Direct and Simultaneous Four-Chamber Volume Estimation by Multi-Output Regression

Xiantong Zhen[1], Ali Islam[3], Mousumi Bhaduri[4], Ian Chan[4], and Shuo Li[1,2]

[1] The University of Western Ontario, London, ON, Canada
[2] GE Healthcare, London, ON, Canada
[3] St. Joseph's Health Care, London, ON, Canada
[4] London Health Sciences Centre, ON, Canada

Abstract. Cardiac four-chamber volumes provide crucial information for quantitative analysis of whole heart functions. Conventional cardiac volume estimation relies on a segmentation step; recently emerging direct estimation without segmentation has shown better performance than than segmentation-based methods. However, due to the high complexity, four-chamber volume estimation poses great challenges to these existing methods: four-chamber segmentation is not feasible due to intensity homogeneity of ventricle and atrium without implicit boundaries between them; existing direct methods which can only handle single or bi-ventricles are not directly applicable due to great combinatorial variability of four chambers. In this paper, by leveraging the full strength of direct estimation, we propose a new method for direct and simultaneous four-chamber volume estimation using multi-output regression that can disentangle complex relationship of image appearance and four-chamber volumes via statistical learning. To accomplish accurate and efficient estimation, we propose using a supervised descriptor learning (SDL) algorithm to generate a compact and discriminative feature representation. By casting into generalized low-rank approximations of matrices with a supervised manifold regularization, the SDL jointly removes irrelevant and redundant information by feature reduction and extracts discriminative features directly related to four chambers via supervised learning, which overcomes the high complexity of four chambers. We evaluate the proposed method on a cardiac four-chamber MR dataset from 125 subjects including both healthy and diseased cases. The experimental results show that our method achieves a high correlation coefficient of up to 91.5% with manual segmentation obtained by human experts. Our method for the first time achieves simultaneous and direct four-chamber volume estimation, which enables more efficient and accurate functional assessment of the whole heart.

1 Introduction

Cardiac four-chamber volumes offer comprehensive measurement for heart functional assessment by capturing the dynamic pattern of the whole heart [1]. The left/right ventricles (LV/RV), which have been extensively studied, play a critical role in heart disease diagnosis, while the left/right atrium (LA/RA) volumes

N. Navab et al. (Eds.): MICCAI 2015, Part I, LNCS 9349, pp. 669–676, 2015.
DOI: 10.1007/978-3-319-24553-9_82

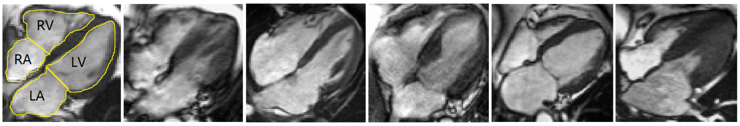

Fig. 1. Cardiac four-chamber MR images from different subjects with different temporal frames

are strongly associated with heart functions and indicate severity of diastolic dysfunctions and cardiovascular disease burden; together, four chambers provide crucial information for quantitative functional analysis of the whole heart. However, the LA and RA have long been overlooked due to the difficulty in measuring their volumes. Efficient simultaneous four-chamber volume estimation would enable more accurate and comprehensive cardiac functional analysis.

Four-chamber volume estimation poses great challenges to existing methods due to the huge complexity of four chambers stemming from highly complex contours, temporal deformations, anatomical interdependency of chambers, low tissue contrast and large patient variability as shown in Fig. 1. Conventional segmentation-based methods for cardiac volume estimation mainly focus on the LV with few on the RV. LA and RA volume estimation has not yet been addressed, not to mention simultaneous four-chamber volume estimation. Although automatic segmentation becomes more reliable, accurate and less time-consuming, four-chamber segmentation is still a challenging task and far from being used in clinical practice. Whole heart segmentation [1] potentially offers a solution to simultaneous four-chamber volume estimation; however, it is currently unable to segment four chambers separately to obtain their individual volumes due to the fact that ventricle and atrium are of intensity homogeneity with vague boundary between them, and two atriums are mostly connected with a very thin wall as shown in Fig. 1.

