OPTIMIZING IMAGE QUALITY USING TEST SIGNALS: TRADING OFF BLUR, NOISE AND CONTRAST

Bart Goossens, Hiêp Luong, Ljiljana Platiša, and Wilfried Philips

Ghent University, Dept. Telecommunication and Information Processing (TELIN-IPPI-BBT)
St. Pietersnieuwstraat 41, 9000 Ghent, Belgium

ABSTRACT

Objective image quality assessment (QA) is crucial in order to improve imaging systems and image processing techniques. In medical imaging, model observers that estimate signal detectability, have become widespread and promising as a means to avoid costly human observer experiments. However, signal detectability alone does not give the complete picture: one may also be interested in optimizing several independent quality factors (e.g., contrast, spatial resolution, noise).

In recent work, we have proposed the channelized joint observer (CJO), to jointly detect and estimate random parametric signals in images, a so-called signal-known-statistically (SKS) detection task. In this paper, we show how the estimation capabilities of the CJO can be exploited to estimate several image quality factors in degraded images, through signal insertion. By fixing the signal detectability, we illustrate how to benefit from the trade-offs that exist between the different quality factors. Our method is in the first place intended to aid medical image reconstruction techniques and medical display design, although the technique can also be useful in a much wider context.

1. INTRODUCTION

To improve medical imaging systems and processing techniques, objective assessment of image quality is an important factor. In medical image processing, the goal is not to create aesthetically pleasing images, but rather to be maximally useful for a specific purpose, e.g., the diagnosis of a disease. For this reason, task-based quality assessment is mostly adopted: first, the task of interest is specified, and next, it is quantitatively determined how well this task has been performed [1]. Most considered tasks are detection tasks, in which abnormalities in images (e.g., tumors, vein calcifications, lesions...) are being detected. Image quality can then be expressed objectively in terms of the detectability of abnormalities.

While signal detectability can be a good indicator of the image quality, in some circumstances, we are also interested in the reason why the image quality is determined to be good or bad. Even though the overall quality in terms of detectability may be the same, the individual quality factors may differ.

As an example, consider the design of a tomographic reconstruction method. During the past decades, iterative reconstruction methods (e.g., [2, 3, 4]) have found to give a superior reconstruction performance compared to the more traditional analytical methods, especially for low-dose or limited view images. Quite often, iterative methods depend on a set of parameters $\beta$ that need to be determined in order to optimize image quality. Since in many cases, the parameters are even content-dependent, and optimization of $\beta$ is a difficult design problem. Moreover, even though for two possible sets of parameter values, the signal detectability may be the same, it is theoretically speaking still possible to weight the noise level obtained after reconstruction, contrast and spatial resolution against each other. More concretely, consider a disk of diameter $d$ and contrast $c$ in a reconstructed image. The signal detectability of this disk can be expressed by the signal-to-noise-ratio (SNR) [5]:

$$\text{SNR} = \frac{c}{\sigma_n},$$

where $\sigma_n$ is the standard deviation of the noise in the reconstructed image, averaged over a circular region of diameter $d$. Now, when both the contrast and noise standard deviation are multiplied with the same factor, the signal detectability in terms of SNR remains the same. Note that the direct measurement of $c$ and $\sigma_n$ from an image is not as trivial as it may seem: first the background noise may be correlated or non-stationary, which makes the estimation very difficult [6]. Second, we are mostly interested in measuring the contrast of a signal inserted in a realistic background (rather than a uniform background). In this case, the background should be taken into account when measuring $c$. Another deficiency of (1) is that spatial resolution effects (e.g., blurring, aliasing) are not directly taken into account.

