Masked FFT Registration

Dirk Padfield
GE Global Research
One Research Circle, Niskayuna, NY 12309
padfield@research.ge.com

Abstract

Registration is a ubiquitous task for image analysis applications. Generally, the requirements of registration algorithms include fast computation and large capture range. For these purposes, registration in the Fourier domain using normalized cross correlation is well suited and has been extensively studied in the literature. Another common requirement is masking, which is necessary for applications where certain regions of the image that would adversely affect the registration result should be ignored. To address these requirements, we have derived a mathematical model that describes an exact form for embedding the masking step fully into the Fourier domain. We also provide an extension of this masked registration approach from simple translation to also include rotation and scale. We demonstrate the computational efficiency of our algorithm and validate its correctness on several synthetic images and real ultrasound images. Our framework enables fast, global, parameter-free registration of images with masked regions.

1. Introduction

Image registration is a fundamental task in image analysis whereby two or more images are aligned by finding a transformation that minimizes some distance between the transformed target image and the reference image. Registration algorithms can be divided into a large number of categories [5, 8], but, in this paper, we are concerned with the class of registration algorithms based on the Fourier transform [1, 7, 15]. This class is of particular interest for various reasons. First, only one iteration of the algorithm is needed to provide a global registration result that calculates the metric for all possible transforms between the images. This is in contrast to algorithms that start from a given transform, test the metric after moving in a particular direction, and iterate this process. Such algorithms only calculate the metric for a small region of the overlap, can become stuck in local optima, and can take a long time to converge. This is important in applications where the images are significantly misaligned so that few constraints can be placed on the transformation between them. Second, Fourier domain algorithms are fast because they require only a small number of forward and backward transforms along with simple mathematical operations. Third, such approaches require few or no parameters. Furthermore, FFT approaches are applicable not only to translational registration; they have also been extended to rotational transforms [7], and further extended to also include scale [15, 6, 23, 17, 16]. This is accomplished by making use of the rotation and scale properties of Fourier transforms and their duals in the spatial domain.

This paper focuses on the use of the correlation metric since the cross-correlation of two signals has a dual in the Fourier domain. Correlation is the most widely used method for similarity detection [3] and the automatic determination of translation [2, 13]. For images from the same imaging modality, the normalized cross correlation (NCC) metric is appropriate because it is insensitive to multiplicative factors between the two images and produces a cost function with sharp peaks and well-defined minima. In this context, it is also worth noting that the use of phase correlation with FFTs for registration has also been extensively studied [11, 9].

During registration, it may be desirable to restrict evaluation of the metric within a specified region so that undesirable regions do not contribute to the metric calculation. For example, in the registration of ultrasound images (see Figure 1), the background region surrounding the “pie-slice” containing the actual image information should be ignored. If the pixels in such regions contribute to the metric calculation, this can lead to incorrect registration results. Various approaches have been proposed for such masking. Kaneko et al. in [10] propose an algorithm that incorporates a binary mask into the NCC equation, but, in the process of masking the images before correlation, the masked regions still affect the registration by lowering the sums used for normalization. This problem also exists for the approach of Lei in [12] where the images are masked by multiplying them with the binary masks. Alternatively, approaches such as those
Fixed image  Moving Image  Mask Image  Standard FFT NCC  Masked FFT NCC

Figure 1. **Registration example requiring region masking.** The first and second images are the fixed and moving images to be registered. The third image is the mask, where white indicates the region of interest and black indicates the region to be ignored during registration. The fourth image shows the result of the standard FFT NCC algorithm, which is unable to find the correct transform because of the influence of the zero values around the cone beam. The final image shows the result of the masked FFT NCC algorithm proposed in this paper, which calculates the transforms correctly with a correlation score of 1.

Table 1. **Advantages and disadvantages of NCC registration methods.** The three classes of NCC methods are shown in the rows. The different approaches are compared relative to the three requirements listed in the column headings: global transform computation, fast computation, and ability to mask image. A “+” means the approach is able to meet the requirement, and a “−” means it is not.

