www.ietdl.org

Published in IET Computer Vision
Received on 16th May 2012
Revised on 11th December 2012
Accepted on 7th January 2013
doi: 10.1049/iet-cvi.2012.0081

ISSN 1751-9632

Shape and appearance priors for level set-based left ventricle segmentation

Ronghua Yang[1], Majid Mirmehdi[1], Xianghua Xie[2], David Hall[3]

[1]*Department of Computer Science, University of Bristol, Bristol BS8 1TH, UK*
[2]*Department of Computer Science, University of Swansea, Swansea SA2 8PP, UK*
[3]*Department of Medical Physics and Bioengineering, University Hospitals Bristol NHS Foundation Trust, UK*
E-mail: majid@cs.bris.ac.uk

Abstract: The authors propose a novel spatiotemporal constraint based on shape and appearance and combine it with a level-set deformable model for left ventricle (LV) segmentation in four-dimensional gated cardiac SPECT, particularly in the presence of perfusion defects. The model incorporates appearance and shape information into a 'soft-to-hard' probabilistic constraint, and utilises spatiotemporal regularisation via a maximum a posteriori framework. This constraint force allows more flexibility than the rigid forces of shape constraint-only schemes, as well as other state of the art joint shape and appearance constraints. The combined model can hypothesise defective LV borders based on prior knowledge. The authors present comparative results to illustrate the improvement gain. A brief defect detection example is finally presented as an application of the proposed method.

1 Introduction

Single-photon emission computed-tomography (SPECT) is a commonly used approach for clinical evaluation of coronary artery disease. It visualises functional information of the left ventricle (LV) in three-dimension (3D), such as discontinuous perfusion scintigrams because of scarred or ischaemic myocardium. Gated cardiac SPECT also provides cardiac motion information along the time dimension.

The segmentation of LV epicardial and endocardial borders allows quantitative analysis of perfusion defects and cardiac function, and hence is of significant importance for diagnosis and clinical study. In SPECT, normal perfusion has high brightness in good contrast against a dark background, whereas perfusion defects have low or no brightness and hence low or no contrast against the background. Moreover, SPECT imaging can suffer from motion artefacts because of long imaging times [1], which is not a trivial task to compensate. As a result, accurate determination of the LV borders around the defect is difficult as there are few or no defined image features, such as gradients. In addition, the image data very often contains hot structures in the proximity of the LV, such as the liver, considered as perfusion artefacts that impose similar difficulties for correct LV segmentation. Furthermore, although resolution recovery techniques, such as [2], can be applied to increase SPECT spatial resolution, SPECT images have far lower resolution, as well as exhibit fewer anatomical landmarks (since the images suffer degradation because of various types of attenuation), in comparison with other modalities, such as Computed tomography (CT) and magnetic resonance imaging (MRI). Methods that are purely data-driven may only work well on SPECT images that do contain those image artefacts and ambiguities caused by perfusion at surrounding tissues, for example, paper [3] which uses a conventional edge based active contour method. Prior knowledge of the LV shape is hence usually required to estimate myocardium borders and defective borders in particular.

Beyond such analysis in static LV data, segmentation of LV in 4D allows analysis of cardiac motion. However, because the low-resolution nature of SPECT imaging produces a partial volume effect, the images at end-systole tend to look brighter than end-diastole – known as the artefact of myocardial brightening [4]. Hence, prior knowledge of cardiac motion is also required to estimate LV borders to guarantee correct motion estimation.

We propose a level-set framework that incorporates a novel shape and appearance-based spatiotemporal constraint into a general data-driven deformable model. The proposed constraint makes use of a priori knowledge in shape, appearance and their correlation, as well as spatiotemporal variations of the object under study, allowing more flexibility and reliability in the overall model. This helps cope with misleading image data such as defects, artefacts and temporal inconsistencies in gated cardiac SPECT.

In the next section, we review previous work on LV segmentation in cardiac SPECT data and other related issues. Section 3 covers charged active contour based on electrostatics (CACE), the basis for the proposed constrained CACE (CCACE) model. Section 4 gives details of CCACE, followed by experimental results of the model in Section 5. Finally, the paper is concluded in Section 6.

2 Previous work

A variety of deformable models have been proposed for LV segmentation in cardiac SPECT (e.g. [5–10]). Although bottom-up methods, for example [9] based on piecewise constant assumption, have been used, they generally suffer from image ambiguities, for instance, caused by under-perfusion. Top-down approaches in comparison are probably more robust, owning to the rigidity of the pre-defined models. For example, in [10], a predefined LV mesh is manually initialised and subsequently deformed to fit into local image gradients. Defective borders in SPECT can be hypothesised with these models as they sustain the shape of an LV during segmentation. However, these models have neglected the time dimension and deal with static LV only, missing the advantages that time consistent segmentation can offer [11]. Others have pursued a sequential segmentation approach, for example [12, 13], by taking the segmentation result of the previous timeframe as an initial guess for the current timeframe, but this generally leads to an underestimation of cardiac motion [14]. Paying more attention to the time dimension, [15] built a statistical model on the motion fields between the images at end-systole and end-diastole for better estimate of cardiac motion. Nonetheless, this approach, as well as other motion tracking algorithms [15, 16–18], track cardiac motion along sampled points in the image, which may not provide accurate segmentation of the LV shape.

There are also some elaborate spatiotemporal methods for better estimates of LV shape and motion on gated cardiac SPECT. Debreuve *et al.* [11] and Charnoz *et al.* [19] implemented a geometric active contour to determine LV borders taking into account all the timeframes of the sequence to cope with background noise and provide more temporally smooth results. Montagnat and Delingette [14] used a pre-determined 4D shape model registered onto the input image sequence under shape and temporal constraints. Similar to [20–22], Kohlberger *et al.* [23] built a statistical constraint through principal component analysis (PCA) on training shapes. The constraint is applied to a level-set model searching LV borders in a Mumford–Shah's formulation [24, 19]. We hereafter refer to this model as SCMS, that is, the Spatiotemporally Constrained Mumford–Shah's model. The model in [25] was extended from [23] and inferred the constraint by utilising a kernel density estimator as suggested in [26] instead of the multivariate Gaussian model.

The main limitations of the above models would become apparent if they dealt with pathological cases. For example, the authors of [11, 19] assumed that the LV has constant pixel values, Montagnat and Delingette [14] relied on its own rigidity to hypothesise on the defective borders, and the constraints in [23, 25] were applied to the models rigidly, that is, with equal strength everywhere. Also, Kohlberger *et al.* [23] only allows local variations that the statistical priors can recover. Such shortcomings can compromise the accuracy of the estimate, especially on LVs with local variations and medium-to-severe perfusion defects. Furthermore, the constraints of these models are shape-based, with no mechanism proposed to incorporate appearance information correlating to the shape in the constraint forces. Since the shape of a perfusion does not always correspond to the LV shape in the presence of defects, we argue that these shape-based constraints still lack flexibility to deal with abnormal perfusion and hence can compromise segmentation accuracy. Appearance information, in addition to shape, should be considered.

In [27], Rousson and Paragios proposed a constraint via finding maximum probability density function (pdf) of the shape and combined it with a level-set model for object detection in optical images with noisy or missing data. This Self-Constrained Geodesic Active Regions model is hereafter referred to as SCGAR. Their probabilistic approach took into account shape variability, producing a soft-to-hard force, which allowed a wider range of shape variations than the models in [23, 25]. We will use this methodology as part of our approach towards LV detection in cardiac SPECT as LV shapes contain large degrees of irregular inter-class variations [28].

For more reliable application in SPECT, we take inspiration from the soft-to-hard probabilistic framework of SCGAR and replace its shape-based pdf objective function with a posterior that considers the shape and appearance 'of the whole sequence', leading to a novel shape and appearance-based spatiotemporal constraint derived with maximum a posteriori (MAP) estimation. Instead of globally aligning shapes when evaluating the pdf in SCGAR, we apply a global-to-local alignment. Meanwhile, as in [23, 25], PCA is employed in our model to reduce data dimension and statistically depict shape variations. However, unlike [23, 25], instead of applying PCA to the whole level-set domain, we apply it to the deformation fields of the level sets, characterised by transformation parameters, to further reduce data space and allow better control over temporal shape variations. Compared with current works on joint shape and appearance constraints such as [29–33], our soft-to-hard constraint allows more flexibility than their rigid forces. In addition, we argue that the probabilistic framework makes better use of the correlation between shape and appearance than those that linearly combine the two, for example [29–33].

In all, we present a model combining our recently proposed general-purpose deformable model for boundary detection, namely CACE [34], with a shape and appearance-based constraint that is capable of incorporating prior knowledge of correlated LV shape and appearance, as well as spatiotemporal variations. This is designed to cope with the irregularities in variations of LV shapes and motions and the extremely fuzzy gradients because of (temporal) perfusion defects and artefacts in cardiac SPECT. We refer to this combined model as CCACE.

3 Review of the CACE model

We now briefly review the CACE model [34] which is the foundation stone for the proposed CCACE method. CACE is a general-purpose active contour model that incorporates particle-based electrostatic interactions into the geometric active contour framework. It can be used for object boundary detection in a variety of images and applications. Embedded in level sets, CACE propagates under the influence of a bi-directional force field that simulates the electrostatic interaction between an image-derived point charge field and a charged contour. The force field leads the contour towards object boundaries and dynamically adapts as the contour evolves. It is a joint electrostatic force consisting of a boundary attraction force $\boldsymbol{F}_{\mathrm{A}}$ and a competition force $\boldsymbol{F}_{\mathrm{C}}$

$$\boldsymbol{J} = \lambda_{\mathrm{A}}\boldsymbol{F}_{\mathrm{A}} + (1 - \lambda_{\mathrm{A}})\boldsymbol{F}_{\mathrm{C}} \tag{1}$$

where

$$\boldsymbol{F}_{\mathrm{A}} = p_j \sum_{i=1,\boldsymbol{r}_i \neq \boldsymbol{s}_j}^{N_{\mathrm{c}}} \frac{q_i}{4\pi\varepsilon_0} \frac{\boldsymbol{s}_j - \boldsymbol{r}_i}{|\boldsymbol{s}_j - \boldsymbol{r}_i|^3}, \quad \boldsymbol{s}_j \in \boldsymbol{X} \tag{2}$$

$$\boldsymbol{F}_{\mathrm{C}} = g(\boldsymbol{s}_j)p_j \sum_{k=1,k \neq j}^{M_{\mathrm{c}}} \frac{f(\boldsymbol{s}_k)p_k}{4\pi\varepsilon_0} \frac{\boldsymbol{s}_j - \boldsymbol{s}_k}{|\boldsymbol{s}_j - \boldsymbol{s}_k|^3}, \quad \boldsymbol{s}_j \in \boldsymbol{X} \tag{3}$$

λ_{A} is a constant balancing the contributions from the two forces. $f(.)$ is the edge map of the input image, and $g(.)$ is an edge stopping function (commonly defined as $g = 1/(1 + f)$). q_i is the negative fixed charge assigned to location $\boldsymbol{r}_i$, with magnitude proportional to its edge strength, that is, $q_i = -f(\boldsymbol{r}_i)$. p_j (or p_k) is the positive charge at location $\boldsymbol{s}_j$ (or $\boldsymbol{s}_k$) on the active contour, and is set to unity for all $\boldsymbol{s}_j$, although p_j can be treated as a variable that can be tuned for specific applications. N_{c} and M_{c} are the number of negative and positive charges, respectively. ε_0 is the permittivity of free space and $\mathbf{X}$ is the set of all possible locations in image domain.

