Measuring image similarity in the presence of noise

Gustavo K. Rohde[a], Carlos A. Berenstein[b,c], Dennis M. Healy Jr.[c]

[a]Applied Mathematics and Scientific Computation Program, University of Maryland, College Park, MD;
[b]Institute for Systems Research, University of Maryland, College Park, MD;
[c]Department of Mathematics, University of Maryland, College Park, MD.

ABSTRACT

Measuring the similarity between discretely sampled intensity values of different images as a function of geometric transformations is necessary for performing automatic image registration. Arbitrary spatial transformations require a continuous model for the intensity values of the discrete images. Because of computation cost most researchers choose to use low order basis functions, such as the linear hat function or low order B-splines, to model the discrete images. Using the theory of random processes we show that low order interpolators cause undesirable local optima artifacts in similarity measures based on the L_2 norm, linear correlation coefficient, and mutual information. We show how these artifacts can be significantly reduced, and at times completely eliminated, by using sinc approximating kernels.

Keywords: image registration, similarity measure, noise, interpolation, random process

1. INTRODUCTION

Image registration is the process of identifying the spatial correspondence between different images. Registration of medical images is an important procedure in many aspects of biomedical research and clinical practice where it is used to fuse information from images of a single subject taken at different times to account for subject motion, or geometric distortions. Image registration methods are also used by neuroscientists to relate images taken from different subjects. In the context of brain neuroanatomy for example, researchers (neuroscientists assisted by imaging engineers, scientific computation experts, mathematicians, and statisticians) have long used image registration methods to study the variation of biological tissue properties, such as shape and composition, described in images across a given population. The aim in such endeavors is usually to provide a quantitative description of a healthy 'normal' population, as opposed to a diseased one.

Though implementations vary, most methods seek to solve the digital image registration problem within an optimization framework where the goal is to find a function $f\colon \mathbf{x} \to \mathbf{x}'$, where $\mathbf{x}$ and $\mathbf{x}'$ are vectors in $\mathbb{R}^d$, that transforms the spatial coordinates $\mathbf{x}$ of a target image $T(\mathbf{x})$ to the spatial coordinates $\mathbf{x}'$ of a source image $S(\mathbf{x}')$. For the source and target images to be spatially aligned the mapping function f should be chosen in such a way as to optimize some cost function (objective function) between the two images. Mathematically, the image registration problem can be stated as a minimization problem:

$$\min_f \Theta(S(f(\mathbf{x})), T(\mathbf{x}), f) \tag{1}$$

where $\Theta(\cdot,\cdot,\cdot)$ represents the objective function being optimized. Note that f is an explicit argument into Θ since (1) is often a constrained optimization problem. To prevent registration methods from producing transformations that violate the intrinsic topology of the images the objective function Θ is usually computed as a sum of two terms.

$$\Theta(S(f(\mathbf{x})), T(\mathbf{x}), f) = I(S(f(\mathbf{x})), T(\mathbf{x})) + C(f) \tag{2}$$

The term $C(f)$ is a 'regularization' term meant to prevent wild spatial oscillations in the transformation f. Popular choices for regularization of registration methods include $C(f) = \int |Df(\mathbf{x})|^2 d\mathbf{x}$ and $C(f) = \int 1/|\det(J(f))| d\mathbf{x}$,

Medical Imaging 2005: Image Processing, edited by J. Michael Fitzpatrick, Joseph M. Reinhardt, Proc. of SPIE Vol. 5747 (SPIE, Bellingham, WA, 2005)
1605-7422/05/$15 · doi: 10.1117/12.594964

where $\det(J(f))$ stands for the determinant of the Jacobian matrix of f and D stands for a differential operator usually chosen based on arguments from continuum mechanics.

The function I represents some distance measure, or equivalently, the negative of a similarity measure between the images being registered. The choice of I is usually determined by the requirements of the application. In instances when images S and T are expected to be nearly identical under optimal alignment, L_2-norms or the sum of squared differences are often employed. When the relationship between the intensity values of S and T is unknown but expected to be linear I can be based on the correlation coefficient of their intensity values. In more complicated situations many researchers have shown the Mutual Information[1] similarity measure to be a good choice for I.

Before an estimate for the similarity measure $I(\cdot,\cdot)$ can be computed for an arbitrary $f(\mathbf{x})$, however, a suitable strategy for computing $S(f(\mathbf{x}))$ from a set of discrete samples $S(\mathbf{i})$, $\mathbf{i} \in \mathbb{Z}^d$ is necessary. Almost unanimously researchers choose to make an image $S(\mathbf{i})$ continuous by modeling it as a linear combination of symmetric basis functions[2–4] determined by interpolation or approximation from the discrete data $S(\mathbf{i})$. In recent years, several researchers have reported that popular methods used to estimate the Mutual Information similarity measure can behave unexpectedly with respect to the spatial transformation f being applied.[5,6] Such artifacts have been attributed to the interpolation or approximation strategies being used. In this work we show that such interpolation artifacts also occur in L_2 and correlation-based similarity measures, in addition to mutual information. We show that such artifacts stem from the fact that in estimating the value of the similarity measure for an arbitrary spatial transformation f one is obliged to interpolate or approximate noisy data, thus inevitably reducing the frequency content of the image.

The remaining is organized as follows. Using the theory of random processes we first describe the effects of spatial transformations on the variance and covariance structure of the interpolated image. We explain that noise covariance distortions can have undesired effects on measures of image similarity, such as the L_2 norm and correlation coefficient. We show that such effects can be eliminated by using sinc approximating kernels. Moreover, we show that low order interpolators also cause the aforementioned 'grid' artefacts when measuring image similarity using mutual information and that sinc approximating kernels can also be used to improve the registration curves of mutual information-based similarity measures. Finally, we test our solutions using both simulated and real magnetic resonance imaging (MRI) data.