<table>
<thead>
<tr>
<th></th>
<th>Global</th>
<th>Fast</th>
<th>Masked</th>
<th>Selected Refs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>[1, 7, 15]</td>
</tr>
<tr>
<td>Mask Spatial</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>[10, 12]</td>
</tr>
<tr>
<td>Mask Iterative</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>[21, 20]</td>
</tr>
</tbody>
</table>

by Thevenaz *et al.* [21, 20] provide a mask under which the registration metric is to be calculated, but this masking is applied in an iterative framework, so a global solution is not calculated. Furthermore, approaches that calculate the masked NCC in the spatial domain for all locations become computationally infeasible for large images. Finally, Stone *et al.* in [18, 19] present a masking step based on correlations for the application of measuring differences between images and patterns.

Table 1 summarizes the advantages and disadvantages of various approaches for the task of calculating global transformations in a computationally efficient manner using masked regions. It is desirable to integrate the advantages of these approaches into a consistent framework.

In order to meet all of the requirements of Table 1, in this paper we present a method to integrate the masking directly into the FFT registration framework. Our goal is neither to develop a new metric nor promote the NCC metric over others but rather to introduce a method for enabling masking in the Fourier domain using the NCC metric. This is accomplished by introducing a mathematical formulation of the masked NCC that is fully described using forward and backward Fourier transforms. In this approach, the masked regions are completely ignored. This is a fundamentally different approach from simply multiplying the binary mask by its corresponding image and then registering, which will lead to registration errors since the zero-values still influence the registration metric (see the fourth image in Figure 1). Our derivations indicate that the computational complexity of the masked NCC is the same as the standard NCC. We further generalize our method to enable full masked translation, rotation, and scale registration. Finally, we demonstrate the ability of the algorithm to correctly ignore the regions outside of the mask on several synthetic images and real ultrasound images. The algorithm is able to yield the correct transformations between the images and outputs high correlation scores. In summary, our approach combines the global nature, speed, and parameter-free advantages of the Fourier transform and the appropriateness of the NCC metric for intra-modality images with the power of masking undesirable regions in the images.

2. **Methods**

In this section, we demonstrate how to derive mathematically the masked registration algorithm in the Fourier domain. First, in Section 2.1, we give the standard form of normalized cross correlation (NCC) of two images and show a form of this equation consisting of sums over the images. Although these sums can be efficiently calculated using a type of integral images, such methods will not work in the case of masked registration. However, we demonstrate that the sums can be represented fully in the Fourier domain and that this representation generalizes to the case of masked registration. Therefore, in Section 2.2, we transform all of these sums into forms that are represented fully using FFTs. We show in Section 2.3 how this representation leads to a simple and general form for masked FFT NCC. Finally, in Section 2.4, we generalize the masked framework to translation, rotation, and scale registration.

2.1. **Spatial Form of Normalized Cross Correlation**

The 2D normalized cross correlation indicates the similarity of two images \( f_1(x, y) \) and \( f_2(x - u, y - v) \), with \( f_2 \) shifted by \((u, v)\). Using this notation, we define that \( f_1 \) is
the fixed image and \( f_2 \) is the moving image. In the process of shifting, we use \( D_{u,v} = D_2(u,v) \cap D_1 \) to represent the region of overlap of the two images, where \( D_1 \) is the domain of \( f_1 \) and \( D_2(u,v) = \{(x,y)|(x-u,y-v) \in D_2\} \) is the domain \( D_2 \) of \( f_2 \) shifted by \((u,v)\). The region of overlap is constantly shifting and represents the overlapping region of the correlation operation. Using this notation, the normalized cross correlation between images \( f_1 \) and \( f_2 \) at a given \((u,v)\) is defined as

\[
\text{NCC}(u,v) = \frac{\text{NCC}_{\text{num}}(u,v)}{\sqrt{\text{NCC}_{\text{den},1}(u,v)} \sqrt{\text{NCC}_{\text{den},2}(u,v)}}
\]

