

Quantifying image distortion based on Gabor filter bank and multiple regression analysis

B. Ortiz-Jaramillo^{a,b}, J.C. Garcia-Alvarez^a, H. Führ^c, S.A. Orjuela-Vargas^b, G. Castellanos-Dominguez^a and W. Philips^b

^aUniversidad Nacional de Colombia, Cra 27 # 64-60, Manizales, Colombia;

^bGent University, TELIN-IPI-IBBT, St-Pietersnieuwstraat 41, B-9000 Gent, Belgium;

^cLehrstuhl A für Mathematik, RWTH Aachen, D-52064 Aachen, Germany

ABSTRACT

Image quality assessment is indispensable for image-based applications. The approaches towards image quality assessment fall into two main categories: subjective and objective methods. Subjective assessment has been widely used. However, careful subjective assessments are experimentally difficult and lengthy, and the results obtained may vary depending on the test conditions. On the other hand, objective image quality assessment would not only alleviate the difficulties described above but would also help to expand the application field. Therefore, several works have been developed for quantifying the distortion presented on a image achieving goodness of fit between subjective and objective scores up to 92%. Nevertheless, current methodologies are designed assuming that the nature of the distortion is known. Generally, this is a limiting assumption for practical applications, since in a majority of cases the distortions in the image are unknown. Therefore, we believe that the current methods of image quality assessment should be adapted in order to identify and quantify the distortion of images at the same time. That combination can improve processes such as enhancement, restoration, compression, transmission, among others. We present an approach based on the power of the experimental design and the joint localization of the Gabor filters for studying the influence of the spatial/frequencies on image quality assessment. Therefore, we achieve a correct identification and quantification of the distortion affecting images. This method provides accurate scores and differentiability between distortions.

Keywords: Image quality assessment, Gabor filter bank, multiple linear regression

1. INTRODUCTION

Image quality assessment has an important role in improving the performance of image-based applications such as face recognition,¹ video coding,² tele-medicine,³ among others. In this type of applications, quality assessment is used for measuring the performance of compression or enhancement algorithms. These studies can be grouped into two main classes: subjective and objective methods.⁴ The former methods are performed by a group of persons (experts), who evaluate compressed or corrupted images according to certain well-defined criteria such as ITU standards.⁵ Therefore, assessments are measured according to the perceived quality, termed Differential Mean Observation Scores (DMOS). However, performing such tests is complex and time and money consuming.^{4,6-9} Hence, objective (automatic) assessment is a desirable alternative.

Objective methods try to estimate the amount of distortion on an image using quantitative operations provided on the spatial or spatial/frequency domain. In this field, standard objective distortion measures such as mean squared error, signal to noise ratio and peak signal to noise ratio are simple methods for providing a measure of the differences between two considered images. However, this kind of measures do not take into account perceptual features. Therefore, the nature of the distortion cannot be identified and quantified. Hence, several works has been developed for quantifying the distortion presented on an image. For instance, Ginesu et. al.⁴ presented a perceptual metric based on three typical distortions produced by compression algorithms. In that work, correlation values up to 0.84 between DMOS and the perceptual metric are achieved. The study developed

Further author information: (Send correspondence to Benhur Ortiz)

Benhur Ortiz-Jaramillo.: E-mail: bortiz@telin.ugent.be, Telephone: +32 9 264 34 12

by Sheikh et. al.⁶ is perhaps the most complete work concerning quality assessment. In that work, a comparison between current objective metrics is performed. Particularly, it is found that the metric based on information theory is more accurate than others compared in the study.

More recent studies have applied the translation of Ginesu metrics to the wavelet domain.^{1,8} In those works, estimated correlation values up to 0.92 between DMOS and the perceptual metric by using wavelet domain are presented. Another approach presented by Mahmoudi et. al.⁷ attempts to classify and to identify the distortion of images. In that work, correlations up to 0.84 between DMOS and the perceptual metric are achieved. That method is also capable to differentiate white noise within a set of five types of distortion.

The design of current assessments assumes that the nature of the distortion is known. For instance, there are assessments dedicated to measuring *Blocking Artifacts*, which affect the quality of JPEG-compressed images,⁴ or to measure *Blurring Artifacts* for evaluating JPEG2000 compressors for tele-medicine applications.⁸ The restriction to these types of artifacts limits the use of these studies for practical applications, since in a majority of cases the distortions in the image are unknown. Thus, current image quality assessment techniques assume that the nature distortion, resulting in severe limitations of generalization.¹⁰

In any case, we believe that the current methods of image quality assessment should be adapted in order to identify and quantify the distortion of images at the same time, which should help to develop more efficient and accurate algorithms for problems such as enhancement, restoration, compression, transmission, among others. This provides the main motivation for the work presented in this paper.

This work presents a perceptual assessment studying the influence of the spatial/frequencies on image quality assessment. This ability allows the automatic optimization processes inside of image processing parameters. For instance, it can be used to optimize the parameters of a codec. Our proposed assessment method is a linear combination of Gabor filter responses. Therefore, this assessment is adequate to give a close representation of the human vision model because Gabor filters have proved to be closely related to the visual perception.¹¹

The rest of this paper is structured as follows. In Section 2 Gabor filters, multiple regression analysis, quadratic Bayes normal classifiers and multi-class performance estimation are introduced. Afterwards, the experimental setup and results are explored in Sections 3 and 4, respectively. Finally, in Section 5 conclusions and future work are provided.

