

A Dynamical Appearance Model Based on Multiscale Sparse Representation: Segmentation of the Left Ventricle from 4D Echocardiography^{*}

Xiaojie Huang¹, Donald P. Dione⁴, Colin B. Compas²,
Xenophon Papademetris^{2,3}, Ben A. Lin⁴, Albert J. Sinusas^{3,4},
and James S. Duncan^{1,2,3}

Departments of

¹ Electrical Engineering

² Biomedical Engineering

³ Diagnostic Radiology

⁴ Internal Medicine,

Yale University, New Haven, CT, USA

xiaojie.huang@yale.edu

Abstract. The spatio-temporal coherence in data plays an important role in echocardiographic segmentation. While learning offline dynamical priors from databases has received considerable attention, these priors may not be suitable for post-infarct patients and children with congenital heart disease. This paper presents a dynamical appearance model (DAM) driven by individual inherent data coherence. It employs multiscale sparse representation of local appearance, learns online multiscale appearance dictionaries as the image sequence is segmented sequentially, and integrates a spectrum of complementary multiscale appearance information including intensity, multiscale local appearance, and dynamical shape predictions. It overcomes the limitations of database-driven statistical models and applies to a broader range of subjects. Results on 26 4D canine echocardiographic images acquired from both healthy and post-infarct subjects show that our method significantly improves segmentation accuracy and robustness compared to a conventional intensity model and our previous single-scale sparse representation method.

1 Introduction

Segmentation of the left ventricle from 4D echocardiography plays an essential role in quantitative cardiac functional analysis. Due to gross image inhomogeneities, artifacts, and poor contrast between regions of interest, robust and accurate automatic segmentation of the left ventricle, especially the epicardial border, is very challenging in echocardiography. The inherent spatio-temporal coherence of echocardiographic data provides important constraints that can be exploited to guide cardiac border estimation and has motivated a spatio-temporal view point of echocardiographic segmentation.

^{*} This work was supported by NIH RO1HL082640.

Following the seminal work of Cootes et al. [1] on statistical shape/appearance modeling, a number of spatio-temporal statistical models (e.g., [2–6]) have been proposed for learning dynamical priors offline from databases. While these models have advantages in different aspects, the problem of forming a database that can handle a wide range of normal and abnormal heart images is still open to our knowledge. The assumption that different subjects have similar shape or motion pattern or their clinical images have similar appearance may not hold for routine clinical images, especially for disease cases, due to natural subject-to-subject tissue property variations and operator-to-operator variation in acquisition [7]. For example, for post-infarct patients, the positions, sizes and shapes of infarcts and thereby the overall heart motion can be highly variable across the population. It is very hard to build a reliable database accounting for all these variations, while such individual uniqueness is essentially desired information in some important applications like motion-based functional analysis. In addition, the tremendous cost of building reliable databases compromises the attractiveness of the database-driven methods.

Exploiting individual data coherence through online learning overcomes these limitations. It is particularly attractive when a database is inapplicable, unavailable, or defective. To this end, a model is indispensable for reliably uncovering the inherent spatio-temporal structure of individual 4D data. Sparse representation is a powerful mathematical framework for studying high-dimensional data. We proposed a 2D single-scale sparse-representation-based segmentation method in [8]. It shows the feasibility of analyzing 2D+t echocardiographic images via sparse representation and online dictionary learning. However, it is difficult to directly apply this method to 4D data. An important limitation is that it utilizes only a single scale of appearance information and requires careful tuning of scale parameters. This compromises segmentation accuracy and robustness. This paper generalizes our previous work [8] and introduces a new 3D dynamical appearance model (DAM) that leverages a full spectrum of complementary multiscale appearance information including intensity, multiscale local appearance, and shape. It employs multiscale sparse representation of high-dimensional local appearance, encodes appearance patterns with multiscale appearance dictionaries, and dynamically updates the dictionaries as the frames are segmented sequentially. The online multiscale dictionary learning process is supervised in a boosting framework to seek optimal weighting of multiscale information and generate dictionaries that are both generative and discriminative. Sparse coding w.r.t. the predictive dictionaries produces a local appearance discriminant. We also include intensity and a dynamical shape prediction to complete the appearance spectrum that we incorporate into a MAP framework.

