An improved 3D shape context registration method for non-rigid surface registration

Di Xiao*, David Zahra, Pierrick Bourgeat, Paula Berghofer, Oscar Acosta Tamayo, Catriona Wimberley, Marie Claude Gregoire, Olivier Salvado

*aThe Australian e-Health Research Center, ICT, CSIRO, Australia
*bRadiopharmaceutical Research Institute, ANSTO, Australia

ABSTRACT

3D shape context is a method to define matching points between similar shapes as a pre-processing step to non-rigid registration. The main limitation of the approach is point mismatching, which includes long geodesic distance mismatch and neighbors crossing mismatch. In this paper, we propose a topological structure verification method to correct the long geodesic distance mismatch and a correspondence field smoothing method to correct the neighbors crossing mismatch. A robust 3D shape context model is proposed and further combined with thin-plate spline model for non-rigid surface registration. The method was tested on phantoms and rat hind limb skeletons from micro CT images. The results from experiments on mouse hind limb skeletons indicate that the approach is robust.

Keywords: 3D Shape context, non-rigid registration, small animal skeleton registration

1. INTRODUCTION

Small animal imaging is increasingly used as a pre-clinical tool to assess effectiveness of therapy. Positron Emission Tomography (PET) scans or combined micro-CT scans can provide some anatomical information and functional information. One challenging step to study a large numbers of individual animals is the spatial normalization. From the information present in the micro-CT, scans need to be warped to a template, before subsequent processing, for example, registering to their respective PET images.

Using 3D shape context algorithm for surface registration is one alternative to intensity based non-linear registration [1]. Given a set of points on a surface, 3D shape context provides the corresponding point locations on a target surface with similar shape. Once the matching is established, warping of one surface to the other can be performed with a smooth deformation model. Shape context was proposed by Belongie et al for 2D graph matching [2], providing an effective way to compute the similarity between two sets of points. It was later extended [3] for measuring similarity between shapes and exploiting it for object recognition. Mori et al. [4] demonstrated its ability for quick search for 2D similar shapes. Carneiro et al. [5] applied it to improve the distinctiveness of local image features. Yusuf et al. [6] applied it for handwritten Urdu digit recognition and got zero percent error on their 28 test digits. Liu et al. [7] extended shape context to shape context pattern and applied it for feature point matching on Chinese character images by eigenvector analysis. Yang et al. [8] integrated shape context based recognition into image segmentation for 2D image segmentation. Efforts were made for improving the performance of shape context. Singh et al. [9] proposed an enhanced shape context by identifying corner points on object contours as landmarks. Urschler et al. [1] incorporated local grey value information from images, normalized cross correlation of neighborhood grey values between points, into shape context.

A first truly 3D shape context was built by Kortgen et al [11] for measuring 3D shape similarity. Urschler et al. [1, 12, 13] extended 2D shape context to 3D and firstly applied it in image registration area for lung surfaces and thoracic registration from CT images. Xiao et al. [14] added surface curvature information from shapes to improve the performance of 3D shape context based surface registration.

*Di.Xiao@csiro.au; phone 61 2 97179893; http://www.aehrc.com/biomedical_imaging/
One of the shape context limitations for point matching is the significant amount of point mismatching that may occur between two objects because of global and local dissimilarity and existence of outliers. The point mismatching issue was pointed out in [15] for 2D object recognition and also existed in 3D shape context point matching [12]. In traditional 3D shape context based registration, a straightforward solution for point mismatching is to remove a percentage of correspondences with the highest cost [1, 12, 14]. In this study, we try to identify and correct the wrong correspondences using criteria other than the costs. The point mismatches are classified into two types: long geodesic distance mismatch and neighbors crossing mismatch. We propose to first correct the long geodesic distance mismatches, before applying a smoothing filter to correct the neighboring mismatches. We tested the improved 3D shape context method on phantoms and applied it to rat and mouse hind limb skeletons registration.

