Entropy of Gabor Filtering for Image Quality Assessment

Esteban Vazquez-Fernandez[1,2], Angel Dacal-Nieto[3], Fernando Martin[2], and Soledad Torres-Guijarro[4]

[1] GRADIANT - Galician R&D Center in Advanced Telecommunications, Spain
evazquez@gradiant.org
[2] Communications and Signal Theory Department, Universidade de Vigo
[3] Computer Science Department, Universidade de Vigo, Spain
[4] Laboratorio Oficial de Metroloxía de Galicia (LOMG), Spain

Abstract. A new algorithm for image quality assessment based on entropy of Gabor filtered images is proposed. A bank of Gabor filters is used to extract contours and directional textures. Then, the entropy of the images obtained after the Gabor filtering is calculated. Finally, a metric for the image quality is proposed. It is important to note that the quality of the image is image content-dependent, so our metric must be applied to variations of the same scene, like in image acquisition and image processing tasks. This process makes up an interesting tool to evaluate the quality of image acquisition systems or to adjust them to obtain the best possible images for further processing tasks. An image database has been created to test the algorithm with series of images degraded by four methods that simulate image acquisition usual problems. The presented results show that the proposed method accurately measures image quality, even with slight degradations.

1 Introduction

Image acquisition is a fundamental stage in every machine vision system. Obtaining the best quality images is critical to ensure a good performance. In this context, it is interesting to have a reliable way to measure the quality of the captured images or, from another point of view, to adjust the system to obtain the best possible images. Image quality assessment plays a fundamental role in this process, as well as in many image processing applications. It can be used to compare the performance of different methods (processing or acquisition) and to select the one which provides the best quality (or less image degradation); it can be used to measure the degradation itself after image processing operations; it also provides a metric to evaluate the performance of compression methods, like JPEG, or the quality of transmission channels (which is not covered in this work).

The most challenging problem in image quality assessment is the subjectivity inherent to perceived visual quality [1]. Several attempts to measure the quality of an image have been made, but it remains an open problem. Methods based on the measurement of Peak Signal to Noise Ratio (PSNR) or Mean Square Error (MSE) have been widely used due to their easy implementation, but the results show that they are not well suited to measure the human observer perceived quality [2]. Methods based on the use of previous knowledge of the Human Visual System (HVS) have shown a better performance in image quality assessment [3], [4]. HVS relays on the assumption that human observers pay more attention to details like structural information, which are more relevant to image quality measurement. Some previous contributions have pointed the use of entropy to measure image quality [5]. However, an entropy measure is unable to distinguish between noise and structural information. To solve this problem, a method based on image anisotropy has been proposed in [6].

Gabor filters have been extensively used in texture analysis and classification [7], [8], [9], but their use in image quality assessment remains little explored [10], [11]. The proposed method uses a bank of Gabor filters to model the linear filtering properties of single cells in visual cortex and to extract image contours and directional textures, which are directly related to HVS. Then, an estimation of the amount of visual information (randomness) perceived is calculated measuring the entropy of the outputs of the filter bank. The entropy value is directly related to the randomness of the image. Poorly defined transitions in the perceived image (Gabor response), which means less image quality, would produce a high entropy value. A metric is calculated by averaging the entropies obtained from the different Gabor filter bank outputs. This value can be used by itself as a reference, or can be normalized in relation to the original reference image, to show whether certain adjustment or process diminishes the image quality.

The paper is organized as follows. A theoretical background and the proposed algorithm are presented in Sect. 2. In Sect. 3 the developed test procedure to validate the method is shown. Results and discussion are presented in Sect. 4. Finally, some conclusions are given in Sect. 5.

2 Algorithm

2.1 Gabor Filters

Gabor filtering for image textural analysis has been introduced by Daugman [12]. The success of Gabor filters in this field is due to their aptitude to model the response of simple cortical cells in the visual system.

A 2D Gabor filter can be thought of as a complex plane wave modulated by a 2D Gaussian envelope and can be expressed in the spatial domain as:

$$\begin{aligned} &G_{\theta,f,\sigma_1,\sigma_2}(x,y) = \exp\left[\frac{-1}{2}\left(\frac{x'^2}{\sigma_1^2} + \frac{y'^2}{\sigma_2^2}\right)\right]\cos\left(2\pi f x' + \varphi\right) \\ &x' = x\sin\theta + y\cos\theta \\ &y' = x\cos\theta - y\sin\theta \end{aligned} \tag{1}$$

where f is the spatial frequency of the wave at an angle θ with the x axis, σ_1 and σ_2 are the standard deviations of the 2D Gaussian envelope, and φ is the phase.

