Learning Varying Dimension Radial Basis Functions for Deformable Image Alignment

Di YANG
RSISE, Australian National University
di.yang@anu.edu.au

Hongdong LI
NICTA & Australian National University
Hongdong.Li@anu.edu.au

Abstract

This paper presents a method for learning Radial Basis Functions (RBF) model with variable dimensions for aligning/registrating images of deformable surface. Traditional RBF-based approach, which is mainly based on a fixed dimension parametric model, often suffers from severe parameter over-fitting and complicated model selection (i.e. select the number and locations of centers determination) problems which lead to inaccurate estimation and unreliable convergence. Our strategy for solving both the parameter over-fitting and model selection problems is through the use of a probabilistic Bayesian inference model to obtain a posterior estimation of the alignment as well as the model parameters simultaneously. To learn the parameters of the Bayesian model, a reversible jump Markov Chain Monte Carlo (RJMCMC) algorithm is employed, allowing us to handle large deformation image registration. Our approach is demonstrated successfully on real image sequences of different deformation types, with results compared favorable against other existing approaches.

1. Introduction

In computer vision, sets of data acquired by sampling the same object at different times, or from different perspectives, will be in different coordinate systems. Image alignment is the process of deforming the different sets of data into one coordinate system. Alignment is necessary in order to be able to compare or integrate the data obtained from different measurements. This task is very important in some application domains, especially for medical image analysis and vision based robot. In this paper, our goal is to computer this dense image non-rigid deformation, mapping pixels from one image to corresponding pixels in the other images. We present a new learning varying dimension radial basis functions model to obtain an optimal estimation of the parameters of deformation model.

1.1. Related Work

Many techniques have been specifically designed to exploit a deformable image alignment methodology, such as Lucas-Kanade algorithm, mutual information method, coarse-to-fine strategy, etc [14]. While in deformable image alignment domain, most simplification approaches to provide a regularized optic-flow with compare intensity patterns in images via correlation metrics.

Demons Algorithm is a famous and widely applied intensity-based approach. Some detailed description of the method and some impressive applications can be found in papers [12]. This algorithm is to compute a regularized optic flow field, by minimizing an free energy function based on data term with Maxwell's demons force, such as a regularizer, encouraging smoothness of the optic-flow. the essence of demons algorithm is combination between feature driven and intensity driven. Driven by the demons forces, Demons Algorithm represents an excellent performance in homogeneous object deformation alignment domain, but insensitive to texture-rich situation.

Direct Estimation Radial Basis Mapping Approach is another impressive texture sensitive algorithm, described in [1], presented by A. Bartoli and A. Zisserman. RBF framework has been applied to establish a parameterized model for deformation of objects. For centers selection, they also provided an impressive dynamic center insertion strategy by employed model selection method as a stopping criterion, *e.g.* Akaike Information Criterion (AIC). Finally, this intensity driven parameterized approach does not only provide the possibility to incorporate feature-based methods in deformable image alignment domain, except some drawbacks need to be overcome, but also is the inspiration of our approach.

1.2. Limitations

In the non-rigid deformation image alignment domain, there are two major tasks need to be accomplished, such as model selection and parameter over-fitting.

Model selection is the task of selecting a statistical model

2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops
978-1-4244-4441-0/09/$25.00 ©2009 IEEE

from a set of potential models, given data. In the RBF framework, the model selection task is interpreted as determination of the number and locations of control centers in the observation. In dynamic centers insertion strategy, locations of centers is chosen under constrains, sort of residual and minimum distance between adjacent centers, while the number of centers is determined by AIC. Although, this strategy can determine the 'best' number of centers, it fail to adaptively insert the centers into effective locations, due to minimum distance need predefined by hand. That means the various minimum distance is set, the different alignment performance will be obtained. The dense distribution of inserted centers mean over-fitting problem, while spares centers distribution causes under-fitting problem.

Whereas, parameters over-fitting problem is actually caused by ill-pose problem in the inversible model. The failure of capture the motion parameters, *i.e.* missing data points, may be led by featureless found in image regions. According to this reason, the inaccurate pseudo inverse solution will make parameterized model suffer with unreliable convergence, getting trapped into local minima, see [2], especially in high degree of freedom situation. This problem has been found by several literatures, *e.g.* [11]. Tikhonov regularization is the most commonly used method of regularization of ill-posed problems to overcome over-fitting problem. And this method has been widely used in image registration domain, *e.g.* [9]. In statistics, it is also known as ridge regression. This method is related to the Levenberg-Marquardt algorithm for non-rigid least squares problems. However, as mentioned perviously, the regularization method, adding a small, positive regularization term the least-squares solution, could not solve this problem fundamentally in parameterized deformable image alignment model, if without appropriate model determination.

1.3. Contributions

We present a new feature-based approach to obtain a posterior determination of appropriate dimensions of the model and more accurate parameters estimation simultaneously under Bayesian learning inference. Through embedding features regression of deformable image alignment, our approach is able to understand the motion of deformation and automatically align corresponding images. Radial basis parameterized model is described in §2. Bayesian prior information is incorporated to make posterior estimation of parameter and determination of dimension in §3. In §4, the reversible jump Markov Chain Monte Carlo algorithm has been proposed to approximate posterior density of this radial basis model. Training data generation and Features extraction and tracking is discussed in §5. Then, we use our approach to deal with image sequences and report experiments results in §6. Finally, we give our conclusions and discuss further work in §7.