2. MATERIALS AND METHODS

2.1 Gabor Filtering

Gabor filtering allows imaging with optimal joint localization in the spatial-spatial/frequency domains.¹¹ Moreover, Gabor filtering is widely used for characterizing the human vision system.¹¹ Therefore, this technique should be appropriate for identifying and quantifying the distortion in image quality assessment tasks.

A two-dimensional Gabor function $g(x, y)$ and its Fourier transform $G(u, v)$ are defined as,¹¹

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + 2\pi ju_0x\right) \quad (1)$$

$$G(u, v) = \exp\left(-\frac{1}{2}\left(\frac{(u - u_0)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right)$$

where, $\sigma_u = (2\pi\sigma_x)^{-1}$ and $\sigma_v = (2\pi\sigma_y)^{-1}$. Here, σ_u and σ_v characterize the band width of the Gabor filter centered at the point $(u_0, 0)$ in the spatial/frequency domain (u, v) . From Equation (1), it is possible to generate a set of Gabor functions by appropriate dilations and rotations, i.e.,

$$\begin{aligned} g_{m,n} &= a^{-m}g(x', y'), \quad a > 1, \quad m, n \in \mathbb{Z} \\ x' &= a^{-m}(x \cos(\theta) + y \sin(\theta)), \quad \text{and} \quad y' = a^{-m}(-x \sin(\theta) + y \cos(\theta)) \end{aligned}$$

where $\theta = n\pi/K$, with $n = 0, \dots, K - 1$, and K is the total number of orientations. Here a^{-m} , with $m = 0, \dots, S - 1$, is the scale parameter, where S is the number of scales. This set of functions are often used as a filter bank in image processing tasks.

Due to the nonorthogonality of the Gabor functions the resulting representation is redundant. To cope with

this shortcoming, Manjunath and Ma¹¹ proposed a strategy to reduce the redundancy presented in the filter bank decomposition. The strategy ensures that the responses of the filters in the spatial/frequency domain are tangent to each other (See Figure 1). Let U_l and U_h denote the lower and upper center frequencies of interest, respectively. Then, the filters are defined in terms of:

$$\begin{aligned}
 a &= \frac{U_h^{\frac{1}{s-1}}}{U_l} \\
 \sigma_u &= \frac{(a-1)U_h}{(a+1)\sqrt{2\ln(2)}} \\
 \sigma_v &= \left(\frac{\pi}{2K}\right) \left(U_h - 2\ln\left(\frac{2\sigma_u^2}{U_h}\right)\right) \left(2\ln(2) - \frac{(2\ln(2))^2\sigma_u^2}{U_h^2}\right)^{-\frac{1}{2}}
 \end{aligned}$$

where $u_0 = U_h$. As shown in Figure 1, the filters responses are rotated versions of the response in Equation (1), where σ_u and σ_v are related to a .

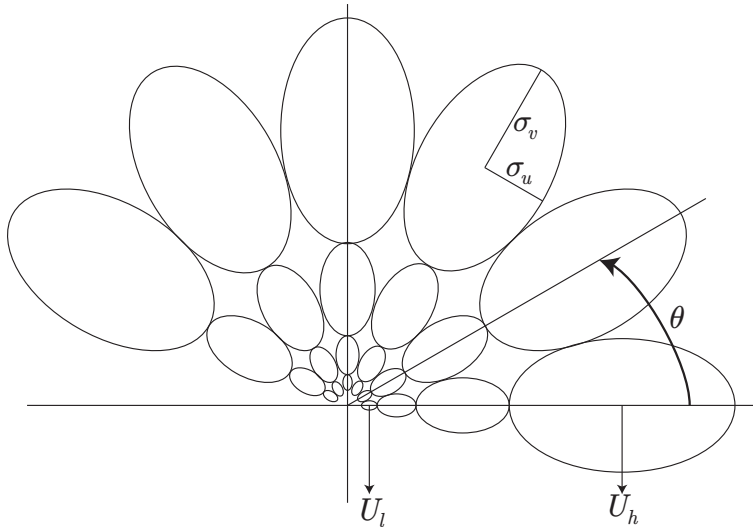


Figure 1. The elliptical contours tangent to each other indicate the response of the Gabor functions.

