PERCEPTUAL IMAGE QUALITY ASSESSMENT BASED ON BAYESIAN NETWORKS

Ronaldo de Freitas Zampolo and Rui Seara

LINSE – Circuits and Signal Processing Laboratory
Department of Electrical Engineering
Federal University of Santa Catarina
88040-900 – Florianópolis – SC – Brazil

ABSTRACT

This paper addresses the issue of perceptual image quality assessment. By using Bayesian Networks, we propose a Bayesian *composed quality measure* (B-CQM). This metric can assess quality in images degraded by combined noise injection and frequency distortion. It presents some advantages with respect to the original CQM approach, such as to uphold the stochastic nature of the subjective quality assessment and easier inclusion of the effect of new experimental data in the metric model by just updating its probability tables. Some examples are provided in order to verify the behavior of the proposed metric.

1. INTRODUCTION

The performance of image processing systems usually depends on some quality assessment, which is associated with the used system model. Commonly criteria based on *mean-square error* (MSE) are used, even though the mismatch between MSE and human visual perception is well known [1, 2]. The previous aspects, along with the fact that many systems are intended for human consumers, have been stimulating research efforts in perceptual image quality metrics, i.e., objective measures that reflect in some sense the human perception of image quality. Early contributions to this field can be traced back to the research of Weber, Fechner and Michelson [3] on quantifying image contrast. Nevertheless, since the image patterns used in their definitions are rather simple, the so-called *Weber-Fechner* and *Michelson contrast* measures do not adequately describe perceived contrast in real-world images [2–4]. Another approach has been proposed by Hess and Pointer, in which a contrast metric is defined in the frequency domain, taking into account the perception dependencies on spatial frequency. However, the *Hess and Pointer metric* cannot capture the local nature of contrast changes [3, 4]. As an attempt to overcome such problems, Peli has developed a metric that combines sequence and frequency domain information [3]. Experimental results indicate a superior performance of Peli's approach as compared with other strategies (Michelson, King-Smith and Kulicowski contrast measures) [5]. In [4], the authors have proposed a strategy to assess the quality of restored images, in which the frequency distortion and noise injection are considered decoupled psychovisual effects. In order to quantify those effects, a *noise quality measure* (NQM) and a *distortion measure* (DM) have been developed. The NQM outperforms the *signal-to-noise ratio* (SNR) as well as *weighted SNR* (WSNR). A recent paper [6] has shown experimental evidences that DM does not perform well. Thus, the authors have proposed a *distortion quality measure* (DQM), which consists of the NQM block function with modified inputs (Fig. 1). Furthermore, a *composed quality measure* (CQM) is derived as a function of DQM and NQM from experimental data. The CQM has been used to assess quality of images degraded both by frequency distortion and noise injection. Although a subjective quality assessment presents an important factor of uncertainty, the CQM has been represented deterministically in the same way as other image quality measures. Such an uncertainty stems from the peculiarities in the neurophysiology and cultural aspects of the enrolled subjects. This paper proposes an alternative approach for the CQM by using Bayesian networks [7] aiming to model this inherent uncertainty in image quality assessment. This paper is outlined as follows. Section 2 summarizes the CQM features and the experimental procedure used in [6]. Section 3 presents the basic concepts about Bayesian netwoks with emphasis on those aspects that are of major interest for the intended application. Section 4 details the proposed approach as well as presents some experimental results. Finally, concluding remarks are presented in Section 5.

2. IMAGE QUALITY ASSESSMENT EXPERIMENT

This section addresses the *composed quality measure* (CQM) as well as the experimental procedure used to derive it, both defined in [6]. The referred procedure is summarized as follows:

i) a set of 81 images, which are degraded by frequency distortion and noise injection in different levels, is obtained from the original "Lena" image;

ii) the NQM and DQM for each test image is assessed as depicted in Fig. 1, where x, $\hat{x}$ and $\hat{x}'$ are the original image, test image and a noiseless version of the test image, respectively;

iii) each of the images is randomly ranked seven times by subjects (20-50 years old, both sexes, corrected or normal vision), according to Table 1;

iv) a 15 inch monitor with 800×600 pixel resolution is used;

v) the viewing distance is about 60 cm.

