Objective Quality Potential Measures of Natural Color Images

Juha Katajamäki† and Hannu Saarelma*

Helsinki University of Technology, Laboratory of Media Technology, FIN-02015 HUT, Finland

This article deals with estimating the quality of natural color images from the digital image itself, for the purpose of determining whether the quality is sufficient for printing or displaying at some specified quality level. The problem of measuring simply defined parameters that correlate reasonably with human perception is discussed. Then an algorithm is described that measures sharpness, stochastic noise, and the blocking artifact of JPEG compression. The sharpness measure is the maximum of local sharpness, defined as the maximum gradient at an edge divided by the edge contrast. The algorithm behaves fairly consistently with respect to the amount of simulated blur and the measure correlates with the subjective sharpness of different images. The output of the stochastic noise algorithm is an unspecified number of noise levels, each of which is associated with the corresponding mean luminance. The algorithm computes a gradient image using the Sobel filters and uses the bivariate histogram of gradient and luminance to find the noise levels. The gradient filters are suitable for estimating the effect of power spectrum on noise visibility, but the discrimination between image details and noise is only partially solved by this method. The blocking artifact measure describes a characteristic spikiness in the histogram of the Sobel gradient.

Journal of Imaging Science and Technology 42: 250–263 (1998)

Introduction

The study reported here belongs to a framework to develop source and device independent and automatic color image correction.[1,2] The aim of this framework is to generate algorithms and systems making it possible to print proper quality pictures from variable quality network originals without interactive image correction. For automatic image processing, a classifier telling whether an individual image needs correction and whether the quality will be sufficient for the intended printing process is of use. The former depends on the quality rendition of an image, the latter on its quality potential. The purpose of this study is to find out basic principles to measure quality potential automatically from digital real scene color image signals.

For a digital image signal, the quality potential of a pixel is, in bits, the bit depth as such. The final quality potential for an individual image may, however, be reduced due to reduced dynamic range, increased point spread, or increased noise. The measurement of dynamic range is straightforward using the gray-level histogram if it is assumed that every natural scene contains both very dark and very bright regions. This assumption is not strictly valid; a more sophisticated estimation of dynamic range is actually an implicit part of the automatic color correction software developed in our laboratory. This article deals with other important subquestions, namely, measurement of line spread, stochastic noise, and deterministic noise caused by lossy compression.

Objective measurement of image quality has been the topic of numerous studies during the past decades. Today image transmission and lossy coding are very active fields of image quality research. The objective is to develop methods to predict the visual difference of an "original" and a "distorted" image.[3–6] Such a method can then be used to evaluate a compression algorithm, given both the original and degraded images.

This principle of comparing two images is not applicable in our case, where only one version of each picture is available during both the development and application of the methods. A human observer can judge an image as having high or low quality without any reference image, so the question arises whether this is possible for a machine too. Some inspection of typical images suggests that the human judgment often involves interpretation at a very high level. For example, one expects an image area representing grass to be "rough," so it is not judged noisy, whereas roughness in a region interpreted as sky is an indication of noisiness.

What are the design criteria of objective quality measurement? The most ambitious goal is that the results should correspond to the judgment of an "average" human observer. It is very difficult to define the average observer. Moreover, the assessment of visual quality requires knowledge of the way in which the digital code values are transformed into a physical light distribution. In the framework of quality potential measurement the final device is not known, instead one of the goals is deciding for which purposes the image is sufficiently good. The picture may also undergo some digital manipulations after the quality measurement; for instance, sharpening and tone rendering can affect the visibility of noise considerably. Hence we cannot adopt a

Original manuscript received July 22, 1997

* IS&T Member

† E-mail: Juha.Katajamaki@hut.fi, Hannu.Saarelma@hut.fi

completely visual approach, instead, in this article *repeatability* is considered the main motivation of objective measurement.

The objectiveness of quality measurement is highest when the measured quantities are purely physical parameters, such as the line spread function (equivalently, the optical transfer function or OTF) and noise. Line spread is closely related to the subjective notion of "sharpness," while (stochastic) noise roughly corresponds to subjective "graininess." For example, the OTF can be estimated from a "knife edge" image[7] (for a nonlinear process this can be regarded as a possible *definition* of OTF). Similarly, the noise power spectrum can be measured from a test patch intended to be uniform. One can therefore try to generalize the measurement of physical parameters to real images assuming that every natural scene contains both sharp edge-like features and uniform regions.

Due to the complexity of real images, the physical properties cannot be measured accurately. Then modeling enters the picture and the problem reduces to estimating a small number of parameters instead of a whole OTF or noise spectrum. This is the approach of Kayargadde and Martens,[8–10] who divide the task into two phases: first develop a method to estimate the physical model parameters, then use models of vision to convert the parameters to perceptual quality indices. This separation has flexibility as its advantage: the visual modeling in the latter phase can be altered or omitted altogether depending on the application.

It is likely that the model assumptions are not valid in some cases. Moreover it is well known that even if the physical parameters could be measured, they would not necessarily predict perceptual quality well. This would favor the perceptual approach over the physical one, but the arguments in the above paragraphs suggest the contrary view. The perceptual effects may also be very difficult to model. Some of them, such as masking of noise by image structure,[3–6] occur at a rather low level in the visual system, while others, such as the dependence of optimal sharpness on image content[11] involve high- level mental processes. Nevertheless, *the methods should desirably be such that they, even though based on physical parameters, give results that correlate positively with human perception whenever the employed model is not valid.*

The methods of this article lie somewhere between the physical and perceptual approaches. The main objective is to find algorithms that measure some simply defined quality parameters and are fast enough to be run on a regular computer. Proper consideration of the filtering in low-level vision is regarded desirable, but no attempt has been made so far toward the modeling of masking effects or image interpretation.

Background and Principles of the Methods

Sharpness. The term "sharpness" means here the aspect of image quality related to small line spread, while "blur" refers to the opposite of sharpness. (We do not deal with "tangential" sharpness, the opposite of which is "raggedness".) Some authors use the word "acutance." The measurement of sharpness has been studied extensively. On the one hand there is the vast literature on adaptive enhancement filters, all of which must implicitly do some kind of sharpness estimation. The problem is referred to as "blind deconvolution." On the other hand, many measures aim to relate known properties of the imaging system to the perceived sharpness.[12] Finally, there are the image fidelity measures mentioned in the Introduction. A brief review of the various methods has been given by Rangayyan and Elkadiki.[13]

Some techniques estimate sharpness using the spectrum.[14,15] Other methods work in the spatial domain and resemble ours.[8,13,16] The common principle of these latter methods as well as ours is to locate the edge areas in an image, compute some kind of local sharpness values at the edges, and combine the results to obtain a global sharpness value.

In the method of Inoue and Tajima,[16] the local sharpness is the absolute value of a band-pass filtered image. The edge areas are extracted by thresholding the Sobel gradient, and the global sharpness is the average of the local sharpness over the edge area. The intended application of this method is to tell which *artificially sharpened* version of an image is optimal when the amount of sharpening is varied. Thus it is unclear how well the method would measure the visual sharpness of arbitrary images. The authors do not explain in detail how the thresholding of gradient is accomplished, even though this seems a crucial point.

The method of Rangayyan and Elkadiki[13] has been shown to give good results with respect to visual sharpness,[17] but the published tests have been conducted on very simple images only. The image is assumed to consist of "object" and "background" regions, and the algorithm locates the edges between them. The local sharpness measure is computed by forming several discrete approximations of gradient in the direction of the edge normal and taking the average of the results. The global acutance measure is obtained as the root mean squared value of the local values. The objective of the authors is somewhat different from ours: not classification of individual images but rather evaluation of imaging devices or image processing operations.

The objectives of Kayargadde and Martens[8,10] match ours in the sense that natural images with only one sample per scene are being measured. Their overall approach is to first estimate the physical blur and then relate this to perceived blur. The method to measure the physical blur models the blurring kernel as a rotationally symmetric Gaussian, whose only parameter is its width. The method is an application of the theory of polynomial transforms, which are essentially filter banks consisting of Gaussian derivatives of all orders. By analyzing certain components of the transform, the locally edge-like features are detected, and a further analysis involving up to third-order derivatives gives the local edge parameters, one of which is the width of the blurring kernel. Assuming the blur is uniform over the image, the global blur value is obtained as the value that minimizes a weighted mean squared error. The authors do not present a method to obtain the global value when the blur is nonuniform. Finally, the physical blur parameter is converted to a perceptual blur index by a nonlinear transformation.

According to the present authors' experience based on hundreds of images, nonuniformity of blur is rather common. Some parts of the scene are better focused than others due to limitations of the camera optics or the photographer's intent to record depth information. Assuming that the best focused region is the basis for quality judgment, one should concentrate on the maximally sharp regions within an image.

The local sharpness measure we use is simply *the maximum edge gradient divided by the total edge height*. If the shape of the line spread function is fixed except for the spread width, then the measure in question is inversely proportional to the width. In this respect the method corresponds to the blur parameter estimation of Kayargadde and Martens.[8,10] If, however, the shape is not fixed but we take the more general (and realistic) assumption that the

line spread function is smooth and symmetric and attains its maximum value at zero, then the measure is directly proportional to the value at zero, which, in turn, is (to a constant factor) equal to the integral of the OTF (which is real and symmetric). The preceding discussion was one dimensional with the implicit assumption that the two-dimensional OTF is rotationally invariant. If this assumption holds, it follows from well-known properties of the Fourier transform that the sharpness measure is proportional to the integral of the OTF cross section along any line that passes through the origin.

A conventional method to measure sharpness is to integrate the modulation transfer function (MTF), which is the absolute value of the OTF, possibly weighted by some function modeling the frequency sensitivity of the eye.[12,18] Hence our method, apart from the weighting, is equivalent to the MTF method if the OTF does not change sign. In the current implementation visual weighting is not used but some low-pass-type weighting will in fact occur due to the finite resolution.

The measurement of normalized maximum gradient is not new in itself; it has been used, e.g., by Lourens, Du Toit, and Du Toit,[19] but it is unclear from their study how the normalization is actually done with real images. Other measures have been proposed to predict perceptual sharpness better. One of these is the acutance as defined by Higgins and Jones and reviewed by Rangayyan and Elkadiki.[13] In the frequency domain it is equivalent to the integral of the squared MTF multiplied by the edge height, so edges of high contrast get a higher sharpness than edges of low contrast. In spite of this and similar models we use the above-mentioned measure because no general agreement exists on which model is the best.