As a general means to estimate the detectability SNR, while alleviating some of the previous problems, numerical observer models have been proposed [1, 7]. Some examples are the (non)pre-whitening matched filter, the Bayesian ideal observer and the (channelized) Hotelling observer. Let $b$ denote a vector of intensities of the background, and let $x$ and $y$ respectively denote the known signal and the reconstructed image. Then, the detection problem is formulated as follows:

$$y = \begin{cases} b & (H_0) \\ b + x & (H_1) \end{cases}$$

The observer model decides whether a signal is present in the considered image ($H_1$), or not ($H_0$). This is done by comparing the decision test statistic $t$ to a given predefined threshold. For Gaussian distributed backgrounds $b$ with mean $0$ and covariance $C_b$, the ideal Bayesian observer has the following test statistic:

$$t = x^T C_b^{-1} y.$$  \hspace{1cm} (3)

From (3), the detectability SNR is computed as [8]:

$$\text{SNR} = \frac{\text{E}[t|H_1] - \text{E}[t|H_0]}{\sqrt{\frac{1}{2} \left( \text{Var}[t|H_1] + \text{Var}[t|H_0] \right)}} \approx \left( x^T C_b^{-1} x \right)^{\frac{1}{2}}.$$  \hspace{1cm} (4)

Bart Goossens is a postdoctoral fellow of the Flemish Research Foundation (FWO) Flanders, Belgium.
Note that the computation of the ideal observer SNR requires that both the signal \( x \) and the background statistics \( C_b \) are exactly known. In the image quality assessment application in the above example, this is not the case, because the signal \( x \) in (4) corresponds to a (possibly non-linearly) tomographically reconstructed signal. The channelized Hotelling observer (CHO) [7] circumvents this problem by estimating \( x \) and \( C_b \) directly from the reconstructed images. To do so, a dimensionality reduction is applied by projecting onto a set of channels. Although the CHO may an accurate estimate of the ideal Bayesian observer SNR [7] and serves as a no-reference based QA method, the CHO is unable to provide estimates of individual image quality factors.

Recently, we have extended the CHO model to signal-known-statistically (SKS) detection tasks [9, 10]. The resulting jointed observer (JCO), is able to jointly estimate signal parameters and detect its presence. In this paper, we build a QA system mostly based on ideas forthcoming from the CJO model, namely that the estimation capabilities of the CJO can be used to measure the signal degradation. Therefore, we assume that the degraded signal has randomly distributed parameters, which are to be estimated. In particular, we consider the measurement of noise, contrast and blur through signal insertion the image (possibly as an off-line procedure). In contrast to [9, 10], to maximize the estimation accuracy, we do not estimate the quality measures on an image basis, but on an overall basis, over an ensemble of images. We point out that the signal detectability in terms of SNR (4) is a function of these quality factors. This means that, when fixing the signal detectability, a trade-off is possible between the independent quality factors. This may be useful to tune display and visualization methods to the users’ needs and comfort.

This paper is structured as follows: In Section 2, the proposed QA method is explained. The method consists from a background and signal model (Subsection 2.1). The trade-off between the individual quality factors is discussed in Subsection 2.2 and the estimation in Subsection 2.3. The computation of the signal detectability is discussed in Subsection 2.4. Finally, experimental results are given in Section 3 and Section 4 concludes this paper.

2. THE PROPOSED QA METHOD

An overview of our QA technique is shown in Figure 1. We assume that an idealized background and signal (which we do not have at our disposal), is corrupted by a degradation process and consequently restored by a reconstruction process, resulting in a “measured” background and signal. This technique is similar to the one recommended in [13], with the difference that our aim is not to estimate the PSNR, but rather individual quality factors. A few examples of degradation processes and the corresponding reconstruction processes are listed in Table 1. Some of the parameters of both processes (denoted by \( \beta \)) can be controlled. The goal is then to jointly compute the detectability in the SNR sense, the blur level \( \sigma_b \), noise level \( \sigma_n \), and contrast \( c \). These quality measures may then be used in turn to update the parameters \( \beta \) (note that this last topic heavily depends on the specific degradation and reconstruction processes and is therefore outside the scope of this paper). To measure the mentioned image quality factors, we first define a signal and background model, and then explain how these factors can be measured from an image, in a reference-free manner.