Recently, direct estimation [2–4] without segmentation has emerged as an effective tool for cardiac ventricular volume estimation [5, 6] which outperforms segmentation-based estimation in terms of both accuracy and efficiency [6]. More importantly, direct estimation allows us to leverage fast evolving state-of-the-art machine learning techniques which makes automatic detection and diagnosis comparable to a well-trained and experienced radiologist [7]. Although direct estimation has gained great success in single and bi-ventricular volume estimation [2–4], existing direct methods are not directly applicable to simultaneous four-chamber volume estimation due to the great combinatorial variability of four chambers and even more complicated relationship between image appearance and four-chamber volumes compared to single or bi-ventricles.

In this paper, we formulate four-chamber volume estimation as a multi-output regression problem. This formulation naturally models four-chamber volumes simultaneously, successfully handles the great challenge of four chambers, and provides clinical more meaningful volume estimation of four chambers. By removing unreliable segmentation, our method enables accurate and convenient functional analysis of the whole heart. To establish compact and discriminative feature representation for accurate and efficient volume estimation, we propose using a supervised descriptor learning (SDL) algorithm [8] formulated as generalized low-rank approximations of matrices with a supervised manifold regularization (SMR). The SDL jointly removes irrelevant and redundant information by feature reduction, i.e., generalized low-rank approximation and extracts discriminative features directly related to four-chamber volumes via supervised learning, i.e., supervised manifold regularization. The obtained cardiac four-chamber image representations by SDL are compact and discriminative, which enables efficient and accurate four-chamber volume estimation.

This work contributes in three folds: **1)** Our method is the first to achieve direct and simultaneous cardiac four-chamber volume estimation, which removes unreliable segmentation and enables more accurate and convenient whole heart functional analysis. The method can be conveniently extended to other clinical direct organ volume estimation; **2)** We formulate four-chamber volume estimation as a multi-output regression problem, which leverages the strength of statistical learning to achieve simultaneous four-chamber volume estimation. Other similar clinical data prediction from medical images can be modeled and solved in the same way; **3)** We propose using a supervised descriptor learning (SDL) algorithm [8] to generate compact and discriminative cardiac image representations, which overcomes the huge complexity of four chambers. The SDL provides a general supervised descriptor learning framework that can be widely used in other clinical multivariate estimation tasks.

2 Cardiac Four-Chamber Volume Estimation via Multi-Output Regression

2.1 Cardiac Image Representations

We are given a set of annotated data $\{X_1, \dots, X_L\}$ and the corresponding multivariate targets $\{Y_1, \dots, Y_L\}$, where L is the number of training samples and $Y_i \in \mathbb{R}^d$ denote four-chamber volumes. We start with matrix representations of four chamber cardiac images, *i.e.*, $X_i \in \mathbb{R}^{M \times N}$, which could be any matrix representations, *e.g.*, raw pixel intensities. We use the gradient orientation matrix (GOM) which is constructed from pyramid histogram of gradients (PHOG) of images by stacking spatial cells in rows and orientation bins in columns. The GOM takes advantages of prior knowledge to capture characteristic spatial layout and local shape which are the key characteristics of four chambers. The GOM is fed into the proposed SDL to learn a compact and discriminative representation of four chambers.

Generalized Low-Rank Approximation. We propose using the generalized low-rank approximation of matrices due to its efficient computation of dimension reduction of matrices [9]. This is to find two transformations: $W \in \mathbb{R}^{M \times m}$ and $V \in \mathbb{R}^{N \times n}$ with $m \ll M$ and $n \ll N$, and L matrices $D_i \in \mathbb{R}^{m \times n}$ such that

WD_iV^{T} is an appropriate approximation of each X_i, $i = 1, \ldots, L$. We solve the following optimization problem of minimizing the reconstruction errors:

$$\mathop{\arg\min}_{\substack{W,V,D_1,\ldots,D_L\\ W^{\mathrm{T}}W=I_m,V^{\mathrm{T}}V=I_n}} \frac{1}{L}\sum_{i=1}^{L} \|X_i - WD_iV^{\mathrm{T}}\|_F^2 \tag{1}$$

where $\|\cdot\|_F$ is the Frobenius norm of a matrix, I_m is an identity matrix of size $m \times m$ and the constraints $W^{\mathrm{T}}W = I_m$ and $V^{\mathrm{T}}V = I_n$ ensure that W and V have orthogonal columns to avoid redundancy in the approximations.