2. METHOD

2.1 3D shape context based non-rigid registration

The shape context of a point is a measure of the distribution of relative positions of neighboring points. The distribution is defined as a joint histogram where each histogram axis represents a parameter in a polar coordinate system. Bins around each point are spatially defined and the number of the neighboring points in these bins is assimilated as a context to the point. In 2D shape context, a point histogram is built based on a 2D polar coordinate system. In a 3D shape context, a point histogram is commonly built based on a 3D spherical coordinate system (radial coordinate, azimuth and elevation). Along the radial direction, bins are arranged uniformly in log-polar space increasing the importance of nearby points with respect to points farther away. If there are \( i \) bins for the radius, \( j \) bins for the azimuth and \( k \) bins for the elevation, there are \( i \times j \times k = L \) bins for the 3D shape context in total. For any 2 points, \( m \) belonging to point set 1 (shape 1) and \( n \) belonging to point set 2 (shape 2), their shape contexts can be expressed as histograms \( h_m(l) \) and \( h_n(l) \), where \( l = 1, 2, ..., L \) is the number of bins. A measure of similarity between the two shape contexts can be computed as a cost \( C_{mn} \) between the two histograms expressed by using the \( \chi^2 \)-distance:

\[
C_{mn} = \frac{1}{2} \sum_{l=1}^{L} \frac{(h_m(l) - h_n(l))^2}{h_m(l) + h_n(l)}
\]

Therefore, a low cost value for two shape contexts means a high similarity between the two points. As the absence of Cartesian coordinates implies in Eq. (1), shape context is translation invariant. Scale invariance can be obtained by changing the bin radius adaptively according to each shape’s mean point distance. If the two point sets are similar with small rotation, as in our application, quasi rotation invariance is also obtained. Otherwise, a shape should be adjusted by a pre-computed rotation factor before calculating its shape contexts.

For surface registration, the shape context method on one pair of points is applied to a pair of point sets from two similar surfaces (for example shape 1 and shape 2) and applied to find point correspondences between the two sets of points. Associated with each correspondence between a point in shape 1 and a point in shape 2 is a cost, defined in Eq. (1). For each point in shape 1, there are corresponding costs with all points in shape 2. Considering all the points from shape 1, one can define a cost matrix between the two shapes. An ideal set of correspondences would minimize the sum of corresponding costs as this is the set of points in shape 1 that best resembles the associated set of points in shape 2. This can be expressed as follows:

\[
H = \min \sum_{m,n} C_{mn}
\]

The solution of Eq. (2) is the one with minimum summed cost found from the cost summations from all possible matched point pairs. A Hungarian algorithm [16], which is a shortest augmenting path algorithm for the linear assignment problem, is used here to minimize this cost function and enforce a one-to-one point matching. In tradition, a fixed percentage of correspondences with low cost are selected as the point sets with high confidence, the remaining points are removed. Therefore, only matched points with high similarity are chosen as final correspondences.

The final correspondences are used in a subsequent non-rigid registration step as landmarks. The landmarks are from the uniform points by mesh decimation method from their original high-density meshes (surfaces), therefore being the subsets of the entire sets of points on the two shapes. After finding the point correspondences between the two sets of
landmarks (named source landmarks and target landmarks), the mapping of the high-density points from the source shape to the target shape is modeled with a thin-plate spline (TPS) [17] to enforce a smooth deformation. The TPS model, describing a transformed 3D point \((x', y', z')\) independently as a function of a source point \((x, y, z)\), has the form:

\[
f(x, y, z) = a_1x + a_2y + a_3z + \sum_{i=1}^{N} w_i U\left(\left\| P_i - (x, y, z) \right\|\right) \tag{3}
\]

where \(U(r) = r^2 \log r\) is the basis function and \(r = \|\) Euclidean distance between two points. \(P_i\) \((i = 1, \ldots, N)\) are assigned by the source landmarks in the non-rigid registration method. \(N\) is the number of the source landmarks. \(a = \{a_1, a_2, a_3, a_4\}\) is the global affine coefficients of the transformation and \(w = \{w_1, \ldots, w_N\}\) the coefficients of the additional non-linear deformation. The corresponding source and target landmarks are used to compute the coefficients of the TPS function by minimizing its bending energy, which is typically in TPS [17] the integral of the squares of the second-order derivatives of Eq. (3). After obtaining the coefficients, a non-rigid mapping function is generated between the source and target points. Therefore, a non-rigid warping can be applied to the high-density source points and registered surface can be obtained with a smooth deformation field.

