Contents

1

Estimating internal respiratory motion from respiratory surrogate signals using correspondence models

Jamie McClelland

Summary It is often difficult or impossible to directly monitor the respiratory motion of the tumour and other internal anatomy during RT treatment. Implanted markers can be used, but this involves an invasive procedure and has a number of other associated risks and problems. An alternative option is to use a correspondence model. This models the relationship between a respiratory surrogate signal(s), such as spirometry or the displacement of the skin surface, and the motion of the internal anatomy. Such a model allows the internal motion to be estimated from the surrogate signal(s), which can be easily monitored during RT treatment. The correspondence model is constructed prior to RT treatment. Imaging data is simultaneously acquired with the surrogate signal(s), and the internal motion is measured from the imaging data, e.g. using deformable image registration. A correspondence model is then fit relating the internal motion to the surrogate signal(s). This can then be used during treatment to estimate the internal motion from the surrogate signal(s). This chapter reviews the most popular correspondence models that have been used in the literature, as well as the different surrogate signals, types of imaging data used to measure the internal motion, and fitting methods used to fit the correspondence model to the data.

1.1 Introduction

The previous chapters have described how to acquire 4DCT images of the patient prior to treatment (cf. Part I) and how to estimate and model the respiratory motion from these images using bio-mechanical models or deformable image registration (cf. Part II). This information about the respiratory motion can be very useful for planning RT treatments, e.g. to define appropriate targets and margins that encom-

Jamie McClelland
Centre for Medical Image Computing, Department of Medical Physics and Bioengineering, University College London, London, UK e-mail: j.mcclelland@ucl.ac.uk

pass the tumour motion, and to account for motion in dose calculations (cf. Chapter 11). However, if the respiratory motion can be monitored during treatment then there is the possibility of actively modifying the treatment to compensate for the respiratory motion, e.g. gated or tracked treatments (cf. Sections 11.5 and 11.6). During gated treatment the radiotherapy beam is switched off when the tumour moves outside the target region (e.g. as the patient breathes in), and switched back on when it returns within the target region [4]. This effectively reduces the motion of the tumour but increases the treatment time. Tracked treatments attempt to make the radiotherapy beam follow the tumour motion, effectively reducing the motion of the tumour without increasing treatment time. There are two main approaches that have been proposed for tracked treatments: using a robotically mounted LINAC [57], and moving the leaves of a dynamic multi-leaf collimator (MLC) so that the collimated radiotherapy beam follows the tumour [49]. Both gated and tracked treatments can potentially compensate for respiratory motion, so they can use a smaller target volume as it does not need to encompass the tumour motion which can reduce the dose to the surrounding healthy tissue.

One of the main challenges with actively modifying the treatments is that it can be very difficult to image the tumour motion during treatment. Some lung tumours may be visible using x-ray imaging from certain angles, but generally it is very difficult to accurately identify the tumour. The only current practical solution to this problem is to use implanted markers. However, this involves an invasive procedure to implant the markers and there are a number of other potential problems including the possibility of marker migration [50, 62]. Also, markers are usually only implanted in or close to the tumour so only provide information on the tumour motion. While this is the most important information for guiding gated or tracked treatments, knowing how the rest of the anatomy is moving may be important for ensuring that dose limits to organs at risk are not exceeded.

Instead of attempting to directly image the motion of the tumour and other internal anatomy during treatment, a correspondence model can be used [47]. This models the relationship between the motion of the internal anatomy and a respiratory surrogate signal, such as spirometry or the displacement of the skin surface, which can be easily measured during treatment. Figures 1.1 and 1.2 illustrate the concept of a correspondence model. Prior to treatment imaging data is acquired simultaneously with one or more respiratory surrogate signals. The internal motion is measured from the imaging data and then a correspondence model is fitted which relates the internal motion to the respiratory surrogate signal(s) (cf. Figure 1.1). The surrogate signal can then be monitored during treatment and used to estimate the motion of the internal anatomy and to guide gated or tracked treatments (cf. Figure 1.2).

Correspondence models have also been proposed for other applications including motion compensated image reconstruction in Cone-Beam CT (CBCT) [37, 54, 69] (cf. Section 14.5) and Positron Emission Tomography (PET) [38] and for other image guided interventions such as High Intensity Focussed Ultrasound (HIFU) and cardiac catheterisation. For example, to perform motion compensated CBCT reconstruction a correspondence model can be built from 4DCT, which relates the motion

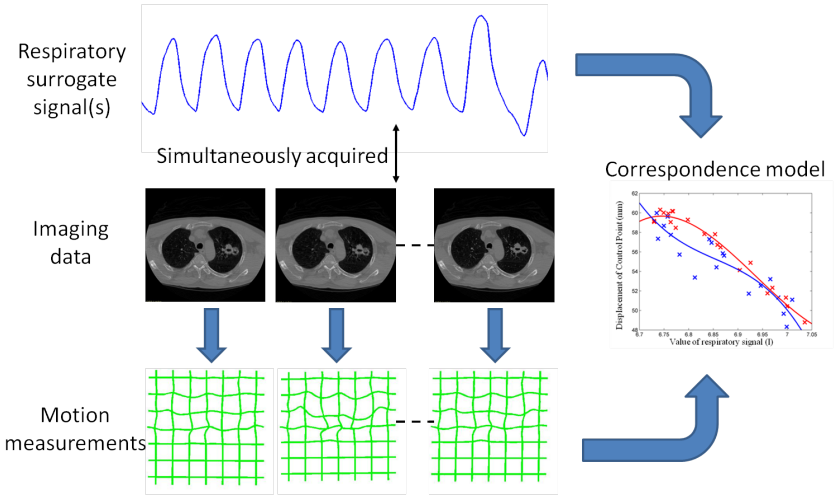

Fig. 1.1 An illustration of how a correspondence model is typically built. Prior to treatment one or more respiratory surrogate signals are acquired simultaneously with imaging data. The motion of the internal anatomy is measured from the imaging data, e.g. using deformable image registration. A correspondence model is then fit which approximates the relationship between the internal motion and the surrogate signal(s).

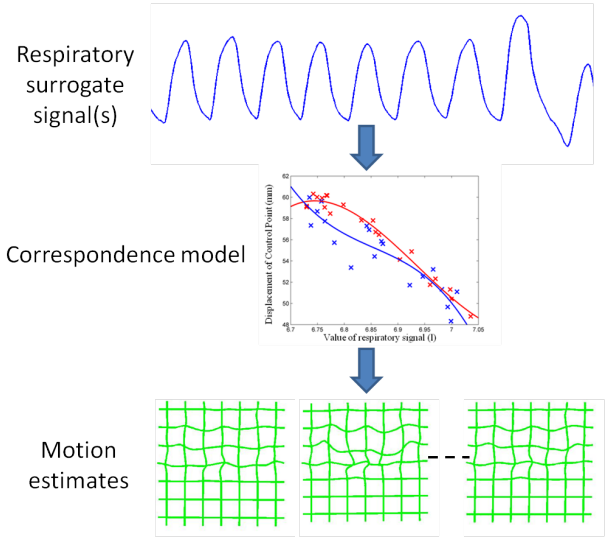

Fig. 1.2 An illustration of how a correspondence model is typically used. During RT treatment the respiratory surrogate signal(s) can be easily measured. The correspondence model can then be used to estimate the internal motion corresponding to the measured surrogate signal(s). These motion estimates could then be used to guide gated or tracked treatments.

of the anatomy to the displacement of the diaphragm. When the CBCT data is acquired the diaphragm motion can be measured from each individual projection. The correspondence model can then be used to estimate the motion for each projection, enabling a motion compensated reconstruction to be performed. A recent review paper [47] which covers all aspects of using correspondence models to estimate respiratory motion includes a description of the various different applications that have used such models.

Establishing accurate and robust correspondence models can be very challenging and is still very much an open area of research. Issues such as hysteresis, where the tumour follows a different path during inhalation and exhalation, and phase offsets between the tumour motion and the surrogate motion, may mean that a simple linear correlation between the surrogate signal and the tumour motion is insufficient to accurately predict the tumour motion [1, 2, 8, 16, 24, 28, 34, 39, 53]. Another problem is that it is well known that there can be breath to breath variations (inter-cycle variation) in the respiratory motion, both during a single fraction of treatment (intra-faction variation) and between fractions of treatment (inter-fraction variation) [24, 48]. Therefore the correspondence models need to account for or adapt to the variations in the respiratory motion.

The rest of the chapter is arranged as follows: Section 1.2 will review different respiratory surrogate signals that have been used for correspondence model in the literature; Section 1.3 will review different methods used to image the internal motion and different representations of the motion used for correspondence models; Section 1.4 will review the correspondence models that have been used in the literature; Section 1.4.2 will review methods of fitting the correspondence models to the data; and Section 1.5 will discuss some of the issues raised by the previous sections and the problems that still need to be overcome before the correspondence models can enter widespread clinical use.

1.2 Respiratory surrogate signals

Many different respiratory surrogate signals have been used for correspondence models. This section will give a brief overview of some of the most common surrogate signals that have been used in the literature. Note, surrogate signals are also used in 4DCT reconstruction to sort the imaging data into coherent volumes (cf. Part I).

Spirometers are used to measure the volume (or flow) of air being inhaled and exhaled by the patient. Spirometry is a popular choice of respiratory surrogate signal [24, 27, 41, 42, 43, 68, 71] as the signal is physiologically related to the respiratory motion and has historically been used for assessing respiratory performance and patterns [3]. However, it has been reported that some patients can have difficulty tolerating spirometry for long periods of time [24], and that spirometry measurements can be subject to time dependent drifts of the end-exhale and end-inhale values due

to escaping air and instrumentation errors [24, 27, 43]. One solution to this problem is to use another 'drift-free' surrogate signal to correct for the drifts [43].