2. Radial Basis Mapping Deformation

We now formulate the problem of interpreting non-rigid deformable image alignment. Our approach concerned the motion estimation procedure as the regression of a target variable Y on an input set of covariates X given the data pairings $\mathcal{D} = \{(\mathrm{y}_1, \mathbf{x}_1), (\mathbf{y}_2, \mathbf{x}_2), \ldots, (\mathbf{y}_n, \mathbf{x}_n)\}$, which is an important topic in data analysis. The regression curve or surface is assumed to be the conditional mean function $m(\mathbf{x}) = E[Y|X = \mathbf{x}]$ and the recorded observations Y to be corrupted with gaussian noise ϵ, so that

$$\mathbf{y}_i = m(\mathbf{x}_i) + \epsilon_i, \quad i = 1, \ldots, n. \tag{1}$$

where ϵ_i are independent and identically distributed (i.i.d.) $\sim N(0, \sigma^2)$. In this paper, we concerned with the approximation of $m(\mathbf{x})$ by an estimate $\hat{m}(\mathbf{x})$ of the form

$$\hat{m}(\mathbf{x}_i) = \sum_{j=1}^{K} w_j \phi_j(\|\mathbf{x}_i - \mu_j\|), \tag{2}$$

where $\{\phi_j(\mathbf{x})\}_{j=1,\ldots,n}$ represents a radial basis function (RBF) and w are the output coefficients (weights). RBF networks have proved a popular method for approximating data, and a review could be found in [10]. The models are feed-forward networks of radial functions where each basis is parameterized by a knot or position vector μ located in the $d-$dimensional covariate space X. Conventionally there are as many basis functions as data points to be approximated with the position vector set to the data values, especially the non-rigid deformable image alignment. The model output $m(\mathbf{x})$ could be rewritten as a linear combination of the basis functions response and a low-order polynomial term,

$$m(\mathbf{x}) = \sum_{j=1}^{K} w_j \phi_j(\|\mathbf{x} - \mu_j\|) + \sum_{n=0}^{p} c_n b_n(\mathbf{x}), \tag{3}$$

where $\|\cdot\|$ denotes a distance metric, usually Euclidean or Mahalanobis, and $c_n(\mathbf{x})$ represents a polynomial of degree n. The coefficients w and c are calculated by least squares where the constraint $\{\sum_{i=1}^{N} w_j c_n(\mathbf{x}_i) = 0\}_{j=0,\ldots,p}$, is imposed to ensure the uniqueness of the solution.

In the domain of non-rigid deformable image alignment, the low-order polynomial term in equation 3 could be interpreted as global affine transformation parameterized model, which has been used successfully for estimating and representing the optic flow of smooth textured surfaces and incorporated in as much literatures, *e.g.*[7]. Additionally, this model has been applied locally within small image regions, and globally, for applications such as tracking, moscaicing and stabilization. Thus the low order polynomial term $\sum_{n=0}^{p} c_n b_n(\mathbf{x})$ could be rewritten as optical flow field $u(\mathbf{x})$ over the positions $\mathbf{x} = (x, y)$ by a weighted sum of basis flow fields:

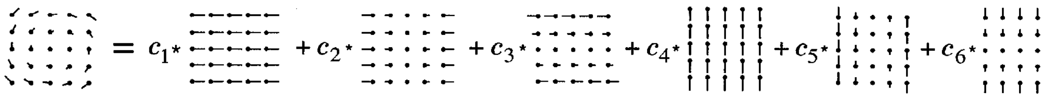


Figure 1. Illustration of affine optic-flow model of order six

$$u(\mathbf{x}) = \sum_{n=1}^{6} c_n \mathbf{b}_n(\mathbf{x}) \tag{4}$$

where $\{\mathbf{b}_n(\mathbf{x})\}_{n=1,\dots,6}$ are six basis sets, while affine model required, encoding horizontal and vertical translation, shearing and scaling encoding respectively. And $c = \{c_1, \dots, c_6\}$ is the vector containing the scalar coefficients. Meanwhile, affine transformation is also modelled by a 2×3 matrix $A = (\bar{A}, \mathbf{d})$, where $\bar{A}$ is the leading 2×2 submatrix and $\mathbf{d} = (\mathbf{d}_x, \mathbf{d}_y)^T$ the last column. A better alternative is to enrich the description of motion, that is, to define a set of possible optical flow field functions $u(\mathbf{x}) = \{u^x(\mathbf{x}), u^y(\mathbf{x})\}^T$. Then, a point $\mathbf{x}$ is mapped from the first image I to point $\mathbf{x}'$ in the second image J. Thus, the affine motion model can be summarized by the following displacement equation:

$$u(\mathbf{x}) = \begin{pmatrix} u^x(\mathbf{x}) \\ u^y(\mathbf{x}) \end{pmatrix} = \bar{A}\mathbf{x} + \mathbf{d}. \tag{5}$$