(1)

where the numerator \( \text{NCC}_{\text{num}}(u,v) \) can be defined as

\[
\sum_{(x,y) \in D_{u,v}} f_1(x,y) f_2(x-u,y-v)
\]

\[
\sum_{(x,y) \in D_{u,v}} f_1(x,y) \sum_{(x,y) \in D_{u,v}} f_2(x-u,y-v)
\]

(2)

the \( \text{NCC}_{\text{den},1}(u,v) \) term can be represented as

\[
\sum_{(x,y) \in D_{u,v}} (f_1(x,y))^2 - \left( \sum_{(x,y) \in D_{u,v}} f_1(x,y) \right)^2
\]

(3)

and the \( \text{NCC}_{\text{den},2}(u,v) \) term is defined analogously.

In Equation 2, the first term is simply the definition of the cross correlation of the images. The numerator of the second term consists of the product of the local sum of \( f_1 \) and the local sum of \( f_2 \), and the denominator is the number of pixels in the overlapping region. The sums in the terms of Equations 2 and 3 can be computationally intensive because they must be computed for each region of overlap of the two images. This amounts to \((q_1 + q_2 - 1) \times (r_1 + r_2 - 1)\) different sums, where \((q_j,r_j)\) are the rows and columns, respectively, of the images, and \( j \in 1,2 \). Although there are algorithms to compute these sums more efficiently by precomputing running sums such as integral images, these algorithms will have difficulty when using masked regions. We therefore seek to represent all of these sums directly in the Fourier domain.

2.2. FFT Form of Normalized Cross Correlation

Our next goal is to define the entire normalized cross correlation operation in the Fourier domain for the standard NCC without masking. We will then show in Section 2.3 that this representation enables a simple definition of masked NCC.

For simplicity of notation, we will rotate the moving image \( f_2 \) by 180°, which we will call \( f_2' \), and we will carry out all operations on this rotated image. We will also define \( f_1 = \mathcal{F}(f_1) \), and \( f_2' = \mathcal{F}(f_2') \), where \( \mathcal{F}(\cdot) \) represents the FFT operation. Also, assuming \( i_1 \) and \( i_2 \) are images of ones the same size as \( f_1 \) and \( f_2 \), respectively, we define \( i_1 = \mathcal{F}(i_1) \) and \( i_2' = \mathcal{F}(i_2') \). Furthermore, the size of each dimension of all of the FFT images in the following equations is set to \( \max(q_1 + q_2 - 1, r_1 + r_2 - 1) \), where \((q_j,r_j)\) represents the number of rows and columns of the images, and the subscripts refer to the two images. This is accomplished by padding the images with zeros before calculating the FFT. Note that all of the FFTs are thus defined as square and that the sizes of the original images could be different from each other.

The first term of Equation 2 is the cross correlation of the two images, and it is well known [4] that correlation in the spatial domain can be calculated in the Fourier domain as \( \text{CC}(f_1,f_2) = \mathcal{F}^{-1}(\mathcal{F}(f_1) \cdot \mathcal{F}(f_2')) \). To accomplish the conversion of the elements of Equation 1 to the Fourier domain, we need to re-define the local sums in Equations 2 and 3 in terms of Fourier transforms, which can be done as follows

\[
\sum_{(x,y) \in D_{u,v}} 1 = \mathcal{F}^{-1}(i_1 \cdot i_2')(u,v)
\]

(4)

\[
\sum_{(x,y) \in D_{u,v}} f_1(x,y) = \mathcal{F}^{-1}(F_1 \cdot I_2')(u,v)
\]

(5)

\[
\sum_{(x,y) \in D_{u,v}} (f_1(x,y))^2 = \mathcal{F}^{-1}(\mathcal{F}(f_1) \cdot F_2')(u,v)
\]

(6)

The first of these local sums \( \sum_{x,y \in D_{u,v}} 1 \) calculates the number of pixels in the overlap region for a particular value of \((u,v)\). Because the correlation operation by definition calculates the sum of the product in the overlap region as the moving image is shifted across the fixed image, we can define this sum in Equation 4 as the correlation of two images whose values are all 1. The second of these sums \( \sum_{x,y \in D_{u,v}} f_1(x,y) \) is the local sum of image \( f_1 \), which can similarly be represented as the correlation of this image with an image of ones in Equation 5. Similarly, in Equation 6, the sum \( \sum_{x,y \in D_{u,v}} (f_1(x,y))^2 \) is the local sum of the element-wise square of image \( f_1 \), which can be represented by the correlation of this image with an image of ones.