Another popular choice of respiratory surrogate signal is to measure the displacement of the patient's chest or abdomen. This is often done using one or more Infra-Red (IR) markers which are tracked optically [2, 8, 11, 9, 10, 12, 14, 13, 15, 21, 23, 25, 38, 37, 43, 45, 46, 49, 57, 58, 60, 59, 64, 66], and there are a number of commercial system that use this technology: e.g. the Real-Time Position Management (RPM) system (Varian, Palo Alto, California, USA), the Cyberknife (Accuray, Sunnyvale, California, USA), and the ExacTrac system (Brainlab, Feldkirchen, Germany). Other methods of tracking the displacement of the chest or abdomen include electromagnetic tracking systems [24], laser tracking systems [4, 28, 61, 55], and respiratory belts which go around the patient's chest or abdomen and stretch with respiration, e.g. the Anzai respiratory gating system (Siemens, Erlangen, Germany).

Another option for measuring the displacement of the chest and abdomen is to acquire a 3D representation of the patient's skin surface over the chest and abdomen. This has been done using stereo imaging techniques (e.g the Align RT system, Vision RT, London, UK)[26, 27, 44, 48], and the use of time of flight cameras has also been proposed [17]. For building the correspondence models the skin surface can be extracted from CT [17, 15, 19, 35, 45, 46] or MR volumes [20], although another method will be required to monitor the surface during treatment. The displacement at different points on the surface or over different areas of the surface can be followed and used as surrogate signals [17, 15, 19, 20]. Alternatively, a single, global, surrogate signal can be produced by calculating the volume underneath the skin surface [44, 48], which has been shown to produce a similar signal to that acquired from a spirometer but without the signal drift that can affect spirometry [35, 27].

Surrogate signals measured from the internal anatomy have also been proposed in the literature. Such signals may be more closely related to the motion of the tumour and other internal anatomy, but are much more challenging to measure during RT treatment. The motion of one or more points on the diaphragm can be used as surrogate signals [5, 6, 7, 32, 33, 54, 70, 69]. During treatment diaphragm position could be followed using x-ray imaging (although imaging dose could be a factor) [6, 7] or using ultrasound imaging [67]. It has also been proposed that the lung surface could be used as a surrogate signal [40], although it is not clear how the full lung surface would be detected during RT treatment delivery.

It should be noted that a number of correspondence models use derived surrogate signals as well as the measured surrogate signals. These include processing the surrogate signal to calculate the respiratory phase, a time-delayed copy of the surrogate signal, and the temporal derivative of the surrogate signal. Using respiratory phase can be useful for modelling hysteresis, but if respiratory phase is used on its own then it is not possible to model any inter-cycle variation. Using the time derivative of a surrogate signal, or a time delayed copy of the surrogate signal, in conjunction with the original signal can enable the modelling of hysteresis and some degree of inter-cycle variation.

Some authors have proposed using full 2D images of the internal anatomy (e.g. CBCT projection data) as surrogate data [36, 65]. When using full images as the

surrogate data, the direct relationship between the surrogate data and the internal motion is not modelled. Rather, an 'in-direct' correspondence model is used to estimate the internal motion from the surrogate data (cf. Section1.4.3).

1.3 Internal motion data

1.3.1 Imaging the internal motion

This section will describe the different imaging modalities that have been used in the literature for imaging the internal respiratory motion. Some papers image the respiratory motion and construct the correspondence models at the same time as planning the RT treatment, i.e. several days before the treatment. Other papers image the respiratory motion just prior to treatment with the patient setup ready for treatment. This means that a new model can be built for each fraction of treatment and can help account for inter-fraction variations in the respiratory motion (cf. Section1.5). Figure 1.3 shows examples of the different types of imaging data that have been used to image the motion of the internal anatomy.

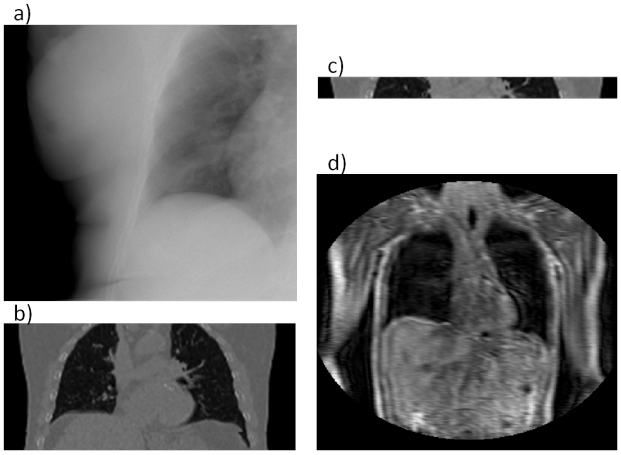

Fig. 1.3 Examples of the different types of imaging data that have been used to image the respiratory motion of the internal anatomy. a) An x-ray projection image, b) a 4DCT volume, c) A Cine CT volume, d) A dynamic MRI volume (acquisition time approx 250 ms for the full volume).

Dynamic x-ray imaging (such as fluoroscopy) has been used to image the internal respiratory motion [1, 4, 6, 7, 11, 9, 10, 14, 25, 24, 28, 29, 44, 55, 57, 58, 60, 59, 61, 64, 66](cf. Figure 1.3a). Its main advantages are that it can be easily acquired just prior to or during RT treatment and can be acquired with a very high temporal resolution so that the images can be considered free of motion artefacts. However,

as noted in the introduction to this chapter, it can be very difficult to detect the tumour in x-ray images due to it being occluded by high intensity structures such as bones or the mediastinum. One solution is to implant small radio-opaque markers into or close to the tumour in order to follow its motion [4, 11, 9, 10, 14, 25, 28, 29, 55, 57, 58, 61, 64]. These markers are easily detectable in the x-ray images and can be automatically tracked, but have a number of drawbacks including an invasive procedure to implant them and the possibility of them moving away from their intended locations [50, 62]. Therefore developing methods to automatically determine the tumour motion from x-ray images without using implanted markers is an active area of research [44, 60]. As x-rays produce 2D projection images two simultaneous x-ray images from different directions are generally required to track the 3D motion. Another option that has been proposed is to use a series of x-ray images acquired with the source and detector rotating around the patient, i.e. the raw projections images from a cone-beam CT acquisition (cf. Chapter 14). The model can then be fit to all of the x-ray images simultaneously, allowing the full 3D motion to be modelled [44].

CT imaging has been widely used in the literature to image the internal respiratory motion. It provides information on the full 3D motion and deformation of the anatomy, enabling 3D deformable registrations to be performed. Many studies have used 4DCT datasets as described in Part I to image the respiratory motion during free breathing [2, 18, 17, 15, 19, 16, 32, 33, 38, 37, 36, 40, 54, 65, 70, 69](cf. Figure 1.3b). These are popular as they provide images over the full region of interest. However, they are a formed by combining data acquired from several different respiratory cycles, and are based on the assumption that the respiratory motion will be the same from one respiratory cycle to the next. Breath-to-breath variations in the respiratory motion will cause artefacts in the 4DCT volumes (cf. Section 1.5.1 and Section 2.5). This also means that 4DCT volumes are not appropriate for studying and modelling the short term breath-to-breath variations in the respiratory motion.

To overcome this problem some studies have used Cine CT data [12, 8, 21, 23, 41, 42, 45, 46, 49, 48, 68, 71](cf. Figure 1.3c) which is essentially the unsorted data used to construct cine 4DCT volumes (cf. Chapter 2). This means that the field of view and number of slices in the superior-inferior direction is limited by the size of the CT detector. Many scanners used clinically only have 12 or 16 slices covering approximately 2-2.5 cm. Some recently developed scanners can have a much larger detector giving greater coverage and many more slices [51], although these are still not generally large enough to cover the whole region of interest for RT planning. Therefore, if motion estimates are required for the whole region of interest it may be necessary to build multiple correspondence models and to combine the results from these models [45, 46, 49, 48, 68]. Cine CT data may be acquired at the same location over more than one respiratory cycle, giving information on the breath-to-breath variations that occur.

MRI data has also been used for constructing correspondence models [5, 20, 34, 39, 53](cf. Figure 1.3d). MRI has the advantage that it does not deliver any radiation dose to the patient, potentially allowing much more data to be acquired so as to study both the intra- and inter-fraction variation, and enabling volunteer studies. It

is possible to acquire 3D MR volumes of a large region of interest (e.g. the lungs) at a temporal resolution fast enough to image the respiratory motion (0.5 seconds per image), although the spatial resolution and image quality is poor compared to 4DCT making accurate image registration much more challenging [5, 20].

1.3.2 Representing the internal motion data

Some papers just model the motion of one or a small number of points of interest. In a clinical setting the point of interest would usually correspond to the tumour. Some papers manually determine the motion of the point of interest but most papers use an automatic method such as template tracking (e.g. [41]). Other correspondence models estimate the motion over an entire region of interest. In these cases image registration is often used to determine the motion from the imaging data. Affine registrations have been used [5], although deformable registrations are more common as these can account for the local deformations that occur during respiration. Several different deformable registration algorithms have been used in the literature, including B-spline (e.g. [46]), demons (e.g. [36]), optical flow/diffusion (e.g. [70]), fluid (e.g. [40]), and surface based registrations (e.g. [33]). For more information on the different registration algorithms please see Part II.

