Respiratory Motion Estimation in Nuclear Medicine Imaging using a Kernel Model-based Particle Filter Framework

A. A. Abd. Rahni, Student Member, IEEE, E. Lewis, and K. Wells

Abstract—The continual improvement in spatial resolution of Nuclear Medicine (NM) scanners has made accurate compensation of patient motion increasingly important. A major source of corrupting motion in NM acquisition is due to respiration. Therefore a particle filter (PF) approach has been proposed as a powerful method for motion correction in NM. The probabilistic view of the system in the PF is seen as an advantage that considers the complexity and uncertainties in estimating respiratory motion. Tests using the XCAT phantom have previously shown the possibility of estimating unseen organ configurations using training data that only consist of a single respiratory cycle. This paper builds upon previous work in two ways: this is the first evaluation of a PF framework using clinical 4D thoracic CT data; and, this implementation uses a kernel density estimation (KDE) representation for the transition model, thus taking advantage of the PF’s ability to use a wider range of stochastic models. The results show some improvement with the use a KDE-based transition model and indicate that the PF should be applicable to clinical data. This novel approach will be more fully evaluated using more 4D datasets that better reflects the inter-patient variations in motion that may be seen in clinical practice.

I. INTRODUCTION

The most sensitive approach currently recognized for functional imaging of the human body is nuclear medicine. As such it is an important tool in oncology, especially in the early detection of cancerous lesions. Although there has been significant technological improvement in system spatial resolution and sensitivity, leading to reduced acquisition time on current NM scanners, scan times are still much longer than the period of a single breath-hold used in X-ray CT (i.e. typically 5-15 minutes for PET and 15-30 minutes for SPECT [1]). This results in an acquisition process that occurs over several breathing cycles and together with the improvement in spatial image resolution, these considerations make patient motion an increasingly important issue to be addressed especially regarding the potential issue of image blurring due to such motion [2].

Respiratory motion dominates motion artefacts when imaging the torso. This paper is based on a motion correction framework that has been proposed in [3]. In this framework, respiratory motion correction is proposed to be applied continuously, enabling the use of all potential data, for example, list mode data in NM before image reconstruction. Respiratory motion has been shown to exhibit complex behaviour [3][4]. Therefore, it seems most suitable to use a probabilistic model for motion estimation and this forms the basis of the particle filter (PF) approach. In this approach, internal organ deformation is inferred from a stereo surface capture of the anterior portion of the external surface of the torso and from previous estimates of organ deformation. This inference is based on a state transition model, which relates the configuration of organs, x_k at a discrete time index k with those at k−1 and a measurement model, which correlates the state x_k to the external observation z_k.

This paper expands upon the previous implementation of the PF approach [5] in two ways:

• This is the first evaluation of a PF framework using clinical CT data.
• This implementation uses kernel density estimation (KDE) for the transition model, thus taking advantage of the PF’s ability to use non-Gaussian models.

In this first analysis of the PF approach on patient data, external observation of marker motion is used to predict organ motion. Using CT data facilitates evaluation to higher spatial accuracy compared to that achievable with NM data.

Section (II) that follows gives an overview of the basic framework and also the KDE implementation of the transition model. Section (III) outlines the data and experiments used for evaluation. Section (IV) discusses the results and is followed by concluding remarks and notes for future work in Section (V).
II. METHODOLOGY

A. Particle Filter Implementation

Motion estimation is viewed as a Bayesian tracking problem which formulates the estimation process as a first-order Markov model. Here the state $x_k$ is a hidden random variable and only correlated to its value at the previous time step $k-1$. At each time step, the hidden state is also correlated to independent observations $z_k$. This HMM structure is shown in Fig. 1 and this generic nonlinear dynamic system has the following state space representation:

Transition probability density

$$x_k = a(x_{k-1}, v_{k-1}) \sim f(x_k \mid x_{k-1})$$ \hspace{1cm} (1)

Measurement probability density

$$z_k = b(x_k, w_k) \sim g(z_k \mid x_k).$$ \hspace{1cm} (2)

where the models $a$ and $b$ have respective stochastic components $v_{k-1}$ and $w_k$ to represent uncertainties. This gives rise to their respective probability densities $f$ and $g$.

At time $k$ the state can then be estimated from the posterior probability density $p(x_k | z_{1:k})$ given the set of all observations up until the present time $k$, $z_{1:k} = \{z_1, \ldots, z_k\}$. Some moment of the posterior can be taken as the actual state estimate, such as its expected value $E[x_k]$. The most general solution is to use a PF which can accommodate the widest range of transition and measurement models. In a PF, the posterior is estimated as

$$p(x_k \mid z_{1:k}) \approx \sum_{i=1}^{N} \omega^i_k \delta(x_k - x^i_k).$$ \hspace{1cm} (3)

where the posterior is sampled along the space of $x_k$ (i.e. possible states) by a weighted impulse train represented by point masses (i.e. particles) $x^i_k$. The particles are Monte Carlo samples of the state space and the probability of the posterior at the locations of the particles are given by their respective weights, $\omega^i_k$. As in [5], the PF is implemented as a sampling importance resampling (SIR) filter. Following the previous implementation, the following adaptation methods are also used to improve particle diversity and make the sampling of the state space more efficient:

1) Dimensionality Reduction using principal component analysis (PCA).
2) Incorporation of estimated respiratory parameters, and
3) Planned sampling of particles.