Measurement of the edge gradient as well as the edge height requires extraction of the edge locations, which is commonly done as a separate step.[8,13] In the method presented here, a very rough edge detection stage based on thresholding of gradient is only performed to find the initial edge candidates. The subsequent algorithm *simultaneously* examines the candidates to extract edge pixels and measures the local sharpness at these pixels.

Stochastic Noise. In this subsection we only deal with noise that is isotropic and at least locally stationary.

Estimation of noise is an essential part of adaptive image enhancement algorithms, where a tradeoff must be made between enhancing the details and suppressing the noise. The design of such algorithms is, however, directed at optimal quality of the final image and the algorithm may be robust with respect to the estimated value, so the noise estimation may have little to do with the original visual quality. Segmentation and machine vision are other fields where noise estimation is needed.[20,21] Many authors have, though, discussed the noise estimation as a separate problem.[9,22–24] A comparison of some methods has been made by Olsen.[25]

The common assumption, because of its mathematical simplicity, is that the noise is additive with zero mean, Gaussian, and spatially uncorrelated (white). The combination of additive and multiplicative noise, the latter arising from the Poisson statistics of photons, is discussed by Lee and Hoppel[22] as well as Brügelmann and Förstner.[21] The method of Kayargadde and Martens[9,10] measures both the variance and the correlation length of noise by modeling the noise as a Gaussian random process with Gaussian power spectrum.

In the case of noise the properties of an analog device are particularly important. The visibility of noise is affected by the resolution in angular units, and this depends both on the pixel frequency of the device and the viewing distance. Although noise can be "white" in the discrete domain, it will always have some finite bandwidth when the image is reproduced on a device for human viewing. Tone reproduction is another crucial issue; for example, a dark region is perceived noisier than a bright region when the noise variances in luminance units are equal.

The simplest objective measure of white noise is some first-order statistic such as the variance. To measure the variance from real images, the image structure must, however, be suppressed by appropriate filtering. By analyzing the histogram of the filtered image, the variance can be estimated. For instance, the magnitude of the gradient of Gaussian white noise follows a Rayleigh distribution[9,20,25] whose variance parameter can be estimated from the gradient histogram of a real image by curve fitting, assuming the image contains sufficiently large homogeneous regions.

However, in practice the noise power spectrum varies from image to image, and with typical image resolutions this variation occurs even within the bandwidth of the human eye. Methods that employ some filter to measure white noise actually measure the variance of the *filtered* noise, so the methods may behave incorrectly in regard to perceptual noise when the whiteness assumption does not hold. This is the case, for example, with the high-pass-type filter proposed by Immerkær,[24] although this filter is able to eliminate the image structure very efficiently and hence produce an accurate estimate when the noise *is* white. Two approaches are possible to measure perceptual noisiness correctly. One approach, used, e.g., by Kayargadde and Martens,[9,10] is to first estimate the parameters of a modeled noise spectrum and then calculate a perceptual noisiness index from the knowledge of the spectrum. The other approach, which is adopted in this article, is to use a filter that mimics the response of the eye in such a way that the standard deviation of the filtered image directly corresponds to perceptual noisiness.

In digitized natural images, noise level often varies between different spatial regions, especially between areas of different brightness. This may be a characteristic of film or scanner or it can result from nonlinear tone reproduction that amplifies noise at some gray levels and suppresses it elsewhere. Unfortunately, not even the multiplicative noise model[21,22] is adequate. For instance, the noise level may be low (or even zero, due to out-of-gamut clipping) in the dark as well as bright regions but higher at the midtones.

Thus the histogram of the filtered image may not be unimodal, as demonstrated in Fig. 1. One can, however, assume that each distinct peak in the histogram corresponds to either the noise in a large homogeneous region or to a textured region. The discrimination between noise and texture is a nontrivial subject as such, and although various approaches are available, the grass versus sky example in the Introduction suggests that the solution is not straightforward. At present our noise estimation method does not recognize the texture in any sophisticated way. Instead it analyses the peaks in the histogram of a filtered image and, on the basis of both the absolute and relative positions of the peaks, it tries to decide which of them are likely to have arisen from noise. The main problem then is that to prevent texture from being misinterpreted as noise, the peaks with *lowest* position should only be accepted, but at the same time the method should be able to identify the *highest* noise levels present in the image, because they are usually significant for overall quality.

One could think of taking, e.g., the maximal detected noise level as an overall measure of noisiness. However,

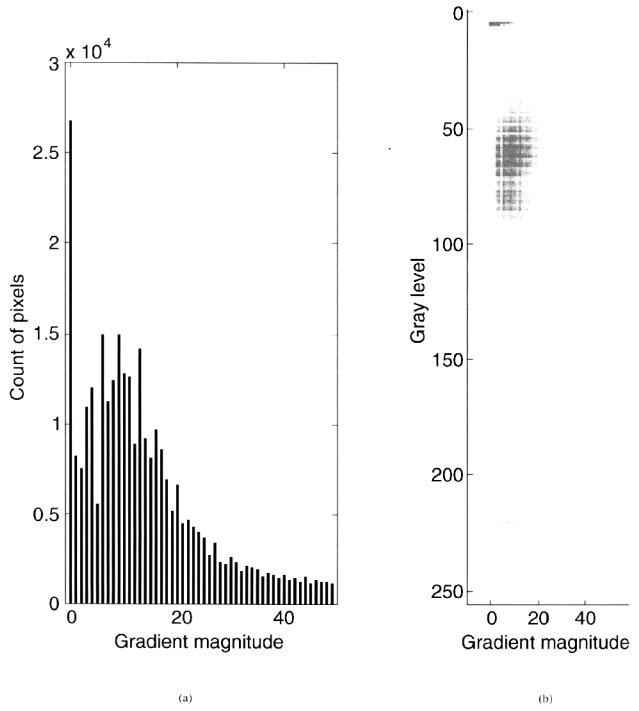


Figure 1. (a) An example of nonunimodal gradient histogram. The histogram has been computed from the magnitude of the gradient defined by the Sobel filters. (b) The bivariate histogram of gradient and luminance reveals that the noise level depends on the luminance. The image contains an almost black region of practically zero noise and a gray and a white region with approximately equal noise levels.

the question is complicated by the fact that the visibility of noise depends considerably on local color, especially on lightness. Thus, unless the final color reproduction is exactly known, it cannot be determined which of the noise levels will be visually maximal. But it is still possible to measure the local color of noisy regions and save it to be used at a later stage where the color reproduction, and hence the correct weight for the noise visibility, will be known. Consequently, *the goal of our noise measurement method is not a single noise index, but a set of noise levels, each of which is associated with the corresponding mean color*.

Currently the method only measures the noise of the luminance component of a color image, and the associated local mean also includes the luminance only. Strictly speaking, this is not enough to predict perceptual noisiness, but this is one of the points where simplicity is favored over accuracy. In practice the luminance noise is often dominant both physically and perceptually and the mean color of large homogeneous regions is frequently close to a neutral gray.

In view of the above discussion, we prefer the "visual" filtering approach over the spectral parameter estimation approach, because it would be complicated to first identify every distinct noise level and then estimate the parameters separately for each. Hence, the main parts of the method are the following:

1. Extract the luminance component of the color image and process it with a filter that approximates the frequency response of the human eye.
2. Analyze the joint statistics of the luminance image and the filtered image in order to identify the noise levels and the mean luminances of significant homogeneous regions, while excluding textured areas.

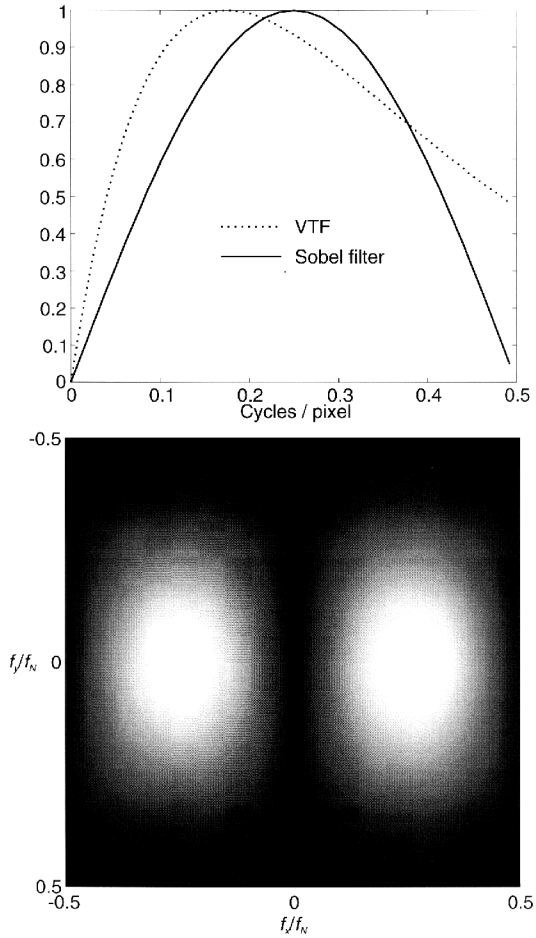


Figure 2. Comparison of the magnitude response of the Sobel filter and a model for the visual transfer function (VTF). The horizontal Sobel filter has response $H(f_x,f_y) = -4\ \sin(\pi f_x/f_N)\ (1+\cos(\pi f_y/f_N))$, where f_N is the Nyquist frequency. The VTF is defined according to Ref. 27 as $5.05\exp(-0.138u)\ (1-\exp(-0.1u))$, where u is the spatial frequency in cycle/degree. Part (a) shows the cross section of $|H/8|$ with $f_y = 0$ and $0 \le f_x \le f_N$ and the VTF evaluated at viewing distance 35 cm and resolution 100 pixel/inch. Part (b) is an image of $|H|$ over the frequency plane.

To select an optimal visual filter, the image resolution, viewing distance, and the OTF of the display device should be fixed. Because this cannot be done in a standardized way and because no standardized model exists for the visual transfer function (VTF), we simply use the Sobel gradient[26] defined as

$$G = \frac{1}{2}\left(|S_x * L| + |S_y * L|\right),\quad S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix},\quad S_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \quad (1)$$

where L denotes the luminance image. The main reasons for using this filter are computational efficiency and the fact that the gradient can be utilized in other ways in quality measurement (cf. Fig. 5). Nevertheless, the magnitude of the Sobel filters in the frequency domain is a reasonable approximation to the visual transfer function at least for low resolutions, as shown in Fig. 2. The preferred orientation of the Sobel filters has no perceptual justification, but under the assumption of noise isotropicity the angular behavior is not as important as the radial behavior.

