# Inference for Distorted Image Reconstruction

B.S. Yandell, C.C. Taylor and B.D. Ripley
U. of Wisconsin–Madison, U. of Leeds and U. of Oxford

**Abstract**

We examine ways to measure to "closeness" of two pixel images, such as comparing a true image to a reconstruction of that image following blur, noise and distortion. Methods based on simple differences of intensity between images may be missleading due to slight shifts of a few pixels. We examine the effect of such local distortions and investigate ways to combine measures of intensity and distance differences. Methods of estimating local distortion, and blur in the presence of shift, are discussed and demonstrated on readily available images. Simulations show the promise of this approach to provide objective methods for comparing image reconstructions.

**Key words:** image restoration, image enhancement, local neighbourhood, penalised likelihood, power spectrum.
**Correspondence address:** Brian S. Yandell, Department of Statistics, University of Wisconsin, 1210 West Dayton Street, Madison, WI 53706, USA.

# 1 Introduction

We consider the problem of determining how well one has reconstructed an image. In essence, this reduces to comparing two images using some objective measure or measures of agreement. Metrics based on simple differencing of images is flawed by not detecting slight shifts or deformations [3,10]. It is also readily apparent that one measure may not be adequate to summarise discrepancies, and that one may have to rely on several measures which elicit different aspects of image reconstruction. No attempt has been made here to comprehensively review literature on image analysis and image reconstruction. The reader is referred to [5,9] and recent issues of *IEEE Trans. PAMI.* Instead, our focus is upon developing inferential tools for comparing images when, in addition to noise, there may be some shift or deformation.

Suppose that we consider an "original image" and an image "reconstructed" from that original. Alternatively, we could consider two reconstructions of a single image. Some authors prefer to use "template" and "degraded image", but we recognize that "template" has many other meanings, and the reconstruction may actually be an enhancement rather than a degradation. There are many situations where a reconstructed image may be "close" to the original, but suffer from some combination of blurring, noise and global shift or local deformation.

We propose a specific model of such a combination and explore several ways of summarizing the deviation of the reconstruction from the original image with the aim of developing inferential tools to select the "best" reconstruction or to order several reconstructions. It is important to keep in mind that there probably is not one "best" reconstruction but rather several, possibly contradictory, ones which would probably have to be weighed subjectively in any particular practical situation. One way to accomplish this it to establish one method of comparison between images, and then consider several linear transforms (or filters) which elicit features of particular interest applied to both original and reconstruction before comparison. The statistic for evaluation of "best" reconstruction would be a weighted combination of these measures based on various filters. This idea was used for boundary detection by [5].

Let the original image be $x = \{(t, x(t)); t \in T, x(t) \in V\}$, where $T$ is the pixel grid and $V$ is the intensity (brightness) range, and the reconstructed image be $y = \{(t, y(t)); t \in T, y(t) \in V\}$. Usually we consider theoretically that $T = [0,1]^2$ and $V = [0,1]$. In practice, $T$ is a 2D matrix indexed by 1 to 256, 512 or 1024, while $V = \{0, \cdots, 255\}$ represents a grey level scale. We consider reconstructions of the form

$$y(t) = [f * x](t+s) + \epsilon(t+s) = \int x(t+s(t)-u)f(u)du + \epsilon(t+s(t)), \tag{1.1}$$

in which $f$ is the blur window convolved ($*$) with the original image $x$, $s = s(t)$ is the shift and $\epsilon$ is the noise. The choice of $f$ as the delta function is equivalent to no blur. It is quite likely that the shift is not uniform over the image, and may be better thought of as local deformation. For completeness, $x$ should be extended to $T(s) = \{t + s(t); t \in T\}$ in a natural way.

