A Novel Error Metric for Parametric Fitting
of Point Spread Functions

Jonathan D. Simpkins and Robert L. Stevenson
University of Notre Dame, 275 Fitzpatrick Hall, Notre Dame, IN 46556, USA

ABSTRACT
Established work in the literature has demonstrated that with accurate knowledge of the corresponding blur kernel (or point spread function, PSF), an unblurred prior image can be reliably estimated from one or more blurred observations. It has also been demonstrated, however, that an incorrect PSF specification leads to inaccurate image restoration. In this paper, we present a novel metric which relates the discrepancy between a known PSF and a choice of approximate PSF, and the resulting effect that this discrepancy will have on the reconstruction of an unblurred image. Such a metric is essential to the accurate development and application of a parameterized PSF model.

Several error measures are proposed, which quantify the inaccuracy of image deblurring using a particular incorrect PSF. Using a set of simulation results, it is shown that the desired metric is feasible even without specification of the unblurred prior image or the radiometric response of the camera. It is also shown that the proposed metric accurately and reliably predicts the resulting deblurring error from the use of an approximate PSF in place of an exact PSF.

Keywords: Point Spread Function, Image Restoration, Noise Metric

INTRODUCTION

If \( Z \) represents an ideal unblurred luminance image, lexicographically ordered as a vector, then the formation of an observed digital image \( y \) can be described as the application of a linear blur operator \( H \), followed by the application of a monotonically non-decreasing function \( R(\cdot) \) which maps luminance to pixel brightness (the radiometric response of a given camera), and finally the addition of noise \( n \) (Eqn. 1).

\[
y = R(HZ) + n \tag{1}
\]

Established work in the literature\(^2,3\) has demonstrated that the ideal image \( Z \) can be reliably estimated from an observation \( y \) through the minimization of the objective function \( J(\hat{Z}) \) (Eqn. 2). The first term, \( F(\hat{Z}) \) is a fidelity penalization (Eqn. 3), \( V(\hat{Z}) \) is a regularization term specific to the particular estimation approach, and \( \lambda \) controls the relative weighting of the two terms.

\[
J(\hat{Z}) = F(\hat{Z}) + \lambda V(\hat{Z}) \tag{2}
\]

\[
F(\hat{Z}) = \|y - R(HZ)\|^2 \tag{3}
\]

Our research seeks to accurately describe the blur operator using a parametric model with finite degrees of freedom,\(^4\) and it is certain that a parametric fit \( \tilde{H} \) can only be a close approximation to the ground-truth measured blur operator \( H \). The question that this paper intends to address is what is a good metric \( M(H, \tilde{H}) \) which should be minimized in order to choose an “optimal” approximation \( \tilde{H} \) of a known blur operator \( H \).

Some care should be taken when considering this problem statement: the discrepancy between \( H \) and \( \tilde{H} \) is not due to measurement error, it is due to the limited number of degrees of freedom in the parametric model which is able to produce \( \tilde{H} \).
PROPOSED MEASURES OF PSF ACCURACY

Since the fidelity function \( F(\cdot) \) is essential to image reconstruction, it is desirable for the score of the fidelity function using the correct unblurred image and PSF to be approximately equal to the fidelity function using the correct unblurred image and the approximate PSF. An error measure that reflects this can be posed as:

\[
\chi_F (H, \hat{H}, Z, \mathcal{R}(\cdot)) = 10 \log_{10} \left( \frac{\text{Var}[\mathcal{R}(HZ)]}{\|\mathcal{R}(HZ) - \mathcal{R}(\hat{HZ})\|^2} \right)
\]

This measure can be interpreted as similar to the image SNR, where the “noise” is the discrepancy between the blurred ground-truth image using the correct blur operator and the blurred ground-truth image using the approximate blur operator.