In the basis form, RBFs define a mapping from $\mathbb{R}^d$ to $\mathbb{R}$, where d is the known data space X dimension. A 2D non-rigid deformable optical flow fields $u(\mathbf{x})$ with affine low order polynomial term is defined by $\mathbb{R}^2 \to \mathbb{R}$ by basis function $\{\phi_i(\|\mathbf{x} - \mu_i\|)\}_{i=1,\dots,K}$, and a set of K centers $\{\mu_i\}_{i=1,\dots,K}$. However, in image alignment, the domain need be mapped to $\mathbb{R}^2 \to \mathbb{R}^2$ radial basis functions. The usual way to construct this mapping m, is to stack 2 RBFs $\delta^x(\mathbf{x})$ and $\delta^y(\mathbf{x})$ sharing their centers. For 2D image motion with global affine deformation, the equation can be written as:

$$m(\mathbf{x}) = \begin{pmatrix} \delta^x(\mathbf{x}) \\ \delta^y(\mathbf{x}) \end{pmatrix} = \bar{A}\mathbf{x} + \mathbf{d} + \sum_{j=1}^{K} \begin{pmatrix} w_j^x \\ w_j^y \end{pmatrix} \phi_j(\|\mathbf{x} - \mu_j\|) \tag{6}$$

where $\{\delta^x(\mathbf{x}), \delta^y(\mathbf{x})\}^T$ and $(w_j^x, w_j^y)^T$ represent encoding horizontal and vertical optical flow fields and related coefficients respectively.

The theory of RBFs specifies many permissible forms that the function $\{\phi_j(\|\mathbf{x} - \mu_j\|)\}_{j=1,\dots,K}$ can take. Here, we choose to analyze thin-plate-spline $\phi(z) = z^2 \log z$, which has an elegant algebra expressing the dependence of physical bending energy of a thin plate on points [3].

For the case of image alignment circumstance, we need to reduce the flexibility of the model, which means stabilize the image alignment algorithm to overcome over-fitting problem. In the neural network community, it is more standard to control complexity by reducing the number of centers to less than the data points, this suggestion first advocated by Broomhead and Lowe [4]. The reduction in the dimension leads to simple models with less variance but greater bias. In extreme with no radial functions, we recover the standard linear model. Reduction of flexibility can also be achieved by adding a small, positive regularization term η to the diagonal terms in the design matrix of the least-squares solution of $\{w_j\}_{j=1,\dots,K}$ in equation 3, suggested by Wahba [13]. The larger the value of η is, the smoother the output of the model will be.

These two approaches share a common motivation: to reduce the complexity of the deformation model given by equation 6 and to overcome over-fitting problem. However, these the regularization is too restrictive in many cases or regions in image and too loose for others. For example, when tracking a face, deformation of the jaw are much more likely than deformations of the nose. Moreover, the weight for this regularization term must be specified by hand. Therefore in next section, we describe a Bayesian learning framework for the RBF network to fundamentally reduce flexibility by explicitly incorporating in a prior density that we define over the model-parameter space Meanwhile, another important key to deformable image alignment with this model is centers selection and centers linear embedding.

3. Bayesian radial basis functions of varying dimension

After establish the radial basis mapping deformation function $m(\mathbf{x})$ and according to equation (1), these points in the image frames are regressed under the data pairing $\mathcal{D}$:

$$\mathbf{y}_i = \bar{A}\mathbf{x}_i + \mathbf{d} + \sum_{j=1}^{K} \begin{pmatrix} w_j^x \\ w_j^y \end{pmatrix} \phi_j(\|\mathbf{x}_i - \mu_j\|) + \epsilon_i \tag{7}$$

where ϵ is 2D zero-mean gaussian noise with covariance $\Lambda = \sigma^2\mathbf{I}$, $\mathbf{I}$ is 2×2 identity matrix. As mentioned in last section, to overcome overfitting, the regularization could hardly solve it fundamentally. Alternative approach is place a user-specified Gaussian prior on the deformation basis and a prior on the deformations based on an initial estimate.

In order to motivate our approach, we restate the above techniques as follows. Bayesian framework is the calculation of a probability distribution on the unknown parameter (weight) vector θ. And in this paper, we also wish to take account of uncertainty between model $\mathcal{M}$ of different dimension *i.e.* number of basis functions. Thus we make this deformable image optical flow field estimation explicit by writing the expectation as

$$E(Y|\mathbf{x},\mathcal{D}) = \sum_{k=0}^{K} \int m(\mathbf{x},\theta_k,M_k)p(\theta_k|M_k,\mathcal{D})p(M_k|\mathcal{D})d\theta_k \tag{8}$$

where $\mathcal{M} = \{M_0,\ldots,M_K\}$ is the set of models entertained, while M_k is a defined model including the number and locations of basis functions $M_k = \{k,\mu_k\}$, and θ_k to represent the coefficients w_k and A_k within this model structure. The posterior distribution $p(M_k,\theta_k|\mathcal{D})$ is given as a combination of likelihood and prior

$$p(M_k,\theta_k|\mathcal{D}) = \frac{p(\mathcal{D}|\theta_k,M_k)p(\theta_k,M_k)}{p(\mathcal{D})}, \tag{9}$$

where $p(\mathcal{D}|\theta_k,M_k)$ is the likelihood function and $p(\theta_k,M_k)$ is the prior density.