Our ultimate aim is to determine how "far" $y$ is from $x$. We proceed by (1) estimating $\hat{x}$ from $y$ assuming no shift or local deformation and (2) estimating $\hat{s}$ by deforming $\hat{x}$ closer to $x$. In some cases, $f$ is of unknown form, or at least has unknown bandwidth. One could assess the "closeness" of the two images by summaries based on $\hat{f}$, $\hat{s}$, and the residuals for $x$

$$\hat{\eta}(t) = x(t) - \hat{x}(t - \hat{s}(t))$$

and for $y$

$$\hat{\epsilon}(t) = y(t - \hat{s}(t)) - [\hat{f} * \hat{x}](t - \hat{s}(t)).$$

We consider several approaches in the present study. Section 2 considers estimation and detection of noise and blur, with emphasis on problems that arise when there is an undetected shift or local deformation. Section 3 examines estimation of global shifts or local deformations of one image relative to the other. Section 4 concerns other approaches to measuring deviation, including the combinations of distance and intensity metrics introduced by Baddeley [2,3] and examination of the power spectrum. Section 5 compares measures applied to several images with artificially added blur, noise and/or local deformation.

## 2 Detecting Noise and Blur

Ignoring shift for the moment, the reconstruction model becomes

$$y(t) = [f * x](t) + \epsilon(t).$$

We want to estimate the blur and noise, arriving at some measure of closeness of the reconstruction to the original image. Natural candidates, explored below, involve tests of the bandwidth in the blur kernel $f$ and of the size of the residual variation $\hat{\epsilon}$. If the kernel integrates to 1 and is sufficiently smooth, then a Taylor series approximation yields

$$y(t) \approx x(t) - \int \dot{x}^u(t) f(u) du + \int \ddot{x}^u(t) f(u) du/2 + \epsilon(t),$$

where $\dot{x}^u = u^{\mathrm{T}}\dot{x} = (u_1\dot{x}_1, u_2\dot{x}_2)$, with $\dot{x}_j = \partial x/\partial t_j$, and $\ddot{x}^u = u^{\mathrm{T}}\ddot{x}^u$, with $\ddot{x}_{ij} = \partial^2 x/\partial t_i \partial t_j$, $i, j = 1, 2$. In the case of symmetric $f$ and continuous $\dot{x}$ near $t$, $y(t) \approx x(t) + \int \ddot{x}^u(t) f(u) du/2$. If, further, $f$ is separable, then

$$y(t) \approx x(t) + \Delta x(t)\sigma_f^2/2 + \epsilon(t), \tag{2.1}$$

where $f(u) = f_1(u_1)f_1(u_2)$, $\sigma_f^2 = \int v^2 f_1(v) dv$ and the Laplacian is $\Delta x(t) = \ddot{x}_{11}(t) + \ddot{x}_{22}(t)$. Note that the bandwidth of $f$ is proportional to $\sigma_f$. Clearly if one knew more about the form of $f$, one could use this information to improve local approximations, *e.g.*, by deconvolution [8].

Regressing $y(t) - x(t)$ on $\Delta x(t)$ with zero intercept provides an estimate of $\sigma_f^2$, and of the noise process. One could then test if the slope were zero (no blur) and determine if the noise were negligible ($\sigma_\epsilon^2 = 0$), and use these or various regression diagnostics to examine the blur and/or noise. An interesting approach to examining patterns in residuals was suggested by Jubb [7]. One could determine the frequency distribution of "groups" of contiguous "informative pixels" (e.g., $\hat{\epsilon}(t)$ at least $k\sigma_\epsilon$), and compare this with Monte Carlo simulations to assess the goodness of fit.

