# Search Strategies for Image Multi-Distortion Estimation

André-Louis Caron[1]          Pierre-Marc Jodoin[1]          Christophe Charrier[1,2]

[1]MOIVRE
Université de Sherbrooke
2500 boulevard de l'Université
Sherbrooke, J1K 2R1, Canada

[2]GREYC, UMR CNRS 6072, Image team
Université de Caen Basse-Normandie
6 boulevard Maréchal Juin
14000 Caen, France

## Abstract

*In this paper, we present a method for estimating the amount of Gaussian noise and Gaussian blur in a distorted image. Our method is based on the MS-SSIM framework which, although designed to measure image quality, is used to estimate the amount of blur and noise in a degraded image given a reference image. Various search strategies such as Newton, Simplex, and brute force search are presented and rigorously compared. Based on quantitative results, we show that the amount of blur and noise in a distorted image can be recovered with an accuracy up to $0.95\%$ and $5.40\%$, respectively. To our knowledge, such precision has never been achieved before.*

## 1 Introduction

In the last decade, various quality metrics have been designed to measure the visual distance between two images. Examples of such metrics are Multi-Scale Structural SIMilarity (MS-SSIM) [11], Wavelet Structural Similarity (WSSI) [7], Visual Signal-to-Noise Ratio (VSNR) [1], and Visual Information Fidelity (VIF) [8] indices, to name a few. These metrics are designed to give a score between a reference image $f$ and a distorted version of it $g$. The aim of these metrics is to provide a score similar to that given by an average human observer. From that score, one can conclude that the quality of an image is from excellent to very bad.

Unfortunately, those quality metrics share a common limitation as they cannot identify the kind (nor the amount) of degradation that has been applied on $f$ to obtain $g$. Furthermore, given a visual score, those methods cannot determine if one or more degradations has been applied to the reference image. Although methods exist to estimate the amount of a single distortion [9], none is capable of estimating simultaneously the kind *and* the amount of multiple distortions.

In this paper, we introduce a new method to estimate the kind and the amount of distortion in an image $g$ given a reference image $f$. We focus on the combination of Gaussian noise ($\sigma$) and Gaussian blur ($\eta$), two distortions often considered by denoising and deconvolution techniques [2]. Our method is based on the MS-SSIM factors which account for luminance distortion ($ld$), contrast distortion ($cd$) and structure distor-

tion ($sd$) between $f$ and $g$. In the upcoming sections, we will show that there is a bijective relation between a pair ($\sigma, \eta$) and a 3D point ($ld, cd, sd$) in the MS-SSIM space. Since the transfer function provided by MS-SSIM is expensive computational wise, we propose a transfer function based on a bi-cubic patch which maps any 2D point ($\sigma, \eta$) to its 3D counterpart in the MS-SSIM space with only 4 matrix multiplications. The problem of recovering ($\sigma, \eta$) from ($f, g$) will then be formalized as a search problem for which four solutions will be proposed.

## 2. The MS-SSIM factors

As mentioned previously, the MS-SSIM index [11] is based on three factors: 1) the luminance distortion ($ld$) 2) the contrast distortion ($cd$) and 3) the structure distortion ($sd$) between an image $f$ and a degraded version of it $g$. The philosophy behind MS-SSIM lies in its representation of an image: an $N \times M$ image is a point in the $\mathbb{R}^{MxN}$ image space where any distortion is modeled by a translational vector added to a reference image (the length of the vector is proportional to the magnitude of the distortion). The two vectors responsible for luminance and contrast distortion span a plane on which lies the reference image. The authors mention that distortions corresponding to a rotation of that plane are associated to structural changes between $f$ and $g$.