Frequently in textural analysis applications, and also in this case, the Gaussian envelop is symmetric, so we have $\sigma = \sigma_1 = \sigma_2$.

A Gabor filter is suited to obtain local frequency information in a specific orientation (given by θ), which is directly related with image contours. A common practice in Gabor texture analysis is to use a bank of Gabor filters with different parameters tuned to capture several orientations and spatial frequencies. Attempts to systematize the design of the bank have been proposed [9], showing that increasing the number of frequencies and orientations has a little effect on the performance of the filter bank. However, the smoothing parameter, σ, is a significant factor to be carefully chosen in the bank design. Unfortunately, most of the times, it needs to be empirically chosen.

2.2 Image Entropy

The concept of entropy is associated with the amount of disorder in a physical system. Shannon redefined the entropy as a measure of the amount of information (uncertainty) in a source [13]. If the source is an image, it can be seen as a 2D array of information. The Shannon entropy is given by:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log_b p(x_i) \tag{2}$$

where $\Pr[X = x_i] = p(x_i)$ is the probability mass distribution of the source. This equation can be used to estimate the global entropy of an image characterized by its histogram:

$$H(I) = -\sum_{i=1}^{N} \mathrm{hist}_{\mathrm{norm}}(L_i) \log(\mathrm{hist}_{\mathrm{norm}}(L_i)) \tag{3}$$

where L_i represents the N intensity levels of the $m \times n$ image $I(x, y)$ and $\mathrm{hist}_{\mathrm{norm}}(L_i)$ is the histogram properly normalized to fit a probability distribution function:

$$\sum_{i=1}^{N} \mathrm{hist}_{\mathrm{norm}}(L_i) = 1 \tag{4}$$

The entropy of an image is an estimation of randomness, and is frequently used to measure its texture. As shown in Fig. 1, entropy can be thought as a measurement of the sharpness of the histogram peaks, which is directly related with a better defined structural information.

2.3 The Proposed Method

A flowchart of the proposed process is shown in Fig. 2. The input image is a grey level one; however the process can be easily applied to planes of a color space (like RGB). A bank of Gabor filters is used to extract contours and textural information. This stage converts the information to the HVS domain (cortex responses). The selected parameters for the filters are the following:

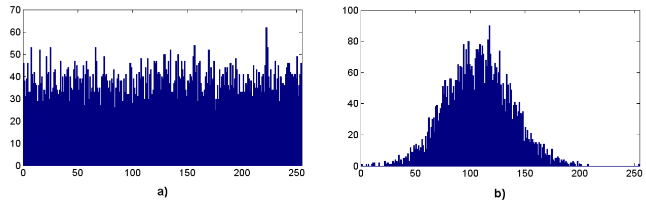

Fig. 1. Example of the entropy of different shape histograms. a) shows a higher entropy than b) ($Ha = 13.9627$; $Hb = 6.9216$)

- Six different orientations are used. However, the empirical tests show that the number of filters and angles does not seem to be crucial:

$$\theta \in \left[0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}\right] \tag{5}$$

- Two phases are used, $\varphi_1 = 0$ for a symmetric filter (on the θ orientation) and $\varphi_2 = \frac{\pi}{2}$ for an anti-symmetric filter. This can be thought as real and imaginary parts of the same filter response.
- Two different spatial frequencies are used: $f_1 = \frac{1}{8}$ (spatial period λ of 8 pixels) and $f_2 = \frac{1}{4}$ (spatial period λ of 4 pixels).
- The standard deviation of the Gaussian envelope is empirically fixed to $\sigma = \frac{\lambda}{2}$ for all the filters.

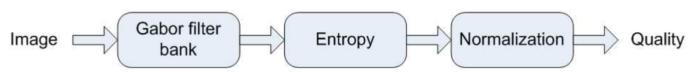

Fig. 2. Flowchart of the process

24 filtered images are obtained, 12 with $\varphi_1 = 0$ and 12 with $\varphi_2 = \frac{\pi}{2}$. Each pair is combined to estimate the energy of the filtered images by:

$$E(x, y) = \sqrt{R_{\varphi_1}(x, y)^2 + R_{\varphi_2}(x, y)^2} \tag{6}$$

where $R_{\varphi_i(x,y)}$ is the Gabor response for the phase φ_i.