From (1), we know that D_i is the low-rank approximation of X_i in terms of the transformations of W and V, and it is worth to mention that the matrices $D_1, \ldots, D_L$ are not required to be diagonal. It is also proven in [9] that given the W and V, for any i, D_i is uniquely determined by $D_i = W^{\mathrm{T}}X_iV$ which is the compact representation of X_i that will reduce regression complexity for efficient multivariate estimation. (1) only minimize the reconstruction error in the low-rank space leading to indiscriminate representations $\{D\}_{i=1}^{L}$.

Supervised Manifold Regularization (SMR). We impose discrimination on the low-rank representation $\{D_i\}_{i=1}^{L}$ by integrating the proposed SMR into (1). To this end, we first construct a weighted graph $G = (V, E)$ using the ϵ-neighborhood method [10], where V and E respectively represent L vertices and edges between vertices. The graph is built on the multivariate targets $(Y_1, \ldots, Y_L)$, *i.e.*, the four-chamber volume values, rather than on inputs in conventional manifold regularization [11], which naturally induces the supervision. We denote $S \in \mathbb{R}^{L\times L}$ as the symmetric similarity matrix with non-negative elements corresponding to the edge weight of the graph G, where each element S_{ij} is computed by a heat kernel with parameter σ: $S_{ij} = \exp\left(\frac{-\|Y_i - Y_j\|^2}{2\sigma^2}\right)$, $i, j = 1, \ldots, L$. We set the diagonal elements of S to be zeros, *i.e.*, $S_{ii} = 0$. In the low-rank space, we would like to minimize the following term

$$\sum_{i,j} \|D_i - D_j\|_F^2 S_{ij}. \tag{2}$$

Since the similarity matrix S characterizes the manifold structure of the multivariate target space, low-rank approximations $\{D_i\}_{i=1}^{L}$ preserve the intrinsic local geometrical structure of the target space and are therefore automatically aligned to their regression targets. The discrimination is then naturally injected into the low-rank representations $\{D_i\}_{i=1}^{L}$.

Feature Learning with SMR. By integrating the SMR term in (2) into (1), we obtain the compact objective function of generalized low-rank approximation of matrices with the supervised manifold regularization (SMR) as follows:

$$\mathop{\arg\min}_{\substack{W,V,D_1,\ldots,D_L\\ W^{\mathrm{T}}W=I_m,V^{\mathrm{T}}V=I_n}} \frac{1}{L}\sum_{i=1}^{L} \|X_i - WD_iV^{\mathrm{T}}\|_F^2 + \beta\sum_{i,j} \|D_i - D_j\|_F^2 S_{ij} \tag{3}$$

where $\beta \in (0, \infty)$ is a tuning parameter to balance the tradeoff between reconstruction errors and discrimination of the low-rank approximations, which also serves to keep the flexibility of the model.

In the objective function of (3), the first term guarantees the reconstruction fidelity in the low-rank approximation while the second SMR term introduces the discrimination to learned new representations. The objective function is solved by an iterative algorithm via alternate optimization: fixing W, solve V and fixing V, solve W.

2.2 Multi-Output Regression with Random Forests

Regression forests, an efficient way of mapping a complex input space to continuous output, started to attract interest in medical image analysis [4,12]. Due to the strong capability of naturally handling multivariate estimation, regression forests offer a best-suited tool for simultaneous cardiac four-chamber volume estimation. They can **1)** effectively handle the non-linear relationship between the cardiac image appearance and four-chamber volumes; **2)** naturally deal with multiple outputs, i.e., four-chamber volumes, by capturing the interdependency among them; and **3)** provide accurate and clinically more meaningful volume estimation without overfitting due to the nature of ensemble learning.

We adopt the adaptive K-cluster regression forests (AKRF) recently proposed in [13] for multivariate estimation. In the AKRF, a novel node splitting method formulated as a classification problem is proposed to replace simple thresholding in conventional regression forests [14], which allows each node to have more than two child nodes. This enhances the ability to handle the complex distributions of four-chamber volumes. It has been shown in [13] that the AKRF significantly outperforms other regressors, *e.g.*, support vector regression (SVR), conventional random forests and kernel partial least squares [13].