Assuming a normally distribution noise term as equation 7, we obtain the following log-likelihood of the RBF network, up to an additive constant,

$$\mathcal{L}[p(\mathcal{D}|\theta_k,M_k)] = -n\log\sigma - \frac{1}{2\sigma^2}\sum_{i=1}^{n}[\mathbf{y}_i - m(\mathbf{x}_i,\theta_k,M_k)]^2 \tag{10}$$

Then fixing the number of basis functions and location vectors is equivalent to setting a point prior on single values of k and μ, that is $\mathcal{M} = \{M_k\}$. After model selection step has been done, the flexibility of this model is controlled through the use of a prior on θ_k. Typically this will be form of shrinkage prior $p(\theta_k) \sim N(\theta_k|0,\sigma_m^2\mathbf{I})$ that penalizes large values. The prior is controlled through the precision parameter σ^2. This has the same effect as the regularization term η described in previous section.

The analysis here accounts for additional uncertainty present in the model selection of M_k. We place a proper prior over the whole of RBF model space $\mathcal{M}$ of a given basis function. Typically, a Poisson prior on k and a uniform prior on μ_k should be chosen, meanwhile setting the mean of this Poisson prior to a small value penalizes networks with a large number of basis functions, *e.g.* a predefined full grid control nodes for whole image. However, this fails to concern the amount of smoothing that is already being achieved by the setting of σ_m^2 in the prior on the output coefficients θ_k. Therefore, we place a function Ψ of k, μ_k and σ_m^2, which is readily computable and measures the amount of fitting that the model achieves. And the details of function ψ will be described in next subsection. Then a gamma prior is placed on Ψ, thus:

$$\begin{aligned} p(\theta_k,M_k) &= p(\theta_k,\mu_k,k) = p(\theta_k)p(\mu_k,k) \\ &= N(\theta_k|0,\sigma_m^2\mathbf{I})\Gamma(\Psi|\alpha,\beta), \end{aligned} \tag{11}$$

where α and β are Gamma prior parameters, the mean of the Gamma distribution is α/β, and variance is α/β^2. And, in d-dimensions, setting the mean equal to $d+1$ indicates a preference for a linear fit.

3.1. Function Ψ

Consider the classic RBF model as linear smoothers, and let ψ denote the design matrix given by the output of basis functions at the data points $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ and θ denote the output coefficients. The linear equations could be written as: $y = \psi\theta$. However, with the over-fitting problems, the regularization matrix R is incorporated in the linear equations to smooth the output. Therefore, the linear equations could represent as $\hat{y} = Ry$. Then the least-square approximation should be:

$$\hat{\theta} = (\psi^T\psi + \sigma_m^2\mathbf{I})^{-1}\psi^T y, \tag{12}$$

where σ_m^2 is the regularization parameter mentioned in previous section. Then, according to $\hat{y} = \psi\hat{\theta}$, $\hat{y}$ could be written as:

$$\hat{y} = \psi(\psi^T\psi + \sigma_m^2\mathbf{I})^{-1}\psi^T y, \tag{13}$$

and hence

$$R = \psi(\psi^T\psi + \sigma_m^2\mathbf{I})^{-1}\psi^T. \tag{14}$$

In accordance with generalized additive model (GAM), we can define function Ψ of RBF as:

$$\Psi = tr(R). \tag{15}$$

This is the sum of the eigenvalues of R, which gives a measure of amount of fitting that R the expected regularization matrix for our model.

Above all, our ultimate aim is to use the densities defined in equation 9 for the predictions given in equation 8 by integrating over the distribution to calculate the deformable image optical flow field. Unfortunately, due to these posterior densities are typically complex and of varying dimension, thus the Reversible Jump Markov Chain Monte Carlo (RJMCMC), presented by Green 1995 [6], is applied to approximate the output of our image deformable alignment model.

4. Reversible Jump MCMC

Bayesian inference involves integration, and Markov Chain Monte Carlo (MCMC) methods form an important tool for approximating the integrals we are interested in. First we should have a brief overview of MCMC method.

As described in previous section, the θ denoted a parameter vector having a parameter space Θ, index family of possible probability distributions $p(\mathcal{D}|\theta)$ describing our observations data pairing $\mathcal{D}$. Assuming $p(\theta)$ is a prior probability density describing θ, then Bayesian formula gives the posterior $\pi(\theta)$ in terms of likelihood and prior:

$$\pi(\theta) = p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{\int p(\mathcal{D}|\theta)p(\theta)d\theta}, \tag{16}$$

In Markov Chain Monte Carlo methods (MCMC), a Markov Chain will be constructed, which has parameter space Θ as its state space and $\pi(\theta)$ as its limiting stationary distribution. That means we have a way of sampling values from posterior distribution $\pi(\theta)$ and therefore make Monte Carlo inference about θ in form of sample averages and density estimates.

