

Interpolation Artifacts in Sub-Pixel Image Registration

Gustavo K. Rohde, *Member, IEEE*, Akram Aldroubi, and Dennis M Healy, Jr

Abstract

We consider the problem of registering (aligning) two images to sub-pixel accuracy by optimization of objective functions constructed from the images' intensity values. We show that some widely used interpolation methods can introduce multiple local optima in the energy of the interpolated image which, if not counter-balanced by other terms, can cause local optima in registration objective functions including the sum of squared differences, cross correlation, and mutual information. We discuss different solutions to address the problem based on high degree B-spline interpolation, low pass filtering the images, and stochastic integration. Numerical examples using synthetic and real signals and images are shown.

Index Terms

Image registration; interpolation; artifacts; mutual information

I. INTRODUCTION

Image registration is the process of computing the spatial correspondence between two or more images. Over the past few decades image registration methods have enabled the extraction of quantitative spatial information from a series of digital images in a variety of settings in the technological and experimental

G.K. Rohde is with the Center for Bioimage Informatics, Biomedical Engineering Department, Carnegie Mellon University, Pittsburgh, PA, 15213 USA. Email: gustavor@cmu.edu. Phone: 412-268-3684. Fax: 412-268-9580.

A. Aldroubi is with the Department of Mathematics, Vanderbilt University, Nashville, TN, 37240 USA. Email: akram.aldroubi@vanderbilt.edu. Phone: 615 322-6656. Fax: 615 343 0215

D.M. Healy, Jr., is with the Department of Mathematics, University of Maryland, College Park, MD, USA. Email: dhealy@math.umd.edu. Phone: 301-405-5047. Fax: 301 314-6710

A portion of this work was previously presented at IEEE ISBI 2007. This work was supported in part by Office of Naval Research Grant ONR-N00014-06-1-0002, and NSF DMS grant 0504788.

sciences. They have been used extensively in remote sensing, computer vision, and medical imaging applications. See [1], [2], [3] for comprehensive reviews.

There exist several different methods for image registration currently under research. In a broad sense, these can be split into manual, semi-automatic, and automatic methods. Within the automatic class, significant research efforts have focused on those which seek to compute the best alignment between two images by optimizing some objective function quantifying the similarity between corresponding intensity values in the two images. Research in the area, however, has revealed that standard approaches to interpolation-based sub-pixel image registration can be sub-optimal due to the presence of so called ‘interpolation artifacts.’ As opposed to efforts towards analyzing the effects of noise on the performance in registration algorithms (see Robinson and Milanfar [4] for an example), researchers have noted that certain image interpolation algorithms cause systematic, artifactual, oscillatory behavior of many objective functions. Ashburner and Friston [5] reported such oscillations in the sum of squared differences (SSD) cost function while Pluim *et al.* reported these artifacts in the mutual information (MI) similarity measure [6]. Naturally, artifactual oscillations induce local optima in the optimization procedure. Local optima in automatic intensity-based registration methods are not desirable because 1) they may prevent the search procedure from arriving at a global optimum and 2) even if the neighborhood of a global optimum can be determined, the artifactual oscillations may prevent sub-pixel accuracy in the alignment procedure.

Several recent examples in the image registration literature present observations about interpolation artifact problems together with ad-hoc techniques for reducing their impact. However, a general mathematical explanation for such artifacts is yet to come forth that would provide quantitative understanding and analyses of the proposed remediation. For example, Pluim *et al.* [6] observed that the artifacts in the mutual information similarity measure are more pronounced when the pixel sizes in both images are equal and proposed a modified sampling scheme to diminish the impact of these artifacts in registration. Their proposal is one of a number of approaches (see also [7], [8]) in which the effects of artifactual oscillations are reduced by computing the image similarity measure from samples taken from spatial coordinates which do not align with the grid of the images. Another suggestion has been blurring (low-pass filtering) the images in order to reduce the magnitude of the artifacts in the sum of squared difference cost function [5]. These suggestions and others found in the literature are effective in varying degrees, but to our knowledge there has been no careful mathematical analysis of the registration artifact phenomenon to help us understand and validate the various proposals for reducing its impact.

In [9], [10] we proposed a unified explanation for the phenomenon and related the oscillatory behavior of different objective functions such as the mutual information (MI), cross correlation (CC), and sum

of squared differences (SSD) based on the energy of the interpolated signal. Our explanation, however, was based on stochastic arguments in the sense that only effects of additive noise were considered. In this work we expand our previous work and offer a more complete analysis of the phenomena, showing that interpolation on noise is not the only cause of the artifacts. We note, however, that our analysis in this paper is limited to showing that common interpolation schemes using the B-spline framework cause oscillatory artifacts in the energy of interpolated signals. We also show that the energy of the interpolated signal is present in the CC, SSD, and MI similarity measures discussed below. Thus, if left unbalanced by oscillations in other terms (the analysis of which is beyond the scope of this paper), these oscillations could cause local optima artifacts in the similarity measure.

We start by introducing the common methodology of image registration using B-spline interpolation as the basis for our arguments. We then proceed to analyze the oscillatory phenomena of the mutual information and related similarity measures. Algorithms for eliminating the artifactual local optima are presented. Finally, results using simulated as well as real image data are presented followed by discussion and conclusions.

II. NOTATION AND PRELIMINARIES

Let L_2 define the standard space of measurable, square-integrable real valued functions $g(x)$. The L_2 norm is derived from the inner product:

$$\langle g, h \rangle_{L_2} = \int_{-\infty}^{\infty} g(x)h^*(x)dx, \quad (1)$$

$$\|g\|_{L_2}^2 = \langle g, g \rangle_{L_2}. \quad (2)$$

The space of square summable discrete signals $b[k]$, $k \in \mathbb{Z}$, is denoted ℓ_2 . The inner product in ℓ_2 is defined by:

$$\langle a, b \rangle_{\ell_2} = \sum_{k \in \mathbb{Z}} a[k]b^*[k] \quad (3)$$

while $\|b\|_{\ell_2}^2 = \langle b, b \rangle_{\ell_2}$. The convolution between two discrete signals $a, b \in \ell_2$ is denoted $a * b$ and given by:

$$a * b[k] = \sum_{j \in \mathbb{Z}} a[j]b[k - j]. \quad (4)$$

The sequence b can be viewed as a discrete convolution operator characterized by its transfer function $B(z) = \sum_{k \in \mathbb{Z}} b[k]z^{-k}$. If B has no zeros on the unit circle, and $b \in \ell_1$ (all of the filters considered in this paper satisfy these requirements), then the inverse operator $(b)^{-1}$ exists and is uniquely defined by:

$$(b)^{-1} \longleftrightarrow 1/B. \quad (5)$$

A. Basic methodology for image registration

We consider the problem of aligning digital images, considered as discrete signals on a d -dimensional sampling lattice: $s[\mathbf{k}]$, $\mathbf{k} \in \mathbb{Z}^d \subset \mathbb{R}^d$. The goal in image registration is to compute the coordinate transformation $f_\theta : \mathbb{R}^d \rightarrow \mathbb{R}^d$ (here f is parameterized by θ) that will bring the two images into spatial alignment. For example, f_θ could be translation or shifting of a function by a fixed offset value θ . Loosely speaking, registration methods try to compute the spatial transformation f which makes the intensity values of two images s and t similar:

$$\tilde{s}(f_\theta(m, n)) \sim t(m, n), \quad (6)$$

where in this case we work with the familiar $d = 2$ images, and the tilde over s indicates that, in general, it is necessary to obtain the values for s on the transformed discrete lattice by some sort of interpolation.

To make sense of such a concept one must quantify the notion of similarity between the intensity values $t(m, n)$ and the transformed image values $\tilde{s}(f_\theta(m, n))$. We focus on those methods which seek the best choice of alignment f_θ by optimizing some objective function (maximizing some similarity measure) between these two sets of intensities with respect to the parameter θ .

This raises two immediate implementation issues. One is the choice of the objective function for assessing similarity between t and the transformed version of s . The second is the mechanism for evaluation of these transformed values $\tilde{s}(f_\theta(m, n))$, as $f_\theta(m, n)$ need not lie on the original sampling lattice \mathbb{Z}^2 . Consider for instance our example of translation by θ with θ a fraction of a pixel, say $\theta = 1/2$. In that case the transformed grid points are completely disjoint from the original grid where our known sample values live. In the following subsections we briefly review some relevant aspects of the standard interpolation approach for spatially transforming an image and then take up some key properties of the standard choices of objective functions.

With these implementation issues in hand, we can give a general problem statement of the registration task: given a pair of digital images $s(m, n)$, $t(m, n)$, $m, n \in \mathbb{Z}$, find the function f_θ , or alternatively, the parameter θ , so that some functional $\Psi(\theta)$ of the digital image data is optimized. The functional $\Psi(\theta)$ usually contains at least two terms:

$$\Psi(\theta) = \Psi_{data}(\theta) + \Psi_{constraint}(\theta). \quad (7)$$

The first term, known as the image force, uses the image similarity objective function mentioned above and represents the influence of the image data over f in the minimization, while $\Psi_{constraint}$ restricts the solution space of allowable transformations so as to make the problem well posed. While both Ψ_{data}

and $\Psi_{constraint}$ are important in determining the solution f in this work we concentrate on the data term Ψ_{data} exclusively. Our first goal for the paper is to show that this term may contain artifactual fluctuations creating false local optima when standard interpolations strategies are used. This explains certain commonly observed difficulties encountered in sub-pixel registration applications. Our second objective is to use our new understanding of this problem to inform the creation of novel registration methods providing improved performance by minimizing these difficulties.