2.1. Signal and background model

First, we will explain the signal and background model used for QA. Consider an “ideal” (reference) background image \( b \), and an “ideal” (reference) signal \( x \). For notational simplicity, the images are stored into a vector, using column stacking. Since our QA method is reference-free, these images will not be available to our method. Now, consider a symmetric Gaussian signal:

\[
\tilde{x} = \exp\left(-r^2 / 2\sigma_b^2\right),
\]

where \( \sigma_b \) is a known scale parameter of the signal, \( r \) is the distance to the center of the signal \( q \) (which is a known location in the image). We also assume that the relationship between the background and the signal is additive. Then, under the hypotheses, the following “ideal” image is synthesized:

\[
\hat{y} = \begin{bmatrix} \tilde{b} \\ \tilde{b} + \tilde{x} \end{bmatrix} (H_0) \quad \begin{bmatrix} \tilde{b} \\ \tilde{b} + \tilde{x} \end{bmatrix} (H_1)
\]

Next, the “ideal” image is corrupted by a process \( f_\beta (\cdot) \), resulting in the observed image \( y \):

\[
y = \begin{cases} f_\beta (\tilde{b}) + n & (H_0) \\ f_\beta (\tilde{b} + \tilde{x}) + n & (H_1) \end{cases}
\]

where \( f_\beta (\cdot) \) is an unknown general nonlinear non-invertible vector function, and \( n \) represents statistical noise introduced during the degradation process (here, we assume that \( n \) is Gaussian distributed with mean 0 and variance \( \sigma_n^2 \)). In the remainder of this paper, we will assume that neither \( b \) or \( \tilde{y} \) are available. Remark that \( \hat{b} \) or \( \hat{y} \) may be available when the process \( f_\beta (\cdot) \) is simulated (e.g., involving Monte-Carlo simulations). This corresponds to full-reference based quality assessment, which is treated more extensively in [14, 15].

To simplify our following analysis, we will perform a first order Taylor series approximation of \( f_\beta (\tilde{b} + \tilde{x}) \) at \( \tilde{x} = 0 \). This gives the following additive relationship between the observed background and signal:

\[
y = \begin{cases} f_\beta (\tilde{b}) + n & (H_0) \\ f_\beta (\tilde{b}) + T\tilde{x} + n & (H_1) \end{cases}
\]

where \( T = \left[Df_\beta (\tilde{b})\right]^T \). Consequently, we arrive at the detection problem from (2), when setting:

\[
b = f_\beta (\tilde{b}) + n \quad \text{and} \quad x = T\tilde{x}.
\]
This equation, which is due to the linearization, has the intuitive interpretation that the background and signal are processed independently, and then added together to yield $y$. Next, we put forward a simple model for the operation $T$ (see Figure 2): we assume that $T$ is composed of an intensity scaling (corresponding to a signal contrast adjustment), and a Gaussian blurring operation.

For the signal from (5) and according to (9), the measured signal $x$ can be calculated analytically:

$$|x|_c = c \cdot \exp\left(-r^2/2\sigma_b^2\right), \text{ with } \sigma_x^2 = \sigma_b^2 + \sigma_c^2.$$  (10)

Now, we are interested in estimating the signal parameters of $x$, directly from the observation of $y$. These estimated signal parameters then indicate how the contrast, spatial resolution are affected by the operation $f_B$ ($\cdot$).

### 2.2. The trade-off between noise, contrast and blur

For the analytical signal (10) and using a simple correlated Gaussian background model, the SNR (4) can easily be calculated. To show the individual relationships between the different signal parameters, we performed an experiment in which the SNR for various parameter signal values was computed, for a correlated Gaussian background. The results are shown in Figure 3. In particular, contours of equal SNR are shown function of the blur, noise and contrast levels. It can be seen that, when increasing the blur level, a higher contrast is needed to preserve the SNR. This effect diminishes however when the SNR reaches 1 (0 dB). Also, the relationship between contrast and blur is nonlinear, the SNR drops rapidly when increasing the blur level.