The luminance distortion is defined as

$$ld(f,g) = \frac{2\mu_f\mu_g + C_1}{\mu_f^2 + \mu_g^2 + C_1}$$

where $\mu_f$ and $\mu_g$ represent the mean intensity of $f$ and $g$, and $C_1$ is a constant to avoid instability when $\mu_f^2 + \mu_g^2 \approx 0$. According to Weber's law [10], the magnitude of a just-noticeable luminance change $\delta L$ is proportional to the background luminance $L$. In that case, $\mu_f = \alpha\mu_g$, where $\alpha$ represents the ratio of the luminance of $g$ versus $f$. Thus, the luminance distortion can also be defined as

$$ld(f,g) = \frac{2\alpha\mu_f^2 + C_1}{(1+\alpha^2)\mu_f^2 + C_1}. \tag{1}$$

Contrast distortion is defined in a similar way *i.e.*:

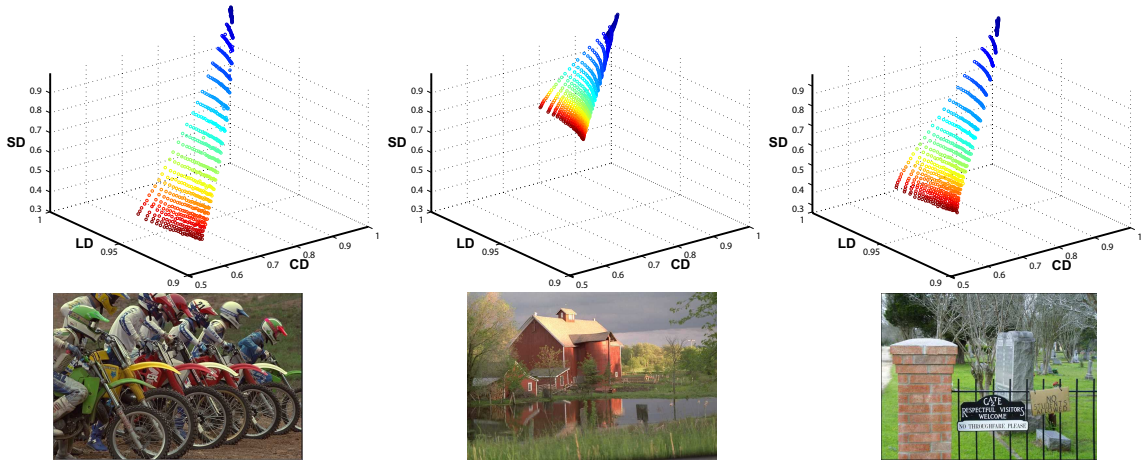$$cd(f,g) = \frac{2\sigma_f\sigma_g + C_2}{\sigma_f^2 + \sigma_g^2 + C_2} \tag{2}$$

**Figure 1.** Three MS-SSIM manifolds obtained from three images. Each point in the (LD,CD,SD) space corresponds to a specific $(\sigma, \eta)$ distortion.

where $C_2$ is a non negative constant and $\sigma_f$ (resp. $\sigma_g$) represents the standard deviation of $f$ (and $g$).

As for structure distortion, it is measured after subtracting the average luminance and normalizing the contrast of both $f$ and $g$. This leads to :

$$sd(f,g) = \frac{2\sigma_{f,g} + C_3}{\sigma_f^2 \sigma_g^2 + C_3} \qquad (3)$$

where $\sigma_{f,g} = \frac{1}{N-1}\sum_{i=1}^{N}(f_i - \mu_f)(g_i - \mu_g)$, and $C_3$ is a small constant. Note that $sd(f,g)$ takes negative values whenever the local image structure is inverted.

According to Eq. (1), (2), and (3), a pair $(f,g)$ (where $g$ is a degraded version of $f$) is associated to a 3D point $(ld, cd, sd) \in [(0,0,0),(1,1,1)]$. A simple but yet fundamental observation can be made here : the distance between $f$ and $g$ has a geometric meaning which states that the closer $(ld, cd, sd)$ is to $(1,1,1)$, the more similar $f$ and $g$ will be.