B. Cardinal B-spline interpolation for registration

As discussed above, sub-pixel accurate image registration requires one to be able to compute the value of a sequence s in ‘between’ sample values $k \in \mathbb{Z}$. To that end we consider the widely used B-spline-based interpolation.

We look at the B-spline family of basis functions, where $\beta^0(x)$ is the centered, normalized, rectangular pulse evaluated at x and $\beta^n(x) = \beta^{n-1} * \beta^0(x)$, while n refers to the degree of the basis function. The corresponding function space is:

$$V(\beta^n) = \left\{ \tilde{s}(x) = \sum_{k \in \mathbb{Z}} c[k] \beta^n(x - k), c \in \ell_2 \right\}. \quad (8)$$

Given a sampled signal $\{s[k]\}_{k \in \mathbb{Z}}$ we can find its interpolant in this space in the classical fashion by solving a linear system to find the coefficients $\{c[k]\}_{k \in \mathbb{Z}}$. Alternatively, the continuous model can also be efficiently obtained through inverse filtering [11], [12] by:

$$\tilde{s}(x) = \sum_{k \in \mathbb{Z}} s[k] \eta^n(x - k). \quad (9)$$

$\eta^n(x)$ is known as the cardinal B-spline interpolator and is given by:

$$\eta^n(x) = \sum_{k \in \mathbb{Z}} (b^n)^{-1}[k] \beta^n(x - k), \quad (10)$$

where $b^n[k] = \beta^n(x)|_{x=k}$. The basis functions η^n are interpolating since

$$\eta^n(k) = \begin{cases} 1 & \text{if } k = 0; \\ 0 & \text{for all other } k \in \mathbb{Z}. \end{cases} \quad (11)$$

Cardinal B-splines of different degrees are displayed in figure 1.

Extension to multiple dimensions is possible through the use of tensor product splines simply obtained by cross multiplication of the one-dimensional functions of the individual index variables:

$$\tilde{s}(x, y) = \sum_{p, q \in \mathbb{Z}} s[p, q] \eta^n(x - p) \eta^n(y - q). \quad (12)$$

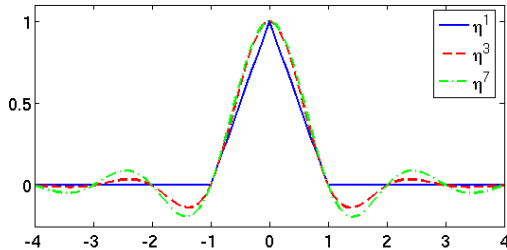


Fig. 1. Cardinal B-splines of degree 1, 3, and 7.

C. Objective functions for registration

Consider two discrete real-valued, one dimensional signals $\{s[k]\}_{k \in \mathbb{Z}}$, $\{t[k]\}_{k \in \mathbb{Z}}$. One natural metric for computing the spatial transformation $f_\theta : \mathbb{R} \rightarrow \mathbb{R}$ would be the ℓ_2 norm of the difference between $\{\tilde{s}(f_\theta(k))\}_{k \in \mathbb{Z}}$ and $\{t[k]\}_{k \in \mathbb{Z}}$:

$$\begin{aligned} \Psi_{data}(\theta) = SSD(\theta) &= \sum_{k \in \mathbb{Z}} (\tilde{s}(f_\theta(k)) - t[k])^2 \\ &= \|t\|_{\ell_2}^2 + \|\tilde{s}(f_\theta)\|_{\ell_2}^2 - 2 \langle t, \tilde{s}(f_\theta) \rangle_{\ell_2} \end{aligned} \quad (13)$$

where $\tilde{s}(f_\theta)$ is short notation for the discrete signal generated by $\{\tilde{s}(f_\theta(k))\}_{k \in \mathbb{Z}}$. Note that in order to register two discrete signals one would minimize the SSD between them. Alternatively, one may register two discrete signals by maximizing their cross correlation:

$$\Psi_{data}(\theta) = CC(\theta) = \frac{\langle \tilde{s}(f_\theta), t \rangle_{\ell_2}}{\|t\|_{\ell_2} \|\tilde{s}(f_\theta)\|_{\ell_2}}. \quad (14)$$

Alternatively, one may also consider a slightly different normalized error described in [13] that has the same optimum as equation (14). In either case, our initial intuition may be that, were we dealing with signal translations for example, the inner product $\langle \tilde{s}(f_\theta), t \rangle_{\ell_2}$ is the only term which varies with f_θ and therefore is the only contributor to the optimization of alignment. The other components are the norms of t and $s(f_\theta)$. When considering signals of infinite length, and thus neglecting considerations due to changes in the regions of overlap, $\|t\|$ is clearly independent of θ and, at first glance, it also appears that the same should hold for $\|\tilde{s}(f_\theta)\|$ for the most commonly used alignments f_θ . Indeed, if we picture registration as aligning two interpolated surfaces with heights given by the respective intensity levels, then aligning with rigid motions f_θ , (rotations and translations) acting on the image surface s , certainly does nothing to change its energy or L_2 norm.

In fact we are misled by this continuum picture: we shall see that even rigid motions induce significant variations in $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ when low degree interpolation is used to evaluate it, and that this does cause oscillatory artifacts in sub-pixel registration.

We shall show a similar phenomenon is obtained in the case of mutual information similarity (MI) measure, the other widely used objective function candidate, despite its significantly different form from correlation and SSD. Let S^θ and T represent random variables associated with the discrete signals $\{\tilde{s}(f_\theta(k))\}_{k \in \mathbb{Z}}$ and $\{t[k]\}_{k \in \mathbb{Z}}$, respectively. The mutual information objective function is given by

$$\begin{aligned} \Psi_{data}(\theta) &= MI(S^\theta, T) \\ &= \int \int p_{S^\theta, T}(\mu, \nu) \log \left(\frac{p_{S^\theta, T}(\mu, \nu)}{p_{S^\theta}(\mu)p_T(\nu)} \right) d\mu d\nu \end{aligned} \quad (15)$$

where integrals are taken from $-\infty$ to ∞ . $p_{S^\theta, T}(\mu, \nu)$ stands for the joint probability density function (pdf) of the intensity values of images S^θ and T , and $p_{S^\theta}(\mu)$ and $p_T(\nu)$ stand for their respective marginal distributions. An alternative, entirely equivalent, formula is given by:

$$MI(S^\theta, T) = H(T) + H(S^\theta) - H(T, S^\theta), \quad (16)$$

where $H(T)$ stands for the differential entropy of random variable T , and $H(T, S^\theta)$ (their respective definitions are given in appendix A) the joint differential entropy between random variables T and S^θ . Estimates of the mutual information similarity measure are usually computed from image pixel values through histograms [14] or Parzen windowing techniques [15]. This similarity measure is based on advanced statistics and information theory [16] and their introduction as a similarity measure for image registration problems is due to Collignon *et al.* [17] and Viola and Wells III [15]. For a review of mutual information-based image registration methods in medical imaging see [18].

Without loss of generality we assume that the mean of each discrete signal is zero, while the standard deviation of the pixel values is one. Let S^θ now represent the random variable associated with a single signal value $\tilde{s}(f_\theta(k))$. If (strict) stationarity is assumed, S^θ can be used to represent the random variable for all samples in the same signal. Similar statements can be made with regards to image (signal) $t[k]$. Let $p_{S^\theta, T}(\mu, \nu)$ in (15) be given by a bivariate Gaussian distribution. Then it can be shown that the mutual information is given by [16]:

$$I(T, S^\theta) = -\frac{1}{2} \log(1 - \rho^2(\theta)), \quad \rho(\theta) = \xi(\theta) / \sigma_T \sigma_{S^\theta} \quad (17)$$

where $\xi(\theta)$ stands for the covariance between random variables S^θ and T , and σ_T^2 and $\sigma_{S^\theta}^2$ represent their respective variances. These quantities can be estimated from a set of samples $\tilde{s}(f_\theta(k))$ and $t[k]$,

$k \in \Omega$:

$$\xi(\theta) = \frac{1}{N} \sum_{k \in \Omega} \tilde{s}(f_\theta(k))t[k], \quad (18)$$

$$\sigma_T^2 = \frac{1}{N} \sum_{k \in \Omega} |t[k]|^2, \quad \text{and} \quad (19)$$

$$\sigma_{S^\theta}^2 = \frac{1}{N} \sum_{k \in \Omega} |\tilde{s}(f_\theta(k))|^2, \quad (20)$$

where Ω represents a set of predefined coordinates and $N = |\Omega|$. Clearly, $\sigma_{S^\theta}^2$ is related to $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ via:

$$\sigma_{S^\theta}^2 = \frac{1}{N} \|\tilde{s}(f_\theta)\|_{\ell_2}^2 - \frac{1}{N} \sum_{k \notin \Omega} |\tilde{s}(f_\theta(k))|^2. \quad (21)$$

Furthermore, $\sigma_{S^\theta}^2 = \frac{1}{N} \|\tilde{s}(f_\theta)\|_{\ell_2}^2$ if $\tilde{s}(f_\theta)$ has compact support in Ω .