The attraction force $\boldsymbol{F}_{\mathrm{A}}$ acts as a vector field that leads each contour point towards the boundaries from both sides. The competition force $\boldsymbol{F}_{\mathrm{C}}$ upon a contour is proportional to the inverted strength of the edge (i.e. $g(\boldsymbol{s}_j)$) covered by this contour. Contours that are on the object boundaries endow most to the competition force with contributions proportional to edge strength. In other words, the competition force exerts most influence once any part of the snake reaches a boundary and repels 'free' contours nearby from reaching the already occupied boundary, forcing them to search for unoccupied boundaries. This repelling force is also designed in a way such that only contours in homogeneous regions are most affected. It means contours that reach object boundaries will exert repellent forces upon other contours, whereas they themselves will be least affected by others and remain on the boundaries. At the same time, contours in homogeneous regions will continue to deform according to both attraction and competition forces.

As the joint force field dynamically adapts by updating itself when a contour reaches a boundary, it brings flexibility in initialisation and better curve propagation towards object boundaries. The level-set ϕ evolution of CACE is

$$\frac{\partial \phi}{\partial t} = \lambda_\kappa g \kappa |\nabla \phi| - (1 - \lambda_\kappa)\boldsymbol{J} \cdot \nabla \phi \tag{4}$$

where λ_κ is a real positive constant and κ denotes the curvature of the contour (which regualrises it). For more details of CACE the reader is referred to [34]. As mentioned earlier, we build CCACE on the CACE framework and effectively CCACE is CACE with application of specific constraints for 4D LV segmentation.

4 Proposed CCACE model

4.1 Overview

As shown in Fig. 1, the proposed model comprises of two stages: training and segmentation. Through the training stage, sets of pixel-wise Gaussian and spatiotemporal priors are obtained. The Gaussian priors include a prior image, a prior shape, an image variation term, a shape variation term and the correlation between image and shape. The spatiotemporal priors are the products of applying PCA to global-to-local transformation parameters of shape variations, and include mean of the parameters, eigenmodes (i.e. modes of variations), weights for the eigenmodes, as well as the covariance of the weights.

During segmentation, an initial surface embedded in level sets is placed in the input image. A constraint force is derived by finding the maximum of the multivariate pdf of the input image and shape (the evolving level sets) based on the Gaussian priors from the training stage. As the level sets evolve, the pdf continuously updates by aligning the prior image and shape with the input image and shape via global-to-local transformations to enable meaningful evaluation of the pdf. The alignment is regularised by prior distribution over a set of 'spatiotemporal parameters' of the transformations. The force derived from this regularised multivariate pdf, effectively a posterior, is a shape and appearance-based spatiotemporal constraint obtained via MAP estimation. The evolving level sets update according to the combination of CACE forces and this constraint to reach (and hypothesise missing or defective) LV borders. Next, we discuss the training and segmentation stages in detail.

4.2 Training

A training set $\mathcal{M} = \{(\grave{\boldsymbol{I}}_i, \grave{\boldsymbol{\phi}}_i): i \in [1, N]\}$ of N-gated SPECT samples is constructed from manually labelled 4D sequences. Each sample $(\grave{\boldsymbol{I}}_i, \grave{\boldsymbol{\phi}}_i)$ comprises of a pair – an image sequence $\grave{\boldsymbol{I}}_i$ and its corresponding LV shape sequence $\grave{\boldsymbol{\phi}}_i$, where $\grave{\boldsymbol{I}}_i$ is the ith training image sequence, and $\grave{\boldsymbol{\phi}}_i$ is the ith training shape sequence embedded in level sets constructed by manual labelling on $\grave{\boldsymbol{I}}_i$. Each training sequence consists of K timeframes, that is, $\grave{\boldsymbol{I}}_i = \{\grave{I}_{i,k}: k \in [1, K]\}$ and $\grave{\boldsymbol{\phi}}_i = \{\grave{\boldsymbol{\phi}}_{i,k}: k \in [1, K]\}$. Any known defect-free sequence can be selected as the reference shape sequence $\mathring{\boldsymbol{\phi}} = \mathring{\boldsymbol{\phi}}_k: k \in [1, K]$. To reduce non-linearity to fit in a Gaussian prior, each training shape $\grave{\boldsymbol{\phi}}_{i,k}$ is globally and locally aligned towards the corresponding timeframe of the reference shape sequence, that is, $\mathring{\boldsymbol{\phi}}_k$, via a global-to-local registration method, whose details are given in the Appendix. The transformations recovered from these shape registrations are then applied to the training images accordingly to sustain the correspondence between training images and shapes. From the aligned training set, we derive two sets of priors: pixel-wise Gaussian priors and spatiotemporal priors.

Gaussian priors – We compute pixel-wise means of the aligned images $\bar{\boldsymbol{I}} = \{\bar{I}_k: k \in [1, K]\}$, image standard deviations $\bar{\boldsymbol{\sigma}}_I = \{\bar{\sigma}_{I_k}: k \in [1, K]\}$, means of the aligned shapes $\bar{\boldsymbol{\phi}} = \{\bar{\phi}_k: k \in [1, K]\}$, shape standard deviations $\bar{\boldsymbol{\sigma}}_\phi = \{\bar{\sigma}_{\phi_k}: k \in [1, K]\}$ and the correlation coefficients between image and shape $\bar{\boldsymbol{\rho}} = \{\bar{\rho}_k: k \in [1, K]\}$. $\bar{\boldsymbol{I}}$ and $\bar{\boldsymbol{\phi}}$ are also referred to as the prior image/appearance and prior shape sequence, respectively. As the training shapes are locally aligned, there is risk of bias in the priors towards the reference $\mathring{\boldsymbol{\phi}}$, whose impact on model performance is however insignificant, as will be mentioned later at the segmentation stage. In practice, we choose a sample with a typical LV shape that has the least local variations from the mean shape of the training set, in order to further reduce the bias. Along with their standard deviations, the mean

doi: 10.1049/iet-cvi.2012.0081

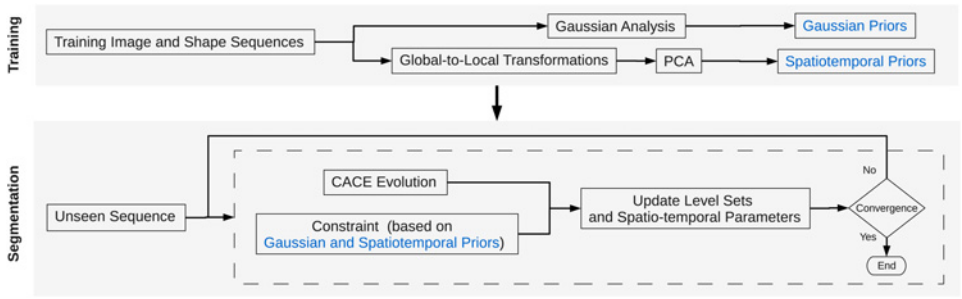

Fig. 1 *Overview of CCACE*

image and shape sequences of our training set, for $K = 8$ and $N = 15$, are shown in Fig. 2.

Spatiotemporal priors – We seek parameterised temporal correlation among the spatial variations by applying two levels of PCA to a set of transformation parameters that describe how the training shape sequences $\grave{\boldsymbol{\phi}}_i$ vary from the prior shape sequence $\bar{\boldsymbol{\phi}}$. These transformation parameters, denoted as $\{\grave{\boldsymbol{\Theta}}_{i,k}: (i, k) \in [1, N] \times [1, K]\}$, are recovered by mapping the prior shape sequence $\bar{\boldsymbol{\phi}}$ into the training shape sequences $\grave{\boldsymbol{\phi}}_i$ for $i \in [1, N]$, again via the global-to-local registration model (25). Note that in this mapping $\bar{\boldsymbol{\phi}}$ is the source shape and $\grave{\boldsymbol{\phi}}_i$ are the target shapes, whereas in the alignment of the training set above, $\grave{\boldsymbol{\phi}}_i$ is the source and $\mathring{\boldsymbol{\phi}}$ the target. The reason for mapping here is that the distribution of $\grave{\boldsymbol{\Theta}}_{i,k}$ can be used to regularise segmentation later during which the prior appearance and shape $(\bar{\boldsymbol{I}}, \bar{\boldsymbol{\phi}})$ are aligned towards the input appearance and shape when evaluating the posterior.

The first-level PCA is applied to the recovered transformation parameters $\{\grave{\boldsymbol{\Theta}}_{i,k}: i \in [1, N]\}$ at each timeframe k individually, each giving a mean $\bar{\boldsymbol{\Theta}}_k$, eigenmodes $\boldsymbol{P}_k$ comprising the m_k most significant modes of variations, and weights for the eigenmodes $\{\grave{\boldsymbol{b}}_{i,k}: i \in [1, N]\}$. The transformation parameters $\grave{\boldsymbol{\Theta}}_{i,k}$ can then be approximated as

$$\grave{\boldsymbol{\Theta}}_{i,k} = \bar{\boldsymbol{\Theta}}_k + \boldsymbol{P}_k \grave{\boldsymbol{b}}_{i,k} \tag{5}$$

This PCA captures the variations among shape transformations. We also consider that there is potential correlation among the global and local transformations which are both included in the PCA.

A further PCA is then applied to the weights of all timeframes, that is, $\{\grave{\boldsymbol{b}}_{i,k}: (i, k) \in [1, N] \times [1, K]\}$, and produces second-level eigenmodes $\boldsymbol{Q} = [\boldsymbol{Q}'_1 \ldots \boldsymbol{Q}'_K]'$ representing the n most significant modes of variations, and weights $\{\grave{\boldsymbol{c}}_i: i \in [1, N]\}$, giving the following representation

$$\grave{\boldsymbol{b}}_{i,k} = \boldsymbol{Q}_k \grave{\boldsymbol{c}}_i \tag{6}$$

The transformation parameters $\hat{\boldsymbol{\Theta}}_{i,k}$ can be therefore represented in $\hat{\boldsymbol{c}}_k$ as

$$\hat{\boldsymbol{\Theta}}_k = \bar{\boldsymbol{\Theta}}_k + \boldsymbol{P}_k \boldsymbol{Q}_k \hat{\boldsymbol{c}}_i \tag{7}$$

Different to the first-level PCA, in this second-level one the entire sequence is taken as one dataset as an observation, hence the weights $\hat{\boldsymbol{c}}_i$ do not relate to time dimension as $\hat{\boldsymbol{b}}_{i,k}$ does, that is, $\hat{\boldsymbol{c}}_i$ is not indexed by k.