2.2 Multiple regression analysis

Multiple regression is a method used to model the relationship between a variable provided by a system or an expert (in statistics this is called response variable) and one or more input variables provided by measuring physical quantities (in statistics this is called predictor variables). In linear regression analysis when the response variable is highly correlated with the predictor variables, it is possible to predict the response variable by knowing the predictor variables. A linear model is defined as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$ (see Equation (2)), where $\mathbf{y} \in \mathbb{R}^{n \times 1}$, $\mathbf{X} \in \mathbb{R}^{n \times p+1}$, $\boldsymbol{\beta} \in \mathbb{R}^{p+1 \times 1}$ and $\boldsymbol{\epsilon} \in \mathbb{R}^{n \times 1}$ are the response variable, predictor variables, unknown parameters and error associated to the model respectively. From this model, it is possible to predict any value in \mathbf{y} having \mathbf{X} and $\boldsymbol{\beta}$, i.e.,

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & \cdots & x_{p,1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1,n} & \cdots & x_{p,n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_p \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad (2)$$

It is common to assume that $\boldsymbol{\epsilon}$ is independent and distributed as a normal random variable with mean zero and variance σ^2 . From this assumption it is possible to find an unbiased estimation of the parameters $\boldsymbol{\beta}$ using least square estimation, i.e., $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{y})$.

Usually, the most useful measure for determining the success of a model is the coefficient of determination, that is defined as $R^2 = \frac{\hat{\beta}^T \mathbf{X}^T \mathbf{y} - (\sum \mathbf{y})^2/n}{\mathbf{y}^T \mathbf{y} - (\sum \mathbf{y})^2/n}$.¹² Despite its usefulness in measuring the model fit, the value of R^2 does not provide information about how much each predictor variable contributes to the model. For this, selection of variables is probably the most fundamental and important issue in regression analysis, because in such way, the contribution of each predictor variable could be computed. In this field several approaches have been applied such as standardized parameters, simple correlation, ANOVA, Pratt's measure, decomposition of R^2 into components associated with the regressors.^{12,13} Particularly, the decomposition of R^2 in components proposed by Genizi¹³ provides a general framework to select variables by using an orthogonal transformation of the data in \mathbf{X} . The main idea of this methodology is that whenever a set of mutually orthogonal predictors is available, the sum of each partial correlation is equal to the coefficient of determination. That is, $R^2 = \sum_{i=1}^p \rho_{\mathbf{y}, \mathbf{x}_i}$, where $\rho_{\mathbf{y}, \mathbf{x}_i}$ is the partial correlation between response and the i th predictor variable. Usually, predictor variables are not mutually orthogonal. In this case, it is possible to achieve a variable selection by using a singular value decomposition of the data correlation matrix as $\Sigma = \mathbf{Q}\mathbf{C}^2\mathbf{Q}^T$. Thus, a set of values representing the relative importance in the model are computed as $v_i = \sum_{k=1}^p a_{i,k}^2 \gamma_k^2$, where $a_{i,k}$ is the element in the position (i, k) of matrix $\mathbf{A} = \mathbf{Q}\mathbf{C}\mathbf{Q}^T$ and $\gamma = \mathbf{A}\mathbf{D}^{-1}\mathbf{y}$ as discussed in.¹³

2.3 Quadratic Bayes normal classifier and multi-class performance estimation

The most widely used classifiers are based on the normal distribution.¹⁴ Particularly, the quadratic Bayes normal classifier (QBNC) is a common tool for classification tasks.¹⁴ In general, a classifier should assign to a new object \mathbf{x} the most likely class c_i . This class can be assigned when the probabilities $p(c_1|\mathbf{x}), \dots, p(c_C|\mathbf{x})$ (for a C class classification problem) are known. Then, the class is assigned to the object \mathbf{x} by determining the maximum over $p(c_1|\mathbf{x}), \dots, p(c_C|\mathbf{x})$. When $p(\mathbf{x}|c_i)$ is given by a Gaussian distribution the classifier is called QBNC, where $p(\mathbf{x}|c_i) = \frac{p(c_i|\mathbf{x})p(\mathbf{x})}{p(c_i)}$ is the decision Bayes rule.¹⁴

The discriminability measurement of a classifier (how well it classifies unseen data) is important in pattern recognition. There are many measures of discriminability, but the most common is the classification rate. Generally, it is difficult to obtain an analytic expression for the classification rate and therefore it must be estimated from the available data.¹⁴ In addition, usually the confusion or misclassification matrix is also computed. The (i, j) th element of that matrix is the number of patterns of class j that are classified as class i by the classifier. This measure is very useful for identifying how the error rate is decomposed.¹⁴

In particular, the holdout method is widely used for classification performance measure that splits the data into two mutually exclusive sets, termed as the training and test sets. The classifier is designed using the training set and performance is evaluated on the test set.¹⁴

2.4 Proposed methodology

In the following paragraphs the methodology for identification and quantification of image distortion is explained.