Finally, when the distribution of the intensity values of the images $p_{S^\theta}(\mu)$ and $p_T(v)$ (and their joint distribution) are not Gaussian, equation (17) does not hold. However, as shown in appendix A, the mutual information can be expanded as:

$$MI(\theta) = I_G(\theta) + \mathcal{R}(\theta) \quad (22)$$

where $I_G(\theta)$ is the Gaussian approximation to the mutual information given in (17). Therefore, in a general sense, $-\frac{1}{2} \log(1 - \rho^2(\theta))$ can be interpreted as an approximation to the mutual information based on the second order (variance and covariance) statistics.

Note that in all cases (SSD, CC, and MI), estimates of the objective function involve the quantity $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$. In the next section we show that estimates of $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ computed using B-spline interpolation are not independent of the parameter(s) θ . Therefore, if these oscillations are left unbalanced by other terms, local optima are likely to be introduced in the objective function being optimized as shown below. However, we note that, although some computational examples concerning the term $\langle \tilde{s}(f_\theta), t \rangle_{\ell_2}$ are shown, the mathematical analysis to follow is not complete in that it neglects the several other terms that are present in the three different similarity measures we discuss (e.g. $\langle \tilde{s}(f_\theta), t \rangle_{\ell_2}$, $\mathcal{R}(\theta)$, $\Psi_{constraint}(\theta)$). In the case of the MI similarity measure, for example, it has been noted previously that it may suffer from artifacts stemming from changes in regions of overlap [19], as well as the number of histogram bins used in the computation [7]. We do not make an attempt to explain such sources of artifacts and a complete analysis of all components of all of the similarity measures discussed above is beyond the scope of this study. Rather, below we concentrate on explaining how oscillations in $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ can contribute to systematic oscillation artifacts in the similarity measures discussed above.

III. OSCILLATION ARTIFACTS

In this section we describe the effects of B-spline interpolation on the objective functions used in image registration described above. For the moment, let $f_\theta(x)$ represent a pure translation in one dimension:

$$f_\theta(x) = x + \theta. \quad (23)$$

Using the B-spline interpolation model above, a translated signal $\{\tilde{s}(k + \theta)\}_{k \in \mathbb{Z}}$ can be computed by:

$$\tilde{s}(k + \theta) = \eta_\theta^n * s[k] \quad (24)$$

where we have used the notation $\eta_\theta^n[k] = \eta^n(k + \theta)$. Note that all three objective functions discussed depend on the quantity $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$:

$$\|\tilde{s}(f_\theta)\|_{\ell_2}^2 = \langle \tilde{s}(f_\theta), \tilde{s}(f_\theta) \rangle_{\ell_2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{s}(\omega)|^2 |\hat{\eta}_\theta^n(\omega)|^2 d\omega \quad (25)$$

where

$$\hat{s}(\omega) = \sum_{k=-\infty}^{\infty} s[k] e^{-i\omega k}. \quad (26)$$

is the discrete time Fourier transform (DTFT) of sampled signal s . The DTFT of the shifted, sampled, cardinal B-spline interpolator is given by:

$$\begin{aligned} \hat{\eta}_\theta^n(\omega) &= \sum_{k=-\infty}^{\infty} \eta^n(k + \theta) e^{-i\omega k} \\ &= \sum_{k \in \mathbb{Z}} \frac{(\phi(\omega - 2\pi k))^{n+1}}{B^n(\omega - 2\pi k)} e^{-j(\omega - 2\pi k)\theta} \end{aligned} \quad (27)$$

where

$$\phi(\omega) = \frac{\sin(\omega/2)}{\omega/2}. \quad (28)$$

and $B^n(\omega)$ is the DTFT of the sampled B-spline of degree n , $b_n[k] = \beta^n(x)|_{x=k}$ which was described in section II. Clearly, because of the interpolation condition (11), we have that $\hat{\eta}_0^n(\omega) = 1$. It is also not hard to check that if $\theta \in \mathbb{Z}$, $|\hat{\eta}_\theta^n(\omega)| = 1$. This implies that for whole pixel uniform translations, the energy of the translated signal as expressed in equation (25) remains constant. This is not the case for sub-pixel translations. Figure 2 contains the frequency response $|\hat{\eta}_\theta^n(\omega)|$ for different translation values θ with $n = 3$ (part a), and different degrees at $\theta = 0.5$ (part b). It is clear that the energy of the translated signal using B-spline interpolation is not constant with respect to the translation value θ . More specifically, signal translation using low degree (1 or 3, for example) interpolation involves low-pass filtering the data, which, depending on the frequency content of the original signal, will affect (diminish)

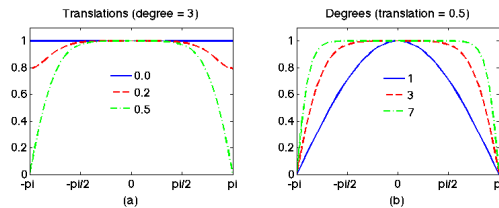


Fig. 2. Frequency response for cardinal B-spline translation filters for different translation values (a) and degrees (b).

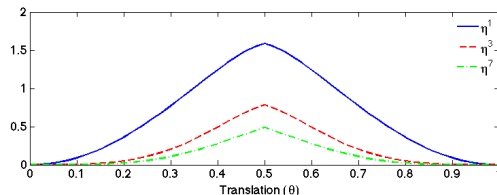


Fig. 3. Plot of $\int_{-\pi}^{\pi} (1 - |\hat{\eta}_{\theta}^n(\omega)|)^2 d\omega$ for different degrees n and different translation values θ .

its ℓ_2 norm. The overall effect will be a periodic (with period one sample or pixel) oscillation in the signal's ℓ_2 norm.

In order to avoid oscillation artifacts due to variation in the signal's ℓ_2 norm it would be necessary that $\hat{\eta}_{\theta}^n$ be as close as possible to the 'all pass' filter, for any θ . Thus one useful figure of merit for evaluating the oscillatory behavior that B-spline interpolators may cause in registration objective functions is

$$\int_{-\pi}^{\pi} (1 - |\hat{\eta}_{\theta}^n(\omega)|)^2 d\omega. \quad (29)$$

Using equation (27) this quantity was evaluated numerically for different translation values θ and different degrees n . Results are shown in Figure 3. As Figure 3 shows, higher degree interpolation can be used to lessen the oscillatory behavior of the objective function. This can be intuitively understood by observing that cardinal B-spline interpolators converge to the ideal 'sinc' interpolator as $n \rightarrow \infty$ (see Theorem 1 in [20]). The frequency response of the ideal sinc interpolator is an all pass filter and thus

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} (1 - |\hat{\eta}_{\theta}^n(\omega)|)^2 d\omega = 0, \quad \forall \theta. \quad (30)$$

IV. STRATEGIES FOR AVOIDING LOCAL OPTIMA

As noted above, the effects of translating a sequence $\tilde{s}(k)$ by θ on the ℓ_2 norm (energy) of the sequence will depend both on the frequency content of the original sequence as well as the degree of

the interpolating cardinal B-spline. Translation of a data sequence, when implemented by resampling a low-order interpolation of the data, tends to attenuate any high frequency component present in the original sequence. If no high frequency components are present in the original data, translation via low degree interpolation may not have a strong effect on its ℓ_2 norm. If, however, high frequency components are present in the original data, other alternatives must be sought. Clearly, ideal interpolation using the sinc kernel ($\text{sinc}_\theta(x) = \sin(\pi(x + \theta))/(\pi(x + \theta))$) would address the issue since it has a perfectly flat frequency response for any value of θ . Sub-pixel accurate registration methods using sinc-type (Dirichlet) interpolation kernels can be implemented efficiently through the use of the Fast Fourier Transform (FFT) in the phase correlation method [21] and more recent extensions [22], [23]. However, such methods are often limited to pure translations and similar solutions for registration problems using more varied spatial transformation (affine, nonrigid) are not available. For these applications one must use truncated versions of the kernel which can be overwhelmingly computationally intensive, especially for nonlinear optimization problems such as nonrigid registration.

Two other techniques can be used to avoid oscillation artifacts. The first technique would be to choose the degree n of the interpolating spline so that $|\hat{\eta}_\theta^n(\omega)| \sim 1, \forall \theta$, over the frequency content of the original data. High degree B-spline interpolation, however, is also computationally expensive. This is due to the fact that the higher the degree of the B-spline, the wider its support. Thus, by increasing the degree of the B-spline being used to interpolate the data, computing each $\tilde{s}(f_\theta(k))$ involves more additions and multiplications. Given a fixed degree n chosen perhaps based on computational complexity considerations, another alternative is to low-pass the sequence of data to be interpolated so that its frequency content falls entirely within regions where $|\hat{\eta}_\theta^n(\omega)| \sim 1, \forall \theta$. Low-pass filtering the data prior to registration would reduce the magnitude of the oscillation artifacts. However it could also throw away high frequency information which could be used in the matching process. Therefore an iterative procedure such as the one described in [24] whereby low resolution (coarse) versions of the images can be used in the first few iterations to find an approximate solution is attractive for avoiding such artifacts. The resolution can then be increased, and optimization reinitialized using the starting point obtained from the previous iteration.

