


# Multivariate Mixture Model for Myocardial Segmentation Combining Multi-Source Images

Xiahai Zhuang 

**Abstract**—The author proposes a method for simultaneous registration and segmentation of multi-source images, using the multivariate mixture model (MvMM) and maximum of log-likelihood (LL) framework. Specifically, the method is applied to the problem of myocardial segmentation combining the complementary information from multi-sequence (MS) cardiac magnetic resonance (CMR) images. For the image misalignment and incongruent data, the MvMM is formulated with transformations and is further generalized for dealing with the hetero-coverage multi-modality images (HC-MMIs). The segmentation of MvMM is performed in a virtual common space, to which all the images and misaligned slices are simultaneously registered. Furthermore, this common space can be divided into a number of sub-regions, each of which contains congruent data, thus the HC-MMIs can be modeled using a set of conventional MvMMs. Results show that MvMM obtained significantly better performance compared to the conventional approaches and demonstrated good potential for scar quantification as well as myocardial segmentation. The generalized MvMM has also demonstrated better robustness in the incongruent data, where some images may not fully cover the region of interest, and the full coverage can only be reconstructed combining the images from multiple sources.

**Index Terms**—Multivariate image, multi-modality, segmentation, registration, medical image analysis, cardiac MRI

## 1 INTRODUCTION

SEGMENTATION from medical images is an essential prerequisite in a number of imaging data assisted medical applications, such as in the localization and quantification of tissues and pathologies, and modeling of anatomical structures [1], [2], [3]. Magnetic resonance (MR) technology provides an important means for imaging anatomical and functional information of the heart, in particular the late gadolinium enhancement (LGE) cardiac magnetic resonance (CMR) sequence which visualizes myocardial infarction (MI), the T2-weighted CMR which images the acute injury and ischemic regions, and the balanced-Steady State Free Precession (bSSFP) cine sequence which captures cardiac motions and presents clear boundaries [4]. Fig. 1 provides an example of the three sequences. Notice that the term *image* in this paper means a 3D *scan*, not a 2D image which is referred to as a *slice*. LGE CMR enhances the infarcted myocardium, to appear with distinctive brightness compared with the healthy tissues, and therefore is effective in determining the presence, location, and extent of MI. To identify the MI regions and subsequently perform the quantitative analysis, which is important in the diagnosis and treatment management of patients, a common method is to first delineate the myocardium from the LGE CMR

images, where the pathologies can be differentiated from the healthy tissues using intensity-based threshold algorithms [4], [5], [6].

Since manual delineation is generally time-consuming, tedious and subject to inter- and intra-observer variations, automating this segmentation is desired in clinical practice, which is however still arduous, particularly due to the pathological myocardium from LGE CMR. Besides the great variations of the heart shape across different subjects, there are three major issues related to the intensity distributions of the images, which challenge the conventional intensity-based classification methods, such as the Gaussian mixture model (GMM)-based segmentation [7], [8], [9]. First, the intensity range of myocardium in LGE CMR overlaps those of its surroundings, leading to indistinguishable boundaries from its adjacent organs. For example, the infarcted myocardium, which is enhanced in LGE CMR, can appear identical to the blood pools, and the intensity range of the healthy myocardium can be similar to that of the adjacent liver or lung. Also, the pathologies result in heterogeneous intensity of the myocardium, making the assumption of a simple distribution invalid; for example the single component Gaussian density function is widely used in the segmentation of bSSFP cine CMR [3], [8], but can not be used for LGE CMR. Finally, the enhancement patterns can be complex. The location, size and shape of infarcted regions vary greatly across different patients, and the microvascular obstruction (MVO), which occurs in some patients, appears as hypo-enhanced areas (due to the lack of contrast agent uptake) within the hyper-enhanced regions. Hence, it is difficult to predict or make assumptions of the location and geometry of the MI regions.

- The author is with the School of Data Science, Fudan University, Shanghai, 200433, China. E-mail: zxh@fudan.edu.cn.

Manuscript received 22 Sept. 2017; revised 27 July 2018; accepted 5 Sept. 2018. Date of publication 10 Sept. 2018; date of current version 31 Oct. 2019. (Corresponding author: Xiahai Zhuang).

Recommended for acceptance by M. Betke.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/TPAMI.2018.2869576

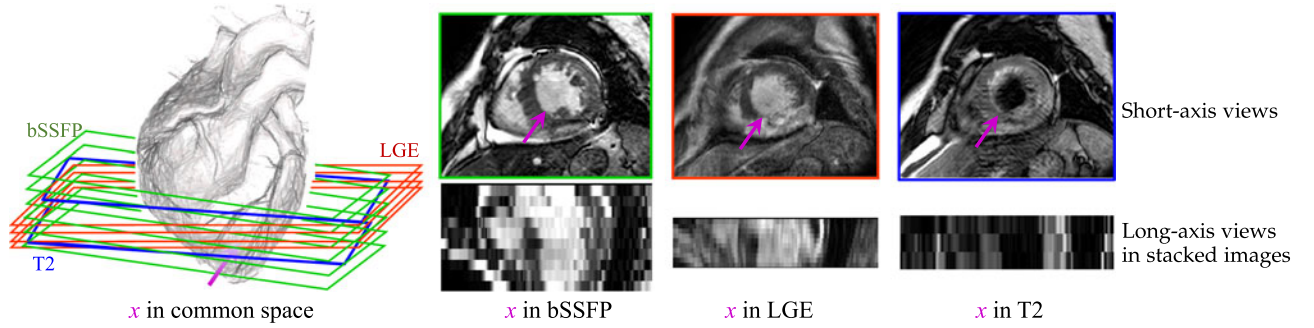


Fig. 1. Illustration of multi-modality images and the common space of a subject. Three CMR images are acquired from the patient, i.e., the bSSFP, LGE, and T2 CMR, which form  $I = [I_{\text{bSSFP}}, I_{\text{LGE}}, I_{\text{T2}}]$ . The pink arrows indicate the same position at the common space and the three images which have different appearances and intensity values: The bSSFP CMR provides good boundary information between the myocardium and blood pools, but one could not see the scars from the bSSFP CMR; LGE CMR visualizes the scars as brighter texture, in contrast to the dark healthy myocardium, but the boundaries between the scars and the blood pools are indistinct; T2 CMR presents the myocardial edema in higher intensity values and can visualize the trabeculations and pupillary muscles in great detail, but the resolution in z-direction is low (generally 3 to 6 slices). Therefore, the author proposes to combine the complementary information from the three types of images within a unified framework for simultaneous registration and segmentation.

### 1.1 Related Work

To the best of our knowledge, little work has been done in the fully automated myocardial segmentation from LGE CMR. To obtain such segmentation, the algorithms generally need to integrate the prior shape information of the myocardium. A number of reported methods used the segmentation result of the bSSFP cine CMR from the same subject, acquired in the same session, as *a priori* knowledge. The LGE CMR segmentation can then be achieved by directly propagating the bSSFP segmentation to the LGE CMR. For this implementation, different registration methods can be used, such as 2D rigid registration based on a shift window [10], affine registration [11], or rigid registration incorporating multi-scale total variation flow [12]. To detect myocardial contours, a 2D geometrical template was proposed in [13], where the myocardium was modeled as a closed ribbon structure with an imaginary centerline and varying width. The myocardial template was divided into four quadrants, of which each was assigned an energy term based on the anatomical prior and knowledge of potential scars, for a deformable adaptation of the template to the target image. In [14], a 3D mesh was built, based on the propagated prior segmentation from the bSSFP, and then deformed towards the myocardial contours in the LGE CMR to compensate for the difference. In [15], the authors proposed a 1D parametric model to detect the paired endocardial and epicardial edge points, where the intensity patterns along the radial rays from left ventricle (LV) center to beyond the epicardium were modeled. They imposed a thickness constraint for the 3D deformation.

In the literature, there is a research topic referred to as multivariate image analysis, which has been proposed to deal with images that have more than one measurement per pixel [16], [17]. This research has particularly focused on the three-channel (RGB) color images or multi-spectral and hyper-spectral images, and the methods have been designed to efficiently compress the highly correlated data and project them onto a reduced dimensional subspace, for example using the principal component analysis [18]. Multivariate image analysis considers only *congruent* data, meaning for each pixel in one image there should be a corresponding pixel in the other images [17]. However, the medical images acquired from different sources commonly

do not have this congruency. These images can be misaligned to each other, and the imaging field-of-views and resolution can vary greatly across different acquisitions. The images having such differences are referred to as *incongruent data* or *hetero-coverage* images in this work.

### 1.2 Contribution of This Work

The related works in the literature for myocardial segmentation from LGE CMR are mainly semi-automatic [14], or to propagate the segmented myocardium from one image as a constraint for the following separate segmentation of the other [11], [15]. The separate segmentation however can be challenged to maintain a consistent and accurate result for the multi-source images or the multi-sequence (MS) CMR in this specific task of myocardial segmentation. The consistency is important since the multi-source images come from the same subject, and it can be obtained when the images are combined and segmented simultaneously in a unified framework.

This work proposes the combined segmentation for multi-source images, meaning the segmentation of the images is performed *simultaneously* in a common space. This combined segmentation can delineate the indistinct boundaries, such as those between the infarcted myocardium and the blood pools in LGE CMR, under the guidance of the T2 and bSSFP CMR, and vice versa. This is achieved by formulating the segmentation using the multivariate mixture model (MvMM) and the maximum of log-likelihood (LL) framework. The MvMM is adopted for modeling the joint intensity distribution of the multi-modality images, i.e., the multi-source images.

Furthermore, in medical images there exists inter- and intra-image misalignment, and the images can vary greatly in sizes and pixel dimensions. Hence, the MvMM is formulated with transformations and is further generalized in the common space for dealing with the hetero-coverage multi-modality images (HC-MMIs). The misalignment is mainly due to the respiratory motions of the subject during acquisition, and it not only happens between images, but also happens among the slices from the same image. The transformations, which are embedded into the LL framework, are assigned to each slice of the images as well as the images. The transformation-embedded LL is then optimized using the iterative conditional mode (ICM) method,

where the MvMM parameters (segmentation) are obtained using the Expectation Maximization (EM) algorithm and the transformation (registration) parameters are updated using the gradient ascent optimization approach.

The generalization of MvMM for HC-MMIs is formulated in a virtual common space, where all the images and misaligned slices are registered and the segmentation is performed. This common space can be divided into a number of sub-regions, each of which contains congruent imaging data and thus can be modeled using a conventional MvMM. Using this generalization, one can model the hetero-coverage images and perform the segmentation of them simultaneously in the common space. This enables the segmentation of the regions with indistinguishable boundaries such as the myocardial infarcts.

The rest of this paper is organized as follows: Section 2 describes the proposed method in detail; Section 3 presents the experimental setup and evaluation results, using the clinical MS CMR images for assessing the myocardial segmentation and using the simulated MS brain MR for parameter studies and comparisons. In Section 4, the author concludes this work and presents insights for potential extension of the proposed method.

## 2 METHOD

The goal is to classify the anatomical structures from the multi-modality CMR images, which provide diverse morphological and pathological information of the heart. By combining the multi-source information and performing the segmentation of the images simultaneously, one aims to obtain a segmentation result with sufficient accuracy.

Let  $I = \{I_i | i=1, \dots, N_I\}$  be the set of  $N_I$  images acquired from the same subject. One can denote the spatial domain of the region of interest (ROI) of the subject as  $\Omega$ , referred to as *the common space*, which is the coordinate of the subject and thus defined by the combination of the images. Fig. 1 illustrates the concept of the common space using the three CMR sequences, where the image variable is  $I = [I_{bSSFP}, I_{LGE}, I_{T2}]$ .

