Shape Registration via the Wavelet Transform

Julie S. Chalfant∗
Massachusetts Institute of Technology

Nicholas M. Patrikalakis†
Massachusetts Institute of Technology

Abstract

We present a system for the automatic global registration of voxelized data using the wavelet transform. The huge size of current data sets makes such multi-resolution techniques appealing; however, the traditional discrete wavelet transform has the drawback of variance with translation and rotation. We employ the dual-tree complex wavelet transform (Kingsbury, 2003), which carefully adds near invariance to translation and reduced sensitivity to rotation in computationally efficient ways, thus enabling the object matching application. We employ the wavelet transform to accomplish registration by extracting features in a multi-resolutional format, thus rapidly detecting and matching feature points. The method is robust to occlusion, clutter and noise. The efficacy of the algorithm is demonstrated through examples from solid modeling and biomedical applications.

Keywords: Three-Dimensional Object Registration, Shape-Based Matching, Wavelets

Index Terms: I.5.2 [Pattern Recognition]: Design Methodology—Feature evaluation and selection

1 Introduction

We present an algorithm that employs the wavelet transform to accomplish global, rigid, shape-based registration of three-dimensional voxelized objects which have undergone translation and rotation. The method accommodates noise, which can be induced in models through events such as the sensing process, transformation and voxelization.

An overview of the algorithm is as follows:

1. Apply a wavelet transform to both objects. We select the Dual-Tree Complex Wavelet Transform designed by Kingsbury [2, 3, 4]. This wavelet transform achieves approximate shift invariance by omitting the downsampling at the first level and interleaving samples at lower levels. It achieves reduced sensitivity to rotation by increasing the number of subbands of information at each decomposition level.

2. Select the large magnitude wavelet coefficients in each model as feature points. The wavelet coefficients provide more distinct, stable and repeatable feature points than the approximation coefficients. See Figure 1.

3. Search within the lists of feature points for matching triplets of points. After eliminating collinear points and small triangles, match triangles with the same angles and leg lengths within a tolerance.

4. For each matching set of triplets, determine the rotation required to match the triangles. Use a histogram voting scheme to select the best rotation. Using the selected rotation, determine the translation.

2 Examples and Applications

In this section we show examples from a variety of applications. Note that no tuning is required: all examples were run using the second and third resolution wavelet coefficients with 30 feature points selected. The voting scheme is a four dimensional histogram with bin size of 0.02 for each quaternion coefficient. Examples were run on a Core 2 Duo machine.

Non-Homogeneous Objects. These objects are fully three dimensional, varying in intensity throughout the object. Both examples shown here are Magnetic Resonance Imaging (MRI) data. However, this description fits many types of data such as Computed Tomography (CT), Ultrasound (US), and functional MRI (fMRI).

The first example is an MRI scan of a brain [5]; see Figure 2. The original model was rotated 28, 37 and -34 degrees in a pitch-roll-yaw Euler angle scheme, then translated -14, 12 and -3 voxels to create the second model. Total elapsed time for registration was 63 seconds.

The second example is an MRI scan of Paul Debevec’s knee, which is available online courtesy of a torn ligament [1]. The original object was rotated by pitch-roll-yaw Euler angles of -20, 4, and
Table 1: Registration results for various models. Angles are the pitch, roll and yaw Euler angles, $\theta$, $\phi$ and $\psi$, required for registration as determined by the algorithm. $\Delta_r$ is a pseudo-Manhattan distance which is the sum of the magnitude of the distance between each recommended angle and the inverse of the induced rotation. Vote is the number of votes received in the averaged shifted histogram voting scheme. Similarly, the translation is the number of voxels in the i, j and k direction required for registration as determined by the algorithm and $\Delta_t$ is the Manhattan distance between the recommended translation and the inverse of the induced translation. Results are rounded to the nearest degree or voxel for presentation. Time is elapsed time on a Core 2 Duo machine.

<table>
<thead>
<tr>
<th>Model</th>
<th>$\theta$</th>
<th>$\phi$</th>
<th>$\psi$</th>
<th>$\Delta_r$</th>
<th>Vote</th>
<th>$\Delta_t$</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>$\Delta_t$</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knee MRI</td>
<td>20</td>
<td>-4</td>
<td>-40</td>
<td>1</td>
<td>5582</td>
<td>-12</td>
<td>-9</td>
<td>-10</td>
<td>2</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>Brain MRI</td>
<td>-27</td>
<td>-36</td>
<td>33</td>
<td>4</td>
<td>297</td>
<td>14</td>
<td>-13</td>
<td>2</td>
<td>2</td>
<td>63</td>
<td></td>
</tr>
<tr>
<td>Full Shrek</td>
<td>-12</td>
<td>-14</td>
<td>14</td>
<td>1</td>
<td>802</td>
<td>-12</td>
<td>3</td>
<td>-10</td>
<td>2</td>
<td>262</td>
<td></td>
</tr>
<tr>
<td>Cropped Shrek</td>
<td>-10</td>
<td>-21</td>
<td>-30</td>
<td>1</td>
<td>820</td>
<td>13</td>
<td>-7</td>
<td>4</td>
<td>4</td>
<td>67</td>
<td></td>
</tr>
</tbody>
</table>

41 degrees, then translated 13, 8 and 11 voxels in the i, j and k directions respectively. Registration was performed in a total elapsed time of 44 seconds. Results are shown in Table 1.

Homogeneous Objects. These objects are solid models which do not vary in the interior. Two tests were run on a 3D model of Shrek [8], voxelized using Binvox [6, 7]. The first was a registration of the full model and the second a cropped portion of the model.

The first test, using the full Shrek model with an induced rotation in pitch-roll-yaw Euler angles of 12, 15 and -14 degrees and an induced translation of 11, -3 and 11 voxels, had an elapsed time of 262 seconds. The second test using the cropped model had an elapsed registration time of 67 seconds. See Figure 3. Results are shown in Table 1.

ACKNOWLEDGEMENTS

This work was supported in part by the National Defense Science and Engineering Graduate (NDSEG) Fellowship to the first author, by U.S. National Science Foundation grant DMI-062933 and by the SMART Center for Environmental Sensing and Modeling (CEN-SAM) sponsored by the Singapore National Research Foundation.

REFERENCES