2. THEORY

2.1. Covariance properties of interpolated signals

Following the approach described in[7] we use the following linear, stochastic, image model in our analysis:

$$S(\mathbf{x}) = \int W(\mathbf{p})\Upsilon(\mathbf{x}-\mathbf{p})d\mathbf{p} + e(\mathbf{x}) \tag{3}$$

where $\mathbf{x}$ and $\mathbf{p} \in \mathbb{R}^d$, $W(\mathbf{p})$ corresponds to the function describing the object being imaged, and $\Upsilon(\mathbf{x})$ is the point spread function of the imaging system. Note that throughout this paper all quantities will be assumed to be real valued. Unless noted otherwise, all integrals shall be evaluated from $-\infty$ to ∞. $e(\mathbf{x})$ refers to a zero mean stochastic process whose covariance structure will soon be described. We shall consider the integral part of equation (3) to be deterministic. The covariance function of the random process (3) can be shown to be:

$$\begin{aligned} R_S(\mathbf{x}_1,\mathbf{x}_2) &= \mathrm{Cov}\{S(\mathbf{x}_1), S(\mathbf{x}_2)\} \\ &= \mathrm{E}\{(S(\mathbf{x}_1) - \overline{S}(\mathbf{x}_1))(S(\mathbf{x}_2) - \overline{S}(\mathbf{x}_2))\} \\ &= \mathrm{E}\{e(\mathbf{x}_1)e(\mathbf{x}_2)\} \\ &= \mathrm{Cov}\{e(\mathbf{x}_1), e(\mathbf{x}_2)\} = R_e(\mathbf{x}_1,\mathbf{x}_2) \end{aligned}$$

where

$$\mathrm{E}\{e(\mathbf{x})\} = \overline{e}(\mathbf{x}) = \int e(\mathbf{x})\mathrm{pr}[e(\mathbf{x})]\, de(\mathbf{x}), \tag{4}$$

and pr$[e(\mathbf{x})]$ stands for the probability density function of the quantity $e(\mathbf{x})$. Thus, without loss of generality, for the purposes of analyzing the covariance of interpolated signals we momentarily assume that $S(\mathbf{x})$ is a zero-mean random process (i.e. $S(\mathbf{x}) = e(\mathbf{x})$ through subtraction of the deterministic part of (3)).

As stated above, before a given similarity measure $I(S(f(\mathbf{x})), T(\mathbf{x}), f)$ can be evaluated for an arbitrary spatial transformation f a continuous model for the digital image $S(\mathbf{i})$ is needed. Most often, researchers choose to model $S(\mathbf{i})$ as a linear combination of symmetric basis functions $h(\mathbf{x})$:

$$S^c(\mathbf{x}) = \sum_{\mathbf{i}\in\mathbb{Z}^d} S(\mathbf{i})h(\mathbf{x}-\mathbf{i}). \tag{5}$$

In this case we have chosen the coefficients of the linear combination to be the sampled image values. Note that the summations above are carried from $-\infty$ to $+\infty$ by extending the sampled signal $S(\mathbf{i})$ with periodic or mirror boundary conditions. When sampled signals are extended using boundary conditions their covariance structure $R_S(\mathbf{x}_1, \mathbf{x}_2)$ must be extended the same way. If we would like $S^c(\mathbf{i}) = S(\mathbf{i})$, the basis function $h(\mathbf{x})$ must obey the following: $h(\mathbf{i}) = 0 \;\; \forall\, \mathbf{i} \in \mathbb{Z}^d \neq 0$ and $h(0) = 1$. Also note that when $d > 1$, the interpolation kernel is generally taken to be separable:

$$\hat{h}(\mathbf{x}) = \prod_{j=1}^{d} h(x_j). \tag{6}$$

As shown by Aldroubi et al.[8] expression (5) can be interpreted as a continuous filtering operation of the initial sampled image values

$$S^c(\mathbf{x}) = \int S_\delta(\mathbf{q})h(\mathbf{x}-\mathbf{q})d\mathbf{q} = (h * S_\delta)(\mathbf{x}), \tag{7}$$

where the notation a_δ represents the tempered distribution consisting of the train of weighted Dirac delta impulses $a_\delta(\mathbf{x}) = \sum_{\mathbf{i}=-\infty}^{+\infty} a(\mathbf{i})\delta(\mathbf{x}-\mathbf{i})$.

Using (7) one can show that the covariance structure of the continuous function (5) $R_{S^c}(\mathbf{x}+\Delta\mathbf{x}, \mathbf{x}) = \mathrm{Cov}\{S^c(\mathbf{x}+\Delta\mathbf{x}), S^c(\mathbf{x})\}$ is:

$$R_{S^c}(\mathbf{x}+\Delta\mathbf{x}, \mathbf{x}) = \int\int h(\mathbf{x}+\Delta\mathbf{x}-\mathbf{q}_1)R_{S_\delta}(\mathbf{q}_1, \mathbf{q}_2)h(\mathbf{x}-\mathbf{q}_2)d\mathbf{q}_1 d\mathbf{q}_2, \tag{8}$$

where $R_{S_\delta}(\mathbf{q}_1, \mathbf{q}_2) = \mathrm{E}\{S_\delta(\mathbf{q}_1)S_\delta(\mathbf{q}_2)\}$. If we assume that the correlations in the random process $S_\delta(\mathbf{x})$ are of such short range that $R_{S_\delta}(\mathbf{q}_1, \mathbf{q}_2)$ can be approximated by $z_\delta(\mathbf{q}_1)\delta(\mathbf{q}_1 - \mathbf{q}_2)$, where $z_\delta(\mathbf{x}) = \mathrm{Var}\{S_\delta(\mathbf{x})\}$, then

$$\begin{aligned} R_{S^c}(\mathbf{x}+\Delta\mathbf{x}, \mathbf{x}) &= \int z_\delta(\mathbf{q}_1)h(\mathbf{x}+\Delta\mathbf{x}-\mathbf{q}_1)h(\mathbf{x}-\mathbf{q}_1)d\mathbf{q}_1 \\ &= \sum_{\mathbf{i}=-\infty}^{+\infty} z(\mathbf{i})h(\mathbf{x}+\Delta\mathbf{x}-\mathbf{i})h(\mathbf{x}-\mathbf{i}). \end{aligned}$$