2 Methods

2.1 Multiscale Sparse Representation

Let Ω denote the 3D image domain. We describe the multiscale local appearance at a pixel $\mathbf{u} \in \Omega$ in frame I_t with a series of appearance vectors $\mathbf{y}_t^k(\mathbf{u}) \in \mathbb{R}^n$

at different appearance scales $k = 1, \dots, J$. $\mathbf{y}_t^k(\mathbf{u})$ is constructed by concatenating orderly the pixels within a block centered at \mathbf{u} . Complementary multiscale appearance information is extracted using a fixed block size at different levels of Gaussian pyramid. Modeled with sparse representation, an appearance vector $\mathbf{y} \in \mathbb{R}^n$ can be represented as a sparse linear combination of the atoms from an appearance dictionary $\mathbf{D} \in \mathbb{R}^{n \times K}$ which encodes the typical patterns of a corresponding appearance class. That is, $\mathbf{y} \approx \mathbf{D}\mathbf{x}$. Given \mathbf{y} , \mathbf{D} , and a sparsity factor T_0 , the sparse representation \mathbf{x} can be solved by sparse coding:

$$\min_{\mathbf{x}} \|\mathbf{y} - \mathbf{D}\mathbf{x}\|_2^2 \text{ s.t. } \|\mathbf{x}\|_0 \leq T_0. \quad (1)$$

A shape s_t in I_t is represented by a level set function $\Phi_t(\mathbf{u})$. We define $\Phi_t^+(\mathbf{u}) = \Phi_t(\mathbf{u}) + \psi_1$ and $\Phi_t^-(\mathbf{u}) = \Phi_t(\mathbf{u}) - \psi_2$. The regions of interest are two band regions $\Omega_t^1 = \{\mathbf{u} \in \Omega : \Phi_t^-(\mathbf{u}) < 0, \Phi_t(\mathbf{u}) \geq 0\}$, and $\Omega_t^2 = \{\mathbf{u} \in \Omega : \Phi_t^+(\mathbf{u}) > 0, \Phi_t(\mathbf{u}) < 0\}$. We define $\Omega_t^* = \{\mathbf{u} \in \Omega : \Phi_{t-1}^+(\mathbf{u}) + \zeta_1 \geq 0, \Phi_{t-1}^-(\mathbf{u}) - \zeta_2 \leq 0\}$. The constants are chosen such that $s_t \in \Omega_t^*$. Suppose $\{\mathbf{D}_t^1, \mathbf{D}_t^2\}_k$ are two dictionaries adapted to appearance classes Ω_t^1 and Ω_t^2 respectively at scale k . They exclusively span, in terms of sparse representation, the subspaces of the respective classes. Reconstruction residues are defined as

$$\{R_t^c(\mathbf{u})\}_k = \|\mathbf{y}_t^k(\mathbf{u}) - \{\mathbf{D}_t^c \hat{\mathbf{x}}_t^c(\mathbf{u})\}_k\|_2 \quad (2)$$

$\forall \mathbf{u} \in \Omega_t^*$, $k \in \{1, \dots, J\}$, and $c \in \{1, 2\}$, where $\hat{\mathbf{x}}_t^c$ is the sparse representation of \mathbf{y}_t^k w.r.t. \mathbf{D}_t^c . It is logical to expect that $\{R_t^1(\mathbf{u})\}_k > \{R_t^2(\mathbf{u})\}_k$ when $\mathbf{u} \in \Omega_t^2$, and $\{R_t^1(\mathbf{u})\}_k < \{R_t^2(\mathbf{u})\}_k$ when $\mathbf{u} \in \Omega_t^1$. Combining the multiscale information, we introduce a local appearance discriminant

$$R_t(\mathbf{u}) = \mathbf{1}_{\Omega_t^*}(\mathbf{u}) \sum_{k=1}^J (\log \frac{1}{\beta_k}) \text{sgn}(\{R_t^2(\mathbf{u})\}_k - \{R_t^1(\mathbf{u})\}_k), \quad (3)$$

$\forall \mathbf{u} \in \Omega$, where β_k 's are the weighting parameters of the J appearance scales.