## 3. Noise and Blur Distortions

A degradation process frequently accounted for is the one associated to most digital cameras [2]. According to this process, the lens of the camera induces blur and the digitizer adds random noise. It is well known that if the degradation process is linear and position invariant, and that the 3D scene is made of objects located roughly at the same distance from the camera, then the degraded image $g$ is obtained as follows

$$g = h * f + \mathcal{N} \qquad (4)$$

where $h$ is a low-pass filter, $\mathcal{N}$ is white noise and $*$ indicates convolution. In this paper, $h$ is a zero-mean Gaussian filter and $\mathcal{N}$ is zero-mean Gaussian white noise. According to this model, distortion depends on two standard deviation values namely $(\sigma, \eta)$.

By combining Eq. (1), (2), (3), and (4), one can associate a distortion $(\sigma, \eta)$ to a 3D MS-SSIM point as

follows

$$\begin{aligned}(\sigma, \eta) \quad \rightarrow \quad & (ld(f, h_\sigma * f + \mathcal{N}_\eta), cd(f, h_\sigma * f + \mathcal{N}_\eta), \\ & sd(f, h_\sigma * f + \mathcal{N}_\eta). \qquad (5)\end{aligned}$$

Interestingly, as shown in Fig. 1, the 3D points associated to $(\sigma, \eta)$ form a smooth manifold in the 3D MS-SSIM space. The manifolds in Fig. 1 were obtained with 676 distortions ranging between $(0,0)$ and $(\sigma_{\text{MAX}}, \eta_{\text{MAX}})$. Out of these plots, we empirically observed that

1. there is a one-to-one mapping between each $(\sigma, \eta)$ pair and its 3D correspondence $(ld, cd, sd)$;

2. the manifolds have two principal directions corresponding to variations of $\sigma$ and $\eta$.

## 4. 2D to 3D Mapping

Since the position on the MS-SSIM manifold is determined by the magnitude of $\sigma$ and $\eta$, they can be seen as parameters allowing to navigate on the 3D manifold. This is very close to the definition of a 3D parametric patch which relates two parameters $s \in [0,1]$ and $t \in [0,1]$ to 3D points $(x, y, z)$[5] :

$$(x(s,t) \; y(s,t) \; z(s,t))^{\mathrm{T}} = S.M.G.M^{\mathrm{T}}.T^{\mathrm{T}} \qquad (6)$$

where, in the case of a bi-cubic patch, $S = (1 \; s \; s^2 \; s^3)$, $T = (1 \; t \; t^2 \; t^3)$, $G$ is a $4 \times 4$ matrix containing 16 control points, and $M$ is the $4 \times 4$ *basis matrix* defining the nature of the patch (a Bezier, an Hermite, a Catmull-Rom or any other patch). The manifolds in the $(LD, CD, SD)$ space being smooth, one can fit a patch on it with a Vandermonde matrix[1] and with 16 control points obtained after uniformly sampling the $(\sigma, \eta)$

---

[1] $\quad M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1/3 & 1/9 & 1/27 \\ 1 & 2/3 & 4/9 & 8/27 \\ 1 & 1 & 1 & 1 \end{pmatrix}^{-1}.$
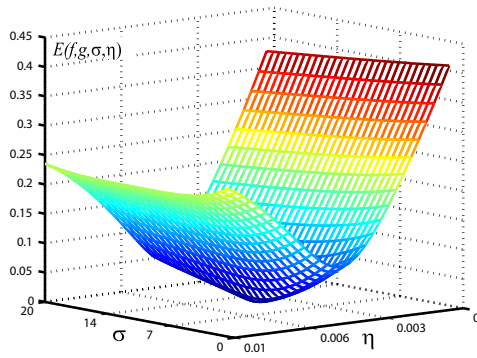
**Figure 2.** Error function $E$ for the *Bikes* image in Fig. 1. The minimum is at $(\frac{\sigma_{\text{MAX}}}{2}, \frac{\eta_{\text{MAX}}}{2})$.

space with a $4 \times 4$ lattice. This allows to map a distortion pair $(\sigma, \eta)$ to its associated $(ld, cd, sd)$ point with very little computational effort.