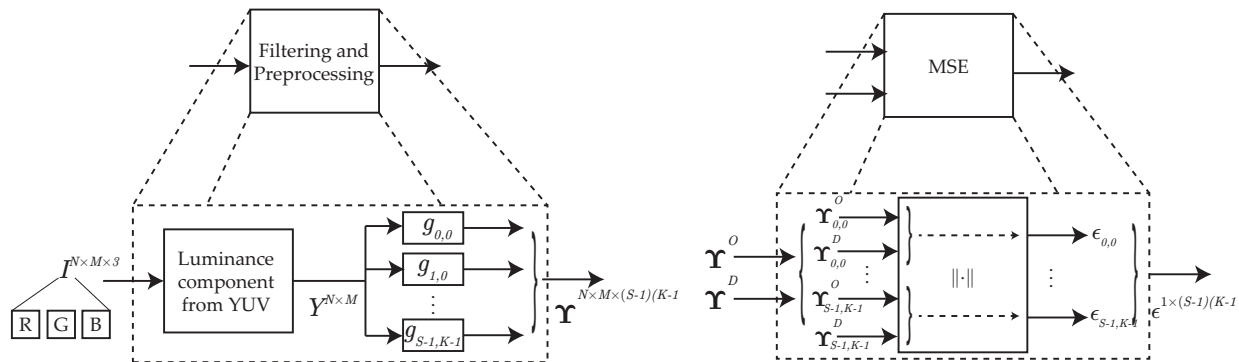


Figure 2. Filtering, preprocessing and MSE by using the set of images obtained by Gabor filter bank decomposition.

First at all, the luminance component from each image is computed ($Y = 0.299R + 0.587G + 0.114B$) as suggested by Furht and Marques.¹⁵ Afterwards, each image is filtered by using the set of $(S - 1)(K - 1)$ filters designed with the methodology explained in previous section. In this paper the following parameters are used: $U_h = 0.7$, $U_l = 0.005$, $K = 6$ and $S = 4$. Thus, each image from the database is represented by a set of $(S - 1)(K - 1)$ images representing a particular band in the spatial/frequency space. In the following, the proposed method to identify and quantify the distortion of the images is explained.

Let I_i^D and $I_i^O \in \mathbb{R}^{N \times M \times 3}$ be the i th distorted image and its corresponding reference image belonging to a particular distortion $C \in \{fastfad, gblur, jp2k, jpeg, wn\}$. First of all, from each image the luminance component is extracted. Afterwards, both images, distorted and original, are filtered by using the set of $(S - 1)(K - 1)$ Gabor filters defined in Section 2 (See Figure 2 left side). That is, $\Upsilon_{i,m,n}^j = Y_i^j * g_{m,n}$, where $j = \{O, D\}$, n, m are the scale and angle variation in the Gabor filter bank and $*$ stands for convolution operation. By this filtering, two sets of $(S - 1)(K - 1)$ images with the same size of the original one are obtained. Those sets represent the spatial/frequency distribution of the distorted $\Upsilon^D \in \mathbb{R}^{N \times M \times (S-1)(K-1)}$ and the original image $\Upsilon^O \in \mathbb{R}^{N \times M \times (S-1)(K-1)}$. Afterwards, the MSE between filtered original and filtered distorted sets is computed. That is, from the set of filtered images the MSE between the original and the distorted band in the Gabor space (see Figure 2 right side) is computed, i.e., $\epsilon_{i,m,n} = \|\Upsilon_{i,m,n}^O - \Upsilon_{i,m,n}^D\|$, where $\|\mathbf{x}\| = \sqrt{\sum_k x_k^2}$. Thus, it is possible to build a set of MSEs ($\epsilon \in \mathbb{R}^{n-1 \times (S-1)(K-1)}$) for representing the difference between each frequency band in the spatial/frequency space for the n images belonging to the database (See Figure 3 left side).

3. EXPERIMENTAL SETUP

The proposed experimental setup for identifying and quantifying image distortion comprises the following stages: a) Luminance component extraction, b) Gabor Filter bank decomposition, c) mean squared error evaluation (MSE), and d) Performance evaluation.

3.1 Database

In this paper the LIVE Image Quality Assessment Database Release 2–2005¹⁶ is used. This database contains 30 source images. Each of those images are processed for acquiring the following sets:

- 169 images coming from degradation of source images by jpeg 2000 compression, labeled as *jp2k*.
- 174 images coming from degradation of source images by jpeg compression, labeled as *jpeg*.
- 144 images coming from degradation of source images by white gaussian noise, labeled as *wn*.
- 144 images coming from degradation of source images by Blurring distortion, labeled as *gblur*.
- 144 images coming from degradation of source images by Fast Fading distortion, labeled as *fastfad*.

In this database a (DMOS) value is assigned to each distorted image.⁶ These DMOS values are assigned by subjective evaluation. Each subjective evaluation offers a reference of comparison for the performance of metrics intending measure the quality of a image.

3.2 Identification and quantification of image distortion

Here, the set of MSEs is used as feature vector in the regression analysis and the classifier. The regression analysis is performed by using the Genizi method. This methodology is used because it allows to identify which bands in the spatial/frequency are most affected by the distortion. I.e., it provides the level of influence from each band in the estimated DMOS. It means, which of the column vectors in ϵ variable are the most important for the regression model. Furthermore, for validating the regression models the adjusted coefficient of determination is computed. The adjusted coefficient of determination is computed by using the subjective DMOS given by human assessment and our proposed objective metric.