2.2 Online Multiscale Dictionary Learning

To obtain the discriminant R_t , $\{\mathbf{D}_t^1, \mathbf{D}_t^2\}_k$ and β_k need to be learned. Leveraging the inherent spatio-temporal coherence of individual data, we introduce an online multiscale appearance dictionary learning process supervised in a boosting framework. We interlace the processes of dictionary learning and segmentation as illustrated in Fig. 1. Similar to the database-driven dynamical shape models [3–5], we also assume a segmented first frame for initialization. It can be achieved by some automatic method with expert correction or purely manual segmentation. We dynamically update the multiscale appearance dictionaries each time a new frame is segmented. For $t > 2$, $\{\mathbf{D}_t^1, \mathbf{D}_t^2\}_k$ are well initialized with $\{\mathbf{D}_{t-1}^1, \mathbf{D}_{t-1}^2\}_k$ and updated with only a few iterations. $\{\mathbf{D}_2^1, \mathbf{D}_2^2\}_k$ are initialized with training samples. To reduce propagation error, we divide a sequence into two subsequences to perform bidirectional segmentation like [8]. The proposed dictionary

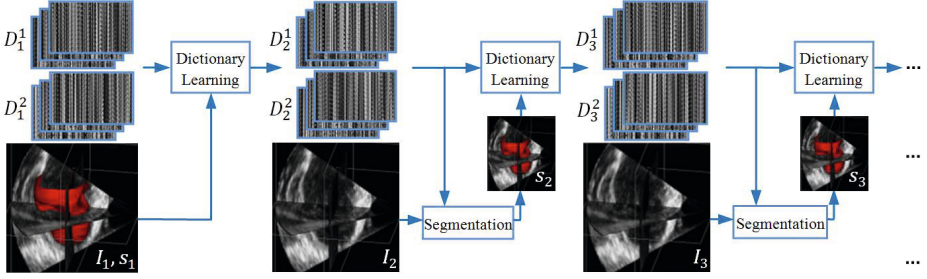


Fig. 1. Dynamical dictionary update interlaced with sequential segmentation

learning algorithm following the structure of the AdaBoost [9] is detailed in Algorithm 1. J dictionary pairs $\{\mathbf{D}_t^1, \mathbf{D}_t^2\}_k$ and weighting parameters β_k are learned from two classes of appearance samples $\{\mathbf{Y}_{t-1}^1\}_k = \{\mathbf{y}_{t-1}^k(\mathbf{u}) : \mathbf{u} \in \Omega_{t-1}^1\}$ and $\{\mathbf{Y}_{t-1}^2\}_k = \{\mathbf{y}_{t-1}^k(\mathbf{u}) : \mathbf{u} \in \Omega_{t-1}^2\}$, $k = 1, \dots, J$. The K-SVD [10] algorithm is invoked to enforce the reconstructive property of the dictionaries. The boosting supervision strengthens the discriminative property of the dictionaries and optimizes the weighting of multiscale information.

2.3 MAP Estimation

We estimate the shape Φ_t in frame I_t given the knowledge of $\hat{\Phi}_{1:t-1}$ and $I_{1:t}$. Different from the single-scale method in [8], we integrate a spectrum of complementary multiscale appearance information including intensity, the multiscale local appearance discriminant, and a dynamical shape prediction Φ_t^* . Since Φ_{t-1} and Φ_{t-2} are both spatially and temporally close, we assume a constant evolution speed during $[t-2, t]$. Within the band domain $\Omega_t^1 \cup \Omega_t^2$ we introduce an approximate shape prediction $\Phi_t^* = \hat{\Phi}_{t-1} + G(\hat{\Phi}_{t-1} - \hat{\Phi}_{t-2})$ to regularize the shape estimation. Here $G(*)$ denotes Gaussian smoothing operation used to preserve the smoothness of level set function. The segmentation is estimated by maximizing the posterior probability:

$$\begin{aligned} \hat{\Phi}_t &= \arg \max_{\Phi_t} p(\hat{\Phi}_{1:t-1}, I_{1:t-1}, I_t | \Phi_t) p(\Phi_t) \\ &\approx \arg \max_{\Phi_t} p(\Phi_t^*, R_t, I_t | \Phi_t) p(\Phi_t) \\ &\approx \arg \max_{\Phi_t} p(\Phi_t^* | \Phi_t) p(R_t | \Phi_t) p(I_t | \Phi_t) p(\Phi_t). \end{aligned} \quad (4)$$

