ABSTRACT

Examining the growth rate of pleural thickenings in consecutive 3D-CT images requires the matching of identical thickenings in lung images acquired at two different points in time. The thickenings can be subject to strong deformations caused by their growth. This implies that position information should play a major role in finding correspondences. Here, a MGRF approach is presented to determine a rigid transformation. It aligns the lung volumes by maximizing the probability of the regarded lung tissue to fit an offline trained model. To ensure a symmetrical matching of lung surfaces this probability is calculated reciprocally. Using precalculation, strong subsampling and a multiscale approach, the required time can be reduced by a factor of about 80, depending on the image resolution. Due to this speed-up, online follow-up assessment is feasible. We show that this approach results in precise registration which can be used for a reliable matching of lung thickenings.

Index Terms—Markov-Gibbs random field, registration, multiscale, pleuramesothelioma, lung, CT

1. INTRODUCTION

Asbestos is known to be a human carcinogen substance, which was widely used until the 1990s. When inhaling the 3-5 µm long persistent fibers, they can cause pleural thickenings shown in fig. 4, and subsequently pleuramesothelioma. The long latency of 14-72 years leads to an expected maximum of diagnosed cases in the year 2018 [1]. High-risk patients undergo a preventive check-up including 3D CT imaging. The medical diagnosis of the resulting volume data is highly time-consuming and subject to strong inter- and intra-reader variances. Algorithms to segment the lung, to detect thickenings and subsequently for classifying the thickenings were developed to realize a more objective assessment. Follow-up CT scans allow the monitoring of the thickenings.

The detection and analysis of the thickenings are performed fully automatically [2]. Up to now, the follow-up assessment requires the manual specification of anatomical fix points. These points are used to calculate a transformation matrix, which is utilized to match the thickenings. In this paper we present a fully automatic solution to align the lung images.

Among image registration methods for medical images, point-feature based registration methods need the manual selection of anatomical features or use automatically determined features. N-SIFT [3], e.g., is a very fast feature extractor, but not suitable to focus on selected image regions such as the lung. The approach from [4] is well suited for the brain surface and might be adaptable to lung surfaces, but the computational expense is very high. Conventional intensity based approaches relying on entropy [5], [6], correlation or differences could be optimized for fast computation, but are not appropriate for the lung surface. The Markov-Gibbs random field (MGRF) technique [7] is well suited for the lung surface, but the computational expense is too high and a symmetric optimization is not ensured. Here, we therefore describe how, in combination with downsampling, multiscaling and suitable precalculations, a bidirectional version of the MGRF approach allows to register pleural thickenings at a drastically reduced computation time.

2. METHODS

The objective of the algorithm is to match the detected thickenings, which are all located on the lung surface. The idea is to find a transformation $T_r$ for the lung mask from the first acquisition $t = t_0$, so that its superposition with the lung volume from the following acquisition $t = t_0 + \Delta t$ contains as much lung tissue as possible and vice versa. This approach is visualized in fig. 1. The conforming of the regarded tissue and the lung model is determined using MGRFs.

2.1. Markov-Gibbs random fields

A 3D image $g$ is a mapping of coordinates $i$ contained in the lattice $\mathcal{R} = \{i = (x, y, z) | 1 \leq x \leq X; 1 \leq y \leq Y; 1 \leq z \leq Z\}$ to gray levels $\mathcal{Q} = \{q | 1 \leq q \leq Q\}$. The voxels $g_i$ are random variables with, in the case of natural images, strong statistical dependencies. Hence, it is useful to consider the probability $Pr(g_i | g')$ for a
The energy is given by

\[ \text{voxel at the position } i \text{ having a special gray value, given the gray values of all other voxels } g^i = g \setminus g_i. \]

For acceptable calculation expenses, only a limited neighborhood \( N = \{ n \mid n \in \{(x'_1, y'_1, z'_1), \ldots, (x'_N, y'_N, z'_N)\} \} \) is considered and \( \Pr(g_i \mid g^i) = \Pr(g_i \mid g_{i+n}; n \in N) \) is assumed. A voxel at the position \( i \) and \( M \) of its neighbors \( i + n \), \( n \in N \) form a clique \( C \) of \( M \)-th order. In accordance to the Hammersley-Clifford theorem \([8]\) the described Markov random field can be interpreted as a Markov-Gibbs random field. The probability for a voxel \( g_i \) is therefore given by

\[
\Pr(g_i \mid g^i) = \frac{1}{Z_i} \exp \left( - \sum_{C \in \mathcal{C}} V_C(g_i, g_j : i, j \in C; i \neq j) \right),
\]

where \( \mathcal{C} \) is the set of all cliques and \( V_C(i, j) \) are the Gibbs potentials modeling the statistical dependency of the voxels in one clique. Only second-order cliques are considered here. These cliques are divided into clique families \( C_n(R) = \{ c \mid c = \{ i, i+n \}; i \in R \} \), which contain all cliques for the neighbor \( n \). Rather than for every single clique, the potentials are defined for every clique family and gray-level pair \( (q, q') \), denoted by \( V_{n,q,q'}(R) \). The relative frequency of occurring gray-level pairs \( (q, q') \) for a neighbor \( n \) is given by \( f_{n,q,q'}(R) \). The probability for an image volume to fit the assumed model is

\[
\Pr(g(R)) = \frac{1}{Z} \exp \left( -E(g(R)) \right).
\]