A 'rule of thumb' for avoiding interpolation artifacts in the SSD objective function, for example, can be devised by looking at the requirement that a local optima exist. Taking the first derivative of the SSD objective function with respect to translation parameter θ :

$$\frac{d\|\tilde{s}(f_\theta)\|^2}{d\theta} = 2 \frac{d\langle t, \tilde{s}(f_\theta) \rangle_{\ell_2}}{d\theta} \quad (31)$$

We now consider an idealized experiment where two identical signals are matched ($t = s$). In the

neighborhood $\theta = 0$ in order to avoid a local optima we must have that

$$2 \frac{d \langle t, \tilde{s}(f_\theta) \rangle_{\ell_2}}{d\theta} > \frac{d \|\tilde{s}(f_\theta)\|^2}{d\theta}. \quad (32)$$

Naturally, the direction of the inequality above will depend on whether θ is positive or negative (which in turn depends on our definition of f_θ). Using (24), the requirement in equation (32) can be stated as

$$\langle A_t, \eta_\theta^{n'} \rangle_{\ell_2} > \langle \eta_\theta^n * A_t, \eta_\theta^{n'} \rangle_{\ell_2} \quad \forall \theta < 0 \quad (33)$$

where

$$\eta_\theta^{n'}[k] = \left. \frac{d\eta^n(x)}{dx} \right|_{x=k+\theta} \quad (34)$$

and $A_t[k]$, $k \in \mathbb{Z}$ is the autocorrelation sequence for t defined as

$$A_t[k] = \sum_{j \in \mathbb{Z}} t[j]t[j-k]. \quad (35)$$

We note that the inequality above can be represented in Fourier domain, providing an alternative interpretation. Using the same assumptions as above we have that

$$\left\langle |\hat{t}|^2, \frac{d\hat{\eta}_\theta^n}{d\theta} (\hat{\eta}_\theta^{n*} - 1) \right\rangle_{L_2, [-\pi, \pi]} > 0. \quad (36)$$

Thus $|\hat{t}|^2$, through the $L_2[-\pi, \pi]$ inner product, defines a positive cone for $\frac{d\hat{\eta}_\theta^n}{d\theta} (\hat{\eta}_\theta^{n*} - 1)$. If these two functions are orthogonal for some choice of θ near $\theta = 0$, then a local optimum occurs.

A similar requirement for the CC objective function can be stated as

$$\langle A_t, \eta_\theta^{n'} \rangle_{\ell_2} > \frac{\langle A_t, \eta_\theta^n \rangle_{\ell_2}}{\langle \eta_\theta^n * A_t, \eta_\theta^n \rangle_{\ell_2}} \langle \eta_\theta^n * A_t, \eta_\theta^{n'} \rangle_{\ell_2}. \quad (37)$$

By identifying $\rho(\theta)$ in equation (17) with $CC(\theta)$ in (14), a similar bound can be derived for the gaussian portion ($I_G(\theta)$) of the MI similarity measure. However, in this case the bound must be thought of as an approximation since $I_G(\theta)$ is only a second order (Gaussian) approximation for $MI(\theta)$.

Equations (33) and (37) can be used to estimate the degree of the B-spline interpolation method that would be necessary for avoiding oscillation artifacts when computing the registration objective function of a signal and a shifted version of the signal, for some predetermined range of θ . Given an image or signal s which is to be translated to match another similar signal, both sides of the inequality (33), for example, can be plotted as a function of θ , for increasing values of n . Specific values for n for which (33) is satisfied (for all θ of interest) can then be used to perform the matching without local optima artifacts.

Finally, we note that since as n tends to infinity, the frequency response of the translation filter η_θ^n tends to the frequency response of the ideal (sinc) interpolator [20], we have that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{d\|\tilde{s}(f_\theta)\|^2}{d\theta} &= \lim_{n \rightarrow \infty} \langle \eta_\theta^n * A_t, \eta_\theta^{n'} \rangle_{\ell_2} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |\hat{t}(\omega)|^2 i\omega d\omega = 0. \end{aligned} \quad (38)$$

Computational examples showing how low-pass filtering the data prior to interpolation, as well as higher degree interpolation, can be used to avoid oscillation artifacts will be shown later. First, however, we describe a technique (stochastic integration) which allows one to use low degree interpolation, even with signals with high frequency components, without significant interference from the oscillation artifacts discussed earlier.

A. Stochastic integration

Here we show that the energy of a signal interpolated using random sampling methods [8] is, in theory, independent of the translation being applied. Expanding $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ we have

$$\|\tilde{s}(f_\theta)\|_{\ell_2}^2 = \sum_{p \in \mathbb{Z}} \sum_{q \in \mathbb{Z}} s[p]s[q] \sum_{k \in \mathbb{Z}} \eta^n(k + \theta - p)\eta^n(k + \theta - q). \quad (39)$$

Clearly, the source of the problem is the fact that $\sum_{k \in \mathbb{Z}} \eta^n(k + \theta - p)\eta^n(k + \theta - q)$ is not independent of θ . Approximating this term with its continuous version would avoid the oscillation artifacts aforementioned since

$$\int_{-\infty}^{\infty} \eta^n(x + \theta - p)\eta^n(x + \theta - q)dx = \int_{-\infty}^{\infty} \eta^n(x - p)\eta^n(x - q)dx. \quad (40)$$

Replacing discrete sums in the SSD and CC objective functions with continuous integrals, however, would significantly add to the computational complexity of all but the most trivial image registration problems. Moreover, in the case of mutual information computations with histograms or Parzen windows, such continuous formulation would be unfeasible.

We now show that Monte Carlo methods can be used to compute approximations of the continuous integral above which do not oscillate as a function of θ , for any given n . Monte Carlo methods are a class of computational algorithms that have found applications in several areas of exact sciences, including integration of multidimensional functions. Suppose we have a bounded function φ and we wish to find out its integral over Ω . The basic result of Monte Carlo integration is that this integral can be estimated by first generating a set of pseudo-random numbers a_k , $k = 1, \dots, M$ uniformly distributed in Ω and then using

$$\frac{|\Omega|}{M} \sum_{k=1}^M \varphi(a_k). \quad (41)$$

In the sense of expectations, the sum above approximates the value of the integral since:

$$\begin{aligned}
\frac{|\Omega|}{M} \sum_{k=1}^M \varphi(a_k) &\approx |\Omega| \mathbb{E}\{\varphi(a)\} \\
&= |\Omega| \int_{\Omega} \varphi(a) p_a(a) da \\
&= \int_{\Omega} \varphi(a) da
\end{aligned} \tag{42}$$

where $p_a(a) = 1/|\Omega|$. The standard deviation of the random variable in (41) is $O(M^{-1/2})$. Therefore the error in the value of the integral computed using this method can be expected to be of this order of magnitude.

Now let $\varphi(a) = \eta^n(a-p)\eta^n(a-q)$. Because a is a random variable uniformly distributed in Ω we may use another random variable $a+\theta$ where θ is a deterministic constant and $a+\theta$ is a random variable uniformly distributed in Ω (ignoring edge effects). Clearly, ignoring edge effects, $\mathbb{E}\{\varphi(a)\} \approx \mathbb{E}\{\varphi(a+\theta)\}$. Therefore we have:

$$\frac{1}{M} \sum_{k=1}^M \varphi(a_k) \approx \frac{1}{M} \sum_{k=1}^M \varphi(a_k + \theta) \text{ and} \tag{43}$$

$$\sum_{k=1}^M \eta^n(a_k - p)\eta^n(a_k - q) \approx \sum_{k=1}^M \eta^n(a_k + \theta - p)\eta^n(a_k + \theta - q). \tag{44}$$

We conclude that if uniform distribution-based stochastic methods are used the oscillations described above do not occur. Thus we define the stochastic sum

$$\langle s, t \rangle_{MC} = \sum_{k \in \mathbb{Z}} \tilde{s}(a_k) \tilde{t}(a_k) \tag{45}$$

where $\tilde{s}(x)$ and $\tilde{t}(x)$ refer to the interpolated value defined in equation (9), and a_k refers to a uniformly distributed random coordinate. As before, we define $\|s\|_{MC}^2 = \langle s, s \rangle_{MC}$ and henceforth whenever we speak of a similarity measure computed using Monte Carlo integration, we replace $\langle s, t \rangle_{\ell_2}$ with $\langle s, t \rangle_{MC}$ where appropriate. Finally, note that all conclusions regarding stochastic integration translate without problems to dimensions 2 and higher, since we are using separable B-spline interpolation.

B. Constant Variance Interpolation

Recent studies have shown that statistical properties (e.g. variance) of images are not constant with respect to interpolation-based spatial transformations of image data and that this could have a deleterious

effect on standard operations such as image registration [9] as well as model fitting [25]. Subsequently researchers have begun tailoring basis functions to meet the constant variance requirement [26], [27] in an effort to avoid such artifacts. In this section we show that the constant variance requirement is necessary but not sufficient for avoiding the oscillations in the energy of the interpolated signal, as described above.