The weights from the first PCA statistically depict shape variations of every timeframe, and the second PCA on them produces parameters that control their temporal variations. This is similar to Cootes *et al.* [35] where an appearance model is built by applying second-level PCA to the results of a first-level PCA on shape and image greylevel data, respectively. These two levels of PCA give an approximation for a new set of transformation parameters

$$\boldsymbol{\Theta}_k = \bar{\boldsymbol{\Theta}}_k + \boldsymbol{P}_k \boldsymbol{Q}_k \boldsymbol{c} \tag{8}$$

where $\boldsymbol{c}$ is a varying parameter to be determined with many fewer dimensions than the transformation parameters $\boldsymbol{\Theta}_k$. As it indirectly controls both spatial and temporal variations of the sequence by affecting its transformations, we refer to $\boldsymbol{c}$ as the 'spatiotemporal parameters'. Assuming normal distribution, the prior distribution of $\boldsymbol{c}$ is

$$P_c(\boldsymbol{c}) = \left(\sqrt{(2\pi)^n |\boldsymbol{U}|}\right)^{-1} \exp\left(-\frac{1}{2}\boldsymbol{c}' \boldsymbol{U}^{-1} \boldsymbol{c}\right) \tag{9}$$

where $\boldsymbol{U}$ is the covariance matrix whose diagonal comprises the eigenvalues obtained from the two-level PCA.

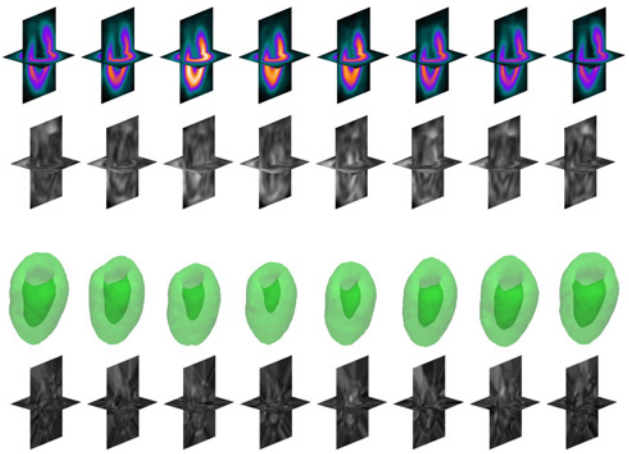

Fig. 2 *Top two rows: mean appearance sequence $\bar{\boldsymbol{I}}$ and its corresponding standard deviations; bottom two rows: mean shape sequence $\bar{\boldsymbol{\phi}}$ and its corresponding standard deviations*

From left: timeframes 1–8

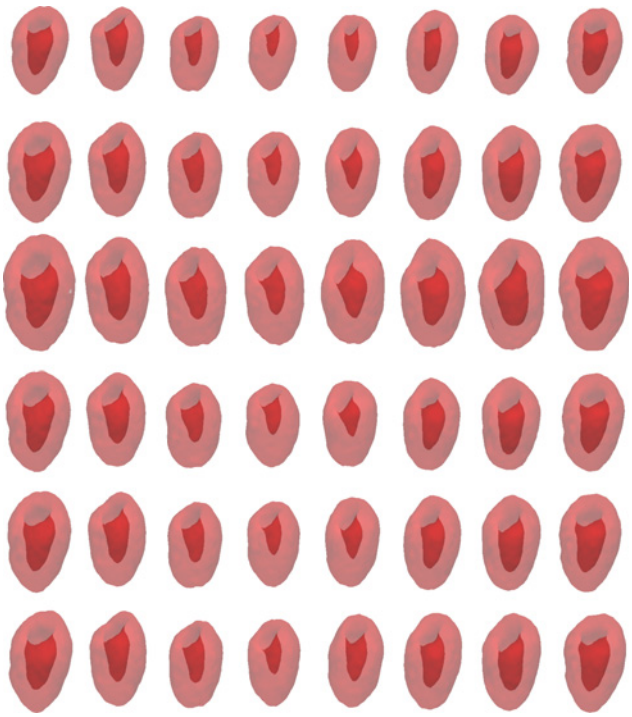

Fig. 3 *LV shape sequences recovered by varying the first two parameters of* $\boldsymbol{c}$

From top, the first parameter set to -15σ, 0, $+1.5\sigma$ (rows 1–3), and the second parameter set to -15σ, 0, $+1.5\sigma$ (rows 4–6)

To show the effect of varying the spatiotemporal parameters, we reconstruct transformations by varying each of the first two spatiotemporal parameters in turn, keeping the others zero, and apply the transformations to the prior shape, as shown in Fig. 3. It can be seen that the first parameter of $\boldsymbol{c}$ has obvious influence over global motions such as scaling and rotation of the LV, as well as local features such as the height of the septum. The second parameter, on the other hand, affects the LV shapes more subtly, for example, in changing the shape of the apical region and the ratio of the apical region over the whole LV size.

In summary, from the training stage, we obtain sets of Gaussian and spatiotemporal priors. The priors that are later used in the segmentation stage are listed in Table 1.

4.3 Shape and appearance-based spatiotemporal constraint

We now outline the derivation of our proposed constraint force based on the priors obtained during training and then give the combined formulation of CCACE in the next section. We use a variational framework to derive the constraint force that helps evolve the level sets to optimise a posterior, defined as the product of a shape and appearance-based multivariate pdf and prior distribution over the spatiotemporal parameters. The posterior measures the probability of an input shape and appearance to occur in a multivariate normal distribution characterised by the Gaussian priors obtained from the training stage, under spatiotemporal regularisation through prior distribution.

Table 1 Gaussian and spatiotemporal priors

Gaussian priors	
$\bar{\boldsymbol{\phi}} = \{\bar{\phi}_k : k \in [1, K]\}$	prior shape
$\bar{\mathbf{I}} = \{\bar{I}_k : k \in [1, K]\}$	prior image
$\bar{\boldsymbol{\sigma}}_\phi = \{\bar{\sigma}_{\phi_k} : k \in [1, K]\}$	shape standard deviation
$\bar{\boldsymbol{\sigma}}_I = \{\bar{\sigma}_{I_k} : k \in [1, K]\}$	image standard deviation
$\bar{\boldsymbol{\rho}} = \{\bar{\rho}_k : k \in [1, K]\}$	correlation between image and shape
spatiotemporal priors	
$\{\boldsymbol{\Theta}_k : k \in [1, K]\}$	mean of transformation parameters
$\{\boldsymbol{P}_k : k \in [1, K]\}$	first-level eigenmodes
$\{\boldsymbol{Q} = \boldsymbol{Q}_k : k \in [1, K]\}$	second-level eigenmodes
$\boldsymbol{U}$	covariance of spatiotemporal parameters

Given the input image sequence $\boldsymbol{I} = I_k : k \in [1, K]$ and the evolving level sets $\boldsymbol{\phi} = \boldsymbol{\phi}_k : k \in [1, K]$ placed on the image, the posterior is defined as

$$P_{\mathrm{A}}(\boldsymbol{\phi}, \boldsymbol{c}) = \prod_{k=1}^{K} \left((2\pi) \sqrt{|\Sigma_k|} \right)^{-1} \exp\left(-\frac{1}{2} \boldsymbol{a}_k' \sum_k^{-1} \boldsymbol{a}_k \right) P_{\mathrm{c}}(\boldsymbol{c}) \tag{10}$$

where

$$\boldsymbol{a}_k = \begin{bmatrix} a_{\phi_k} \\ a_{I_k} \end{bmatrix} = \begin{bmatrix} (1 - \lambda_I)\left(s_k \phi_k - \bar{\phi}_k(T)\right) \\ \lambda_I \left(I_k - \bar{I}_k(T)\right) \end{bmatrix} \tag{11}$$

$$\sum_k = \frac{1}{1 - \bar{\rho}_k^2(T)} \begin{bmatrix} \bar{\sigma}_{\phi_k}^2(T) & \bar{\rho}_k(T)\bar{\sigma}_{\phi_k}(T)\bar{\sigma}_{I_k}(T) \\ \bar{\rho}_k(T)\bar{\sigma}_{\phi_k}(T)\bar{\sigma}_{I_k}(T) & \bar{\sigma}_{I_k}^2(T) \end{bmatrix} \tag{12}$$

and T is the simplified representation of a global-to-local transformation model $T(\boldsymbol{\Theta}_k; \boldsymbol{x})$, details of which are given in the Appendix. The joint priors $(\bar{\boldsymbol{I}}, \bar{\boldsymbol{\phi}})$ are aligned with the input appearance and shape $(\boldsymbol{I}, \boldsymbol{\phi})$ through $T(\boldsymbol{\Theta}_k; \boldsymbol{x})$, for $k \in [1, K]$. The transformations act on the priors so that the derived constraint force applies straightforwardly to the level sets and can be directly combined with CACE. Note that the impact of potential bias in the priors towards the reference $\mathring{\boldsymbol{\phi}}$ is largely alleviated via the alignment here. Positive constant λ_{I} balances the contributions from image and shape. Σ_k is the covariance between image and shape. s_k, a component of $\boldsymbol{\Theta}_k$, is the scaling factor at kth timeframe. The reason for applying s_k to $\boldsymbol{\phi}_k$ is that level-set representation is not invariant to scaling – see [27] for details. As both $\boldsymbol{\phi}$ and $\boldsymbol{c}$ are unknown, a coupled optimisation task can optimise the posterior

$$\langle \tilde{\boldsymbol{\phi}}, \tilde{\boldsymbol{c}} \rangle = \arg\min_{\boldsymbol{\phi}, \boldsymbol{c}} E_{\mathrm{A}}(\boldsymbol{\phi}, \boldsymbol{c}) \tag{13}$$

where $E_{\mathrm{A}}(\boldsymbol{\phi}, \mathbf{c})$ is the posterior-based objective function defined as

$$\begin{aligned} E_{\mathrm{A}}(\boldsymbol{\phi}, \mathbf{c}) &= -\int_X \log\left(P_{\mathrm{A}}(\boldsymbol{\phi}, \boldsymbol{c})\right) \mathrm{d}\boldsymbol{x} \\ &= \sum_{k=1}^{K} \int_X \left(\log\left((2\pi)\sqrt{|\Sigma_k|} \right) + \frac{1}{2} \left(\boldsymbol{a}_k' \sum_k^{-1} \boldsymbol{a}_k \right) \right) \mathrm{d}\boldsymbol{x} \\ &\quad + \log\left(\sqrt{(2\pi)^n |\boldsymbol{U}|} \right) + \frac{1}{2} \boldsymbol{c}' \boldsymbol{U}^{-1} \boldsymbol{c} \end{aligned} \tag{14}$$

$\boldsymbol{X}$ represents all the possible locations in the shape domain [Note that the level-set shape domain coincides with the image domain.]. $E_{\mathrm{A}}(\boldsymbol{\phi}, \boldsymbol{c})$ considers both shape and appearance information and is evaluated with variations

doi: 10.1049/iet-cvi.2012.0081

between the priors $(\bar{\boldsymbol{I}}, \bar{\boldsymbol{\phi}})$ and the input terms $(\boldsymbol{I}, \boldsymbol{\phi})$ taken into account. In particular, only spatiotemporally plausible variations are accounted for owing to regularisation that biases the transformation parameters towards their most probable values. This keeps the model within plausible shapes and motions. The level sets evolve towards the maximum of the posterior to give the best probabilistic estimate over the shape in the input image.