All correspondence models in the literature essentially represent the internal motion data using a motion vector, $\mathbf{m} = [m_1, m_2, \ldots, m_{N_m}]^T$, where there are N_m values representing the internal motion. In some cases the elements of the motion vector correspond to the motion of one or more anatomical points. These could be a point of interest such as the tumour [1, 2, 6, 7, 8, 11, 9, 10, 14, 13, 24, 25, 28, 29, 34, 39, 41, 42, 44, 55, 57, 58, 60, 59, 61, 64, 71], or they could be points defining the surface of an organ [5, 32, 33]. The actual values used could correspond to the absolute position of the points (in some predefined coordinate system) or the displacement of the points relative to a reference position [33]. Some papers only model the 1D motion (usually in the SI direction) of the anatomical points, so $\mathbf{m} = \left[x_1, x_2, \ldots, x_{N_p}\right]^T$ where x_n is the motion of the nth point and N_p is the number of anatomical points. In this case $N_m = N_p$. Some papers model the 2D motion of the anatomical points so $\mathbf{m} = \left[x_1, y_1, x_2, \ldots, x_{N_p}, y_{N_p}\right]^T$ and $N_m = 2N_p$, and some papers model the full 3D motion of the points so $\mathbf{m} = \left[x_1, y_1, z_1, x_2, \ldots, x_{N_p}, y_{N_p}, z_{N_p}\right]^T$ and $N_m = 3N_p$, where x_n is the motion of the nth point in the x direction, y_n is the motion of the nth point in the y direction, and z_n is the motion of the nth point in the z direction.

Other representations of the motion have also been used in the literature. The motion vector could represent a deformation field [17, 15, 19, 38, 37, 36, 40, 54, 65, 68, 70, 69], where each element of $\mathbf{m}$ corresponds to the motion of a voxel in a particular direction, and N_m equals 3 times the number of voxels in the volume. The motion vector could represent a velocity field that defines a diffeomorphic transformation [23, 21], where each element of $\mathbf{m}$ corresponds to the velocity at a voxel in a particular direction, and N_m again equals 3 times the number of voxels in the vol-

ume. Or the motion vector could represent a control point grid defining a B-spline transformation [12, 18, 20, 45, 46, 49, 48], where each element of **m** corresponds to the displacement of one of the control points in one direction and N_m equals 3 times the number of control points that define the B-spline transformation. The size of the motion vector, N_m, will depend on the motion representation being used. It will be small if only one or a few points of interest are being modelled, it could be several thousand if points defining an organ surface are being modelled, it could be many thousand if a B-spline control point grid is being modelled, or it could be millions if a deformation or velocity field is being modelled.

1.4 Correspondence models

1.4.1 Types of model

This section will describe some of the different correspondence models that have been used to relate the internal motion to respiratory surrogate signals. Table 1.1 summarises the different correspondence models that have been used in the literature. Figure 1.4 illustrates the different motion trajectories that can be modelled by the more popular types of correspondence models in Table 1.1.

Type	**Details**	**Examples**
Linear	1 surrogate signal	[1, 2, 6, 7, 8, 11, 9, 16, 24, 25, 28, 34, 39, 53, 55, 57, 58, 60, 59, 61]
	2 or more surrogate signals	[6, 7, 10, 14, 13, 17, 15, 19, 20, 29, 32, 33, 40, 42, 41, 48, 44, 55, 68, 70, 71, 69]
Piece-wise linear	1 surrogate signal	[21, 23]
	Using respiratory phase as surrogate signal	[54]
Polynomial	1 surrogate signal	[13, 25, 45, 55]
	With separate inhalation and exhalation	[5, 14, 13, 25, 48, 61, 64]
	2 surrogate signals	[55, 64]
B-spline	Using respiratory phase as surrogate signal	[12, 45, 46, 49, 48]
	2 or more surrogate signals	[18]
Others	Fourier series	[45]
	Neural networks	[29, 64]
	Fuzzy logic	[64]
	Support vector regression	[14, 13]

Table 1.1 Summary of the different types of correspondence models used in the literature.

1.4.1.1 Linear models

As can be seen in Table 1.1, linear correspondence models are by far the most common in the literature. These model the motion as a linear combination of surrogate signals:

$$\mathbf{m} = \mathbf{C}\mathbf{s} + \mathbf{C_0} \tag{1.1}$$

where $\mathbf{m}$ is the vector of motion estimates (see Section 1.3.2), $\mathbf{s}$ is a vector of surrogate signals (i.e. $\mathbf{s} = [s_1, s_2, \ldots, s_{N_s}]^T$ where s_n is the nth surrogate signal and there are N_s surrogate signals, some of which may be derived surrogate signals), $\mathbf{C}$ is a $N_m \times N_s$ matrix of linear coefficients, and $\mathbf{C_0}$ is a N_m element vector of constant terms.

Several papers use a linear model with just a single surrogate signal (i.e. $N_s = 1$), which is therefore a simple linear correlation between the surrogate signal and the motion. This constrains the motion to follow a straight line during each breath (cf. Figure 1.4a). Such a model was first proposed for RT related applications by Schweikard et al. [57] for guiding tracked RT treatment using the Cyberknife (Accuray, Sunnyvale, California, USA). Although such a model may not be very realistic, it has been shown to be sufficiently accurate in some circumstances, and is used clinically in the Cyberknife system to treat some patients [25, 61].

There have been a number of studies assessing how different factors can affect the correlation between the surrogate signal and the motion [1, 2, 8, 16, 24, 28, 34, 39, 53], e.g. the surrogate signal used [24], the type of breathing the patient is performing (deep or shallow, using their ribs or using their diaphragm) [34, 53], and over what time scales the linear correlations are valid [24]. These studies have had mixed results. In some circumstances a simple linear correlation can approximate the respiratory motion relatively well over a short time frame, but in other circumstances, such as when there is significant hysteresis, a simple 1-D linear correlation is not sufficient.

Other papers have used linear models of two or more (sometimes many more) surrogate signals (i.e. $N_s > 1$). When there are only two signals these usually comprise one measured surrogate signal and one derived surrogate signal, either a time delayed signal or the temporal derivative (e.g. [41]). When more signals are used they are often multiple measured surrogate signals from different parts of the anatomy (e.g. [32]), although sometimes they will use multiple measured and multiple derived signals (e.g. [20]), and sometimes just a single measured signal with multiple derived signals (i.e. a number of time delayed signals with a different length delay for each signal, e.g. [29]). Sometimes the different signals may be highly correlated with each other, e.g. when the signals represent points on the surface of an organ, and this can make the models susceptible to over-fitting [32, 33]. However, the models are very flexible and can potentially model complex motion including both hysteresis and inter-cycle variation (cf. Figure 1.4b). Linear models using multiple surrogate signals were first proposed for RT related applications by Low et al. [41].

One way to model more complex motion with only a single surrogate signal is to use a piece-wise linear model [21, 23, 38, 37, 54]. In this case the motion has been measured from several images (e.g. using deformable image registration), each of which corresponds to a different surrogate signal value. These motion measurements are linearly interpolated to estimate the motion at other surrogate signal values where data has not been acquired. Such models allow the motion to follow more complex paths than a simple straight line, but as with any model using a single surrogate signal, they are limited in the variation they can model. A piece-wise linear model that relates the motion to the measured surrogate signal [21, 23] cannot model hysteresis but can model some limited inter-cycle variation. The motion will always follow the same trajectory during each breath, but can move a different distance along this trajectory depending on how deeply the patient is breathing (cf. Figure 1.4c). A piece-wise linear model that relates the motion to respiratory phase [38, 37, 54] can model hysteresis but cannot model any inter-cycle variation (cf. Figure 1.4d). Furthermore, a piece-wise linear model cannot extrapolate outside the range of surrogate signal values used to build the model.

1.4.1.2 Polynomial models

Polynomial correspondence models have also been widely used. These estimate the motion as a polynomial function of the surrogate signal(s). The equation for a polynomial correspondence model using a single scalar surrogate signal, s_1, is:

$$\mathbf{m} = \sum_{i=0}^{p} \mathbf{C}_i s_1^i \tag{1.2}$$

and the equation for a polynomial model using two scalar surrogate signals, s_1 and s_2, is:

$$\mathbf{m} = \sum_{i=0}^{p} \sum_{j=0}^{n-i} \mathbf{C}_{i,j} s_1^i s_2^j \tag{1.3}$$

where $\mathbf{C}_i$ and $\mathbf{C}_{i,j}$ are vectors of polynomial coefficients and p is the order of the polynomial. Such models are usually 2nd or 3rd order polynomials. Higher orders polynomials have been investigated [45], but they are more likely to over-fit the data and lead to very large extrapolation errors. Polynomial models were first proposed for RT related applications by Blackall et al. [5].

Polynomial models are usually used with a single scalar surrogate signal, but they have also been used with respiratory phase [45], and with two surrogate signals (the second being a precursor of the first) [55, 64]. As with the linear models, if a single surrogate is used the motion is constrained to follow the same trajectory during each breath, although it can move a different distance along this trajectory. However, unlike the linear models the trajectory is no longer constrained to be a straight line (cf. Figure 1.4e). As polynomial models can be susceptible to large extrapolation errors some papers revert to using a linear model when estimating the

motion for values of the surrogate signal that are outside the range used to build the model [14, 13, 61, 64].

A polynomial model of a single surrogate signal cannot model hysteresis as the motion is constrained to follow the same path during inhalation and exhalation. Therefore some papers fit two separate polynomial models, one to the data from inhalation and one to the data from exhalation, as first proposed in Blackall et al [5]. This allows hysteresis to be modelled but can result in discontinuities in the motion estimates when switching from inhalation to exhalation (cf. Figure 1.4f). To try and minimise the discontinuities some papers 'blend together' the polynomial models and a linear model in the regions corresponding to end-exhale and end-inhale, creating a smooth transition from one model to the other and allowing the linear model to be used when extrapolation is required [14, 13].

1.4.1.3 B-spline models

B-spline correspondence models have also been proposed to relate the surrogate signal(s) and the motion. These are usually used with a single derived surrogate signal representing respiratory phase, so a 1D B-spline is used. One paper [18] has proposed using both a derived respiratory phase signal and the actual measured signal so uses a two dimensional B-spline.