For a location  $x \in \Omega$ , the tissue type of  $x$  is determined regardless the appearance of the medical images. One can denote tissue types using label values, namely  $s(x) = k$ ,  $k \in K$ ,  $K$  is the set of labels, and refer to subtypes of a tissue  $k$  in image  $I_i$  as  $z_i(x) = c$ ,  $c \in C_{ik}$ ,  $C_{ik}$  is the set of subtypes. Note that provided the images are all aligned to the common space, the label information for all the images should be the same, but the texture of them and classification of subtypes can be different. In the example of CMR images in Fig. 1, the myocardial scars are visible in LGE CMR, but not in bSSFP. Therefore, one can use two Gaussian components to model the intensity distribution of myocardium in LGE, i.e.,  $C_{LGE\text{myo}} = \{c_{\text{scar}}, c_{\text{normal}}\}$ , while for the myocardium in bSSFP one can use one component, i.e.,  $C_{bSSFP\text{myo}} = \{c_{\text{myo}}\}$ . Here, a *type* (label) is defined when its prior probability is available from a prior probabilistic atlas, such as the LV and myocardium; and a *subtype* (component) is used to denote the tissues or structures, for which we do not have the prior atlas, such as the scars whose prior atlas is impossible to construct, or the background organs whose separated atlases are too costly to build.

In the following, Section 2.1 introduces the MvMM and LL framework. Section 2.2 provides the EM approach and

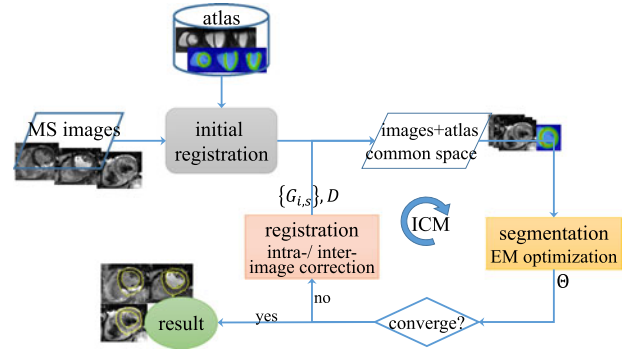


Fig. 2. Flowchart of the proposed myocardial segmentation method from the MS CMR.

the initialization of the parameters. The registration is described in Section 2.4, and the optimization of the two sets of parameters is given in Section 2.5. Finally, Section 2.6 generalizes the MvMM for incongruent data, i.e., the images with different coverage. Fig. 2 presents the flowchart of the proposed segmentation framework.

### 2.1 Multivariate Mixture Model in Log-Likelihood Framework

For a single image, one can use the mixture of Gaussian to model the intensity distributions, namely the GMM method, where the intensity probability density function (PDF) of one tissue is given by a Gaussian function. For a tissue with multiple subtypes, the multi-component GMM can be used [19]. For the multi-source image segmentation, one can use the MvMM.

The likelihood ( $LH$ ) of the MvMM parameters  $\theta$  is given by  $LH_{\Omega}(\theta; I) = p(I|\theta)$ , similar to the GMM for classification, where  $I = \{I_1, \dots, I_{N_I}\}$  is the multi-modality image vector. Assuming independence of each location (pixel), one gets  $LH_{\Omega}(\theta; I) = \prod_{x \in \Omega} p(I(x)|\theta)$ . The label at each location and component information at each image are hidden data. Note that for the convenience and conciseness of denotation, the author sometimes uses the short terms  $k_x$  and  $c_{ix}$  (or  $c_{ikx}$ ) to indicate  $s(x) = k$  and  $z_i(x) = c$  (or  $z_i(x) = c_{ik}$ ) when no confusion is caused.

The conditional probability of an image vector at location  $x$  given the model parameters can be computed as follows,

$$p(I(x)|\theta) = \sum_{k \in K} \pi_{kx} p(I(x)|s(x) = k, \theta). \quad (1)$$

Here,  $\pi_{kx} = p(s(x) = k|\theta)$  is the label proportion, and  $\pi_{kx} = p(k) = \pi_k$  when no spatial constraint is considered. The graphical representation of this MvMM probabilistic model is provided in Fig. 3a.

When the tissue type of a position is known, the intensity distributions of different images become independent, namely

$$p(I(x)|s(x) = k, \theta) = \prod_{i=1, \dots, N_I} p(I_i(x)|k_x, \theta). \quad (2)$$

The intensity PDF of an image,  $p(I_i(x)|k_x, \theta)$ , is given by the conventional multi-component GMM, as follows,

$$p(I_i(x)|k_x, \theta) = \sum_{c \in C_{ik}} \tau_{ikc} \Phi_{ikc}(\mu_{ikc}, \sigma_{ikc}, I_i(x)), \quad (3)$$



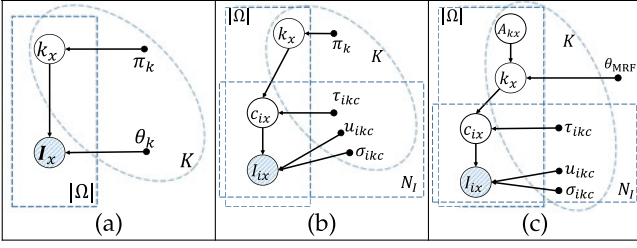


Fig. 3. The graphical representation of the multivariate mixture model in three formulations. Readers are referred to the text for details.

where,  $\tau_{ikc} = p(c_{ikx}|k_x, \theta)$ , s.t.  $\sum_c \tau_{ikc} = 1$ , is the component proportion, and  $\Phi_{ikc}(\cdot) = p(I_i(x)|c_{ikx}, \theta)$  is the Gaussian function modeling the intensity PDF of a tissue subtype  $c$  of tissue  $k$  in the image  $I_i$ . The graphical representation of this probabilistic model is provided in Fig. 3b.

## 2.2 Expectation Maximization Algorithm

The LL of the complete data, which includes both the observed images and the hidden data, is as follows,

$$\begin{aligned} LL_{com} &= \sum_{x \in \Omega} \sum_{k \in K} \delta_{s(x), k} \left\{ \log \pi_{kx} \right. \\ &\quad \left. + \sum_i \sum_{c_{ik}} \delta_{z_i(x), c_{ik}} (\log \tau_{ikc} + \log \Phi_{ikc}(I_i(x))) \right\} \\ &= \sum_x \sum_k \delta_{s(x), k} \log \pi_{kx} \\ &\quad + \sum_x \sum_k \sum_i \sum_{c_{ik}} \delta_{s(x), k} \delta_{z_i(x), c_{ik}} \text{Term}_{LTP}, \end{aligned}$$

where  $\delta_{a,b}$  is the Kronecker delta function, and  $\text{Term}_{LTP} = (\log \tau_{ikc} + \log \Phi_{ikc}(I_i(x)))$ . The conditional expectation of  $LL_{com}$ , given the current parameter  $\theta^{[m]}$  at the  $m$ -th step and the observed images, is  $Q(\theta|\theta^{[m]}) = E(LL_{com}|I, \theta^{[m]})$ . Therefore, one can employ the following E-step and M-step iteratively to obtain the model parameters and then the segmentation variables.

**E-Step.** One can obtain the expectation of  $LL_{com}$  by computing the expectation of  $\delta_{s(x), k}$  and  $\delta_{s(x), k} \delta_{z_i(x), c_{ik}}$  given the observed images and current estimation of model parameters,

$$\begin{aligned} P_{kx}^{[m+1]} &= E_{(I, \theta^{[m]})}(\delta_{s(x), k}) \\ &= p(s(x) = k | I, \theta^{[m]}) \\ &= \frac{p(I(x)|k_x, \theta^{[m]}) \pi_{kx}^{[m]}}{\sum_{l \in K} p(I(x)|l_x, \theta^{[m]}) \pi_{lx}^{[m]}}, \end{aligned} \quad (4)$$

and,

$$\begin{aligned} P_{ikcx}^{[m+1]} &= E_{(I_i, \theta^{[m]})}(\delta_{s(x), k} \delta_{z_i(x), c_{ik}}) \\ &= p(s(x) = k, z_i(x) = c_{ik} | I, \theta^{[m]}) \\ &= p(c_{ikx}|k_x, I, \theta^{[m]}) P_{kx}^{[m+1]} \\ &= \frac{p(I(x)|c_{ikx}, k_x, \theta^{[m]}) \tau_{ikc}^{[m]}}{p(I(x)|k_x, \theta^{[m]})} P_{kx}^{[m+1]} \\ &= \frac{\Phi_{ikc}^{[m]}(\sigma_{ikc}^{[m]}, I_i(x)) \tau_{ikc}^{[m]}}{p(I_i(x)|k_x, \theta^{[m]})} P_{kx}^{[m+1]}, \end{aligned} \quad (5)$$

where,  $p(I(x)|c_{ikx}, k_x, \theta^{[m]})$  is computed using  $p(I_i(x)|c_{ikx}, \theta^{[m]}) \prod_{j \neq i} p(I_j(x)|k_x, \theta^{[m]})$ , which is based on the same assumption

for  $p(I(x)|k_x, \theta^{[m]})$  in (2) that the intensity distributions of the image vector becomes independent when the label and component information are known, and the common part,  $\prod_{j \neq i} p(I_j(x)|k_x, \theta^{[m]})$ , in the numerator and denominator is then cancelled out;  $\pi_{kx}^{[m]}$  and  $\tau_{ikc}^{[m]}$  are calculated in the previous M-Step. Here,  $P_{kx}^{[m+1]}$  and  $P_{ikcx}^{[m+1]}$  are also referred to as the estimation results of the hidden data.

**M-Step.** The model parameters are updated by analytically maximizing  $Q(\theta|\theta^{[m]})$ ,

$$\begin{aligned} \tau_{ikc}^{[m+1]} &= \frac{\sum_{x \in \Omega} P_{ikcx}^{[m+1]}}{\sum_{d \in C_{ik}} \sum_{x \in \Omega} P_{ikdx}^{[m+1]}} \\ \mu_{ikc}^{[m+1]} &= \frac{\sum_{x \in \Omega} I_i(x) P_{ikcx}^{[m+1]}}{\sum_{x \in \Omega} P_{ikcx}^{[m+1]}} \\ (\sigma_{ikc}^{[m+1]})^2 &= \frac{\sum_{x \in \Omega} (I_i(x) - \mu_{ikc}^{[m+1]})^2 P_{ikcx}^{[m+1]}}{\sum_{x \in \Omega} P_{ikcx}^{[m+1]}}. \end{aligned} \quad (6)$$

For  $\pi_k^{[m+1]}$ , one can compute  $\frac{\partial Q(\theta|\theta^{[m]})}{\partial \pi_k} = 0$ , and when no spatial regularization is applied, the update is given by,  $\pi_k^{[m+1]} = (\sum_x P_{kx}^{[m+1]}) / (\sum_l \sum_x P_{lx}^{[m+1]})$ .

## 2.3 Spatial Regularization

The finite mixture model, using (1) and  $\pi_{kx} = p(k)$ , describes the statistical intensity distribution of the pixels without considering spatial information. However, pixels with the same intensity distribution in medical images can come from different structures. For example, the left ventricle blood pool generally has the same intensity distribution as the right ventricle. Here, the proposed method uses the spatial regularization to deal with this problem. This section first introduces the constraint from an atlas for computing prior probabilities, then describes the Markov Random Fields (MRFs) to incorporate the contextual dependency.