## 5. Distortion Estimation

As mentioned previously, our goal is to estimate the amount of blur $\sigma^*$ and noise $\eta^*$ there is in image $g$ given image $f$. Knowing that the pair $(f, g = f * h_{\sigma^*} + \mathcal{N}_{\eta^*})$ corresponds to a unique 3D point $(lc^*, cd^*, sd^*)$ in the MS-SSIM space, the goal is to find a pair $(\hat{\sigma}, \hat{\eta})$ such that $(f, f * h_{\hat{\sigma}} + \mathcal{N}_{\hat{\eta}})$ corresponds to a 3D point $(\hat{lc}, \hat{cd}, \hat{sd})$ close to $(lc^*, cd^*, sd^*)$. The estimation of $(\sigma^*, \eta^*)$ thus becomes a search problem that we formalize as follows

$$(\hat{\sigma}, \hat{\eta}) = \underset{\substack{\sigma \in [0, \sigma_{\text{MAX}}] \\ \eta \in [0, \eta_{\text{MAX}}]}}{\arg\min} \ E(f, g, \sigma, \eta)$$

where $E(.) = \sqrt{(lc - lc^*)^2 + (cd - cd^*)^2 + (sd - sd^*)^2}$. In order to illustrate the shape of this error function, we computed its value over 10000 samples given that the solution $(\sigma^*, \eta^*)$ is $(\frac{\sigma_{\text{MAX}}}{2}, \frac{\eta_{\text{MAX}}}{2})$. As shown in Fig. 2, the error function is smooth and has only one minimum which is global. In the upcoming subsections, we will introduce four strategies to find that minimum.

### 5.1 Brute Force Search (BFS)

The simplest way to find $(\hat{\sigma}, \hat{\eta})$ given $(f, g)$ is by considering a large number of $(\sigma, \eta)$ values and keep the one whose 3D point $(lc, cd, sd)$ is the closest to $(lc^*, cd^*, sd^*)$ (*i.e.* the one with the lowest error $E(.)$). In this paper, we uniformly sample the $(\sigma, \eta)$ space with a $26 \times 26$ lattice ranging from $(0, 0)$ to $(\sigma_{\text{MAX}}, \eta_{\text{MAX}})$.

### 5.2 Simplex Search (SS)

As one would expect, considering a large number of $(\sigma, \eta)$ values (here 676) is prohibitive computational wise. The reason being that computing $(lc, cd, sd)$ with $f * h_{\sigma} + \mathcal{N}_{\eta}$ followed by Eq.(1), (2), and (3) is a time consuming procedure. To reduce the computational effort of the search, one needs to visit a smaller number of

$(\sigma, \eta)$ values before to reach the global minimum. Unfortunately, since the 3D shape of $E(f, g, \sigma, \eta)$ is unknown *a priori*, gradient descent strategies are not applicable here. In fact, such problem is an unconstrained optimization problem without derivatives for which the simplex search is a typical solution [4]. Simplex starts with 3 solutions called *vertices* which form a triangle in the $(\sigma, \eta)$ space. New positions are then iteratively identified inside and outside the triangle. The error value at these new positions is compared with the error values at the vertices of the triangle and, usually, one of the vertices is replaced by the new point, leading to a new triangle. This is repeated until the diameter of the triangle is below a specified tolerance or when the number of points visited has reached a pre-defined threshold. In this paper, the initial triangle is centered at the middle of the $(\sigma, \eta)$ space, *i.e.* at position $(\frac{\sigma_{\text{MAX}}}{2}, \frac{\eta_{\text{MAX}}}{2})$ and the algorithm stops when 50 points have been visited.

### 5.3 Simplex Search with Patch (SBS)

Although the simplex search is drastically faster than the brute force search, its processing time can be further reduced. This can be done by using the bicubic patch introduced in section 4 to map any 2D point $(\sigma, \eta)$ to its related 3D position $(lc, cd, sd)$. Such a patch needs 16 3D control points that can only be obtained after computing $g = f * h_{\sigma} + \mathcal{N}_{\eta}$ followed by Eq.(1), (2), and (3). However, once these control points have been computed, the 2D-to-3D mapping requires very little computational effort (only 4 matrix multiplications) and thus speeds up the search procedure. Here the simplex search has been modified to account for this 2D-to-3D mapping procedure.