Given this notation, Equation 2 becomes

\[
\text{NCC}_{\text{num}} = \mathcal{F}^{-1}(F_1 \cdot F_2') - \frac{\mathcal{F}^{-1}(F_1 \cdot I_2') \cdot \mathcal{F}^{-1}(I_1 \cdot F_2')}{\mathcal{F}^{-1}(I_1 \cdot I_2')},
\]

(7)
Equation 3 becomes

\[ \text{NCC}_{\text{den},1} = \mathcal{F}^{-1}(\mathcal{F}(f_1 \cdot f_1) \cdot I_2^2) - \frac{(\mathcal{F}^{-1}(f_1^2) \cdot I_2^2)}{\mathcal{F}^{-1}(I_1 \cdot I_2^2)}, \] (8)

and NCC\text{den,2} becomes

\[ \text{NCC}_{\text{den,2}} = \mathcal{F}^{-1}(I_1 \cdot \mathcal{F}(f_1^2 \cdot f_2^2)) - \frac{(\mathcal{F}^{-1}(I_1 \cdot f_2^2))}{\mathcal{F}^{-1}(I_1 \cdot I_2^2)}. \] (9)

Note that, whereas Equations 1, 2, and 3 are defined for a particular value of \((u,v)\), Equations 7, 8, and 9 are defined for all values of \((u,v)\).

This construction requires the calculation of 6 forward FFTs (only 5 are required if the image size of \(f_1\) and \(f_2\) are the same since \(I_1\) and \(I_2\) will then be equal) and 6 backward FFTs. However, since the 6 forward FFTs all have as their input real images, we embed them in pairs into complex images, calculate 3 FFTs, and then use odd-even separation to extract their FFTs, thus saving 3 FFT computations. These forward and backward FFTs make up the majority of the computational complexity.

### 2.3. Masked FFT NCC

In this section, we extend the FFT NCC formulation to enable the correlation of images that have associated masks. The beauty of representing the NCC completely in the Fourier domain in Section 2.2 is that its generalization to the masked FFT NCC becomes straightforward. We will demonstrate how this formulation enables the pixels in the masked regions to be totally ignored so that they have no effect in the NCC metric.

Assume \(m_1\) and \(m_2\) are mask images of the same size as \(f_1\) and \(f_2\), respectively. In these masks, regions of interest are given by a value of 1, and regions to be ignored are given a value of 0. To account for the masked regions, in the process of shifting, we use \(D_{u,v,m} = D_{2,m}(u,v) \cap D_{1,m}\) to represent the region of overlap of the two images where neither of the masked images are 0. Here \(D_{1,m}\) is the domain of \(f_1\) in the areas that \(m_1 \neq 0\), and \(D_{2,m}(u,v) = \{x,y|(x-u,y-v) \in D_{2,m}\}\) is the domain \(D_{2,m}\) of \(f_2\) shifted by \((u,v)\) in the areas that \(m_2 \neq 0\). Note that this notation ensures that the regions in the masks that are set to 0 have no influence in the overlap region. This is in contrast to simply masking the fixed and moving images and then correlating them using Equation 1, which would result in the zero values influencing the calculation (Figure 2 demonstrates the fundamental difference of these two approaches).