Another important measure to consider is the degradation of restoration performance, using the approximate blur operator instead of the ground-truth blur operator. Let a function \( \Psi(y, H) \) be defined as the optimal estimator of the unblurred image \( Z \), given a particular image prior \( \mathcal{V}(\hat{Z}) \) and a weighting parameter \( \lambda \):

\[
\Psi(y, H) = \arg \min_{\hat{Z}} \left( \|y - \mathcal{R}(H\hat{Z})\|^2 + \lambda \mathcal{V}(\hat{Z}) \right)
\]

The degradation of restoration performance due to using an approximate blur operator can then be posed as:

\[
\chi_\Psi (H, \hat{H}, Z, \mathcal{R}(\cdot)) = 10 \log_{10} \left( \frac{\|\mathcal{R}(Z) - \mathcal{R}(\Psi(\mathcal{R}(HZ), \hat{H}))\|^2}{\|\mathcal{R}(Z) - \mathcal{R}(\Psi(\mathcal{R}(HZ), H))\|^2} \right)
\]

This measure can be interpreted as the difference between the SNR of the restored image, using the true blur operator in restoration, and the SNR of the restored image using the approximate blur operator. In this paper, the two measures of this type that will be considered are \( \chi_{\Psi, LR} \), denoting image restoration using the Lucy-Richardson algorithm, and \( \chi_{\Psi, RD} \), denoting regularized filter deconvolution. These restorations algorithms were implemented using, respectively, the MATLAB Image Processing Toolbox functions “deconvlucy” and “deconvreg.”

PROPOSED METRIC COMPONENTS

The proposed metric \( \mathcal{M}(H, \hat{H}) \) for each of the three error measures is constructed from a combination of 4 different component functions. These functions came from a candidate set of 110 different functions designed to assess different disparities between \( H \) and \( \hat{H} \), and these 4 were selected based on a thorough analysis of the simulation results discussed later in this paper.

The first component \( \sigma_C(H, \hat{H}) \) is designed to measure the correlation of the “noise” \( N = H - \hat{H} \) with the values in both \( H \) and \( \hat{H} \). If \( \rho_{N,H} \) denotes the Pearson correlation coefficient between \( N \) and \( H \), then \( \sigma_C(H, \hat{H}) \) is defined as:

\[
\sigma_C(H, \hat{H}) = \log \left( (1 - \rho_{N,H}) * (1 - \rho_{N,\hat{H}}) \right)
\]

The second component \( \sigma_S(H, \hat{H}) \) is designed to measure how spatially “spread out” the approximate blur kernel and the difference between the exact and approximate blur kernels are. A function \( S(K) \) which provides a description of the spatial spread of a 2D filter \( K \) is defined as:

\[
S(K) = \left( \frac{\sum_{x,y} |K_{x,y}|((x - \mu_x)^2 + (y - \mu_y)^2)}{\sum_{x,y} |K_{x,y}|} \right)
\]

\[
\mu_x = \frac{\sum_{x,y} x |K_{x,y}|}{\sum_{x,y} |K_{x,y}|}
\]

\[
\mu_y = \frac{\sum_{x,y} y |K_{x,y}|}{\sum_{x,y} |K_{x,y}|}
\]
\[
\mu_y = \frac{\sum_{x,y} y |K_{x,y}|}{\sum_{x,y} |K_{x,y}|}
\]  

(10)

If \( K_H \) is the 2D spatial blur kernel corresponding to the blur operator \( H \), then \( \sigma_S(H, \tilde{H}) \) is defined as:

\[
\sigma_S(H, \tilde{H}) = S(K_\tilde{H}) * S(K_H - K_\tilde{H})
\]

(11)

The third and fourth components \( \sigma_{SNR}(H, \tilde{H}) \) and \( \sigma_{AltSNR}(H, \tilde{H}) \) measure the power of \( N \) relative to the power of \( H \) and \( \tilde{H} \), respectively:

\[
\sigma_{SNR}(H, \tilde{H}) = \log \left( \frac{\sum_i (H_i)^2}{\sum_i (H_i - \tilde{H}_i)^2} \right)
\]

(12)

\[
\sigma_{AltSNR}(H, \tilde{H}) = \log \left( \frac{\sum_i (\tilde{H}_i)^2}{\sum_i (H_i - \tilde{H}_i)^2} \right)
\]

(13)

The complete proposed metric is a 6-parameter weighted combination of these components, where the weights \( \{w_1, w_2, \ldots, w_6\} \) vary depending on which error measure the metric is designed to predict:

\[
\mathcal{M}(H, \tilde{H}) = w_1 \sigma_C(H, \tilde{H}) + w_2 \sigma_S(H, \tilde{H}) + w_3 \sigma_{SNR}(H, \tilde{H}) + w_4 \sigma_{AltSNR}(H, \tilde{H}) + w_5 \left( \sigma_{SNR}(H, \tilde{H}) * \sigma_{AltSNR}(H, \tilde{H}) \right) + w_6
\]