Moreover, if we are dealing with a constant variance random process $z_\delta(\mathbf{x}) = \sigma^2$, the variance of the interpolated signal $S^c(\mathbf{x})$ is given by:

$$\mathrm{Var}\{S^c(\mathbf{x})\} = R_{S^c}(\mathbf{x}, \mathbf{x}) = \sigma^2 \sum_{\mathbf{i}=-\infty}^{+\infty} [h(\mathbf{x}-\mathbf{i})]^2. \tag{9}$$

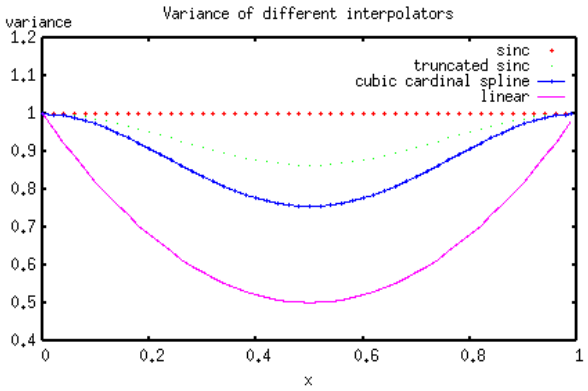

Figure 1. Plot of equation (9), with $\sigma^2 = 1$, for several interpolating basis functions typically used in medical imaging.

The purpose of the exercise above is to characterize the effects of the signal interpolation model (5) on the covariance structure of the signal. To illustrate this effect we have plotted equation (9), with $\sigma^2 = 1$, for several interpolating basis functions currently used in medical imaging in figure 1. See appendix A for the definition of the interpolators used. The truncated sinc basis function was computed using $W = 6$. As shown in figure 1, given a constant variance 'white' discrete random process, the continuous model expressed in (5) produces a function whose variance at each point in space is no longer uniform. This is especially true for low order interpolators such as the linear 'hat' function, but in general it is also true for any kind of interpolating basis function other than sinc. To understand why, it is enough to verify that $\sum_{\mathbf{i}=-\infty}^{+\infty}[\mathrm{sinc}(\mathbf{x}-\mathbf{i})]^2 = 1 \; \forall \; \mathbf{x}$. Thus the covariance structure of a digital image that undergoes a geometric transformation via

$$S^c(f(\mathbf{x})) = \sum_{\mathbf{i}\in\mathbb{Z}^d} S(\mathbf{i})h(f(\mathbf{x})-\mathbf{i}) \tag{10}$$

is dependent on the function f. The covariance structure of an interpolated image can be computed by substituting $\mathbf{x}$ for $f(\mathbf{x})$ in equation (7).

To illustrate this concept, we have computed the following simulation. A series of 200 digital images was created using a random number generator such that the signal of each pixel was zero and the variance one. The sample variance (in each pixel) is displayed on the left panel of figure (2). As expected, this image is fairly uniform depicting a constant-variance random process. Next, each of the 200 simulated images was rotated about its center by 4 degrees counter clock-wise using bilinear interpolation. The sample variance of the newly created series of rotated images was computed for each pixel and is displayed in the center panel of figure (2). As can be expected, the variance becomes non-uniform as a function of image coordinate and acquires a certain 'striped' configuration. Lastly, the original series of 200 images of random noise was again rotated about its center counter clock-wise by 4 degrees. This time, however, truncated sinc interpolation (with $W = 30$, see appendix A) was used. The sample variance of the rotated image series for each pixel is displayed on the right panel of figure (2). As can be expected, the variance of the rotated image series using truncated sinc interpolation is almost perfectly uniform.

2.2. Optimization of L_2-based similarity measures

We have shown above how the covariance properties of digital images are modified according to the spatial transformation f being used during registration. We now show that this dependence can be detrimental to the image registration process when commonly used similarity measures such as the L_2 norm and the linear correlation coefficient are used. In order to facilitate analysis we use the following vector notation for the source and target digital images. That is, $\mathbf{S} = \{S_0, S_1, \cdots, S_N\}^T$, with $N = m \times n$ where m, n are the dimensions of the two-dimensional image. Using (3) we can write $\mathbf{S} = \tilde{\mathbf{W}}_S + \mathbf{e}_S$, with $\tilde{\mathbf{W}}_S$ representing the deterministic part of equation (3) and $\mathbf{e}_S$ representing the noise vector. Similarly $\mathbf{T} = \tilde{\mathbf{W}}_T + \mathbf{e}_T$.