### 2.3.1 Probabilistic Atlases

An atlas with prior probabilities of labels, also referred to as probabilistic atlases, are commonly used for computation of the label proportion  $\pi_{kx}$ , such that

$$\pi_{kx} = p(s(x) = k | \theta) = \frac{\pi_k p(A_{kx})}{NF}, \quad (7)$$

where  $p(A_{kx}) = p_A(s(x) = k)$  is the atlas prior probability,  $NF$  is the normalization factor [8]. This scheme provides a constraint to penalize the misclassification of pixels which are far away from the cluster of true labels. Fig. 4 provides the example of an atlas of CMR and the probability maps of the four labels.

For this formulation, one needs to update  $\pi_k$  in M-step by computing,

$$\begin{aligned} \frac{\partial Q(\theta|\theta^{[m]})}{\partial \pi_k} &= \sum_x \frac{P_{kx}^{[m+1]}}{\pi_k} - \sum_x \sum_{j \in K} \frac{P_{jx}^{[m+1]} p_A(s(x) = k)}{\sum_{l \in K} p_A(s(x) = l) \pi_l} \\ &= 0. \end{aligned}$$

Since there is no closed form for the solution, one can use numerical methods such as the iterative gradient ascent method to search the optimal solution for  $\pi_k^{[m+1]}$ . Alternatively,

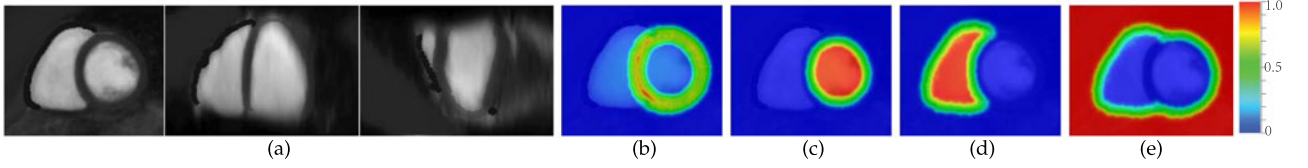


Fig. 4. The three orthogonal views of the atlas intensity image (a), and the short-axis views of the four probabilistic atlases of myocardium (b), left ventricle blood pool (c), right ventricular blood pool (d) and background (e). The probability maps are superimposed onto the intensity image, and the color bars indicate the mapping between the probability values and displayed colors.

one can regard  $C_x^{[m]} = \sum_{l \in K} p(A_{lx}) \pi_l$  as a constant using  $\pi_l = \pi_l^{[m]}$  in the  $[m+1]$ th iteration, which results in

$$\pi_k^{[m+1]} = \frac{\sum_x P_{kx}^{[m+1]}}{\sum_x (p(A_{kx})/C_x^{[m]})}. \quad (8)$$

This guarantees an improvement of the likelihood of the parameters with respect to the images in each iteration and leads to a generalized EM scheme [20].

Initialization of  $\pi_k^{[0]}$ ,  $\tau_{ikc}^{[0]}$ ,  $\mu_{ikc}^{[0]}$  and  $\sigma_{ikc}^{[0]}$  are computed based on the atlas prior probabilities, which are propagated by the atlas-to-target image registration step [21],

$$\pi_k^{[0]} = \frac{\sum_x p(A_{kx})}{\sum_{l \in K} \sum_x p(A_{lx})}, \tau_{ikc}^{[0]} = \frac{1}{|C_{ik}|}$$

$$\mu_{ikc}^{[0]} = \begin{cases} \mu_{ik}^{[0]} + a\sigma_{ik}^{[0]}, & |C_{ik}| \geq 2 \\ \mu_{ik}^{[0]}, & |C_{ik}| = 1 \end{cases}, \left(\sigma_{ikc}^{[0]}\right)^2 = \frac{\left(\sigma_{ik}^{[0]}\right)^2}{|C_{ik}|}$$

where  $a$  is a proportion factor to uniformly sample  $|C_{ik}|$  values from  $[-1, 1]$  for the  $|C_{ik}|$  components of label  $k$  in image  $I_i$ ;  $\mu_{ik}^{[0]} = \frac{\sum_x I_i(x)p(A_{kx})}{\sum_x p(A_{kx})}$ , and  $\left(\sigma_{ik}^{[0]}\right)^2 = \frac{\sum_x (I_i(x) - \mu_{ik}^{[0]})^2 p(A_{kx})}{\sum_x p(A_{kx})}$ .

### 2.3.2 Markov Random Fields in Multivariate Mixture Model

The MRF can be used to incorporate spatial and anatomical constraints [7], [22], [23], [24], [25], [26], [27]. Specifically, it models the neighborhood dependencies for labeling the pixels, such that the probability of a pixel belonging to a label depends on the labeling of its neighbors. The label proportion is then defined as  $\pi_{kx} = f(s(x) = k | s(\mathcal{N}_x), \theta)$ , where  $\mathcal{N}_x$  is the set of neighboring pixels to  $x$ ,  $s(\mathcal{N}_x) = \{s(x') | x' \in \mathcal{N}_x\}$  denotes the labeling of the set of pixels in  $\mathcal{N}_x$ .

According to the Hammersley-Clifford theorem, an MRF can be equivalently characterized by a Gibbs distribution,

$$P(s(\Omega) | \theta_s) = \frac{1}{\mathcal{Z}} e^{-U_{\text{mrf}}(s | \theta_s)}, \quad (9)$$

where  $\mathcal{Z}$  is the normalization constant (also referred to as partition function),  $U_{\text{mrf}}()$  is the energy function.

Following the literature, one can use the Potts model to represent the spatial interactions [7], [8],

$$U_{\text{mrf}}(s | \theta_s) = \sum_x \hat{e}(x)^T G \hat{g}(x), \quad (10)$$

where,  $\hat{e}(x)$  is a vector indicating the label type of  $x$  such that if  $s(x) = k$ , then the  $k^{\text{th}}$  element of  $\hat{e}(x) = \hat{e}_k$  is one and other elements are zero;  $\hat{g}(x)$  is the sum of the label vectors of the neighborhood, i.e.,  $\hat{g}(x) = \sum_{x' \in \mathcal{N}_x} \hat{e}(x')$ ; and the  $K \times K$  matrix  $G$  encodes the MRF parameters, which can be

solved based on the same mean field idiom and using the least squares fit procedure. Readers are referred to the original papers of [7], [22] for details. Furthermore, in the MvMM framework the tissue type ( $s(x) = k$ ) is determined in the virtual common space, which is not limited to the physical resolution of any image and can be defined isotropically. Hence, the contributions of neighbors from the three directions are considered equally, and the energy function  $U_{\text{mrf}}()$  only takes one set of parameters ( $G$ ) for all the three directions.

In the proposed framework, the label prior can be solely defined in the common space, thus the MRF configuration is irrelevant to the physical images in the MvMM. Hence, the label prior  $\pi_{kx}^{[m]}$  in (4) can be directly redefined to the neighbor-dependent MRF estimation, as follows,

$$\pi_{kx}^{[m+1]} = f(s(x) = k | P_{\mathcal{N}_x}^{[m]}, \theta^{[m]})$$

$$= \frac{e^{-U_{\text{mrf}}(\hat{e}_k | P_{\mathcal{N}_x}^{[m]}, \theta^{[m]})}}{\sum_{l \in K} e^{-U_{\text{mrf}}(\hat{e}_l | P_{\mathcal{N}_x}^{[m]}, \theta^{[m]})}},$$

where  $U_{\text{mrf}}(\hat{e}_k | P_{\mathcal{N}_x}^{[m]}, \theta^{[m]}) = \hat{e}_k^T G \hat{g}(x)$ , and the estimation results of the hidden data is given in the previous E-step by  $P_{\mathcal{N}_x}^{[m]} = \{P_{kx'}^{[m]} | x' \in \mathcal{N}_x\}$ .

In practice, the constraint using an atlas in (7) can be similarly applied with the MRF configuration by multiplying the prior probability from the MRF estimation with the prior in the atlas,

$$\pi_{kx}^{[m+1]} = \frac{f(k_x | P_{\mathcal{N}_x}^{[m]}, \theta^{[m]}) p(A_{kx})}{NF}. \quad (11)$$

The graphical representation of this model is provided in Fig. 3c.

## 2.4 Registration in Multivariate Mixture Model

The MvMM is initialized and regularized by the prior probabilities from an atlas which can be registered to the common space of the target images using the conventional methods in the atlas-based segmentation framework [28].

However, there exist two types of misalignment. First, the motion shift of each slice is commonly seen in the multi-slice CMR, besides the misalignment of the whole image to the common space of the subject. Fig. 5 visualizes the motion shifts in the *in vivo* CMR and the corrected images. Second, the atlas, providing the prior probabilities, can be mis-registered to the common space in some local details.

The motion shift of a slice is modeled by a rigid transformation, leading to reformulation of the intensity PDF of a subtype tissue as follows,

$$p(I_i(x) | c_{ik}, \theta, G_{i,s}) = \Phi_{ikc}(I_i(G_{i,s}(x))), \quad (12)$$

where  $\{G_{i,s}\}$  are the transformations for correcting slices.

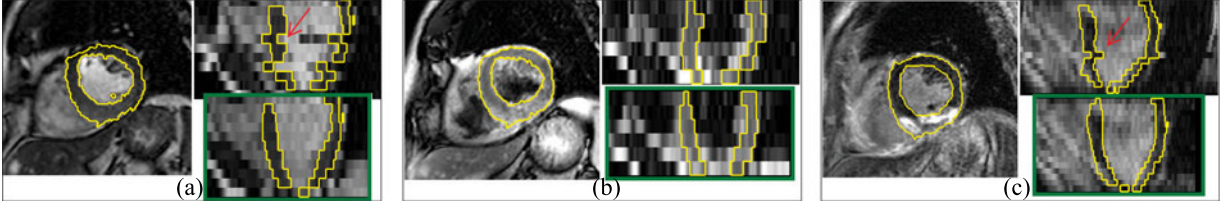


Fig. 5. Segmentation results of the three CMR sequences, (a) bSSFP, (b) T2-weight, (c) LGE; myocardial boundaries are highlighted in yellow color; the motion shifts are pointed out by the red arrows. The images in the green boxes of (a)-(c) are the shift corrected images.

The atlas deformation, denoted using  $D$ , can be embedded into the atlas prior probabilities for correcting the local mis-registration,

$$p_A(s(x) = k|D) = p_A(s(D(x)) = k) = A_k(D(x)), k = 1 \dots K, \quad (13)$$

which are the probabilistic atlas images. With the deformation embedded prior  $\pi_{kx|D} = p(s(x) = k|D)$ , the LL becomes,

$$\begin{aligned} LL(\theta, D, \{G_{i,s}\}) &= \sum_{x \in \Omega} \log LH(x) \\ &= \sum_{x \in \Omega} \log \left\{ \sum_k \pi_{kx|D} \cdot p(I(x)|k_x, \theta, \{G_{i,s}\}) \right\} \\ &= \sum_{x \in \Omega} \log \left\{ \sum_k \pi_{kx|D} \prod_i p(I_i(x)|k_x, \theta, G_{i,s}) \right\} \\ &= \sum_{x \in \Omega} \log \left\{ \sum_k \pi_{kx|D} \prod_i \sum_{c_{ik}} \tau_{ikc} \Phi_{ikc}(I_i(G_{i,s}(x))) \right\}. \end{aligned} \quad (14)$$

Here, the short form of the likelihood term on each location,  $LH(x)$ , is introduced for convenience of denotation; the prior  $\pi_{kx|D}$  is computed using (7) or (11) (with MRF).