Using the domain \(D_{u,v,m}\), we can re-define the sums from Equations 4, 5, and 6. As in the case of correlating the images with images of ones in Equations 7 and 8, we define \(M_1 = \mathcal{F}(m_1)\) and \(M_2 = \mathcal{F}(m_2')\), where \(m_2'\) is \(m_2\) rotated by 180°. Then the sums can be converted to FFTs as follows

\[
\sum_{(x,y) \in D_{u,v,m}} 1 = \mathcal{F}^{-1}(M_1 \cdot M_2')(u,v) \quad (10)
\]

\[
\sum_{(x,y) \in D_{u,v,m}} f_1(x,y) = \mathcal{F}^{-1}(F_1 \cdot M_2')(u,v) \quad (11)
\]

\[
\sum_{(x,y) \in D_{u,v,m}} (f_1(x,y))^2 = \mathcal{F}^{-1}(F_1 \cdot F_1 \cdot M_2')(u,v) \quad (12)
\]

Given these sums, the masked version of the numerator in Equation 7 becomes

\[
\text{NCC}_{\text{num}} = \mathcal{F}^{-1}(F_1 \cdot F_2') - \frac{\mathcal{F}^{-1}(F_1 \cdot M_2') \cdot \mathcal{F}^{-1}(M_1 \cdot F_2')}{\mathcal{F}^{-1}(M_1 \cdot M_2')}, \quad (13)
\]

Equation 8 becomes

\[
\text{NCC}_{\text{den},1} = \mathcal{F}^{-1}(\mathcal{F}(f_1 \cdot f_1) \cdot M_2') - \frac{(\mathcal{F}^{-1}(F_1 \cdot M_2'))^2}{\mathcal{F}^{-1}(M_1 \cdot M_2')}, \quad (14)
\]

and Equation 9 becomes

\[
\text{NCC}_{\text{den},2} = \mathcal{F}^{-1}(M_1 \cdot \mathcal{F}(f_2' \cdot f_2')) - \frac{(\mathcal{F}^{-1}(M_1 \cdot F_2'))^2}{\mathcal{F}^{-1}(M_1 \cdot M_2')}. \quad (15)
\]

The advantages of representing the NCC fully in the FFT domain are clear from Equations 7, 8, and 9: it enables the representation of the masked NCC by simply replacing the \(I_1\) and \(I_2\) in the equations with \(M_1\) and \(M_2\). These equations also demonstrate that the computational complexity for the masked algorithm is exactly the same as the complexity for the standard NCC.

To summarize, the masked FFT NCC can be calculated using a few simple steps. First, rotate \(f_2\) and \(m_2\) by 180° and then calculate the FFTs \(F_1, F_2', M_1,\) and \(M_2'\) using zero padding to make the resulting images of size \((q_1 + q_2 - 1, r_1 + r_2 - 1)\). Next, calculate \(\text{NCC}_{\text{num}}\) using Equation 13, \(\text{NCC}_{\text{den},1}\) using Equation 14, and \(\text{NCC}_{\text{den},2}\) using Equation 15. Finally, plug these expressions into Equation 1 (ignoring the \((u,v)\) indexing) to yield the masked NCC.

For completeness, we here make a brief mention of how the result of the NCC yields the translation transform between the images. The result of the NCC operation is an image of size \((q_1 + q_2 - 1, r_1 + r_2 - 1)\) with values between -1 and 1, where 1 indicates a perfect correlation. The center of this NCC image represents a translation of \((0,0)\), so the true translation between the images is found by finding the offset from the center to the location with the highest correlation score. Practically, we first zero out values on the border of the NCC image because in these regions there is insufficient overlap between the images for a stable computation of NCC. We use the image of the number of overlap pixels \(\mathcal{F}^{-1}(M_1 \cdot M_2')\) for this purpose: values in the overlap
demonstrating the correctness and effectiveness of our

3. Results

When the moving image is rotated and scaled, the

differences in their framework even for non-masked registration applica-

tions: when the moving image is rotated and scaled, the

Our full algorithm for finding the transforms using

Our full algorithm for finding the transforms using

Our full algorithm for finding the transforms using

Algorithm 1: Masked image registration algorithm for translation, rotation, and scale.