This process results in 12 energy images. The histogram of these energy images is computed, and their entropy estimated through Eq. (3). Entropy measures the amount of information or, in other words, the randomness of the image histogram.

This procedure combines the benefits of objective and subjective measurements. On the one hand, Gabor filtering provides features inherent to the visual perceived quality by modelling the behaviour of visual cells. On the other hand, this information is quantified by the use of entropy.

However, the amount of information in an image depends on its content as well as on its quality. E.g. there is less information in an image of a white sheet than in a written one. For this reason, the entropy of the Gabor filtered image is not an absolute quality measurement, unless compared to a reference image. This is not a problem for the applications proposed in Section 1, in which the interest lies in comparing the quality of images of the same scene, or the effect of certain processing.

Taking this into account, the proposed relative quality metric (Q_r) is computed averaging the entropy of the energies of the 12 Gabor filtered images. The result is inverted and multiplied by the entropy of the reference image, obtained the same way:

$$\begin{aligned} H_{ref} &= \tfrac{1}{12} \textstyle\sum_{i=1}^{12} H_{ref_i} \\ H &= \tfrac{1}{12} \textstyle\sum_{i=1}^{12} H_i \\ Q_r &= \tfrac{H_{ref}}{H} \end{aligned} \tag{7}$$

with H being the calculated entropy and H_{ref} the entropy of the reference image. As the entropy increases, the quality of the image decreases, so a $Q_r \in (0, 1)$ means the quality of the image is lower than the reference (e.g. after the transmission through a noisy channel). If the resultant $Q_r > 1$, the quality of the image is higher than the reference (e.g. a noisy image which is enhanced by a median filtering, a blurred image which is enhanced by a fine tuning of the acquisition system, etc.).

3 Test Design

Two different test procedures have been developed to validate the performance of the proposed metric. The fist one is intended to model subtle variations in the image acquisition system. This is an objective quality test. The second one is intended to compare the proposed metric with the quality perceived by human observers. This is a subjective test.

3.1 Objective test

For this purpose, an image database of natural scenes has been created. It is composed of 1100 images of 2136×1424 pixels. It was originated by 25 original images (see Figure 3) progressively degraded in 10 steps following 4 different procedures (see below), which becomes in $25 \times (10+1) \times 4$ quality tagged images.

The degradations introduced to the original images in this first database are:

- Blur: Gaussian blur has been applied by increasing the filter size in 10 steps (from 3×3 to 21×21 pixel blocks).
- Noise: Zero mean Gaussian noise has been added by increasing its standard deviation in 10 steps (from 5 to 25 in 8 bits per pixel grey scale images).