Let $\zeta(x)$ represent an interpolating basis function with which one may implement a pure translation by an arbitrary shift θ . Assuming a stationary, zero mean, random field, the constant variance requirement for translating the field can be expressed as $\sum_{k \in \mathbb{Z}} [\zeta(k + \theta)]^2 = 1$, for all θ (see [9] for a derivation and for plots of this quantity for several popular basis functions). Using Parseval's relation the constant variance requirement can be written as

$$\|\hat{\zeta}_\theta\|_{L_2[-\pi, \pi]}^2 = 2\pi \quad (46)$$

for all θ and where, once again $\hat{\zeta}_\theta(\omega) = \sum_{k=-\infty}^{\infty} \zeta(k + \theta)e^{-i\omega k}$. Clearly, basis functions that have a flat frequency response (that is $|\hat{\zeta}_\theta| = 1$) for all θ satisfy this requirement but so do many other basis functions (an example is shown below). As we have shown above, basis function that do not have a flat frequency response for all translation values θ run the risk of changing the value of $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ as a function of θ thus possibly introducing oscillatory artifacts in the objective function if these oscillations are not balanced by oscillations in other terms. Naturally, constant variance interpolation methods will be suitable (maintain the energy of the signal) for translating signals whose whose spectrum is flat (i.e. $|\hat{s}| = c$ with c some constant) as can be seen by inserting eq. (46) into (25). White noise satisfies this requirement, for example. However it is not customary for either natural or biomedical images to possess a perfectly flat frequency spectrum. Thus we conclude that constant variance interpolation, though necessary, is not sufficient to avoid interpolation artifacts in $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ for arbitrary signals (images) s . To demonstrate this, below we show a computational example involving the use of the following basis function:

$$\zeta(x) = \begin{cases} a_0(x) & , 0 \leq |x| \leq \frac{1}{2}; \\ 1 + 2c(a_1(x)) - a_1(x) & , \frac{1}{2} < |x| \leq 1; \\ -c(a_2(x)) & , 1 < |x| \leq \frac{3}{2}; \\ -c(a_3(x)) & , \frac{3}{2} < |x| \leq 2; \\ 0 & , 2 < |x|. \end{cases} \quad (47)$$

The definitions above are: $a_0(x) = 1 - \kappa x^2$, $a_1(x) = 1 - \kappa(1 - |x|)^2$, $a_2(x) = a_1(x)$, $a_3(x) = 1 - \kappa(2 - |x|)^2$, $\kappa = 3 - \frac{\sqrt{12}}{2}$, and

$$c(a) = \frac{-4(1-a) + \sqrt{12(1-a)^2 - 48(a^2 - a)}}{12}. \quad (48)$$

The basis function defined above can be shown to satisfy both the constant variance requirement as well as the partition of unity requirement $\sum_{k \in \mathbb{Z}} [\zeta(k + \theta)] = 1, \forall \theta \in \mathbb{R}$, it is continuous, has compact support in $[-2, 2]$, is interpolating, but it is not differentiable everywhere. However, as clarified in the computational example shown below, these properties are not sufficient for avoiding the interpolation artifacts mentioned above.

V. COMPUTATIONAL EXAMPLES

Here we include a series of computational examples, both with synthetic and real signals and images, aimed at exploring the oscillation artifacts mentioned above, as well as the several remedies described. We note that in all of the computational examples provided in this section we have not considered variations in the region of overlap between the signals (images) being registered. In our computations we have avoided artifacts due to variations in the region of overlap when pure translations and rotations are involved by ensuring the signals being used are equal to zero near the borders. Thus, little or no loss in signal energy is caused by translation or rotation of an image due to image samples falling out of the region of integration. Naturally, many real world imaging experiments do not permit such assumptions. For a discussion of such artifacts in the mutual information similarity measure, please refer to Studholme et al. [19].

A. 1-D synthetic examples

The concepts discussed above are illustrated with the aid of simple computations done with synthetic one dimensional data. Figure 4 contains results of computations involving translating a ‘broad-band’ (a spike or signal containing a broad range of frequencies in its spectrum) signal (top left) with respect to itself and displays the SSD objective function as a function of translation using $\eta^1(x)$ (top right) and η^7 (bottom right) as interpolating basis functions. In order to elucidate the frequency component of the signal and translation filter (for translation value of 0.5 pixels) being used in this simulation, we have plotted the magnitude of the discrete Fourier transform (DFT) in the bottom left panel of Figure 4. We note that although the translation filter being used in this case is an infinite impulse response (IIR) filter of fast decay, the signal contained an abundant amount of zeros near the border. Therefore, the use of the DFT in this case can provide a good approximation of its frequency content. As shown in this figure, at translation 0.5 pixels, interpolation with $\eta^1(x)$ severely attenuates the energy of the signal, introducing

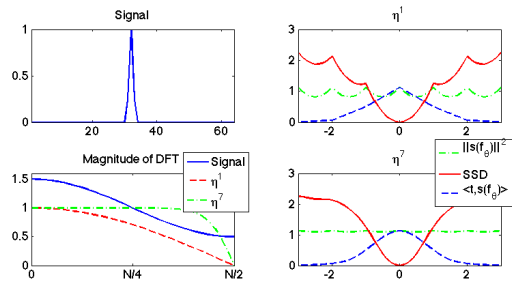


Fig. 4. Top left: signal. Bottom left: DFT of signal and interpolating basis functions $\eta^1(x)$ and η^7 sampled at translation value 0.5. Top right: SSD between original signal and translated signal using $\eta^1(x)$. Note the translation (x axis) values are in pixels (samples). Bottom right: SSD between original signal and translated signal using $\eta^7(x)$.

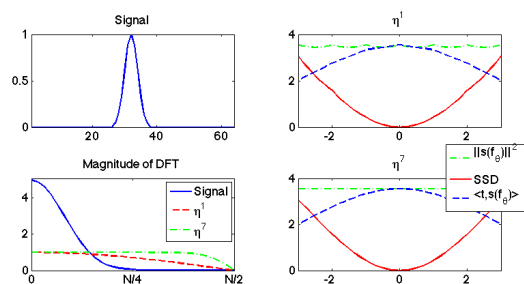


Fig. 5. Top left: signal. Bottom left: DFT of signal and interpolating basis functions $\eta^1(x)$ and η^7 sampled at translation value 0.5. Top right: SSD between original signal and translated signal using $\eta^1(x)$. Bottom right: SSD between original signal and translated signal using $\eta^7(x)$.

the so called oscillation artifact in the cost function. The figure also includes the term $\langle t, \tilde{s}(f_\theta) \rangle_{\ell_2}$ which turns out to be artifact free demonstrating that, in this specific simulation, the oscillatory artifacts in the objective function are entirely due to the fluctuations in signal energy described above.

The situation changes dramatically if a ‘narrow-band’ signal whose spectrum is composed of a few low frequencies is used instead. As shown in the bottom right panel of Figure 5, the energy of the signal is concentrated on low frequency components, where the responses of the translated, sampled interpolating cardinal B-splines of low and high degree are similar. As a result the SSD curve between the signal and its translated versions is similar whether it is computed using $\eta^1(x)$ or $\eta^7(x)$. This also explains, in part, why it is beneficial to use multiresolution image matching schemes such as the ones described in [24]. In addition to avoiding local optima due to fine image detail in the first few iterations, these approaches have the added benefit of also reducing or removing local optima artifacts due to low degree interpolation

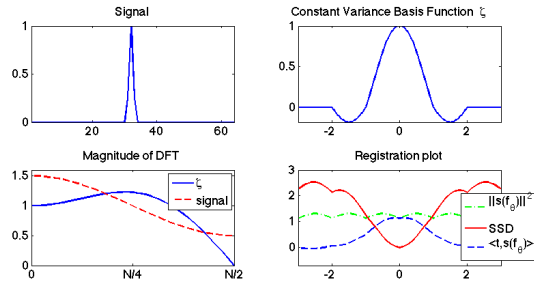


Fig. 6. Top left: signal. Bottom left: DFT of signal and constant variance basis function ζ sampled at translation value 0.5. Top right: Constant variance basis function $\zeta(x)$ in equation 47. Bottom right: SSD between original signal and translated signal using $\zeta(x)$.

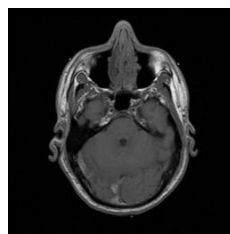


Fig. 7. Magnetic resonance image (8 bits/pixel, 256x256 samples) of the human head used in the pure translation experiments.

of the image data.

We also reproduce the same experiment using the constant variance basis function (47). Results are shown in Figure 6. As results show, the constant variance requirement is not sufficient for avoiding changes in the energy of the translated signal. In fact, since the frequency spectrum of the basis function at translation value $\theta = 0.5$ (see bottom left panel of Figure 6) is greater than one for some frequencies, an increase in energy of the translated signal is observed at half pixel translation values (as opposed to a decrease when using low degree B-splines).