Because of the possibility of noise nonuniformity, we do not use any parametric model for the histogram of the

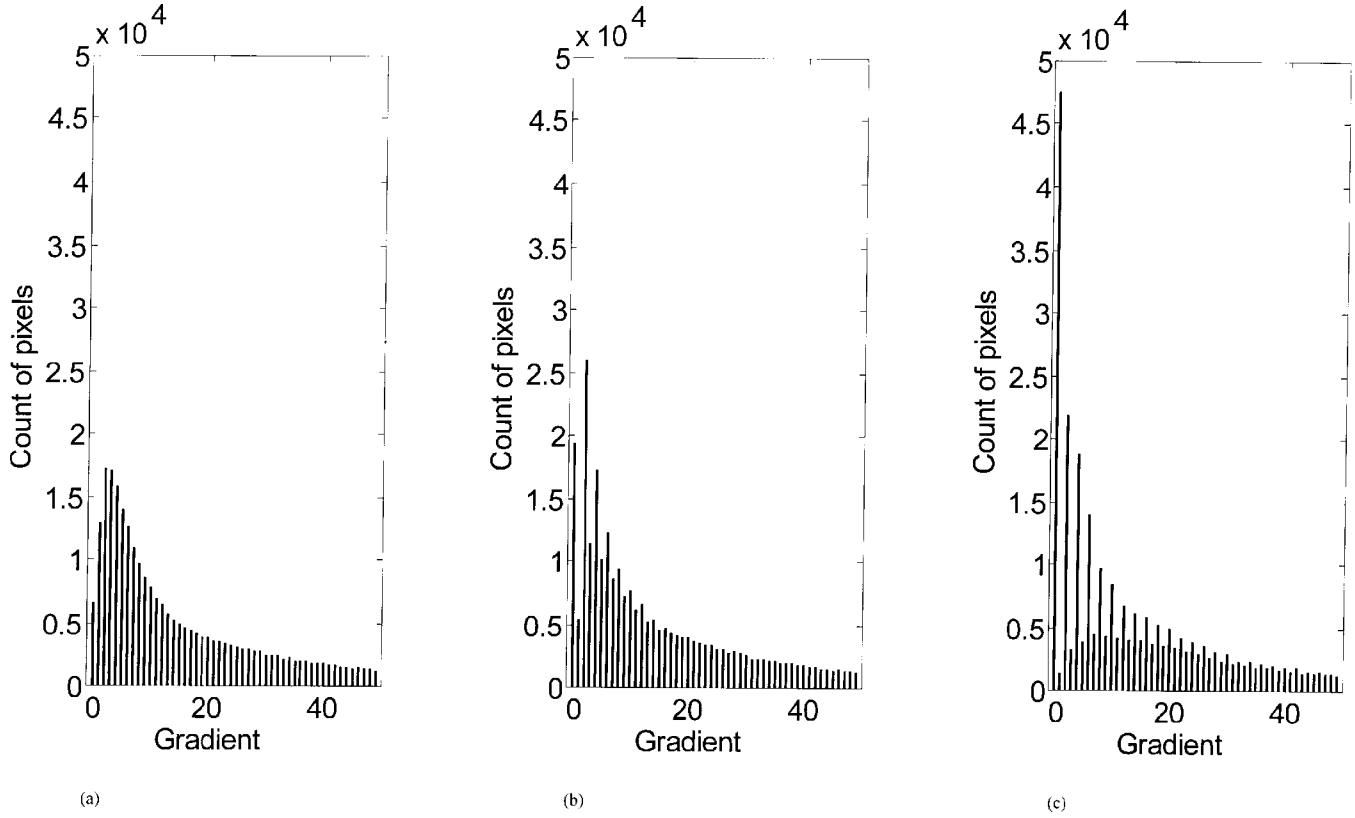


Figure 3. Effect of JPEG compression on the histogram of Sobel gradient (Eq. 1). The lower part of the histograms is only shown. (a) Uncompressed image. (b) Compressed image with moderate compression ratio. (c) Compressed image with high compression ratio.

gradient image. After the gradient image has been computed, it is smoothed. The smoothing reduces the variance of the gradient in homogeneous regions, making multiple peaks in the histogram of the gradient more discernible from each other and facilitating their separation from signal gradients. Because the smoothing also reduces the skewness of the peaks, the position of a peak after smoothing will be approximately equal to the mean absolute Sobel gradient (either horizontal or vertical) over a homogeneous area. (This mean value is also approximately proportional to the standard deviation of gradient, because the Sobel filter output of noise can be assumed nearly Gaussian; more specifically, the mean of the absolute value of a Gaussian variable is $\sqrt{2/\pi}$ times the standard deviation.)

The luminance image is also smoothed. Then the bivariate histogram of the luminance and the gradient images is constructed. The 2-D locations of pronounced peaks in this histogram are found and recorded as representing the potential noise levels of the image with their associated local luminances. This sequence of peaks is further analyzed to exclude those likely to have arisen from texture. The algorithm is based on the assumption that if there are several peaks with approximately equal luminance values, then only the one with the lowest gradient is likely to represent noise. This assumption is a necessary tradeoff, but of course it is not always valid, because regions of the same luminance but different color may have different noise levels. The method of Lee and Hoppel[22] is similar in that it utilizes the joint distribution (scattergram instead of histogram) of local mean and local variance, but the dependence of the variance on the mean is assumed a combination of multiplicative and additive noise.

Coding Artifacts. In images from digital sources, deterministic noise due to lossy coding can be a very significant factor of quality. The block discrete cosine transform (BDCT), while not being the best coding method visually, is widely used because of the JPEG compression standard. Images coded with low bit rate BDCT have a typical blocky appearance. A lot of research has been done to measure objectively the image quality by comparing the uncompressed and reconstructed images. There are also published methods to estimate the strength of the "blocking artifact" from the reconstructed image only.[28,29]

In the general framework of this article it not necessarily known whether a digital image has been compressed at some stage in its history. It is important to be able to recognize BDCT coded images, because the unusual statistics of such images may cause a systematic error in the sharpness and noise measurement algorithms. A strong compression artifact is an indication of poor quality as such and detection of it allows the other quality measurements to be interpreted correctly or even omitted. We do not elaborate on visual models to estimate the blockiness, but we present a very simple method that is a by-product of the other quality estimation procedures, namely, a method based on the histogram of the Sobel gradient.

Figure 3 shows how the histogram of gradient defined by Eq. 1 changes as the bit rate of the JPEG compression is reduced. The color image has 24 bits per pixel and the gradient has been computed from the luminance component. It is seen that even values of the gradient dominate over odd values the lower the bit rate is, especially at low gradients. (It is easy to show that with integral pixel values the gradient values are always integers despite the division by two in Eq. 1.) This phenomenon can

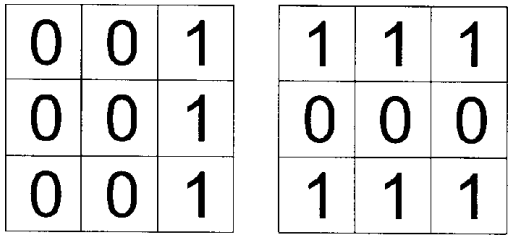


Figure 4. Examples of bit blocks of size 3×3 invariant in either horizontal or vertical direction. These occur with large frequencies in highly compressed images. The output of the gradient filter (Eq. 1) is even for all of these.

be explained by examining how the least significant bit (LSB) of the gradient value depends on the local 3×3 block covered by the filter masks. First of all, the LSB of the gradient can be shown to depend on the LSBs of the pixels only. There are $2^9 = 512$ possible combinations of the bits, and one can construct a 512-bin histogram by counting the number of occurrences of each combination over the whole image. For a typical uncompressed image each combination will occur with practically equal frequency, provided that there is no completely noiseless area. For a JPEG compressed image, however, "one-dimensional" bit configurations such as those in Fig. 4 are more frequent than others. This results probably both from the block structure and the fact that oblique spatial frequencies are quantized to a smaller number of bits than horizontal and vertical ones.

Therefore, to detect the BDCT artifact, the histogram of gradient is analyzed to calculate an index describing the amount of disproportion between even and odd values. So the method does not measure any visible artifact directly, but assumes that in practice this kind of phenomenon is an indication of *some* kind of quality degradation. Note that the method, unlike methods designed specifically for JPEG, is applicable even to images resized after reconstruction.

The method by Datcu[28] is similar, but it is based on the "spikiness" of gray level instead of gradient histograms. However, spikiness in the Sobel gradient seems more easily detectable in complex images. Of course, no fundamental reason exists to use the Sobel filters. For example, the above-mentioned histogram of LSB configurations could