## 2.5 Iterative Conditional Mode Optimization

There is no closed form solution for minimization of (14). Since the Gaussian parameters depend on the values of the transformation parameters, and vice versa, one can use the ICM approach to solve this optimization problem, which is eventual a coordinate ascent method in this formulation [29]. The ICM scheme optimizes one group of parameters while keeping the others unchanged at each iteration. The different groups of parameters are alternately optimized and this alternation process iterates until a local optimum is found. Fig. 2 provides the flowchart of the framework, where the segmentation parameters  $\Theta$ , i.e., the MvMM parameters and the hidden data, are updated using the EM approach (Section 2.2), and the registration parameters, i.e., the transformations  $\{G_{i,s}\}$  and  $D$ , are optimized using the gradient ascent method.

The derivatives of LL with respect to the rigid transformations and atlas deformation are respectively given by,

$$\begin{aligned} \frac{\partial LL}{\partial G_{i,s}} &= \sum_x \frac{1}{LH(x)} \sum_k \pi_{kx|D} \prod_{j \neq i} \left\{ p(I_j(x)|k_x, \theta, G_{j,s}) \right. \\ &\quad \cdot \left. \sum_c (\tau_{ikc} \Phi'_{ikc} \nabla I_i(y) \nabla G_{i,s}(x)) \right\}, \end{aligned} \quad (15)$$

where  $y = G_{i,s}(x)$ , and

$$\frac{\partial LL}{\partial D} = \sum_x \frac{1}{LH(x)} \sum_k \frac{\partial \pi_{kx|D}}{\partial D} p(I(x)|k_x, \theta, \{G_{i,s}\}). \quad (16)$$

The computation of  $\frac{\partial \pi_{kx|D}}{\partial D}$  is related to  $\frac{\partial A_k(D(x))}{\partial D}$ , and

$$\frac{\partial A_k(D(x))}{\partial D} = \nabla A_k|_{y=[y_1, y_2, y_3]} \times \begin{bmatrix} \frac{\partial y_1}{\partial \phi_d}, \frac{\partial y_2}{\partial \phi_d}, \frac{\partial y_3}{\partial \phi_d} \end{bmatrix}^T.$$

Here,  $y = D(x)$ , and  $\{\phi_d\}$  are the free-form deformation parameters [30], [31].

During the optimization of registration, the partial derivative of the LL w.r.t the rigid transformation parameters and the nonrigid free-form deformation (FFD) parameters are respectively computed using (15) and (16). For the FFD, the derivative of the bending energy is also computed [30]. The update in each iteration step becomes,

$$\begin{aligned} G_{i,s}^{[k+1]} &= G_{i,s}^{[k]} + l_G \cdot \frac{\partial LL}{\partial G_{i,s}} \\ D^{[k+1]} &= D^{[k]} + l_D \cdot \left( \frac{\partial LL}{\partial D} + \lambda \cdot \frac{\partial C_{smooth}}{\partial D} \right). \end{aligned} \quad (17)$$

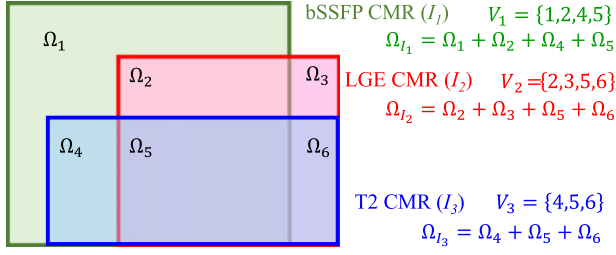
Here,  $l_G$  and  $l_D$  are the learning rates, also referred to as step length;  $C_{smooth}$  is the smoothness term, i.e., bending energy of the FFD transformation;  $\lambda$  is the weighting parameter controlling the tradeoff between the alignment of images and the smoothness of deformation. In this setting, the objective function becomes  $LL + \lambda C_{smooth}$ . The smoothness term is irrelevant to other parameters, either rigid transformations or segmentation parameters in the MvMM, the whole framework therefore is not affected by this regularization.

## 2.6 Hetero-Coverage Multi-Modality Images

In medical imaging, data from different acquisitions can generally have different resolutions and coverage of the ROI. The combination of these images is referred to as incongruent data, or hetero-coverage multi-modality images. For example, the bSSFP CMR sequence generally covers the whole ventricle, while the LGE and T2 sequences may only acquire the data from major part of the main ventricles without covering the apex or base, as Fig. 1 shows. Also, the missing information between the slices may be inappropriate to obtain by image interpolation techniques due to the thick slices and the large gap between them, i.e., low resolution in the inter-slice dimension. This section generalizes the formulation of MvMM for such HC-MMIs.

In the new formulation, the ROI of the subject is divided into  $N_{sr}$  non-overlapping sub-regions,  $\{\Omega_v|_{v=1 \dots N_{sr}}\}$ , and  $\Omega = \bigcup_{v=1}^{N_{sr}} \Omega_v$ . For image  $I_i$ , the volume is denoted as  $\Omega_{I_i}$ , and the set of indices of sub-regions belonging to  $\Omega_{I_i}$  is denoted as  $V_i$ . Therefore,  $\Omega_{I_i}$  is the sum of the set of sub-regions belonging to  $I_i$ , i.e.,  $\Omega_{I_i} = \bigcup_{v \in V_i} \Omega_v$ . Also, each sub-region  $\Omega_v$  fully covers a sub-volume of  $N_v$  images,  $N_v \in [1, N_I]$ , and thus one can apply a conventional  $N_v$ -variate mixture model for  $\Omega_v$ .





- $LL_{\Omega_1}$  and  $LL_{\Omega_3}$  use univariate mixture model ( $N_v = 1$ )
- $LL_{\Omega_2}$ ,  $LL_{\Omega_4}$  and  $LL_{\Omega_6}$  use bivariate mixture model ( $N_v = 2$ )
- $LL_{\Omega_5}$  uses multivariate mixture model ( $N_v = 3$ )

Fig. 6. Illustration of multi-modality images with different coverage, where the log-likelihood  $LL_{HC} = \sum_{v=1}^6 LL_{\Omega_v}$ .

Fig. 6 illuminates the sub-regions of HC-MMIs, and as it shows, one can rewrite the LL as follows,

$$LL_{HC} = \sum_{v=1}^{N_{sr}} LL_{\Omega_v} = \sum_{v=1}^{N_{sr}} \log LH_{\Omega_v}(\theta; I_v), \quad (18)$$

where,

$$LL_{\Omega_v} = \sum_{\Omega_v} \log \sum_{k \in K} \left\{ \pi_{kx} \prod_{i=1 \dots N_v} \left( \sum_{c \in C_{ik}} \tau_{ikc} \Phi_{ikc}(I_i(x)) \right) \right\}.$$

Here,  $I_v$  could be one image, such as solely for the bSSFP image when neither of the other two images covers sub-region  $\Omega_v$ , which results in a conventional univariate GMM.

Since  $LL_{HC}$  is the sum of the LL terms from a set of standard MvMM models, the optimization of the parameters is similar to that of LL in Section 2.2.

*E-Step.* For  $x \in \Omega_v$  and the images in  $\Omega_v$  is  $I_v$ , the estimation of the hidden label in (4) is computed as follows,

$$\begin{aligned} P_{kx}^{[m+1]} &= p(s(x) = k | I_v, \theta^{[m]}) \\ &= \frac{p(I_v(x) | k_x, \theta^{[m]}) \pi_{kx}^{[m]}}{\sum_{l \in K} p(I_v(x) | l_x, \theta^{[m]}) \pi_{lx}^{[m]}}, \end{aligned} \quad (19)$$

where  $p(I_v(x) | k_x, \theta^{[m]}) = \prod_{i=1 \dots N_v} p(I_i(x) | k_x, \theta^{[m]})$  is computed following (2) and  $\pi_{kx}^{[m]} = \frac{\pi_k^{[m]} p(A_{kx})}{N_F}$ . Having  $P_{kx}^{[m+1]}$  for  $\forall x \in \Omega$ , one can obtain  $P_{ikcx}^{[m+1]} = \frac{\Phi(\mu_{ikc}^{[m]}, \sigma_{ikc}^{[m]}, I_i(x)) \tau_{ikc}^{[m]}}{p(I_i(x) | k_x, \theta^{[m]})} P_{kx}^{[m+1]}$  for  $\forall x \in \Omega_{I_i}$ , similar to (5).

*M-Step:* By maximizing the  $Q$  function, one obtains,

$$\begin{aligned} \tau_{ikc}^{[m+1]} &= \frac{\sum_{v \in V_i} \sum_{x \in \Omega_v} P_{ikcx}^{[m+1]}}{\sum_{d \in C_{ik}} \sum_{v \in V_i} \sum_{x \in \Omega_v} P_{ikdx}^{[m+1]}} \\ &= \frac{\sum_{x \in \Omega_{I_i}} P_{ikcx}^{[m+1]}}{\sum_{d \in C_{ik}} \sum_{x \in \Omega_{I_i}} P_{ikdx}^{[m+1]}} \\ \mu_{ikc}^{[m+1]} &= \frac{\sum_{x \in \Omega_{I_i}} I_i(x) P_{ikcx}^{[m+1]}}{\sum_{x \in \Omega_{I_i}} P_{ikcx}^{[m+1]}} \\ (\sigma_{ikc}^{[m+1]})^2 &= \frac{\sum_{x \in \Omega_{I_i}} (I_i(x) - \mu_{ikc}^{[m+1]})^2 P_{ikcx}^{[m+1]}}{\sum_{x \in \Omega_{I_i}} P_{ikcx}^{[m+1]}}. \end{aligned} \quad (20)$$

The computation of  $\pi_{kx}^{[m+1]}$  is solely related to the atlas prior probability and the estimation of the hidden data or MRF estimation, which are all obtained within the common space using the different multivariate models. Therefore,

one can compute  $\pi_{kx}^{[m+1]}$  using (7) and (8), or the similar updating formula with MRF in (11).

Initialization of the parameters and optimization of the registration parameters are similar to the computation for the congruent data in Section 2.2 and 2.4, except that they are based on the sum of the sub-regions  $\{\Omega_v | v \in V_i\}$  for the variables of image  $I_i$ .

### 3 EXPERIMENTS AND RESULTS

This section consists of three experiments to evaluate the proposed MvMM segmentation method. Section 3.1 evaluates the performance of the proposed MvMM for myocardial segmentation from MS CMR. The information of the CMR data, evaluation metrics and implementation details are presented. The inter-observer and inter-sequence variations are also studied here. Section 3.2 compares the performance of different segmentation methods, including the conventional approaches and the MvMM with alternative registration schemes. The study focuses on the segmentation of LGE CMR, because it is challenging and is the main focus of this work. Section 3.3 is a parameter study, investigating the performance of the proposed MvMM with different dimensions of the multivariate variable and HC-MMIs. This study uses the simulated images from Brain Web [32]. The cplusplus implementation of MvMM is publicly available from [www.sdspeople.fudan.edu.cn/zhuangxiahai/0/zxhproj/](http://www.sdspeople.fudan.edu.cn/zhuangxiahai/0/zxhproj/).

#### 3.1 Segmentation Combining Multi-Sequence CMR

##### 3.1.1 Data

CMR. The CMR data from 35 patients, who underwent cardiomyopathy, had been collected from Shanghai Renji hospital with institutional ethics approval and had been anonymized. Each patient had been scanned using the three CMR sequences, i.e., the LGE, T2 and bSSFP. The three CMR sequences were all breath-hold, multi-slice, acquired in the ventricular short-axis views.

The LGE CMR was a T1-weighted, inversion-recovery, gradient-echo sequence, consisting of 10 to 18 slices, covering the main body of the ventricles. The typical parameters are as follows, TR/TE: 3.6/1.8 ms; slice thickness: 5 mm; in-plane resolution: reconstructed into  $0.75 \times 0.75$  mm.