B. Image translation

Figure 7 displays a magnetic resonance image (MRI) of the human head. This image was translated with respect to itself using the B-spline interpolation method described above. As in the 1D experiments above, the image is translated with respect to itself and the value of the objective function is computed for each (sub-pixel) translation value. In addition to straight forward B-spline interpolation, we also compute the objective function using the Monte Carlo (MC) approach of Section IVA, implemented using linear interpolation. Objective function plots computed using a ‘smooth’ version of the image

(computed by convolving the image with a 5x5 uniform filter) are also included. Finally, we have also included computations using the Dirichlet interpolating basis function (see, for example, [21]) computed in the discrete Fourier transform (DFT) domain via multiplication (phase addition) of the DFT of the signal with the complex exponential $\exp(i2\pi\theta k/N)$, where N is the length (size) of the signal, k is the frequency sample, and θ is the value by which the image is being translated in the horizontal dimensional. We denote this method of signal translation as the Fourier method in subsequent figures. As done in [8], translation is performed in one direction (in our case horizontal) while keeping a fixed translation value in the other direction (in our case 3 pixels in the vertical direction). This is done to accentuate any interpolation artifacts that may be present.

As far as computational complexity, for the FFT algorithm we discussed, if N is a power of 2, the computational complexity for producing an interpolated image is roughly $N^d + 2dN^d \log N$, where as before d refers to the dimensionality of the image. For comparison purposes using $N = 256$, for example, linear interpolation would cost approximately 1.3×10^8 computations, FFT-based interpolation 5.7×10^8 computations, while cubic B-spline based image interpolation would cost roughly 1.1×10^9 .

Results for the SSD, cross correlation, and MI objective functions are shown in Figures 8 and 9. Note that the curves computed using fifth degree B-spline and Dirichlet kernel interpolation methods are nearly indistinguishable. The several components of the SSD objective function computed using linear and fifth degree B-spline interpolation are shown in Figure 8 part (b). These plots indicate that, in this computational example, oscillations in the interpolated signal energy are entirely responsible for the local optima encountered in the SSD cost function. The MI objective function was computed using the joint histogram method (32 bins for each image).

C. Local translation

As shown in [9], interpolation artifacts of the type discussed above also occur in nonrigid registration. We demonstrate the performance of the computational algorithms described above by deforming the image shown in Figure 10 according to the following radial basis function model:

$$f_{\theta}(x, y) = \begin{bmatrix} x + \theta \Upsilon \left(\frac{x-x_0}{B}, \frac{y-y_0}{B} \right) \\ y \end{bmatrix} \quad (49)$$

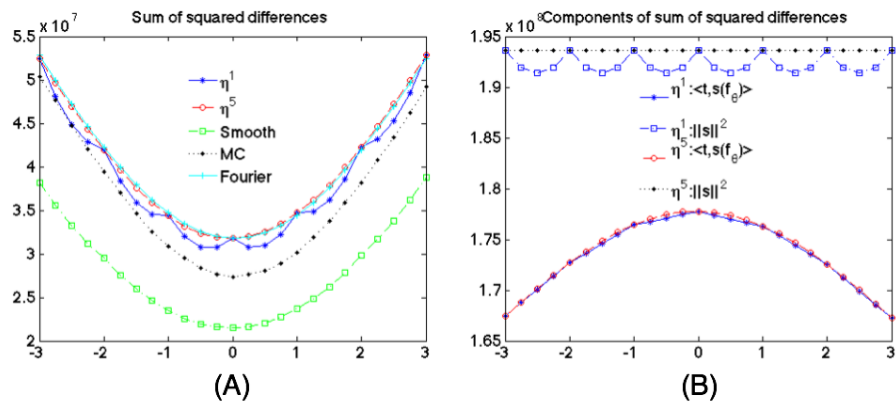


Fig. 8. Part (a): SSD objective function with respect to translation computed using different interpolation strategies. Part (b) different components of SSD objective functions computed using different interpolating basis functions. Fourier interpolation was computed through complex exponential multiplication in DFT domain (see text for more details).

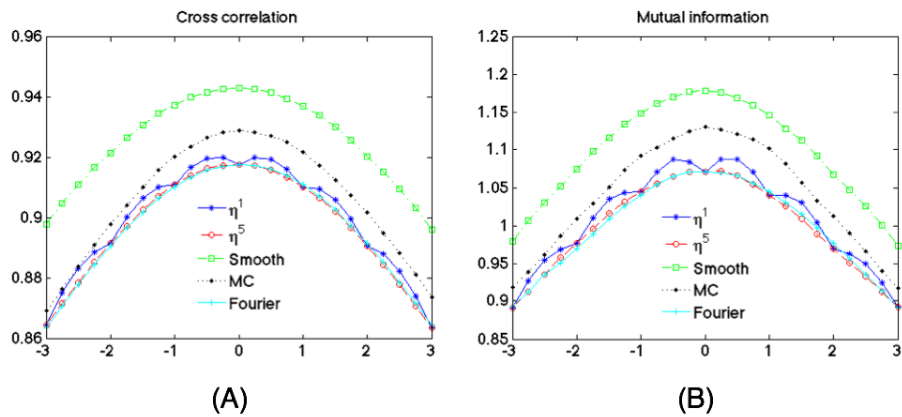


Fig. 9. Part (a): CC objective function computed using several interpolation strategies. Part (b): MI objective function computed using different interpolation strategies. Fourier interpolation was computed through complex exponential multiplication in DFT domain (see text for more details).



Fig. 10. Image (1024x768) used in local translation experiments.

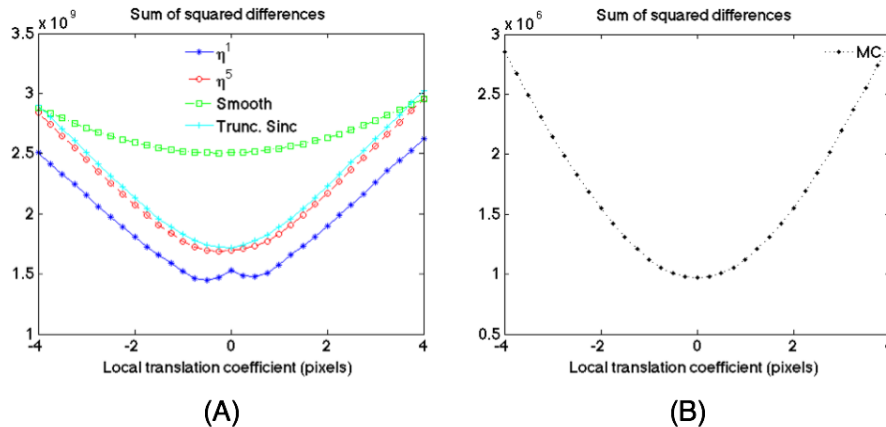


Fig. 11. SSD objective function with respect to local translation. Part (a) shows the SSD similarity measure as a function of local translation for linear, fifth degree B-spline, sinc, and low-pass with linear interpolation methods. Part (b) shows the results using the Monte Carlo integration strategy described earlier.

with x_0, y_0 representing the center of the compactly supported radial basis function

$$\Upsilon(x, y) = [\mu(1 - r)]^4 (4r + 1); \quad r = \sqrt{x^2 + y^2}. \quad (50)$$

and $\mu(x)$ the unit step function. Because $\Upsilon(0, 0) = 1$, when $\theta = 1$, $\tilde{s}(f_\theta(m, n))$ represents a local translation of the pixel coordinate x_0, y_0 by 1 pixel to the left. Neighboring pixels are translated according to equation (49). For large B and small θ , pixels in the immediate vicinity of x_0, y_0 are translated approximately uniformly by θ . Thus B-spline interpolation will modify the energy of the signal, locally, as described earlier.

We repeat the experiments of the earlier section to show that low degree interpolation will cause the introduction of local optima in the SSD objective function (results with other objective functions are omitted for brevity). Other interpolation strategies (high degree, sinc, stochastic sampling, and blurring the data prior to computations) do not produce objective functions with multiple local optima. Results for the sinc interpolation kernel were computed using the standard sinc kernel $\sin(\pi x)/(\pi x)$, truncated at $x = \pm 30$. Results are shown in Figure 11 (boundary conditions were not an issue since we omitted from the computations pixels along the borders of the image). Note that the curves computed using truncated sinc and fifth degree B-spline interpolation are nearly identical. As with the uniform translation experiments, local translation was computed by substituting y in (49) with $y + 3\Upsilon\left(\frac{x-x_0}{B}, \frac{y-y_0}{B}\right)$ so as to accentuate oscillation artifacts. B was set to $768/2$.

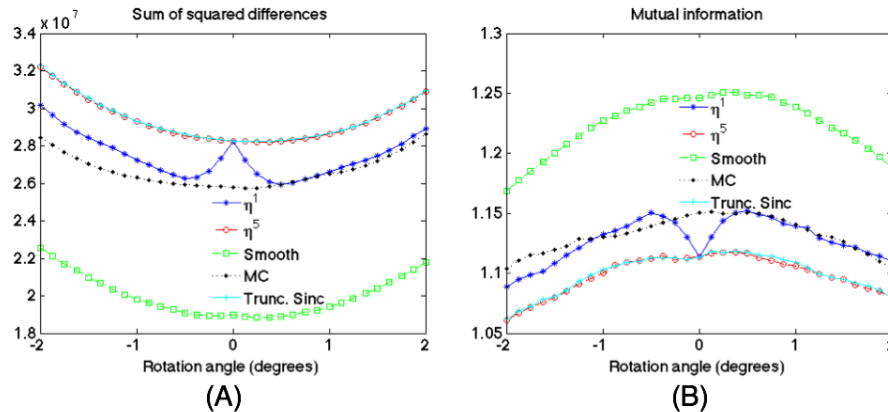


Fig. 12. SSD (a) and MI (b) objective functions with respect to rotation computed using several interpolation algorithms. See text for details.