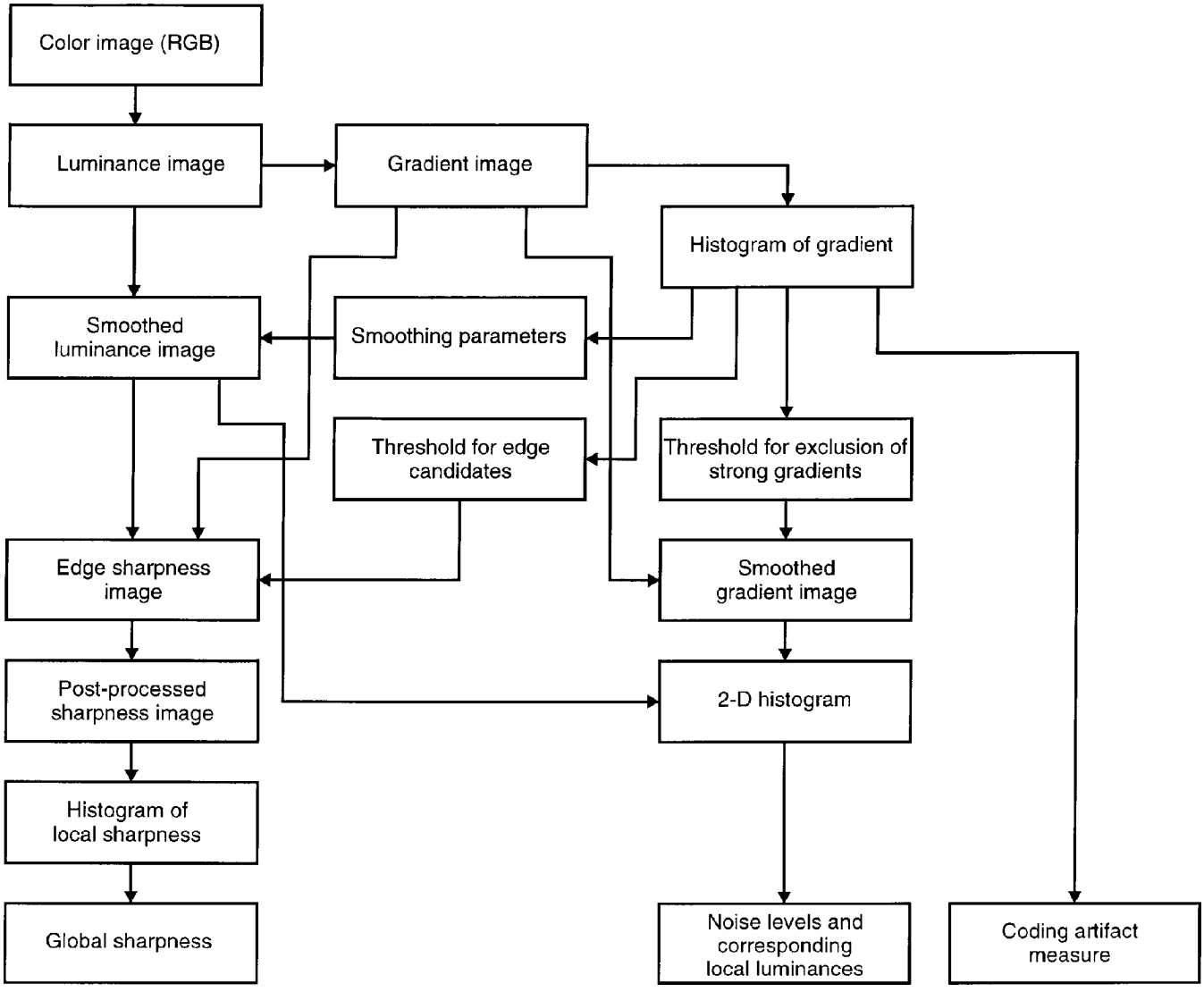


Figure 5. Data flow of the quality measurement algorithm. The boxes represent various computation results and the arrows indicate the input data used to compute a result.

be used; it would actually reveal some other types of coding artifacts as well.

Quality Measurement Algorithms

Overall Structure. Figure 5 is a diagram of the overall algorithm, showing the auxiliary images and histograms common to the measurements. First the luminance component of an RGB image is computed by

$$L = k_R R + k_G G + k_B B \quad (k_R + k_G + k_B = 1) \tag{2}$$

and all subsequent processing is done for the luminance image. The gradient image is computed according to Eq. 1 and its histogram is constructed. This histogram itself is the input to the coding artifact measurement. The noise measurement uses the gradient image and its histogram to exclude strong gradients, and the sharpness measurement uses them to select the potential edges.

The noise algorithm is executed last so that the smoothed gradient image used in the noise algorithm can replace the original gradient buffer. The sharpness algorithm smooths the luminance image to suppress noise, overwriting the original luminance buffer. This same smoothed image together with the smoothed gradient image is used by the noise algorithm to construct the 2-D histogram, in which the noise levels with associated gray levels are looked for. The sharpness algorithm produces an image representing the local sharpness values, and the histogram of this image after postprocessing serves as the basis for the global sharpness index.

All the auxiliary images have a precision of 8 bits.

The principles of the algorithms were already discussed in the previous section. The following subsections will give more detailed descriptions of the nonobvious computations in Fig. 5.

Smoothing of the Luminance Image. The method to smooth the luminance image is aimed at the suppression of noise in the sharpness measurement, although for efficiency the same image is used in the noise measurement as well. Because we prefer methods that are simple to implement, a combination of median filtering and binomial low-pass filtering is used. The problem with noise is twofold: first, noise may ruin the operation of the edge analysis algorithm (see below), and second, noise increases the variance of the local sharpness values, thus increasing the maximum sharpness which is our interest. For sharp edges, where the luminance transition occurs practically within one pixel, median filtering suffices. For blurred edges median filtering does not adequately suppress noise over the whole edge zone, which is several pixels wide. Therefore the histogram of gradient is used to obtain a preliminary sharpness estimate. For high values of this estimate only median filtering is used and for low values only low-pass filtering is used. In the intermediate range a smooth transition is made by taking a weighted combination of the original and median filtered image and by varying the size of the low-pass mask. Whenever low-pass filtering is performed, it is necessary to correct the bias caused in the sharpness values. This is done by modeling the original blurring kernel as Gaussian with spread parameter b:

$$K_b(x,y) = \frac{1}{b\sqrt{2\pi}} \exp\left(-\frac{x^2+y^2}{2b^2}\right). \tag{3}$$

The binomial filtering is also a discrete approximation to Gaussian blurring with some spread parameter b_{lp}. The spread parameters of repeated Gaussian blurring combine in a "root of sum of squares" manner. Hence an estimate for the original blur, given the estimate $\hat{b}_1$ obtained from the filtered image, is

$$\hat{b} = \sqrt{\hat{b}_1^2 - b_{lp}^2}, \tag{4}$$

and the corresponding sharpness estimate is the inverse of this value. But this low-pass method with correction would not be applicable to very sharp images, because it assumes the resulting measured blur is greater than the blur of the filter, which might not be the case in practice.

The implementation of the median filtering is based on updating the histogram of the pixels in a gliding window, which is considerably more efficient than a straightforward method. The binomial filter is also very efficient, because it is separable and can be implemented using additions followed by a shift operation.

Edge Analysis for Local Sharpness. As stated in the Background and Principles Section, the quantity to be measured is

$$\text{local sharpness} = \sqrt{G_x^2 + G_y^2} \,/\, H, \tag{5}$$

where G_x and G_y are the components of the gradient at an edge and H is the total height of the edge. To apply Eq. 5 we need to find those locations in an image that are really "edges" and not "lines" or "points." We assume that the edge is sufficiently straight locally and that the cross section and its second derivative have the qualitative shape shown in Fig. 6. Let the cross section be denoted by $f(x)$ and let a point x_0 be fixed. The decision whether x_0 is an edge point is done by evaluating the difference function

$$D(s) = \left| f(x_0 + s) - f(x_0 - s) \right| \tag{6}$$

for values of s ranging from zero to an upper limit $s_{\max}$. The following is true when the model of Fig. 6 is valid:

1. D is a monotonically increasing function of s for every fixed x_0. The value D tends to H as $s \to \infty$.
2. There is a distance a_1 such that when the distance from x_0 to the edge center is smaller than a_1, then D is a

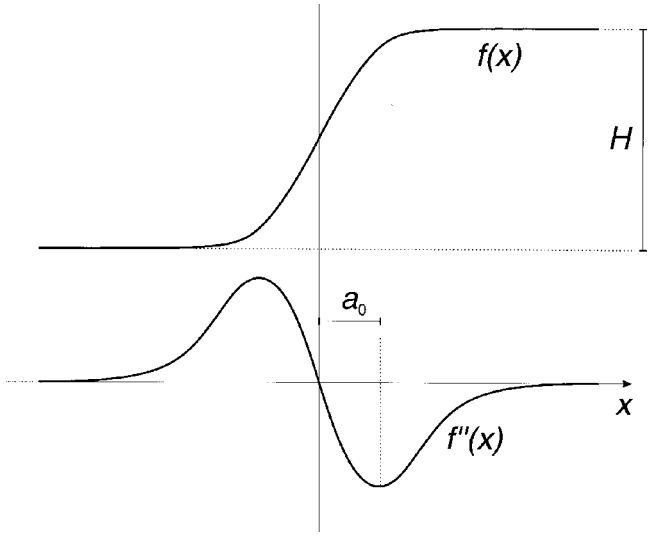
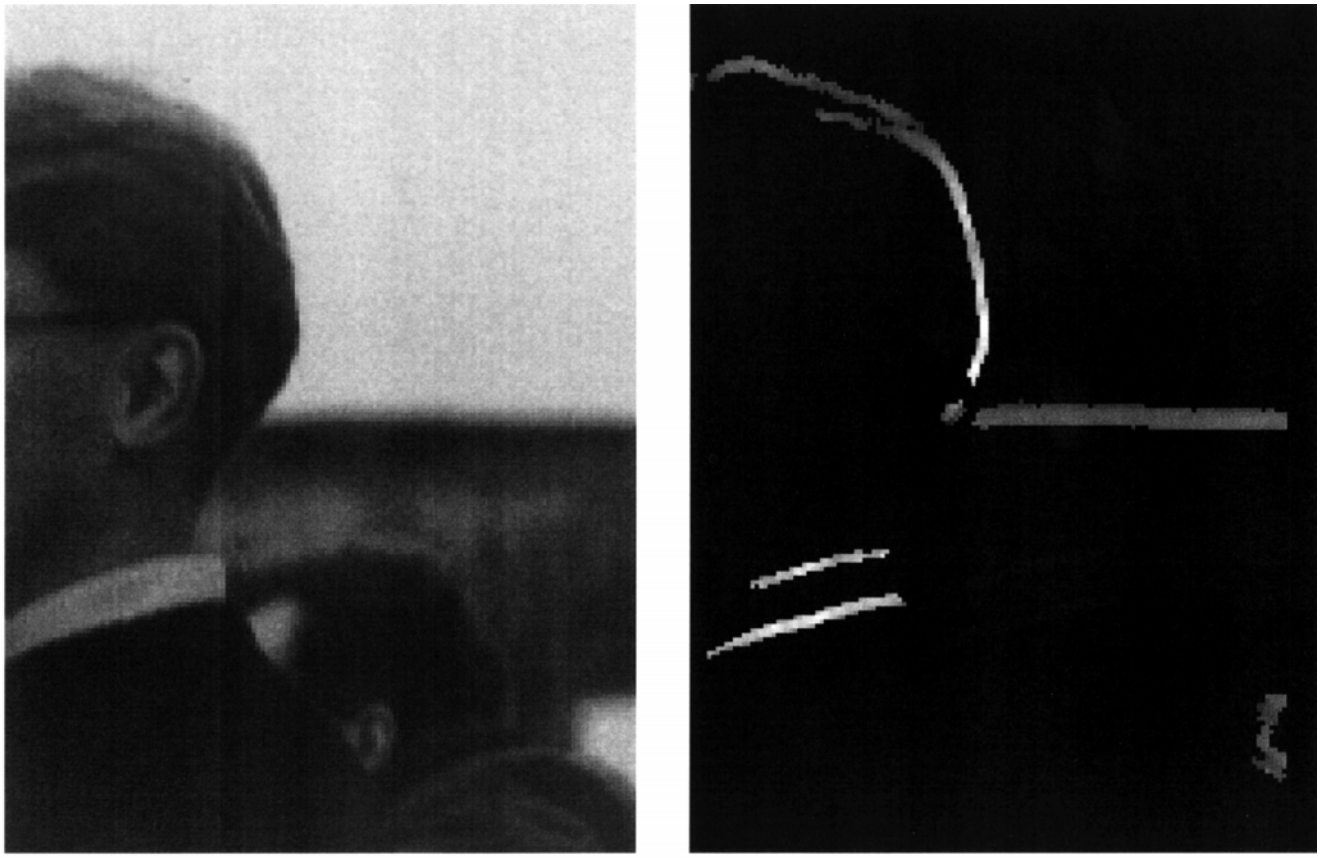


