Classification of Image Distortions using Image Quality Metrics and Linear Discriminant Analysis

Aladine Chetouani[1], Mohamed Deriche[2], and Azeddine Beghdadi[1]

[1]L2TI, Institut Galilée, Université Paris13, France
[2]EE. Dept., King Fahd University of Petroleum & Minerals, Saudi Arabia
email: aladine.chetouani@univ-paris13.fr, mderiche@kfupm.edu.sa, azeddine.beghdadi@univ-paris13.fr,

ABSTRACT

Numerous Image Quality Measures (IQMs) have been proposed in the literature. While some are based on structural analysis of images, others rely on the characteristics (or limitations) of the Human Visual System (HVS). However, none of the existing IQMs is shown to be robust across all types of degradations. Indeed, some IQMs are more efficient for a given artifact (such as blurring or blocking) but inefficient for others. In this paper, we propose to circumvent this limitation by adding a preprocessing step before measuring image quality. We propose to detect the type of the degradation contained in the image, then use the most "relevant" IQM for that specific type of degradation. The classification of different degradations is performed using simple Linear Discriminant Analysis (LDA). The performance of the proposed method is evaluated in terms of classification accuracy across different types of degradations and shown to outperform different IQMs when used independently of the degradation type.

1.　INTRODUCTION

During the last decade, we have witnessed an increasing demand for quality multimedia material. This is essentially due to the development of advanced image/video production technologies. Indeed, the progress achieved in these domains is unprecedented. Despite such a progress, quantifying and reducing image degradation continues to be a challenging problem. A typical example is that of image degradation due to blocking effects in JPEG compression. This artifact manifests in the image as disturbing horizontal and vertical boundaries.

In recent years, research efforts in image quality have resulted in the development of a number of IQMs. The most common metric is the Peak Signal to Noise Ratio (PSNR). Unfortunately, PSNR provides poor results in terms of correlation with subjective measures such as the Mean Opinion Score (MOS). Some methods were proposed to improve the PSNR by adding Human Visual System (HVS) characteristics such as the PNSR-HVS [1]. This metric takes into account the limitations of the HVS in discerning fine details in an image. This limitation is simulated using the Contrast Sensitivity Function

(CSF). A more recent version was developed which takes into account masking effects [2] in the DCT domain [3].

Some IQMs are essentially based on Human Visual System (HVS) characteristics such as the VDP [4]. However, this IQM is complex. Other metrics, such as the SSIM [5], use structural characteristics, or mutual information concepts [6,14] to quantify image quality.

Despite all these available IQMs, there is no metric aim that can predict or measure image quality accurately across all degradations. Indeed, the efficiency of a given IQM may be excellent for a given degradation but poor for others. This is essentially due to the fact that generally Full Reference (FR) IQMs cannot take into account the type of the visual distortion contained in an image.

To overcome this limitation, one of the solutions is to identity first the degradation type contained in image then measure the quality of the image, using the most appropriate IQM. Here, we do not focus on the particular artifacts such as blocking effects or ringing effects, but use a statistical framework that covers all possible degradations.

In this paper, we propose to use Linear Discriminant Analysis (LDA) to model the statistical behaviour of a number of estimated IQMs (seen as features) across all types of degradations. After a training stage, the LDA is used as a classifier where the classes are the different types of degradations.

The paper is organized as follows: Section 2 describes the proposed method and the image database used in our experiments. Our experimental results are discussed in section 3 followed by some concluding remarks in section 4.

2.　PROPOSED METHOD

The overall block diagram of the proposed algorithm is presented in Fig. 1. We propose to detect the type of degradation contained in a given test image before quantifying the quality of such image using an appropriate IQM.
We show the importance of knowing the type of distortion contained in an image before measuring its quality though the following simple experiment.
We selected 2 degradation types Blur and JPEG compression. For each type of degradation, we ranked the different IQMs using the Pearson's Correlation Coefficient (PCC) between each IQM index and the Mean

Opinion Score (MOS). The *IQMs* used here cover structural analysis, HVS-based, MI-based measures and classical MSE-based metrics. These are: *VIF, VIFP* [6], *VIF, VIFP* [6], *PSNR-HVS* (PSNRH) [1], *PSNRHVS-M* (PSNRM) [3], *SSIM* and *MSSIM* [5], *UQI* [7], *IFC* [8], *WSNR* [9], *VSNR* [10], *NQM* [11], *XYZ* [12], *SNR* and *PSNR*. Other measures can also be used.
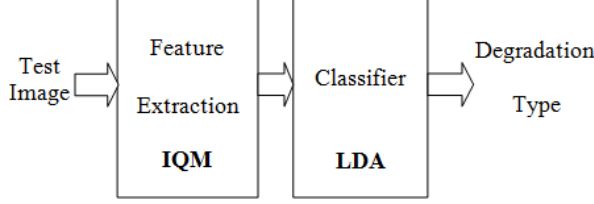

Figure 1 – Flowchart of the proposed method.