The T2 CMR was a T2-weighted, black blood Spectral Presaturation Attenuated Inversion-Recovery (SPAIR) sequence, generally consisting of a small number of slices. For example, among the 35 cases, 13 have only three slices, and the others have five (13 subjects), six (8 subjects) or seven (one subject) slices. The typical parameters are as follows, TR/TE: 2000/90 ms; slice thickness: 12-20 mm; in-plane resolution: reconstructed into  $1.35 \times 1.35$  mm.

The bSSFP CMR was a balanced steady-state, free precession cine sequence. Since both the LGE and T2 CMR were scanned at the end-diastolic phase, the same cardiac phase of the bSSFP cine data was selected for this study. The bSSFP images generally consist of 8 to 12 contiguous slices, covering the full ventricles from the apex to the basal plane of the mitral valve, with some cases having several slices beyond the ventricles. The typical parameters are as follows, TR/TE: 2.7/1.4 ms; slice thickness: 8-13 mm; in-plane resolution: reconstructed into  $1.25 \times 1.25$  mm.

The parameters of the three CMR sequences are summarized in Table 1.

TABLE 1  
Information of the CMR Sequences

CMR	TR/TE (ms)	Pixel size	Slice thickness	No. of slices
LGE	3.6/ 1.8	0.75 × 0.75 mm	5 mm	10 - 18
T2	2000/ 90	1.35 × 1.35 mm	12-20 mm	3 - 7
bSSFP	2.7/ 1.4	1.25 × 1.25 mm	8-13 mm	8 - 12
Number of patients: 35				

*Gold Standard.* Each of the CMR images had been manually delineated three times, by three independent, well-trained observers who were not aware of the methodology of this work. The manual segmentation was performed slice-by-slice, using the brush tool in the ITK-SNAP [33], and the gold standard segmentation was achieved by averaging the three manual delineations using the shape-based approach [34].

*Atlas.* The atlas was built from *another* set of bSSFP images acquired from *healthy* subjects. The bSSFP images were first manually segmented, and then nonrigidly registered to a selected reference using a comprehensive registration method [21]. The atlas intensity image and label probabilities can be computed based on the mean of all the transformed images and label information [8]. To enhance the right ventricular (RV) boundary of the atlas intensity image, which was blurred due to the intensity averaging, the author manually added a low-intensity region to the atlas image, to simulate an intensity-distinct epicardial boundary of the RV. To obtain a more uniform and less biased probability map of labels, the probabilistic atlases were simulated by applying Gaussian convolution to the labels of the atlas and then normalizing the values as label probabilities. The CMR atlas has four labels, including, the myocardium of left ventricle, left ventricle blood pool, right ventricular blood pool and background. Fig. 4 illustrates the constructed atlas.

### 3.1.2 Evaluation Metrics

To evaluate the accuracy of segmentation results, the Dice metric, the average contour distance (ACD) and Hausdorff distance between the automatic segmentation and the corresponding gold standard are calculated:

Dice: Let  $V_{\text{seg}}$  represent the volume segmented by an algorithm and  $V_{\text{GD}}$  be the gold standard. The Dice is defined related to the overlap of the two volumes, i.e.,  $\frac{2|V_{\text{seg}} \cap V_{\text{GD}}|}{|V_{\text{seg}}| + |V_{\text{GD}}|}$ . Here,  $|V|$  indicates the volume size. Dice scores range from 0, meaning no overlap between the two volumes, to 1, implying perfect overlap and similarity.

ACD. The ACD metric computes the average euclidean distance between the two corresponding contours of the segmentation results. For every point in one contour, the euclidean distance to the nearest point in the other is calculated. To obtain a symmetric measure, one can compute the ACD based on the average of the computed distance of all the points from the two contours. The lower value of the ACD, the better accuracy of the segmentation is.

Hausdorff distance: This metric measures how far two subsets of a metric space are from each other. Let  $X$  and  $Y$  be two sets of contour points, and  $d(x, y)$  indicate the distance between two points. The Hausdorff distance is defined by,

$$HD(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}.$$

The lower the Hausdorff distance, the better the segmentation is.

### 3.1.3 Implementations

In the atlas-based segmentation, the atlas is registered to the target image using a hierarchical registration scheme, which is specifically designed for cardiac images and consists of three levels of transformations, i.e., affine, locally affine, and free-form deformations [21]. In the MvMM myocardial segmentation, the atlas is first registered to the bSSFP which covers the whole ventricles, and the registered atlas probabilities are propagated to the common space of the MS CMR.

For the multi-component modeling, both GMM and MvMM assign two components to the myocardium for the LGE and T2 CMR, respectively modeling the normal and abnormal tissues. For the background of all the three sequences, two components are used. For the other classes, such as the left ventricular blood pool or the right ventricular blood pool, one component is used.

The algorithms such as registration and EM iteration were implemented based on single thread using C++ code. These tools were then wrapped up using Matlab scripts to form different segmentation pipelines. The experiments were run on a Lenovo D30 ThinkStation which had 3.30 GHz Intel Xeon E5-2667 V2 CPU and 64 GB main memory.

### 3.1.4 Results of MS CMR Segmentation

Table 2 presents the average Dice scores, ACD and Hausdorff distance of the proposed MvMM, which segments the three CMR sequences simultaneously. One can see that all the three sequences have been well segmented, particularly the LGE CMR which is the most challenging and important task in cardiac segmentation. This is mainly attributed to the combination of the complementary information, particularly from the T2 and bSSFP sequences which provide the critical boundary information for the segmentation of the LGE CMR on the scar regions.

One can notice that the Dice scores and ACD of the T2 segmentation are not as good as those of bSSFP, which seems to contradict the fact that T2 images generally have better contrast, and a better accuracy should be expected. This is mainly due to the fact that the trabeculae carneae of the heart are more visible and the myocardium thus appears to be thicker in the T2 CMR images. Our observers also tended to include the trabeculae carneae into the myocardium in the manual segmentation, leading to segmentation inconsistency between the T2 and the other two sequences. This inconsistency affects the evaluation of the combined segmentation of the MvMM.

The author hence further applied a GMM segmentation to the T2 images where the prior probabilities were computed from the resultant T2 segmentation of the MvMM. This segmentation, referred to as T2(+GMM), relies on the appearance of the T2 images, and hence the evaluation results should be expected to better than the MvMM. One can find from Table 2 that the myocardium segmentation by



TABLE 2  
The Dice, ACD and Hausdorff Distance of the Proposed MvMM from the Three CMR Sequences in Details;  
the Segmentation of T2 by GMM Based on the Results of MvMM Are Provided for Comparisons in T2(+GMM)

CMR	Dice			ACD (mm)		Hausdorff distance (mm)	
	Endocardium	Epicardium	Myocardium	Endocardium	Epicardium	Endocardium	Epicardium
LGE	$0.866 \pm 0.063$	$0.896 \pm 0.036$	$0.717 \pm 0.076$	$2.54 \pm 1.00$	$2.62 \pm 0.91$	$10.6 \pm 4.67$	$11.2 \pm 4.06$
T2	$0.794 \pm 0.124$	$0.908 \pm 0.043$	$0.717 \pm 0.129$	$3.75 \pm 2.18$	$2.46 \pm 1.27$	$11.9 \pm 5.90$	$9.94 \pm 5.94$
bSSFP	$0.903 \pm 0.048$	$0.917 \pm 0.027$	$0.764 \pm 0.064$	$2.06 \pm 0.96$	$2.16 \pm 0.81$	$9.23 \pm 5.06$	$10.7 \pm 4.56$
T2(+GMM)	$0.827 \pm 0.094$	$0.878 \pm 0.046$	$0.744 \pm 0.094$	$2.88 \pm 1.69$	$2.46 \pm 1.13$	$10.6 \pm 5.99$	$12.1 \pm 5.47$

TABLE 3  
The Dice, ACD and Hausdorff Distance of the Inter-Observer (IOb) and Inter-Sequence (ISq) Variations

Study: CMR	Dice			ACD (mm)		Hausdorff distance (mm)	
	Endocardium	Epicardium	Myocardium	Endocardium	Epicardium	Endocardium	Epicardium
IOb: LGE	$0.876 \pm 0.069$	$0.903 \pm 0.041$	$0.757 \pm 0.083$	$2.31 \pm 1.19$	$2.38 \pm 0.95$	$10.6 \pm 4.65$	$12.5 \pm 5.38$
IOb: T2	$0.894 \pm 0.052$	$0.906 \pm 0.056$	$0.824 \pm 0.069$	$2.07 \pm 1.36$	$2.57 \pm 1.75$	$8.97 \pm 6.66$	$11.0 \pm 7.06$
IOb: bSSFP	$0.919 \pm 0.029$	$0.927 \pm 0.028$	$0.812 \pm 0.046$	$1.61 \pm 0.60$	$1.88 \pm 0.68$	$9.64 \pm 4.81$	$12.3 \pm 10.6$
ISq: T2-LGE	$0.825 \pm 0.100$	$0.898 \pm 0.050$	$0.734 \pm 0.115$	$3.51 \pm 1.88$	$3.33 \pm 1.69$	$11.7 \pm 6.98$	$11.7 \pm 6.57$
ISq: T2-bSSFP	$0.778 \pm 0.115$	$0.910 \pm 0.040$	$0.681 \pm 0.106$	$4.46 \pm 2.05$	$2.86 \pm 1.23$	$12.6 \pm 4.54$	$10.1 \pm 4.45$
ISq: LGE-bSSFP	$0.819 \pm 0.084$	$0.875 \pm 0.055$	$0.660 \pm 0.103$	$4.11 \pm 1.72$	$3.70 \pm 1.45$	$12.9 \pm 4.68$	$11.7 \pm 4.67$

T2(+GMM) has indeed become better, mainly thanks to the improvement of the endocardium segmentation, but the epicardium segmentation is worse (in Dice and Hausdorff distance) due to the indistinctness of boundaries and lack of assistance from the other two sequences.

Finally, the mean runtime of the MvMM for registering and segmenting the three CMR sequences is  $9.17 \pm 3.95$  min.

### 3.1.5 Variation Study: Inter-Observer and Inter-Sequence

Table 3 gives the *inter-observer* variations from the three manual delineations and the *inter-sequence* variations among the three sequences. In computing the inter-sequence variations, the manual segmentation of the high-resolution images was transformed onto the image space with lower resolution, after a global affine registration to correct the global misalignment due to body motions, and the Dice scores and ACD were solely computed on the overlapped regions of the two CMR sequences. In this study, both the LGE and bSSFP CMR were transformed onto the T2 image space, and the LGE images were transformed onto the bSSFP image space.

The inter-sequence inconsistency is evident, mainly due to two reasons. One is that the intra-image motion shift between slices of the images were not corrected, though the global motion shift was compensated using affine registration. This inter-slice shift of two CMR sequences can be very different. For example, one of the first sequence can be shifted to the left side of the whole heart, while the corresponding slice in the other sequence can be shifted entirely to the opposite direction, i.e., the right side. The other reason is that there can exist large deformation of the heart between two scans, while in this study only affine transformations were assumed for correcting inter-sequence motions, leading to big difference of the segmentation results from two CMR sequences.

### 3.1.6 Scar Segmentation

In clinics, evaluation of the scars is one of the main goals. The quantification is commonly performed based on the myocardial segmentation, which can be done by an automatic method, such as the proposed MvMM, or by manual delineation, referred to as ManMyo. Based on the initial myocardial segmentation, one can classify the scars and non-scar tissues, using either the automatic approaches or the manual delineation of the scars (MDS). Here, the automatic Otsu threshold approach is studied [35]. This leads to four different schemes for comparisons, namely *MvMM+Otsu*, *ManMyo+Otsu*, *MvMM+MDS* and *ManMyo+MDS*.