Figure 6. Assumed shape of an edge cross section and of its second derivative. See text for explanation of the symbols.



(a) (b)

Figure 7. (a) Part of a luminance image. (b) The corresponding sharpness image, where brightness is proportional to sharpness. Note that there are both sharp and blurred edges.

concave function of s, i.e., $\partial^2 D/\partial s^2$ is negative. This is seen from Fig. 6 by noting that $\partial^2 D/\partial s^2$ is equal to $-|f''(x_0+s)-f''(x_0-s)|$. For example, a_1 can be defined as $a_0/2$, where a_0 is the position of the extremum of f''.

3. There is a distance a_2 such that when the distance from x_0 to the edge center is at least a_2, there exist values of s where D is not concave. For example, let $a_2 = a_0 + \varepsilon$ with ε positive; then $\partial^2 D/\partial s^2$ is positive whenever $0 < s < \varepsilon$.

These properties are applied pixel by pixel. The direction for evaluating D in the neighborhood of a given pixel must first be determined. To simplify the implementation, one of four directions, namely, the horizontal, vertical, and the two diagonal directions, is selected on the basis of local gradient. Then D is evaluated for several values of s starting at zero. If D turns out not to be monotonic or concave at some value of s, it is known by properties 1 and 2 that the current pixel cannot be an edge point. But if D is monotonic and concave for every s, the current pixel is accepted as an edge point. For the accepted pixels, the latest computed value of D is taken as an estimate for the height H, and using this with the gradient at x_0, the local sharpness estimate is computed from Eq. 5. According to property 2, the edge area detected in this way contains a narrow zone with width roughly proportional to the line spread. But property 3 guarantees that the accepted area is not too large. Having an edge zone instead of edge line causes no problems, because the maximal sharpness over an image is only of interest. The sharpness is locally maximal at the precise edge location, because the numerator of (5) is highest there while the denominator H is approximately constant over the zone.

In reality the monotonicity and concavity conditions must be relaxed because of noise; they are only required to hold to a certain tolerance. To reduce the possibility of false edge detection there is also a lower limit on the acceptable values of H. The parameter $s_{\max}$ is large enough to allow the correct measurement of H even for the least sharp edges occurring in practice.

To store the local sharpness image with 8 bits, the sharpness values from Eq. 5 are multiplied by 100 and the nonedge pixels are marked with zeros. When the edge profile approaches a step function, the sharpness saturates to 100, because the resolution imposes an upper limit on the discrete gradient. This corresponds to low-pass filtering of an analog image before the sharpness estimation, so the method is qualitatively consistent with the corresponding filtering effect of the eye. Due to the tolerance mentioned in the preceding paragraph, the method also has capability of measuring "overshoot" edges, which are common in, e.g., artificially sharpened images. The larger sharpness value obtained for overshoot edges is consistent with perceptual sharpness as well. (For very sharp edges with overshoot the local sharpness can be even greater than 100.)

To speed up the algorithm it is not run for every pixel, but candidates for edge pixels are first selected by taking a fixed percentile from the high end of the gradient distribution. Because the maximum gradient at an edge depends not only on the line spread but on the edge height as well, some edges of interest may be missed with this method. However, for most practical images enough edges are included.

Postprocessing of the Sharpness Image. The monotonicity and convexity properties of the function D discussed

in the previous subsection are only necessary, not sufficient, conditions for the presence of an edge. For example, if any function that is symmetric with respect to x_0 is added to f, the values of D are not altered. Because of this and the required relaxation of the conditions, there will usually be some pixels misinterpreted as edges. These are often scattered over the image and in the current implementation they are removed by finding the connected regions among all pixels that were accepted as edges, and deleting very small regions.

Figure 7 shows a portion of an image and the corresponding sharpness image.

Determination of Global Sharpness. According to the discussion in the second section of this report, global sharpness is defined as the maximal local sharpness. In practice the strict maximum would be too nonrobust, so a fixed percentile from the high end of the local sharpness histogram is used instead. The percentile is calculated over the edge pixels only (i.e., the nonzeros in the sharpness image).

Smoothing of the Gradient Image. The smoothing of the gradient required in the noise measurement is basically spatial averaging over a small area. The area should be as large as possible, but simple averaging over a large area would result in "spreading" of strong edge gradients, causing the reduction of homogeneous regions, which may already be small in some images. To prevent this, a binary mask image is first constructed to separate very strong edges from other regions. The binary image is initially created by thresholding the unsmoothed gradient, and is subsequently "cleaned" by median filtering. The threshold selection currently utilizes the histogram of the gradient only. It first transforms the histogram frequencies by a nonlinear function that lowers large frequencies in relation to small frequencies and then computes a "percentile" from the transformed histogram.

The spatial averaging is performed using a fixed size square window, but the average is calculated over only those pixels in the window that were not marked as edges in the binary image. At the same time the edge pixels in the output image are marked with value 255 so that the gradient/luminance histogram can be computed over the whole image without disturbing the analysis of peaks. In this method of smoothing, spreading can still occur between two homogeneous regions of *different* mean gradients and separated by an edge, if the width of the edge is smaller than the window size, but in practice this effect is much smaller than the spreading of the edge gradient.

The smoothing operation allows very efficient implementation, in which the sum over the nonedge pixels and the number of them within the averaging window is updated as the window glides over the image.

Detection of Peaks Representing Noise in the Gradient/Luminance Histogram. The task is to analyze the bivariate histogram of smoothed gradient and smoothed luminance to find peaks likely to have arisen from large smooth regions and to discard peaks that probably represent texture. The coordinates of an accepted peak are measures of the noise level and the mean luminance of the corresponding region.

The algorithm to find the peaks contains many details and is still a potential subject for development, so we do not give a complete description of it. There are two main objectives: firstly, small gradients should be favored but simultaneously the existence of different gradients at different luminance levels should be recognized; secondly, even small homogeneous regions should be detected because these may be the only possibility of noise measurement for some images. The structure of the histogram is analyzed on the basis of the 1-D marginal distributions, but during the course of the algorithm partial histograms are examined, which makes it possible to detect weak peaks that could not be discerned in the global marginal histograms.

After the execution of this algorithm there can still be some false noise recordings left. Although perfect discrimination between noise and textures is not possible by this method, the probability of error can be reduced with rules that constrain both the absolute and the relative location of the noise peaks in the gradient/luminance plane. Currently the following post-processing algorithm is used: The

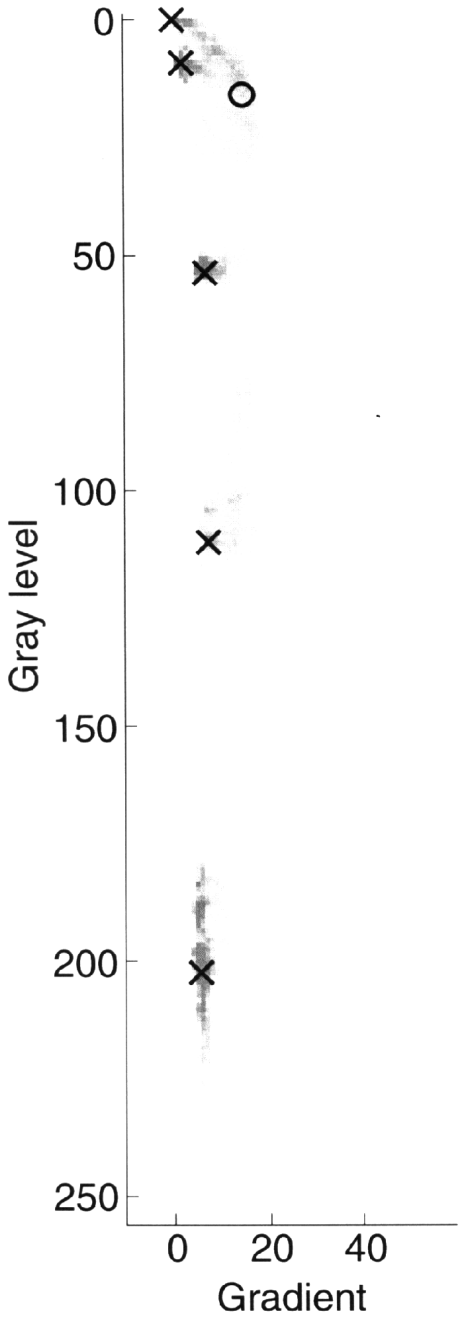
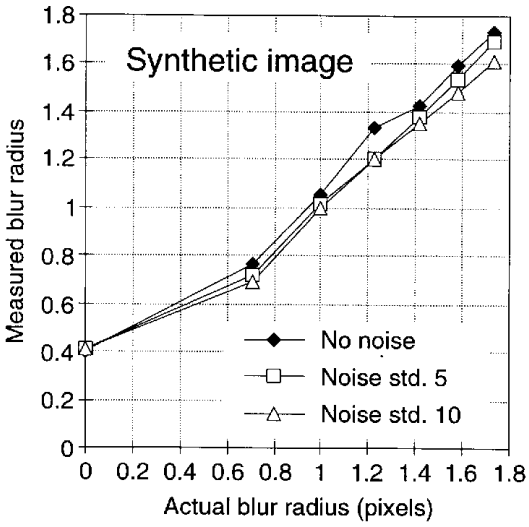
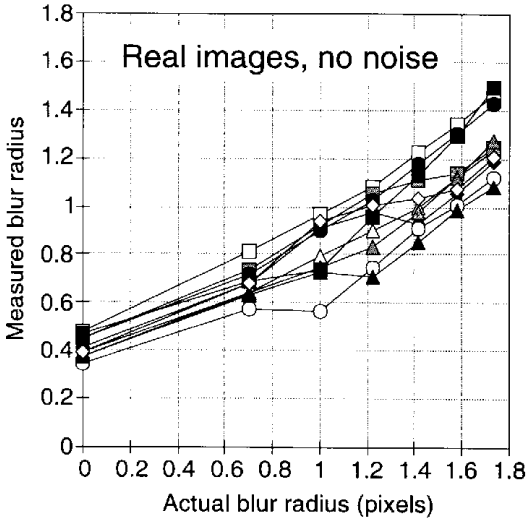


Figure 8. A bivariate histogram of smoothed gradient and smoothed luminance. The crosses indicate peak candidates that were accepted as noise. The circle is a peak not accepted because its gradient value was too different from that of its neighbors.