A Markov Chain is described by a transition kernel procedure that gives for each state θ^t the probability distribution for the chain to move to state $\tilde{\theta}^{t+1}$ by step size $d\theta$. For ease of exposition, it is assumed that for each distribution the corresponding density function exists and denote the transition density as $p(\theta^t, \tilde{\theta}^{t+1})$. MCMC methods produce chains that are aperiodic, irreducible and fulfill a reversibility condition, also called detailed balance equation:

$$\pi(\theta^t)p(\theta^t, \tilde{\theta}^{t+1}) = \pi(\tilde{\theta}^{t+1})p(\tilde{\theta}^{t+1}, \theta^t). \tag{17}$$

This equation means if $\pi(\theta^t)$ is the initial distribution of the starting state, then the intensity of going from state θ^t to $\tilde{\theta}^{t+1}$ is as same as that of going from $\tilde{\theta}^{t+1}$ to θ^t. Meanwhile, the reversibility could directly lead:

$$\int \pi(\theta^t)p(\theta^t, \tilde{\theta}^{t+1})d\theta = \pi(\tilde{\theta}^{t+1}), \tag{18}$$

which means that $\pi(\theta^t)$ is in fact the stationary distribution of the chain and we can used a sample from the chain as a random sample from distribution $\pi(\theta^t)$. Therefore, in Metropolis-Hastings algorithm, which is presented by Metropolis, Rosenbluth *et.al* in 1953 [8], we generate a Markov Chain with transition density:

$$p(\theta^t, \tilde{\theta}^{t+1}) = q(\theta^t, \tilde{\theta}^{t+1})\alpha(\theta^t, \tilde{\theta}^{t+1}) \tag{19}$$

then for some proposal density $q(\theta^t, \tilde{\theta}^{t+1})$, which is commonly chosen to be $N(0, \Sigma)$, and for acceptance probability $\alpha(\theta^t, \tilde{\theta}^{t+1})$. Therefore, the proposed sample is accepted with probability:

$$\alpha(\theta^t, \tilde{\theta}^{t+1}) = \mathbf{min}\left\{1, \frac{\pi(\tilde{\theta}^{t+1})q(\tilde{\theta}^{t+1}, \theta^t)}{\pi(\theta^t)q(\theta^t, \tilde{\theta}^{t+1})}\right\} \tag{20}$$

If the proposed state is accepted, then the next point in the chain, θ^{t+1} is set to $\tilde{\theta}^{t+1}$; otherwise it is set to the pervious point $\theta^{t+1} = \theta^t$. The new sample θ^{t+1} now forms the starting point for next proposed state, and the algorithm is iterated until a large number of samples have been drawn. Each updated state is just a function of previous state, and they are hence referred to as Markov Chains. However, the MCMC methods were mainly restricted to densities of fixed dimension. For our approach of deformable image alignment, the fixed dimension optimization method could not be applied to fundamentally solve the over-fitting and accurate alignment problem by only regularization the parameters (weights) in deformable image alignment RBF network. Therefore, the Reversible Jump MCMC is chosen to sample from $p(\theta^t)$, while θ^t is of unknown dimension, *i.e.* unknown model M_k.

4.1. Reversible Jump Radial Basis Functions

The detailed balance equation can also be formulated in general state space. For the Metropolis-Hastings algorithm to work it must only accept states from where there is a positive probability to do a reversible move back to the original state. For the reversible jump Metropolis-Hastings algorithm the state space is written as:

$$E = \{M_k, \theta_k | M_k \in \mathcal{M}, \theta_k \in \Theta_k\} \tag{21}$$

where $\mathcal{M}$ is enumerable model space and $\theta_k \in \Theta_k$ is the parameter space of model $M_k = \{k, \mu_k\}$. The dimension of θ_k varies with k. Reversible jump algorithm proceeds by augmenting the usual proposal step of a conventional Metropolis-Hastings algorithm sampler with a number of other possible move types surrounding a change in the dimension of the density. At each iteration, in addition to the possibility of moving within a particular parameter subspace, the sampler can propose to switch dimension, either up or down, by adding or removing a basis radial function from network. To improve the convergence speed in non-rigid deformable image alignment, we implement the birth and death proposition kernel, presented by L. Garcin *et.al.* [5]. The probability of attempting a birth or death step when the current state has k basis functions is given by b_k and d_k, respectively, which is set as $d_1 = 0$, $b_n = 0$, and $b_k = d_k = 0.1$ for all other values of k.

It is common when jumping between dimensions to generate some random vector dE that augments the current state E (defined in equation 21) to form E', *i.e.* generally we can write $E' = E + dE$. For a birth step, this vector dE is a datum vector drawn at random from those points that do not already have a basis function located on them. The jump move is accepted with probability:

$$\alpha(E, E') = \mathbf{min}\left\{1, \frac{p(E'|\mathcal{D})\tau_m(E')}{p(E|\mathcal{D})\tau_m(E)q(dE)}\left|\frac{\partial(E')}{\partial(E, dE)}\right|\right\} \tag{22}$$

where $\tau_m(E)$ is the probability of choosing a jump of type M when the current state is E and $q(dE)$ is the density function of dE. The final term, a Jacobian arises from the change of variables from (E, dE) to E'. For the radial basis mapping model, the accepting probability could be interpreted as (likelihood ratio)×(prior ratio)×(proposal ratio).