3 Experiments

3.1 Dataset and Implementation Details

The dataset contains four-chamber cardiac MR images from 125 subjects each of which has 25 frames across a temporal cardiac cycle. Images were acquired on a 1.5T scanner with fast imaging employing steady-state acquisition (FIESTA) image sequence mode, using these acquisition parameters: TR=35.5 ms, TE=1.2 ms and slice thickness=6mm. The performance of the proposed method is quantitatively evaluated by comparing with ground truth by manual segmentation. The correlation coefficient between the ground truth and the estimation is used as measurement to evaluate estimation performance as in [4,5], and higher correlation coefficient indicates better performance. The leave-one-subject-out cross validation is used for evaluation.

We estimate cavity areas of four chambers in MR images, and the volumes are computed by integrating cavity areas in the sagittal direction. Note that we use the normalized areas as the targets, *i.e.*, the number of pixels in a chamber divided by the total number of pixels of the images. A region of interest (ROI) is placed to enclose four chambers in an MR image according to the method in [3]. We use a three-level pyramid HOG (PHOG) obtaining a matrix of size 84×31 from an image of 64×64 pixels. To show the advantage of our SDL algorithm, we have also compared with popular descriptors, *e.g.*, GIST and histogram of LBP both of which are implemented with a similar spatial pyramid to the PHOG descriptor, and dimensionality reduction methods, *e.g.*, generalized principal component analysis (GPCA) [15] and principal component analysis (PCA). In the implementation of adaptive K-clustering random forests (AKRF) [13], we use 20 trees to construct the regression forests, which can keep low computational cost with satisfactory performance.

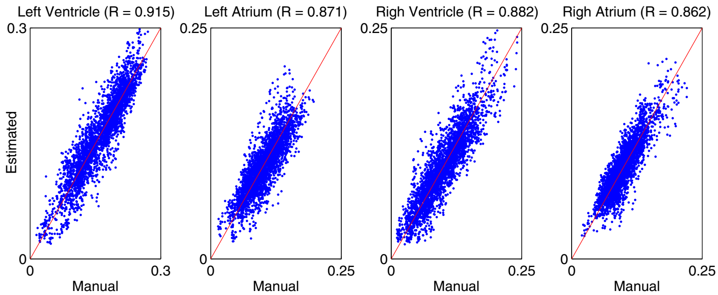

Fig. 2. The correlation coefficients between estimated and manually obtained volumes for four chambers. R is the correlation coefficient.

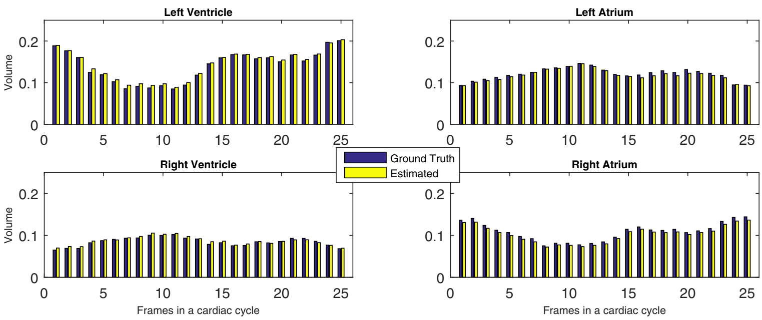

Fig. 3. The illustration of ground truth against estimation by the proposed method for 25 frames in cardiac cycle averaged over subjects

Table 1. The comparison results for cardiac four-chamber volume estimation

Methods	Left Ventricle	Left Atrium	Right Ventricle	Right Atrium
SDL	**0.915**	**0.871**	**0.882**	**0.862**
PHOG	0.869	0.819	0.832	0.811
GPCA	0.885	0.838	0.843	0.822
PCA	0.871	0.812	0.825	0.807
LBP	0.868	0.799	0.827	0.794
GIST	0.864	0.828	0.815	0.843

3.2 Simultaneous Four-Chamber Volume Estimation

The proposed method for the first time achieves simultaneous four-chamber volume estimation and produces high estimation accuracy for all the four chambers despite of the great challenge of four chambers, especially for the LV with a high correlation coefficient of 91.5% as illustrated in Fig. 2. Although the boundary between ventricle and atrium is mostly invisible and not supported by edge and region homogeneity, the proposed method can successfully predict four-chamber volumes due to the use of multi-output regression via statistically learning. Moreover, LA and RA volumes not measured previously due to their complex anatomical geometry are successfully predicted by our method with high accuracy. The results are clinically significant showing the great potential in clinical use [4, 6].