### 2.2 Issues in point mismatching in shape context model

In practice, point mismatching is often happened while applying 3D shape context method. Large dissimilarity of shapes is one reason of the point mismatching. Another reason is that the search algorithm for bipartite graph matching guaranteed to find a global minimum. Two kinds of point mismatches were identified in this paper and defined as follows:

1. **Long geodesic distance mismatch (LGDM):** a point from a source shape is matched to a wrong point (in a target shape), which has a large geodesic distance from the correct point (e.g. points A and B in Fig. 1) supposed to be matched. This kind of mismatch can cause a long or short Euclidean distance surface stretching after implanting surface warping.

2. **Neighbors crossing mismatch (NCM):** in a correspondences field of two point sets, there exist some local crossing correspondences because of point mismatching among neighboring points (Fig. 2). The local crossing correspondences can cause local surface folding and stretching after implementing surface warping. NCM usually affects the accuracy of the registered surface.

We propose topological structure correction method and smoothing correspondences method for correcting those issues as described in the following two sections.

### 2.3 Topological structure correction (TSC) method

In a triangulated surface (mesh), a triangle is a fundamental cell of the surface as shown in Fig. 3. Any point in the surface has a specific location, which has both spatial and topological relationship with its neighboring points. 3D shape context model only uses the spatial position information of points from a mesh regardless of the topological constraints of the points. We propose a method of preserving the topological constraints of meshes in shape context model and using geodesic distance to identify the long geodesic distance mismatching points to solve the LGDM issue.
When the points from shape 1 are matched to the points in shape 2, the topology in shape 1 is preserved. This preserved topological structure and the matched points in shape 2 will generate a new mesh. For each corresponding point, both on its original mesh (shape 1) and new mesh, a geodesic distance pair between the point and all its topological neighboring points (as point A and its neighbors shown in Fig. 3) are computed as:

\[
\begin{align*}
(d_O(A, j), & \quad j = 1, \ldots, n \\
(d_N(A, j), & \quad j = 1, \ldots, n)
\end{align*}
\]

(4)

where \(d_o(A, j)\) is the geodesic distance between point \(A\) and its neighboring point \(j\) (a one-ring selection of the neighbors is used here) on the original mesh and \(d_N(A, j)\) the corresponding geodesic distance on the new mesh.

For example, if \(j = 8\), there will be 8 geodesic distance pairs. For each corresponding \(d_O(A, j)\) and \(d_N(A, j)\), if they fulfill the condition

\[
|d_N(A, j) - d_O(A, j)| > d_O \cdot \alpha, \quad j = 1, \ldots, n
\]

(5)

the current \(d_N(A, j)\) is considered as long geodesic distance change with its specific neighboring point \(j\) compared to its original geodesic distance \(d_o(A, j)\). \(\alpha\) is a scaling parameter determined from our experiments. The point \(A\) in the new mesh is classified as a LGDM point, if the number of the long geodesic distance changes as a percentage of the total number of neighboring points is larger than a threshold \(\beta\). All identified LGDM correspondences are removed from the correspondence field. The method is effective for single point or sparse points mismatching in a homogeneous correspondence field, while its neighborhood points keep unchanged with correct matching.

2.4 Correspondence field smoothing (CFS) method

Neighbors crossing mismatches may contain some local topological mismatches. From a vector field point of view, a field of crossing correspondences is a small patch of inhomogeneous vectors in a relatively large homogeneous vector field. Therefore, a correspondence field smoothing approach is proposed for the NCM correction. A correspondence field (vector image) is generated from the sparse correspondences using thin-plate spline interpolation. A median smoothing is applied for correcting the correspondence field. The kernel size of the smoothing filter will be determined experimentally in order to remove the crossing correspondences without affecting the remaining homogenous field substantially. The approach includes the following steps:

- An empty uniform vector image with specific pixel resolutions and image dimensions is generated from the bounding box of the source mesh.
- The correspondences from the source points to the matched target points are obtained by 3D shape context method with TSC correction and assigned to the corresponding positions in the vector image as correspondence vectors.
- A TPS transformation is computed from the available correspondences. All other pixels’ vectors are computed by the TPS transformation for performing vector field interpolation.
- A median smoothing filter is applied to process the vector field and a smooth correspondence vector image is obtained.
- The positions of the wrong points are replaced by the ones from the smoothed vector field. A new corrected smooth correspondence field is finally obtained.