To illustrate this model, the birth step, from $E = \{k, \mu_k, \theta_k\}$ to $E' = \{k+1, \mu_{k+1}, \theta_{k+1}\}$ should be considered. In actually, the death steps are just an inversion of the the prior ratio for a birth. Then, the prior prior ratio for birth is:

$$\frac{\Gamma(\Psi_{M_{k+1}})\,[(k+1)!(n-k+1)!/n!]\,p(\theta_{k+1})}{\Gamma(\Psi_{M_k})\,[k!(n-k)!/n!]\,p(\theta_k)}, \tag{23}$$

where $\Gamma(\Psi)$ and $p(\theta)$ are taken from equation 11, and the term $k!(n-k)!/n!$ represents the probability of choosing the k basis locations from n data pairing points.

Meanwhile, the birth proposal ration could be given by:

$$\frac{d_{k+1}p(\theta_k|\mathcal{D})/(k+1)}{b_k p(\theta_{k+1}|\mathcal{D})/(n-k)}, \quad (24)$$

where the full conditional of $p(\theta|\mathcal{D})$ is used as proposed distribution for θ, which could be presented by parameters calculated in function Ψ:

$$p(\theta|\mathcal{D}) = N(\theta|(\psi^T\psi + \sigma_m^2\mathbf{I})^{-1}\psi^T y, \sigma^2\psi^T\psi) \quad (25)$$

where σ_m^2 is the prior precision for θ, σ^2 is the noise variance for and ψ is the design matrix of outputs from RBFs and polynomial terms, which is described in section 3.1.

Finally, the likelihood ratio is taken from equation 10. Thus, the reversible jump MCMC algorithm for deformable image alignment radial basis mapping model has been established as above. The detailed procedure of this iterative algorithm could be written in pseudocode in Table 1.

5. Generation of Training Data Pairing Set $\mathcal{D}$

Thus far, we have established the whole reversible jump Markov Chain Monte Carlo algorithm to automatically estimate more accurate parameters θ_k and M_k of the Bayesian Radial Basis Function Mapping model for deformable image alignment model. However, the entire learning model, built in previous sections, is based on the assumption that data pairing set $\mathcal{D}$ has been obtained. In this section, we will discuss how to generate the data pairing set from intensity of deformable image frames.

Generation of training data pairing set $\mathcal{D}$ is image feature extraction and tracking problem. There are some well-studied feature extraction and tracking algorithms. In our approach, the good feature tracking method, presented by J. Shi and C. Tomasi in [11], has been concerned to generate the training data pairing set $\mathcal{D}$. Then, parameters θ_k and M_k are learned from this training data pairing set $\mathcal{D}$ by reversible Markov Chain Monte Carlo approach. Thus, the accuracy of alignment is totally depends on the accuracy of features extraction and tracking. Additional and more accurate features been confirmed, more detailed and more exact deformable image alignment results we can get.

After parameters θ_k and M_k are learnt out, the deformable image alignment has become an missing data points filling procedure. Bayesian RBF mapping model Equation 7 with given parameters will be recalculated on each pixel of input image to get dense deformation field.

6. Experiments

This section reports experimental results on three real motion sequences with different types of deformable surface alignment problems, namely namely warping cloth sequence (Figure 2a and 2b), Claire speech sequence (Figure 2c and 2d) and heart beating sequence (Figure 2e and 2f). The goal of these experiments is to validate our approach and compared with the existing deformable registration algorithms, such as Demons Algorithm with fixed grid (*i.e.* Demon **I** appeared in [12]) and Direct Estimation Radial Basis Mapping with Dynamic Centers Insertion (D.C.I. as abbreviation) method, (presented Adrien Bartoli *et.al.*, [1]).

The alignment results and quantitative evaluation in

Input: Problem Eq.7, data pairings set $\mathcal{D}$
Output: Stationary distribution of Markov Chain: $p(\theta_k, M_k|\mathcal{D})$

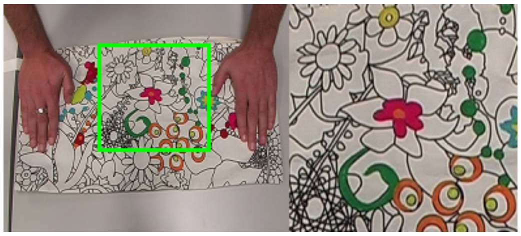
Initialization:

1.Define variance σ^2 from its prior probability: $\Gamma(\sigma^{-2}|10^{-3}, 10^{-3})$

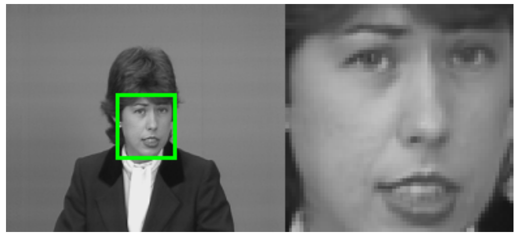
2.Draw the out put coefficients θ by used the full conditionals of θ probability $p(\theta|\mathcal{D})$ (*i.e.* equation 25)

Repeat:
(1). Draw a uniform random variable $dE \sim U(0,1)$.
(2). Propose the next state of the chain as follows:
  if $(dE < b_k)$
    then perform BIRTH step.
  else if $(dE < b_k + d_k)$
    then perform DEATH step.
  else
    then perform MOVE step.
  end
(3). Redraw the coefficients θ' as intialization step 2.
(4). Draw a uniform random variable $dE \sim U(0,1)$.
  if $\{dE < \alpha(E, E')\}$ (*i.e.* equation 22)
    then accept the proposed state.
  else
    then set the next state to be the current state.
  end
(5). Draw the noise variance σ^2 from its prior probability: $\Gamma(\sigma^{-2}|10^{-3} + n/2, 10^{-3} + \varepsilon^2/2)$, where n is number of data points and ε^2 is the sum of squared residuals for the current model.
(6). Until convergence is assumed
End