On the other hand, by using a QBNC the type of distortion in the image is identified. Here, each row vector of ϵ variable is used as sample. Each of these samples are labeled by the type of distortion used to create the distorted

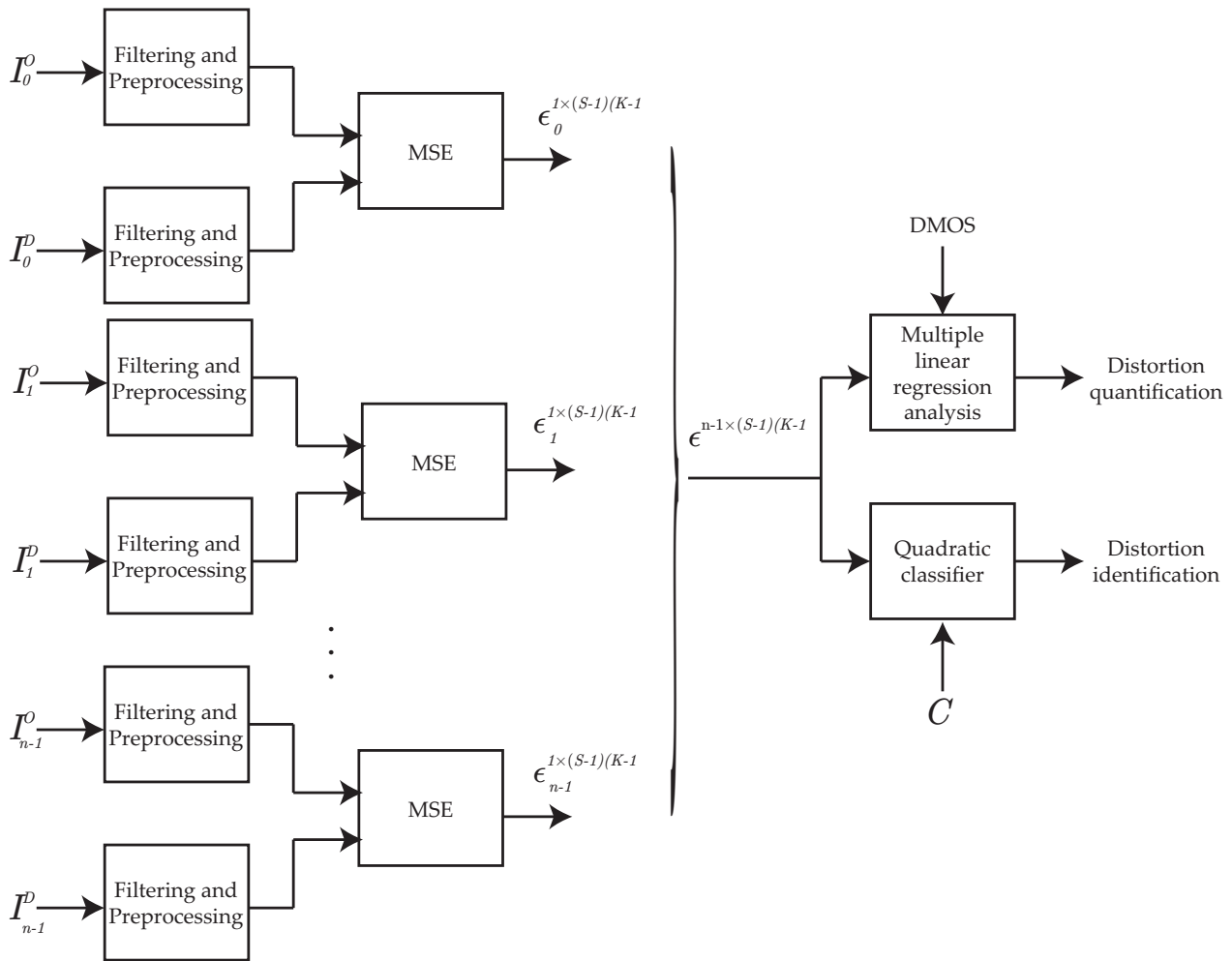


Figure 3. Proposed methodology for quantifying and identifying image distortion.

image. In this work, 60% of the samples are used as training samples, and the rest are used as validation samples. From this analysis a correct classification rate and a confusion matrix are computed for validating distortion identification procedure.

4. RESULTS AND DISCUSSION

Since Gabor filters have proved to be close related to the human perception, we expect the Gabor filter bank approach to be well suited for image quality assessment. The Genizi method is used to determine which band of the spatial/frequency space is more affected by a particular type of distortion; the decomposition of R^2 into components was then performed by using the DMOS. The R^2 components therefore measure the amount of variation accounted for each band in the spatial/frequency space. The bands with bigger v value are the most affected for the distortion. For instance, assume that an image is distorted by using *jpeg* compression and the greatest v value is located in the spatial/frequency band generated by the filter with parameters (m, n) . Then, it is possible to say that such spatial/frequency band is the most affected by *jpeg* compression.