The accuracies, assessed by the average Dice scores of the scar segmentation against the gold standard, are respectively  $0.3910 \pm 0.2287$  (*MvMM+Otsu*),  $0.4093 \pm 0.1790$  (*ManMyo+Otsu*),  $0.4457 \pm 0.2158$  (*MvMM+MDS*), and  $0.4779 \pm 0.1855$  (*ManMyo+MDS*). It should be pointed out that for either the automatic Otsu or the MDS, there was no statistically significant difference, in terms of Dice scores, between the two initialization methods. The p-values were respectively 0.5049 (*MvMM+Otsu* versus *ManMyo+Otsu*) and 0.4090 (*MvMM+MDS* versus *ManMyo+MDS*).

The gold standard scar segmentation was obtained by computing the shape mean of three or four manual labeling results [34]. The inter-observer variation of manual scar segmentation, in terms of Dice overlap, was  $0.5243 \pm 0.1578$ .

## 3.2 LGE CMR Segmentation Using Different Schemes

### 3.2.1 Comparisons with Conventional Methods

This section compares four conventional methods which perform the segmentation of each CMR sequence separately, including,

- 1) *Atlas*: The atlas-based segmentation using the initial result computed directly from the registered probabilistic atlases.

TABLE 4

This Table Provides the Dice Scores of the LGE CMR Segmentation by the Four Separate Segmentation Methods, Four Combined Segmentation Schemes, and the MvMM with Other Spatial Regularization (Spatial Regu) Approaches

Separate Seg	Atlas	GMM	Atlas+bSSFP	GMM+bSSFP
Endocardium	$0.7335 \pm 0.1661$	$0.7321 \pm 0.1549$	$0.8307 \pm 0.0740$	$0.8357 \pm 0.0713$
Epicardium	$0.8160 \pm 0.1149$	$0.7972 \pm 0.1103$	$0.8842 \pm 0.0422$	$0.8710 \pm 0.0557$
Myocardium	$0.4950 \pm 0.1959$	$0.5163 \pm 0.1989$	$0.6203 \pm 0.1117$	$0.6352 \pm 0.1202$
Combined Seg	Mvmm <sup>⊖</sup>	Mvmm <sup>⊖</sup> <sub>regAtlas</sub>	Mvmm <sup>⊖</sup> <sub>regSC</sub>	MvMM
Endocardium	$0.8552 \pm 0.0649$	$0.8569 \pm 0.0640$	$0.8654 \pm 0.0638$	$0.8657 \pm 0.0633$
Epicardium	$0.8919 \pm 0.0384$	$0.8912 \pm 0.0396$	$0.8958 \pm 0.0356$	$0.8958 \pm 0.0356$
Myocardium	$0.6971 \pm 0.0864$	$0.7008 \pm 0.0851$	$0.7168 \pm 0.0762$	$0.7169 \pm 0.0760$
Spatial Regu	MvMM <sub>a,MRF</sub>	MvMM <sub>no atlas</sub> <sub>MRF</sub>	MvMM <sub>no atlas</sub> <sub>MRF+SSM</sub>	MvMM <sub>+SSM</sub>
Endocardium	$0.8708 \pm 0.0725$	$0.4789 \pm 0.1305$	$0.7819 \pm 0.1257$	$0.8672 \pm 0.0498$
Epicardium	$0.9047 \pm 0.0403$	$0.5472 \pm 0.1095$	$0.8478 \pm 0.0575$	$0.9106 \pm 0.0278$
Myocardium	$0.7134 \pm 0.0938$	$0.4308 \pm 0.1610$	$0.5499 \pm 0.1290$	$0.7297 \pm 0.0705$

The MvMM in the table uses probabilistic atlases for spatial regularization. Please refer to the text for more details.

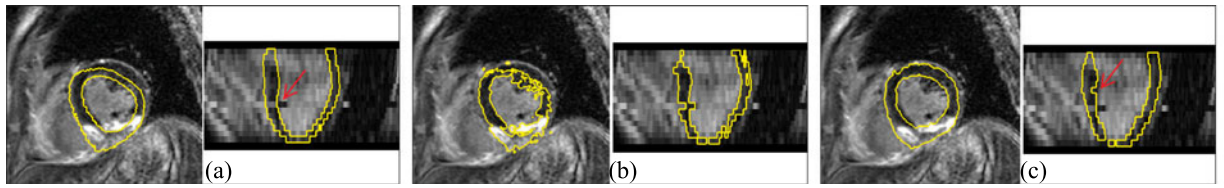


Fig. 7. The LGE CMR segmentation results using the atlas-based segmentation (a), GMM segmentation (b), and the Mvmm<sup>⊖</sup> without registration correction (c). The segmentation result using the proposed MvMM is presented in Fig. 5.

- 2) **GMM**: The conventional GMM segmentation based on the work of J. Ashburner and K. Friston [20], which is a univariate mixture model. This GMM segmentation is initialized by the propagated probabilities from the atlas. To avoid the blood pool of the atlas being registered to the scar tissue in the LGE CMR, the registration refinement is not adopted.
- 3) **Atlas+bSSFP**: The atlas-based segmentation where the bSSFP image of the same subject is used to assist the atlas-to-target registration;
- 4) **GMM+bSSFP**: The GMM segmentation initialized from the result of Atlas+bSSFP.

The Dice scores of the four separate segmentation for the LGE CMR sequence are presented in the sub-table, *Separate Seg*, of Table 4. Compared with the proposed MvMM which combines and segments the three sequences simultaneously, all the four separate segmentation schemes obtained evidently and significantly worse myocardium Dice scores ( $p < 0.01$ ). The advantage of including complimentary information from other sequences can be confirmed by comparing the results between the four separate segmentation methods: with the assistance of the bSSFP sequence, both *Atlas+bSSFP* and *GMM+bSSFP* achieved significantly better myocardium Dice scores,  $p < 0.01$ , than the conventional *Atlas* and *GMM* methods.

Fig. 7 displays the segmentation results using *Atlas* (a) and *GMM* (b). One can see that the registration could not correctly align the atlas and the target image in the motion shifted slice, as the red arrow point out in Fig. 7a. *GMM* segmentation could correct part of the misclassification, but the delineation was poor in the scar area, where no boundary constraint from other images was applied, as Fig. 7b shows. By contrast, the proposed MvMM produced a much better myocardial segmentation on the LGE CMR with the

guidance and constraints from the other two sequences, as Fig. 5c shows.

### 3.2.2 Comparisons Using Different Registration Schemes

This section studies four MvMM segmentation methods which combine the complimentary information from the three CMR sequences for simultaneous segmentation. These methods adopt the probabilistic atlases for spatial regularization, as described in Section 2.3.1, and the registration schemes are different, as follows,

- a) **Mvmm<sup>⊖</sup>**: the MvMM scheme without any registration correction;
- b) **Mvmm<sup>⊖</sup><sub>regAtlas</sub>**: the Mvmm<sup>⊖</sup> plus nonrigid registration for atlas;
- c) **Mvmm<sup>⊖</sup><sub>regSC</sub>**: the Mvmm<sup>⊖</sup> plus shift correction registration for the motion shifted slices;
- d) **MvMM**: the proposed MvMM with both atlas registration and shift correction.

The Dice scores of the LGE CMR sequence by the four methods are presented in the sub-table, *Combined Seg*, of Table 4. Here, the two types of registration correction are studied separately, i.e., (1) the nonrigid registration for atlas, (2) the rigid registration for inter-slice motion shift. One can find from the results that the registration improved the segmentation performance of Mvmm<sup>⊖</sup>. The improvement on the myocardium Dice scores was evident and statistically significant for Mvmm<sup>⊖</sup><sub>regSC</sub> (0.6971 versus 0.7168,  $p = 0.005$ ), and was significant but marginal for Mvmm<sup>⊖</sup><sub>regAtlas</sub> (0.6971 versus 0.7008,  $p = 0.008$ ). Also, the difference between the Dice scores of Mvmm<sup>⊖</sup><sub>regSC</sub> and MvMM was trivial and non-significant ( $p = 0.665$ ). This indicates that the nonrigid registration for atlas, which accounts for the major computation cost, did

not significantly improve the performance of the combined segmentation, as the initial atlas-to-target registration could provide good alignment between them [30]. By contrast, the shift correction had significant difference and should be the main concern of registration correction.

Fig. 7c provides an example to illustrate that  $Mvmm^\ominus$  could accurately include the scars into the myocardium, but still erroneously delineated the shifted slices, such as in the slice pointed out by the red arrow in (a). Fig. 7 also shows that  $Mvmm^\ominus$  misclassified one slice, pointed out by the red arrow in (c). This was due to the shifted slice in the bSSFP image, which is pointed out by the red arrow in Fig. 5a.

### 3.2.3 Study of Spatial Regularization

This study explores the performance of MvMM with different spatial regularization schemes. The MvMM with spatial regularization solely by the atlas is referred to MvMM in the previous study. MvMM is compared with another four methods, including

- i)  $MvMM_{a,MRF}$ : the MvMM with both atlas and MRF for spatial regularization;
- ii)  $MvMM_{MRF}^{no\ atlas}$ : the MvMM with MRF regularization but without atlas;
- iii)  $MvMM_{MRF+SSM}^{no\ atlas}$ : the  $MvMM_{MRF}^{no\ atlas}$  segmentation result is post-processed by projecting the segmentation onto a preconstructed statistical shape model (SSM) for shape regularization;
- iv)  $MvMM_{+SSM}$ : the MvMM segmentation result is post-processed by projecting the segmentation onto the SSM.

Here, a SSM has been built using a public data set of cardiac images from MMWHS Challenge 2017, which is available online.<sup>1</sup> For the detail of SSM construction and the application for shape regularization, readers are referred to the publications from literature [36], [37]

The Dice scores of the four different spatial regularization schemes are presented in the sub-table, *Spatial Regu*, of Table 4. Here, the author not only investigated the two types of conventional spatial constraints by an atlas and MRF, but also applied the SSM as a post-processing step for further shape regularization. One can see from the results that with the atlas, adding the MRF did not make significant difference in terms of the segmentation accuracy measured by the Dice score. For example, the p-value between the myocardial segmentation of  $MvMM_{a,MRF}$  and  $MvMM$  was  $p = 0.6388$ . The shape prior from a SSM could significantly improve the segmentation of  $MvMM_{MRF}^{no\ atlas}$ , which did not have any prior information for shape regularization ( $p < 0.01$  for all endocardium, epicardium and myocardium segmentation). However, when the atlas regularization was used  $MvMM_{+SSM}$  only marginally improved the myocardial segmentation, and the improvement was not statistically significant ( $p = 0.2708$ ).

## 3.3 Parameter Study of the Image Vector with Different Dimensions and Variant Coverage of ROI

### 3.3.1 Data and Experimental Setup

In order to perform the parameter studies, the author applied the MvMM to the task of brain MR segmentation [38]. The studies used the three simulated MR sequences from Brain

Web, including the T1-weighted, T2-weighted and proton density (PD) MR. The original images and corresponding segmentation were downloaded with image size  $181 \times 217 \times 181$  and voxel size  $1 \times 1 \times 1$  mm. The probabilistic atlases of the interested labels of the brain were generated by applying a Gaussian smoothing operation on the extracted labels, followed by a normalization process as the probabilities of labels on each pixel.

To simulate thirty different cases for each experiment, the author added simulated random motion artifacts and deformation fields to the original brain MR images. Each of the deformation fields, resulting in a misaligned image to the probabilistic atlases, was generated using a FFD transformation with  $20 \times 20 \times 20$  mm mesh spacing. The FFD transformation had each control point randomly displaced following a normal distribution of 0 mean and 2 mm standard deviation. For each case, the three MR sequences were deformed using three different transformations, simulating the mis-registration between the images of different sequences acquired from the same subject.