The shape regularization is given by $p(\Phi_t^* | \Phi_t) p(\Phi_t) \propto \exp\{-\gamma \int_{\Omega_t^1 \cup \Omega_t^2} (\Phi_t - \Phi_t^*)^2 d\mathbf{u}\} \exp\{-\mu \int_{\Omega} \delta(\Phi_t) |\nabla \Phi_t| d\mathbf{u}\}$. We assume i.i.d. normal distribution of R_t : $p(R_t | \Phi_t) \propto \prod_{\mathbf{u} \in \Omega_t^1} \exp\{-\frac{[R_t(\mathbf{u}) - c_1]^2}{2\omega_1^2}\} \prod_{\mathbf{u} \in \Omega_t^2} \exp\{-\frac{[R_t(\mathbf{u}) - c_2]^2}{2\omega_2^2}\}$, and i.i.d. Rayleigh density of I_t : $p(I_t | \Phi_t) = \prod_{\mathbf{u} \in \Omega_t^1} \frac{I_t(\mathbf{u})}{\sigma_1^2} \exp\{-\frac{I_t(\mathbf{u})^2}{2\sigma_1^2}\} \prod_{\mathbf{u} \in \Omega_t^2} \frac{I_t(\mathbf{u})}{\sigma_2^2} \exp\{-\frac{I_t(\mathbf{u})^2}{2\sigma_2^2}\}$. Since intensity is not helpful for epicardial discrimination, $p(I_t | \Phi_t)$ is dropped in the epicardial case. The overall segmentation energy functional is given by:

$$\begin{aligned} E(\Theta, \Phi_t) &= \int_{\Omega_t^1} I_t^2 / 2\sigma_1^2 + \log(\sigma_1^2 / I_t) d\mathbf{u} + \int_{\Omega_t^2} I_t^2 / 2\sigma_2^2 + \log(\sigma_2^2 / I_t) d\mathbf{u} \\ &+ \int_{\Omega_t^1} (R_t - c_1)^2 / 2\omega_1^2 d\mathbf{u} + \int_{\Omega_t^2} (R_t - c_2)^2 / 2\omega_2^2 d\mathbf{u} \\ &+ \gamma \int_{\Omega_t^1 \cup \Omega_t^2} (\Phi_t - \Phi_t^*)^2 d\mathbf{u} + \mu \int_{\Omega} \delta(\Phi_t) |\nabla \Phi_t| d\mathbf{u}, \end{aligned} \quad (5)$$

where $\Theta = [c_1, c_2, \omega_1, \omega_2, \sigma_1, \sigma_2]$. We minimize the energy functional as follows: (a) Initialize Φ_t^0 with Φ_{t-1} , $\tau = 0$; (b) Compute the maximum likelihood estimate of $\Theta(\Phi_t^\tau)$; (c) Update $\Phi_t^{\tau+1}$ by gradient descent; (d) Reinitialize $\Phi_t^{\tau+1}$ after every few iterations; (e) Stop if $\|\Phi_t^{\tau+1} - \Phi_t^\tau\|_2 < \xi$, otherwise, $\tau = \tau + 1$, go to (b).

Algorithm 1. Multiscale Appearance Dictionary Learning

Input: appearance samples $\{\mathbf{Y}_{t-1}^1\}_k = \{\mathbf{y}_{1,i}^k\}_{i=1}^{M_1}$ and $\{\mathbf{Y}_{t-1}^2\}_k = \{\mathbf{y}_{2,j}^k\}_{j=1}^{M_2}$, initial dictionaries $\{\mathbf{D}_{t-1}^1, \mathbf{D}_{t-1}^2\}_k$, $k = 1, \dots, J$, $\mathbf{w}_1^1 = \{w_{1,i}^1\}_{i=1}^{M_1} = \mathbf{1}$, $\mathbf{w}_2^1 = \{w_{2,j}^1\}_{j=1}^{M_2} = \mathbf{1}$.

Output: dictionary pairs $\{\mathbf{D}_t^1, \mathbf{D}_t^2\}_k$, weighting parameters β_k , $k = 1, \dots, J$.