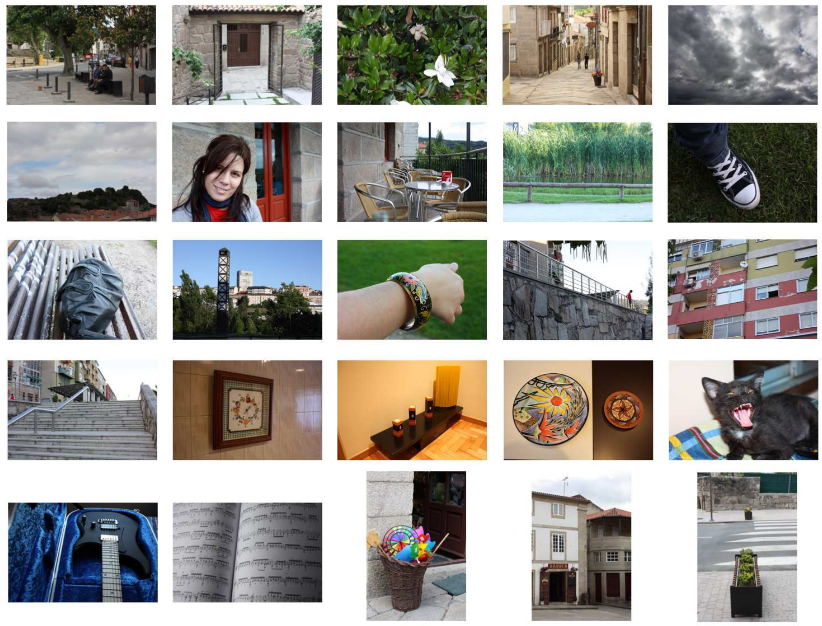

Fig. 3. Original images of natural scenes used to create the image database

– Blur & Noise: Gaussian blur has been applied, followed by adding Gaussian noise (10 steps). It models the effect of sensor noise after an out of focus imaging.
– Noise & Blur: Gaussian Noise has been added, followed by Gaussian blur (10 steps). It models the effect of software blurring operations after a noisy image acquisition (sensor noise).

The combination of noise and blur effects in different order, allows to simulate the effects of different acquisition systems, preprocessing operations, etc. [14]. Gaussian blur simulates the blur in an out of focus image. The Gaussian noise models the electronic noise which is produced in the camera sensor if the illumination, exposure time and gain parameters are not properly set. Figure 4 shows the effect of the 4 degradation procedures. This first database is intended to test the performance of the metric in the presence of subtle degradations.

3.2 Subjective test

For the second test procedure, images from the LIVE Image Database [14] have been used. LIVE database contains images obtained by several distortion procedures, which have been subjectively classified and scored by human observers. The scores have been scaled and shifted to a 1 to 100 range, and a Difference Mean Opinion Score (DMOS) was computed. For our test, images distorted

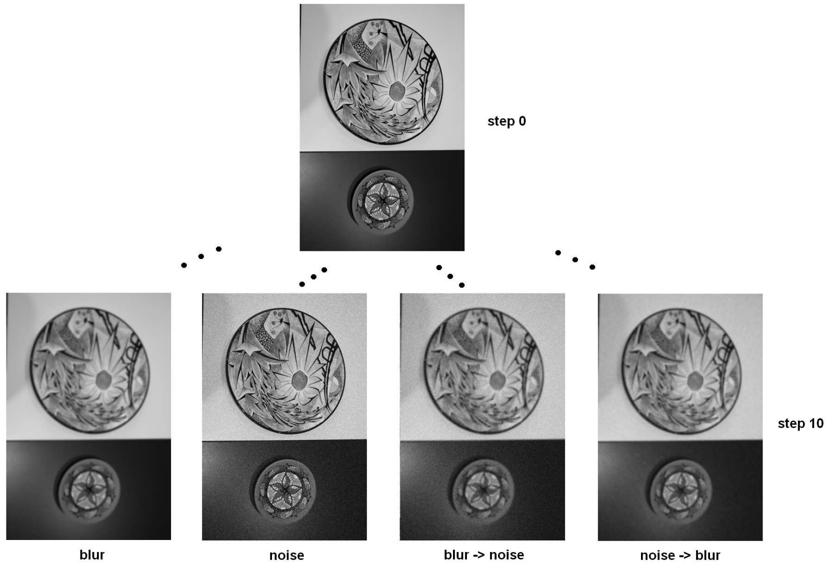

Fig. 4. Some steps for the blur, noise, blur & noise, and noise & blur degradation process of an example image

with white noise and Gaussian blur have been used. The database also contains images affected by JPEG compression, but it is not the aim of the proposed algorithm to test compression formats. The test is performed in a similar way to the first one. Images from the same scene have been sorted by their DMOS value (original and distorted ones). Then, the Q_r metric has been computed.

4 Results and Discussion

Figure 5 shows the results of the quality evaluation for the progressively degraded image shown in Figure 4. Similar results are obtained for the rest of the images used for the objective quality test.

It is interesting to note that most of the image quality assessment algorithms are tested using images that have a broad variation in quality. This is adequate when the objective is to model the quality perceived by an observer, e.g. to evaluate the performance of a compression algorithm or a transmission system. In these situations, the evaluation algorithm can be less precise (more tolerant), since variations in quality which are not perceived by the observer are not critical for the system.

