	0.938
C4	0.925	0.926
VIF	0.916	0.911
Our method	0.972	0.968

5. Conclusion

In this paper, a new approach for image quality evaluation is introduced. The combination of the fuzzy logic-based similarity measures through a neural network offers good results in terms of correlation with subjective evaluation as expressed in the image quality database. Experimental results show that a neural network can be trained to accurately predict the MOS values using five fuzzy similarity measures computed from the reference and watermarked images.

Acknowledgments

The authors would like to thank Professor William Puech at LIRMM, Universite Montpellier II in France for his educational and scientific support.

References

[1] K. Seshadrinathan, R. Soundararajan, and A. C. Bovik, “Study of Subjective and Objective Quality Assessment of Video,” IEEE Transactions on Image Processing, June 2010, vol. 19(6), pp. 1427-1441.
[2] Z. Wang, A. C. Bovik, H. Sheikh, and E.Simoncelli, “Image Quality Assessment: From Error Visibility to Structural Similarity,” IEEE Transactions on Image Processing, April 2004, vol. 13(4), pp. 600-612.
[3] K. Vlachos and D. Sergidis, “Image Quality Indices Based on Fuzzy Discrimination Information Measures,” Aristotle University of Thessaloniki, Faculty of Technology, Greede, 2010.
[4] M. Nachtegael, S. Schulte, V. Witte, T. M´elange, and E. Kerre, “Image Similarity – From Fuzzy Sets to Color Image Applications,” LNCS, 2007, vol. 4781, pp. 26–37.

[5] L. Zhai and X. Tang “Fuzzy Comprehensive Evaluation Method and its Application in Subjective Quality Assessment for Compressed Remote Sensing Images,” Chinese Academy of Surveying and Mapping, 2006.
[6] H. Tizhoosh, Fuzzy Image Processing, Springer-Verlag, 2002.
[7] T. Acharya and A. Ray, Image Processing Principles and Applications, Published by John Wiley & Sons, Inc., Hoboken, New Jersey, pp.242–246, 2005.
[8] P. Le Callet and F. Autrusseau, “A Robust Image Watermarking Technique based on Quantization Noise Visibility Thresholds,” Elsevier Signal Processing, June 2007, vol. 87(6), pp. 1363–1383.
[9] Z. Wang and A. C. Bovik, “Universal Image Quality Index,” IEEE SP letters, March 2002, vol. 9, pp. 81-84.
[10] Z. Wang, E. Simoncelli and A. C. Bovik, “Multi-Scale Structural Similarity for Image Quality Assessment,” IEEE Asilomar Conference on Signals, Systems and Computers, 2003, vol. 2, pp. 1398-1402.
[11] I. Avcıbas, B. Sankur and K. Sayood, “Statistical Evaluation of Image Quality Measures,” Journal of Electronic Imaging, 2002, vol.11, no. 2, pp. 206–223.
[12] H. Sheikh and A. C. Bovik, “Image Information and Visual Quality,” IEEE Transactions on Image Processing, 2006, vol.15, no.2, pp. 430-444.

Methaq Gaata received the B.S. in computer sciences from Faculty of Sciences, Al- Mustansiriyah University, IRAQ in 2002. He received the M.Sc. in computer sciences from computer sciences department, Technology University, IRAQ in 2005. He now preparing for Ph.D. degree in computer sciences. His interests include signal processing, image processing, watermarking and information security.

Sattar Sadkhn is a member of the IEEE and Chairman of IEEE IRAQ Section, Chairman of URSI - IRAQ Community, Editor -in-Chief of IJACT in South Korea, Member of Scientific Research Committee-IMPACT- Malaysia. His interests include signal processing, image processing, watermarking, digital communication systems and Information Security.

Saad Hasson is currently a professor and head of computer sciences Department, Faculty of Sciences, Babylon University, IRAQ. His research interests focus on digital signal processing, image processing and mathematical modeling.