The details have been outlined in [5].

B. KDE representation of the transition model

Previously, a second-order autoregressive process, AR(2), was chosen as the transition model $a$, as it postulated that this would reasonably represent the pseudo-oscillatory nature of respiratory motion [6]. As the AR(2) model tracks organ configuration in a variable $c_k$ with the state $x_k$ consisting of configuration parameters from two time points, $c_k$ and $c_{k-1}$, the transition model $a$ is still a first-order Markov model. Consequently its corresponding probability density $f$ will be Gaussian and was made non-Gaussian only with the adaptation methods listed in subsection II(A) [5].

This however, limits the capabilities of the PF to draw on a much wider range of stochastic models. In lieu of this, $f$ is taken to be a general probability distribution by simply being the conditional density of $\gamma(y_k)$ which is represented by the summation of kernel densities located at samples $y_i$:

$$\gamma(y_k) = \frac{1}{M} \sum_{i=1}^{M} k(y_k \mid y_i)$$ \hspace{1cm} (4)

where $y_k$ is the joint variable of $x_k$ and $x_{k-1}$. A Gaussian is used as the kernel, hence $k = N(y_i, \Sigma_{y_i})$, with

$$\Sigma_{y_i} = \begin{pmatrix} \Sigma_{x_i} & 0 \\ 0 & \Sigma_{x_{i-1}} \end{pmatrix}.$$

Therefore, $f(x_k | x_{k-1}) = f(y_k | x_{k-1})$. The samples $y_i$ and covariance $\Sigma_{y_i}$ are obtained from the training dataset [7]. Additionally, a more optimal kernel covariance can be found using the plugin method which adjusts the determinant of the covariance matrix, $|\Sigma_{y_i}|$ [8].

Fig. 2 illustrates the density $\gamma(y)$ when $c_k$ is a scalar (using PCA) with the values of $\gamma$ plotted on three border planes according to the colourbar. The magenta line is a particular value of $x_{k+1}$ in such a case. With that value, the conditional density of $c_k$ can found along the line and is plotted in Fig. 3. With this conditional density, an estimate of $c_k$ can be found such as the maximum likelihood which is indicated by the magenta asterisk (*). Such an estimate can be found by the PF through Monte Carlo simulation of the transition density $f$. 

![Fig. 1. Diagram of the system in state-space form illustrating the structure as a first order hidden Markov model.](image-url)
C. Organ segmentation and tracked state

Organs of interest were obtained from the 4D CT dataset using semiautomatic segmentation based on connected component analysis of thresholded images. As the aim is for semiautomatic segmentation, moderate accuracy is accepted using the above methods and any inaccuracy is corrected manually. It is found that this procedure is still faster than full manual segmentation, in agreement to similar comparisons [9]-[11], although those applications were intended for clinical segmentation of particular structures in the body. Renders of the segmented organs for one of the datasets used are shown in Fig. 4.

The organs chosen for tracking are the major airways (trachea and largest bronchi), spine, heart, liver, lungs, ribcage, kidneys, and shoulder bones (clavicles and scapulas). The organ configurations that are tracked, $c_{k}$, are represented by piece-wise affine transformation parameters for each organ $o$ at time $k$ for registration back to a baseline configuration $\theta_{b}$ selected from the training dataset [5]. The transformation parameters for organs that are not fully within the CT scan volume (i.e. field of view, FOV) were constrained so that they were rigid. The transformation parameters are found from iterative closest point (ICP) registration as outlined in [12].

D. Observables and the measurement model

The observables used in this paper to infer organ configuration are anterior markers on the external surface of the body which are available in the 4D CT datasets used. They can be categorized into artificial and natural markers. The artificial markers are the optical markers and buckle of the belt used to monitor respiration in 4D CT acquisition. The natural markers used are the nipples, approximated by the location of the underlying mammary glands, obtained during the semiautomatic segmentation procedure in subsection II(D). Fig. 5 illustrates the location of the optical and natural markers in relation to the renders of part of the external surface of the body and segmented organs for one of the datasets used.
The measurement density \( g \) is still chosen to be Gaussian. Its generative form \( b \), is a linear map of the state with a stochastic component. The covariance of the measurement model noise, \( \Sigma_w \), accounts for the estimated inaccuracy of the map constants and observation noise. In this paper, the observation noise is associated with the process of segmenting the physical abdominal surface markers from the 4D CT dataset. The parameters for \( b \) are found from least squares estimation on a training dataset.

### III. EVALUATION

Two 4D thoracic CT datasets were used for evaluation, one obtained from the Royal Surrey County Hospital (RSCH) [12] and another from the POPI model study [13]. Both datasets consist of 10 respiratory phases, assumed to be equally spaced in time. They also have in plane resolutions of 0.98 mm and have optical markers present in the CT volumes. All 12 pairs of ribs are present but the lower ribs are truncated. The other properties of these datasets are listed in Table I.