Then a set of candidate functions is optimized to fit the resulting surface of averaged classifications (Fig. 2). The bidimensional Gaussian function with non-correlated variables (1) has been chosen among the considered functions due to its correlation coefficient (0.935770) and variance (0.174635) values with respect to experimental data.

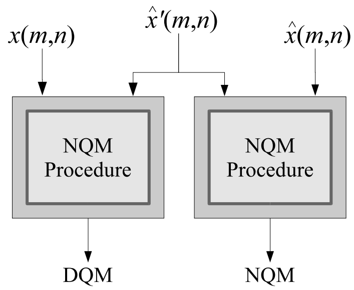

Fig. 1. Blocks for assessing DQM and NQM (x, $\hat{x}$ and $\hat{x}'$ are the original image, test image and a noiseless version of the test image, respectively).

Table 1. Values used to rank images

Rank	Perceived Quality
4	Excellent
3	Good
2	Regular
1	Bad
0	Unacceptable

The CQM model [6] is given by

$$CQM = k \cdot \exp\left[-\frac{(N-\mu_1)^2}{2\sigma_1^2} - \frac{(D-\mu_2)^2}{2\sigma_2^2}\right], \qquad (1)$$

where N and D denote NQM and DQM, respectively; μ_1, μ_2, σ_1^2 and σ_2^2 denote the means and variances of N and D, respectively; and k is a normalization constant used to keep CQM values within the 0 to 4 range (Table 1). In this deterministic approach, after a CQM model has been obtained, the stochastic aspect of the experiment becomes implicit. The data variance has been estimated, but in practical terms it is no longer considered. Furthermore, if additional experimental data are collected, the whole optimization procedure must be repeated in order to estimate the new μ_1, μ_2, σ_1 and σ_2 in (1). Such an approach can be justified by the need of having a suitable quality assessment function to be used for parameter optimization in image processing systems.

In the next sections, we propose a Bayesian CQM (B-CQM) that permits an easier inclusion of new statistical data, preserves the random aspects of a subjective image quality assessment and is also useful in optimization tasks.

3. BAYESIAN NETWORK BACKGROUND

Bayesian networks are directed acyclic graphs (DAG), which are commonly used in expert systems and artificial intelligence applications due to their capability for modeling uncertainty by randomness and easily responding to changing conditions. The knowledge of a particular field can be modeled by a Bayesian network, in which each node represents a meaningful variable. From the experimental data collected and/or stochastic assumptions, *a priori* conditional probabilities are assessed to define the Bayesian network. When an evidence is observed, the *a priori* conditional probabilities are used for determining the *a posteriori* probabilities. Thus, a Bayesian network explicitly represents the conditional dependencies between different knowledge components and provides a graphic visualization of the knowledge modeled [7]. Bayesian network basis rests on the well-known Bayes' theorem [8]:

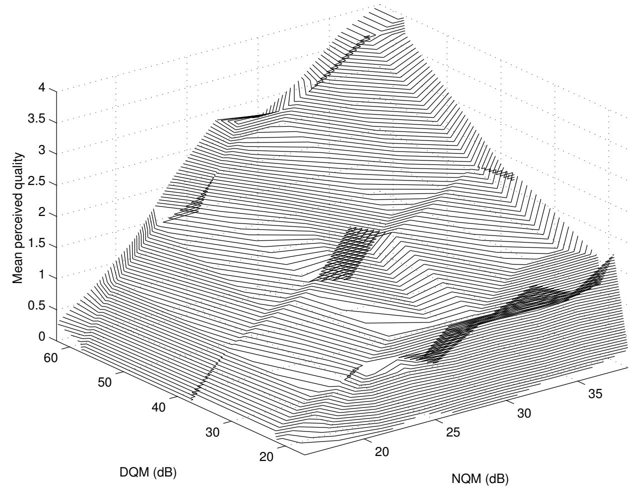

Fig. 2. Mean perceived quality as a function of the NQM and DQM.