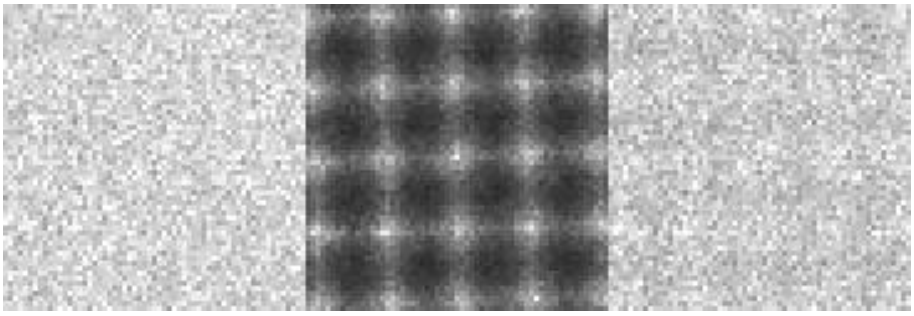

Figure 2. Variance of a constant variance random process before and after rotation using different interpolators. The panel on the left shows the variance of the random process prior to any transformation. The middle panel show the variance of the random process after rotation of the images about their center using bilinear interpolation. The panel on the left shows the variance of the random process after the same rotation of the images this time computed using sinc approximating basis functions.

In this paper we will use rigid body spatial transformations defined by $f_{\theta,\mathbf{t}}(\mathbf{x}) = \mathbf{R}_\theta(\mathbf{x}-\mathbf{c})+\mathbf{c}+\mathbf{t}$, where $\mathbf{t}\in\mathbb{R}^d$ represents a translation vector, $\mathbf{c}$ represents the center coordinate of the image, and

$$\mathbf{R}_\theta = \begin{vmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{vmatrix}, \tag{11}$$

represents a rotation matrix, in the two dimensional case. For shorthand notation we write $\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}$ to mean the operation of applying a spatial transformation $f_{\theta,\mathbf{t}}$ to digital image $S(\mathbf{i})$ through (10). Note that since (10) is a linear operation, $\mathbf{F}_{\theta,\mathbf{t}}$ is a linear operator, though in general it is not shift invariant.

We look at the discrete L_2-norm similarity measure, defined by $\|\mathbf{a}\|^2 = \langle \mathbf{a},\mathbf{a}\rangle = \frac{1}{N}\sum_{i=1}^{N} a_i^2$, of the difference between the source and target images being registered as a function of the transformation parameters:

$$I(\theta,\mathbf{t}) = \|\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}-\mathbf{T}\|^2 = \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}-\mathbf{T}, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}-\mathbf{T}\rangle \tag{12}$$

Expanding all terms we write:

$$I(\theta,\mathbf{t}) = Q_1(\theta,\mathbf{t}) + Q_2(\theta,\mathbf{t}), \tag{13}$$

where

$$Q_1(\theta,\mathbf{t}) = \langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S\rangle + 2\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\rangle - 2\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{T}\rangle - 2\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{T}\rangle + \langle \mathbf{T},\mathbf{T}\rangle \tag{14}$$

and

$$Q_2(\theta,\mathbf{t}) = \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\rangle. \tag{15}$$

To illustrate the behavior of $Q_1(\theta,\mathbf{t})$ and $Q_2(\theta,\mathbf{t})$ in the presence of noise we have computed the following simulation. Normally distributed spatially uncorrelated noise was added to a digitally manufactured 'phantom' image (shown in the left panel of figure 3) such that the signal to noise ratio, defined to be mean signal divided by the standard deviation, was 5. Two such images were generated with identical signal and different noise vectors, thus simulating the situation of registering a source image to a target image. $Q_1(\theta,\mathbf{t})$ and $Q_2(\theta,\mathbf{t})$ were then computed with respect to translation along the x direction (figure 3, right panel) using bilinear interpolation. Here optimal alignment is defined to be at zero pixel translation. Evidently $Q_1(\theta,\mathbf{t})$ is able to determine the transformation parameters that optimally align the images when noise is present. $Q_2(\theta,\mathbf{t})$ on the other hand varies with respect to the spatial transformation being applied in such a way that it is independent of the alignment of the signal part of the images. As a consequence, L_2-based similarity measures

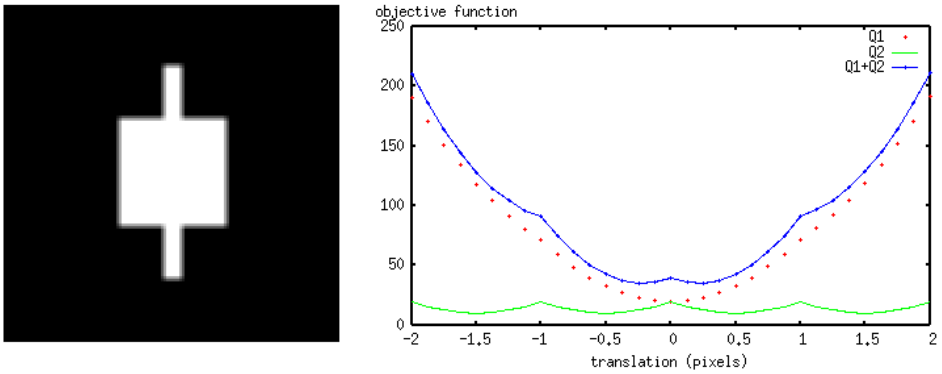

Figure 3. Left panel: phantom image (without noise) used in computing the registration curves shown on the right panel.