4.4 Level-set evolution of CCACE

Combining the constraint component with CACE, the motion equations for CCACE are given in (15) and (16), through which $\boldsymbol{\phi}$ and $\boldsymbol{c}$ update iteratively and simultaneously until equilibrium

$$\frac{\partial \phi_k}{\partial t} = \lambda_c \Bigg(\underbrace{\lambda_\kappa g \kappa |\nabla \phi_k| - (1-\lambda_\kappa) J \cdot \nabla \phi_k}_{\text{CACE}} \Bigg) + (1-\lambda_c) \frac{s_k}{1-\bar{\rho}_k^2(T)} \Bigg(\underbrace{-(1-\lambda_I) \frac{s_k \phi_k - \bar{\phi}_k(T)}{\bar{\sigma}_{\phi_k}^2(T)}}_{\text{Shape}} - \underbrace{\lambda_I \bar{\rho}_k(T) \frac{I_k - \bar{I}_k(T)}{\bar{\sigma}_{\phi_k}(T) \bar{\sigma}_{I_k}(T)}}_{\text{Appearance}} \Bigg) \tag{15}$$

$$\frac{\partial \boldsymbol{c}}{\partial t} = -\left(\frac{\partial E_A(\boldsymbol{\phi}, \boldsymbol{c})}{\partial \Theta} \right)' \left[\boldsymbol{Q}'_1 \boldsymbol{P}'_1 \quad \dots \quad \boldsymbol{Q}'_K \boldsymbol{P}'_K \right] - \boldsymbol{U}^{-1} \boldsymbol{c} \tag{16}$$

where $\boldsymbol{\Theta} = \boldsymbol{\Theta}_k$: $k \in [1, K]$. In (15), the first term represents the data-driven evolution of CACE, whereas the second is the constraint force, derived based on (13), with λ_κ balancing the contributions from the two. The constraint force consists of a shape and an appearance term. $\boldsymbol{c}$ also updates according to (16) at each iteration of $\boldsymbol{\phi}'$s evolution to keep the prior image and shape aligned with the input image and shape.

To compute the partial derivatives $((\partial E_A(\boldsymbol{\phi}, \boldsymbol{c}))/\partial \boldsymbol{\Theta})$, we denote the transformation parameter as $\boldsymbol{\Theta}_k = \theta_{k,\,j}$: $j \in [1, N_\theta]$, where N_θ is the length of $\boldsymbol{\Theta}_k$, so that $\theta_{k,j}$ refers to the jth component of $\boldsymbol{\Theta}_k$. Thus, $((\partial E_A(\boldsymbol{\phi}, \boldsymbol{c}))/\partial \boldsymbol{\Theta})$ can be formulated as in (17).

$$\begin{aligned}
\frac{\partial E_A(\boldsymbol{\phi}, \boldsymbol{c})}{\partial \theta_{k,j}} &= \int_{\mathbf{X}} \left[\frac{1}{|\Sigma_k|} \frac{\partial |\Sigma_k|}{\partial \theta_{k,j}} + \frac{1}{2} \boldsymbol{a}_{k'} \frac{\partial \Sigma_k^{-1}}{\partial \theta_{k,j}} \boldsymbol{a}_k + \boldsymbol{a}_{k'} \Sigma_k^{-1} \frac{\partial \boldsymbol{a}}{\partial \theta_{k,j}} \right] \mathrm{d}\boldsymbol{x} \\
&= \int_{\boldsymbol{X}} \left[\frac{\partial a_{\phi_k}}{\partial \theta_{k,j}} \left(\frac{a_{\phi_k} \bar{\sigma}_{I_k}(T) - \bar{\rho}_k(T) a_{I_k} \bar{\sigma}_{\phi_k}(T)}{(1-\bar{\rho}_k^2(T)) \bar{\sigma}_{\phi_k}^2(T) \bar{\sigma}_{I_k}(T)} \right) \right. \\
&\quad + \frac{\partial a_{I_k}}{\partial \theta_{k,j}} \left(\frac{a_{I_k} \bar{\sigma}_{\phi_k}(T) - \bar{\rho}_k(T) a_{\phi_k} \bar{\sigma}_{I_k}(T)}{(1-\bar{\rho}_k^2(T)) \bar{\sigma}_{I_k}^2(T) \bar{\sigma}_{\phi_k}(T)} \right) \\
&\quad + \frac{\partial \bar{\sigma}_{\phi_k}(T)}{\partial \theta_{k,j}} \frac{1}{\bar{\sigma}_{I_k}^2(T) \bar{\sigma}_{\phi_k}^3(T)} (\bar{\sigma}_{\phi_k}^2(T) \bar{\sigma}_{I_k}^2(T) \\
&\quad + a_{\phi_k} a_{I_k} \bar{\rho}_k(T) \bar{\sigma}_{\phi_k}(T) \bar{\sigma}_{I_k}(T) - \bar{\sigma}_{I_k}^2(T) a_{\phi_k}^2) \\
&\quad + \frac{\partial \bar{\sigma}_{I_k}(T)}{\partial \theta_{k,j}} \frac{1}{\bar{\sigma}_{\phi_k}^2(T) \bar{\sigma}_{I_k}^3(T)} (\bar{\sigma}_{I_k}^2(T) \bar{\sigma}_{\phi_k}^2(T) \\
&\quad + a_{I_k} a_{\phi_k} \bar{\rho}_k(T) \bar{\sigma}_{I_k}(T) \bar{\sigma}_{\phi_k}(T) - \bar{\sigma}_{\phi_k}^2(T) a_{I_k}^2) \\
&\quad + \frac{\partial \bar{\rho}_k(T)}{\partial \theta_{k,j}} \left(\frac{(\boldsymbol{a}'_k \Sigma_k^{-1} \boldsymbol{a}_k) \bar{\rho}_k(T) \bar{\sigma}_{\phi_k}(T) \bar{\sigma}_{I_k}(T)}{\bar{\sigma}_{\phi_k}^2(T) \bar{\sigma}_{I_k}^2(T) (1-\bar{\rho}_k^2(T))} \right. \\
&\quad \left. \left. - \frac{a_{\phi_k} a_{I_k} \bar{\sigma}_{\phi_k}(T) \bar{\sigma}_{I_k}(T) (1-\bar{\rho}_k^2(T)) - \bar{\rho}_k(T) \bar{\sigma}_{\phi_k}^2(T) \bar{\sigma}_{I_k}^2(T)}{\bar{\sigma}_{\phi_k}^2(T) \bar{\sigma}_{I_k}^2(T) (1-\bar{\rho}_k^2(T))} \right) \right] \mathrm{d}\boldsymbol{x}
\end{aligned} \tag{17}$$

where

$$\frac{\partial a_{\phi_k}}{\partial \theta_{k,j}} = (1-\lambda_I) \left(\phi_k - \nabla \bar{\phi}_k(T) \cdot \frac{\partial T}{\partial \theta_{k,j}} \right), \quad \theta_{k,j} = s_k \tag{18}$$

$$\frac{\partial a_{\phi_k}}{\partial \theta_{k,j}} = -(1-\lambda_I) \nabla \bar{\phi}_k(T) \cdot \frac{\partial T}{\partial \theta_{k,j}}, \quad \theta_{k,j} \in \boldsymbol{\Theta}_k, \; \theta_{k,j} \neq s_k \tag{19}$$

$$\frac{\partial a_{I_k}}{\partial \theta_{k,j}} = -\lambda_I \nabla \bar{I}_k(T) \cdot \frac{\partial T}{\partial \theta_{k,j}}, \quad \theta_{k,j} \in \boldsymbol{\Theta}_k \tag{20}$$

$$\frac{\partial \bar{\sigma}_{\phi_k}}{\partial \theta_{k,j}} = \nabla \bar{\sigma}_{\phi_k}(T) \cdot \frac{\partial T}{\partial \theta_{k,j}}, \quad \theta_{k,j} \in \boldsymbol{\Theta}_k \tag{21}$$

$$\frac{\partial \bar{\sigma}_{I_k}}{\partial \theta_{k,j}} = \nabla \bar{\sigma}_{I_k}(T) \cdot \frac{\partial T}{\partial \theta_{k,j}}, \quad \theta_{k,j} \in \boldsymbol{\Theta}_k \tag{22}$$

$$\frac{\partial \bar{\rho}_k}{\partial \theta_{k,j}} = \nabla \bar{\rho}_k(T) \cdot \frac{\partial T}{\partial \theta_{k,j}}, \quad \theta_{k,j} \in \boldsymbol{\Theta}_k \tag{23}$$

Note that because of the presence of scaling factor in front of the level sets in (14), the partial derivatives of a_{ϕ_k} w.r.t s_k in (18) take a different form from other global and local transformation parameters in (19).

According to (15), the level sets evolve under the confluence of the force field in CACE and the proposed constraint force. The shape term drives the level sets towards the aligned prior shapes. In the denominator, the squared deviation allows shape variability, hence the shape driving force is applied in a soft-to-hard manner. Meanwhile, the appearance term penalises or strengthens the constraint force depending on the correlation between image and shape. The more distant the input image from the aligned prior image, the stronger the term is, where image variability is also taken into account. It plays the role of an extra criterion for constraining the level sets and brings further flexibility to the constraint.

4.5 Initial conditions

Although CCACE (or CACE) is less sensitive to initial placement, it is still not entirely independent from the initial conditions. Hence, we devise an automated initialisation method for CCACE for higher efficiency, reliability and practicality of the model.

Given the fact that the closer the initial shape to the final shape to be recovered, the more efficient the segmentation process is, we aim to find a rough shape estimate over the

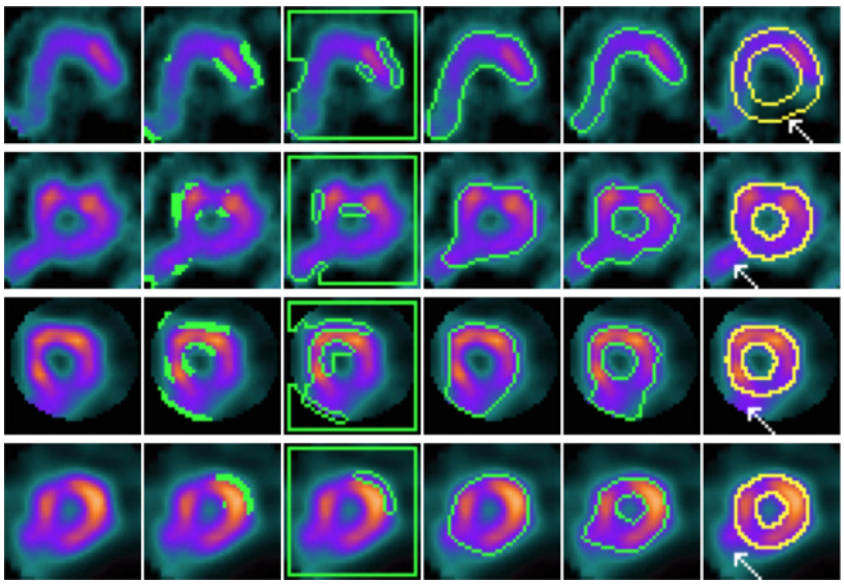

Fig. 4 *CACE and other models on mid-slices containing perfusion defects and artefacts (marked by white arrows)*
From left, each column: input slice, results of CPM, geodesic snake, GVF geodesic snake, CACE and ground truth

input image as an initial shape for CCACE to start with. Towards this end, we set $\lambda_I = 1$ in (14), so that the objective function $E_A(\boldsymbol{\phi}, \boldsymbol{c})|_{\lambda_I=1}$ takes information of greylevel intensities only (with shape information ignored), and measures the probability of the input image to occur. Note that the correlation coefficients $\bar{\rho}_k$ can be assumed $\mathbf{0}$ and $\boldsymbol{\phi}$ can be ignored for $E_A(\phi, \boldsymbol{c})|_{\lambda_I=1}$. We then find the minimum of $E_A(\boldsymbol{\phi}, \boldsymbol{c})|_{\lambda_I=1}$, by updating the spatiotemporal parameters $\boldsymbol{c}$, starting from $\mathbf{0}$, according to a motion equation taking the same form as (16), except that the partial derivatives in it are rewritten as

$$\left.\frac{\partial E_A(\phi, \boldsymbol{c})}{\partial \theta_{k,j}}\right|_{\lambda_I=1} = \int_{\boldsymbol{X}} \left[\frac{\partial \bar{\sigma}_{I_k}(T)}{\partial \theta_{k,j}} \left(\frac{1}{\bar{\sigma}_{I_k}(T)} - \frac{a_{I_k}^2}{\bar{\sigma}_{I_k}^3(T)} \right) + \frac{\partial a_{I_k}}{\partial \theta_{k,j}} \frac{a_{I_k}}{\bar{\sigma}_{I_k}^2(T)} \right] \mathrm{d}\boldsymbol{x} \tag{24}$$