Table 1 shows the values of the decomposition of R^2 over different bands. Each distortion type is sorted from lowest to highest v value. This v values are very useful to determine which spatial/frequencies are more affected by a particular type of distortion. Therefore, using this localization could be appropriate for recognizing different spatial/frequencies areas of images with different distortion, and can lead to restoration of the image by

Table 1. Contribution of each predictor variable to the linear model. The analysis is performed by using the Genizi sum of components. Noteworthy that $R^2 = \sum_i v_i$ where i changes across the rows. (m, n) pair are the parameters of scale and angle used to build the Gabor filter bank.

i	<i>jp2k</i>		<i>jpeg</i>		<i>wn</i>		<i>gblur</i>		<i>fastfad</i>	
	v	(m, n)	v	(m, n)	v	(m, n)	v	(m, n)	v	(m, n)
0	0.0341	(0, 1)	0.0292	(0, 5)	0.0369	(1, 0)	0.0293	(0, 0)	0.0325	(3, 0)
1	0.0341	(0, 5)	0.0292	(0, 1)	0.0370	(1, 5)	0.0294	(0, 5)	0.0326	(2, 0)
2	0.0341	(0, 0)	0.0292	(0, 0)	0.0370	(1, 1)	0.0294	(0, 1)	0.0337	(1, 0)
3	0.0342	(0, 2)	0.0292	(0, 4)	0.0370	(1, 3)	0.0294	(0, 4)	0.0340	(3, 1)
4	0.0342	(0, 4)	0.0292	(0, 2)	0.0371	(1, 4)	0.0294	(0, 2)	0.0341	(2, 1)
5	0.0343	(0, 3)	0.0293	(0, 3)	0.0371	(1, 2)	0.0295	(0, 3)	0.0350	(3, 5)
6	0.0355	(1, 0)	0.0363	(3, 0)	0.0380	(2, 0)	0.0362	(1, 1)	0.0351	(1, 1)
7	0.0355	(1, 5)	0.0364	(2, 0)	0.0382	(2, 3)	0.0364	(1, 0)	0.0352	(2, 5)
4	0.0357	(3, 0)	0.0368	(3, 5)	0.0383	(3, 0)	0.0384	(1, 5)	0.0360	(0, 0)
9	0.0358	(2, 0)	0.0368	(3, 1)	0.0383	(2, 5)	0.0387	(1, 2)	0.0360	(0, 5)
10	0.0359	(3, 5)	0.0370	(2, 1)	0.0383	(0, 3)	0.0389	(3, 0)	0.0360	(0, 1)
11	0.0360	(1, 4)	0.0370	(2, 5)	0.0383	(0, 0)	0.0389	(2, 0)	0.0362	(0, 4)
12	0.0360	(2, 5)	0.0373	(3, 4)	0.0384	(2, 1)	0.0391	(3, 1)	0.0362	(0, 2)
13	0.0362	(1, 1)	0.0373	(1, 0)	0.0384	(0, 5)	0.0391	(2, 1)	0.0363	(0, 3)
14	0.0365	(3, 1)	0.0373	(3, 2)	0.0384	(0, 1)	0.0409	(1, 4)	0.0364	(1, 5)
15	0.0365	(3, 4)	0.0373	(3, 3)	0.0384	(0, 4)	0.0411	(3, 5)	0.0384	(3, 2)
16	0.0366	(2, 1)	0.0374	(2, 4)	0.0384	(0, 2)	0.0411	(2, 5)	0.0385	(2, 2)
17	0.0366	(1, 3)	0.0375	(2, 2)	0.0384	(2, 4)	0.0417	(1, 3)	0.0390	(1, 2)
18	0.0366	(2, 4)	0.0375	(2, 3)	0.0384	(3, 3)	0.0421	(2, 2)	0.0396	(3, 4)
19	0.0368	(1, 2)	0.0378	(1, 1)	0.0385	(2, 2)	0.0421	(3, 2)	0.0397	(2, 4)
20	0.0370	(3, 3)	0.0380	(1, 3)	0.0385	(3, 5)	0.0442	(2, 4)	0.0405	(1, 4)
21	0.0371	(2, 3)	0.0380	(1, 5)	0.0386	(3, 1)	0.0442	(3, 4)	0.0412	(3, 3)
22	0.0371	(3, 2)	0.0381	(1, 2)	0.0386	(3, 4)	0.0450	(2, 3)	0.0413	(2, 3)
23	0.0372	(2, 2)	0.0382	(1, 4)	0.0387	(3, 2)	0.0450	(3, 3)	0.0420	(1, 3)
R^2	0.86		0.85		0.91		0.91		0.89	

adaptive methods. Later, in following paragraphs, the distortion localization on the spatial/frequency domain considering the results showed in Table 1 is provided.

First at all, goodness of fit of the proposed metric is compared with the values presented in the state of the art. For instance, in^{4,9} coefficients of determination up to 0.84 were obtained, i.e., the same performance achieved by our methodology. Nevertheless, our methodology is tested in a database with more images than the images tested in these two approaches. In^{1,8} values up to 0.92 of adjusted coefficient of determination were obtained. However, those two works cannot differentiate between distortion types because they consider one distortion of a time. Afterwards, in^{6,7} coefficients of determination up to 0.94 were presented. For instance, in Ref.⁷ a classification between types of distortion is proposed. However, the metric proposed by these sources is only designed to differentiate between *wn* and other type of distortions. On the other hand, our proposed methodology has a coefficient of determination up to 0.85. This is lower than some other measures such as VIF, SSIM⁶ and RF factor,⁷ but our methodology it is capable to detect which spatial/frequency component is more affected by the distortion. For instance, by checking Table 1 from bottom to top, it is possible to extract the components from the spatial/frequency space that are more relevant for DMOS characterization. For example, $(m, n) = \{(3, 3), (2, 3), (3, 4)\}$ are the parameters of scale and angle that have more importance for DMOS estimation in the distortion type *gblur*.