#### TABLE I. PROPERTIES OF THE TWO DATASETS USED FOR EVALUATION

<table>
<thead>
<tr>
<th>Property</th>
<th>RSCH</th>
<th>POPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acquisition machine</td>
<td>GE LightSpeed( ^{TM} ) RT16</td>
<td>Philips Brilliance Big Bore</td>
</tr>
<tr>
<td>Slice spacing (mm)</td>
<td>2.5</td>
<td>2.0</td>
</tr>
<tr>
<td>Optical markers present</td>
<td>Anterior middle and either side of thorax</td>
<td>One marker at anterior right</td>
</tr>
<tr>
<td>Respiratory belt present?</td>
<td>No</td>
<td>Buckle present at left in inferior slices</td>
</tr>
<tr>
<td>Kidneys in FOV?</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Partial organs</td>
<td>Liver, kidneys and stomach</td>
<td>Liver, spleen and stomach</td>
</tr>
<tr>
<td>Contrast enhanced?</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

As the PF framework is intended to accommodate for only patient specific variability in respiratory motion, the two datasets are used separately in two respective sets of experiments. Each dataset is thereby divided into two subsets, consisting of odd and even phases respectively. One subset is then considered to be a training respiratory cycle used to construct the PF models while the other is considered to be a test cycle. This assignment can be swapped resulting in two experiments for each dataset as listed in Table II. In partitioning the dataset, the models constructed from the training cycle has no prior knowledge of organ configurations during the test cycle. As the reference phase \( \theta \) is taken from an odd phase (phase 1), for experiment B, the test cycle consists of only four phases (excluding phase 1).

#### TABLE II. SUBSETS USED FOR THE TWO EXPERIMENTS OF EACH DATASET

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Odd phases</td>
<td>Even phases</td>
</tr>
<tr>
<td>B</td>
<td>Even phases</td>
<td>Odd phases</td>
</tr>
</tbody>
</table>

IV. RESULTS

The accuracy of PF motion estimation is quantified by the Euclidean distance errors of voxel centres on a per organ basis compared to motion estimated by affine ICP registration. Figs. 6 and 7 compares the mean organ motion (from ICP) to the mean errors from the PF framework using the KDE representation of the transition model for both datasets respectively. The results are averaged over all respiratory phases (excluding the reference phase) and error bars show the standard deviation over those phases.

![Fig. 6. Average motion ('without correction') and PF error ('with correction') for the RSCH dataset. The chosen organs are: mja: major airways, spn: spine, hrt: heart, lvr: liver, spl: spleen, stm: stomach, rl and ll: right and left lungs, rbc: ribcage, shd: shoulder bones and rk and lk: right and left kidneys.](image1)

![Fig. 7. Average motion ('without correction') and PF error ('with correction') for the POPI dataset. The chosen organs are the same as those in Fig. 6 (excluding the kidneys as they are out of the FOV).](image2)

The RSCH dataset has an average organ motion of 7.78 mm whereas the POPI dataset has a lower average organ motion of 2.63 mm. The average error of the PF framework for each dataset is respectively 1.77 mm and 0.97 mm. The improvement of the PF framework using the KDE transition model over an AR(2) based transition model is 26.2% and
8.5% respectively. This suggests that there is a larger improvement when there is more motion present.

When the seven organs with the largest mean motions are considered, the average organ motion is 10.40 mm and 3.28 mm respectively. The improvement of the PF framework using the KDE transition model over an AR(2) based transition model in this case is 29.6% for the RSCH dataset but is still 8.5% for the POPI dataset. This concurs with the previous postulate of amount of improvement depending on the motion present. In both cases, the average error over all organs is comparable with the uncertainty of the position of voxels centres (1.43 mm and 1.22 mm respectively, taking quadrature sums of uncertainty along each dimension). However, the accuracy of the PF framework needs to be assessed with more datasets for a more definite deduction.

As the errors here are from a comparison to motion estimated by ICP, it may have advantage of distinguishing the source of error to those which are solely associated with the PF estimation process. The errors in ICP registration themselves are discussed in [12].

In the final clinical application, motion estimation will be performed during NM acquisition. Estimated organ motion can then potentially be used to reduce the effect of respiratory motion for example, by correcting list mode data.

V. CONCLUSION AND FURTHER WORK

This preliminary evaluation of the PF framework using two clinical 4D thoracic CT datasets is promising and the average error seems to be associated with the actual CT resolution. The results even suggest a better amount of motion compensation relative to the original amount of motion when the latter is larger. This trend is also somewhat true of the amount of improvement when using the KDE representation of the transition model over an AR(2) based model. However as noted in the previous section, more data is needed to evaluate such trends. One such approach is to use a 4D MR sequence that offer volumetric data over time. Another approach is to extend the PF framework to cope with inter-patient variation in respiratory motion. Both approaches are currently being assessed. Additionally the observable will make use of the full anterior surface of the body as originally intended in the proposal of the PF framework [3].

ACKNOWLEDGMENT

The authors would like to thank J. Jones for his implementation of ICP.

REFERENCES