The PCC expressed in this context is given:

$$CORR_{ij} = \frac{\sum_{k=1}^{K} (IQM(k)_{ij} - \overline{IQM_{ij}}).(MOS(k)_i - \overline{MOS_i})}{\sqrt{\sum_{k=1}^{K} (IQM(k)_{ij} - \overline{IQM_{ij}})^2}.\sqrt{\sum_{k=1}^{K} (MOS(k)_i - \overline{MOS_i})^2}}$$
(1)

where i and j designate the ith degradation and the jth IQM, respectively. The index k stands for the kth image, and K is the total number of images.

Table 1 shows the final IQM ranking. Note that the ranking is different between the two types of degradations. Indeed, the best metrics for blur and JPEG compression are respectively VIFP and PSNRH.

Table 1. IQM ranking for blur and JPEG distortions

IQM ranking	Degradation type	
	Blur (8)	**JPEG (10)**
1	VIFP	PSNRH
2	VIF	PSNRM
3	WSNR	VIF
4	VSNR	WSNR
5	PSNRM	NQM
6	PSNRH	VIFP
7	SSIM	VSNR
8	UQI	SSIM
9	NQM	XYZ
10	IFC	SNR
11	SNR	UQI
12	XYZ	IFC

The estimated PCCs obtained for VIFP and PSNRH for each degradation, are shown in Table 2. The results confirm that it would be more efficient to first identify the type of distortion before measuring the image quality.

Table 2. PCC for blur & JPEG distortions.

Degradation type	Person correlation	
	VIFP	**PSNRH**
JPEG	**0.9416**	0.9120
Blur	0.9188	**0.9543**

In what follows, we present the image database used in our experiments. Then, after explaining the feature extraction stage, modelling using LDA is discussed.

2.1 The TID 2008 image database

Figure 2: Sample images from the TID2008 database.

In order to train and test the efficiency of the proposed approach, we need both training and test sets. TID 2008 image database [13] is chosen for this purpose. This image database contains 17 types of degradations with 100 images per distortion using 25 reference images (i.e. 4 distortion levels per image and per degradation). Fig.2 shows some reference images from the TID 2008 database. Table 3 lists the degradation types available in the database. The MOS values for all the degraded images are also available.

Table 3. Types of degradation in the TID2008 database

Degradation	Type
1	Additive Gaussian noise
2	Additive noise in color components
3	Spatially correlated noise
4	Masked noise
5	High frequency noise
6	Impulse noise
7	Quantization noise
8	Gaussian blur
9	Image denoising
10	JPEG compression
11	JPEG2000 compression
12	JPEG transmission errors
13	JPEG2000 transmission errors
14	Non eccentricity pattern noise
15	Local block-wise distortions of different intensity
16	Mean shift (intensity shift)
17	Contrast change

For each degradation, we divide the images from the database into two sets:

- Training set: used only in the learning process.
- Testing set: used only to test the efficiency of the proposed method.

2.2 Features extraction

In this paper, we propose to use the different IQMs cited above (see section 2) as features of inetrest. We assume that we have access to both original and degraded images (i.e. Full Reference metrics are thus considered). Some of the IQMs are HVS-based (such as the PSNRHVS)

while others are based on pixel-wise differences (MSE-based metrics), or local structural information like the SSIM. The total number of features (or IQMs) used is 12, these were listed in Table 1.

2.3 Classification using LDA

LDA has been used successfully in many applications including face recognition, microarray gene expression data analysis and others [15].

In this work, we propose to use the Linear Discriminant Analysis (LDA) approach for identifying the different degradations. Each degradation is considered as a class with a total of M=17 classes. The IQMs estimated from a given image are seen as the extracted feature vectors (of dimension n=12).