| Input: Fixed image \( f_1(x, y) \), moving image \( f_2(x, y) \),
| fixed mask \( m_1(x, y) \), moving mask \( m_2(x, y) \) |
| Output: Registered moving image \( \hat{f}_2(x, y) \) and transform parameters \( x_0, y_0, \theta, \tau \) |
| 1 Calculate the Fourier transforms \( F_1 \) and \( F_2 \); |
| 2 Calculate the log-polar images \( N_1(\log \rho, \theta) \) and \( N_2(\log \rho, \theta) \); |
| 3 Correlate the log-polar images using NCC to find \( \theta_0 \) and \( \tau \); |
| 4 Transform the moving image with \( \theta_0 \) and \( \tau \) to find \( \tilde{f}_2(x, y) \); |
| 5 Transform the moving mask with \( \theta_0 \) and \( \tau \) to find \( \tilde{m}_2(x, y) \); |
| 6 Correlate \( f_1(x, y) \) with \( \tilde{f}_2(x, y) \) using masked NCC with \( m_1(x, y) \) and \( \tilde{m}_2(x, y) \) to find \( x_0 \) and \( y_0 \); |
| 7 Transform \( \tilde{f}_2(x, y) \) with \( x_0 \) and \( y_0 \) to find \( \hat{f}_2(x, y) \); |

image that are less than a reasonable value are zeroed out in the

2.4. Masked FFT NCC Translation, Rotation, and

Scale Registration

The methods presented thus far have been concerned pri-

arily with masked FFT NCC, which enables only translation

registration. However, they are generalizable to translation,

rotation, and scale transforms using a framework similar to that of Reddy and Chatterji [15] and Lucchese et al. [14]. The method is founded on the principles that the Fourier transform is translation invariant and that its conver-
sion to log-polar coordinates converts the scale and rotation differ-

ces to vertical and horizontal offsets.

Our full algorithm for finding the transforms using

masked registration is given in Algorithm 1. The main dif-
fERENCE from [15] is that the translation component of the

registration is found using masked NCC instead of standard

phase correlation. This overcomes an important limitation of their framework even for non-masked registration applica-
tions: when the moving image is rotated and scaled, the

background values must be set to some value, normally 0.

These 0 values then have an effect on the correlation step and
can lead to errors. However, using masked registration, those background regions will be ignored. Thus, even if
masking of the original images is not used, the output of this
algorithm will be more accurate than that described in

[15].

3. Results

Here we provide several synthetic and real results
demonstrating the correctness and effectiveness of our

masked FFT NCC algorithm as compared with the standard

FFT NCC algorithm. It is not our intention to claim that

using the FFT algorithm with the NCC is the best approach,

but rather to show the correctness and utility of our embed-
ing of the masking step in the Fourier domain. For this

purpose, these results utilize a registration application, but

there are many other applications for our masked algorithm

such as template matching, object tracking, and image re-

trieval in the presence of undesired regions.

3.1. Synthetic Masked Registration Results

To demonstrate the capabilities of the masked registra-
tion algorithm, Figure 2 shows results of registering vari-

ous objects in two synthetic images. In these images, three

objects move independently of one another, and there is a

border of constant value. Noise with a uniform distribution

has been added.

To demonstrate the flexibility to define any arbitrary

mask, we show three different masks in the second column,
each one centered around a different object of the fixed im-
age. These are the fixed masks, and for this experiment the

moving masks are set to all 1 (no masking). Using these

masks in the masked registration algorithm is very different

from simply multiplying the image with the mask before the

registration step. The difference lies in the fact that masked

registration actually ignores all values outside of the mask

whereas the standard FFT step will still try to align the val-

ues outside of the mask. To demonstrate this difference, the

third column shows the result of first multiplying the fixed

image with the masks and then running standard FFT NCC

registration. The first two images fail completely, and the

third succeeds only because of the large size of the third

object. Furthermore, the correlation scores are quite low at

0.31, 0.42, and 0.62, which is undesirable given that corre-

lation scores are important as registration quality metrics.

In contrast, the fourth column displays the results of

masked registration, where the masks are passed as fixed

image masks, and the moving image masks are set to 1. In

this case, all of the objects are perfectly registered. Also, the

correlation scores are far more informative than those of the

third column at 0.99, 0.99, and 0.99 (they are not exactly 1

because of the added noise).