As the internal motion is related to respiratory phase, ϑ, most of the papers have modified the standard B-spline function so as to make it periodic (or cyclic), so that there are no discontinuities between one breath cycle and the next:

$$\mathbf{m} = \sum_{i=0}^{3} \mathbf{C}_{i+k \bmod N} B_i(j) \tag{1.4}$$

where $j = \frac{\vartheta}{\delta} - \lfloor \frac{\vartheta}{\delta} \rfloor$, $k = \lfloor \frac{\vartheta}{\delta} \rfloor - 1$, ϑ is the respiratory phase (between 0 % and 100 %), B_i is the ith B-spline basis function (see Section 6.1.2.1 or [56]), $\mathbf{C}_0, \ldots, \mathbf{C}_N$ are vectors of B-spline control points (with N_m elements), N is the number of control points, and δ is the control point spacing ($\delta = \frac{100\%}{N}$). Note, although the equation does directly relate the estimated motion to a measured surrogate signal, the values of the B-spline basis functions B_i depend on the value of the respiratory phase, ϑ, which is derived from the measured surrogate signal.

Periodic B-spline correspondence models were first proposed for RT related applications by McClelland et al. [45, 46]. A 1D periodic B-spline correspondence model constrains the motion to follow a loop shaped trajectory. This means that hysteresis can be modelled, but inter-cycle variation cannot be modelled (cf. Figure 1.4g). The 2D B-spline model can potentially model inter-cycle variation and hysteresis, but as there are many more model parameters to fit there is a much higher risk of over-fitting the data.

1.4.1.4 Others

Other correspondence models have also been proposed in the literature, including models based on Fourier series [45], neural networks [29, 64], fuzzy logic [64], and support vector regression [13, 14]. However, these models have not been widely used in the literature so they will not be covered in detail here. Interested readers should refer to the papers referenced above for more detail on these methods.

1.4.2 Fitting the correspondence models

A number of different methods and techniques have been used to fit the correspondence models to the data. This section will describe some of the more widely used methods.

To fit a correspondence model the motion of the internal anatomy and the surrogate signals must both be simultaneously sampled a number of times. Each sample will usually represent a specific timepoint, although is some papers they represent an entire breath cycle [32, 33]. The motion of the internal anatomy at each sample, t, is represented by the motion vector $\mathbf{m}_t = [m_{1,t}, m_{2,t}, \ldots, m_{N_x,t}]^T$, where there are N_m values representing the internal motion (e.g. N_m is 3 times the number of voxels if the motion is represented as a deformation field, cf. Section 1.3.2). The surrogate signals at each sample, t, are represented by another vector $\mathbf{s}_t = [s_{1,t}, s_{2,t}, \ldots, s_{N_s,t}]^T$, where there are N_s surrogate signals. All of the samples of the motion data can be then concatenated into a matrix, $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_{N_t}]$ and samples of the surrogate signal can be concatenated into another matrix, $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{N_t}]$ where N_t is the number of training samples, i.e. the number timepoints that the internal motion has been imaged (or the number of breath cycles imaged). N_t is typically 8 or 10 if 4DCT has been used to image the internal motion (one sample for each 4DCT image), but may be more if a different modality (e.g. x-ray imaging) has been used.

1.4.2.1 Linear least squares

By far the most commonly used method is linear least squares [5, 11, 9, 10, 12, 14, 13, 25, 41, 42, 45, 46, 49, 48, 55, 57, 58, 60, 59, 61, 64, 68]. The correspondence model can be formulated as a linear multivariate regression problem:

$$\mathbf{m}_t = \mathbf{C}\mathbf{s}_t \tag{1.5}$$

where $\mathbf{C}$ contains the correspondence model parameters, which can be found from the training data using ordinary least squares by minimising the squared difference between $\mathbf{M}$ and $\mathbf{CS}$, i.e.

$$\arg\min (\mathbf{M} - \mathbf{CS})^2 \tag{1.6}$$

this is done by:

$$\mathbf{C} = \mathbf{M}\mathbf{S}^T \left(\mathbf{S}\mathbf{S}^T\right)^{-1} = \boldsymbol{\Sigma}_{\mathbf{MS}}\boldsymbol{\Sigma}_{\mathbf{SS}}^{-1} \tag{1.7}$$

where $\boldsymbol{\Sigma}_{\mathbf{SS}}$ is the covariance matrix of $\mathbf{S}$ and $\boldsymbol{\Sigma}_{\mathbf{MS}}$ the cross-covariance matrix of $\mathbf{M}$ and $\mathbf{S}$.

A motion estimate, $\mathbf{m}_{est}$, can then be made from new surrogate signal values, $\mathbf{s}_{new}$, using the fitted correspondence model parameter, $\mathbf{C}$:

$$\mathbf{m}_{est} = \mathbf{C}\mathbf{s}_{new} \tag{1.8}$$

Note, linear least squares can be used for polynomial models and B-spline models as well as the linear models by considering simple functions of the surrogate signals (such as raising to a power or calculating the B-spline basis functions) as just another surrogate signal. Also, if an additional 'surrogate signal' with a constant value of 1 is used when fitting the linear model this will allow the constant terms $\mathbf{C_0}$ to be fit as well. E.g. to fit a 2nd order polynomial correspondence model with one surrogate signal, s_1, the surrogate signal vector for each sample, t, will be $\mathbf{s}_t = \left[s_{1,t}^2, s_{1,t}, 1\right]^T$.

1.4.2.2 Principal Component Analysis

Principal Component Analysis (PCA) is a mathematical technique which can be used to help analyse and interpret multi-variate data [30]. PCA is performed by finding the eigenvectors of the data covariance matrix. The result of PCA is a set of linearly uncorrelated variables, the principal components. These are ordered such that the first principal component has the largest possible variance (i.e. accounts for the most variation in the data possible), and each successive component has the highest possible variance while being orthogonal to the previous principal components. The total number of principal components will be the same as the number of variables in the data. However, it is usually possible to represent the majority of the variance in the data using substantially fewer principal components than there were original variables. Therefore PCA can be used as a dimensionality reduction tool. The original data can be re-expressed in terms of the PCA weights, that is the values by which each principal component is multiplied to approximate the original data. If all the principal components are used then the principal components multiplied by the PCA weights will equal the original data exactly. If fewer principal components are used this will approximate the original data, although a good approximation will often be achieved if a sufficient amount of the variance (usually 90-95 %) is included in the principal components that are used).

PCA has been utilised in a number of different ways when fitting respiratory motion models.

Applied to internal motion data

It has been applied to the internal motion data prior to fitting a correspondence model [20, 69]. This can help remove unwanted noise from the motion data due to imaging artefacts and/or registration errors. First the mean motion vector, $\overline{\mathbf{m}}$, is subtracted from each motion vector to give the mean centred motion vectors, $\widetilde{\mathbf{m}}_t = \mathbf{m}_t - \overline{\mathbf{m}}$, and these are concatenated to give the mean centred motion matrix $\widetilde{\mathbf{M}}$. This is then approximated using K_m principal components:

$$\widetilde{\mathbf{M}} \approx \mathbf{E}_m \mathbf{W}_m \tag{1.9}$$

where the matrix $\mathbf{E}_m = [\mathbf{e}_1, \ldots, \mathbf{e}_{K_m}]$ consists of the first K_m principal components and matrix $\mathbf{W}_m$ of size $K_m \times N_t$ gives the principal component weights for each of the training motion samples.

The correspondence model now relates the surrogate signal(s) to the principal component weights rather than to the motion vectors themselves. Ordinary least squares can be used as above to fit the correspondence model [69]:

$$\mathbf{C} = \mathbf{W}_m \mathbf{S}^T \left(\mathbf{S}\mathbf{S}^T\right)^{-1} = \boldsymbol{\Sigma}_{\mathbf{W}_m \mathbf{S}} \boldsymbol{\Sigma}_{\mathbf{SS}}^{-1} \tag{1.10}$$

where $\boldsymbol{\Sigma}_{\mathbf{W}_m \mathbf{S}}$ is the cross-covariance matrix of $\mathbf{W}_m$ and $\mathbf{S}$.

A motion estimate can be made from new surrogate signals values by first estimating the principal component weights, $\mathbf{w}_{m\_est}$:

$$\mathbf{w}_{m\_est} = \mathbf{C}\mathbf{s}_{new} \tag{1.11}$$

and then multiplying by the principal components and adding the mean motion vector to get the motion estimate:

$$\mathbf{m}_{est} = \overline{\mathbf{m}} + \mathbf{E}_m \mathbf{w}_{m\_est} = \overline{\mathbf{m}} + \mathbf{E}_m \mathbf{C}\mathbf{s}_{new} \tag{1.12}$$

Note, alternative fitting methods, such as canonical correlation analysis [20, 40], can be used instead of ordinary least squares to calculate the model parameters, $\mathbf{C}$, which relate the principal component weights to the surrogate signals.

Applied to surrogate signal data

PCA has also been applied to the surrogate signal data before fitting a correspondence model [33]. The correspondence model then relates the PCA weights (which now represent the surrogate signals) to the internal motion data. This approach is known as principal component regression (PCR) and it can help provide a more stable and robust fit by removing co-linearities in the surrogate data. This can be particularly useful if there are a large number of surrogate signals that are highly correlated with each other, e.g. when the surrogate signals are from points on an organ/skin surface, and can help prevent over-fitting in such circumstances. The mean

centred surrogate signals, $\widetilde{\mathbf{S}}$ are represented using K_s principal components:

$$\widetilde{\mathbf{S}} \approx \mathbf{E}_s \mathbf{W}_s \tag{1.13}$$

where the matrix $\mathbf{E}_s$ contains the first K_s principal components and matrix $\mathbf{W}_s$ the principal component weights for each of the training surrogate samples.