3. EXPERIMENTS AND DISCUSSIONS

3.1 Tests of TSC and CFS methods

A common technique for eliminating the mismatching points in a traditional shape context model in 2D applications consist of removing correspondences with high cost values and labeling those points as “dummy” nodes [3]. Typically the 20% to 25% correspondences with the highest cost are eliminated from the total correspondences. In 3D, the same method has been applied [12]. This method is referred as standard (or traditional) error correction in the traditional shape context method in this paper. A series of tests were designed for validating the performance of our proposed TSC and CFS methods using rat hind limb skeletons. In the experiments, the method in [18] and a surface measurement technique...
[19] were used to calculate MAD error (mean absolute distance of all sampled points) and face RMS error (root mean square error of all the subdivided triangular faces) between two surfaces.

Fig. 4. Performance of topological structure correction method ((a)-(c) two rat hind limb skeletons and their correspondences; (d) correspondences after keeping the 76% lowest cost correspondences; (e) correspondences after applying TSC method; (f) registration result from high cost removal method (LGDMs pointed by arrows); (g) registration result from TSC method). (Please refer to the color picture of the paper on CD-ROM from proceedings volumes).

Fig. 5. Performance of CFS method by observing registered surfaces ((a) and (c): registered surface without applying CFS; (b) and (d): registered surface by applying CFS). (Please refer to the color picture of the paper on CD-ROM from proceedings volumes).

Fig. 4(a) shows two aligned rat hind limb skeletons (decimated meshes) before point matching. Fig. 4(b) shows a full point matching result and Fig. 4(c) shows a separate correspondence field. A few obvious mismatching correspondences are visible. Fig. 4(d) shows the result after keeping the 76% lowest cost correspondences. Fig. 4(e) shows the resulting correspondences after applying TSC method. Typically, the standard correction method could eliminate most of topologically mismatched correspondences and create a good correspondence field; however the TSC method keeps a more evenly distributed correspondence field with much fewer errors left because of identifying the LGDMs and removing them efficiently. As shown in another test (Fig. 4(f) and 11(g)), for a pair of hind limbs (1000 points each), with the standard correction (778 points preserved), there were 3 LGDMs (with short Euclidean distance) remaining (Fig. 4(f) pointed by arrows). Most of the high cost correspondences were present in the right calcaneus. Removing the 24% highest cost correspondences removed the correspondences in the right calcaneus, therefore causing the 3 LGDMs
(with relatively low costs) left in the pubis. As a result, the right calcaneus and the pubic bones were not well registered. By using TSC method, much better matching was obtained as shown in Fig. 4(g).

Fig. 5 demonstrates performance of CFS method by applying it to register two rat hind limb skeletons. Fig. 5(a) (zoomed in Fig. 5(c)) is a result without applying CFS method, and Fig. 5(b) (zoomed in Fig. 5(d)) a result with applying CFS method. Compared to Fig. 5(c), at the background of the target surface (light black), the registered surface (green color) in Fig. 5(d) preserves the detail well, and avoids local stretching and folding.

### 3.2 Validation of the improved 3D shape context method

(a) Correspondence field generated.
(b) Correspondence field after adding 8 LGDMs and 140 NCMs.
(c) Resulting correspondence field after 20% highest cost removal.
(d) Resulting correspondence field after TSC and CFS corrections.

Fig. 6. Comparison of topological structure correction (TSC) method and high cost removal method for a LGDM phantom.

Fig. 7. Statistical results of registration errors from 4 methods (100% cost preservation, 80% low cost preservation, CFS method, and TSC plus CFS method).