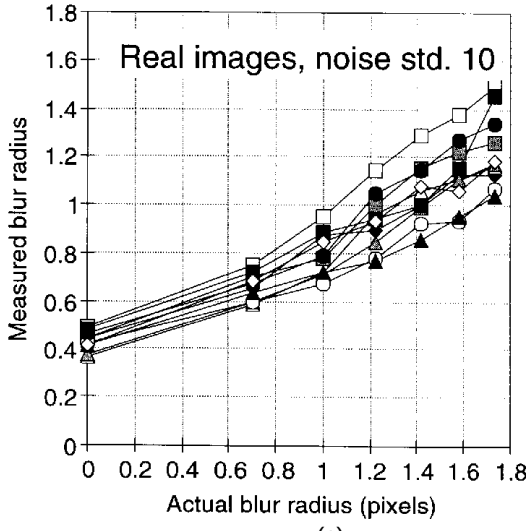


Figure 9. Testing of the sharpness algorithm with artificial blur. (a) Synthetic edge image. (b–c) Natural images without and with added noise. Each curve represents the various versions of a single image.

noise peak candidates are scanned in ascending order of gradient. If the gradient value is greater than a threshold, or if it is too large in relation to the gradient of those candidates that are the nearest neighbors along the luminance axis, the peak is rejected. The rejected peaks are excluded from all subsequent examination. It is possible that after this algorithm there no peaks are left, i.e. the image does not contain any reliable region for noise measurement.

Figure 8 shows a sample 2-D histogram and the accepted and rejected peak positions.

Calculation of the Compression Artifact Measure. To describe the disproportion between even and odd gradient values, two indices are first calculated from the histogram of the gradient. One of the indices measures the amount of disproportion and the other the upper limit of the gradient range where the phenomenon is present.

Each histogram frequency is compared with the interpolated frequency obtained as the average of the two immediate neighbors of the current histogram bin. When the artifact is present, the averaged value is greater than the actual value at odd gradients and less than the actual value at even gradients. The histogram is scanned starting from the low gradient and stopping when the mentioned conditions fail to hold. During the scanning the absolute differences between the actual and interpolated counts are summed and this sum, properly normalized, defines the first of the indices. The gradient value at which the scanning stops is the second index.

The final measure for compression artifacts is a linear combination of the two indices, with the relative weight selected optimally according to experimental results.

Performance of the Algorithms. The justification of the methods described in this article lies mainly in the theoretical background discussed in the Background and Principles Section. The main purpose of experimental testing is to find out how nonidealities occurring in real images affect the measures. There are also some requirements regarding the perceptual meaningfulness of the measures. Results concerning both of these aspects are presented in this section.

Sharpness. The effect of noise and other nonidealities on the performance of the sharpness estimation algorithm

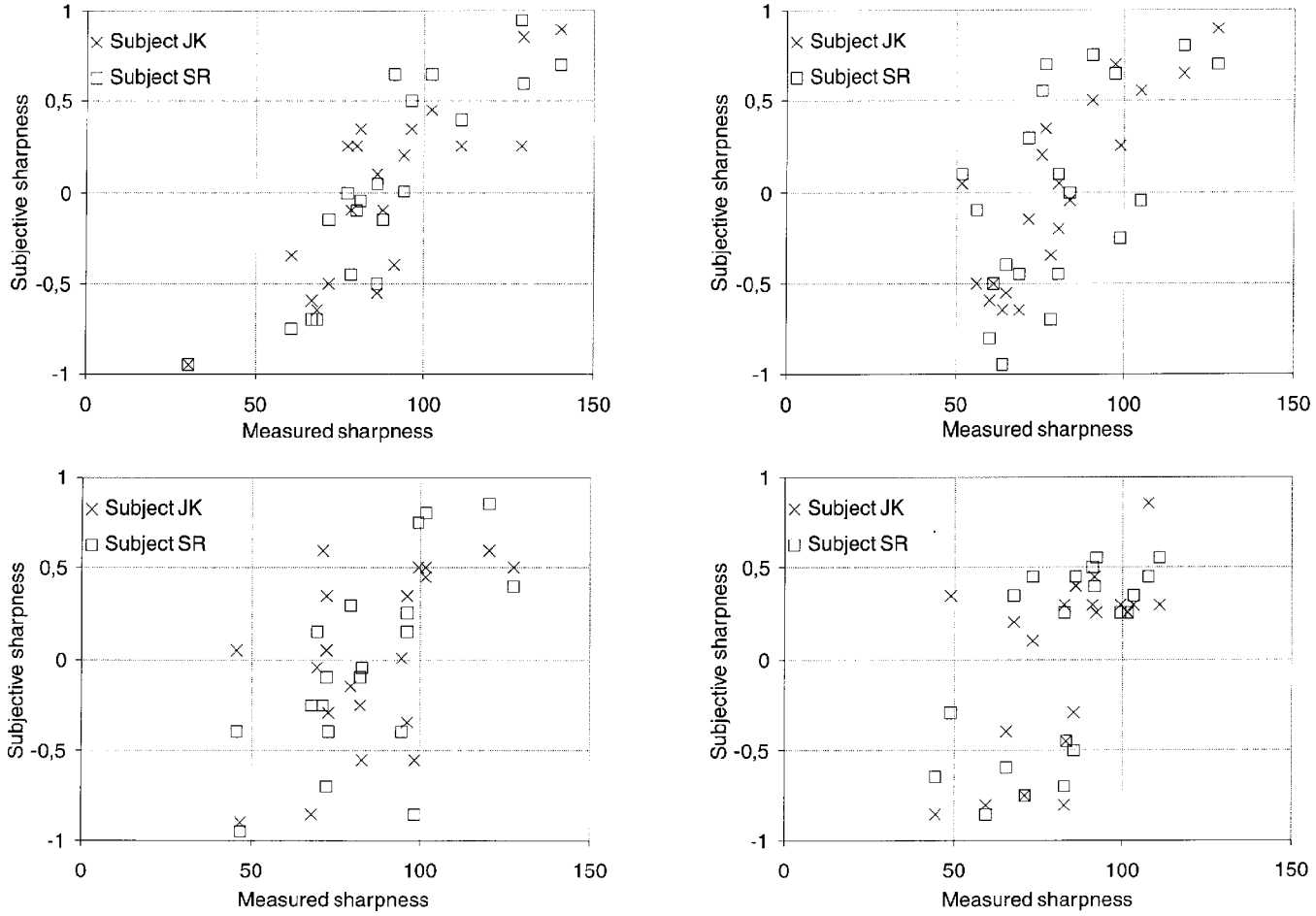


Figure 10. Subjective vs. measured sharpness for four separate sets of 20 different real images.

has been tested both with a synthetic edge image and 10 different natural images. Each image was in 24-bit RGB format. The edge image had size 128×128 and consisted of two adjacent uniform neutral regions of gray levels 50 and 200 separated by a vertical edge. The natural images were originally very sharp and their dimensions were about 400 to 800 pixels.

From each original image, six blurred versions were generated with increasing amount of blur, and more versions were generated by adding noise to the original and the blurred versions. The blurring was performed by filtering each color channel with a filter that is the separable product of 1-D binomial filters of equal radii. The radii used were 1 to 6 pixels. The filter with radius r is a discrete approximation to the Gaussian filter in Eq. 3 with

$$b = \sqrt{r/2}. \tag{7}$$

The noise was white and Gaussian with standard deviation of 5 or 10 gray levels and was added independently to each channel.

The global sharpness values measured with the algorithm were inverted and scaled by $\sqrt{2\pi}$ to be expressed in the same scale as b in Eq. 7. Figure 9 shows this measured blur versus the approximate true blur computed from Eq. 7. (Note: because the luminance image is a linear combination of the RGB channels, it is blurred in the same way as the channels are.)

The nonzero blur value measured for the unblurred images is primarily a consequence of the saturation due to finite resolution discussed above in the subsection Edge Analysis for Local Sharpness. Slight overestimation of sharpness due to noise is present with the synthetic images, but with the natural images this is overshadowed by a larger overestimation due to the image structure. There are also systematic differences between scenes. These result probably both from slight differences in blur between the original versions and differences in the ideality of edge features.

To test the visual reasonableness of the sharpness value between *different* images, subjective pair comparison experiments were conducted. Each single experiment was done with 20 original 24-bit images of different scenes and of varying quality. There are $20 \times 19 / 2 = 190$ possible pairs of the images, and each pair was displayed in random order on a calibrated monitor with a gamma of 1.0 and resolution of about 100 pixel/inch. The subject's response to each pair was one of the alternatives "picture 1 sharper," "picture 2 sharper," and "no difference." The comparison results were gathered into quality scores, scaled between –1 and 1, which tell the subjective ordering of the 20 images. The test was executed by two subjects and repeated for several picture sets. The results in Fig. 10 show a correlation between perceptual and measured sharpness but it is not remarkably high. Because comparison of different images is very inexact by nature, a perfect correlation cannot be expected. Nevertheless, the sharpness algorithm still does have some problems caused by the structure of real images.