Figure 4 shows the set of Gabor filters in the spatial/frequency domain. In those bands, the noise localization according with the last three rows of Table 1 is shown. Noteworthy, every type of distortions has its spatial/frequency bands. Specifically, it is possible to see that *jpeg* compression affects the second scale of the

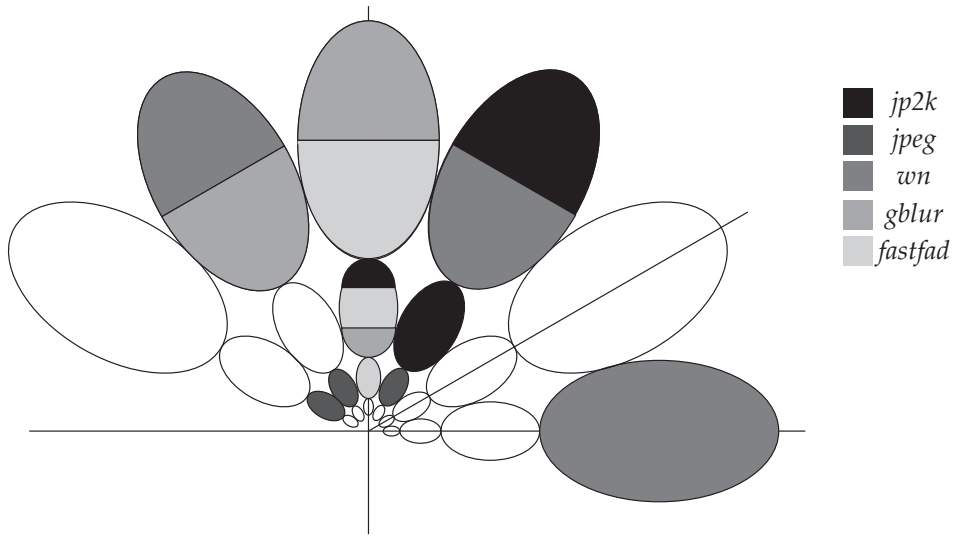


Figure 4. Gabor filter bank in the spatial/frequency domain with the noise localization according to the last three rows of Table 1. Here, those contours with more than one color mean that two or more distortions are located in the same spatial/frequency band.

bands generated by the Gabor filter bank, i.e., the most important variables for determining the quality of the image are obtained when $m = 1$. Additionally, it is possible to see that *wn* has great influence when $m = 3$, i.e., as it is expected *wn* should affect the high spatial/frequency bands. Another remarkable fact is the presence of the *fastfad* distortion over the vertical spatial/frequencies, i.e., $n = 3$ or $\theta = 90^\circ$. It means that such distortion could be more affected by the superposition of vertical edges in every scale. On the other hand, *jp2k* and *gblur* distortions do not exhibit a particular scale or angle to evaluate. It means that this type of distortion is distributed in several scales and angles over the spatial/frequency domain. On the other hand, the scale $m = 0$ is the least important band for the quality measure.

Table 2. Confusion matrix (CM) for the distortion identification by using a QBNC in the validation set. The accuracy of the classifier is 83%.

(a) Classification performance by using method proposed in⁷

		Predicted class				
		<i>jp2k</i>	<i>jpeg</i>	<i>wn</i>	<i>gblur</i>	<i>fastfad</i>
True class	<i>jp2k</i>	65	0	0	3	0
	<i>jpeg</i>	52	16	1	0	1
	<i>wn</i>	27	5	26	0	0
	<i>gblur</i>	22	0	3	20	13
	<i>fastfad</i>	41	2	0	11	4

(b) Classification performance by using our proposed methodology

		Predicted class				
		<i>jp2k</i>	<i>jpeg</i>	<i>wn</i>	<i>gblur</i>	<i>fastfad</i>
True class	<i>jp2k</i>	56	0	3	0	9
	<i>jpeg</i>	3	67	0	0	0
	<i>wn</i>	2	2	54	0	0
	<i>gblur</i>	1	0	12	40	5
	<i>fastfad</i>	10	2	0	4	42

Table 2 shows the confusion matrix obtained after applying a classifier in the validation set. For this experiment a set of 60% of the samples were used to train the classifier. Then, with the rest of the samples the classifier was tested. We perform this test based on the assumption that each type of distortion is located in a particular spatial/frequency band. Then, by using the differences (MSE) between distorted and original images would be possible to differentiate between distortion types. Indeed, from the Table 2(b) it is possible to extract the correct classification rate as $100 \frac{\sum_i CM(i,i)}{\sum_{i,j} CM(i,j)} = 83\%$, which is better than the results in the state of art Table 2(a) (42%). For instance, our methodology is compared with the work proposed by Mahmoudi et. al.⁷ From Table 2(a) is possible to discern that all distortion types are located in the same location because each distortion is confused with *jpeg*. On the other hand, from Table 2(b) it is noteworthy that the most accurate classifications are achieved by *jpeg* and *wn*. This could be due to the fact that this kind of distortion has the most relevant information in only one scale according with Figure 4. On the other hand, *jpeg*, *gblur* and *fastfad* distortions shares spatial/frequency bands doing the classification rate lower than in *jpeg* and *wn* distortions. However, the results illustrate that the spatial/frequency distribution of the distortion is distributed in particular bands specific to the distortion type.