$I(\theta,\mathbf{t}) = ||\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}-\mathbf{T}||^2 = Q_1(\theta,\mathbf{t}) + Q_2(\theta,\mathbf{t})$ are not able to determine the transformation parameters that optimally align the images. Thus a computer program that registers images based on the minimization of equation $||\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}-\mathbf{T}||^2$ would not be able to produce accurate results. Looking at $Q_2(\theta,\mathbf{t})$ more closely, we see that, when $\mathbf{e}_S$ is normally distributed for example, $\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\rangle$ is the maximum likelihood estimator for the expectation $\mathrm{E}\{(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S)^2\}$. It is easy to see that $\mathrm{E}\{(\mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S)^2\} \sim \mathrm{Var}\{S^c(f(\mathbf{x}))\}$, where $\mathrm{Var}\{S^c(f(\mathbf{x}))\}$ is given by equation (9). The oscillations in $Q_2(\theta,\mathbf{t})$ shown in figure 3 are thus equal (up to a scaling factor) to the oscillations shown in the linear portion of figure 1. Note also that the oscillations shown in figure 3 are very similar to the 'grid' effects discussed in by Pluim and Tsao et al..[5,6]

2.3. Optimization of correlation-based similarity measures

In this section we look at the effects of system noise on correlation based similarity measures. For convenience, we look at the cross correlation (zero-mean correlation coefficient) between two N-dimensional vectors often used in image registration $I(\theta,\mathbf{t}) = \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}, \mathbf{T}\rangle / (||\mathbf{T}||||\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}||)$. As done earlier, we use the linear stochastic image model $\mathbf{S} = \tilde{\mathbf{W}}_S + \mathbf{e}_S$ to expand the term $||\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}||$ into

$$||\mathbf{F}_{\theta,\mathbf{t}}\mathbf{S}|| = \sqrt{\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S\rangle + 2\langle \mathbf{F}_{\theta,\mathbf{t}}\tilde{\mathbf{W}}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\rangle + \langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\rangle} \tag{16}$$

where again we have the undesirable, though inevitable, term $\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\rangle$ which depends solely on the noise properties of the source image. The same translation simulations done above reveal that correlation based cost functions suffer from the same problems as L_2 based ones: namely, the optimal value of the cost function will not reveal the transformation parameters that optimally align the images. The results of the simulation are omitted here for brevity.

2.4. Optimization of mutual information

The 'grid' effects observed in mutual information-based optimization are slightly different in nature from the effects demonstrated for the L_2 and correlation objective functions in the sense that noise variance is not the only factor to be considered. In this case, we have also to consider the intensity value distribution of the object, in addition to the distribution of noise. Image registration via maximization of mutual information relies on measuring the statistical dependence in the co-occurrence of intensity values of images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$. Let $\mathrm{pr}_{S_f}[s]$ and $\mathrm{pr}_T[t]$ represent the probability of occurence of a pixel with intensity values s and t in images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$, respectively. Their joint probatility is denoted $\mathrm{pr}_{S_f,T}[s,t]$. With these quantities defined the mutual information between images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$ is given by:

$$I\left(S(f(\mathbf{x})), T(\mathbf{x})\right) = \int\int \mathrm{pr}_{S_f,T}[s,t] \log\left(\frac{\mathrm{pr}_{S_f,T}[s,t]}{\mathrm{pr}_{S_f}[s]\mathrm{pr}_T[t]}\right) dsdt = H(T) + H(S_f) - H(T,S_f), \tag{17}$$

where $H(T) = \mathrm{E}\{\log(\mathrm{pr}_T[t])\}$, $H(S_f) = \mathrm{E}\{\log\left(\mathrm{pr}_{S_f}[s]\right)\}$, and $H(S_f, T) = \mathrm{E}\{\log\left(\mathrm{pr}_{S_f,T}[s,t]\right)\}$. If, for example, the distribution of intensity values of images $S(f(\mathbf{x}))$ and $T(\mathbf{x})$ can be described as a Gaussian probability density function, the mutual information between these two images can be computed analytically[1] by $I(S(f(\mathbf{x})), T(\mathbf{x})) = -\frac{1}{2}\log(1-\rho^2)$ where ρ is the linear correlation coefficient between the intensity values of the images being registered. In this situation we would expect that any reasonably accurate estimate of (17) as a function of spatial transformation f would also contain the artifacts in the correlation coefficient similarity measure demonstrated earlier. Naturally, tomographic images are seldom globally Gaussian distributed. Locally, however, the normal distribution assumption may be more realistic. Nonetheless, we also expect that the 'grid' artifacts in the mutual information-based registration curves should be reduced when higher order sinc-approximating kernels are used instead of low order interpolators such as the hat function. Arguments that will hold under more general assumptions can be used to explain this phenomena.[9] Similar simulation experiments as decribed above were conducted to confirm that 'grid' artifacts are also present in mutual information-based objective functions. For brevity, however, these are also not shown here.

3. METHODS

The theory presented above suggests that the interpolation artifacts seen in L_2 and correlation-based cost functions are entirely due to the presence of noise in the images being registered. One obvious strategy to mitigate such artifacts is to reduce the variance $\mathrm{Var}\{S(\mathbf{i})\}$ in the original source image by performing local averaging. This can be achieved via digital convolution of the image with a digital filter $G(\mathbf{i})$: $\tilde{S}(\mathbf{i}) = \sum_{w\in\mathbb{Z}^d} S(\mathbf{w})G(\mathbf{i}-\mathbf{w})$. If we are concerned with a stationary random process, $\mathrm{Var}\{S(\mathbf{i})\} = \sigma^2 \;\forall\; \mathbf{i}$, using

$$G(\mathbf{x}) = \frac{1}{9}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \tag{18}$$

for example, would reduce the variance of the image to $\mathrm{Var}\{\tilde{S}(\mathbf{i})\} = 0.11\sigma^2$. This would reduce the oscillatory behavior of the cost function due to the term $\langle \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S, \mathbf{F}_{\theta,\mathbf{t}}\mathbf{e}_S\rangle$ in computing L_2 and correlation based cost functions. This may not always be an optimal strategy for excessive smoothing can blur image boundaries that are important for guiding the registration process. As suggested earlier, an alternative approach is to use higher order sinc approximating kernels instead of low order kernels such as the hat function.