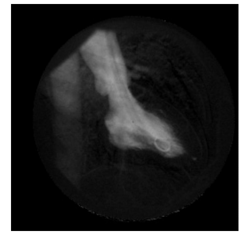
Table 1. Pseudo-code of Reversible Jump MCMC procedure for Bayesian RBF mapping of deformable image alignment algorithm: The MOVE, BIRTH and DEATH steps are simple. MOVE selects a basis function at random and resets its location vector to another datum drawn randomly from data set $\mathcal{D}$. BIRTH adds another basis function at a randomly selected point in the data set $\mathcal{D}$ that does not already contain one. DEATH selects just one basis at random and removes it.



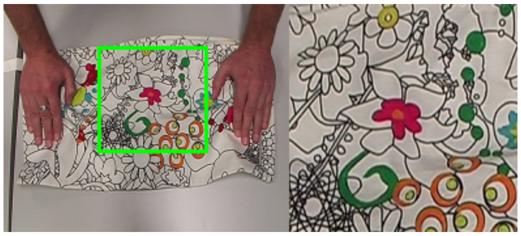
a. Unwarped Texture cloth (Frame:144)



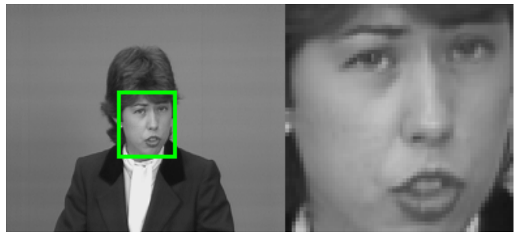
c. Claire Speech sequence (Frame:55)



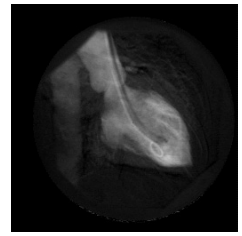
e. Heart (Frame:12)



b. Warped Texture cloth (Frame:159)



d. Claire Speech sequence (Frame:57)



f. Heart (Frame:23)

Figure 2. Three different kinds of deformable surface: (a-b) Two frames of warping texture cloth with folding motion; (c-d) Two frames in Claire Speech Sequence, the motion include lips motion and head pose change; (e-f) Two frames of heart beating sequence. (a-b provided by Mathieu Salzmann, EPFL; c-f downloaded from OSU/SAMPL Database-URL `http://sampl.ece.ohio-state.edu`)

Method	Demons I	D.C.I. method	Our approach
Cloth	23.37	15.36	12.17
Claire	8.12	6.48	6.03
Heart	11.18	9.97	9.21

Table 2. Results of alignment error (measured by mean square error over 256 gray-levels).

mean squared error are presented in Figure 3 and Table 2, respectively. Compared with the first row in Figure 3, we find, for the warping cloth sequence, the D.C.I. method and our approach give more satisfied registration results than Demons **I**. It is easy to see that our approach is the best of three for warping cloth sequence. Meanwhile, We are noted there is a massive over-fitting phenomenon appearing in warped cloth image generated by Demons **I**, due to demons algorithm focused on contours alignment by Maxwell demon force. Therefore, for complex cloth folding motion, it can hardly handel this situation.

In the Claire speech sequence, the second row of Figure 3, under-fitting and over-fitting phenomenons appear in lips area of Demon **I** and D.C.I. But for head pose recovery, excellent effects have been given. The reason for under-fitting in Demons **I** is caused by featureless homogeneous regions in human faces and local complex expression motion. Therefore, the deformation of local expression, especially lips moving, is hardly described by fixed grid in Demons **I**. Although, D.C.I applied dynamic centers insertion algorithm to solve this problem, without quantitative determination of centers' density, it is easy to cause under-fitting with sparse centers set, and over-fitting under dense centers situation. Whereas, with incorporating prior information and learning framework, our approach could easily balance global transformation and local complex motion by reasonable determining the density and location of centers set in images to overcome under-fitting and over-fitting problems.

Finally, for the heart beating sequence, all of three deformable image alignment approach get satisfied registered images. In some extent in medical image alignment, which contains homogeneous regions and without local complex deformation, D.C.I. and our approach could be interpreted as equivalent to Demon algorithm with contours nodes, i.e. Demon **II** [12], due to almost all of centers located near the contours of ventricular and atrial in D.C.I. and our approach. That is why all of these three algorithms get such similar alignment results.

7. Conclusions

We have proposed an learning variable dimension Radial Basis Function for the computation of deformable image alignment problems, through features extraction and tracking to generate the training data pairing set, learning variable dimension algorithm is able to make the 'best' model selection and optimize model coefficients for registration simultaneously, that equivalently be viewed as a way to overcome over-fitting and enhance convergence stabilization by model complexity reduction and parameters regularization. Amongst of the possible avenues for further research, experimenting with more efficient and elegant variable dimensions approximation algorithm would improve the computation speed, as well as computing super-resolution.