Ordinary least squares is then used to fit the correspondence model relating the internal motion to the principal component weights:

$$\mathbf{C} = \mathbf{M}\mathbf{W}_s^T \left(\mathbf{W}_s \mathbf{W}_s^T\right)^{-1} = \boldsymbol{\Sigma}_{\mathbf{M}\mathbf{W}_s} \boldsymbol{\Sigma}_{\mathbf{W}_s\mathbf{W}_s}^{-1} \tag{1.14}$$

where $\boldsymbol{\Sigma}_{\mathbf{W}_s\mathbf{W}_s}$ is the covariance matrix of $\mathbf{W}_s$ and $\boldsymbol{\Sigma}_{\mathbf{M}\mathbf{W}_s}$ the cross-covariance matrix of $\mathbf{M}$ and $\mathbf{W}_s$. A motion estimate is made from new surrogate signal values by first calculating the corresponding principal component weights, $\mathbf{w}_{s\_new}$:

$$\mathbf{w}_{s\_new} = \mathbf{E}_s^T \left(\mathbf{s}_{new} - \bar{\mathbf{s}}\right) \tag{1.15}$$

where $\bar{\mathbf{s}}$ is the mean of the training surrogate vectors, and then estimating the motion using the correspondence model:

$$\mathbf{m}_{est} = \mathbf{C}\mathbf{w}_{s\_new} = \mathbf{C}\mathbf{E}_s^T \left(\mathbf{s}_{new} - \bar{\mathbf{s}}\right) \tag{1.16}$$

Applied to motion and surrogate data

PCA can also be applied to both the motion data and the surrogate data separately, before fitting a correspondence model relating the motion PCA weights to the surrogate data PCA weights [40]. This should both help to remove noise from the motion data and remove co-linearities from the surrogate data. The motion estimate is then made using:

$$\mathbf{m}_{est} = \overline{\mathbf{m}} + \mathbf{E}_m \mathbf{C}\mathbf{E}_s^T \left(\mathbf{s}_{new} - \bar{\mathbf{s}}\right) \tag{1.17}$$

Used to fit the correspondence model

PCA has also been used to actually fit the correspondence model $\mathbf{C}$ [6, 7, 17, 15, 19, 70, 32]. All the mean centred motion vectors and surrogate vectors are combined into a single data vector, $\mathbf{z}_t = \left[\widetilde{\mathbf{m}}_t^T, \widetilde{\mathbf{s}}_t^T\right]^T$ with length $N_m + N_s$. PCA is then performed on the $(N_m + N_s) \times N_t$ matrix $\mathbf{Z}$ which contains all the combined data vectors from all the N_t training samples.

$$\mathbf{Z} \approx \mathbf{E}_z \mathbf{W}_z \tag{1.18}$$

where the matrix $\mathbf{E}_z$ consists of the first K_z principal components. This equation can be split into two separate equations:

$$\widetilde{\mathbf{M}} \approx \mathbf{E}_{zm}\mathbf{W}_z \tag{1.19}$$

$$\widetilde{\mathbf{S}} \approx \mathbf{E}_{zs}\mathbf{W}_z \tag{1.20}$$

where $\mathbf{E}_{zm}$ and $\mathbf{E}_{zs}$ are constructed from the upper N_m and lower N_s rows of $\mathbf{E}_z$ respectively. Assuming the inverse matrix $\mathbf{E}_{zs}^{-1}$ exists $\mathbf{W}_z$ can be eliminated from the equations:

$$\widetilde{\mathbf{M}} \approx \mathbf{E}_{zm}\mathbf{E}_{zs}^{-1}\widetilde{\mathbf{S}} \tag{1.21}$$

and a motion estimate can be made from new surrogate signal values using:

$$\mathbf{m}_{est} = \overline{\mathbf{m}} + \mathbf{E}_{zm}\mathbf{E}_{zs}^{-1}\left(\mathbf{s}_{new} - \overline{\mathbf{s}}\right) \tag{1.22}$$

Note, for $\mathbf{E}_{zs}^{-1}$ to exist there must be at least as many rows as columns in $\mathbf{E}_{zs}$, i.e. the number of principal components used, K_z, cannot exceed the number of surrogate signals, N_s.

1.4.2.3 Other Fitting Methods

There have been a number of other fitting methods proposed in the literature. These include: ridge regression [32], canonical correlation analysis [20, 40], support vector regression [13, 14], multi-level B-spline approximation [18], Nelder-Mead optimisation [71], the Levenberg-Marquardt algorithm [29, 64], Fuzzy logic methods [64], and digital filtering [65]. Some of these fitting methods are general methods that can be used with many different correspondence models, and others are specialised methods designed for specific correspondence models. However, they have not been widely used and will not be discussed in detail. Interested readers should refer to the papers referenced above for more detail on these methods.

1.4.2.4 Fitting the model directly to the imaging data

Some authors have have proposed methods which iterate between fitting the correspondence model directly to the (raw) imaging data and performing a motion compensated image reconstruction. This approach was originally proposed for MR imaging [52], where the correspondence model is fitted directly to the k-space data. The approach has also been used for 4DCT imaging [21, 23], where the model is fitted to the unsorted Cine CT volumes, and for CBCT imaging [44], where the model is fitted directly to the CBCT projection data.

For example, in the case of CBCT it can take a minute or more to acquire all the projections images that are required to reconstruct a full 3D volume. This means that the reconstructed volume will contain artefacts due to the respiratory motion, e.g. the tumour can appear blurred. If a surrogate signal is simultaneously acquired with the projection data then a correspondence model can be used to estimate the motion that occurred when each projection was acquired. This can be used to per-

form a motion compensated image reconstruction which should not contain any motion artefacts (cf. Section 14.4). The correspondence model can be fitted using prior imaging data [37, 54, 69], but this assumes that the motion is the same during the CBCT acquisition as it was when the prior imaging data was acquired. Alternatively an iterative method can be used which fits the correspondence model directly to the CBCT projection data [44]. In this case a image reconstruction is first performed assuming no motion. This reconstructed volume can then be 'animated' according to the correspondence model, and projections can be simulated from the animated reconstructed volume and compared to the original projections. The correspondence model parameters are optimised to give the best match between the simulated projections and the original projections. The correspondence model can then be used to perform a motion compensated image reconstruction. This whole procedure can then be iterated, using the motion compensated image reconstruction to get a better estimate of the correspondence model parameters, and then using the better estimate of the correspondence model to improve the motion compensated image reconstruction.

1.4.2.5 Adaptive fitting

If extra images of the internal motion can be acquired during a fraction of RT treatment then they can be used for 'adaptive fitting' [11, 9, 10, 25, 29, 57, 58, 60, 59, 61, 64]. For example the Cyberknife system (Accuray, Sunnyvale, California, USA) can acquire x-ray images of the patient in the treatment position, allowing images to be acquired during treatment as well as prior to treatment (for building the initial correspondence model). In theory x-ray images could be acquired continuously throughout treatment, removing the need for a correspondence model, but this would greatly increase the radiation dose to the patient due to imaging. Adaptive fitting means that the correspondence model is intermittently updated as new data becomes available. This allows the model to adapt to gradual changes in the relationship between the internal motion and the surrogate. This is usually done by discarding the oldest data in favour of the newly acquired sample and re-fitting the model. The newly acquired sample can also be used to check the accuracy of the current model, and if it drops below a predetermined tolerance the treatment can be paused while the model is completely rebuilt using new data. This approach can help the model react to more sudden changes in the motion-surrogate relationship.

1.4.3 Indirect correspondence models

Unlike the correspondence models described in the previous sections, indirect correspondence models do not directly relate the motion to the surrogate data. Instead, they parameterise the motion using one or more internal variables, and make estimates of the surrogate data as well as the motion data. When surrogate data is

acquired during a procedure the internal variables are optimised to give the best match between the estimated surrogate data and the measured surrogate data. This can be expressed as:

$$\mathbf{x} = \phi(\mathbf{v}) \tag{1.23}$$

$$\mathbf{v} = \arg\max_{\mathbf{v}} \; Sim\left(F\left(T\left(I,\mathbf{x}\right)\right),\mathbf{s}\right) \tag{1.24}$$

where $\phi(\mathbf{v})$ is an indirect correspondence model that estimates the motion, $\mathbf{x}$, from some internal variables, $\mathbf{v}$. I is a reference image, T is a function that transforms the reference image according to the motion estimate $\mathbf{x}$, F is a function which simulates the surrogate data from the transformed reference image, and Sim is a measure of similarity between the simulated surrogate data and the measured surrogate data, $\mathbf{s}$. The function F can vary: for the RT related models in the literature [36, 65] the function F simulates a CBCT projection from CT data, but for other applications F can simulate different types of data (e.g. an Ultrasound slice from MR data for cardiac interventions[31]).

The model $\phi(\mathbf{v})$ can also vary. Two models have been used in the literature: a linear model [36], and a B-spline model which incorporates an additional linear scaling of the motion estimates to allow for variations in the depth of breathing [65]. The linear model is fit to the motion data using PCA and the internal variables, $\mathbf{v}$, are the principal component weights. For the B-spline model the internal variables represent the respiratory phase and the depth of breathing. The B-spline model is fit to the motion data using digital filtering.

At treatment time when the surrogate data is available the internal variables are optimised to estimate the motion that best matches the CBCT projections. The two papers use different strategies to help the optimisation. Li et al [36] use a prediction model (cf. Chapter 12) to obtain good initial guesses for the internal variables. Vandemeulebroucke et al [65] assume the respiratory phase and depth of breathing vary smoothly with time, and then optimise their values for the CBCT projections from an entire breath cycle simultaneously.