Stochastic Noise. Meaningful perceptual scaling of natural images with respect to noisiness is even more difficult

TABLE I. Noise Types Used in the Artificial Noise Experiment

Number of noise type	Expression of power spectrum (f_N = Nyquist frequency)	Total standard deviation of noise
1	$S(f_x,f_y) = \exp((f_x^2 + f_y^2)/2\sigma^2)$, $\sigma = 0.5f_N$	2, 4, 6, and 8 gray levels
2	$S(f_x,f_y) = \exp((f_x^2 + f_y^2)/2\sigma^2)$, $\sigma = 0.75f_N$	2, 4, 6, and 8 gray levels
3	$S(f_x,f_y) = \exp((f_x^2 + f_y^2)/2\sigma^2)$, $\sigma = 1.0f_N$	2, 4, 6, and 8 gray levels
4	$S(f_x,f_y) = a^2/((a^2 + f_x^2 + f_y^2)$, $a = 0.5f_N$	2, 4, 6, and 8 gray levels
5	$S(f_x,f_y) = a^2/((a^2 + f_x^2 + f_y^2)$, $a = 0.75f_N$	2, 4, 6, and 8 gray levels
6	$S(f_x,f_y) = a^2/((a^2 + f_x^2 + f_y^2)$, $a = 1.0f_N$	2, 4, 6, and 8 gray levels

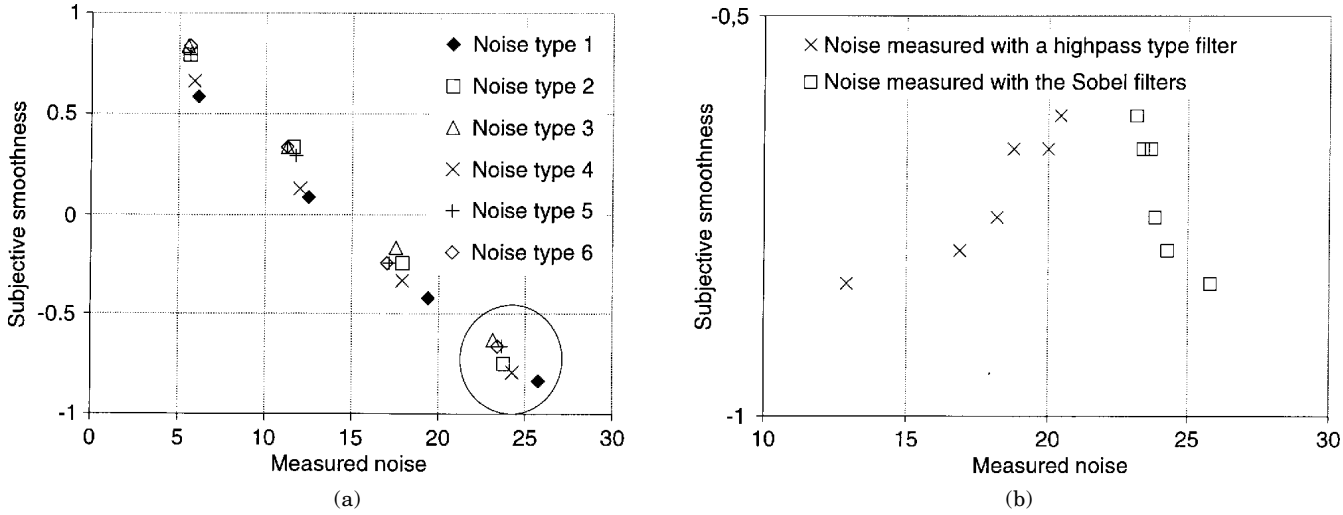


Figure 11. (a) Subjective smoothness versus Sobel filter based noise measurement for synthetic gray patches. The numbers refer to Table I. (b) Comparison of the Sobel filter method and the method of Ref. 24 applied to the encircled samples in part (a).

than the sharpness case. The visibility of noise is very dependent on tone reproduction and viewing conditions, and even if these were fixed, there may be large variation between individuals in the relative subjective importance of various noisy regions. In fact this is one of the reasons for objective quality measurement. Therefore we do not attempt to examine correlations between measured and subjective noisiness. Instead we present two relevant results concerning the noise measurement algorithm: first, a verification of the visual meaningfulness of the filters used and, second, a test of the ability to separate between noise and image structure.

To test the visual meaningfulness of noise measurement using the Sobel filters, uniform 8-bit gray patches of size 128×128 pixels and with different types of additive Gaussian noise were prepared. The mean gray value was 128. Six different types of noise power spectrum were used and four instances of each spectrum were created by varying the total power only. This makes a total of 24 samples, the parameters of which are given in Table I. Each image was processed according to the noise algorithm described above, with the exception that the exclusion of edge pixels was omitted as it would not work for completely edgeless images.

The samples were displayed on the same monitor used in the sharpness test, but with gamma 1.8. To reduce the influence of the monitor MTF, the images were magnified to size 512×512 and the viewing distance was correspondingly larger than normal, namely, 1.8 m. With this arrangement there were roughly 30 pixels of the original (unmagnified) image per one degree of the visual field. The magnification was done by simple pixel replication, which means that by magnifying the image we wished to approximate a fictitious "standard" monitor that displays a single pixel as a square of uniform intensity.

The subjective test was executed by comparing each of the 276 possible pairs on a three-alternative scale. Only one subject was used. Figure 11 shows the subjective results versus the measured noise. The samples are divided into four groups corresponding to constant total powers (this was not obvious a priori, though), so the monotonic relation between groups is not surprising. A more important result is the correlation within each group. The subjective noisiness increases as the noise bandwidth is decreased while keeping the total power constant, which is in accordance with earlier results.[9,10] The Sobel filter approximates this visual effect and, although not optimal, it is significantly better than a high-pass-type filter such as that in Ref. 24. Note that the subjective results are not on a metric scale, so the points are not expected to fall on a straight line. The subjective results, of course, are only valid for the particular angular resolution 30 pixel/degree. This resolution is, however, on a typical range because it corresponds, e.g., to digital printing at 600 dpi and with a dithering block of 6×6 pixels viewed at 45 cm.

To test the ability of the noise algorithm to detect noisy regions and discard textured regions, the algorithm was applied to 187 real images and the resulting noise levels were checked by visually inspecting both the original images and the smoothed gradient images used in the algorithm. None of these images were used during the development of the algorithm. The inspection was done on a monitor with commercial image editing software, which has a feature of displaying numerical pixel values so that the actual noise levels could be identified from the gradient image. The result of this test is necessarily rather imprecise because the test is tedious to conduct and because often no unique way exists to segment the image into uniform parts. The results in Table II indicate that

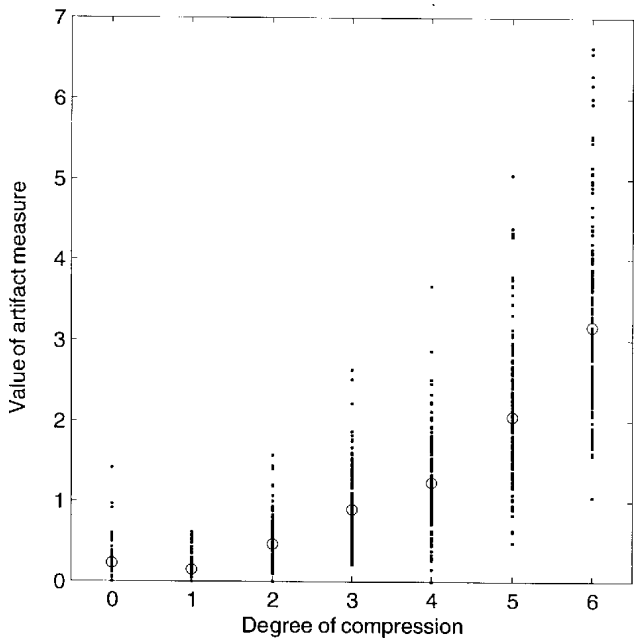


Figure 12. Scatter plot of the compression artifact measure versus degree of compression. The circles show the average of the measure at each compression level.

TABLE II. Performance of the Noise Algorithm for Real Images

Result of noise measurement	Number of images
(All images)	187
All noise regions recognized correctly	72
One noise region missing	32
Two noise regions missing	9
Three noise regions missing	1
One noise region partially interpreted as edges	8
Two noise regions partially interpreted as edges	2
Three noise regions partially interpreted as edges	1
One false noise region	49
Two false noise regions	17
Three false noise regions	6
Nonzero noise at zero gray level	5

while the algorithm works correctly in many cases, it should be considerably improved unless relatively simple images are only used as input. The missing and false noise levels are mainly due to decision errors in the 2-D histogram analysis, and the partial interpretation of noise regions as edges results from errors in the edge exclusion technique. The illogical results of nonzero noise with zero mean gray level are due to certain implementation details.

Coding Artifacts. The algorithm for blocking artifact detection was tested by compressing 170 real 24-bit RGB images with the baseline JPEG algorithm[30] using six grades of quality. (The uncompressed versions of these pictures had been utilized previously, together with some compressed but different pictures, to optimize the parameters of the algorithm.) Because the compression was done with commercial software, the quantization tables and other details of the compression were not known. It is known, however, that the compression was performed in a luminance–chrominance color space and in the three lowest grades the chrominance was undersampled by a factor of 2.

The blocking measure was computed for each version of each picture. Because we have not evaluated the subjective quality of the compressed images, the artifact measure is only compared against the degree of compression in Fig. 12. The average of the measured value increases with increasing compression, except at the weakest compression levels, but the variance at each compression level is rather high. Good correlation with subjective quality would, however, be the most important objective, and although this has not been tested systematically yet, it seems that among images with a given degree of compression, those with a low value of the artifact measure are usually either textured or have a high level of stochastic noise, while those with a high value contain large smooth areas. Because the artifacts are less visible in "busy" images than in smooth images, the artifact measure is expected to correlate somewhat better with subjective quality than with the technical degree of compression.