5. CONCLUSIONS AND FUTURE WORK

An approach based on the power of the experimental design to explore effects between components and the joint localization of Gabor filters for splitting spatial/frequency components was presented. Therefore, a correct identification and quantification of the distortion affecting images were performed. From this combination accurate scores and appropriate differentiability between distortion natures were achieved. Within the results we found coefficients of determination up to 85%. This results are lower than the presented in the state of art. However, our method unlike others does not assume that the distortion is know a-priori. Therefore, our method is useful for various kinds of distortion. Besides, the different spatial/frequency bands affecting the image for each type of distortion were identified. These results could be very useful in restoration and denoising tasks. Hence, from the spatial/frquency localization of the distortion it is possible to use specific algorithms for specific distortions. Thereby, those specific algorithms can be designed based in the distortion nature incrementing the image quality.

The exploration of algorithms for image restoration and denoising applied over this decompositions remains as future work. Also the improvement of the filter bank decomposition for achieving more reliable results in the identification and quantification of the distortion still as future work. Finally, we believe that future development concerning image quality assessment must be performed in the spatial/frequency domain due to the fact that human vision is closely related to this kind of analysis. Therefore, it is possible to say that the results presented in this work could offer a good starting point for future development in image quality assessment and restoration.

ACKNOWLEDGMENTS

This research is partially supported within the framework of 2007–2009 ALECOL–DAAD Fellowship between Universidad Nacional de Colombia and RWTH Aachen; project number 20201006594 funded by Universidad Nacional de Colombia and Universidad de Caldas; project number 20201006570 funded by Universidad Nacional de Colombia; and Gent University, TELIN-IPI-IBBT.

REFERENCES

- [1] B. Ortiz-Jaramillo, J. García-Alvarez, and G. Castellanos-Dominguez, “Traslating distortion measures to a multiresolution space for image analysis,” *Ingenieria y competitividad* **12**, pp. 43–55, 2010.
- [2] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, “Study of subjective and objective quality assessment of video,” *IEEE Transactions on Image Processing* **19**, pp. 1427–1441, 2010.
- [3] J. García-Alvarez, “Telemedicine integrated-services perspectives,” *Revista del congreso internacional de ingeniería electrónica* **11**, pp. 1–4, 2007.
- [4] G. Ginesu, F. Massidda, and D. D. Giusto, “A multi-factors approach for image quality assessment based on a human visual system model,” *Signal Processing: Image Communication* **21**, p. 316333, 2006.

- [5] ITU-R-Recommendation-BT.500-11, “Methodology for the subjective assessment of the quality of television pictures.” ITU, Geneva, Switzerland, 1998.
- [6] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, “A statistical evaluation of recent full reference image quality assessment algorithms,” *IEEE Transactions on Image Processing* **12**, pp. 3441–3452, 2006.
- [7] A. Mahmoudi-Aznavah, A. Mansouri, F. Torkamani-Azara, and M. Eslami, “Image quality measurement besides distortion type classifying,” *Optical review* **16**, p. 3034, 2009.
- [8] J. García-Alvarez, H. Fuhr, B. Ortiz-Jaramillo, and G. Castellanos-Dominguez, “Image distortion mappings in wavelet representation space,” in *Proceedings of the International conference on broadband communications, information technology and biomedical applications*, pp. 1–6, 2009.
- [9] J. A. Redi and I. Heynderickx, “Image quality and visual attention interactions: towards a more reliable analysis in the saliency space,” in *Proceedings of the Third International Workshop on Quality of Multimedia Experience*, pp. 2001–2006, 2011.
- [10] A. K. Moorthy and A. C. Bovik, “A two-step framework for constructing blind image quality indices,” *IEEE Signal Processing Letters* **17**, pp. 513–516, 2010.
- [11] B. S. Manjunath and W. Y. Ma, “Texture features for browsing and retrieval of image data,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**, pp. 837–842, 1996.
- [12] J. H. Zar, *Biostatistical analysis*, Prentice Hall, New Jersey, USA, 2009 (5th edition).
- [13] A. Genizi, “Decomposition of r^2 in multiple regression with correlated regressors,” *Statistica Sinica* **3**, pp. 407–420, 1993.
- [14] A. R. Webb, *Statistical Pattern Recognition*, John Wiley & Sons, Ltd., West Sussex, England, 2002 (2nd edition).
- [15] B. Furht and O. Marques, *Handbook of Video Databases: Design and Applications*, CRC Press, Boca Raton, Florida, USA, 2003 (1st edition).
- [16] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, “Live image quality assessment database release 2.” <http://live.ece.utexas.edu/research/quality>, 2005.