1.5 Discussion and Conclusions

This chapter has reviewed the various different options that have been used in the literature for constructing respiratory motion model that relate the motion of the internal anatomy to one or more easily measured surrogate signals. There are a number if issues that should be considered when designing and using such models which will now be discussed.

Intra-fraction variation

If the models will be used to estimate the respiratory motion during a fraction of RT treatment then the models needs to be able to account for intra-fraction variations in the motion. The model can try and estimate the variations in the internal motion from variations in the surrogate signals. This can potentially be done by using two or more surrogate signals (one of which may be a derived surrogate signal such as the temporal derivative or a time delayed copy of a measured surrogate signal). However, the ability of such models to accurately estimate the motion and its variation over a time frame corresponding to a fraction of treatment has yet to be demonstrated.

The models can also be made to adapt to the intra-fraction variations by using an adaptive fitting approach as described in Section 1.4.2.5. This requires intermittent measurements of the internal motion during treatment, which may be difficult to acquire. An adaptive fitting approach is used in the Cyberknife system from Accuray, one of the only commercial products currently in routine clinical use that utilises correspondence models to estimate the internal respiratory motion. The fact that the accuracy of the models can be verified and the models can be updated during treatment may be one of the main reasons that the models are considered accurate and robust enough to be used in a clinical setting. However, it should be noted that in order to acquire intermittent measurements of the internal motion it is usually necessary to implant internal markers in or around the tumours so that they can be imaged during treatment.

Inter-fraction variation

If the models will be used to estimate the respiratory motion during different fractions of RT treatment then the models need to be able to account for inter-fraction variations. As with intra-fraction variation, the models could try and estimate the inter-fraction variations from variations in the surrogate signals. However, this could be very challenging as between different fractions of treatment there can be anatomical changes such as weight loss/gain and growing/shrinking tumour, and the patient can change the way that they breathe (e.g. from chest breathing to abdominal breathing). These inter-fraction variations can effect the relationship between the surrogate signal(s) and the internal motion, and can invalidate models built during a previous fraction of RT or from data acquired at planning [48].

Another approach for dealing with inter-fraction variations is to check the validity of the correspondence model prior to each fraction of treatment, and if necessary update or rebuild the model. This requires measurements of the internal motion just prior to each fraction of treatment, which may be difficult or impractical to acquire. Additionally, if the anatomy and/or motion have changed considerably from that seen during planning then the treatment plan may no longer be suitable and a new plan may be required. This approach of 'Adpative RT' is an active area of research

[63], and correspondence models that can be updated to account for inter-fraction variations could be a valuable tool for this work.

Reducing variations

Breath-coaching and audio and/or visual feedback can be used to try and reduce intra- and inter-fraction variations in the respiratory motion. A number of methods have been presented in the literature [22, 26], and although they have generally shown promising results, they have only entered limited clinical use. The methods have usually been assessed using surrogate data, and it has been shown that they can reduce the variation of the surrogate signals. However, the effects of these methods on the internal motion, (and on the stability of the correspondence models) has not been fully assessed. The studies also show that using feedback is not beneficial for all patients, and in some cases the feedback can make the variation worse. Additionally, these methods will be of limited or no help with inter-fraction variations due to anatomical changes, as these cannot be accounted for just by having the patient breathing in a reproducible way. In conclusion, breath-coaching and feedback methods could potentially be very useful, particularly for gated or tracked treatments, but more work is required before firm conclusions can be drawn.

Amount of data used to build and validate models

Ideally the internal motion data used to fit the motion model would sample all the types of variation that the model will be used to estimate (i.e. intra-fraction and inter-fraction variation). However, if ionising radiation is used for the imaging then this limits the number of images that can be acquired. Even if dose is not an issue (e.g. if MRI used) it may not be practical or possible to acquire data over the time frames required to sample all of the variation. The amount of data required will also be affected by the use of adaptive fitting methods. If extra data is acquired to update the models to adapt to intra- or inter-fraction variation then it is not necessary to sample all the variation in the initial data. More studies are required to determine the optimal amount of data to acquire in different circumstances in order to find a good trade off between scanning time and resources and model accuracy.

The data used to validate the models should also sample all expected types of variations in the motion. But, as with the data used to build the models it is often difficult to acquire sufficient data to thoroughly validate the models. When designing and investigating models they can be validated using retrospectively acquired data. However, if the models are to be used clinically then it is desirable to have some way of validating them prior to and during use. The difficulty in acquiring good validation data, and hence the lack of confidence in the accuracy of models, could

be one of the main reasons that such models have not yet been more adopted in clinical use.

Computation time

The computation time required to build the motion models can vary greatly. It will depend heavily on the way the motion is measured from the imaging data and represented in the models, e.g. if deformable registrations of several large 3D volumes is required this can take from a few minutes to several hours depending on the algorithms and implementations used. The model fitting itself is generally quite fast but obviously depends on the method used and the amount of data being fit. Often in the literature the imaging data used to build the models is acquired several days before the model is needed (i.e. at treatment time), so the time required to build the models is not a major concern. However, if the models are to be updated or rebuilt for every fraction of treatment in order to account for inter-fraction variations then the time required to construct the models will be an important factor.

If the models are to be used to guide gated or tracked treatments then the motion estimates are required in 'real-time' so the computation time required to calculate them from the surrogate data will be important. Generally this can be done in a fraction of a second for most models, even if estimating a dense deformation field. Indirect models (see Section 1.4.3) may take longer as they need to optimise the motion estimate to best match the surrogate data. However, even if the motion estimates can be calculated quickly they do require some time, as does acquiring and processing the surrogate data and determining how to modify the treatment, so there will always be some degree of latency to treatment which will negatively affect the accuracy. To account for this latency methods that can predict the surrogate signals or the internal motion a short time in the future based on current and past measurements need to be used. These are discussed in detail in Chapter 12. One interesting approach is to combine the predict ahead method and the correspondence model into a single model which can predict the motion estimate a short time in the future from the current and past surrogate signals [29].

Conclusion

In conclusion respiratory motion models that can estimate the internal respiratory motion from easily acquired surrogate signals clearly have many potential uses. There have been a large number of different but closely related methods presented in the literature. This chapter has detailed and discussed the most commonly used methods from the literature, and considered some of the issues relevant to designing and using such motion models for RT. At the time of writing these models remain very much as research proposals, with relatively little clinical use or commercial

products. The continued output of research papers proposing or using such models, both for radiotherapy and other applications, would indicate that the models hold some promise, but more work is required to establish the accuracy and robustness of the models in clinical circumstances before they become more widely adopted.

References

1. Ahn, S., Yi, B., Suh, Y., Kim, J., Lee, S., Shin, S., Shin, S., Choi, E.: A feasibility study on the prediction of tumour location in the lung from skin motion. The British Journal of Radiology **77**, 588–596 (2004)
2. Beddar, A.S., Kainz, K., Briere, T.M., Tsunashima, Y., Pan, T., Prado, K., Mohan, R., Gillin, M., Krishnan, S.: Correlation between internal fiducial tumor motion and external marker motion for liver tumors imaged with 4D-CT. Int. J. Radiation Oncology Biol. Phys. **67**, 630–638 (2007)
3. Benchetrit, G.: Breathing pattern in humans: diversity and individuality. Respiration Physiology **122**, 123–129 (2000)
4. Berbeco, R.I., Nishioka, S., Shirato, H., Chen, G.T.Y., Jiang, S.B.: Residual motion of lung tumours in gated radiotherapy with external respiratory surrogates. Physics in Medicine and Biology **50**, 3655–3667 (2005)
5. Blackall, J.M., Ahmad, S., Miquel, M.E., McClelland, J.R., Landau, D.B., Hawkes, D.J.: MRI-based measurements of respiratory motion variability and assessment of imaging strategies for radiotherapy planning. Physics in Medicine and Biology **51**, 4147–4169 (2006)
6. Cervino, L.I., Chao, A.K.Y., Sandhu, A., Jiang, S.B.: The diaphragm as an anatomic surrogate for lung tumor motion. Physics in Medicine and Biology **54**, 3529–3541 (2009)
7. Cervino, L.I., Jiang, Y., Sandhu, A., Jiang, S.B.: Tumor motion prediction with the diaphragm as a surrogate: a feasibility study. Physics in Medicine and Biology **55**, N221–N229 (2010)
8. Chi, P.C.M., Balter, P., Luo, D., Mohan, R., Pan, T.: Relation of external surface to internal tumor motion studied with cine CT. Medical Physics **33**, 3116–3123 (2006)
9. Cho, B., Poulsen, P.R., Keall, P.J.: Real-time tumor tracking using sequential kV imaging combined with respiratory monitoring: a general framework applicable to commonly used IGRT systems. Physics in Medicine and Biology **55**, 3299–3316 (2010)
10. Cho, B., Poulsen, P.R., Sawant, A., Ruan, D., Keall, P.J.: Real-time target position estimation using stereoscopic kilovoltage/megavoltage imaging and external respiratory monitoring for dynamic multileaf collimator tracking. Int. J. Radiation Oncology Biol. Phys. **79**, 269–278 (2011)
11. Cho, B., Suh, Y., Dieterich, S., Keall, P.J.: A monoscopic method for real-time tumour tracking using combined occasional x-ray imaging and continuous respiratory monitoring. Physics in Medicine and Biology **53**, 2837–2855 (2008)
12. Colgan, R., McClelland, J., McQuaid, D., Evans, P.M., Hawkes, D., Brock, J., Landau, D., Webb, S.: Planning lung radiotherapy using 4D CT data and a motion model. Physics in Medicine and Biology **52**, 5815–5830 (2008)
13. Ernst, F., Bruder, R., Schlaefer, A., Schweikard, A.: Correlation between external and internal respiratory motion: a validation study. International Journal of Computer Assisted Radiology and Surgery (2011). DOI 10.1007/s11548-011-0653-6
14. Ernst, F., Martens, V., Schlichting, S., Beširević, A., Kleemann, M., Koch, C., Petersen, D., Schweikard, A.: Correlating chest surface motion to motion of the liver using ε-SVR – a porcine study. In: Proceedings of Medical Image Computing and Computer-Assisted Interventions (MICCAI), pp. 356–364 (2009)
15. Fayad, H., Clement, J.F., Pan, T., Roux, C., Rest, C.C.L., Pradier, O., Visvikis, D.: Towards a generic respiratory motion model for 4D CT imaging of the thorax. In: IEEE Nuclear Science Symposium Conference Record (2009)