However, if the objective is to select the best imaging system (or adjust it at its best) for a machine vision application, we have to be more strict in the performance of the method in a narrow error interval around the best possible image, which we call *Critical Peak*. In other words, we need to measure the quality of the

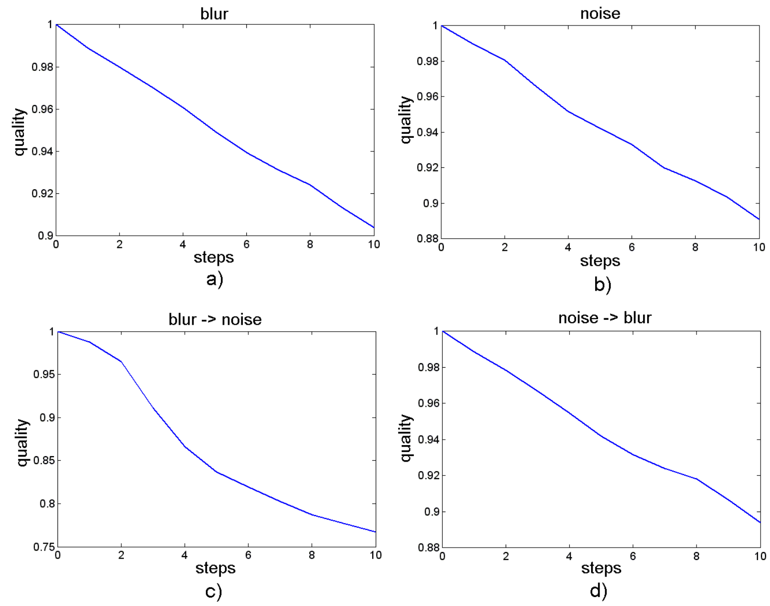

Fig. 5. Measured quality of the 10 step degraded images: a) blur; b) noise; c) noise after blur; d) blur after noise

image with a sufficiently high precision to obtain a strictly crescent/decreasing function.

As can be seen again in the example of Figure 4, the degradation applied is kindly subtle (low noise variance and small blurring mask) to test the Critical Peak performance. Figure 6 shows that all tested images have a strictly de-crescent function for their measured qualities in the test. The slope of the quality function varies significantly from image to image, because the degradation depends on the introduced distortion, as well as on the image content.

For the second part of the test (LIVE images), crescent curves have been obtained, without taking into account the differences in scale (Q_r ranges from 0 to 1, where 1 is the higher quality value). The results are shown in Figure 7. As can be seen, the proposed metric also correlates with the quality perceived by human observers. Note that a higher DMOS value means less quality (larger difference to the reference image).

In this case, most of the curves are strictly crescent. However, there are some anomalies in some functions, which can be due to several factors (inherent to the image database): variance in the perceived qualities by different observers or from one day to another; the DMOS scaling system grades every image in a 1 to 5 discrete scale, which means that the minimum DMOS values for a distorted image is always higher than 20 (in a 1 to 100 re-scaled range); in addition, every single image is evaluated by comparison with the original one, but not with the whole sequence of distorted images.

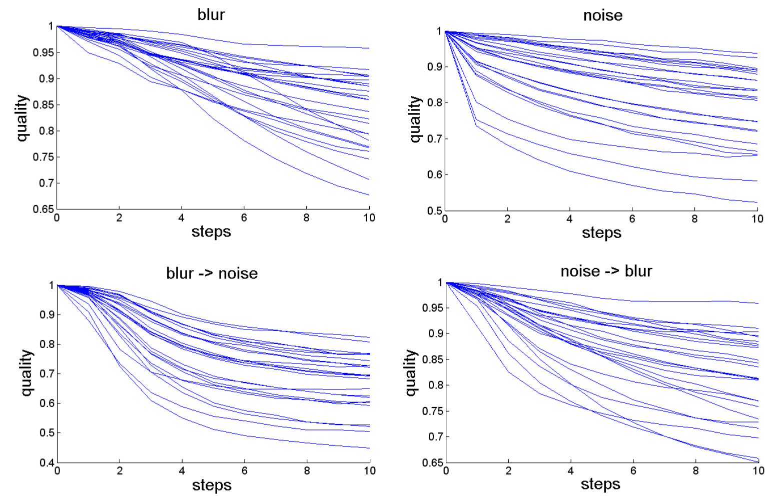

Fig. 6. Quality graphics of the test procedure. All images shown a strictly de-crescent function through the degradation procedures

There are also differences in the slopes, which produce dispersion in the curves of different images, due to the dependency of image content (besides its degradation). However, we obtain a good performance in quality evaluation for degraded sequences of the same image, which is the aim of the method.