For $k = 1, \dots, J$:

- **Resampling:** Draw sample sets $\tilde{\mathbf{Y}}_1^k$ from $\{\mathbf{Y}_{t-1}^1\}_k$ and $\tilde{\mathbf{Y}}_2^k$ from $\{\mathbf{Y}_{t-1}^2\}_k$ based on distributions $\mathbf{p}_1^k = \{p_{1,i}^k\}_{i=1}^{M_1} = \frac{\mathbf{w}_1^k}{\sum_{i=1}^{M_1} w_{1,i}^k}$ and $\mathbf{p}_2^k = \{p_{2,j}^k\}_{j=1}^{M_2} = \frac{\mathbf{w}_2^k}{\sum_{j=1}^{M_2} w_{2,j}^k}$.
- **Dictionary Learning:** Apply the K-SVD to learn $\{\mathbf{D}_t^1, \mathbf{D}_t^2\}_k$ from $\tilde{\mathbf{Y}}_1^k$ and $\tilde{\mathbf{Y}}_2^k$:

$$\min_{\mathbf{D}_t^c, \mathbf{X}} \|\tilde{\mathbf{Y}}_c^k - \mathbf{D}_t^c \mathbf{X}\|_2^2 \quad \text{s.t.} \quad \forall i, \|\mathbf{x}_i\|_0 \leq T_0; \quad c \in \{1, 2\}. \quad (6)$$

- **Sparse Coding:** $\forall \mathbf{y} \in \{\mathbf{Y}_{t-1}^1, \mathbf{Y}_{t-1}^2\}_k$, solve the sparse representations w.r.t. $\{\mathbf{D}_t^1\}_k$, and $\{\mathbf{D}_t^2\}_k$ using the OMP [11], and get residues $R(\mathbf{y}, \mathbf{D}_t^1)_k$ and $R(\mathbf{y}, \mathbf{D}_t^2)_k$.
 - **Classification:** Make a hypothesis $h_k : \mathbf{y} \in \{\mathbf{Y}_{t-1}^1, \mathbf{Y}_{t-1}^2\}_k \rightarrow \{0, 1\}$: $h_k(\mathbf{y}) = \text{Heaviside}(R(\mathbf{y}, \mathbf{D}_t^2)_k - R(\mathbf{y}, \mathbf{D}_t^1)_k)$. Calculate the error of h_k : $\epsilon_k = \sum_{i=1}^{M_1} p_{1,i}^k |h_k(\mathbf{y}_{1,i}^k) - 1| + \sum_{j=1}^{M_2} p_{2,j}^k h_k(\mathbf{y}_{2,j}^k)$. Set $\beta_k = \epsilon_k / (1 - \epsilon_k)$.
 - **Weight Update:** $w_{1,i}^{k+1} = w_{1,i}^k \beta_k^{1 - |h_k(\mathbf{y}_{1,i}^k) - 1|}$, $w_{2,j}^{k+1} = w_{2,j}^k \beta_k^{1 - h_k(\mathbf{y}_{2,j}^k)}$.
-

3 Experiments and Results

We acquired 26 3D canine echocardiographic sequences from both healthy and post-infarct subjects using a Phillips iE33 ultrasound imaging system with a frame rate of ~ 40 Hz. Each sequence spanned a cardiac cycle. The sequential segmentation was initialized with a manual tracing of the first frame. Both endocardial and epicardial borders were segmented throughout the sequences. Fig. 2 shows typical segmentation examples by our method. 100 frames were randomly drawn from ~ 700 frames for manual segmentation and quality assessment. We evaluated automatic results against expert manual tracings using the following segmentation quality metrics: Hausdorff Distance (HD), Mean Absolute Distance (MAD), Dice coefficient (DICE), and Percentage of True Positives (PTP).

Benefit from the Dynamical Appearance Model. When the dynamical appearance components are turned off, our model reduces to a conventional ultrasound intensity model: the Rayleigh model [12]. Comparison with the Rayleigh method clearly shows the added value of the proposed DAM. Since the Rayleigh method is generally sensitive to initial contours, we initialized its segmentation of each frame with the first frame manual tracing. Fig. 3(a) compares typical

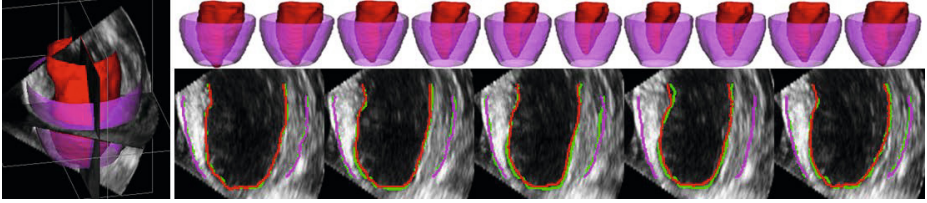


Fig. 2. Typical segmentations by our method (red, purple) and manual tracings (green)

segmentation examples by the Rayleigh method and our method. We observed that the Rayleigh method was easily trapped by misleading intensity information (e.g., image inhomogeneities and artifacts), while our approach produced accurate segmentations. Fig. 2 qualitatively shows the capability of the DAM in estimating reliably 3D left ventricular borders throughout the whole cardiac cycle. The Rayleigh method did not generate acceptable segmentation sequences in the experiment. Table 1 demonstrates that the DAM significantly outperformed the Rayleigh model. Better means (higher DICE and PTP, and lower MAD and HD) and lower standard deviations show the remarkable improvement of segmentation accuracy and robustness achieved by employing the DAM.