The energy is given by

\[
E(g(R)) = \sum_{n \in N} E_n(g(R))
\]

and simply sums up the neighbor energies

\[
E_n(g(R)) = \sum_{(q,q') \in \mathbb{Q}^2} |R| V_{n,q,q'} f_{n,q,q'}(g(R)).
\]

### 2.2. Offline training

The potentials \( V_{n,q,q'} \) for known lung tissue \( \hat{g}(R) \), below denoted by \( \hat{g} \), can be determined analytically. An approximation of the log-likelihood of eq. 2 is maximized and results in \([9]\)

\[
V_{n,q,q'} = \frac{Q^2}{(Q - 1)} \left( \frac{Q - 1}{Q} - f_{n,q,q'}(\hat{g}) \right).
\]

The values of the potential in eq. (5) are high for gray-level pairs and neighbor combinations occurring often and low for combinations appearing rather seldom.

The next step identifies neighbors \( n \), which result in high energies \( E_n \) on lung tissue \( \hat{g}(R_L) \), denoted by \( \hat{g}_L \), and in low energies on non-lung tissue \( \hat{g}(R_N) \), denoted by \( \hat{g}_N \). Classifying neighbors using a histogram of \( E_n(\hat{g}_L) \) and \( E_n(\hat{g}_N) \) as suggested in \([7]\) does not work in all cases. It fails, when the potentials trained on lung tissue result in higher energy on non-lung tissue than on lung tissue. As a solution we choose a set of neighbors producing high energy values \( E_n(\hat{g}_L) \) and simultaneously producing low energy values \( E_n(\hat{g}_N) \). Thus the \( K \) neighbors \( N_K \subseteq N \) with the highest energy differences

\[
\Delta E_n(\hat{g}_L, \hat{g}_N) = E_n(\hat{g}_L) - E_n(\hat{g}_N)
\]

are chosen.

### 2.3. Registration

To register the image data of two CT scans \( g_1(R_1) \) and \( g_2(R_2) \) from the same lung, but from two different points in time, a rigid transformation \( T_r \) is determined. The required lattices \( R_1 \) and \( R_2 \) are known from the previous thickening extraction \([10]\) and are extracted by a two-step supervised range-constrained Otsu thresholding \([2]\). The transformation is applied on the lattice \( R_1 \) and results in a lattice for the volume image \( g_2 \). The probability \( \Pr(g_1(Tr(R_{L,2}))) \) is calculated using eq. 3. The final transformation is determined by

\[
T_r^* = \arg \max_{T_r} \Pr(g_1(Tr(R_{L,2})))|Tr|.
\]

The volume \( R_{L,2} \) should not be shrunk excessively by \( T_r \) to cover solely the lung tissue in \( g_1 \). Vice versa, to symmetrize the process, the lung tissue in \( g_2 \) covered by \( T_r \) should be considered. Thus, we maximize \( \Pr(g_1(Tr(R_{L,2}))) \cdot \Pr(g_2(Tr^{-1}(R_{L,1}))) \), where \( Tr^{-1} \) denotes the inverse transformation of \( T_r \). For this purpose we introduce the so called bidirectional Gibbs energy

\[
E_b(g_1, g_2, Tr) = E(g_1(Tr(R_{L,2}))) + E(g_2(Tr^{-1}(R_{L,1}))),
\]

which can be maximized.

The translation, scaling and rotation are optimized by using a gradient search algorithm. However, this class of algorithms is sensitive to local maxima. Therefore, a multiscale approach is applied to ensure the localization of the global
maximum. Using a scaling factor $s$ the lattice $\mathbf{R}$ is reduced to a sub-sampled lattice $\mathbf{R}^s = \{i = (s^{-1}x, s^{-1}y, s^{-1}z) \mid 1 \leq x \leq sX; 1 \leq y \leq sY; 1 \leq z \leq sZ\} \subseteq \mathbf{R}$. The neighbors $N^s_k$ and the associated potentials are determined offline for each scaling level. Choosing a random subset $\mathbf{R}^{s*} \subseteq \mathbf{R}^s$ containing $W$ voxels leads to a considerable speed-up, when calculating $E_b(\mathbf{g}_1, \mathbf{g}_2, \mathbf{Tr})$.

2.4. Matching of thickenings

Each thickening $h_1 \in H_1$ is matched with the best fitting one $h_2 \in H_2$, where $H_1$ and $H_2$ are the sets of all thickenings at the particular point in time $t_0$ and $t_0 + \Delta t$. The Euclidean distance $d(h_1, h_2, \mathbf{Tr})$ of their registered coordinates and the difference $\Delta(k, l)$ of the mean CT number in HU are combined in a cost function with the parameter $w$. This function is minimized. The correspondences are obtained by solving

$$\arg \min_{h_1 \in H_1, h_2 \in H_2} \left( d(h_1, h_2, \mathbf{Tr}) + w ||\Delta(h_1, h_2)|| \right), \quad (9)$$

where $w = 10$ was empirically determined.