With the recovered spatiotemporal parameters, denoted as $\hat{\boldsymbol{c}}$, and subsequently the recovered transformation parameters $\hat{\boldsymbol{\Theta}}_k$, a probabilistic image estimate for input image I_k can be obtained and denoted as $\bar{I}_k(T(\hat{\boldsymbol{\Theta}}_k; \mathbf{x}))$. Its corresponding shape estimate is therefore $\bar{\boldsymbol{\phi}}_k(T(\hat{\boldsymbol{\Theta}}_k; \mathbf{x}))$. We set $\phi_k = \bar{\phi}_k(T(\hat{\boldsymbol{\Theta}}_k; \boldsymbol{x}))$ and $\boldsymbol{c} = \hat{\boldsymbol{c}}$ as the initial conditions for the coupled evolution task in (15) and (16). λ_I is then set to 0.5 so that both shape and appearance are considered equally during the evolution.

5 Experimental results

We used 15 sets of gated cardiac SPECT sequences of healthy LVs with normal perfusion for training, and applied CCACE to another 14 sequences for performance evaluation, two of which are pathological cases and present perfusion defects. All patient studies were clinical studies acquired following standard protocols. The resolution of the data was $64 \times 64 \times \alpha \times 8$ where α varied between 25 and 58. Throughout

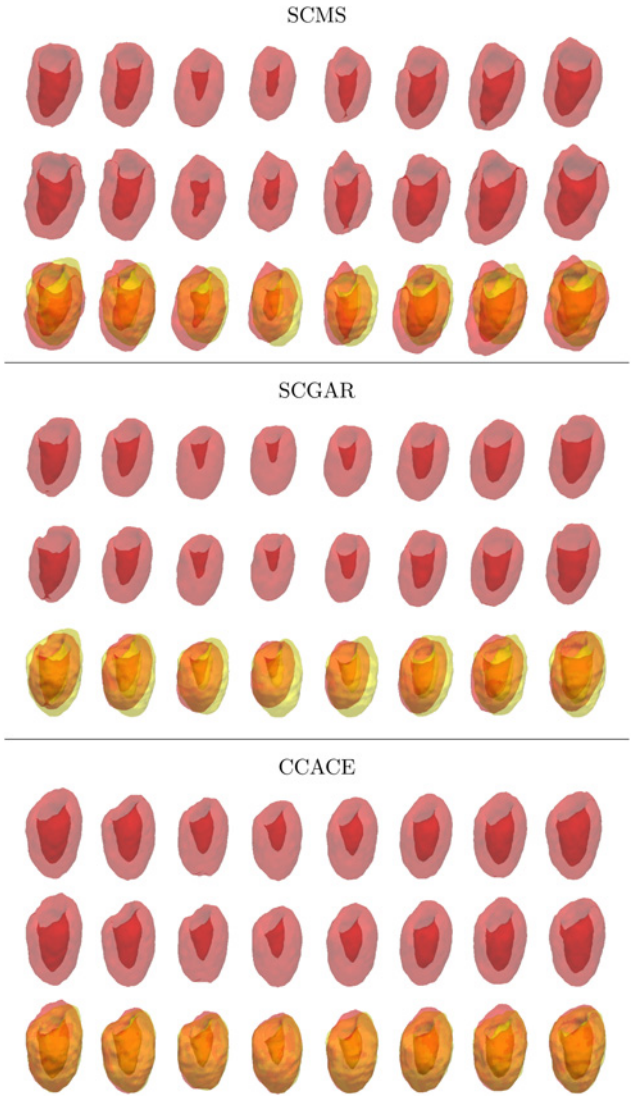

Fig. 5 *Iterations of SCMS (rows 1–3), SCGAR (rows 4–6) and CCACE (rows 7–9) on gated cardiac SPECT with artefacts and severe perfusion defects*
Final row of each model shows the results overlaid on the ground truth

doi: 10.1049/iet-cvi.2012.0081

the experiments, we set $\lambda_c = 0.4$ and $\lambda_I = 0.5$ which were determined empirically. For each $\boldsymbol{P}_k$ and $\boldsymbol{Q}$, the 10 most significant modes were chosen to explain 95% of the variations, that is, $m_k = 10$ for $k \in [1, K]$ and $n = 10$.

For comparison, we also implemented the SCMS and SCGAR models based on [23] and [27], respectively. Since no particular initialisations are specified for these two models, we initially set them as the prior shapes for each example. Note, unlike the 4D nature of SCMS and CCACE, SCGAR deals with each timeframe individually.

We use the following colour scheme in the results that follow: model results in red, ground truth in yellow and localised defect regions in cyan.

5.1 Subjective assessment

To illustrate the necessity of the use of priors, in Fig. 4 we show typical results from a number of data driven methods, including charged particle model (CPM) [36], geodesic active contour [37], GVF geodesic [38] and our base model, CACE. Although CACE has successfully recovered the shapes of the perfusion, it has not outlined the true LV borders according to the ground truth, as the perfusion state does not reflect the true shape of the LV in the presence of defects or artefacts. This is because CACE evolves towards any image gradients it can capture, without considering the anatomical features of the object it is dealing with. Any

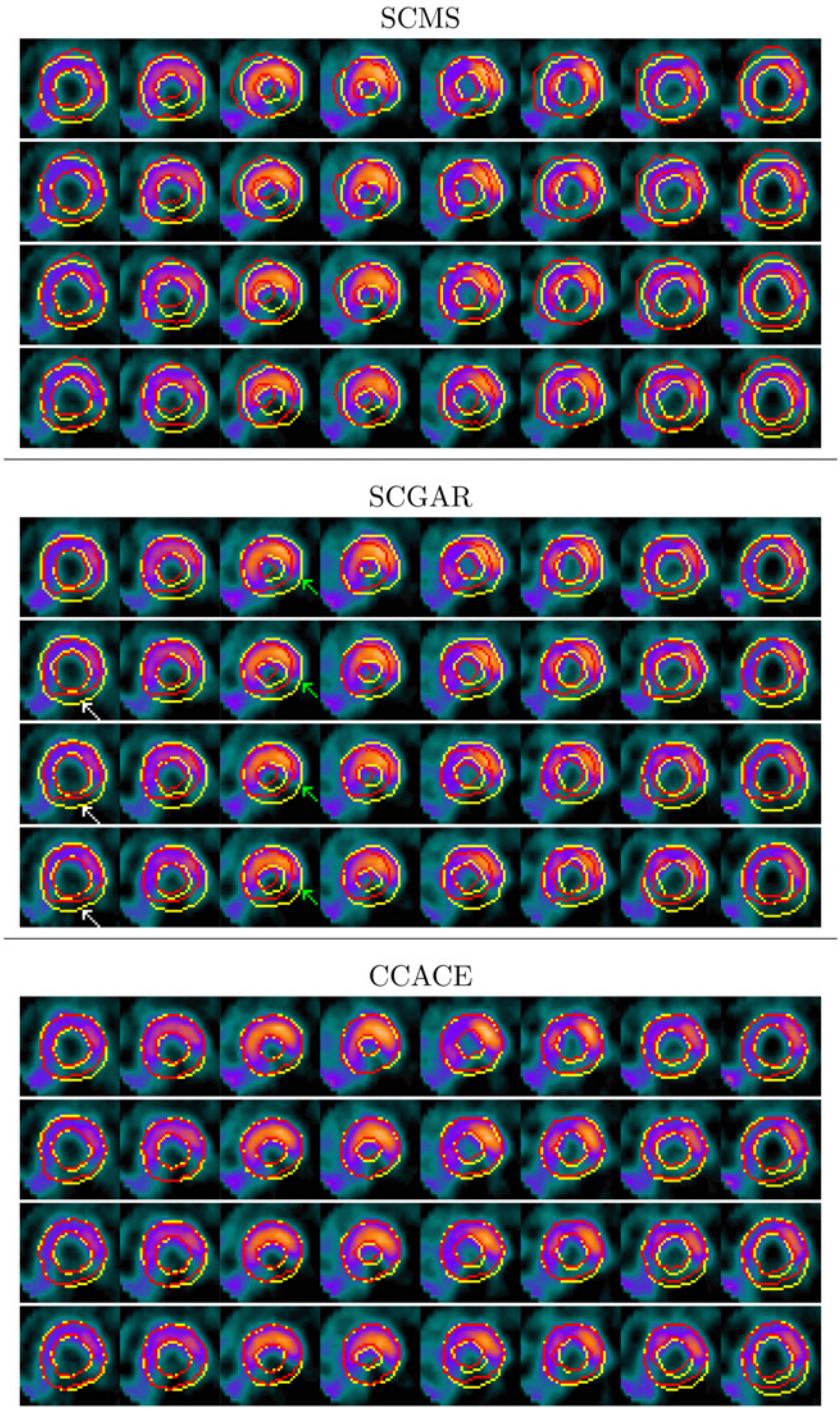

Fig. 6 *SCMS (rows 1–4), SCGAR (rows 5–8) and CCACE (rows 9–12) on slices selected from the sequence*
From top (each group): four slices at different locations within middle region of the LV