These results also demonstrate another powerful aspect

of the approach: if we know the location of objects in the

fixed image and mask them, we don’t need to know their

location in a different image in order for the registration to

work correctly. Such a feature could be useful, for example,

for tracking where the user knows the location in the first

frame but not necessarily in others.

3.2. Ultrasound Masked Registration Results

We used ultrasound images to demonstrate registration

results on real images because the masked region is an ir-
Figure 2. Masked registration results for synthetic data using different masks. The second column shows three different masks that define regions in the fixed image that should be aligned in the moving image; the moving masks are all set to 1. The last two columns show overlay results, where the red is the fixed image and the green is the transformed moving image. The third column shows the overlay result of standard FFT NCC where the fixed image is simply multiplied by each fixed mask before registration; it is clear that the zero regions outside of the mask adversely affect the result. The fourth column shows the result of masked registration, which yields perfect transformations and demonstrates the accuracy of the approach. The numbers in parentheses under these images are the correlation score.

regular pie-shape, which demonstrates the ability of the algorithms to work on arbitrary mask shapes. We used the ultrasound image from Figure 1, transformed the image information by a known amount, and then masked both the fixed and the moving image with a new window that contains only image information contained in both images. Note that if the entire image would have been transformed instead, this would not test the masking because the masked region would move with the image.

The results of registering these images with the standard FFT NCC and the masked FFT NCC are given in Figure 3. The moving images in each row are transformed with different known combinations of translations, rotations, and scaling and then registered to the fixed image. The third column shows the results of the standard FFT NCC. In these images, the background region has such a large effect on the registration values that it tries to maintain the alignment of the pie-shape, resulting in poor correlation scores between 0.527 and 0.743. However, the masked results in the last column show that the background region is entirely ignored, and perfect registration results are achieved even in cases where the transform is extreme. For all of the masked results, the calculated transform corresponds exactly with the ground truth values and the calculated correlation score is a perfect 1 (except for slight interpolation errors), which is possible since the actual image information is the same in
Figure 3. **Registration results comparing masked FFT NCC with standard FFT NCC.** Each row represents images with different known transforms, which are shown on the left of each row with the format (x translation, y translation, rotation, scale). The first column shows the fixed image, the second shows the moving image (transformed by the given transform), the third shows the results using standard FFT NCC, and the last shows the results using masked FFT NCC. In the registration result images in the last two columns, the correlation score for the methods are shown in white. The standard FFT NCC, influenced by the zero values around the cone beam, often fails to find the correct transform and results in poor correlation scores, whereas the masked FFT NCC calculates the transforms correctly with a perfect correlation score of 1 (except for minor effects from interpolation).
the fixed and moving images except that the moving image pixels are transformed. Note that, in the extreme transform cases, algorithms based on iterative registration would have great difficulty and would likely become stuck in local minima and not converge to the correct solution.

4. Conclusions and Future Work

We have presented a mathematical framework that enables the efficient registration in the Fourier domain of images that have associated masks. The definition of the masked registration in the Fourier domain takes advantage of the fast, global, and parameter-free characteristics of Fourier domain registration. Furthermore, we demonstrated that the computation of the masked registration is as efficient as the standard FFT NCC registration, requiring only 3 forward FFTs, 6 backward FFTs, and a number of element-wise matrix multiplications. The results demonstrate the correctness of the algorithm through its ability to find the exact transformation between images with a perfect correlation score of 1. This algorithm for masked registration has a large range of applications beyond the medical application demonstrated in the results. For example, it could be used for 3D medical datasets such as those in [22], for registering photographs with occluded regions, and for tracking in video applications that require the masking of regions that adversely affect the results.

For future work, we plan to extend the algorithm to work on weighted masks rather than binary masks. Weighted masks will have even broader application since they will enable the weighting of each pixel in the image independently based on, for example, the confidence in the importance of a particular pixel.

The masked FFT NCC code is included in the supplemental material and is available upon request. Please send us an e-mail if you would like a copy.

Acknowledgment

We would like to thank Feng Lin, Navneeth Subramanian, Anand Narasimhamurthy, and Paulo Mendonça for many helpful discussions.

References