2.5. Image data

For training purposes nine different CT scans from five different patients were used. The registration itself was performed on eight of these CT scans, which were taken from four patients at two different points in time. A leave-one-out validation was applied to avoid overfitting of the model. Random rigid transformations were applied on the image volumes, which increases the number of registration cases. For a more realistic scenario these transformations include not only translation, scaling and rotation, but also include random shearing, which is not explicitly accounted for the registration model. Ten different cases are generated for each patient.

3. RESULTS

All registrations were performed using three scale levels. The relative non-overlapping volume of the masks $G_W = \frac{1}{2} (\frac{|R_{L,1} \setminus \text{Tr}(R_{L,2})|}{|R_{L,1}|} + \frac{|\text{Tr}(R_{L,2}) \setminus R_{L,1}|}{|R_{L,2}|})$ is used as a quality indicator while regarding the calculation time $t_W$, both depending on the number of used voxels $W$. Furthermore the resulting error $G_{\text{full}}$ and required time $t_{\text{full}}$, when using the complete set of voxels $\mathbf{R}^{s*} = \mathbf{R}^s$, and the error before registration are considered independently of $W$. In fig. 2 the error and computation time are visualized for one patient. Qualitatively this result is identical for the patients 2 and 3. Table 1 presents the results only for the fixed number of $W = 5000$ voxels compared to the results, when using all voxels. In addition the available number of CT slices is given for each patient.

A visualization of the errors $R_{L,1} \setminus \text{Tr}(R_{L,2})$ (light) and $R_{L,2} \setminus \text{Tr}(R_{L,1})$ (dark) using 5000 voxels is shown in fig. 3.

![Fig. 2. Registration error $G$ and computation time $t$ using $W$ voxels, patient 1](image)

### Table 1. Registration error $G$ and computation time $t$

<table>
<thead>
<tr>
<th>patient</th>
<th>slices</th>
<th>$t_{5000}$ [s]</th>
<th>$t_{\text{full}}$ [s]</th>
<th>$G_{5000}$</th>
<th>$G_{\text{full}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>695</td>
<td>2.55</td>
<td>236.72</td>
<td>4.15%</td>
<td>4.03%</td>
</tr>
<tr>
<td>2</td>
<td>397</td>
<td>1.97</td>
<td>171.18</td>
<td>5.31%</td>
<td>5.23%</td>
</tr>
<tr>
<td>3</td>
<td>411</td>
<td>1.76</td>
<td>140.93</td>
<td>5.56%</td>
<td>5.39%</td>
</tr>
<tr>
<td>4</td>
<td>57</td>
<td>1.50</td>
<td>30.93</td>
<td>16.88%</td>
<td>37.78%</td>
</tr>
</tbody>
</table>

Table 2. Evaluation by physician

<table>
<thead>
<tr>
<th>patient</th>
<th># thickenings</th>
<th>identical</th>
<th>matched</th>
<th>wrong</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13</td>
<td>69.23%</td>
<td>30.77%</td>
<td>0.00%</td>
</tr>
<tr>
<td>2</td>
<td>102</td>
<td>94.12%</td>
<td>5.88%</td>
<td>0.00%</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>50.00%</td>
<td>50.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

4. DISCUSSION

The registration using all voxels leads to the lowest error $G_{\text{full}}$ of about 5%. In this case the registration takes approx. 200-300 seconds. The reduced number of 1,000-9,000 voxels is a very strong subsampling compared to the full number of about 10,000,000 - 20,000,000 voxels. The error in fig. 2 decreases with an increasing number of voxels and when using at least 5,000 voxels the error is nearly the same compared to the result obtained using all voxels. The original computation time $t_{\text{full}}$ can be drastically reduced to 2-3 seconds ($t_{5000}$), without increasing the error $G$ significantly. The result for...
patient 4 differs strongly from the other ones and the registration turns out to be very unreliable. This is caused by the low number of only 57 CT slices compared to the approx. 400-700 slices in the other datasets. Transforming the binary mask volumes with such a low resolution results in artifacts. In these cases a minimal rotation generates major changes in the mask. Here the subsampling has in fact a positive effect on the registration error. This special case has no genuine relevance, because today’s CT data for the examination of pleural thickenings consists of approx. 700 slices.

While the reported error is just an indicator of the registration quality, the thickening matching is crucial for the clinical evaluation. All correspondences were evaluated to be correct. However, the ability to identify a matching pair of thickenings to be identical strongly depends on the performance of the previous thickening segmentation.

5. SUMMARY AND PROSPECTS

We described a contour and gray-level based approach using MGRF to align lung volumes from different points in time. Owing to the consequent precalculations of the characteristic neighborhood and the related potentials, together with subsampling of the image data, a fast alignment considering the lung surface could be achieved. The result of this rigid registration is well suited to match the thickenings.

The consistent segmentation of the thickening in volume images at two points in time is a challenge for future research. A solution for this also addresses a precise non-rigid registration at each thickening locally.

6. REFERENCES