### 2.1 Neighbourhood Intensity Differences

It is necessary to consider the ramifications of ignoring possible shift when blur and noise are present. We examine differences in intensity between two images which have been displaced slightly. That is, for every pixel on one image, one compares the intensity with pixels in the neighbourhood on the other image:

$$d_r(t) = d_{r,x,y}(t) = \min_{0 \le |\delta| \le r} \{d_V(x(t), y(t+\delta))\}.$$

We consider some of the implications of this for small $r$. Suppose that $d_V(a, b) = |a - b|$, and let $\bar{d}_r(t) = d_{r,y,x}(t)$. Note that $\bar{d}_0 = d_0$. Throughout the discussion below we use a local Taylor series expansion of the intensity value near $x(t)$ as

$$\begin{aligned} x(t+\delta) &= x(t) + \delta^{\mathrm{T}}\dot{x}(t) + \delta^{\mathrm{T}}\ddot{x}(t)\delta/2 + o(|\delta|^2) \\ &= x(t) + \dot{x}^\delta(t) + \ddot{x}^\delta(t)/2 + o(|\delta|^2), \end{aligned}$$

assuming for the discussion that higher order derivatives of $x$ are uniformly bounded over the image; that is, we impose a prior belief that the original image curvature does not change too rapidly anywhere.

### 2.2 Neighbourhoods of Noise and Blur

A reconstructed image which is simply the original plus (independent) noise, $y(t) = x(t) + \epsilon(t)$, yields differences $d_0(t) = |\epsilon(t)|$. Thus, with no shift, one could simply examine the residual difference of the two images, estimate the magnitude of noise (MSE) and test for spatial patterns, and evaluate closeness on the basis of such tests. If, however, there has been some deformation, it may be important to examine differences in the neighbourhood of $t$. Thus

$$d_r(t) = \min(|\epsilon(t)|, \min_{0 < |\delta| \le r} \{|x(t) - x(t+\delta) - \epsilon(t+\delta)|\})$$

$$\approx \min(|\epsilon(t)|, \min_{0<|\delta|\leq r} \{|\dot{x}^\delta(t) + \ddot{x}^\delta(t)/2 + \epsilon(t+\delta)|\}).$$

The derivatives can be bounded by $|\dot{x}^\delta| \leq r\dot{M}$ and $|\ddot{x}^\delta| \leq r^2\ddot{M}$ for some $M(\cdot)$ and, if the image is sufficiently smooth, this bound can be uniform. This leads to a crude upper bound

$$d_r(t) \leq \min(|\epsilon(t)|, [r\dot{M}(t) + r^2\ddot{M}(t)/2 + \min_{0<|\delta|\leq r} \{|\epsilon(t+\delta)|\}]),$$

or, for $r = 1$,

$$d_1(t) \leq \min(|\epsilon(t)|, [\dot{M}(t) + \ddot{M}(t)/2 + \min_{|\delta|=1} \{|\epsilon(t+\delta)|\}]).$$

If $Pr(|\epsilon(t)| > \beta) = \alpha$, $|\dot{M}(t)| \leq m_1$, and $|\ddot{M}(t)| \leq m_2$, then for a square pixel grid with 4 nearest neighbours, $Pr(d_0(t) > \beta) = \alpha$ and $Pr(d_1(t) > m_1 + m_2/2 + \beta) \leq \alpha^5$. If we reverse the sense of $x$ and $y$, we find that

$$\bar{d}_r(t) \approx \min(|\epsilon(t)|, \min_{0<|\delta|\leq r} \{|\epsilon(t) - \dot{x}^\delta(t) - \ddot{x}^\delta(t)/2|\}).$$

Thus only one value of $\epsilon(t)$ enters and one does not have the exponential decay with $r$ as found for $d_r$. Reconstructions which merely add noise to original images with mild gradient and curvature should yield $\bar{d}_r$ which decays very slowly, while $d_r$ decays very rapidly.