(14)

**SIMULATION APPROACH**

The proposed error measures depend on specific choices of the unblurred image \( Z \) and the camera radiometric response function \( R(\cdot) \). The set of unblurred images used in the simulations were grayscale copies of a set of four canonical images from image processing: “Lena,” “Cameraman,” “Mandrill,” and “Peppers” (Fig. 1, left). Each image is 512 by 512 pixels, and stored in 8-bit grayscale. These images, however, were treated in each simulation as \( R(Z) \), so the inverse function \( R^{-1}(\cdot) \) of the given radiometric response was used to simulate the corresponding luminance image. The set of radiometric response functions used in the simulations were 6 curves taken from the CAVE DoRF database\(^5\) (Fig. 1, right), representing film, digital, and synthetic radiometric camera responses.

Figure 1. The four unblurred images (left) and the six radiometric response curves (right) used in the simulations.
For each simulation, a ground-truth PSF and approximate PSF were created, and then the measures $\chi_F$, $\chi_{\Psi,LR}$, and $\chi_{\Psi,RD}$ were computed for each of the 24 unique combinations of image and radiometric response. The ground-truth PSFs generated were elliptical Gaussian kernels, with standard deviations along the principal axes of between 0.4 and 5.0 pixels, and a uniformly random orientation of the major principal axis. The “approximate” PSF in each simulation was generated by adding non-white noise to the ground truth, and rescaling the result to preserve non-negativity and summation to unity. The non-whiteness of the noise was randomly chosen to increase noise variance in the presence of PSF signal, the absence of signal, or some combination of the two (Fig. 2).

![Figure 2](image)

Figure 2. An example ground-truth PSF (left), and simulated “approximate” PSFs, featuring noise accentuated in the absence of signal (center left), the presence of signal (right), or some combination of the two (center right).

For each simulated $H$, 50 different simulation runs were executed, each with a different random $\tilde{H}$; this was done to ensure a variety of approximate PSFs and corresponding error measures for each ground-truth PSF.

RESULTS

The simulation results consist of 1,125 different ground-truth PSFs, 56,250 corresponding approximate PSFs, and 1.3 million different sets of $\chi_F$, $\chi_{\Psi,LR}$, $\chi_{\Psi,RD}$ calculated error measures.

Verification that Metric Can Exist

A preliminary question that should be asked is if the proposed error metric can even be expected to exist: each of the error measures $\chi_F$, $\chi_{\Psi,LR}$, $\chi_{\Psi,RD}$ depends not only on $H$ and $\tilde{H}$, but also on $R(\cdot)$ and $Z$, so is it even possible to have a metric $M(H, \tilde{H})$ that does not depend on the unblurred image or the camera’s radiometric response? To begin addressing this question, the linear correlation coefficient is considered between corresponding results for each error measure, differing only in the selection of $R(\cdot)$ and $Z$ (Fig. 3).

![Figure 3](image)

Figure 3. The linear (Pearson) correlation coefficients between resulting vectors of 56,250 error measures differing only in the selection of $R(\cdot)$ and $Z$. The combination index spans the 4 choices of $Z$ and 6 choices of $R(\cdot)$, and the error measures are $\chi_F$ (left), $\chi_{\Psi,LR}$ (center), and $\chi_{\Psi,RD}$ (right).

From these results, it is seen that each error measure for a choice of $R(\cdot)$ and $Z$ is well related by an affine function to what the error measure would have been under a different choice of $R(\cdot)$ and/or $Z$. The desired metric $M(H, \tilde{H})$ can therefore be expected to exist in a way that is affinely related to one of the error measures for any particular choice of unblurred image and radiometric response.