Table 2. Quality probabilities

Perceived Quality (Q)	$P(Q)$
Excellent	0.080
Good	0.130
Regular	0.230
Bad	0.320
Unnacceptable	0.240

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B|A_1) + P(B|A_2) + \cdots + P(B|A_n)}, \qquad (2)$$

where $P(A)$ is the probability of an event A, the set $[A_1 A_2 \ldots A_n]$ is a partition of the *certain event* and B is an arbitrary event.

Details about Bayesian networks can be found in [7, 9].

4. BAYESIAN COMPOSED QUALITY MEASURE

In order to use Bayesian networks for deriving a B-CQM, we have used the experimental data obtained in [6] for the original CQM, dividing both the NQM and DQM ranges into nine non-overlapping sections. The limits of the referred sections have been chosen in such a way that each section contains approximately the same number of test images. The non-overlapping sections are labeled by using numbers in an increasing order following the increasing values of NQM and DQM.

Then, the quality probability is estimated (Table 2) as well as the *a priori* conditional probabilities for the NQM and DQM sections (Tables 3 and 4, respectively). Note that the referred probabilities have been estimated using relative frequencies obtained from the experimental data and not from a stochastic model. We have used the *Netica*^{TM} *Application v1.12* [10] along with the probability values shown in Tables 2, 3 and 4 to implement the

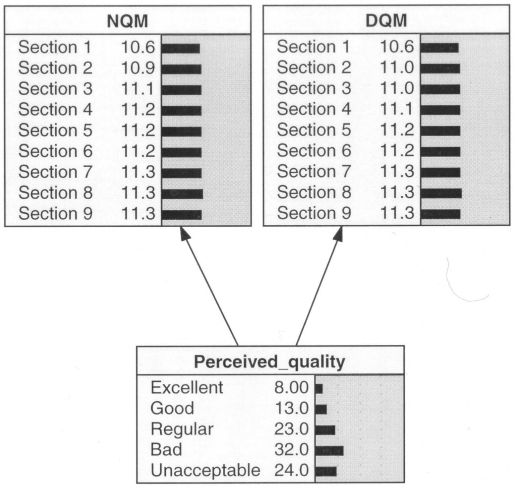

Fig. 3. Bayesian network initial state.

Table 3. NQM *a priori* probabilities – $P(S_N|Q)$

Section # (S_N)	Excell.	Good	Reg.	Bad	Unnaccep.
1	0.000	0.000	0.000	0.028	0.403
2	0.000	0.000	0.016	0.179	0.201
3	0.000	0.000	0.101	0.218	0.076
4	0.000	0.014	0.186	0.162	0.063
5	0.000	0.264	0.116	0.101	0.076
6	0.023	0.236	0.163	0.084	0.063
7	0.163	0.236	0.147	0.078	0.042
8	0.372	0.181	0.109	0.078	0.042
9	0.442	0.069	0.163	0.073	0.035

Table 4. DQM *a priori* probabilities – $P(S_D|Q)$

Section # (S_D)	Excell.	Good	Reg.	Bad	Unnaccep.
1	0.000	0.000	0.000	0.028	0.403
2	0.000	0.000	0.132	0.156	0.125
3	0.000	0.000	0.016	0.240	0.125
4	0.000	0.000	0.140	0.190	0.076
5	0.000	0.125	0.202	0.112	0.056
6	0.023	0.083	0.209	0.112	0.063
7	0.326	0.250	0.124	0.039	0.056
8	0.349	0.264	0.093	0.056	0.049
9	0, 302	0.278	0.085	0.067	0.049

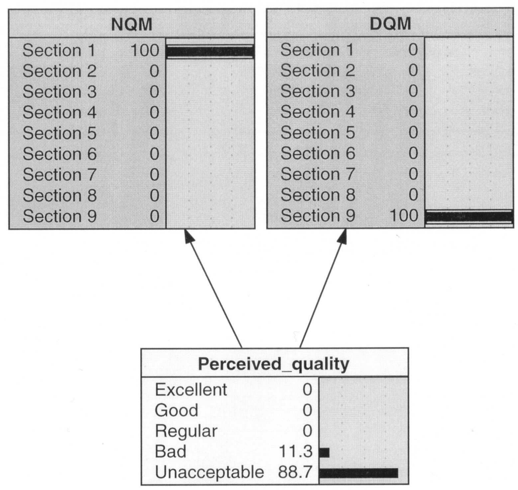

Fig. 4. B-CQM results for a test image highly contaminated by noise and slightly frequency distorted.