Blurring yields slightly different neighbourhood metrics, namely, $d_0 \approx |\Delta x(t)|\sigma_f^2/2$,

$$d_r \approx \min(|\Delta x(t)|\sigma_f^2/2, \min_{0<|\delta|\leq r} \{|r\dot{x}^\delta(t) + r^2\ddot{x}^\delta(t)/2 + \Delta x(t+\delta)\sigma_f^2|/2\}),$$

$$\bar{d}_r \approx \min(|\Delta x(t)|\sigma_f^2/2, \min_{0<|\delta|\leq r} \{| - r\dot{x}^\delta(t) - r^2\ddot{x}^\delta(t)/2 + \Delta x(t)\sigma_f^2|/2\}).$$

Note that $\Delta x(t+\delta) \approx \Delta x(t)$ provided the (uniformly bounded) third derivatives of $x$ are small. Thus $\bar{d}_r$ and $d_r$ should behave similarly for blurred reconstructions for which the original has mild gradient and curvature, with only mild decay as $r$ increases.

# 3 Estimating the Shift

If the shift is unknown, then we must estimate it. We consider three approaches: (1) global estimation of constant shift, (2) local (neighbourhood) search using minimal deviations, and (3) penalised likelihood estimation of local deformations. The third approach is intermediate between a simple (global) model and fidelity to the data found in the second approach. For this section, we assume there is no blur or noise, or rather that these have been estimated and removed using methods of the previous section.

Reconstructions which involve only deformations of the original image, $y(t) = x(t+s)$ with $s = s(t)$, are approximately

$$y(t) \approx x(t) + s^{\mathrm{T}}\dot{x}(t) + s^{\mathrm{T}}\ddot{x}(t)s/2. \tag{3.1}$$

One could estimate the global shift (if any) by non-linear least squares applied to (3.1), or more crudely by ordinary least squares, dropping the $\ddot{x}^s$ term. Alternatively, one could simply shift the images relative to one another and find that shift $s$ which gives the "best" fit, say in terms of minimizing $\sum[y(t) - x(t+s)]^2$. We have used this latter approach with a cascade algorithm along the lines of Jin and Mowforth [6].

If one examines neighbourhoods with $r \geq \max_{t\in T} ||s(t)||$, then $d_r(t) = \bar{d}_r(t) = 0$ in the absence of noise or blur. The gradient and curvature can play an important part in how fast agreement is determined for $r < |s|$. If $|s|$ is small, then one can quickly find the shift. If $|s|$ is large, then

$$y(t+\delta) - x(t) \approx (s+\delta)^{\mathrm{T}}\dot{x}(t) + (s+\delta)^{\mathrm{T}}\ddot{x}(t)(s+\delta)/2$$

$$\approx \dot{x}^\delta(t) + \ddot{x}^\delta(t)/2 + \delta^{\mathrm{T}}\ddot{x}(t)s + \dot{x}^s(t) + \ddot{x}^s(t)/2.$$

Clearly, for shifts, $d_r = \bar{d}_r$ (allowing for the boundary), but it may be difficult to distinguish deformation from blur, since blur introduces the Laplacian $\Delta x$ which can be confused with $\ddot{x}^s$. Thus it is important to remove any blur *before* estimating local deformation. In the case of local deformation, some nonparametric approach might be preferred, allowing $s(t)$ to change smoothly over $t \in T$ [1,6]. These are explored in the next two subsections.

### 3.1 Penalised Likelihood for Local Deformations

If one considers the original problem with blur, shift and noise, the problem can be formulated as trying to characterize the displacement $s$ and the underlying image intensity $z$, if one knows the blur function $f$. The simultaneous solution can be formulated as finding $z(t)$ and $s(t)$ to minimise

$$\sum_{t\in T}[y(t) - f * z(t)]^2 + \sum_{t\in T}[f * z(t) - f * x(t + s(t))]^2 + \lambda_V P_V(z) + \lambda_T P_T(s),$$

with $P_V$ and $P_T$ penalties to ensure smooth intensity and displacement surfaces, respectively. While one may choose to estimate $s$ and $z$ simultaneously, we have implicitly proposed above to separate this into the following two problems: (1) find $z$ to minimize