On the other hand, the relationship between noise and contrast is almost linear. This is not a surprise, due to (1), although the contours do not intersect in the origin (as would be expected in the light of (1)). This is because the background (see Subsection 2.1) is explicitly taken into account here. Hence, when fixing the SNR, the “optimal” contrast does not only depend on the noise or blur, but also on the background. In Section 3, we will perform a more extensive experiment, in which we optimize the parameters of an algorithm according to the above trade-offs.

### 2.3. Local image quality measures

In the following, we will assume that both signal-present (denoted by $y|H_1$) and signal-absent images ($y|H_0$) are given, and that the signal presence is known (hence skipping the detection task). Under these circumstances, the observed degraded signal can be estimated simply by:

$$\hat{x} = \langle y|H_1 \rangle - \langle y|H_0 \rangle,$$  (11)

where $\langle \cdot \rangle$ denotes the sample mean. Next, the signal parameters $(c, \sigma_x)$ can be jointly estimated by least-squares fitting of (11) to (10):

$$c, \sigma_x = \arg \min_{(c,\sigma_x)} \|x(c,\sigma_x) - \hat{x}\|^2$$

$$= \arg \max_{(c,\sigma_x)} \left(\langle x(c,\sigma_x) \rangle - \frac{1}{2} \hat{x}\right).$$  (12)

Note that the signal $x(c,\sigma_x)$ depends on $(c,\sigma_x)$ in a nonlinear way. To facilitate the estimation, we first make use of a linear dimension reduction (similar to the CJO in [9, 10]). The original and estimated signals are linearly projected onto a small set of channels:

$$x'(c,\sigma_x) = U^T x(c,\sigma_x) \text{ and } \hat{x}' = U^T \hat{x},$$  (13)

where $U$ is an $N \times K$ projection matrix containing $K$ column-stacked channels. The number of channels $K$ is typically very small, e.g. $K = 5$. Here, $N$ is the length of the vector $x$ (i.e., the number of pixels in the image). In channel space, (12) becomes:

$$c, \sigma_x = \arg \max_{(c,\sigma_x)} x'(c,\sigma_x) \left[\hat{x}' - \frac{1}{2} x'(c,\sigma_x) \right].$$  (14)

The main benefit of this approach is that both the detection and estimation take place in channel space, which is computationally very efficient. Here, our goal is to estimate both the size (scale) of the
signal and its contrast. The channels should be designed carefully such that the solution of (14) approximates the solution to (12) as much as possible. To aid the signal shape parameter estimation, we use channels that are shiftable in scale \([10]\). These channels can be composed in polar frequency coordinates (where the origin of the polar grid is again at the center of the signal):

\[
f(\sigma)(\omega, \varphi) = \text{sinc}(\text{sign}(\omega) \log_2 |2^T \omega|) |\omega > 0|, \quad \sigma = 0, \ldots, K - 1
\]

with \(\text{sinc}(x) = \sin(\pi x) / (\pi x)\) the sinc function and with \(1 \cdot \) the indicator function. These channels are illustrated in Figure 4. Fixing a reference signal as \(x_0 = x(1, 1)\), it can be shown that, in the scale-shiftable channel domain, any signal \(x(c, \sigma_x)\) can be computed by applying a linear transform \(A_{c, \sigma_x}\) to the reference signal [10]:

\[
x'(c, \sigma_x) = A_{c, \sigma_x} x_0,
\]

where elements \(m, n\) of \(A_{c, \sigma_x}\) are given by:

\[
[A_{c, \sigma_x}]_{mn} = \frac{\sigma_{c}}{c} 2^{-(m-n)} \text{sinc}(\log_2 \sigma_x - (m-n)).
\]