D. Image rotation

As discussed in previous studies, artifactual local optima also occur when rotating an image. Strictly speaking, the mechanism through which local optima artifacts in rotation experiments are generated differs from the explanation given earlier which was based on pure translations. As a consequence, the inequalities (33) and (37) are not useful for determining the degree n of the B-spline interpolant necessary for rotating a given image. However, we note that rotations can be approximated through sequential pure translations in each direction, as described in Unser et al [28], and because the discrete energy of the image $\|\tilde{s}(f_\theta)\|_{\ell_2}^2$ varies with respect to this translation, oscillations caused by pure translations can be used to understand, intuitively, the cause of local optima artifacts in rotation experiments. Clearly, the effect of the interpolation method on the discrete image energy will be most prominent for rotation values which yield a close to regular sampling grid (0, 90, 180, and 270 degrees), as these, in a broad sense, contain the least amount of interpolation of the data. As the rotation angle deviates continuously from these rotation values, the discrete energy of the rotated signal will decrease due to the interpolation performed. For rotation angles that do not yield a regular sampling grid, changes in angle rotation do not affect the discrete energy of the interpolated image as much. Computational studies of rotation of the MRI image of the human head shown earlier reveal that the behavior of the cost functions in rotation experiments computed with the strategies presented earlier is the same as in the translation case.

VI. DISCUSSION AND CONCLUSIONS

We have discussed oscillation artifacts caused by interpolation algorithms on three different objective functions often used in image registration problems: sum of squared differences, cross correlation, and mutual information. Previous works that have considered interpolation artifacts on mutual information-based objective functions (see [14], [6], [7] for examples) fail to note that the interpolation artifacts are general and affect most intensity-based image registration problems. We have shown that interpolation artifacts also occur in SSD and cross-correlation objective functions. To our knowledge, this is the first time, other than our preliminary work [9], [10] that interpolation artifacts were discussed in a general sense (for several similarity measures simultaneously), and that a precise explanation for them in terms of the energy of the interpolated signal was given. We note that though our analysis is based on B-spline interpolation methods, interpolation artifacts are not exclusive to this family of basis functions. In our experience, such artifacts also occur with other interpolation methods (e.g. Hermite). Our choice of B-splines has to do with the fact that this family includes many popular basis functions (linear, cubic, and in the limit, sinc) and provide a standard framework for analyzing the effects on the energy of the signals being interpolated.

We have explained the source of the artifacts in terms of oscillations in the energy of interpolated, translated, and sampled sequences in ℓ_2 . We have shown that $\|s(f_\theta)\|_{\ell_2}$ oscillations can be mitigated using three different approaches: blurring (low passing) the data prior to computations, higher degree and sinc interpolation, as well as stochastic sampling. Naturally, combinations of these approaches may also be used. We note that sinc (truncated sinc and Dirichlet kernel-based) interpolation and fifth degree B-spline interpolation yielded nearly indistinguishable registration objective functions. Stochastic sampling may be used as a computationally inexpensive alternative to linear interpolation which can avoid the shortcomings of linear interpolation discussed above. Finally, we note that in a typical registration problem image rotation is optimized in conjunction with translation. Therefore, for rotation angles that are different from zero, changes in translation values may not cause any significant oscillations in the objective function since sampling grids at rotation values different from 0, 90, 180, and 270 will not be regular in the sense that the point-wise distances between each resampling grid point to its nearest point in the original image grid will be approximately random. Finally, also note that in the experiments with real image data shown above, the cost function was computed by translating one of the images in the horizontal direction, while displacing the image in the vertical direction so as to augment any oscillations that may be present. The plot of the functions near the actual global optima, for the images used in this study, are smooth and

oscillation artifacts (in the overall objective function) are much reduced in comparison to results shown here.

The results and conclusions presented here are in good agreement with other experimental results in the image processing literature. Plum *et al.* [6] note that oscillation artifacts in the mutual information object function increase with increased level of white noise. White noise, as well known, has a flat frequency spectrum. Thus, an image with significant amounts of noise contains more energy in higher frequency bands than an image with less noise. Low degree interpolation will then modify the energy of the interpolated image more significantly. The authors in [6] also show that by performing a slight resampling of the image can decrease the magnitude of oscillation artifacts. Such resampling operation has two effects on the data. Firstly, it decreases the frequency content of the data, which, as discussed earlier, will tend to diminish the oscillation artifacts. Secondly, the resampling operation as described in [6] has the effect of introducing an irregular sampling of the original data, which may implement the randomization effect described earlier. In addition, based on experimental arguments, Ashburner and Friston [5] describe a method for decreasing the oscillation artifacts on the SSD objective function by first blurring the input data. Unser *et al.* [8], [29] as well as Gan *et al.* [30] show, using computational examples, that stochastic sampling can help in reducing oscillation in the MI similarity measure artifacts, though a precise explanation for the source of artifacts is not included.

Tsao [7] describes a method for computing the mutual information using a ‘jittered’ sampling (akin to the stochastic sampling method described earlier) of the image. Computational examples are included in that work to show that such method can help decrease oscillation artifacts in the mutual information similarity measure. However, Tsao’s main conclusion that higher degree interpolation does not lead to an overall decrease in local optima artifacts in the MI similarity measure is not in agreement with our theoretical and computational results. Although slight improvements of cubic B-splines over linear interpolation for MI computations using a certain number of histogram bins are seen in that work, the improvement did not persist as for higher resolution histograms. In our own experience, the improvements of high degree interpolation over low degree interpolation persist over a reasonable range of histogram sizes (resolutions) with the exception of MI computations using very high resolution histograms, on the order of the same bit resolution as the input images. High degree interpolation methods may introduce many intensity values not originally present in the interpolated image. High resolution histograms will be more sensitive to such changes in image content and have other effects on the overall objective function that our theory is not able to predict.

We also note that several researchers have shown that it is possible to construct estimates of joint

histograms for computing the mutual information between two images via partial volume interpolation (see [14], [8], [31] for examples). These, to some extent, can be used to diminish the number of local optima in mutual information-based registration. It is not clear how the methodology we developed in this paper relates to artifacts observed using partial volume interpolation other than the specific case when zero degree B-spline interpolation is used for pure translations. In this case both methods are equal (with the possible exception of edge artifacts) and will show the same 'staircase' artifacts described in [7].

We conclude with a note of caution clarifying that our main contribution in this work is to show that interpolation-based translation can cause oscillations in the energy of the similarity function which, if left unbalanced by other terms, can contribute to oscillatory (local optima) behavior in several similarity measures described above. Thus our investigation of interpolation artifacts for intensity-based image registration is not exhaustive since it does not investigate many other possible effects that could affect the other terms (e.g. $\mathcal{R}(\theta)$, $\langle t, \bar{s}(f_\theta) \rangle_{\ell_2}$, $\Psi_{constraint}(\theta)$) typically present in such problems. Some of these effects could be due to changes in the region (amount) of overlap between the images, nonlinearities introduced due to histogram binning and number of bins used for computing the MI, bias in higher order moments of the intensity distribution of the interpolated values, to name a few.

APPENDIX A

Here we show that the mutual information similarity measure can be expanded as in (22). We use the Gram-Charlier A series expansion of the multivariate probability density function $p_{S^\theta, T}(\mu, \nu)$. Let V represent the two dimensional random variable $V = \{T, S^\theta\}$. Then

$$p_V(\mathbf{v}) = \phi_V(\mathbf{v}) \{1 + \varphi_V(\mathbf{v})\}, \quad (51)$$

where

$$\varphi_V(\mathbf{v}) = \frac{1}{3!} \sum_{i,j,k} \lambda^{i,j,k} h_{i,j,k}(\mathbf{v}) + \frac{1}{4!} \sum_{i,j,k,l} \lambda^{i,j,k,l} h_{i,j,k,l}(\mathbf{v}) + \dots \quad (52)$$

with $h_{i,j,k}$ the ijk th Hermite tensor, $\lambda^{i,k,k}$ are joint cumulants [32], and $\phi_V(\mathbf{v})$ is the bivariate Gaussian density with correlation coefficient $\rho(\theta)$, and standard deviations σ_{S^θ} , σ_T . These quantities may be estimated from the image samples $t(\mathbf{i})$ and $s(f_\theta(\mathbf{i}))$ as in equations (18),(19),(20). The sums above are

computed for all combination of indexes. The differential entropy of V is then

$$\begin{aligned}
H(V) &= - \int \int p_V(\mathbf{v}) \log(p_V(\mathbf{v})) d\mathbf{v} \\
&= - \int \int \phi_V(\mathbf{v}) \log(\phi_V(\mathbf{v})) d\mathbf{v} - \\
&\quad \int \int \phi_V(\mathbf{v}) \log(1 + \varphi_V(\mathbf{v})) d\mathbf{v} - \\
&\quad \int \int \phi_V(\mathbf{v}) \varphi_V(\mathbf{v}) \log(\phi_V(\mathbf{v})(1 + \varphi_V(\mathbf{v}))) d\mathbf{v}.
\end{aligned} \tag{53}$$