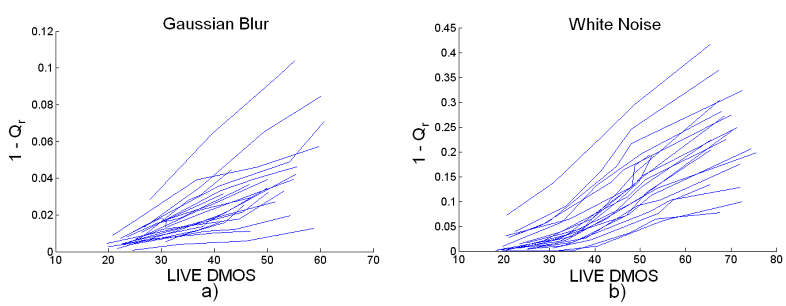

Fig. 7. LIVE images test. a) shows Gaussian blur distorted images; b) shows white noise distorted images

5 Conclusions

Image quality assessment is an important tool that allows the user to perform a comparison between variations of an image. This can be useful when developing image processing algorithms and when designing imaging systems. A method based on the entropy of Gabor filtered images has been developed. It combines objective measures (entropy) with subjective ones (HVS).

An image database has been created to test our metric, by means of an original set of natural images and applying different degradation methods to this initial set. With these degradations, real world behaviours present in image acquisition and image processing systems are modeled. When tested with this database, the proposed metric works properly even in narrow ranges, which can be checked by its strictly de-crescent charts. A test using LIVE Image Database also confirms it is well suited to human observer perception of image quality.

To conclude, the combination of subjective characteristics, modelled by Gabor filtering, and objective features, like entropy, provide a useful and powerful starting point for further developments on Image Quality Assessment.

References

1. Wang, Z., Bovik, A.C., Lu, L.: Why Is Image Quality Assessment so Difficult? Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing 4, 3313–16 (2002)
2. Wang, Z., Sheikh, H.R., Bovik, A.C.: No-reference perceptual quality assessment of JPEG compressed images. Proc. Int. Conf. on Image Processing 1, 477–80 (2002)
3. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Processing 13, 600–12 (2004)
4. Suresh, S., Babu, R.V., Kim, H.J.: No-reference image quality assessment using modified extreme learning machine classifier. Applied Soft Computing 9(2), 541–52 (2009)
5. Kirsanova, E.N., Sadovsky, M.G.: Entropy approach in the analysis of anisotropy of digital images. Open Syst. Inf. Dyn. 9, 239–50 (2004)
6. Gabarda, S., Cristóbal, G.: Blind image quality assessment through anisotropy. Journal of the Optical Society of America 24(12), 42–51 (2007)
7. Jain, A.K., Farrokhnia, F.: Unsupervised texture segmentation using Gabor filters. Pattern Recognition 24(12), 1167–86 (1991)
8. Jain, A.K., Ratha, N.R., Lakhsmanan, S.: Object detection using Gabor filters. Pattern Recognition 30(2), 295–309 (1997)
9. Bianconi, F., Fernández, A.: Evaluation of the effects of Gabor filter parameters on texture classification. Pattern Recognition 40, 3325–35 (2007)
10. Taylor, C.C., Pizlo, Z., Allebach, J.P., Bouman, C.A.: Image Quality Assessment with a Gabor pyramid model of the human visual system. Proc. SPIE Int. Symposium on Electronic Imaging Science and Technology 3016, 58–69 (1997)
11. Zhai, G., Zhang, W., Yang, X., Yao, S., Xu, Y.: GES: A new image quality assessment metric based on energy features in Gabor Transform Domain. IEEE Proc. Int. Symposium on Circuit and Systems, 1715–1718 (2006)
12. Daugman, J.G.: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. Journal of the Optical Society of America 2(7), 1160–69 (1985)
13. Shannon, C.E.: The Mathematical Theory of Communication. The Bell System Technical Journal 27, 379–423, 623–656 (1948)
14. Sheikh, H.R., Wang, Z., Cormack, L., Bovik, A.C.: LIVE Image Quality Assessment Database Release 2 http://live.ece.utexas.edu/research/quality