This gives the following joint optimization problem to solve:

\[
\begin{align*}
\hat{c}, \hat{\sigma_x} & = \arg \max_{c, \sigma_x} x_0^T A_{c, \sigma_x}^T \left[ x' - \frac{1}{2} A_{c, \sigma_x} x_0 \right], \quad (18) \\
\hat{\sigma_b} & = \sqrt{\text{max} (0, \sigma_x^2 - \sigma_b^2)}, \quad (19)
\end{align*}
\]

where the maximum is used to account for possible negative numbers due to estimation errors. Because the objective function in (19) is differentiable in \(\sigma_b\), the maximum can easily be found using Gauss-Newton optimization techniques. Once the signal shape parameter is estimated, the contrast of the signal can be determined:

\[
\hat{\sigma} = \frac{x_0^T A_{\hat{c}, \hat{\sigma_x}}^T \hat{x} - \hat{x}' \hat{x}^T}{x_0^T A_{\hat{c}, \hat{\sigma_x}} x_0},
\]

which is, given \(\hat{\sigma_x}\), a linear function of \(\hat{x}'\). Now we turn to the estimation of the noise level \(\sigma_n\). According to (8), we have that:

\[
\begin{align*}
\text{Var}[y|H_0] = \text{Var}[y|H_1] = \text{Var}[f_\beta(\hat{b})] + \sigma_n^2 I. \\
\end{align*}
\]

In case the reference background variance \(C_b\) is \(\text{Var}[f_\beta(\hat{b})]\) can be estimated, we find the following noise level estimate:

\[
\hat{\sigma}_n^2 = \max \left(0, \frac{1}{N} \text{trace} \left( \frac{1}{2} \left( \text{Var}[y|H_0] + \text{Var}[y|H_1] \right) - C_b \right) \right)
\]

24. Estimating the signal detectability

After estimating the signal parameters, we can compute the observed signal as \(x(\hat{c}, \hat{\sigma}_x)\). This signal can then be fed into the model observers, for example the simple matched filter [1], which has the test statistic:

\[
t = [x(\hat{c}, \hat{\sigma}_x)]^T y.
\]

Alternatively, the CHO can be used, with test statistic:

\[
t = [x'(\hat{c}, \hat{\sigma}_x)]^T \left( U^T (C_b + \sigma_n^2 I) U \right)^{-1} U^T y,
\]

where \(C_b + \sigma_n^2 I\) is then estimated in the training phase of the model[7]. Either test statistic (23) or (24) can then be used in order to estimate the area under the ROC (AUC) as a measure of detectability, e.g., using the Wilcoxon-AUC test [16]. Giving a set of \(M\) signal-present images (with test statistics \(t_{1}\)), and a set of \(M\) signal-absent images (with test statistics \(t_{0}\)), the AUC is estimated as:

\[
\tilde{\text{AUC}} = \frac{1}{M^2} \sum_{i=1}^{M} \sum_{j=1}^{M} I[t_{ij} \geq t_{0}].
\]

In the following, we will use the matched filter (23) to calculate the AUC, for simplicity.

3. EXPERIMENTAL RESULTS

As a first experiment, we investigate the accuracy of the estimated image quality factors. Therefore, we extract 2000 patches of size \(100 \times 100\) from the chest radiography database of [17], and 1000 signal-free patches, and 1000 patches with a signal inserted in the middle. The images are then artificially degraded using a Gaussian blur with parameter \(\sigma_b\), and next corrupted with noise with variance \(\sigma_n^2\). The observation model is then simply (7), with \(f_\beta(\hat{b})\) a Gaussian blurring operation. As an illustration, 6 images from each set are shown in Figure 5. Next, we apply (20), (21) and (25) to jointly estimate the blur, contrast and AUC of the degraded images. The results are shown in Figure 6. It can be noted that the presence of the background has limited influence on the estimation results, and that the estimated values are quite accurate.