Four segmentation methods are compared here, including two well-developed algorithms, which are based on the univariate mixture model (UvMM), and two MvMM segmentation schemes:

- 1) *UvMM-JA*: This is the GMM segmentation proposed by J. Ashburner and K. Friston [20], which is a UvMM. Since the simulated data were free from intensity non-uniformity, the related function and parameters were not used.
- 2) *UvMM-KVL*: This is the conventional GMM segmentation proposed by Leemput et al. [7], which is a UvMM. The author implemented the same MRF strategy for spatial regularization and the same parameter values for MRF matrix reported in [7].
- 3) *MvMM-bivariate*: This is MvMM using two images where the image vector is a bivariate vector.
- 4) *MvMM-trivariate*: This is MvMM using three images where the image vector is a trivariate variable.

### 3.3.2 Image Vector with Different Dimensions

One can see from Fig. 8 that the segmentation performance, indicated by the Dice scores, can be improved with respect to the increased dimensions of the variate, i.e., the number of images to form the image vector. In general, the improvement from the univariate to bivariate is more evident than that from the bivariate to trivariate, indicating the performance tends to converge after certain number of images are combined for simultaneous segmentation.

### 3.3.3 Hetero-Coverage Multi-Modality Images: Variant Coverage of ROI

In this experiment, the author manually reduced the ROI of an image by deleting a number of slices to truncate the image volume, to simulate the HC-MMIs. Given an image, this truncation was done randomly at one end of the three dimensions, for example one image may have the first  $l$  slices at x-dimension truncated, while the other may have the end  $k$  slices at the z-dimension truncated. The sizes of the truncated regions are indicated using the sum of thickness of all the deleted slices.

Fig. 8 displays the Dice scores of the segmented white matter and gray matter using the three segmentation

<sup>1</sup> <http://www.sdspeople.fudan.edu.cn/zhuangxiahai/0/mmwhs>



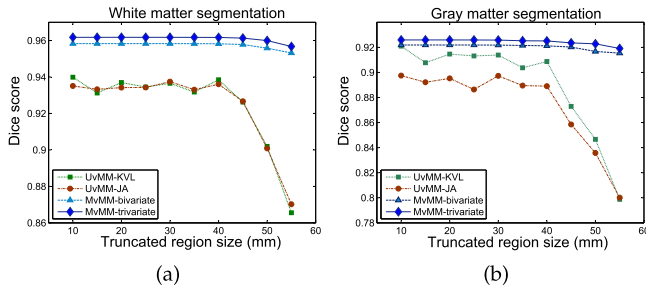


Fig. 8. The Dice scores of the white matter and gray matter segmentation combining multi-sequence brain MR. The UvMM-JA and UvMM-KVL are two well-developed algorithms for brain MR segmentation, respectively from [20] and [7].

schemes. The Dice scores of *MvMM-bivariate* and *MvMM-trivariate* are much higher than those of the two UvMM methods. Particularly, the two MvMM schemes still achieved good performance even when the images were randomly truncated more than 40 mm stack of slices, which compares differently with the UvMM approaches. This confirms that the combined segmentation can improve the performance when none of the multi-source images can provide the full information of the target subject.

#### 4 CONCLUSION AND DISCUSSION

The author has presented a new method, i.e., MvMM, for simultaneous segmentation of multi-source images, particularly for the myocardial segmentation combining the complementary information from multi-sequence CMR. In this method, the MS CMR of the same subject are aligned to a common space, which is virtual, and the segmentation of them is performed simultaneously in this virtual common space. To correct the motion shift of slices and the mis-registration of images, the MvMM and LL are formulated with transformations, and the ICM approach is employed to update the two groups of parameters. The model parameters (segmentation) are optimized using the EM algorithm and the transformation variables (registration) are updated using the gradient ascent method. To tackle the incongruent data, which are commonly seen in medical images, the MvMM framework is further generalized as the sum of a set of conventional MvMMs or UvMMs. Each of the conventional models is used to represent the congruent data in a sub-region of the virtual common space.

In the experiments, the author first investigated the performance of the proposed MvMM for myocardial segmentation of CMR sequences. The average Dice scores for LGE CMR segmentation were respectively 0.866 (endocardium), 0.896 (epicardium), and 0.717 (myocardium); and the ACD for the endocardium and epicardium were 2.54 mm and 2.62 mm, respectively. These results are promising, by comparing them with the corresponding values from the inter-observer variations, which were 0.876 (endocardium), 0.903 (epicardium) and 0.757 (myocardium) in Dice, and 2.31 mm (endocardium) and 2.38 mm (epicardium) in ACD. Generally, the bSSFP can obtain good accuracy thanks to the high quality of images, which also contributes to reducing the inconsistency between manual segmentation (inter-observer variation). The Dice scores and ACD of T2 CMR were not as good as those of bSSFP, mainly due to the appearance of thicker myocardium in T2. A simple follow-up GMM

segmentation based on the results of MvMM had shown to improve the accuracy of T2 CMR segmentation. Finally, the scar segmentation study showed that the automatic segmentation using MvMM could provide good initialization, with no significant difference compared with the manual delineation, for following scar classification.

To compare with the conventional methods, in Section 3.2.1 the author included the atlas-based segmentation and the conventional GMM method for separate segmentation on the LGE CMR. Both the Dice scores of these two methods were significantly worse than the proposed MvMM ( $p < 0.01$ ). The author further included the bSSFP CMR for the initial registration of the *Atlas* and *GMM* segmentation. The usage of bSSFP significantly improved the Dice scores of the segmentation ( $p < 0.01$ ). However, the separate segmentation, either the *Atlas*+bSSFP or *GMM*+bSSFP, was still significantly worse than MvMM using combined segmentation ( $p < 0.01$ ).

In Section 3.2.2, the author investigated the MvMM with different registration schemes, including MvMM without registration, MvMM with nonrigid atlas-to-target registration, and MvMM with rigid registration for correcting motion shift. The registration in each step could improve the segmentation performance, but the improvement from the atlas-to-target registration was not as evident as the correction of motion shift.

Section 3.2.3 compares the performance of MvMM with different spatial regularization algorithms. The regularization from the atlas has been shown to be effective and efficient. The MvMM with both the MRF and atlas performed well, but did not show significant difference compared with MvMM. However, when the constraint from the probabilistic atlases was not included, *MvMM<sub>MRF</sub><sup>noatlas</sup>* performed poorly. Finally, adding the shape regularization from a SSM, as a post-processing step, could improve the segmentation.

Section 3.3 studies the gain of segmentation performance by using an image vector with higher dimensions and investigates the performance of MvMM in HC-MMIs. The two MvMM schemes of combined segmentation were evidently better than the two conventional methods which are both of UvMM segmentation. The performance of *MvMM-bivariate* was comparable to *MvMM-trivariate*, indicating a convergence of segmentation performance by the high-dimensional MvMM after using two-source images. It is worth mentioning that the HC-MMI experiment showed that the combined MvMM segmentation was much more robust than the separate UvMM segmentation in the applications when some images had less coverage of the ROI.

In the literature, there have been few works focusing on the fully automatic LGE CMR segmentation, which is however an important prerequisite in clinical applications of cardiology. Two conference papers [11], [13] have reported about 2.25 mm ACD (42 slices) and 2.1 mm ACD (27 subjects) for LGE CMR myocardial segmentation, respectively. The authors in [15] implemented *manual interaction* to correct the mis-registration when the automatic translational registration failed on certain slices, and the average Dice score was 0.82 on 21 subjects. In these three works, the segmentation of LGE CMR was propagated from the segmentation of the corresponding bSSFP image. The bSSFP segmentation provides strong prior for the segmentation of LGE CMR. However, the application of the shape from

bSSFP is not fully automatic, nor integrated within a unified framework for simultaneous and combined segmentation. In the preliminary conference paper [39], the author tested the MvMM on 18 selected data sets, which had good image quality and small inter-sequence variation. The average Dice score of myocardial segmentation could reach up to 0.74. It should be noted that an objective inter-study comparison can be difficult, since the data sets and implementation can vary greatly across different studies.

Deep learning (DL) techniques have been recently intensively investigated for medical image analysis. A survey of this topic can be found in [40]. Particularly, DL-based segmentation methods can achieve good performance in the applications, where the targets have well-defined boundaries and a large number of annotated training data are available. It is worth mentioning that there exist six public CMR data sets [41], [42], [43], [44], [45], [46]. However, none of them includes the LGE CMR for studies of myocardial pathologies, though all of them have the bSSFP CMR images. Also, the proposed MvMM method aims at simultaneously registration and segmentation combining multi-source images, and no work has been reported on tackling this task using DL techniques, to the best of our knowledge.

To investigate the potential of DL techniques for automatic segmentation of LGE CMR, we implemented two state-of-the-art methods, i.e., U-Net [47] and Deeplab [48]. The evaluation was performed using the 3-fold cross validation scheme, and the author augmented the training data by flipping and rotating each of the 2D slices, resulting in over 2.7 thousand training samples. The average Dice scores were respectively  $0.734 \pm 0.196$  (U-Net) and  $0.495 \pm 0.166$  (Deeplab). The U-Net method was originally designed for biomedical image segmentation and had been successfully applied to the myocardial segmentation of bSSFP CMR images, while Deeplab was not specifically designed for medical image segmentation and did not perform well in this task. U-Net obtained a similar average Dice score as the proposed  $MvMM_{+SSM}$ . However, one could see that the standard deviation was more than twice that of  $MvMM_{+SSM}$ , indicating the unstable performance of the U-Net method. The author visually assessed each of the results, and found that U-Net could indeed achieve good segmentation in most of the cases, thanks to the regular shape of the myocardium and the augmented training set. Particularly, it could accurately delineate the structures when the boundaries were distinct, such as the example shown in Fig. 9a. However, when it came to the pathological myocardium (scars), U-Net could be challenged. Particularly, it could generate results with unrealistic shapes, such as the examples shown in Fig. 9b, 9c, 9d.

DL techniques have great potentials in tackling the challenging segmentation tasks. However, for LGE CMR segmentation it is important to develop mechanisms for integrating prior shape knowledge, to maintain a realistic shape of the segmentation result. Alternatively, one can develop a scheme to register and incorporate the clear boundary information from other-modality images. For DL, the simultaneous registration and segmentation combining multi-source images is still an open question. This is considered as the future work of this study.

There are three limitations of this work. First, the adoption of T2 CMR may be a sub-optimal choice, as T2 CMR generally

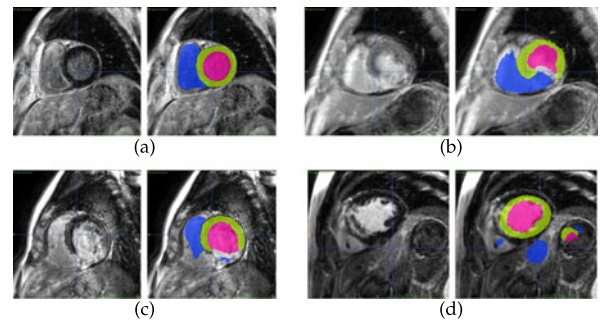


Fig. 9. Illustration of the segmentation results from U-Net: (a) is a typical good result, where the myocardium has clear boundaries; (b) is a typical result where the scars induce misclassification and a resulting unrealistic myocardium shape; (c) is a typical result where the healthy myocardium is accurately delineated, but the scars and blood pools confuse the deep learning method; (d) shows a result visually assessed as a failure, where other organs are misclassified as ventricular labels.

only has a few slices, such as three in many of our clinical data; and it images myocardium which appears to be thicker than the myocardium in the LGE and bSSFP CMR. Second, there can exist nonrigid misalignment between different sequences, for example due to the respiratory motions during acquisitions. Therefore, efficient models for such deformations and effective algorithms to correct them should be considered in the future work. Finally, notably the myocardial segmentation task from MS CMR is a selected example for demonstrating the application of the proposed MvMM, which nevertheless is generally applicable to the segmentation and classification tasks where multi-source images are available and can be combined. For example, in Section 3.3 the MvMM was applied to the segmentation of MS brain MR. Therefore, the future work should include extension of the proposed MvMM and the multi-modality image analysis for other clinical applications.