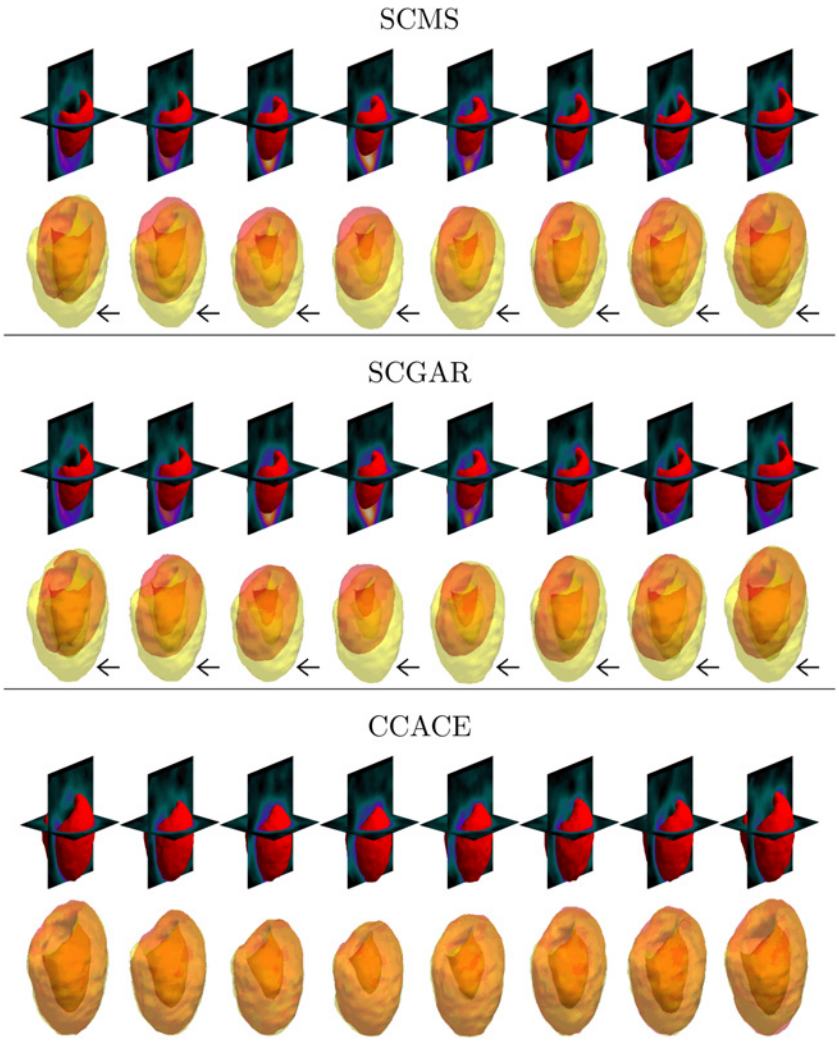

Fig. 7 *SCMS (rows 1–2), SCGAR (rows 3–4) and CCACE (rows 5–6) on a large LV with long apical region*
Black arrows indicate significant areas of mismatch against ground truth

other general data-driven method that only considers image features is likely to suffer the same problem, as demonstrated in the examples. Hence, the rest of the comparative analysis is focused on methods that are using prior-based segmentation.

Fig. 5 shows iterations of SCMS (rows 1–3), SCGAR (rows 4–6) and CCACE (rows 7–9) for an LV with severe myocardial infarction in the inferior wall (consistent perfusion defects present throughout the sequence as well as perfusion artefacts because of uptake in the liver). For each model snapshots are shown in the first two rows, whereas final results against ground truth are shown in the third row. For visual assessment, clearly the less yellow (or red) is seen in the third row for each case better the results. Fig. 6 shows the results in selected 2D slices.

Large deviations between the SCMS results arise because SCMS only allows global scaling and translation and local variations that have occurred among the training sets, thus is strictly restricted to only a pre-defined range of variations. Meanwhile, SCGAR is more flexible than SCMS, as it takes into account rotation in addition to scaling and translation, and allows the model to locally deform based on region forces. It can therefore cope with more variations than SCMS and achieves higher accuracy in many areas. However, the SCGAR model tends to merge over the defect areas as marked by white arrows in Fig. 6 (1st column of SCGAR results), because of its data-driven forces, whereas CCACE sustains a plausible shape and successfully differentiates the epicardial from the endocardial borders in these areas. SCGAR has large deviations from the borders in the lateral segments as marked by green arrows (3rd column of SCGAR results), whereas CCACE converges more closely around them. This shows CCACE has more flexibility as its constraint is more alert to defect regions and less sensitive to lateral segments.

The limitation of SCGAR is due to the fact that the model only removes global variations between the prior and the input shape which means there remain large deviations between them in areas containing local variations such as the defect region and lateral segments. Hence, the ratio between the deviations over plausible variations (e.g. in lateral segments) and implausible variations (e.g. in defect areas), and the resulting constraint forces which are proportional to these deviations, tend to be small. In contrast, as plausible variations between the joint priors and the input shape and image are accounted for through the regularised global-to-local transformations in CCACE, the deviations between them are diminished over plausible variations and kept large over implausible ones. As a result, the magnitudes of the constraint forces over the defects are

doi: 10.1049/iet-cvi.2012.0081

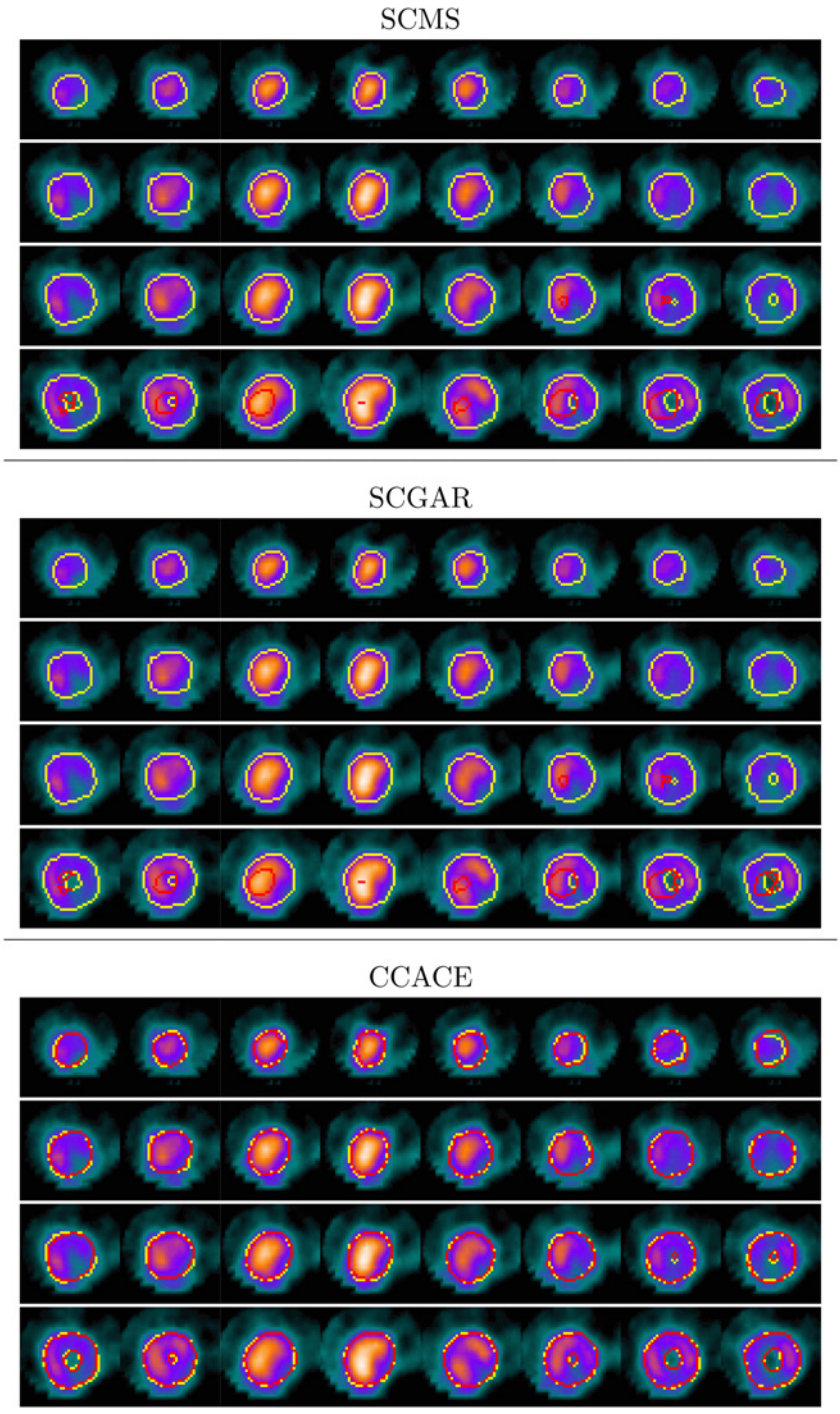

Fig. 8 *SCMS (rows 1–4), SCGAR (rows 5–8) and CCACE (rows 9–12) on slices selected from the sequence*
From top (each group): four slices at different locations within the apical region of the LV

significantly stronger than the lateral segments and other normal regions. CCACE can therefore give a good estimate over the defective borders without much compromiseof its accuracy in other regions.

In addition, the appearance term in CCACE makes crucial contributions to the constraint by preventing the model from merging over the defect regions, because there is significant difference between the intensities of the defect and the aligned prior image, leading to a large appearance term. Meanwhile, in normal perfusion areas the appearance term exerts little influence because of small intensity differences. Hence, this term plays the role of a meaningful criterion that brings more flexibility than shape-based constraints only.

Figs. 7 and 8 show the models' results on an example of a large LV with a long apical region in particular. SCMS and SCGAR fail to produce accurate estimates especially in the apical regions as marked by black arrows, whereas CCACE has achieved much better results.

5.2 Quantitative evaluation

The Jaccard coefficient is a commonly used statistic for similarity (or diversity) measure between two sets. It is the ratio between the size of the intersection and the size of the union of two sets, A and B, under comparison: $|A \cap B|/|A \cup B|$. We compute the Jaccard coefficient between the