$$\sum_{t\in T}[y(t) - f * z(t)]^2 + \lambda_V P_V(z),$$

and (2) given $z$, find $s$ to minimize

$$\sum_{t\in T}[z(t) - x(t + s(t))]^2 + \lambda_T P_T(s).$$

It may be important for some applications to consider a broader definition of deformation [1]. Let $r_0$ be a "template" from some set of objects $R$ which is deformed as $r = \tau(r_0)$. One actually observes $y = \gamma(r)$, with $\gamma$ mapping objects $r \in R$ into the image space $I$. That is, the original image is $x = \gamma(r_0)$, and the reconstructed image is $y = \gamma(\tau(r_0))$. A special case has the deformation as a local shift of the original image, $y(t) = x(t + s(t))$, although more general types of deformations are possible, subject to suitable regularity conditions. Amit considers in general the displacement of the object,

$$r(t) = \tau(r_0(t)) = r_0(t + s(t)),$$

which is finally seen via the map as $y(t) = \gamma(r_0(t + s(t)))$. Thus $s(T) = \{s(t) = (h(t), v(t)), t \in T\}$ is a deformation of the original pixel grid, with $h(T)$ the horizontal and $v(T)$ the vertical displacement. If we extend $x(\cdot)$ to be zero outside $T$, our mapping is well defined.

To ensure that the problem is well-posed, we consider $S = \{(S(t)), t \in T\}$ as a zero-mean continuous bivariate Gaussian random field and construct a reproducing kernel Hilbert space [1,12] which corresponds to a model regarding the unit square as an elastic body subject to random forces through linearisation. The covariance matrix field is specified implicitly by the reproducing kernel, $L(s) = (L_h(s), L_v(s))$, with

$$L_h(s) = L_h((h, v)) = \Delta h + (\beta - 1)(\ddot{h}_{11} + \ddot{v}_{12}) = \ddot{h}_{22} + \beta\ddot{h}_{11} + (\beta - 1)\ddot{v}_{12}$$

and similarly for $L_v(s)$ exchanging 1 and 2, with $\beta \geq 0$. Note that the choice $\beta = 1$ yields independent displacement, which may hold in some limited situations.

The estimation of $s = (h, v)$ is rather involved, with a number of approximations presented by Amit *et al.* The basic problem revolves around selecting $s$ to minimise

$$2\sum_{t\in T} V_t\{y(t), \gamma(r_0(t + s(t)))\} + \lambda P(s),$$

with norm $P(s) = ||L(s)||^2$. Here $\lambda$ controls the tradeoff between smoothness imposed by the penalty norm $P$ and fidelity to the degraded image $y$. The $V_t$ are the (non-negative) potentials, evaluated independently at each pixel location, *e.g.* squared deviation, corresponding to a Gaussian distribution for $Y(t) = y(t)$ given $S(t) = s(t)$.

Computationally, this is a difficult problem, as one has a 2-D spline over a pixel grid which is at least $256^2$ and potentially much larger. It may be preferable to follow some local smoothing approach [4], with appropriate modification of the penalty. Instead, we chose to implement the "cascade" algorithm of [6].

## 3.2 Coarse to Fine Deformation

An alternative to the linearisation above is to estimate local deformations using a cascade of coarse to fine filters. Loosely speaking, one creates very coarse blurring (large bandwidth) of both images and shifts one relative to the other on the scale of the blurring. The process uses estimates of local curvature (either Laplacians or differences of Gaussians) to highlight edges, the key features for detecting local deformation. The scale of blurring and of deformation is progressively reduced until finally one is examining the unblurred images on a pixel-by-pixel basis. At each stage, one corrects for the deformation detected at the previous scale, and progressively refines the estimate of deformation. This idea is a standard approach for image enhancement, but its application to comparing images appears to be due to [6]. The idea follows most nonparametric regression schemes of extracting the signal (low order, coarse features of deformation) from noise (high order, fine features).