Fig. 8. Rat hind limb skeletons registered results (green skeletons from (a) to (d)) to the study data (light black skeletons from (a) to (d)) from improved 3D shape context method. (Please refer to the color picture of the paper on CD-ROM from proceedings volumes).
In order to quantitatively evaluate the performance of the improved 3D shape context method (TSC plus CFS) on small animal skeletons, an experiment was designed to test the general performance of the method in some correspondence fields simulating real situation. In the experiment, a rat hind limb skeleton (decimated to 1000 points) was selected as shape 1, and shape 2 was generated by firstly decimating the original high-density mesh to a 1050-point mesh and secondly applying translation (10 mm) and rotation (15 degree) operation on it. Then a correspondence field (as shown in Fig. 6(a)) between the two shapes was generated by performing a 3D shape context method without any cost removal. The new generated correspondence field had realistic cost values ranging from 0.01 to 0.35. A series of phantoms were generated by adding LGDMs and NCMs at random position (one phantom shown in Fig. 6(b)). In these phantoms, some cost values of the newly added LGDMs and NCMs may be less than that of the existing correspondences, thereby simulating a real situation of correspondence field. Registration errors were calculated for comparing four methods: 1) 100% correspondence preservation; 2) 80% lowest cost correspondence preservation; 3) only applying CFS method and 4) applying TSC and CFS methods together.

Fig. 6(c) and 6(d) show the resulting correspondence fields after applying the traditional correction and improved (TSC plus CFS) correction methods. Fig. 7 illustrates the statistical results of the registration errors for the four methods. The results represent only applying CFS method is not enough for removing LGDMs, therefore the registration error is still large. The 80% lowest cost correspondence preservation can effectively remove most LGDMs to reach good registration result. However, the improved 3D shape context method (TSC and CFS methods) gets better registration result compared to other three methods.

In the above two experiments, the settings of two parameters $\alpha$ and $\beta$ of TSC method were also tested by assessing the removed correspondences and computing the registration errors. The range of $\alpha$ was set from 0.4 to 0.8 and that of $\beta$ from 0.3 to 0.7. $\alpha=0.6$ and $\beta=0.55$ were chosen as optimal values by the experiments and used throughout this paper.

Table 1. MAD and face RMS errors from high-density surfaces from rat hind limb skeletons registration. Bold are best results.

<table>
<thead>
<tr>
<th>Errors (mm)</th>
<th>HL1</th>
<th>HL2</th>
<th>HL3</th>
<th>HL4</th>
<th>HL5</th>
<th>HL6</th>
<th>HL7</th>
<th>HL8</th>
<th>HL9</th>
<th>HL10</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICP alignment</td>
<td>MAD</td>
<td>1.01</td>
<td>1.03</td>
<td>1.06</td>
<td>1.14</td>
<td>1.71</td>
<td>1.04</td>
<td>1.01</td>
<td>1.55</td>
<td>1.23</td>
</tr>
<tr>
<td>Face RMS</td>
<td>1.35</td>
<td>1.39</td>
<td>1.40</td>
<td>1.71</td>
<td>2.52</td>
<td>1.40</td>
<td>1.38</td>
<td>2.11</td>
<td>1.92</td>
<td>2.01</td>
</tr>
<tr>
<td>Traditional shape context</td>
<td>MAD</td>
<td>0.44</td>
<td>0.29</td>
<td>0.30</td>
<td>0.37</td>
<td>0.48</td>
<td>0.29</td>
<td>0.27</td>
<td>0.57</td>
<td>0.55</td>
</tr>
<tr>
<td>Face RMS</td>
<td>0.91</td>
<td>0.46</td>
<td>0.43</td>
<td>0.64</td>
<td>0.74</td>
<td>0.40</td>
<td>0.36</td>
<td>1.03</td>
<td>0.85</td>
<td>0.72</td>
</tr>
<tr>
<td>Improved shape context</td>
<td>MAD</td>
<td>0.31</td>
<td>0.23</td>
<td>0.25</td>
<td>0.31</td>
<td>0.47</td>
<td>0.27</td>
<td>0.25</td>
<td>0.45</td>
<td>0.49</td>
</tr>
<tr>
<td>Face RMS</td>
<td>0.43</td>
<td>0.31</td>
<td>0.33</td>
<td>0.45</td>
<td>0.66</td>
<td>0.36</td>
<td>0.34</td>
<td>0.75</td>
<td>0.81</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Table 2. MAD and face RMS errors of mouse hind limb skeletons registration.