We will compare these interpolation and approximation methods for measuring image similarity in real magnetic resonance images. The images used here were taken from a standard single-shot spin-echo echo planar imaging (EPI) sequence acquisition on a 1.5 T GE Signa system. Because the images were reconstructed by taking the magnitude of the Fourier transform of the time domain signals, they do not necessarily constitute a stationary random process. In fact through the work described in[10,11] it is known that if Gaussian distributed noise is added to the receiver coils, the variance of the magnitude reconstructed image in regions of zero signal is $(2-\pi/2)\sigma^2$ while the variance of the magnitude reconstructed image in regions with relatively high signal approaches σ^2. The variance in each of these domains however, is usually assumed to be constant. In addition, because of several linear filtering steps performed during analog to digital conversion, to remove 'ringing' artifacts for example, the magnitude reconstructed image is usually spatially correlated. However, these correlations are usually small.

As done in by Pluim and Tsao et al.,[5,6] the registration curves were computed by rotating and translating one image with respect to the other, much like in the previously discussed simulations. However, following the approach described by Unser et al.,[12] in order to amplify the interpolation artifacts we choose to vary the spatial transformation at a slightly misaligned state. That is, angle θ is varied using a fixed translation vector $\mathbf{t} = \{0, 2\}^T$, and translation in the x direction is computed using a fixed translation in the y direction $\mathbf{t} = \{t, 2\}^T$. Note that in this case the exact alignment parameters are unknown since we are using real data. As seen in figure 4, however, the images we use seem to be fairly well aligned.

In the experiments relating to the mutual information, the joint probability density function of the images was estimated using both the joint histogram and the Parzen window techniques. For brevity, however, we only

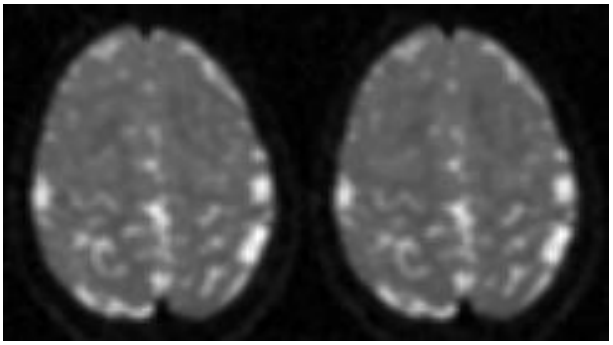

Figure 4. Real T_2 weighted echo planar images used for computing the performance of different interpolation methods for image registration. The images were acquired in rapid succession and are likely to be fairly well aligned.

report the results obtained with PArzen window-based estimates of the joint probability density functions of the images.

In addition, in order to show that the effects described above are general and not limited to affine or 'global' transformations we have also conducted tests using localized radial basis functions. In this example we used the following parameterization for the spatial transformation:

$$f_{\mathbf{k}}(\mathbf{x}) = \mathbf{x} + \mathbf{k}\Phi(\frac{||\mathbf{x}-\mathbf{q}||}{r}), \tag{19}$$

where $\mathbf{k} = \{k_x, k_y\}^T$ are the basis functions coefficients, $\mathbf{q}$ defines the center of the basis function, r its radius, and:

$$\Phi(x) = \frac{1}{4}(1-x)_+^4(4+16x+12x^2+3x^3). \tag{20}$$

As in the example described earlier, we have plotted the variation of the cost function with respect to the k_x coefficient, using a fixed $k_y = 2$ coefficient.

4. RESULTS

Figure 5 shows the sum of squared differences similarity measure of the real images displayed in figure 4 with respect to translation and rotation using the bilinear and truncated sinc interpolation methods. The registration curves computed using bilinear interpolation present what is commonly refereed to as the griding artifact while such artifacts are seemingly inexistent in the registration curve computed using the truncated sinc interpolation method. Truncated sinc interpolation was computed using $W = 6$.

The mutual information similarity measure as a function of image translation is shown in figure 6. The 'grid' effects are evident when linear interpolation is used to produce the geometrically transformed image. The 'grid' effects become nearly undetectable when sinc approximating kernels are used. The width of the sinc approximating kernel used here was $W = 50$.

The results of the nonrigid experiments are displayed in figure 7. Here truncated sinc interpolation was computed using $W = 6$. The location of the basis function is indicated by the bright circle in the image shown in the left panel. The radius of the basis function was chosen to be $r = 20$ pixels. The curve computed using bilinear interpolation again shows multiple local optima values. In contrast, the curve computed using truncated sinc interpolation shows only one. In addition, the global optima for the cost function computed using truncated sinc basis functions and bilinear ones differ.

Lastly, the qualitatively similar results for the experiments above were computed using the cross correlation similarity measure. These, however, are omitted here for brevity.