The algorithm proceeds as follows. First determine and apply any global shift using methods of the previous subsection. Then set $s(t) = \mathbf{0}$. Let

$$G(t,\sigma) = \nabla^2\phi(t;0,\sigma) = -\pi^{-1}\sigma^{-4}(1-|t|^2/2\sigma^2)\exp(-|t|^2/2\sigma^2)$$

be the Laplacian of the two-dimensional Gaussian filter. Apply this to both $x$ and $y$ for some initial $\sigma_G$ to create $x_G(t,\sigma_G) = [G(\cdot,\sigma_G) * x](t)$ and similarly $y_G(t,\sigma_G)$. For each $t$ find $\delta(t)$ which yields the "best" agreement between $x_G(t,\sigma_G)$ and $y_G(t,\sigma_G)$. This agreement might be measured in terms of locally smoothed correlations, as in [6], or some other measure such as smoothed deviations:

$$\sum_{u\in T}[x_G(t+u,\sigma_G) - y_G(t-s(t)-\delta(t)+u,\sigma_G)]^2 w(u,\sigma_w), \tag{3.2}$$

with $w$ some density and $\sigma_w$ suitably chosen. Update $s(t)$ to $s(t)+\delta(t)$, reduce $\sigma_G$ and $\sigma_w$ by a factor of 2 and repeat the procedure until $\delta(\cdot)$ is suitably small. Jin and Mowforth point out that replacing (3.2) by

$$\sum_{u\in T}[x_G(t+u,\sigma_G) - y_G(t+u-s(t+u)-\delta(t+u),\sigma_G)]^2 w(u,\sigma_w),$$

is computationally unstable. We have observed this in practice, with fine features irretrievably lost at a very coarse level.

It may be important for some purposes to use different kinds of filters in place of the difference of Gaussian or the Laplacian. For instance, pictures dominated by texture features might suggest some other filter to detect shifts, perhaps along the lines of [5].

# 4 Other Approaches

## 4.1 Combining Distance and Intensity

Baddeley [3] proposed a metric for comparing images which combines distance and intensity metrics. Let $d_T(s,t)$ be a distance metric (e.g., Euclidean distance), and let $d_V(x(t),y(s))$ be an intensity metric (e.g., $|x(t)-y(s)|$). For convenience, we only consider symmetric metrics, although the discussion could be easily generalised to include the nonsymmetric case. Baddeley defined the "$\lambda$-metric" using information generated at each pixel. Letting

$$\delta_\lambda(t) = \delta_{\lambda,x,y} = \inf\{a \geq 0 : \exists s \in T, d_T(s,t) \leq \lambda a, d_V(x(t),y(s)) \leq a\},$$

Baddeley suggested the "$\lambda$-metric" as the supremum of $\delta_\lambda(t)$, $\Delta_\lambda = \sup_{t\in T}\delta_\lambda(t)$. As pointed out by Baddeley, $\Delta_\lambda$ is a metric which suffers from a lack of robustness to outliers. We propose instead to consider $\alpha$-quantiles $\Delta_\lambda(\alpha)$ of $\{\delta_\lambda(t) : t \in T\}$. It can be readily shown that $\Delta_\lambda(\alpha)$ is a metric, following an argument similar to Baddeley.

Baddeley noted that $\Delta_\lambda$ is a decreasing function of $\lambda$, with the limit, corresponding to the $L^\infty$ metric, at some $\lambda \leq \sup_{s,t\in T} d_T(s,t)$. Note also, the different types of images may have different average levels and ranges of intensities. Therefore, it seems reasonable to select a neighbourhood size $d$ and determine $\lambda$ based on intensity differences in the neighbourhood of $t \in T$. For each $t$, consider the non-increasing stochastic process $V(d;t) = \inf\{d_V(x(t), y(s) : s \in T, d_T(s,t) \leq d\}$. Baddeley's metric is $\delta_\lambda(t) = \max(V_\lambda(t), T_\lambda(t)/\lambda)$, with

$$V_\lambda(t) = \sup\{v : v = V(d;t) \geq d/\lambda\},\ T_\lambda(t) = \inf\{d : v(d;t) = V_\lambda(t)\}.$$

Thus one might determine $\lambda$ from the $\alpha$-quantile of selecting $\lambda$ based on a quantile of the intensities in the set $V(d) = \{V(d;t) : t \in T\}$. Letting $V(d)(\alpha)$ be the $\alpha$-quantile, we choose $\lambda = V(d;\alpha)/d$. We have found that for many $256^2$ images with moderate blurring or noise, a neighbourhood of size $d = 6$ seems fairly robust.