Linear Discriminant Analysis (LDA) is a popular method for dimensionality reduction and classification that projects high-dimensional data onto a low dimensional space where the data achieves maximum class separability. The resulting features in LDA are linear combinations of the original features, where the coefficients are obtained using a projection matrix **W**. The optimal projection or transformation is obtained by minimizing *within-class-distance* and maximizing *between-class-distance* simultaneously, thus achieving maximum class discrimination. The optimal transformation is readily computed by solving a generalized eigenvalue problem.

The initial LDA formulation, known as the Fisher Linear Discriminant Analysis (FLDA) was originally developed for binary classifications. The key idea in FLDA is to look for a direction that separates the class means well (when projected onto that direction) while achieving a small variance around these means. Discriminant Analysis is generally used to find a subspace with *M* - 1 dimensions for multi-class problems, where *M* is the number of classes in the training dataset.

Contrary to Principal Component Analysis (PCA) which considers each observation vector as a class on its own, the objective of LDA is to perform dimensionality reduction while preserving as much of the class discriminatory information as possible. Linear Discriminant Analysis searches for those vectors in the underlying space that best discriminate among classes (rather than those that best describe the data as in PCA).

More formally, for the available samples from the database, we define two measures: (i) *within-class* scatter matrix, given by:

$$S_w = \sum_{j=1}^{M} \sum_{i=1}^{N_i} (x_i^j - \mu_j)(x_i^j - \mu_j)^T$$

where x_i^j (dimension nx**1**) is the ith sample vector of class j, μ_j is the mean of class j, M is the number of classes, and N_i is the number of samples in class j.

The second measure (ii) is called between-class scatter matrix and is defined as:

$$S_b = \sum_{j=1}^{M} (\mu_j - \mu)(\mu_j - \mu)^T$$

where μ is mean vector of all classes.

The goal is to find a transformation **W** that maximizes the between-class measure while minimizing the with-in-class measure. One way to do this is to maximize the ratio det(S_b)/det(S_w). The advantage of using this ratio is that if S_w is a non-singular matrix then this ratio is maximized when the column vectors of the projection matrix, **W**, are the eigenvectors of $S_w^{-1}.S_b$ [15]. It should be noted that: (i) there are at most *M-1* nonzero generalized eigenvectors, and so an upper bound on reduced dimension is M-1, and (ii) we require at least *n* (size of original feature vectors) + M samples to guarantee that S_w does not become singular.

3. RESULTS AND DISCUSSION

To evaluate the efficiency of the proposed method for degradation classification, numerous experiments were carried covering over 400 natural images different from those used during the learning stage.

The experimental procedure is quite simple and requires the original and distorted images. In particular, the following main steps are performed:

1. The images from the training set are processed to obtain the feature vectors. Such feature vectors are then used to find the transformation matrix **W**.
2. For each test image, we extract the feature vector consisting of the 12 IQMs, which is then projected onto **W** and compared to the feature vectors from the images in the training set.
3. The class index (degradation type) corresponding to the minimum distance is declared as the detected class or degradation type.

In our work, we used the Euclidian Distance since the results were comparable to those obtained using the Mahanalobis distance but with a much lower computational cost.

Since the focus is on the degradation type, the performance of the proposed method was evaluated in terms of degradation classification accuracy for the different test images.

We present in Fig. 3 the percentage of good classification for each type of degradation. Note that for all types of degradations, we achieve accuracies of over 90%. Actually, we obtained 100% accuracy for the majority of degradations considered in our experiments. For all of the 17 types of degradations, we obtained an average of 98.11% classification accurcay.
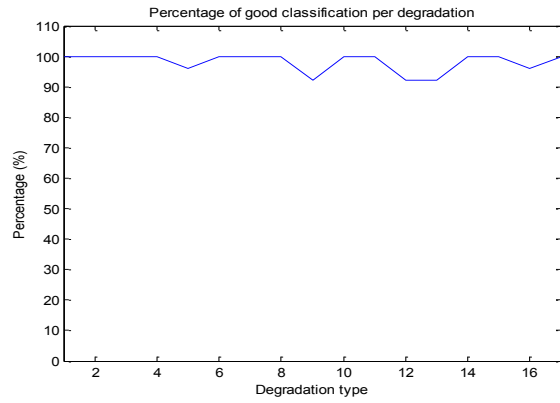
Figure 3 – Percentage of good classification per degradation.