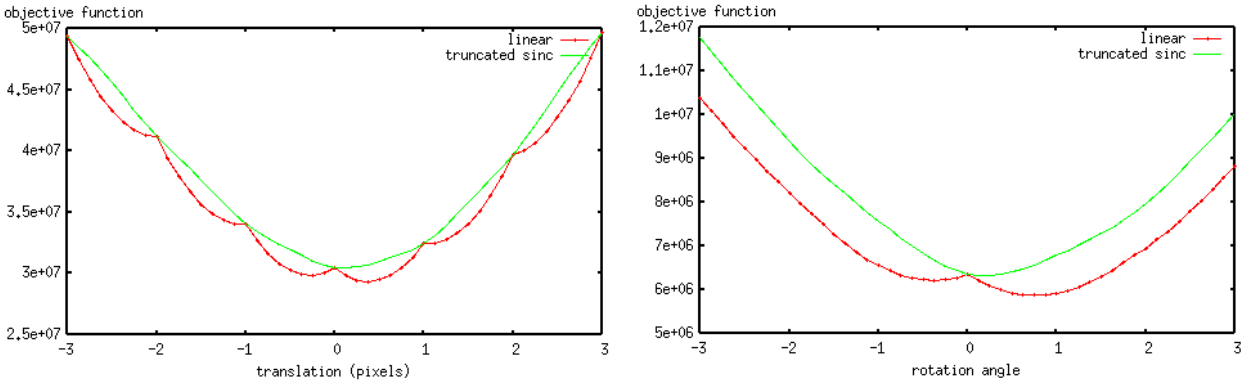

Figure 5. Registration curves computed from real data. Left: sum of squared differences similarity measure w.r.t. translation. Right: s.s.d. similarity w.r.t. rotation angle in degrees.

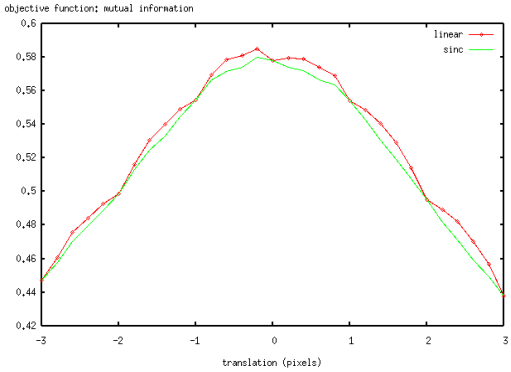

Figure 6. Mutual information similarity measure as a function of translation. The dotted line was computed using linear interpolation while the solid line was computed using truncated sinc interpolation.

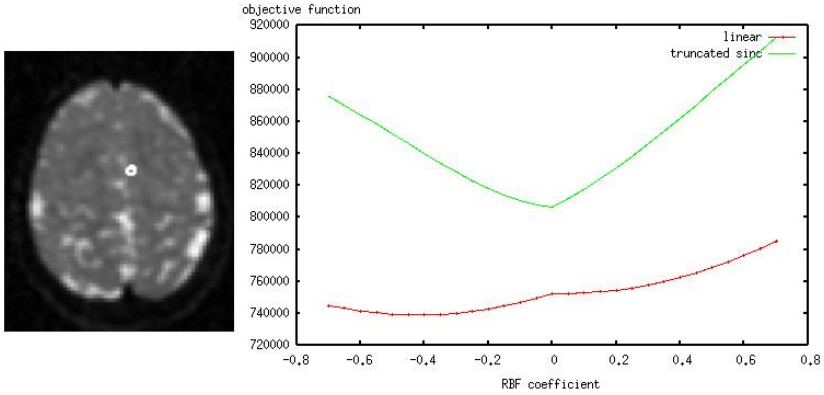

Figure 7. Left: white dot indicates placement of the basis function for computing the registration curves shown on the right panel.

5. DISCUSSION

The experiments with real MR images discussed above depict the effects of low order interpolation methods on popular registration cost functions such as the sum of squared differences, linear correlation coefficient, and mutual information. Though the optimal alignment parameters for the real data experiments were not known, in all cases tried, sinc approximating basis functions performed visibly better than linear ones. Linear interpolation often produced registration curves that contained the aforementioned 'grid' (multiple local optima) artifacts. Thus registration programs based on optimization of L_2, correlation, or mutual information based cost functions using linear interpolation are likely to converge slowly to the global optima solution, or in some cases, converge only to a local optima, significantly degrading the quality of the results. Image interpolation using sinc approximating basis functions, on the other hand, did not produce the artifacts mentioned.

Local, parametric, models for deformation fields are often used in nonrigid registration methods.[13–16] The results produced by using the local nonrigid transformation model (19) show that the interpolation artifacts in registration curves are general, and not necessarily restricted to affine or global transformation models. Results also showed that sinc approximating kernels can help increase the accuracy of such registration algorithms.

We attribute the increase in performance gained by sinc approximating kernels to the fact that their use in computing geometric image transformations causes the least amount of changes in the covariance properties of the image being interpolated. This was particularly evident when we compared sinc approximating basis functions with linear basis functions in our simulations. We have experimented with different appodisation windows for truncating sinc basis functions and found that the support W of the windowing function, and not necessarily the form of $v(x)$, was most significant in determining the covariance properties of interpolated images.[9] As can be expected, the 'grid' artifacts experienced in L_2 norm and correlation based similarity measures are dependent on the amount of noise present in the images. These effects may be negligible in images with very high signal to noise ratios.

6. SUMMARY AND CONCLUSIONS

Image similarity measures based on the L_2 norm, linear correlation and mutual information have been widely employed for rigid-body, affine, and elastic registration of medical and other images. We have shown that such similarity measures contain systematic local optima artifacts when low order interpolators such as the hat function are used for computing geometric transformations. We showed that such interpolation artifacts are entirely caused by the effects of the spatial transformations being used as well as the basis functions used to make the images continuous.