Various plots suggest themselves to elicit the tradeoff of distance and intensity. One could plot $\{(t, T_\lambda(t)), t \in T\}$ and $\{(t, V_\lambda(t)), t \in T\}$ to show where discrepancies are found, indicating the relative importance of distance and intensity. Instead, we propose plotting bivariate histograms of $\{(T_\lambda(t), V_\lambda(t)), t \in T\}$ on a logarithmic scale.

### 4.2 Power Spectrum

As expected, blurring eliminates high frequencies, and uncorrelated noise introduces a uniform background level at all frequencies. Distortion has minimal effect. Here, feature detail may be lost, and one can only examine aspects involving low frequency signals versus high frequency noise. The main thing we have examined here is the cumulative distribution of power. It appears from examination of power spectra for many images that one can distinguish pure noise from images with structure, but it is difficult to differentiate among a range of dissimilar images. Figure 1 shows bivariate power functions derived from FFTs of a number of images along with four types of noise. Notice how the cumulative distribution of power, normalized by volume and superimposed on the FFT, is nearly flat for noise, but strongly curved for most images. One can show that blurring tends to reduce the higher frequency components, and noise tends to increase power uniformly (Figure 2), but finer distinctions seem unlikely.

## 5 Simulations and Image Comparisons

### 5.1 Distance and Intensity

The first and second derivatives of images were examined with scatter plots of the bivariate distribution of the maximum gradient and Hessian at each point using `lgradient` from `scilaim` (Figure 3). The greatest difficultly of detecting added blur and noise was found with images with large ranges of $\dot{M}$ and $\ddot{M}$, such as the `mandrill`.

We examined graphical summaries for a number of images using $V_\lambda(t)$ and $T_\lambda(t)$. In general, we found for blurring, noise and deformation: (1) as blurring increases, the intensity difference increases. (2) noise tends to yield small distance and intensity differences, except for shot noise which has occasional large differences. (3) as deformation (global shift or local deformation) increases, the distance difference increases. At this point, we have mainly considered prefiltering images with gradients (edge and contrast detection) and detail filters (Laplacian and difference of Gaussians). Detail filters applied to blurred images primarily found $V_\lambda(t) \geq \lambda T_\lambda(t)$. Neither detail nor gradient filters were much modified by noise; distance and intensity values were both low.

Bivariate histograms of $\{(T_\lambda(t), V_\lambda(t)), t \in T\}$ were plotted on a logarithmic scale with a box for $(d = 6, V(d;\alpha))$ and a line $v = \lambda t$ show the relative tradeoff of distance and intensity. Figure 4 shows how intensity differences $V_\lambda(t)$ tend to increase as blurring increases, regardless of the image (here `girl, spheres` and `dart`). Figures 5 and 6 show the effect of noise and blur on several filters as summarized by the distance-by-intensity histograms.

Noise by itself tended to produce very short distance and intensity differences, as would be expected. Gaussian, uniform and Cauchy noise seemed to elicit similar responses; shot noise (a small percent of very large values) sometimes required very large distances to match intensities (not shown). The images used here and elsewhere in this paper are shown in Figure 7.

### 5.2 Estimating Blur and Noise by Regression

In some experiments estimates of $\sigma_f^2/2$ were found by regressing $y(t) - x(t)$ on $\Delta x(t)$; see equation (2.1). For the cases when the image was blurred with a Gaussian filter (of various sizes) a positive slope was obtained which was highly significant. For the case when independent Gauusian noise was added, a negative estimate of $\sigma_f^2$ was obtained, indicating no blur, since the tests are one-tailed. In the cases where both blur and noise were present, the success of the inference procedure was dependent on the relative amounts of error introduced by these two deformations.