B-CQM approach. The resulting network initial state is shown in Fig. 3. Thus, given an image with particular values of NQM and DQM, we can assess the *a posteriori* perceived quality probability by selecting the proper sections as in the examples depicted in Figs. 4, 5 and 6. In Fig. 4, a test image is highly contaminated by noise (the assessed NQM belongs to Section 1) and slightly frequency distorted (the assessed DQM belongs to Section 9). The resulting network output node indicates an expectation of 88.7% of this image to be classified as *unacceptable* and 11.3% as *bad*. As the noise contamination decreases, maintaining the same frequency distortion condition, we can verify that the resulting outputs show a change in expectation bias towards better quality (Figs. 5 and 6). The B-CQM output predicts the subjective quality assessment from the experimental data available along with NQM and DQM values, showing the probabilities of a particular image being classified according to the quality rank and experimental conditions adopted (Table 1 and Section 2, respectively).

The effect of new experimental data can be easily included in the metric model by updating the probability tables, which is much simpler than re-optimizing the parameters of the CQM expression (1) for any change in the data set.

The B-CQM can be applied to parameter optimization in image processing systems by defining a performance measure as a function of the B-CQM outputs. The following linear multivariable function (3) is shown as an example of a performance measure based on B-CQM.

$$J = \sum_{Q=0}^{4} k_i P(Q|S), \tag{3}$$

for $S = \left\{ S_N \big|_{S_a} \right\} \cap \left\{ S_D \big|_{S_b} \right\}$, where Q, S_a and S_b denote the quality rank and the assessed NQM and DQM sections, respectively; k_i are weighting constants.

For validation purposes, 8 images have been randomly chosen from the set used in the CQM derivation experiment and subjected to quality assessment, following the procedure previously described in Section 2 items (iii) to (v). The correlation between the B-CQM output (predicted quality) and resulting subjective quality can be verified in Tables 5 and 6, respectively. In both tables, the boldfaced quantities point out the perceived quality classes of larger occurrence. Table 6 shows the quality assessment provided by 10 subjects with the same characterization of the group used in the CQM experiment, but distinct from that one. Through the result analysis, we can confirm the good agreement between the proposed model (B-CQM) and the perceived image quality, in terms of the most expected quality classes, considering the features of the experiment carried out.

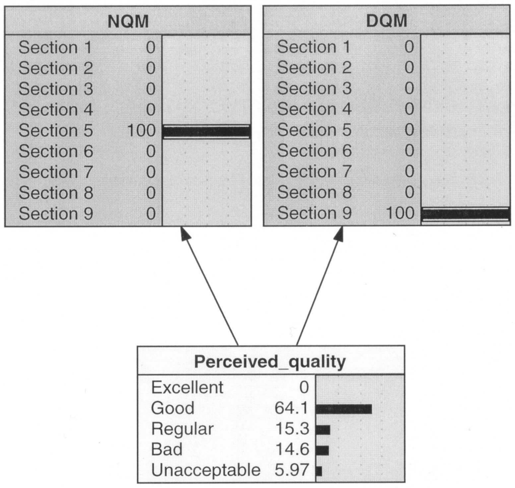

Fig. 5. B-CQM results for a test image moderately contaminated by noise and slightly frequency distorted.