### 5.4 Newton-Raphson Search (NRS)

The use of a parametric patch allows for another formulation of the problem that is : given a 3D point $A = (lc, cd, sd)$ associated to $(f, g)$, find its projection $(s, t)$ on the patch such that the distances between $A$ and $B(s, t) = (x(s, t), y(s, t), z(s, t))$ is minimum. In other words, find the best $(s, t)$ such that the Euclidean distance between $A$ and $B(s, t)$ ( namely $||A - B(s, t)||$ ) is minimum. Unfortunately, when dealing with bicubic patches, there is no known deterministic solution to that problem as it requires to find the roots of a fifth-degree polynomial. As explained by Plass and Stone [6], a solution is to assign an initial approximation of $s$ and $t$ and solve it using a conventional fixed-point scheme. Since we want to minimize $||A - B(s, t)||$ we assume that the best $(s, t)$ is the one for which $\frac{\partial}{\partial s}||A - B(s, t)|| = 0$ and $\frac{\partial}{\partial t}||A - B(s, t)|| = 0$. According to the Newton-Raphson formula

$$s^{[k+1]} = s^{[k]} - \frac{F_s}{F_s'} \qquad t^{[k+1]} = t^{[k]} - \frac{F_t}{F_t'}$$

where $k$ is an iterator, $F_s = \frac{\partial}{\partial s}||A - B||$ and $F_t = \frac{\partial}{\partial t}||A - B||$. In our implementation, Newton-Raphson
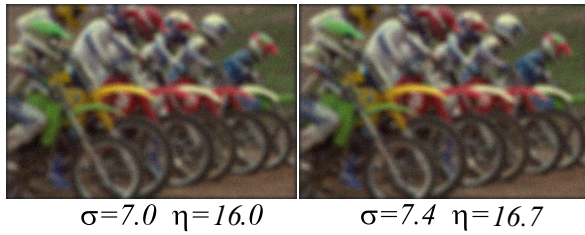
$\sigma=7.0 \quad \eta=16.0$       $\sigma=7.4 \quad \eta=16.7$

**Figure 3.** Typical result obtained by a search strategy. Given an image $f$ (here *Bikes*) and a degraded image $g$ whose degradation parameters are $(7.0, 16.0)$, the recovered parameters $(7.4, 16.7)$ give a new figure $\hat{g}$ (on the right) that is very similar to $g$ (on the left).

stops when both $|s^{[k+1]} - s^{[k]}|$ and $|t^{[k+1]} - t^{[k]}|$ are below 0.0001. Note that in our implementation, the initial approximation of $s$ and $t$ is obtained by averaging the $s$ and $t$ values of the 4 nearest control points. At convergence, $\hat{\sigma}$ and $\hat{\eta}$ are recovered by multiplying $s$ and $t$ by $\sigma_{\text{MAX}}$ and $\eta_{\text{MAX}}$.

## 6. Results

In order to gauge performances, we tested the four search strategies on the 29 real-life images of the LIVE database [3]. These images, whose size is between $610 \times 488$ and $768 \times 512$, are applied a total of $225$ distortions $(\sigma, \eta)$ linearly distributed between $(0, 0)$ and $(\sigma_{\text{MAX}}=20, \eta_{\text{MAX}}=25)$. This leads to a total of $6525$ degraded images. For every degraded image, the estimated amount of degradation $(\hat{\sigma}, \hat{\eta})$ is compared to the actual amount $(\sigma^*, \eta^*)$. The overall results are presented in Fig. 4 in which the mean and standard deviation of errors for every algorithm is presented.