16. Fayad, H., Pan, T., Clement, J.F., Visvikis, D.: Technical note: Correlation of respiratory motion between external patient surface and internal anatomical landmarks. Medical Physics **38**, 3157–3164 (2011)
17. Fayad, H., Pan, T., Roux, C., Rest, C.C.L., Pradier, O., Clement, J.F., Visvikis, D.: A patient specific respiratory model based on 4D CT data and a time of flight camera (TOF). In: IEEE Nuclear Science Symposium Conference Record (2009)
18. Fayad, H., Pan, T., Roux, C., Rest, C.C.L., Pradier, O., Visvikis, D.: A 2D-spline patient specific model for use in radiation therapy. In: Proceedings of the International Symposium on Biomedical Imaging (ISBI), pp. 590–593 (2009)
19. Fayad, H., Pan, T., Roux, C., Visvikis, D.: A generic respiratory motion model for motion correction in PET/CT. In: IEEE Nuclear Science Symposium Conference Record, pp. 2455–2458 (2010)
20. Gao, G., McClelland, J., Tarte, S., Blackall, J., Hawkes, D.: Modelling the respiratory motion of the internal organs by using canonical correlation analysis and dynamic MRI. In: The First International Workshop on Pulmonary Image Analysis held at MICCAI 2008 (2008)
21. Geneser, S., Hinkle, J., Kirby, R., Wang, B., Salter, B., Joshi, S.: Quantifying variability in radiation dose due to respiratory-induced tumor motion. Medical Image Analysis **15**, 640–649 (2011)
22. George, R., Chung, T.D., Vedam, S.S., Ramakrishnan, V., Mohan, R., Weiss, E., Keall, P.J.: Audio-visual biofeedback for respiratory-gated radiotherapy: Impact of audio instruction and audio-visual biofeedback on respiratory-gated radiotherapy. International Journal of Radiation Oncology*Biology*Physics **65**(3), 924 – 933 (2006). DOI 10.1016/j.ijrobp.2006.02.035. URL http://www.sciencedirect.com/science/article/pii/S0360301606003713
23. Hinkle, J., Fletcher, P.T., Wang, B., Salter, B., Joshi, S.: 4D MAP image reconstruction incorporating organ motion. In: Proceedings of the International Conference on Information Processing in Medical Imaging (IPMI), pp. 676–687 (2009)
24. Hoisak, J.D.P., Sixel, K.E., Tirona, R., Cheung, P.C.F., philippe Pignol, J.: Correlation of lung tumour motion with external surrogate indicators of respiration. Int. J. Radiation Oncology Biol. Phys. **60**, 1298–1306 (2004)
25. Hoogeman, M., briac Prévost, J., Nuyttens, J., Levendag, J.P.P., Heijmen, B.: Clinical accuracy of the respiratory tumor tracking system of the cyberknife: assessment by analysis of log files. Int. J. Radiation Oncology Biol. Phys. **74**, 297–303 (2009)
26. Hughes, S., McClelland, J., Tarte, S., Blackall, J., Liong, J., Ahmad, S., Hawkes, D., Landau, D.: Assessment of respiratory cycle variability with and without training using a visual guide. Cancer Therapy **6**, 945–954 (2008)
27. Hughes, S., McClelland, J., Tarte, S., Lawrence, D., Ahmad, S., Hawkes, D., Landau, D.: Assessment of two novel ventilatory surrogates for use in the delivery of gated/tracked radiotherapy for non-small cell lung cancer. Radiotherapy and Oncology **91**, 336–341 (2009)
28. Ionascu, D., Jiang, S.B., Nishioka, S., Shirato, H., Berbeco, R.I.: Internal-external correlation investigations of respiratory induced motion of lung tumors. Medical Physics **34**, 3893–3903 (2007)
29. Isaksson, M., Jalden, J., Murphy, M.J.: On using an adaptive neural network to predict lung tumor motion during respiration for radiotherapy applications. Medical Physics **32**, 3801–3809 (2005)
30. Jolliffe, I.T.: Principal Component Analysis (Springer Series in Statistics). Springer (2002)
31. King, A.P., Rhode, K.S., Ma, Y., Yao, C., Jansen, C., Razavi, R., Penney, G.P.: Registering preprocedure volumetric images with intraprocedure 3-D ultrasound using an ultrasound imaging model. IEEE Transactions on Medical Imaging **29**(3), 924–937 (2010)
32. Klinder, T., Lorenz, C., Ostermann, J.: Free-breathing intra- and intersubject respiratory motion capturing, modeling, and prediction. In: Proceedings of SPIE Medical Imaging 2009: Image Processing (2009)
33. Klinder, T., Lorenz, C., Ostermann, J.: Prediction framework for statistical respiratory motion modelling. In: Proceedings of Medical Image Computing and Computer-Assisted Interventions (MICCAI) 2010, pp. 327–334 (2010)

34. Koch, N., Liu, H.H., Starkschall, G., Jacobson, M., Forster, K., Liao, Z., Komaki, R., Stevens, C.W.: Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: Part i correlating internal lung motion with skin fiducial motion. Int. J. Radiation Oncology Biol. Phys. **60**, 1459–1472 (2004)
35. Li, G., Arora, N.C., Xie, H., Ning, H., Lu, W., Low, D., Citrin, D., Kaushal, A., Zach, L., Camphausen, K., Miller, R.W.: Quantitative prediction of respiratory tidal volume based on the external torso volume change: a potential volumetric surrogate. Physics in Medicine and Biology **54**, 253–267 (2009)
36. Li, R., Lewis, J.H., Jia, X., Gu, X., Folkerts, M., Men, C., Song, W.Y., Jiang, S.B.: 3D tumor localization through real-time volumetric x-ray imaging for lung cancer radiotherapy. Medical Physics **38**, 2783–2794 (2011)
37. Li, T., Schreibmann, E., Yang, Y., L, X.: Motion correction for improved target localization with on-board cone-beam computed tomography. Physics in Medicine and Biology **51**, 253–267 (2006)
38. Li, T., Thorndyke, B., Schreibmann, E., Yang, Y., Xing, L.: Model-based image reconstruction for four-dimensional PET. Medical Physics **33**(5), 1288–1298 (2006)
39. Liu, H.H., Koch, N., Starkschall, G., Jacobson, M., Forster, K., Liao, Z., Komaki, R., Stevens, C.W.: Evaluation of internal lung motion for respiratory-gated radiotherapy using MRI: part II - margin reduction of internal target volume. Int. J. Radiation Oncology Biol. Phys. **60**, 1473–1483 (2004)
40. Liu, X., Oguz, I., Pizer, S.M., Mageras, G.S.: Shape-correlated deformation statistics for respiratory motion prediction in 4D lung. In: Proceedings of SPIE Medical Imaging 2010 (2010)
41. Low, D.A., Parikh, P., Lu, W., Dempsey, J., Wahab, S., Hubenschmidt, J., Nystrom, M., Handoko, M., Bradley, J.: Novel breathing motion model for radiotherapy. Int. J. Radiation Oncology Biol. Phys. **63**, 921–929 (2005)
42. Low, D.A., Zhao, T., White, B., Yang, D., Mutic, S., Noel, C.E., Bradley, J.D., Parikh, P.J., Lu, W.: Application of the continuity equation to a breathing motion model. Medical Physics **37**, 1360–1364 (2010)
43. Lu, W., Low, D.A., Parikh, P.J., Nystrom, M.M., El Naqa, I.M., Wahab, S.H., Handoko, M., Fooshee, D., Bradley, J.D.: Comparison of spirometry and abdominal height as four-dimensional computed tomography metrics in lung. Medical Physics **32**, 2351–2357 (2005)
44. Martin, J., McClelland, J., Thomas, C., Wildermuth, K., Landau, D., Ourselin, S., Hawkes, D.: Motion modelling and motion compensated reconstruction of tumours in cone-beam computed tomography. In: Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA) (2012)
45. McClelland, J.R., Chandler, A.G., Blackall, J.M., Ahmad, S., Landau, D., Hawkes, D.J.: 4D motion models over the respiratory cycle for use in lung cancer radiotherapy planning. In: Proceedings of SPIE Medical Imaging 2005: Visualization, Image-Guided Procedures, and Display (2005)
46. McClelland, J.R., Chandler, A.G., Blackall, J.M., Tarte, S., Hughes, S., Ahmad, S., Landau, D., Hawkes, D.J.: A continuous 4D motion model from multiple respiratory cycles for use in lung radiotherapy. Medical Physics **33**, 3348–3358 (2006)
47. McClelland, J.R., Hawkes, D.J., Schaeffter, T., King, A.P.: Respiratory motion models: A review. Medical Image Analysis (in press 2012)
48. McClelland, J.R., Hughes, S., Modat, M., Qureshi, A., Ahmad, S., Landau, D.B., Ourselin, S., Hawkes, D.J.: Inter-fraction variations in respiratory motion models. Physics in Medicine and Biology **56**, 251–272 (2011)
49. McClelland, J.R., Webb, S., McQuaid, D., Binnie, D.M., Hawkes, D.J.: Tracking differential organ motion with a breathing multileaf collimator: magnitude of problem assessed using 4D CT data and a motion-compensation strategy. Physics in Medicine and Biology **52**, 4805–4826 (2007)
50. de Mey, J., de Steene, J.V., Vandenbroucke, F., Verellen, D., Trappeniers, L., Meysman, M., Everaert, H., Noppen, M., Storme, G., Bossuyt, A.: Percutaneous placement of marking coils before stereotactic radiation therapy of malignant lung lesions. Journal of Vascular and Interventional Radiology **16**(1), 51 – 56 (2005). DOI 10.1097/01.RVI.0000142599.