<table>
<thead>
<tr>
<th>Errors (mm)</th>
<th>Mouse 1 (M1): registered by its first scan</th>
<th>Mouse 2 (M2): registered by its first scan</th>
<th>Mouse 3 (M3): registered by its first scan</th>
<th>Mouse 1 registered by M3 first scan</th>
<th>Mouse 2 registered by M1 first scan</th>
<th>Mouse 3 registered by M2 first scan</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scan 2</td>
<td>MAD</td>
<td>0.16</td>
<td>0.17</td>
<td>0.14</td>
<td>0.15</td>
<td>0.19</td>
</tr>
<tr>
<td>Scan 3</td>
<td>MAD</td>
<td>0.18</td>
<td>0.13</td>
<td>0.20</td>
<td>0.13</td>
<td>0.17</td>
</tr>
<tr>
<td>Scan 4</td>
<td>MAD</td>
<td>0.14</td>
<td>0.15</td>
<td>0.14</td>
<td>0.26</td>
<td>0.15</td>
</tr>
</tbody>
</table>

A practical experiment was designed to compare the performances of the following different techniques: ICP alignment, traditional shape context with cost removal and our improved shape context. 10 rat hind limbs (HLs) (4 rats at different time points) were selected as testing data. Typical results from our method are shown in Fig. 8(a) to 8(d). We evaluated each method by: (1) observing the registered surfaces; and (2) computing errors from high-density sample points on the
surfaces. Mismatching correspondences removal applied in our improved method included three steps: (1) applying topological structure correction method with $\alpha = 0.6$ and $\beta = 0.55$; (2) implementing high cost removal method with 98% correspondences preservation; and (3) applying correspondence field smoothing method. Iterative procedure (iterative number equal to 2) was applied for both the traditional shape context method and the improved method.

For the traditional method and the improved method, 10 hind limbs were registered well (as femur and fibula parts; even for relatively larger different posture) with only the calcaneus and metatarsus region having local distortions on two studies. The maximum errors occurred at the calcaneus for most cases. After visual assessment of the registered surfaces, all the registration results were found acceptable. A further calculation of the registered errors (MAD and face RMS) on the high-density meshes was made. Table 1 shows that our proposed correction scheme provided lower errors in most cases. HL10 has large registration errors which were mainly from the metatarsus parts and due to a relative large difference in the animal postures.

Fig. 9. Some results of mouse hind limb skeletons registration (red-source skeleton; yellow-target skeleton; blue-registered skeleton). (Please refer to the color picture of the paper on CD-ROM from proceedings volumes).

The proposed method was also applied for mouse hind limb skeleton registration. 3 mice were scanned by CT. With a defined protocol, each mouse with similar posture underwent a CT scan once every 4 days. 12 mice hind limb skeletons (3 mice x 4 time points) were chosen for the test. The skeletons from all mice’s first CT scans were used as the references to register to the selected skeletons (listed in Table 2). For each registration, we used $\alpha = 0.6$, $\beta = 0.55$ and 98% lowest cost preservation. 800 landmarks and 4 iterations of registration. The registration results were evaluated by MAD and face RMS and reported in Table 2. The registration errors were small for all subjects based on the resolution of our CT images (0.22mm x 0.22mm x 0.22mm). All the registration results were visually acceptable.

Fig. 9 shows two examples of registration results. Errors from Table 2 show that the proposed method provided the good performances for the mouse hind limb skeletons registration.

4. CONCLUSIONS

We propose an improved 3D shape context model with point mismatching correction. Combined with TPS transformation, a non-rigid registration method was described. The method was applied and tested on phantoms, rat hind limbs and mouse hind limbs registration. Experimental results demonstrated good performance and robustness of the proposed method to mismatching errors inherent with shape context technique. The experimental results on rat hind limb skeletons and mouse hind limb skeletons registration demonstrated the application of the method for small animal non-rigid registration.

REFERENCES