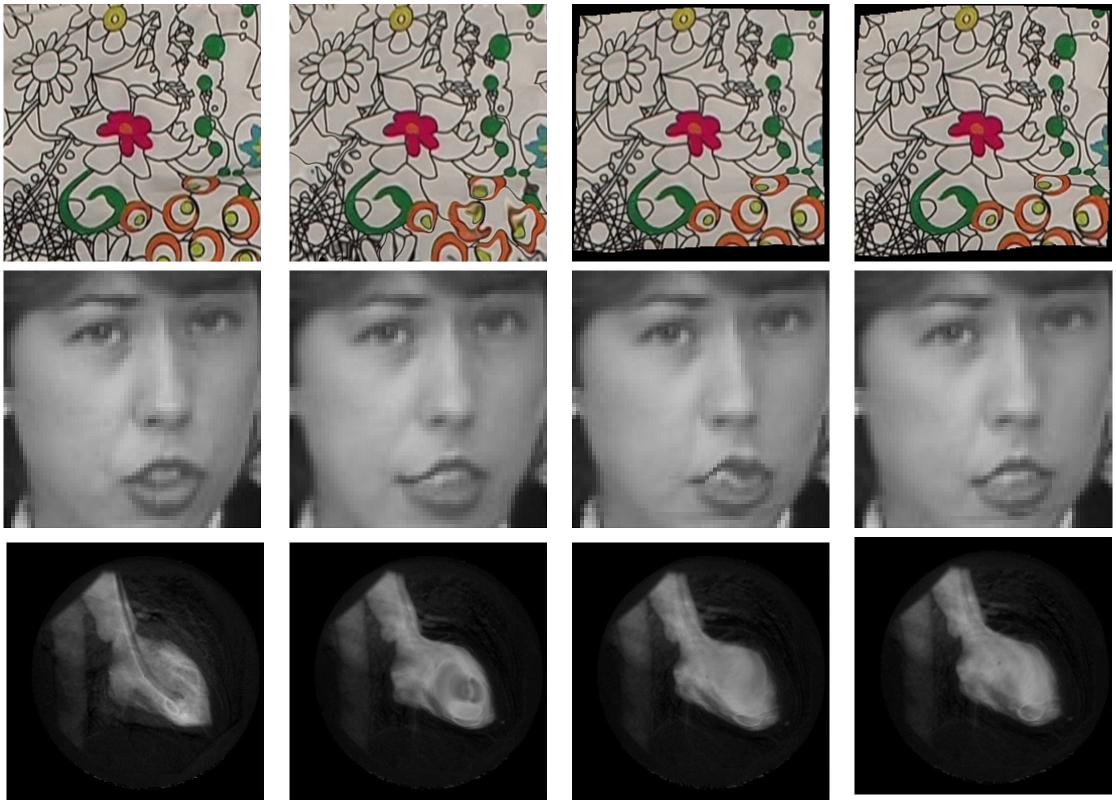


Figure 3. The templates and aligned images of all three approaches, three images in A column are templates of warping cloth, claire speech and heart beating, respectively. Images in B column are registered results of Demons **I**, as same, images in C and D columns are alignment results of D.C.I. method and Our approach, respectively.

References

[1] A. Bartoli and A. Zisserman. Direct estimation of non-rigid registrations. In *proceedings of the 15th British Machine Vision Conference,*, volume 2, pages 899–908, London, UK, September 2004.

[2] J. Bergen, P. Anandan, K. Hanna, and R. Hingorani. Hierarchical model-based motion estimation. In *proceedings of the 2nd European Conference on Computer Vision*, pages 237–252, Santa Margherita Ligure, Italy, 1992.

[3] F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(6):567–585, June 1989.

[4] D. Broomhead and D. Lowe. Multi-variable functional interpolation and adaptive networks. *Complex System*, 2:321–355, 1988.

[5] L. Garcin, X. Descombes, H. L. Men, and J. Z. and. Building detection by markov object processes. In *Proceedings of 2001 IEEE International Conference on Image Processing*, volume 2, pages 565–568, Thessaloniki, Greece, October 2001.

[6] P. Green. Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82:711732, 1995.

[7] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *proceedings of International Joint Conference on Artificial Intelligence*, pages 674–679, Vancouver, Canada, 1981.

[8] N. Metropolis, A. Rosenbluth, M. Rosenbluth, and A. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092, June 1953.

[9] N. D. Molton, A. J. Davison, and I. D. Reid. Parameterisation and probability in image alignment. In *proceedings Asian Conference on Computer Vision*, Jeju, Korea, Jan 2004.

[10] M. Powell. Radial basis functions for multivariate interpolation: A review. In J. C. M. . M. G. Cox, editor, *Algorithms of approximation*, page 143167. Clarendon Press, Oxford, 1987.

[11] J. Shi and C. Tomasi. Good feature to track. pages 593–600, Seattle, Washington, 1994.

[12] J. Thirion. Image mathching as a diffusion process: an analogy with maxwell's demons. *Medical Image Analysis*, 2(3):243–260, 1998.

[13] G. Wahba. *Spline models for observational data*. SIAM, Philadelphia,PA, September 1990.

[14] B. Zitov and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21:9771000, 2003.