Discussion

The remaining problems in the sharpness and noise algorithms are somewhat different by nature. For the sharpness algorithm, potential improvements have mainly to do with better prediction of subjective quality, but not much can be done in this direction without exact knowledge of the picture reproduction. One shortcoming of the sharpness algorithm is that it measures the maximum sharpness without examining where the sharpest parts of the image are. If the main object in a picture is unintentionally less sharp than the background, the algorithm fails to predict subjective sharpness correctly. Because the object is usually in the middle of the picture, some kind of weighting of the local sharpness according to the spatial location could be an adequate solution.

In the noise algorithm, discrimination between noise and image structure seems to be the main problem. Interpretation of texture as noise usually increases the maximum measured noise level in an image, which is particularly disadvantageous from the visual quality viewpoint, because texture makes noise *less* visible. There are some conceivable improvements in the 2-D histogram analysis to reduce this problem; for example, the constraints on the relative positions of the histogram peaks could be more sophisticated. Despite this, solution of the problem seems to require estimation of local orientation, using, e.g., steerable filters[31] or a simpler method similar to that of Nyberg,[32] to prevent signal gradient from contributing to the filter output. In the noise measurement too there are perceptual effects to consider. The most pronounced of these is the masking by signal, another is that the noise of a small homogeneous region is more important in the absence than in the presence of another considerably larger region. Finally, color information should also be utilized to include chromatic noise in the measure, and color could also serve as an important cue in the decision between noise and texture.

The coding artifact measure is very simple; thus it is not expected to predict subjective quality very accurately. Nevertheless it makes possible a rough two-way classification; the decision boundary can be, e.g., about 1.0 according to Fig. 12. Local mean luminance is a very important factor for the visibility of coding noise as well as stochastic noise. Although the described simple method does not take the luminance into account in itself, the noise algorithm does measure the luminance of homogeneous regions, which can in principle be utilized to estimate the visibility of the blocking artifact as well.

Currently the sharpness and noise measurement are separate algorithms despite the utilization of some common intermediate results. Thresholding of gradient to extract edges is performed both in the sharpness and the noise algorithm using different methods. Unifying these into a single operation is worth consideration. Textured regions must be excluded both from sharpness and noise measurement; therefore the partition of the image into

edges, smooth regions, and textures, which is a commonly used technique in image processing, would be a convenient approach.

To be visually meaningful, a quality measure should take the intended pixel resolution and viewing distance into account. These may not be known at the time of quality measurement, which is one of the unavoidable problems. If the viewing parameters are available, then a potential way to incorporate them into the algorithm is to resample the images, preceded by appropriate low-pass filtering, to a "standard" angular resolution prior to application of the quality measurement. This resolution should be such that the corresponding Nyquist frequency is beyond the visual resolution limit. For images of very high resolution the resampling is also useful to reduce computation time.

No method with precisely the same objective as ours has been found in the literature, so a comparison with other methods cannot be made. Typically in the literature an algorithm has been tested on just a few selected images or the population of images has been implicitly very restricted. The results in this article must be viewed against the background that the test images have been taken from an archive of hundreds of pictures from various sources among which very many types of images are represented. Although the quality measurement problem is by no means completely solved, the approach of this article already covers various nonidealities of real pictures, such as nonuniform blur and nonwhiteness and luminance dependency of noise.

Conclusions

A computational procedure has been described to estimate quality potential components of natural color images with no reference image. The algorithm measures sharpness, stochastic noise, and the blocking artifact of JPEG compression. The aim has been to find simply computable and yet visually reasonable quality measures.

The sharpness measure is defined as the global maximum of local sharpness, where local sharpness is defined as maximum edge gradient normalized by the edge contrast. Under reasonable assumptions this measure is equivalent to the one-dimensional integral of the modulation transfer function. Experiments show that the algorithm predicts the spread parameter of simulated blur reasonably well and that the measure correlates with the subjective sharpness of different images.

The stochastic noise algorithm aims at identifying one or more noise levels in an image together with their mean luminances. The motivation for this is the potential dependence of noise on luminance on the one hand, and the influence of the mean luminance on perceived noisiness on the other. The method is based on filtering the image with Sobel gradient filters, which simultaneously suppress most of the image structure and roughly mimic the response of the human eye to noise of various spectra. Homogeneous regions in the image are identified as peaks in the bivariate histogram of smoothed gradient and smoothed luminance, and the location of such a peak indicates the amount of noise and the respective mean luminance. The discrimination between image details and noise is not entirely solved, however, and masking of noise and measurement of chromatic noise are still open questions.

The blocking artifact measure is a convenient supplement for the sharpness and noise algorithms. It describes a spikiness in the histogram of the gradient image characteristic of JPEG compressed images.

References

1. P. Laihanen et al., Automatic color correction, *The Second IS&T/SID Color Imaging Conference: Color Science, Systems, and Applications,* IS&T, Springfield, VA, 1994, pp. 97–101.
2. H. Saarelma, P. Laihanen, and S. Rouhiainen, Source and device independent color correction for printing, *IS&T's 11th International Congress on Advances in Non-Impact Printing Technologies,* IS&T, Springfield, VA, 1995, pp. 451–453.
3. B. Girod, The information theoretical significance of spatial and temporal masking in video signals, *Proc. SPIE* **1077**, 178–187 (1989).
4. C. Zetzsche and G. Hauske, Multiple channel model for the prediction of subjective image quality, *Proc. SPIE* **1077**, 209–216 (1989).
5. S. Daly, The visible differences predictor: an algorithm for the assessment of image fidelity, *Proc. SPIE* **1666**, 2–15 (1992).
6. C. J. van den Branden Lambrecht, A working spatio-temporal model of the human visual system for image restoration and quality assessment applications, *The 1996 International Conference on Acoustics, Speech, and Signal Processing (ICASSP '96),* Vol. **4**, IEEE, Piscataway, NJ, 1996, pp. 2291–2294.
7. A. P. Tzannes, J. M. Mooney, Measurement of the modulation transfer function of infrared cameras, *Opt. Eng.* **34**(6), 1809–1817 (1995).
8. V. Kayargadde and J. B. Martens, Estimation of edge parameters and image blur using polynomial transforms, *CVGIP: Graph. Mod. Image Process.* **56**(6), 442–461 (1994).
9. V. Kayargadde and J. B. Martens, An objective measure for perceived noise, *Sig. Process.,* **49**(3), 187–206 (1996).
10. V. Kayargadde and J. B. Martens, Perceptual characterization of images degraded by blur and noise: model, *J. Opt. Soc. Am.* ***A*** **13**(6), 1178–1188 (1996).
11. H. Kusaka, Consideration of vision and picture quality – psychological effects induced by picture sharpness, *Proc. SPIE* **1077**, 50–55 (1989).
12. P. G. J. Barten, Evaluation of subjective image quality with the square-root integral method, *J. Opt. Soc. Am. A* **7**(10), 2024–2031 (1990).
13. R. M. Rangayyan and S. G. Elkadiki, Algorithm for the computation of region-based image edge profile acutance, *J. Electron. Imaging* **4**(1), 62–70 (1995).
14. N. B. Nill and B. H. Bouzas, Objective image quality measure derived from digital image power spectra, *Opt. Eng.* **31**(4), 813–825 (1992).
15. D. M. Berfanger and N. George, Automatic image quality assessment, *IS&T's 47th Annual Conference/ICPS 1994,* Vol. **2**, IS&T, Springfield, VA, 1994, pp. 436–438.
16. A. Inoue and J. Tajima, Adaptive image sharpening method using edge sharpness, *IEICE Trans. Inf. Syst.* **E76-D**(10), 1174–1180 (1993).
17. S. Delgado Olabarriaga and R. M. Rangayyan, Subjective and objective evaluation of image sharpness: behavior of the region-based image edge profile acutance measure, *Proc. SPIE* **2712**, 154–162 (1995).
18. S. Bosch, I. Juvells, F. Abbad, and J. Campos, Use of the transfer function for evaluating imaging instruments with a nonrotationally symmetrical point spread function (P.S.F.): application to optical systems, *Proc. SPIE* **702**, 179–182 (1986).
19. J. S. Lourens, T. C. Du Toit, and J. B. Du Toit, Assessing the focus quality of television pictures, *Proc. SPIE* **1077**, 35–41 (1989).
20. R. Bracho and A. C. Sanderson, Segmentation of images based on intensity gradient information, *Proc. CVPR-85 Conf. Computer Vis. Patt. Recognition,* IEEE, Piscataway, NJ, 1985, pp. 341–347.
21. R. Brügelmann and W. Förstner, Noise estimation for color edge extraction, in *Robust Computer Vision,* W. Förstner & S. Winter, Eds., Wichmann, Karlsruhe, 1992, pp. 90–106.
22. J. S. Lee and K. Hoppel, Noise modeling and estimation of remotely sensed images, *Proc. 1989 Int. Geoscience and Remote Sensing Symp. (IGARSS '89),* 1989, Vol. 2, pp. 1005–1008.
23. P. Meer, J-M. Jolion, and A. Rosenfeld, A fast parallel algorithm for blind estimation of noise variance, *IEEE Trans. Patt. Analy. Mach.. Intell.* **12**(2), 216–223 (1990).
24. J. Immerkær, Fast noise variance estimation, *Computer Vision and Image Understanding* **64**(2), 300–302 (1996).
25. S. I. Olsen, Estimation of noise in images: an evaluation, *CVGIP: Graph. Mod. Image Process.* **56**(4), 319–323 (1993).
26. R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* Addison-Wesley, New York 1992, p. 199.
27. K. Kagitani, M. Hino, and S. Imakawa, Image noise evaluation method for color hardcopy, *IS&T's 12th International Conference on Digital Printing Technologies (NIP12),* IS&T, Springfield, VA, 1996, pp. 173–176.
28. M. Datcu, Method for quality evaluation of the block transform compressed images, *Proc. SPIE* **2418**, 118–126 (1995).
29. H. R. Wu, A new distortion measure for video coding blocking artifacts, *1996 Inter. Conf. Communication Technology Proceedings (ICCT '96),* IEEE, Piscataway, NJ, 1996, Vol. 2, pp. 658–661.
30. R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* Addison-Wesley, New York 1992, pp. 395–404.
31. W. T. Freeman and E. H. Adelson, The design and use of steerable filters, *IEEE Trans. Patt. Analy. Mach. Intell.* **13**(9), 891–906 (1991).
32. S. Nyberg, On image restoration and noise reduction with respect to subjective criteria, Report No. FOA-B-30045-E1, Försvarets forskningsanstalt, Stockholm, 1981.