Suppose that the vector of metric scores for a “universal” choice of $Z$ and $R(\cdot)$ is denoted as $\Gamma = \{\gamma_1, \gamma_2, \ldots\}$, that the corresponding vector of metric scores for a particular choice of $Z$ and $R(\cdot)$ is denoted as $\Delta_j = \{\delta_{1,j}, \delta_{2,j}, \ldots\}$, and that the corresponding elements of the two vectors are ideally related as $\delta_{i,j} = \alpha_j \gamma_i + \beta_j$. 
The vector $\Delta_j$ represents the set of actual measured errors for a particular choice of $Z$ and $R(\cdot)$, where in the simulation results, $j \in [1, 24]$. The vector $\Gamma$ is therefore determined from the set of measured error vectors as:

$$\arg \min_{\Gamma, \alpha, \beta} \sum_{j=1}^{24} \sum_i (\alpha_j \gamma_i + \beta_j - \delta_{i,j})^2$$

(15)

where $\alpha = \{\alpha_1, \alpha_2, \ldots\}$ and $\beta = \{\beta_1, \beta_2, \ldots\}$. $\alpha_1$ is chosen to be unity, $\beta_1$ is chosen to be zero, and $\Gamma$ is then found as the optimum according to this arbitrary scaling factor and shift.

For the three error measures $\chi_F, \chi_{\Psi,LR}, \chi_{\Psi,RD}$, the corresponding vectors $\Gamma_F, \Gamma_{\Psi,LR}, \Gamma_{\Psi,RD}$ were determined for the results in this way from the set of all 24 combinations of $Z$ and $R(\cdot)$. Each $\Gamma$ was highly correlated with the corresponding set of 24 $\chi_j$ result vectors, with Pearson correlation coefficients $\rho_j$ above 0.9 in all cases. The inaccuracy resulting from using the “universal” error measure $\Gamma$ in place of the “specific” measures $\chi_j$ was determined as the root-mean-squared (RMS) error between $(\alpha_j \gamma_i + \beta_j)$ and $\delta_{i,j}$ over all $i, j$ (Table 1).

Table 1. The observed similarity between each “universal” error prediction vector $\Gamma$ and the corresponding set of measured error vectors $\{\chi_1, \chi_2, \ldots, \chi_{24}\}$.

<table>
<thead>
<tr>
<th>Predicted Error Measure</th>
<th>RMS Error of Metric (dB)</th>
<th>Observation Variance Accounted For</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Gamma_F$</td>
<td>1.359</td>
<td>93.5%</td>
</tr>
<tr>
<td>$\Gamma_{\Psi,LR}$</td>
<td>1.337</td>
<td>85.2%</td>
</tr>
<tr>
<td>$\Gamma_{\Psi,RD}$</td>
<td>2.161</td>
<td>78.2%</td>
</tr>
</tbody>
</table>

The RMS errors of each are reasonable when compared to the overall distribution of resulting values for each of the error measures (Fig. 4).

Figure 4. The histograms of resulting values for $\Gamma_F$ (left), $\Gamma_{\Psi,LR}$ (center), and $\Gamma_{\Psi,RD}$ (right).

These results indicate that the error measures of interest (in particular, $\chi_F$ and $\chi_{\Psi,RD}$) can be accurately generalized to what the error measure may have been under a “universal” choice of $Z$ and $R(\cdot)$, which implies that it is feasible for the desired metric $\mathcal{M}(H, \tilde{H})$ to exist without functional dependence on $Z$ or $R(\cdot)$.

Verification of Proposed Metrics

The weights $\{w_1, w_2, \ldots, w_6\}$ in the proposed metric (Eqn. 14) were fit to each of the three error measures $\Gamma_F$, $\Gamma_{\Psi,LR}$, $\Gamma_{\Psi,RD}$ using 10-fold cross validation. The corresponding average RMS error of each is listed in Table 2, along with the average fraction of the observed variance accounted for by each predictor.

Table 2. The average performance of each metric, over the 10 folds of the cross-validation.
Looking at the relative frequency of actual vs. predicted value of the error measures (Fig. 5), it is seen that the metric is capturing the major trends in the error measures, and that it is most accurate at representing $\Gamma_F$.

![Figure 5. The relative frequency of $M(H, \tilde{H})$ predicting a value for an error measure versus the actual measurement.](image)

**CONCLUSIONS**

This paper has proposed and presented a novel metric which predicts several different measures of error that will result from using an approximate PSF in place of an exact one. The proposed metric depends only on the actual and approximate PSFs, and was verified through more than a million simulations using a variety of actual radiometric camera responses and unblurred images. In future work, the metric will be used to determine a set of optimal parameters in a parameterized PSF model.

**REFERENCES**