As far as precision is concerned, the brute force algorithm produces the best results with $\hat{\sigma} = 0.95\%$ error for blur and $\hat{\eta} = 5.40\%$ error for noise. Note that those results can be slightly improved by using more samples. Fig. 4 also underscores the fact that smart search strategies can retrieve $(\hat{\sigma}, \hat{\eta})$ values close to that obtained by brute force but with less CPU effort. As shown in the third graphic of Fig. 4, simplex search (SS), simplex over bicubic patch search (SBS), and Newton-Raphson search (NRS) are an order of magnitude faster than the brute force search (BFS). In particular, NRS exhibits a 43-fold speed improvement with respect to BFS, yet attaining $\hat{\sigma} = 6.35\%$, $\hat{\eta} = 11.55\%$.

We observed that it is somewhat difficult to recover the noise level $\eta$ with high precision. The reason being that the manifold is usually cramped on this axis near the origin $(1, 1, 1)$ as shown in Fig. 1. High proximity of values in these areas hinder even the brute force algorithm's performance: compare $\hat{\sigma} = 0.95\%$ versus $\hat{\eta} = 5.40\%$.

## 7. Conclusion

In this paper, we presented a method to simultaneously estimate the amount of Gaussian blur and Gaussian noise in a distorted image $g$ given a reference image
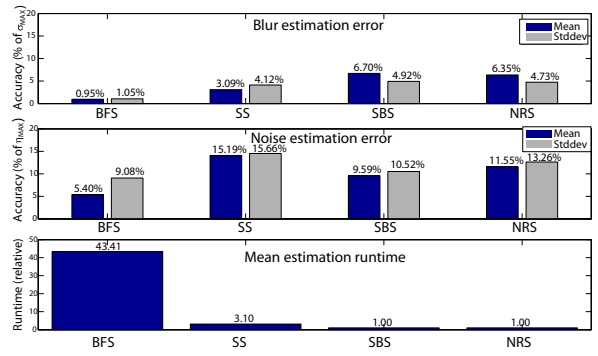


**Figure 4.** Performance and accuracy comparison of four search algorithms.

$f$. We have shown that, provided a multi-dimensional quality metric (MS-SSIM), distortion values $(\sigma, \eta)$ are associated to 3D points which form a manifold in the MS-SSIM space. We formalized the estimation problem as a search problem for which $4$ different search algorithms have been proposed. The simple brute force search algorithm is the slowest approach but returns the most precise solutions. The fastest methods are those involving a bi-cubic patch whose shape is close to the manifold. Interestingly, although a patch is an approximation of the real manifold, its nonetheless reduces the noise error when compared to simplex search.

## References

[1] D. M. Chandler and S. S. Hemami. VSNR: A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Tr. Image Process.*, 16(9):2284–2298, 2007.

[2] R. Gonzalez and R. Woods. *Digital image processing, 2nd Ed.* Prince Hall, 2001.

[3] Laboratory for Image & Video Engineering, University of Texas. LIVE Image Quality Assessment Database. *http://live.ece.utexas.edu/research/Quality*, 2002.

[4] J. Lagarias, J. Reeds, M. Wright, and P. Wright. Convergence properties of the nelder-mead simplex method in low dimensions. *SIAM J. Opt.*, 9(1):112–147, 1998.

[5] L. Piegl and W. Tiller. *The NURBS book (2nd ed.).* Springer-Verlag, Inc., New York, USA, 1997.

[6] M. Plass and M. Stone. Curve-fitting with piecewise parametric cubics. *SIGGRAPH*, 17(3):229–239, 1983.

[7] S. Rezazadeh and S. Coulombe. A novel approach for computing and pooling structural similarity index in the discrete wavelet domain. In *International Conference on Image Processing*, Cairo, Egypt, Nov. 2009.

[8] H. R. Sheik and A. C. Bovik. A visual information fidelity measure for image quality assessment. *IEEE Tr. Image Process.*, 15(2):430–444, 2006.

[9] S.-C. Tai and S.-M. Yang. A fast method for image noise estimation using laplacian operator and adaptive edge detection. In *International Symposium on Communications, Control and Signal Processing*, St Julians, June 2008.

[10] B. A. Wandell. *Foundations of Vision.* Sinauer Associates, Sunderland, Massachusetts, 1995.

[11] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: From error visibility to struc-

tural similarity. *IEEE Tr. Image Process.*, 13(4):600–612, 2004.