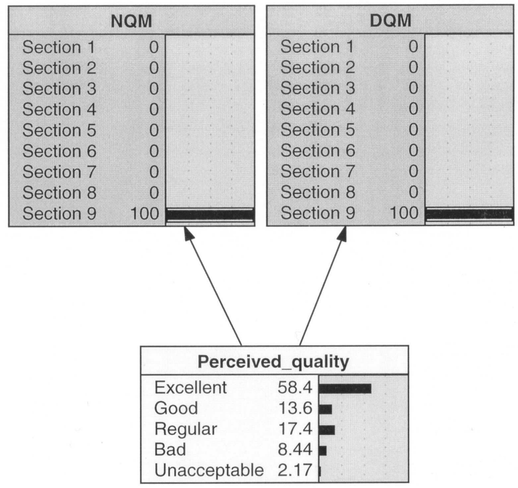

Fig. 6. B-CQM results for a test image almost noiseless and slightly frequency distorted.

Table 5. B-CQM predicted quality (%)

Image #	Excell.	Good	Reg.	Bad	Unnaccep.
1	3.76	**57.50**	21.60	12.20	4.96
2	0.00	5.94	**49.20**	35.80	9.09
3	0.00	0.00	0.00	14.20	**85.80**
4	0.00	0.00	**45.50**	**44.40**	10.00
5	0.00	0.00	33.10	**54.50**	12.40
6	0.00	25.90	**44.50**	24.60	4.94
7	0.00	0.00	5.08	**78.40**	16.50
8	0.00	0.00	0.00	19.20	**80.80**

Table 6. Validation experiment

Image #	Excell.	Good	Reg.	Bad	Unnaccept.
1	3/10	**6/10**	1/10	0/10	0/10
2	0/10	2/10	**6/10**	2/10	0/10
3	0/10	0/10	0/10	1/10	**9/10**
4	0/10	0/10	1/10	**9/10**	0/10
5	0/10	0/10	1/10	**9/10**	0/10
6	0/10	**5/10**	**5/10**	0/10	0/10
7	0/10	0/10	0/10	**8/10**	2/10
8	0/10	0/10	0/10	1/10	**9/10**

5. CONCLUSIONS

In this paper a Bayesian approach for the *composed quality measure* (B- CQM) is proposed. The B-CQM presents some interesting features in terms of modeling: the stochastic nature of visual quality perception remains evident and new experimental data can be included by just updating the *probability tables*. The B-CQM can also be used for parameter tuning in image processing systems, by defining performance measures as a function of B-CQM outputs. In addition, an experimental validation of the B-CQM is presented.

6. REFERENCES

[1] P. C. Teo and D. J. Heeger, "Perceptual image distortion," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, Nov. 1994, pp. 982–986.

[2] N. J. Nadenau, J. Reichel, and M. Kunt, "Performance comparison of masking models based on a new psychovisual test method with natural scenery stimuli," *Signal Processing: Image Communication*, vol. 17, pp. 807–823, Nov. 2002.

[3] E. Peli, "Contrast in complex images," *J. Opt. Soc. Am. A*, vol. 7, no. 10, pp. 2032–2040, 1997.

[4] N. Damera-Venkata, T. D. Kite, W. S. Geisler, *et al.*, "Image quality assessment based on a degradation model," *IEEE Trans. Image Processing*, vol. 9, no. 4, pp. 636–650, Apr. 2000.

[5] E. Peli, "In search of a contrast metric: Matching the perceived contrast of gabor patches at different phases and bandwidths," *Vision Res.*, vol. 37, no. 23, pp. 3217–3224, Oct. 1990.

[6] R. F. Zampolo and R. Seara, "A measure for perceptual image quality assessment," in *Proc. IEEE Int. Conf. Image Processing*, Sept. 2003.

[7] Bayesian networks. I.D.I.S Artificial Intelligence Resources. Last access in Dec. 2003. [Online]. Available: http://excalibur.brc.uconn.edu/ baynet/

[8] A. Papoulis, *Probability, Ramdom Variables, and Stochastic Processes*, 3rd ed., ser. Electrical & Electronic Engineering. McGraw-Hill, 1991.

[9] A brief introduction to graphical models and bayesian networks. I.D.I.S Artificial Intelligence Resources. Last access in Dec. 2003. [Online]. Available: http://www.ai.mit.edu/ murphyk/Bayes/bayes.html

[10] Netica™ application v1.12. Norsys Software Corp. Last access in Dec. 2003. [Online]. Available: www.norsys.com