To better visualize our results, we present the confusion matrix for different classes in Fig.4. Note that the lowest results are obtained for classes 9, 12 and 13 (92%).

ESTIMATED CLASS

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
T	**1**	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
R	**3**	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0
U	**4**	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
E	**5**	0	0	0	4	96	0	0	0	0	0	0	0	0	0	0	0	0
	6	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
C	**8**	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
L	**9**	0	0	0	0	0	0	0	0	92	4	4	0	0	0	0	0	0
A	**10**	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0
S	**11**	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
S	**12**	0	0	0	0	0	0	0	0	0	0	0	92	8	0	0	0	0
	13	0	0	0	0	0	0	0	0	0	0	0	8	92	0	0	0	0
	14	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0
	16	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	96	0
	17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100

Figure 4 – Confusion matrix for 1700 images from TID2008.

The proposed method has been successfully evaluated on various images. Now, the quality of a given distorted image can be better measured using the more appropriate IQM. The detailed results from such analysis are currently being summarized for future publication.

4. CONCLUSIONS

In this study, we proposed a new pre-processing approach for a more robust estimation of image quality. In particular, we discussed an LDA-based technique for classifying degradations before estimating image quality. The classification stage uses the different IQMs estimated from a given image as features. Our experimental results show that the type of degradation can be estimated with more than 90% accuracy. Such knowledge can be used for determining the types of IQMs that need to be used for evaluating quality.

5. ACKNOWLEDGEMENTS

REFERENCES

[1] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti and M. Carli, "New full-reference quality metrics based on HVS", CD-ROM Proceedings of the Second International Workshop on VPQM, 2006.

[2] G. Legge and J. Foley, "Contrast masking in human vision", Vol.70, pp. 1458-4471, 1980.

[3] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola and V. Lukin, "On between-coefficient contrast masking of DCT basis functions", CD-ROM Proceedings of the Third Int. Workshop on VPQM, 2007.

[4] S. Daly, "The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity", Digital Images and Human Vision, Watson A.B.(Ed.). MIT Press, ch14, pp. 179-206, 1993.

[5] Z. Wang, E.P. Simoncelli and A.C Bovik, "Multiscale structural similarity for image quality assessment", IEEE Asilomar Conference on Signals, Systems and Computers, 2003.

[6] H.R. Sheikh and A.C Bovik, "Image information and visual quality", IEEE Transactions on Image Processing, Vol.15, no.2, pp. 430-444, 2006.

[7] Z. Wang and A.C. Bovik, "A universal image quality index", IEEE Signal Processing Letters, Vol. 9, pp. 81-84, 2002.

[8] H.R. Sheikh, A.C. Bovik and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics", IEEE Trans. on Image Processing, Vol.14, no.12, pp. 2117-2128, 2005.

[9] T. Mitsa and K. Varkur, "Evaluation of contrast sensitivity functions for the formulation of quality measures incorporated in halftoning algorithms", IEEE-ICASSP, pp. 301-304, 1993.

[10] D.M. Chandler and S.S. Hemami, "VSNR: A Wavelet-Based Visual Signal-to-Noise Ratio for Natural Images", IEEE Transactions on Image Proc., Vol. 16, pp. 2284-2298, 2007.

[11] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans and A.C. Bovik, "Image Quality Assessment Based on a Degradation Model", IEEE Transactions on Image Processing, Vol. 9, pp. 636-650, 2000.

[12] B. W. Kolpatzik and C. A. Bouman, "Optimized Universal Color Palette Design for Error Diffusion", Journal Electronic Imaging, Vol. 4, pp. 131-143, 1995.

[13] N. Ponomarenko, M. Carli, V. Lukin, K. Egiazarian, J. Astola and F. Battisti, "Color Image Database for Evaluation of Image Quality Metrics", Proceedings of International Workshop on Multimedia Signal Processing, Australia, pp. 403-408, 2008.

[14] M. Deriche, A. Beghdadi, A. Chetouani, "A Distortion-Based Ranking of Image Quality Indices using Mutual Information", 10th Information Science, Signal processing, and their Applications, ISSPA 2010.

[15] N. Peter, J. Belhumeur, P. Hespanha, D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Trans. on Pattern Analysis and Machine Intel., pp. 711-720, July, 1997.