For the second experiment, our goal is to optimize a simple denoising post-processing algorithm, in order to 1) investigate whether denoising affects the signal detectability and 2) determine the influence of the denoising parameters on the individual image quality factors. First, artificial white Gaussian noise with "initial" noise standard deviation \(\sigma_{n,0} = 20\) is added to an MRI image of the human brain (see Figure 7). Note that in practice, noise in MRI magnitude images is Rician rather than Gaussian, but here we use Gaussian noise 1) to keep the setup of the study simple (so that the
results are not affected by the presence of signal-dependent noise), and 2) noise in high SNR regions is generally approximated well using a Gaussian distribution. Then, a Gaussian signal with amplitude $a_0 = 120$ and scale $a_1 = 1$ is inserted at various positions in the image. Next, noise is suppressed from the patches using shearlet-domain [18] hard-thresholding:

$$ f_{\beta, \lambda} (\tilde{y}) = (1 - \beta) \tilde{y} + \beta S^T \left[ \text{hardthreshold} \left( S \tilde{y}, \lambda \right) \right], \quad (26) $$

where $\lambda$ is a fixed threshold to be optimized, $\beta$ controls the amount of filtering ($\beta = 0$ corresponds to no filtering, $\beta = 1$ to “maximal” filtering). By setting the parameter $\beta < 1$ it is possible to not suppress all of the noise. This can be useful to avoid the over-smoothing when the threshold parameter $\lambda$ is chosen too large. The function hardthreshold ($\cdot; \lambda$) applies the hard threshold $\lambda$ to the shearlet coefficients, i.e., coefficients with magnitude smaller than $\lambda$ are set to zero. We opt for hard instead soft thresholding, to preserve sharpness as much as possible. Using the proposed technique, we can now assess the quality of the denoised image, in terms of noise, contrast, blur and signal detectability, as a function of the parameters $\lambda$ and $\beta$. This process is repeated 10 times with different random noise. Subsequently, 200 patches of size $64 \times 64$ are extracted out of the resulting images, of which 100 are signal-free and 100 are signal-present. The results are shown in Figure 9. First, post-processing by denoising affects the signal detectability, but fortunately not significantly: the AUC drops from 0.9 ($\beta = 1, \lambda = 1$) to 0.88 for ($\beta = 2, \lambda = 2$). By looking at Figure 9(b)-(d), we can attribute the detection performance decrease to the introduced smoothing (blur) and reduced signal contrast. We also note that denoising greatly decreases the amount of noise in the image, especially for $\beta = 1, \lambda > 1.8$. Moreover, when plotting the AUC as a function of noise, blur and contrast, it can be seen that the surfaces of equal AUC are highly nonlinear and form a manifold in the 3D space (see Figure 10). Using these findings, we can easily trade-off the various quality factors, e.g., using:

$$ (\tilde{\beta}, \tilde{\lambda}) = \arg \max_{(\beta, \lambda)} w_n \left( 1 - \frac{\hat{\sigma}_b}{\sigma_{b, 0}} \right) + w_c \hat{c} + w_b \tilde{\sigma}_b + w_a \tilde{\text{AUC}}, \quad (27) $$

where $w_n, w_c, w_b, w_a$ are weights that determine the importance of respectively the noise, contrast, blur and AUC detection performance in the overall image quality, and where $\sigma_{b, 0}$ is the reference blur level. For example, if we choose $w_n = w_c = w_b = w_a$ and reference blur level $\sigma_{b, 0} = 1$, we find $\beta = 0.54, \lambda = 1.84$. In this case, a “partially” denoised image offers the best image quality (according to the selected weights). Some visual illustrations are provided in Figure 8.

4. CONCLUSION

In this paper, we have presented an objective image quality approach to jointly measure image quality factors (blur, contrast, noise and signal detectability), based on inserting test-signals in the image. We have demonstrated that various trade-offs exist among the quality factors, when one of them is fixed. Our approach can be used directly to optimize image reconstruction techniques in a general way. When applied to a very simple denoising algorithm, we find that denoising
itself does not necessarily improve the signal detectability, this is mainly due to blur and contrast loss. In future work, we will investigate if this technique can be used to optimize quality in real-time (assuming the inserted signals are weak and invisible to the human eye). Furthermore, it becomes possible to tune existing image reconstruction techniques according to human preferences, while maintaining signal detectability.

5. REFERENCES