## ACKNOWLEDGMENTS

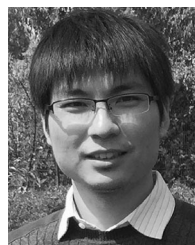
This work was supported by the National Natural Science Foundation of China (61971142), and the Science and Technology Commission of Shanghai Municipality (17JC1401600). The author would like to thank Chen Liu and Fuping Wu for their helps and discussions of the deep learning algorithms.

## REFERENCES

- [1] D. L. Pham, C. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annu. Rev. Biomed. Eng.*, vol. 2, pp. 315–337, 2000.
- [2] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imaging*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.
- [3] C. Petitjean and J. N. Dacher, "A review of segmentation methods in short axis cardiac MR images," *Med. Image Anal.*, vol. 15, no. 2, pp. 169–184, Apr. 2011.
- [4] H. W. Kim, A. Farzaneh-Far, and R. J. Kim, "Cardiovascular magnetic resonance in patients with myocardial infarction: current and emerging applications," *J. Amer. College Cardiology*, vol. 55, no. 1, pp. 1–16, 2009.
- [5] A. Kolipaka, G. P. Chatzimavroudis, R. D. White, T. P. O'Donnell, and R. M. Setser, "Segmentation of non-viable myocardium in delayed enhancement magnetic resonance images," *Int. J. Cardiovascular Imaging*, vol. 21, no. 2/3, pp. 303–311, 2005.
- [6] A. S. Flett, J. Hasleton, C. Cook, D. Hausenloy, G. Quarta, C. Ariti, V. Muthurangu, and J. C. Moon, "Evaluation of techniques for the quantification of myocardial scar of differing etiology using cardiac magnetic resonance," *JACC: Cardiovascular Imaging*, vol. 4, no. 2, pp. 150–156, 2011.



- [7] K. V. Leemput, F. Maes, D. Vandermeulen, and P. Suetens, "Automated model-based tissue classification of MR images of the brain," *IEEE Trans. Med. Imaging*, vol. 18, no. 10, pp. 897–908, Oct. 1999.
- [8] M. Lorenzo-Valdes, G. I. Sanchez-Ortiz, A. G. Elkington, R. Mohiaddin, and D. Rueckert, "Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm," *Med. Image Anal.*, vol. 8, pp. 255–265, 2004.
- [9] M. A. Balafar, "Gaussian mixture model based segmentation methods for brain MRI images," *Artif. Intell. Rev.*, vol. 41, no. 3, pp. 429–439, 2014.
- [10] R. Berbari, N. Kachenoura, F. Frouin, A. Herment, E. Mousseaux, and I. Bloch, "An automated quantification of the transmural myocardial infarct extent using cardiac DE-MR images," in *Proc. IEEE Conf. Eng. Med. Biol. Soc.*, vol. 1, 2009, pp. 4403–4406.
- [11] E. Dikici, T. O'Donnell, R. Setser, and R. White, "Quantification of delayed enhancement MR images," in *Proc. Int. Conf. Medical Image Comput. Comput.-Assisted Intervention*, 2004, vol. 3216, pp. 250–257.
- [12] R. S. Xu, P. Athavale, Y. Lu, P. Radau, and G. A. Wright, "Myocardial segmentation in late-enhancement MR images via registration and propagation of cine contours," in *Proc. 10th Int. Symp. Biomed. Imaging*, 2013, pp. 856–859.
- [13] C. Ciofolo, M. Fradkin, B. Mory, G. Hautvast, and M. Breeuwer, "Automatic myocardium segmentation in late-enhancement MRI," in *Proc. 5th IEEE Int. Symp. Biomed. Imaging: From Nano Macro*, 2008, pp. 225–228.
- [14] M. Rajchl, J. Yuan, J. White, E. Ukwatta, J. Stirrat, C. Nambakhsh, F. Li, and T. Peters, "Interactive hierarchical-flow segmentation of scar tissue from late-enhancement cardiac MR images," *IEEE Trans. Med. Imaging*, vol. 33, pp. 159–172, 2014.
- [15] D. Wei, Y. Sun, S.-H. Ong, P. Chai, L. L. Teo, and A. Low, "Three-dimensional segmentation of the left ventricle in late gadolinium enhanced mr images of chronic infarction combining long- and short-axis information," *Medical Image Anal.*, vol. 17, pp. 685–697, 2013.
- [16] P. Geladi and H. F. Grahn, *Multivariate Image Anal.* Hoboken, NJ, USA: Wiley, 1996.
- [17] J. M. Prats-Montalbán, A. De Juan, and A. Ferrer, "Multivariate image analysis: A review with applications," *Chemometrics Intell. Laboratory Syst.*, vol. 107, no. 1, pp. 1–23, 2011.
- [18] M. Bharati and J. MacGregor, "Multivariate image analysis for real-time process monitoring and control," *Ind. Eng. Chemistry Res.*, vol. 37, no. 12, pp. 4715–4724, 1998.
- [19] W. Shi, X. Zhuang, H. Wang, S. Duckett, D. Oregan, P. Edwards, S. Ourselin, and D. Rueckert, "Automatic segmentation of different pathologies from cardiac cine MRI using registration and multiple component EM estimation," in *Proc. Int. Conf. Functional Imaging Model. Heart*, 2011, pp. 163–170.
- [20] J. Ashburner and K. J. Friston, "Unified segmentation," *NeuroImage*, vol. 26, no. 3, pp. 839–851, 2005.
- [21] X. Zhuang, K. Rhode, R. Razavi, D. J. Hawkes, and S. Ourselin, "A registration-based propagation framework for automatic whole heart segmentation of cardiac MRI," *IEEE Trans. Med. Imaging*, vol. 29, no. 9, pp. 1612–1625, Sep. 2010.
- [22] S. Z. Li, "Markov random field models in computer vision," in *Proc. Eur. Conf. Comput. Vis.*, 1994, pp. 361–370.
- [23] X. Descombes, J.-F. Mangin, E. Peckersky, and M. Sigelle, "Fine structure preserving markov model for image processing," in *Proc. 9th Scandinavian Conf. Image Anal.*, 1995, pp. 349C–356.
- [24] K. Held, E. R. Kops, B. J. Krause, W. M. Wells, R. Kikinis, and H. W. Muller-Gartner, "Markov random field segmentation of brain MR images," *IEEE Trans. Med. Imaging*, vol. 16, no. 6, pp. 878–886, Dec. 1997.
- [25] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm," *IEEE Trans. Med. Imaging*, vol. 20, no. 1, pp. 45–57, Jan. 2001.
- [26] Z. Liang and S. Wang, "An em approach to map solution of segmenting tissue mixtures: A numerical analysis," *IEEE Trans. Med. Imaging*, vol. 28, no. 2, pp. 297–310, Feb. 2009.
- [27] C. Nikou, A. C. Likas, and N. P. Galatsanos, "A Bayesian framework for image segmentation with spatially varying mixtures," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2278–2289, Sep. 2010.
- [28] X. Zhuang and J. Shen, "Multi-scale patch and multi-modality atlases for whole heart segmentation of MRI," *Med. Image Anal.*, vol. 31, pp. 77–87, 2016.
- [29] S. J. Lee, "Accelerated coordinate descent methods for Bayesian reconstruction using ordered subsets of projection data," in *Proc. SPIE 4121 Math. Model. Estimation Imaging*, vol. 4121, no. 9, pp. 170–181, 2000.
- [30] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Med. Imaging*, vol. 18, pp. 712–721, Aug. 1999.
- [31] X. Zhuang, S. Arridge, D. J. Hawkes, and S. Ourselin, "A nonrigid registration framework using spatially encoded mutual information and free-form deformations," *IEEE Trans. Med. Imaging*, vol. 30, no. 10, pp. 1819–1828, Oct. 2011.
- [32] R. K.-S. Kwan, A. C. Evans, and G. B. Pike, "MRI simulation-based evaluation of image-processing and classification methods," *IEEE Trans. Med. Imag.*, vol. 18, no. 11, pp. 1085–1097, 1999.
- [33] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.
- [34] T. Rohlfing and C. R. Maurer, "Shape-based averaging," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 153–161, Jan. 2007.
- [35] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.
- [36] C. Hoogendoorn, N. Duchateau, D. Sanchez-Quintana, T. Whitmarsh, F. M. Sukno, M. D. Craene, K. Lekadir, and A. F. Frangi, "A high-resolution atlas and statistical model of the human heart from multislice CT," *IEEE Trans. Med. Imaging*, vol. 32, no. 1, pp. 28–44, Jan. 2013.
- [37] D. Rueckert, A. F. Frangi, and J. A. Schnabel, "Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration," *IEEE Trans. Med. Imaging*, vol. 22, no. 8, pp. 1014–1025, Aug. 2003.
- [38] M. Balafar, A. Ramli, M. Saripan, and S. Mashohor, "Review of brain MRI image segmentation methods," *Artif. Intell. Rev.*, vol. 33, no. 3, pp. 261–274, 2010.
- [39] X. Zhuang, "Multivariate mixture model for cardiac segmentation from multi-sequence MRI," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2016, pp. 581–588.
- [40] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. van der Laak, B. van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.
- [41] C. Petitjean and J. N. Dacher, "Cardiac MR left ventricle segmentation challenge," (2009). [Online]. Available: [http://smial.sri.utoronto.ca/LV\\_Challenge/Home.html](http://smial.sri.utoronto.ca/LV_Challenge/Home.html)
- [42] A. Suinesiaputra, "Left ventricular segmentation challenge," (2011). [Online]. Available: [www.cardiacatlas.org/challenges/lv-segmentation-challenge](http://www.cardiacatlas.org/challenges/lv-segmentation-challenge)
- [43] C. Petitjean, "RV segmentation challenge in cardiac MRI," (2012). [Online]. Available: [www.litislab.fr/?projet=1rvsc](http://www.litislab.fr/?projet=1rvsc)
- [44] B. A. Hamilton, "Data science bowl cardiac challenge data," (2014). [Online]. Available: [www.kaggle.com/c/second-annual-data-science-bowl](http://www.kaggle.com/c/second-annual-data-science-bowl)
- [45] M. Pop, M. Sermesant, P. Jodoin, A. Lalande, X. Zhuang, G. Yang, A. Young, and O. Bernard, "Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges," LNCS, vol. 10663, Springer, 2018.
- [46] C. Sudlow, "The UK biobank imaging study: Introduction," (2018). [Online]. Available: <http://imaging.ukbiobank.ac.uk/>
- [47] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2016, pp. 234–241.
- [48] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2017.



**Xiahai Zhuang** achieved the bachelor's degree from the Department of Computer Science, Tianjin University, the MS degree in computer science from Shanghai Jiao Tong University, and the doctorate degree from University College London. He is an associate professor with the School of Data Science, Fudan University. His research interests include medical image analysis, image processing, and computer vision. His works have been nominated twice for the MICCAI Young Scientist Awards (2008, 2012).

► For more information on this or any other computing topic, please visit our Digital Library at [www.computer.org/publications/dlib](http://www.computer.org/publications/dlib).