48497.6B. URL http://www.sciencedirect.com/science/article/pii/S1051044307605995
51. Mori, S., Kanematsu, N., Mizuno, H., Sunaoka, M., Endo, M.: Physical evaluation of CT scan methods for radiation therapy planning: comparison of fast, slow and gating scan using the 256-detector row CT scanner. Physics in Medicine and Biology **51**, 587–600 (2006)
52. Odille, F., Vuissoz, P.A., Marie, P.Y., Felblinger, J.: Generalized reconstruction by inversion of coupled systems (GRICS) applied to free-breathing MRI. Magnetic Resonance in Medicine **60**(1), 146–157 (2008)
53. Plathow, C., Zimmermann, H., Fink, C., Umathum, R., Schbinger, M., Huber, P., Zuna, I., Debus, J., Schlegel, W., peter Meinzer, H., Semmler, W., ulrich Kauczor, H., Bock, M.: Influence of different breathing maneuvers on internal and external organ motion: use of fiducial markers in dynamic MRI. Int. J. Radiation Oncology Biol. Phys. **62**, 238–245 (2005)
54. Rit, S., Wolthaus, J.W.H., van Herk, M., Sonke, J.J.: On-the-fly motion-compensated cone-beam CT using an a priori model of the respiratory motion. Medical Physics **36**, 2283–2296 (2009)
55. Ruan, D., Fessler, J.A., Balter, J.M., Berbeco, R.I., Nishioka, S., Shirato, H.: Inference of hysteretic respiratory tumour motion from external surrogates: a state augmentation approach. Physics in Medicine and Biology **53**, 2923–2936 (2008)
56. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: Application to breast MR images. IEEE Transactions on Medical Imaging **18**(8), 712–721 (1999)
57. Schweikard, A., Glosser, G., Bodduluri, M., Murphy, M.J., Adler, J.R.: Robotic motion compensation for respiratory movement during radiosurgery. Computer Aided Surgery **5**, 263–277 (2000)
58. Schweikard, A., Shiomi, H., Adler, J.: Respiration tracking in radiosurgery. Medical Physics **31**, 2738–2741 (2004)
59. Schweikard, A., Shiomi, H., Adler, J.: Respiration tracking in radiosurgery without fiducials. Int J Medical Robotics and Computer Assisted Surgery **1**, 19–27 (2005)
60. Schweikard, A., Shiomi, H., Fisseler, J., Dötter, M., Berlinger, K., Gehl, H.B., Adler, J.: Fiducial-less respiration tracking in radiosurgery. In: Proceedings of Medical Image Computing and Computer-Assisted Interventions (MICCAI) 2004, pp. 992–999 (2004)
61. Seppenwoolde, Y., Berbeco, R.I., Nishioka, S., Shirato, H., Heijmen, B.: Accuracy of tumour motion compensation algorithm from a robotic respiratory tracking system: A simulation study. Medical Physics **34**, 2774–2784 (2007)
62. Shirato, H., Harada, T., Harabayashi, T., Hida, K., Endo, H., Kitamura, K., Onimaru, R., Yamazaki, K., Kurauchi, N., Shimizu, T., et al.: Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy. International Journal of Radiation Oncology, Biology, Physics **56**(1), 240–247 (2003). URL http://linkinghub.elsevier.com/retrieve/pii/S0360301603000762
63. Sonke, J.J., Belderbos, J.: Adaptive radiotherapy for lung cancer. Seminars in Radiation Oncology **20**(2), 94 – 106 (2010). DOI 10.1016/j.semradonc.2009.11.003. URL http://www.sciencedirect.com/science/article/pii/S1053429609000770. ¡ce:title¿Adaptive Radiotherapy¡/ce:title¿
64. Torshabi, A.E., Pella, A., Riboldi, M., Baroni, G.: Targeting accuracy in real-time tumor tracking via external surrogates: A comparative study. Technology in Cancer Research and Treatment **9**, 1–11 (2010)
65. Vandemeulebroucke, J., Kybic, J., Clarysse, P., Sarrut, D.: Respiratory motion estimation from cone-beam projections using a prior model. In: Proceedings of Medical Image Computing and Computer-Assisted Interventions (MICCAI) 2009 (2009)
66. Vedam, S.S., Kini, V.R., Keall, P.J., Ramakrishnan, V., Mostafavi, H., Mohan, R.: Quantifying the predictability of diaphragm motion during respiration with a noninvasive external marker. Medical Physics **30**, 505–513 (2003)
67. Xu, Q., Hamilton, R.: A novel respiratory detection method based on automated analysis of ultrasound diaphragm video. Medical Physics **33**, 916–921 (2006)

68. Yang, D., Lu, W., Low, D.A., Deasy, J.O., Hope, A.J., Naqa, I.E.: 4D-CT motion estimation using deformable image registration and 5D respiratory motion modeling. Medical Physics **35**, 4577–4590 (2008)
69. Zhang, Q., Hu, Y.C., Liu, F., Goodman, K., Rosenzweig, K.E., Mageras, G.S.: Correction of motion artifacts in cone-beam CT using a patient-specific respiratory motion model. Medical Physics **37**, 2901–2909 (2010)
70. Zhang, Q., Pevsner, A., Hertanto, A., Hu, Y.C., Rosenzweig, K.E., Ling, C.C., Mageras, G.S.: A patient-specific respiratory model of anatomical motion for radiation treatment planning. Medical Physics **34**, 4772–4781 (2007)
71. Zhao, T., Lu, W., Yang, D., Mutic, S., Noel, C.E., Parikh, P.J., Bradley, J.D., Low, D.A.: Characterization of free breathing patterns with 5D lung motion model. Medical Physics **36**, 5183–5189 (2009)

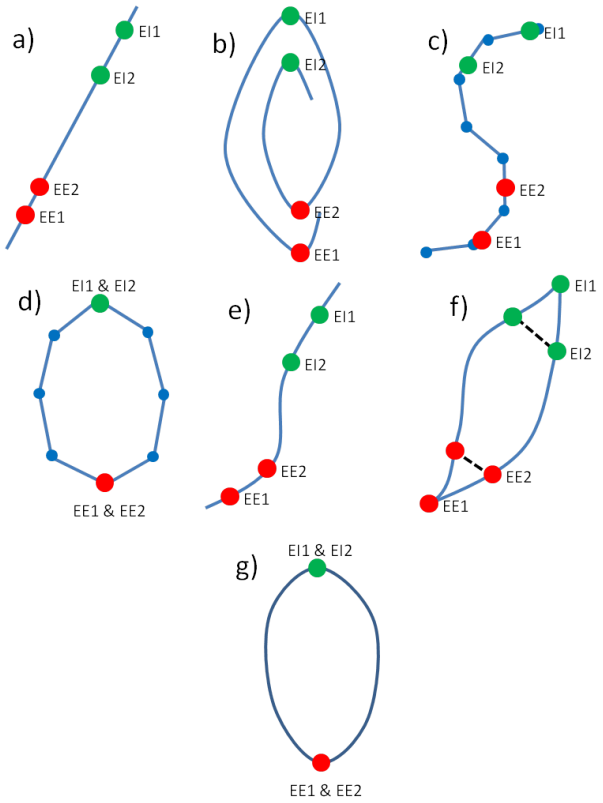

Fig. 1.4 Examples of the different motion trajectories that can be produced by the different types of correspondence model. The points corresponding to end-exhalation (red circles) and end-inhalation (green circles) from two breath cycles are marked (EE1, EI1, EE2, EI2). a) A linear correspondence model using one surrogate signal constrains the motion to follow a straight line during every breath cycle. The estimated motion may move a different distance along the line during each breath, depending on how deep the breathing is (i.e. EE1 $\neq$ EE2 and EI1 $\neq$ EI2). b) A linear (or other) correspondence model that uses two or more surrogate signals can model complex motion including hysteresis (a different trajectory during inhalation and exhalation) and inter-cycle variation (EE1 $\neq$ EE2 and EI1 $\neq$ EI2). c) A piece-wise linear model produces a motion trajectory that is made up of several straight line segments. The blue circles correspond to known motion measurements made from the imaging data. As with the linear model, the estimated motion is constrained to follow the same path during every breath, but may move a different distance along the path depending on how deeply the individual is breathing. d) If the surrogate signal used for the piece-wise linear model is a derived respiratory phase signal rather than the actual measured surrogate signal then the line segments will form a loop shaped trajectory. In this case the estimated motion will go around the same loop during every breath cycle. This means that hysteresis can be modelled, but inter-cycle variation cannot (i.e. EE1 $=$ EE2 and EI1 $=$ EI2). e) A polynomial correspondence model also constrains the motion to follow the same path during every breath, but now the path is a curve rather than one or more straight lines. f) In order to model hysteresis some authors have proposed using two polynomial models, one for inhalation and one for exhalation. While this does allow hysteresis to be modelled it may result in discontinuous motion estimates when switching between the two models, as shown at EE2 and EI2. g) Using respiratory phase as the surrogate signal and a periodic B-spline correspondence model gives a similar trajectory to using phase with a piece-wise linear model, but the trajectory is now a smooth curve rather than being composed of a number of straight line segments.

Index