Table 2 Evaluation of model results on cardiac SPECT datasets

D	*M*	Jaccard coefficient								
		O	Timeframes							
1	*A*	53.1	47.3	58.7	53.2	52.2	47.6	42.9	46.7	49.4
	B	55.6	47.8	59.3	55.0	51.4	47.8	44.7	50.3	52.2
	C	**94.2**	**92.9**	**93.6**	**94.6**	**95.3**	**95.1**	**94.2**	**93.1**	**94.5**
2	*A*	59.2	64.9	66.3	64.5	59.7	53.4	50.0	54.4	60.3
	B	63.2	63.0	67.3	69.3	59.1	55.0	54.7	61.2	64.6
	C	**93.5**	**92.8**	**93.4**	**94.0**	**94.4**	**94.3**	**93.8**	**93.2**	**92.0**
3	*A*	63.8	65.4	74.9	68.2	62.7	57.0	52.7	61.8	67.5
	B	62.1	62.6	73.6	65.5	57.5	52.9	53.3	63.9	67.3
	C	**93.5**	**92.0**	**94.3**	**93.8**	**94.4**	**94.5**	**93.7**	**92.7**	**92.5**
4	*A*	61.2	63.0	68.4	62.9	64.6	57.9	53.3	55.7	64.1
	B	65.2	63.9	69.6	66.0	66.1	61.1	60.0	63.2	69.8
	C	**93.3**	**92.2**	**93.2**	**94.3**	**95.0**	**94.5**	**93.7**	**93.2**	**90.6**
5	*A*	58.4	64.3	68.5	62.3	55.5	48.4	51.9	57.5	58.5
	B	57.7	53.8	66.1	63.1	54.6	50.1	51.1	60.0	59.3
	C	**93.2**	**93.9**	**93.0**	**94.2**	**95.1**	**94.5**	**93.7**	**91.3**	**90.3**
6 (Fig. 7)	A	48.3	41.7	48.3	51.1	42.7	43.8	48.4	49.5	41.2
	B	48.9	41.8	48.5	51.1	42.5	44.0	48.8	50.4	41.8
	C	**90.6**	**88.0**	**89.3**	**90.1**	**93.0**	**92.0**	**91.3**	**90.1**	**90.7**
7*	*A*	63.0	73.7	67.0	58.8	53.4	60.0	58.4	62.1	70.9
	B	62.5	57.9	57.8	58.4	50.9	61.0	72.4	73.2	68.6
	C	**87.3**	**85.4**	**89.0**	**90.2**	**91.0**	**85.0**	**85.4**	**85.6**	**85.6**
8* (Fig. 5)	A	63.1	71.7	67.2	63.1	62.7	60.4	60.5	55.8	63.4
	B	62.3	62.4	63.1	66.3	58.8	55.0	65.3	64.6	63.0
	C	**85.8**	**83.6**	**85.6**	**86.7**	**86.4**	**83.5**	**84.4**	**87.9**	**88.6**
9	*A*	59.8	66.0	63.7	60.8	55.2	56.1	55.4	59.7	61.6
	B	63.2	62.2	64.1	65.9	57.3	62.0	64.1	67.3	62.5
	C	**85.4**	**86.7**	**81.8**	**88.5**	**85.8**	**83.6**	**87.0**	**85.1**	**84.8**
10	*A*	55.2	54.7	60.1	54.3	59.7	54.4	48.3	47.8	52.9
	B	68.4	66.9	70.4	64.1	66.0	65.3	65.7	64.9	69.9
	C	**84.7**	**85.6**	**88.4**	**81.5**	**89.8**	**85.5**	**80.0**	**83.7**	**83.4**
11	*A*	57.0	57.3	58.8	53.1	56.1	55.1	51.8	49.8	59.1
	B	61.5	58.9	61.1	55.1	56.4	56.7	58.6	54.4	61.6
	C	**84.6**	**79.6**	**84.4**	**89.7**	**85.3**	**85.1**	**87.3**	**85.0**	**80.4**
12	*A*	70.9	74.2	79.9	76.0	72.1	66.8	63.3	64.2	70.5
	B	78.8	81.9	85.1	81.9	72.9	71.4	75.3	80.6	81.3
	C	**82.6**	**81.5**	**81.8**	**82.4**	**85.4**	**84.3**	**83.1**	**82.7**	**79.9**
13	*A*	49.8	43.2	51.7	54.2	47.1	46.7	50.2	53.8	51.8
	B	40.8	31.0	41.1	44.6	42.5	40.7	40.5	42.2	39.3
	C	**82.3**	**68.8**	**83.7**	**83.2**	**89.2**	**84.8**	**85.8**	**83.5**	**79.6**
14	*A*	48.8	48.5	55.0	55.3	49.5	43.1	41.3	47.5	50.4
	B	47.1	41.5	51.5	54.4	47.9	42.7	41.6	47.4	47.8
	C	**80.9**	**76.1**	**75.5**	**84.6**	**85.8**	**82.5**	**82.6**	**77.5**	**82.6**

Column '*D*': dataset indices. Column '*M*': model names, where '*A*', '*B*' and '*C*' represent SCMS, SCGAR and CCACE, respectively, and CCACE results are highlighted in bold. Column '*O*': the overall accuracy for each sequence

model results and the ground truth as the accuracy measure for the SCMS, SCGAR and CCACE models, as given in Table 2. The results for the examples shown above are also labelled in the table. The datasets marked with stars in the 'D' column are pathological cases with perfusion defects. The average accuracy of CCACE across the whole dataset is 88.0%, with $\sigma = 4.9\%$, and that of SCMS and SCGAR is 58.0% with $\sigma = 6.5\%$ and 59.8% with $\sigma = 9.5\%$, respectively.

From Table 2 it can be seen that CCACE has achieved higher accuracy than SCMS and SCGAR on every dataset. It also exhibits more consistent performance along the time dimension than the other two models, attributed to its spatiotemporal regularisation. The best performance of CCACE is 94.2% on a normal perfusion sequence with mild attenuation. In this example, the CCACE model has converged closely onto the LV borders in overall, with minor errors mainly resulting in the apical and basal areas. The worst result of CCACE is 80.9% on a normal sequence with artefacts, attenuation and large local variations across the whole LV, where CCACE has managed to outline the LV shapes, but errors occur not only in the apical and basal but also in middle regions with large local variations.

5.3 Computation time

Using a 2.8 GHz Linux PC running uncompiled Matlab code, the computational times for the SCMS, SCGAR and CCACE models are given in Table 3. The training stage of CCACE takes significantly longer than the other two models, because of the extra local transformations involved and the fact that the global-to-local transformations are applied to each dataset twice (the first one for alignment and the second one for spatiotemporal

Table 3 Computational times for the models

Models		SCMS	SCGAR	CCACE
time	training	51 min	53 min	223 min
	segmentation	56 s	95 s	149 s

doi: 10.1049/iet-cvi.2012.0081

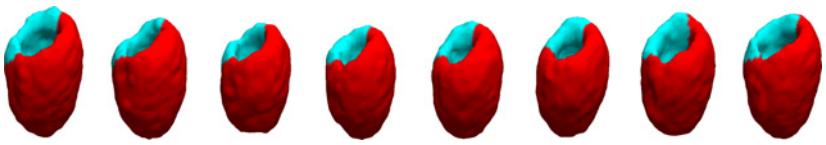

Fig. 9 *Automated defect detection by applying CACE to the segmented image based on the CCACE result*

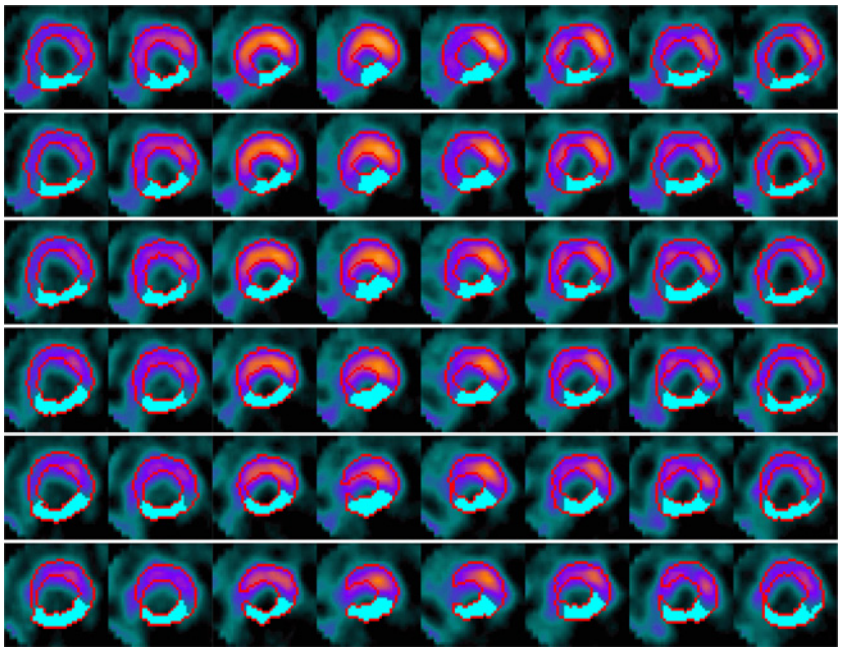

Fig. 10 *Results of defect detection in 2D slices*

Table 4 Evaluation of defect detection

D	*M*	Jaccard coefficient								
		O	Timeframes							
7*	*A*	74.3	81.1	88.2	80.3	89.4	76.9	47.3	51.4	79.9
	B	50.2	22.6	37.2	67.3	51.9	37.3	47.5	44.0	35.8
	C	78.2	78.6	83.0	82.6	81.4	71.4	77.2	70.2	81.0
8* (Fig. 9)	*A*	53.8	63.9	44.2	45.3	45.4	47.5	10.0	51.6	46.4
	B	38.8	41.0	39.6	38.7	39.9	36.9	40.7	34.2	29.7
	C	63.0	75.7	73.4	67.5	67.9	54.5	58.5	44.3	54.8

analysis). Each global-to-local transformation in CCACE training requires 54s on average, whereas the global transformation during SCMS and SCGAR training needs 25 s. The segmentation stage of CCACE takes more time than the other two models also because of the extra local transformations as well as the more complex computation of energy minimisation. In summary, CCACE clearly pays a price for greater accuracy.

5.4 Defect detection

In this section, we present a brief defect detection study to illustrate the potential of the proposed method. One of the ultimate goals of LV segmentation is to assist diagnosis of heart disease, e.g. finding myocardial infarction by localising perfusion defect. With the image correctly segmented, defect detection is fairly simple since the scope of detection is within the myocardium and the defects present much lower brightness than normal perfusion. We accomplish this detection task on the two example image sequences with defects) by running CACE on the segmented images obtained from CCACE's results. Other active contours with robust handling of weak gradients could also have been used.

We first define the segmented image as $H(\tilde{\phi}_k)I_k$ where $\tilde{\phi}_k$ is the level-set LV shape recovered by CCACE and $H(.)$ is a Heaviside function. The joint electrostatic force field is then derived for the segmented image according to (1) to guide CACE evolution towards defect boundaries. To automate the detection, we use the divergence-based initialisation method proposed in [39], by finding divergent points of the initial force field and selecting those in the dark areas

(potential defects) where the intensities are lower than 25% (determined empirically) of the maximum. Initial CACE snakes are then automatically placed around the selected divergent points to expand under the vector forces towards the boundaries of the defect regions.

Figs. 9 and 10 show our preliminary results of defect detection by applying CACE based on the segmentation results of CCACE, from which it can be seen that the defects are successfully localised. The same detection procedure was also applied to the segmentation results of SCMS and SCGAR for comparison. The Jaccard coefficients for the defect regions (as measured against the manually labelled ground truth) on the two defective datasets are given in Table 4.

6 Conclusions

In this paper we have proposed the CCACE model which recovers LV on gated cardiac SPECT by probabilistically determining its boundaries based on image gradients under a shape and appearance-based spatiotemporal constraint within a MAP framework. CCACE a) makes use of global-to-local rather than global registration to account for larger range of variability, b) applies appearance information in addition to shape information and c) has MAP spatiotemporal regularisation that takes the temporal domain into consideration. The first two aspects allow more flexibility in the constraint, whereas the third brings more robustness to irregularities in temporal variations of the input images.

The model has shown promising results on gated cardiac SPECT even on data with artefacts or severe perfusion defects. Further improvements may be achieved, for example, by utilising the kernel density estimator as suggested in [26] towards better evaluation of prior distribution of spatiotemporal parameters, and by applying a more robust optimisation scheme such as the trust region method.

7 References

1 Schumacher, H., Modersitzki, J., Fischer, B.: 'Combined reconstruction and motion correction in SPECT imaging', *IEEE Trans. Nucl. Sci.*, 2009, **56**, (1), pp. 73–80

2 Bronnikov, A.: 'SPECT imaging with resolution recovery', *IEEE Trans. Nucl. Sci.*, 2012, **59**, (4), pp. 1458–1464

3 Faro, A., Giordano, D., Spampinato, C., Ullo, S., Stefano, A.D.: 'Basal ganglia activity measurement by automatic 3-D striatum segmentation in SPECT imaging', *IEEE Trans. Instrum. Meas.*, 2009, **60**, (10), pp. 3269–3280

4 Depuey, E., Garcia, E., Berman, D.: 'Cardiac SPECT imaging' (Lippincott Williams and Wilkins, Philadelphia, 2000, 2nd edn.)