Using a linear stochastic model for the image data we showed that in addition to being functions of the spatial transformations being applied, L_2 and correlation based similarity measures are also functions of the covariance structure of the interpolated images. The covariance structure of geometrically transformed images is described as follows. Because of inevitable system noise stemming from thermal and other effects, the images being registered are viewed as random processes. Since image interpolation, at each point in space, can be interpreted as a linear filtering procedure we used the second order theory of random processes to describe the effects of different interpolating basis functions on the covariance structure of the images. In our experience, the 'grid' effects on mutual information registration curves are not only caused by changes in the covariance structure of the image as it relates to noise but also due to the loss of high-frequency information when images are interpolated using low order interpolators.[9]

We showed that sinc approximating basis functions are optimal for generating continuous approximations of image data in registration problems since they are least likely to cause significant changes in the covariance structure of interpolated images. Thus, in addition to the more traditional properties based on linear approximation theory described in,[3] we argue that the criterion $\sum_{\mathbf{i}=-\infty}^{+\infty}[h(\mathbf{x}-\mathbf{i})]^2 = 1 \; \forall \; \mathbf{x}$ should be kept in mind when choosing basis functions for computing spatial transformations of images in registration problems.

Future work in this area could include investigating further which basis functions best preserve the covariance properties of the images during image registration while keeping in mind their computational cost. In addition, a similar approach could be used to better characterize the interpolation artifacts already observed with mutual information type cost functions. Lastly, we note that continuous approximations of discretely sampled image

data are not only necessary in solving the image registration problem but also in other areas such as image segmentation using deformable models. Future work could also include demonstrating the effects of low order interpolators in these applications.

Appendix A

Here we give the definition of the several interpolating basis functions used throughout this paper. The linear basis function, also known as the B-spline of order 1 or hat function, is given by:

$$h(x) = \begin{cases} 1 - |x| & \text{if } |x| \leq 1; \\ 0 & \text{if } |x| > 1. \end{cases} \tag{21}$$

The cubic cardinal spline function refereed to above is given by:

$$h(x) = \frac{-6\alpha}{(1-\alpha^2)} \sum_{k=-\infty}^{+\infty} \alpha^{|k|} \beta^3(x-k), \tag{22}$$

where $\alpha = 0.2679$, and $\beta^3(x)$ is the popular B-spline of order 3:

$$\beta^3(x) = \begin{cases} \frac{2}{3} - \frac{1}{2}|x|^2(2-|x|) & , 0 \leq |x| < 1; \\ \frac{1}{6}(2-|x|)^3 & , 1 \leq |x| < 2\ ; \\ 0 & , 2 \leq |x|\ . \end{cases} \tag{23}$$

Finally, the popular sinc basis function is given by:

$$h(x) = \mathrm{sinc}(x) = \frac{\sin(\pi x)}{\pi x}. \tag{24}$$

The truncated version of (24) is given by multiplication of (24) with a window function $w(x)$:

$$w(x) = \begin{cases} v(x) & \text{if } |x| < W; \\ 0 & \text{if } W \leq |x|. \end{cases} \tag{25}$$

In this work we use the Hann function (sometimes referred to as Hanning) defined as:

$$v(x) = 0.5 + 0.5\cos(\frac{\pi x}{W}). \tag{26}$$

ACKNOWLEDGMENTS

We would like to thank Dr. Carlo Pierpaoli of the Section on Tissue Biophysics and Biomimetics, Laboratory of Integrative and Medical Biophysics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA for providing some of the images used in this study.

Berenstein's work has been partly supported by NSF grant DMS0400698.

REFERENCES

1. S. Kullback, *Information Theory and Statistics*, Dover Publications, 1968.
2. T. Lehmann, C. Gonner, and K. Spitzer, "Survey: Interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging* **18**, pp. 1049–1075, 1999.
3. P. Thévenaz, T. Blu, and M. Unser, "Image interpolation and resampling," in *Handbook of Medical Imaging, Processing and Analysis*, I. Bankman, ed., ch. 25, pp. 393–420, Academic Press, San Diego CA, USA, 2000.
4. E. Meijering, W. Niessen, and M.A.Viergever, "Quantitative evaluation of convolution-based methods for medical image interpolation," *Medical Image Analysis* **5**, pp. 111–126, 2001.

5. J. Pluim, J. Maintz, and M. Viergever, "Interpolation artefacts in mutual information-based image registration," *Computer Vision and Image Understanding* **77**, pp. 211–232, 2000.
6. J. Tsao, "Interpolation artifacts in multimodality image registration based on maximization of mutual information," *IEEE Transactions on Medical Imaging* **22**, pp. 854–964, 2003.
7. K. Astrom and A. Heyden, "Stochastic analysis of image acquisition, interpolation and scale-space smoothing," *Adv. in Appl. Probab.* **31**, pp. 855–894, 1999.
8. A. Aldroubi, M. Unser, and M. Eden, "Cardinal spline filters: Stability and convergence to the ideal sinc interpolator," *Signal Processing* **28**, pp. 127–138, 1992.
9. G. K. Rohde, "Ph.D. thesis work. Applied Mathematics and Scientific Computation program, University of Maryland," 2005.
10. R. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," *Medical Physics* **13**, pp. 232–233, 1985.
11. H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magnetic Resonance in Medicine* **34**, pp. 910–914, 1995.
12. M. Unser and P. Thévenaz, "Stochastic sampling for computing the mutual information of two images," in *Proceedings of the Fifth International Workshop on Sampling Theory and Applications (SampTA'03)*, pp. 102–109, (Strobl, Austria), May 26-30, 2003.
13. D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging* **18**, pp. 712–721, 1999.
14. J. Kybic, P. Thévenaz, A. Nirkko, and M. Unser, "Unwarping of unidirectionally distorted EPI images," *IEEE Transactions on Medical Imaging* **19**, pp. 80–93, February 2000.
15. J. Kybic and M. Unser, "Fast parametric elastic image registration," *IEEE Transactions on Image Processing* **12**, pp. 1427–1442, November 2003.
16. G. K. Rohde, A. Aldroubi, and B. M. Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration," *IEEE Transactions on Medical Imaging* **22**, pp. 1470–1479, 2003.