Similar expansions can be written for the univariate probability density functions of random variables S^θ and T :

$$p_{S^\theta}(\mathbf{v}) = \phi_{S^\theta}(\mathbf{v})(1 + \varphi_{S^\theta}(\mathbf{v})), \text{ and} \tag{54}$$

$$p_T(\mathbf{v}) = \phi_T(\mathbf{v})(1 + \varphi_T(\mathbf{v})), \tag{55}$$

with $\phi_{S^\theta}(\mathbf{v})$ and $\phi_T(\mathbf{v})$ representing Gaussian pdfs with variances $\sigma_{S^\theta}^2$ and σ_T^2 , respectively, and $\varphi_{S^\theta}(\mathbf{v})$ and $\varphi_T(\mathbf{v})$ derived from the univariate Gram-Charlier expansions of the pdfs of S^θ and T . The differential entropies are:

$$\begin{aligned}
H(S^\theta) &= - \int \phi_{S^\theta}(\mathbf{v}) \log(\phi_{S^\theta}(\mathbf{v})) d\mathbf{v} - \\
&\quad \int \phi_{S^\theta}(\mathbf{v}) \log(1 + \varphi_{S^\theta}(\mathbf{v})) d\mathbf{v} - \\
&\quad \int \phi_{S^\theta}(\mathbf{v}) \varphi_{S^\theta}(\mathbf{v}) \log(\phi_{S^\theta}(\mathbf{v})(1 + \varphi_{S^\theta}(\mathbf{v}))) d\mathbf{v}
\end{aligned} \tag{56}$$

and

$$\begin{aligned}
H(T) &= - \int \phi_T(\mathbf{v}) \log(\phi_T(\mathbf{v})) d\mathbf{v} - \\
&\quad \int \phi_T(\mathbf{v}) \log(1 + \varphi_T(\mathbf{v})) d\mathbf{v} - \\
&\quad \int \phi_T(\mathbf{v}) \varphi_T(\mathbf{v}) \log(\phi_T(\mathbf{v})(1 + \varphi_T(\mathbf{v}))) d\mathbf{v}.
\end{aligned} \tag{57}$$

Substituting equations (53), (56) and (57) into (16) we arrive to (22) with

$$\begin{aligned}
I_G(\theta) = & - \int \phi_{S^\theta}(\mathbf{v}) \log(\phi_{S^\theta}(\mathbf{v})) d\mathbf{v} - \\
& \int \phi_T(\mathbf{v}) \log(\phi_T(\mathbf{v})) d\mathbf{v} + \\
& \int \int p_V(\mathbf{v}) \log(p_V(\mathbf{v})) d\mathbf{v}
\end{aligned} \tag{58}$$

and

$$\begin{aligned}
\mathcal{R}(\theta) = & \int \int \phi_V(\mathbf{v}) \log(1 + \varphi_V(\mathbf{v})) d\mathbf{v} + \\
& \int \int \phi_V(\mathbf{v}) \varphi_V(\mathbf{v}) \log(\phi_V(\mathbf{v})(1 + \varphi_V(\mathbf{v}))) d\mathbf{v} - \\
& \int \phi_{S^\theta}(\mathbf{v}) \log(1 + \varphi_{S^\theta}(\mathbf{v})) d\mathbf{v} - \\
& \int \phi_{S^\theta}(\mathbf{v}) \varphi_{S^\theta}(\mathbf{v}) \log(\phi_{S^\theta}(\mathbf{v})(1 + \varphi_{S^\theta}(\mathbf{v}))) d\mathbf{v} - \\
& \int \phi_T(\mathbf{v}) \log(1 + \varphi_T(\mathbf{v})) d\mathbf{v} - \\
& \int \phi_T(\mathbf{v}) \varphi_T(\mathbf{v}) \log(\phi_T(\mathbf{v})(1 + \varphi_T(\mathbf{v}))) d\mathbf{v}.
\end{aligned} \tag{59}$$

REFERENCES

- [1] L.G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, pp. 326–376, 1992.
- [2] J.B. Aintoine Maintz and M.A. Viergever, "A survey of medical image registration," *Medical Image Analysis*, vol. 2, pp. 1–36, 1998.
- [3] B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.
- [4] D. Robinson and P. Milanfar, "Fundamental performance limits in image registration," *IEEE Trans. on Image Processing*, vol. 13, pp. 1185–1199, 2004.
- [5] J. Ashburner and K.J. Friston, "Image Registration," in *Functional MRI*, C.T.W. Moonen and P.A. Bandettini, Eds. Springer-Verlag, 1999.
- [6] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Computer Vision and Image Understanding*, vol. 77, pp. 211–232, 2000.
- [7] J. Tsao, "Interpolation artifacts in multimodality image registration based on maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 854–964, 2003.
- [8] 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)*, Strobl, Austria, May 26–30, 2003, pp. 102–109.
- [9] G.K. Rohde, C.A. Berenstein, and D.M. Healy Jr., "Measuring image similarity in the presence of noise," in *Proceedings of the SPIE Medical Imaging: Image Processing*, J.M. Fitzpatrick and J.M. Reinhardt, Eds., San Diego, USA, February 2005, vol. 5747, pp. 132–143.

- [10] G.K. Rohde, Jr. D.M. Healy, C.A. Berenstein, and A. Aldroubi, "Measuring image similarity to sub-pixel accuracy," in *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*, Washington, DC, USA, April 2006, pp. 638–641.
- [11] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I-theory," *IEEE Transactions on Signal Processing*, vol. 41, pp. 821–833, 1993.
- [12] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part II-efficient design and applications," *IEEE Transactions on Signal Processing*, vol. 41, pp. 834–848, 1993.
- [13] J. R. Fienup, "Invariant error metrics for image reconstruction," *Appl. Opt.*, vol. 36, pp. 8352–8357, 1997.
- [14] F. Maes, A. Collignon, D. Vandermeulen, G. Marcha, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Transactions on Medical Imaging*, vol. 16, pp. 187–198, 1997.
- [15] W.M. Wells III, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, pp. 35–51, 1996.
- [16] S. Kullback, *Information Theory and Statistics*, Dover Publications, 1968.
- [17] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multimodality image registration using information theory," 1995, *Information Processing in Medical Imaging*, pp. 263–274, Kluwer Academic Publishers.
- [18] J.P. Pluim, J. B. Mainz, and M.A. Vergever, "Mutual-information-based registration of medical images: a survey," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 986–1004, 2003.
- [19] C. Studholme, D.L.G. Hill, and D.J. Hawkes, "An overlap invariant entropy measure of 3D medical image alignment," *Pattern Recognition*, vol. 32, pp. 71–86, 1999.
- [20] A. Aldroubi, M. Unser, and M. Eden, "Cardinal spline filters: Stability and convergence to the ideal sinc interpolator," *Signal Processing*, vol. 28, pp. 127–138, 1992.
- [21] H. Foroosh, J.B., Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, pp. 188 – 200, 2002.
- [22] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup, "Efficient subpixel image registration algorithms," *Opt. Lett.*, vol. 33, pp. 156–158, 2008.
- [23] E. Sidick, J. J. Green, R. M. Morgan, C. M. Ohara, and D. C. Redding, "Adaptive cross-correlation algorithm for extended scene shack-hartmann wavefront sensing," *Opt. Lett.*, pp. 213–215, 2008.
- [24] P. Thévenaz, Urs E. Ruttimann, and M. Unser, "A pyramid approach to subpixel registration based on intensity," *IEEE Transactions on Image Processing*, vol. 7, no. 1, pp. 27–41, January 1998.
- [25] G.K. Rohde, A.S. Barnett, P.J. Basser, and C. Pierpaoli, "Estimating intensity variance due to noise in registered images: applications to DT-MRI," *Neuroimage*, vol. 26, pp. 673–684, 2005.
- [26] O. Salvado and D. Wilson, "Removal of local and biased global maxima in intensity-based registration," *Medical Image Analysis*, vol. 11, pp. 183–196, 2007.
- [27] P. Thévenaz, T. Blu, and M. Unser, "Short basis functions for constant-variance interpolation," in *Proceedings of the SPIE Medical Imaging: Image Processing*, in Press, 2008.
- [28] M. Unser, P. Thévenaz, and L. Yaroslavsky, "Convolution-based interpolation for fast, high-quality rotation of images," *IEEE Transactions on Image Processing*, vol. 4, pp. 1371–1381, 1995.
- [29] P. Thévenaz, M. Bierlaire, and M. Unser, "Halton sampling for image registration based on mutual information," *Sampling Theory in Signal and Image Processing*, vol. In Press, 2006.

- [30] R. Gan, J. Wu, A.C.S. Chung, S.C.H. Yu, and W.M. Wells III, "Multiresolution image registration based on Kullback-Leibler distance," in *MICCAI*, C. Barillot, D.R. Haynor, and P. Hellier, Eds., Berlin, 2004, vol. 3216 of *LNCIS*, pp. 599–606, Springer-Verlag.
- [31] H. Chen and P.K. Varsheny, "Mutual information-based CT-MR brain image registration using generalized partial volume joint histogram estimation," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 1111–1119, 2003.
- [32] P. McCullagh, *Tensor Methods in Statistics*, Chapman and Hall Ltd, 1987.