5 Loats, H.: 'CT and SPECT image registration and fusion for spatial localization of metastatic processes using radiolabeled monoclonals', *J. Nucl. Med.*, 1993, **34**, pp. 562–566

6 Montagnat, J., Delingette, H., Malandain, G.: 'Cylindrical Echocar-diographic images segmentation based on 3D deformable models' (MICCAI, 1999) pp. 168–175

7 Sermesant, M., Forest, C., Pennec, X., Delingette, H., Ayache, N.: 'Deformable biomechanical models: application to 4D cardiac image analysis', *Med. Image Anal.*, 2003, **7**, (4), pp. 475–488

8 Dornheim, L., Tönnies, K.D., Dixon, K.: 'Automatic segmentation of the left ventricle in 3D SPECT data by registration with a dynamic anatomic model' (MICCAI, 2005), vol. 8, pp. 335–342

9 Choi, S., Kim, H., Oh, J., Kang, T., Sun, K., Kim, M.: 'Segmentation of the left ventricle in myocardial perfusion SPECT using variational level set formulation'. Nuclear Science Symp. Conf. Record, 2007, vol. 4, pp. 3060–3064

10 Hentschke, C., Engel, K., Schafer, S., Tonnies, K.: 'Segmentation of the left ventricle in SPECT by an active surface' (MIUA, 2009) pp. 1–5

11 Debreuve, E., Barlaud, M., Aubert, G., Laurette, I., Darcourt, J.: 'Space-time segmentation using level set active contours applied to myocardial gated SPECT', *IEEE Trans. Med. Imaging*, 2001, **20**, (7), pp. 643–659

12 McInerney, T., Terzopoulos, D.: 'A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis', *Comput. Imaging Graph.*, 1995, **19**, pp. 69–83

13 Bardinet, E., Cohen, L., Ayache, N.: 'Tracking and motion analysis of the left ventricle with deformable superquadrics', *Med. Image Anal.*, 1996, **1**, pp. 129–149

14 Montagnat, J., Delingette, H.: '4D deformable models with temporal constraints: application to 4D cardiac image segmentation', *Med. Image Anal.*, 2005, **9**, (1), pp. 87–100

15 Chandrashekara, R., Rao, A., Sanchez-Oritz, G., Mohiaddin, R., Rueckert, D.: 'Construction of a statistical model for cardiac motion analysis using non-rigid image registration', *Inf. Process. Med. Imaging*, 2003, **18**, pp. 599–610

16 O'Connor, M., Kanal, K., Gebhard, M., Rossman, P.: 'Comparison of four motion correction techniques in SPECT imaging of the heart: a cardiac phantom study', *J. Nucl. Med.*, 1998, **39**, pp. 2027–2034

17 Laading, J., McCulloch, C., Johnson, V., Gill, D., Jaszczak, R.A.: 'Hierarchical feature based deformation model applied to 4D cardiac SPECT Data'. Int. Conf. Information Processing in Medical Imaging 1999, (*LNCS*, **1613**), pp. 266–279

18 Brankov, J., Yang, Y., Wernick, M.: 'Spatiotemporal processing of gated cardiac SPECT images using deformable mesh modeling', *Med. Phys.*, 2005, **32**, (9), pp. 2839–2849

19 Charnoz, A., Lingrand, D., Montagnat, J.: 'A levelset based method for segmenting the heart in 3D + T gated SPECT images', Int. Workshop on Functional Imaging and Modeling of the Heart, 2003, vol. 2674, pp. 87–100

20 Leventon, M., Grimson, W., Faugeras, O.: 'Statistical shape influence in geodesic active contours' (CVPR, 2000) pp. 316–323

21 Tsai, A., Yezzi, A., Wells, W., *et al.*: 'Model-based curve evolution technique for image segmentation' (CVPR, 2001) pp. 463–468

22 Cremers, D., Tischhauser, F., Weickert, J., Schnorr, C.: 'Diffusion snakes: introducing statistical shape knowledge into the Mumford-Shah functional', *Comput. Vis.*, 2002, **50**, (3), pp. 295–313

23 Kohlberger, T., Cremers, D., Rousson, M., Ramaraj, R., Funka-Lea, G.: '4D shape priors for a level set segmentation of the left myocardium in SPECT sequences' (MICCAI, 2006), vol. 9, pp. 92–100

24 Chan, T., Vese, L.: 'Active contour without edges', *IEEE Trans. Image Process.*, 2001, **10**, (2), pp. 266–277

25 Kohlberger, T., Funka-Lea, G., Desh, V.: 'Soft level set coupling for LV segmentation in gated perfusion SPECT' (MICCAI, 2007), vol. 10, pp. 327–334

26 Rousson, M., Cremers, D.: 'Efficient kernel density estimation of shape and intensity priors for level set segmentation' (MICCAI, 2005), vol. 8, pp. 335–342

27 Rousson, M., Paragios, N.: 'Shape priors for level set representations' (ECCV, 2002) pp. 78–92

28 Fisher, V., Lee, R., Gourin, A., Bolooki, H., Stuckey, J., Kavaler, F.: 'Muscle fiber length: a determinant of left ventricular contraction pattern', *Am. J. Physiol.*, 1966, **211**, pp. 301–306

29 Gleason, S., Abidi, M., Sari-Sarraf, H.: 'Probabilistic shape and appearance model for scene segmentation' (ICRA, 2002), vol. 3, pp. 2982–2987

30 Yang, J., Duncan, J.: '3D image segmentation of deformable objects with shape-appearance joint prior models' (MICCAI, 2003) pp. 573–580

31 Huang, X., Li, Z., Metaxas, D.: 'Learning coupled prior shape and appearance models for segmentation' (MICCAI, 2004), pp. 60–69

32 Litvin, A., Karl, W.C., Shah, J.: 'Shape and appearance modeling with feature distributions for image segmentation' (ISBI, 2006), pp. 1128–1131

33 Ben-Ari, R., Aiger, D.: 'Geodesic active contours with combined shape and appearance priors' (ACIVS, 2008) pp. 494–505

34 Yang, R., Mirmehdi, M., Xie, X.: 'A charged active contour based on electrostatics' (ACIVS, 2006) pp. 173–184

35 Cootes, T., Edwards, G., Taylor, C.: 'Active appearance models' (ECCV, 1998), vol. 2, pp. 484–498

36 Jalba, A.C., Wilkinson, M.H.F., Roerdink, J.B.T.M.: 'CPM: a deformable model for shape recovery and segmentation based on charged particles', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2004, **26**, (10), pp. 1320–1335

doi: 10.1049/iet-cvi.2012.0081

37 Caselles, V., Kimmel, R., Sapiro, G.: 'Geodesic active contours' (ICCV, 1995), pp. 694–699
38 Paragios, N., Mellina-Gottardo, V., Ramesh, O.: 'Gradient vector flow fast geodesic active contours', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2004, **26**, (3), pp. 402–407
39 Yang, R., Mirmehdi, M., Hall, D.: 'Charged contour model for cardiac SPECT segmentation' (MIUA, 2006) pp. 171–175
40 Huang, X., Paragios, N., Metaxas, D.: 'Shape registration in implicit spaces using information theory and free form deformations', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2006, **28**, (8), pp. 1303–1318

8 Appendix 1

The global-to-local registration method proposed for CCACE combines rigid transformation and incremental free form deformation (IFFD) [40] and applies a sum of squared differences (SSD) criterion with a smoothness constraint, defined as

$$E_T(\boldsymbol{\Psi};\boldsymbol{x}) = (1-\lambda_s)\int_{\mathbf{X}} \underbrace{\left(s\phi_D - \phi_S(T(\boldsymbol{\Psi};\boldsymbol{x}))\right)^2}_{\mathbf{SSD}} + \lambda_s \underbrace{\left(\left\|\frac{\partial \delta T_l}{\partial x}\right\|^2 + \left\|\frac{\partial \delta T_l}{\partial y}\right\|^2 + \left\|\frac{\partial \delta T_l}{\partial z}\right\|^2\right)}_{\mathbf{Smoothness}} \mathrm{d}\boldsymbol{x} \quad (25)$$

where $\boldsymbol{\phi}_S$ is the source shape, which is to be registered onto the target shape $\boldsymbol{\phi}_D$. The constant λ_s balances between the SSD measure and the smoothness constraint. The transformation model $T(\boldsymbol{\Psi};\boldsymbol{x})$ is defined as

$$T(\boldsymbol{\Psi};\boldsymbol{x}) = sA_R(\boldsymbol{x}+\delta T_l) + \begin{bmatrix} h_x \\ h_y \\ h_z \end{bmatrix} \quad (26)$$

where

$$A_R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\omega_x & \sin\omega_x \\ 0 & -\sin\omega_x & \cos\omega_x \end{bmatrix} \begin{bmatrix} \cos\omega_y & 0 & -\sin\omega_y \\ 0 & 1 & 0 \\ \sin\omega_y & 0 & \cos\omega_y \end{bmatrix} \begin{bmatrix} \cos\omega_z & \sin\omega_z & 0 \\ -\sin\omega_z & \cos\omega_z & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (27)$$

$$\delta T_l = \sum_{u=0}^{3}\sum_{v=0}^{3}\sum_{w=0}^{3} B_i(\alpha)B_j(\beta)B_k(\gamma) \begin{bmatrix} \delta P^x_{i+u,j+v,k+w} \\ \delta P^y_{i+u,j+v,k+w} \\ \delta P^z_{i+u,j+v,k+w} \end{bmatrix} \quad (28)$$

$\boldsymbol{\Psi}$ is the global-to-local transformation parameters defined as s, ω_x, ω_y, ω_z, h_x, h_y, h_z, δP. The rigid transformation is constituted of scaling s, rotation $(\omega_x, \omega_y, \omega_z)$ and translation (h_x, h_y, h_z). δP is the displacement of an $L_p \times M_p \times N_p$ mesh grid of control points overlaid on the domain with uniform spacings δ_x, δ_y and δ_z, and $\delta P = \{(\delta P^x_{l,m,n}, \delta P^y_{l,m,n}, \delta P^z_{l,m,n}):(l, m, n) \in [1, L_p] \times [1, M_p] \times [1, N_p]\}$. The indices i, j, k denote the position of the control point at the top left corner of the cell that contains $\boldsymbol{x}$, and α, β, γ are the relative positions of x, y, z within the cell, that is, $i = \lfloor x/\delta_x \rfloor - 1$, $j = \lfloor y/\delta_y \rfloor - 1$, $k = \lfloor z/\delta_z \rfloor - 1$ and $\alpha = x/\delta_x - \lfloor x/\delta_x \rfloor$, $\beta = y/\delta_y - \lfloor y/\delta_y \rfloor$, $\gamma = z/\delta_z - \lfloor z/\delta_z \rfloor$. $B_i(\alpha)$ is the ith basis function of the cubic B-spline

$$B_0(\alpha) = (1-\alpha)^3/6 \quad (29)$$

$$B_1(\alpha) = (3\alpha^3 - 6\alpha^2 + 4)/6 \quad (30)$$

$$B_2(\alpha) = (-3\alpha^3 + 3\alpha^2 + 3\alpha + 1)/6 \quad (31)$$

$$B_3(\alpha) = \alpha^3/6 \quad (32)$$

$B_j(\beta)$ and $B_k(\gamma)$ are similarly defined.