There have been many investigations into the removal of blur via deconvolution, and we will not dwell on this subject. Rather, we concentrate on the issue of detecting a shift or local deformation in the presence of noise and in the presence of noise and blur. We focus primarily on Gaussian noise and shot noise, and on blurring due to a Gaussian window.

### 5.3 Estimating Shift and Local Deformation

Our experience is that global shifts are fairly easy to detect using the cascade algorithm, and are not affected by either noise or blur. It is only when one considers the possibility of local deformation that one detects any variation in estimation. Here we have found that the coefficient of variation (CV) for the ratio of deviations between the original and deconvolved vs. the original and reconstructed images,

$$\sum[y(t - \hat{s}(t)) - x(t)]^2 / \sum[y(t) - x(t)]^2,$$

is about 1% over a range of images subjected to Gaussian noise (Table 1) and about 3% for images with shot noise.

We also examined an image of a `mandrill` (primate) along with a deformation which makes the nose wider and moves it to the right, and shifts the eyes upwards. Our reconstructions recover the major feature deformations, but, depending on the actual algorithm, tend to obliterate some detail or other (not shown).

We are concerned that our cascade algorithm is not very precise, missing some features entirely and much rougher than we would expect. In fact, for some combinations of initial and final values of $\sigma_w$ and $\sigma_G$ in the cascade algorithm (see equation (3.2)) the deconvolved images are worse than the original! (Table 1a, `molecule`). This is partly due to the inherent inaccuracy imposed upon us by the current computing environment, `scilaim` using pixel $(0 - 255)$ and integer arithmetic where floating point arithmetic would be more appropriate. Our experience with the cascade algorithm suggests that an initial value of $\sigma_G$ should be 127 or 63 and that $\sigma_w$ should be set to $(\sigma_G + 1)/4 - 1$ or $(\sigma_G + 1)/8 - 1$. These combinations seem to work adequately for a variety of images.

## 6 References


1. Y. Amit, U. Grenander and M. Piccioni, Structural image restoration through deformable templates, Brown U., 1989.

2. A.J. Baddeley, Statistical image analysis, *STATCOMP 87, LaTrobe U., Melbourne, Australia, 14-15 May*, 1987.

3. A.J. Baddeley, A class of image metrics, *ANZAAS Congress, Townsville, Australia, 24-28 Aug.*, 1987.

4. J.E. Besag, On the statistical analysis of dirty pictures (with discussion), *J. R. Statist. Soc. B*, 48, 259-302, 1986.

5. D. Geman, S. Geman, C. Graffigne and P. Dong, Boundary detection by constrained optimization, *IEEE Trans. Pattern Anal. Mach. Intel.*, 12, 609-628, 1990.

6. Z.P. Jin and P. Mowforth, A discrete approach to signal matching, *TIRM-89-036, The Turing Institute, Glasgow*, 1989.

7. M.D. Jubb, Image reconstruction, *Ph.D. dissertation, Dept. of Statist., U. of Bath*, 1989.

8. R. Molina and B.D. Ripley, Using spatial models as priors in astronomical image analysis, *J. Appl. Statist.*, 16, 193-206, 1989.

9. T. Pavlidis, A critical survey of image analysis methods, *Proc. 8th Intl. Conf. Pattern Recog., Paris, Oct.*, 1986.

10. B.D. Ripley, Statistics, images and pattern recognition, *Can. J. Statist.*, 14, 83-111, 1986.

11. C.C. Taylor, Measure of similarity between two images, *Spatial Statistics and Imaging, IMS Lecture Notes*, 1988.

12. G. Wahba, *Spline Models for Observational Data*, SIAM, CBMS-NSF 59, 1990.