Table 1. Sample means and standard deviations of segmentation quality measures

expressed as mean \pm std		DICE (%)	PTP (%)	MAD (mm)	HD (mm)
Endocardial	Rayleigh [12]	74.9 \pm 18.8	83.1 \pm 16.3	2.01 \pm 1.22	9.17 \pm 3.37
	DAM	93.6 \pm 2.49	94.9 \pm 2.34	0.57 \pm 0.14	2.95 \pm 0.62
	SSDM [5]	—	95.9 \pm 1.24	1.41 \pm 0.40	2.53 \pm 0.75
Epicardial	Rayleigh [12]	74.1 \pm 17.4	82.5 \pm 12.0	2.80 \pm 1.55	16.9 \pm 9.30
	DAM	97.1 \pm 0.93	97.6 \pm 0.86	0.60 \pm 0.19	3.03 \pm 0.76
	SSDM [5]	—	94.5 \pm 1.74	1.74 \pm 0.39	2.79 \pm 0.97

Advantages over Single-scale Sparse Representation. We compared our DAM to the single-scale sparse representation model (SSR) in [8]. The SSR was extended to 3D and performed at 5 appearance scales ranging from low scale $3.5 \times 3.5 \times 3.5 \text{mm}^3$ to high scale $15.5 \times 15.5 \times 15.5 \text{mm}^3$, while the DAM utilized

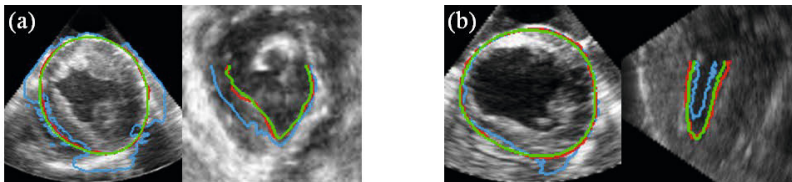


Fig. 3. Segmentation examples. (a) Manual (Green), DAM (Red), Rayleigh (Blue). (b) Manual (Green), DAM (Red), SSR (Blue).

multiscale appearance information. Fig. 3(b) presents end-systolic segmentation examples showing that the use of DAM resulted in lower propagation error and higher segmentation accuracy compared to the SSR. Fig. 4 presents the quantitative results of the comparison study. We observed that the performance of the SSR varied with the scale, which implies its sensitivity to the appearance scale. The SSR required scale tuning to get better results. The DAM achieved the best results in almost all the metrics for both endocardial and epicardial segmentations, which demonstrates the advantages of DAM over SSR. By summarizing complementary multiscale appearance information, the DAM consistently produced accurate segmentations without careful parameter tuning.

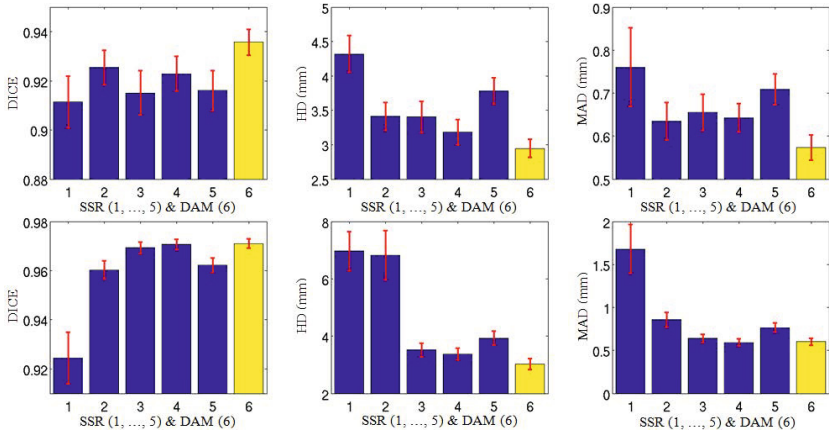


Fig. 4. Means and 95% confidence intervals obtained by the SSR (blue, scales 1, ..., 5) and the DAM (yellow, 6) in endocardial (top row) and epicardial (bottom row) cases.