The average estimation results by our method against ground truth is shown in Fig. 3 for 25 frames (aligned across a cardiac cycle) over subjects. Our method can produce very close estimations with low errors for frames of all the four chambers to the ground truth manually obtained by human experts. The volume change pattern in a cardiac cycle is successfully captured by our method providing further information for cardiac pathologies, e.g., diastolic dysfunction, which indicates its practical use in clinical cardiac functional assessment and enables much wider clinical applications.

The strength of the proposed SDL algorithm for cardiac four-chamber image representation is also demonstrated by comparing with other methods as shown in Table 1. The SDL substantially outperforms both state-of-the-art descriptors, *e.g.*, LBP and GIST and dimensionality reduction techniques, *e.g.*, GPCA and PCA by up to 7.2% showing the effectiveness of the SDL for continuous multivariate estimation. The generalized low-rank approximation removes redundant information while the supervised manifold regularization extracts most discriminative features that are directly related to four-chamber volumes. By integrating them, we obtain compact and discriminative image representations for efficient and accurate cardiac four-chamber volume estimation.

4 Conclusion

In this paper, we proposed a new method for direct and simultaneous cardiac four-chamber volume estimation. Without depending on an intractable four-chamber segmentation step, we formulate four-chamber volume estimation as a multi-output regression problem. To achieve accurate and efficient estimation, we proposed using a supervised descriptor learning (SDL) algorithm to generate compact and discriminative cardiac image representation. Experimental results show that our method can produce highly accurate four-chamber volume estimation close to that obtained by human experts.

References

1. Zheng, Y., Barbu, A., Georgescu, B., Scheuering, M., Comaniciu, D.: Four-chamber heart modeling and automatic segmentation for 3-d cardiac ct volumes using marginal space learning and steerable features. IEEE TMI 27(11), 1668–1681 (2008)
2. Afshin, M., Ayed, I.B., Islam, A., Goela, A., Peters, T.M., Li, S.: Global assessment of cardiac function using image statistics in MRI. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part II. LNCS, vol. 7511, pp. 535–543. Springer, Heidelberg (2012)
3. Wang, Z., Ben Salah, M., Gu, B., Islam, A., Goela, A., Li, S.: Direct estimation of cardiac bi-ventricular volumes with an adapted bayesian formulation. IEEE TBME, 1251–1260 (2014)
4. Zhen, X., Wang, Z., Islam, A., Bhaduri, M., Chan, I., Li, S.: Direct estimation of cardiac bi-ventricular volumes with regression forests. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014, Part II. LNCS, vol. 8674, pp. 586–593. Springer, Heidelberg (2014)
5. Wang, Z., Salah, M., Ayed, I., Islam, A., Goela, A., Li, S.: Bi-ventricular volume estimation for cardiac functional assessment. In: RSNA (2013)
6. Zhen, X., Wang, Z., Islam, A., Chan, I., Li, S.: A comparative study of methods for cardiac ventricular volume estimation. In: RSNA (2014)
7. Wang, S., Summers, R.M.: Machine learning and radiology. Medical Image Analysis 16(5), 933–951 (2012)
8. Zhen, X., Wang, Z., Yu, M., Li, S.: Supervised descriptor learning for multi-output regression. In: CVPR (2015)
9. Ye, J.: Generalized low rank approximations of matrices. Machine Learning 61(1-3), 167–191 (2005)
10. He, X., Niyogi, P.: Locality preserving projections. In: NIPS, vol. 16, p. 153 (2004)
11. Belkin, M., Niyogi, P., Sindhwani, V.: Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. JMLR 7, 2399–2434 (2006)
12. Criminisi, A., Robertson, D., Konukoglu, E., Shotton, J., Pathak, S., White, S., Siddiqui, K.: Regression forests for efficient anatomy detection and localization in computed tomography scans. Medical Image Analysis 17(8), 1293–1303 (2013)
13. Hara, K., Chellappa, R.: Growing regression forests by classification: Applications to object pose estimation. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part II. LNCS, vol. 8690, pp. 552–567. Springer, Heidelberg (2014)
14. Breiman, L.: Random forests. Machine Learning 45(1), 5–32 (2001)
15. Ye, J., Janardan, R., Li, Q.: GPCA: an efficient dimension reduction scheme for image compression and retrieval. In: ACM SIGKDD, pp. 354–363 (2004)