Comparison with Database-Driven Dynamical Models. Table 1 compares the HD, MAD and PTP achieved by our model and that by a state-of-the-art database-driven dynamical shape model SSDM reported in [5]. The database-free DAM achieved comparable results with the SSDM, and outperformed the SSDM in segmenting epicardial borders. It is worth noticing that the DAM does not require more human interaction at the segmentation stage than the database-driven dynamical models such as [3–5] which also need manual tracings of the first or first few frames for initialization.

4 Discussion and Conclusion

We have proposed a 3D dynamical appearance model that exploits the inherent spatio-temporal coherence of individual echocardiographic data. It employs multiscale sparse representation, online multiscale appearance dictionary learning, and a spectrum of complementary multiscale appearance information including intensity, multiscale local appearance, and shape. Our method resulted in

significantly improved accuracy and robustness of left ventricular segmentation compared to a standard intensity method and our previous single-scale sparse representation method. The DAM achieved comparable results with a state-of-the-art database-driven statistical dynamical model SSDM. Since the DAM is database-free, it overcomes the limitations introduced by the use of databases. The DAM can be applied to the cases (e.g., the post-infarct subjects in this study) where it is inappropriate to apply database-based a priori motion or shape knowledge. Even when the priors are effective, the DAM can be a good choice for complementing the database and relaxing the reliance of statistical models (e.g., [2–6]) on database quality. Future work includes extensions to human data, other modalities, and an integrated online and offline learning framework to exploit their complementarity.

References

1. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. *IEEE TPAMI* 23(6), 681–685 (2001)
2. Bosch, J.G., Mitchell, S.C., Lelieveldt, B.P.F., Nijland, F., Kamp, O., Sonka, M., Reiber, J.H.C.: Automatic segmentation of echocardiographic sequences by active appearance motion models. *IEEE TMI* 21(11), 1374–1383 (2002)
3. Jacob, G., Noble, J.A., Behrenbruch, C.P., Kelion, A.D., Banning, A.P.: A shape-space based approach to tracking myocardial borders and quantifying regional left ventricular function applied in echocardiography. *IEEE TMI* 21(3), 226–238 (2002)
4. Sun, W., Çetin, M., Chan, R., Reddy, V., Holmvang, G., Chandar, V., Willsky, A.S.: Segmenting and Tracking the Left Ventricle by Learning the Dynamics in Cardiac Images. In: Christensen, G.E., Sonka, M. (eds.) *IPMI 2005*. LNCS, vol. 3565, pp. 553–565. Springer, Heidelberg (2005)
5. Zhu, Y., Papademetris, X., Sinusas, A.J., Duncan, J.S.: A Dynamical Shape Prior for LV Segmentation from RT3D Echocardiography. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) *MICCAI 2009, Part I*. LNCS, vol. 5761, pp. 206–213. Springer, Heidelberg (2009)
6. Yang, L., Georgescu, B., Zheng, Y., Meer, P., Comaniciu, D.: 3D ultrasound tracking of the left ventricle using one-step forward prediction and data fusion of collaborative trackers. In: *CVPR* (2008)
7. Noble, J.A., Boukerroui, D.: Ultrasound image segmentation: a survey. *IEEE TMI* 25(8), 987–1010 (2006)
8. Huang, X., Lin, B.A., Compas, C.B., Sinusas, A.J., Staib, L.H., Duncan, J.S.: Segmentation of left ventricles from echocardiographic sequences via sparse appearance representation. In: *MMBIA*, pp. 305–312 (January 2012)
9. Freund, Y., Schapire, R.: A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting. In: Vitányi, P.M.B. (ed.) *EuroCOLT 1995*. LNCS, vol. 904, pp. 23–37. Springer, Heidelberg (1995)
10. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE TSP* 54(11), 4311–4322 (2006)
11. Tropp, J.: Greed is good: algorithmic results for sparse approximation. *IEEE TIT* 50(10), 2231–2242 (2004)
12. Sarti, A., Corsi, C., Mazzini, E., Lamberti, C.: Maximum likelihood segmentation of ultrasound images with rayleigh distribution. *IEEE TUFFC* 52(6), 947–960 (2005)