

Machine Learning for Cardiac Ultrasound Time Series Data

Baichuan Yuan^a, Sathya R Chitturi^{b,a}, Geoffrey Iyer^a, Nuoyu Li^a, Xiaochuan Xu^{d,a}, Ruohan Zhan^{c,a}, Rafael Llerena^e, Jesse T. Yen^e, and Andrea L. Bertozzi^a

^aUniversity of California, Los Angeles, USA

^bPomona College, USA

^cPeking University, China

^dJilin University, China

^eUniversity of Southern California, USA

ABSTRACT

We consider the problem of identifying frames in a cardiac ultrasound video associated with left ventricular chamber end-systolic (ES, contraction) and end-diastolic (ED, expansion) phases of the cardiac cycle. Our procedure involves a simple application of non-negative matrix factorization (NMF) to a series of frames of a video from a single patient. Rank-2 NMF is performed to compute two end-members. The end members are shown to be close representations of the actual heart morphology at the end of each phase of the heart function. Moreover, the entire time series can be represented as a linear combination of these two end-member states thus providing a very low dimensional representation of the time dynamics of the heart. Unlike previous work, our methods do not require any electrocardiogram (ECG) information in order to select the end-diastolic frame. Results are presented for a data set of 99 patients including both healthy and diseased examples.

Keywords: Non-negative Matrix Factorization, ECG Independent, Ultrasound Video, Echocardiography

1. INTRODUCTION

Echocardiography is an inexpensive, non-invasive medical imaging technique routinely used to assess cardiovascular health. In general, current analysis of cardiac ultrasound videos is performed manually by highly trained specialists. Although this is the gold standard method for identifying heart abnormalities, cost, time and geographic constraints necessitate research into automatic detection of key markers that signal heart dysfunction.

Of central importance to the automation problem is the selection of frames corresponding to end-systole and end-diastole. Current clinical determination of these frames involves an experienced cardiologist manually observing a given video for changes in left ventricular volume, mitral valve position and ECG spikes.¹ Recently proposed automation techniques for end-systole and end-diastole detection fall into two classes: tracking mitral valve motion and manifold learning techniques.

Mitral valve methodologies are based on prior information regarding mitral valve motion during the cardiac cycle. One approach relies on rapid-mitral valve opening at the start of diastole.² The authors compute mean intensity variation in a region of interest that is selected using three landmarks; according to their analysis, the highest intensity measure is end-systole. Although the authors achieved reasonable results, the method still requires significant manual intervention, is sensitive to left ventricular longitudinal motion and is not robust to poorer quality images where the mitral valve is not clearly delineated.

Recently, various different manifold learning techniques have been used to automatically select end-systolic and end-diastolic frames. The analysis in Ref. 3 used a local linear embedding (LLE) algorithm to embed information in the high dimensional frame domain onto a 2-dimensional manifold. This work was further extended by applying isometric feature mapping together with LLE and ISOMAP.⁴ These techniques preserve local configurations such as the periodicity of the cardiac cycle and can help elucidate three key separate phases of the heartbeat. Although both methods yield very accurate results, the similarity method in Ref. 3 does not appear

Send correspondence to

Baichuan Yuan: byuan@math.ucla.edu.

to be robust to larger, more noisy datasets. Thus both works require some ECG information to find the start marker of a cardiac cycle, adding to both patient discomfort and cost.

In this paper we present a fully automated, ECG independent and ultimately simpler method to detect end-systole and end-diastole based on Rank-2-NMF.

2. MATERIAL AND METHOD

2.1 End-Systole and End-Diastole

The cardiac cycle is a medical term that describes the periodicity of the heart beat.⁵ One cardiac cycle is defined from the beginning of one heartbeat to the start of the next. Although there are seven well-characterized phases of the cardiac cycle, in this paper we consider only the two most biologically important phases, end-systole and end-diastole, which refer to the stages of the heart in maximum contraction and expansion respectively.

2.2 Non-negative Matrix Factorization

NMF is a powerful matrix factorization method that is widely used in computer vision to learn parts of objects⁶ and in machine learning for automated text sorting.⁷ In particular it is a dimensionality reduction technique that seeks to find lower paramertizations for high dimensionality data. In our analysis, Rank-2 NMF was implemented on each cardiac ultrasound video for the apical 4 view, the apical 2 view, the short axis view and the long axis view by using the following procedure.

First a video that contains N frames is concatenated into a data matrix \mathbf{X} by reshaping each frame into a column vector \mathbf{x}_i (1).

$$\mathbf{X} = [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \dots \quad \mathbf{x}_N], \quad \mathbf{x}_i \in \mathbb{R}^D. \quad (1)$$

The Rank-2 non-negative matrix factorization for \mathbf{X} is formulated as the following optimization problem (2)

$$\min_{\mathbf{W}, \mathbf{H}} \frac{1}{2} \|\mathbf{W}\mathbf{H}^T - \mathbf{X}\|_F^2, \quad \mathbf{W} \in \mathbb{R}_+^{D \times 2}, \mathbf{H} \in \mathbb{R}_+^{N \times 2}, \quad (2)$$

$$(3)$$

where \mathbf{W} and \mathbf{H} are non-negative matrices denoting two eigenimages (represented as two vectors) that approximate all frames by linear combination and the matrix of coefficients respectively.

To complete this minimization, \mathbf{W} and \mathbf{H} are iteratively updated based on the framework in Ref. 8:

$$\mathbf{W}^{(k+1)} := \arg \min \frac{1}{2} \|\mathbf{H}^{(k)} \mathbf{W}^T - \mathbf{X}^T\|_F^2, \quad \mathbf{W} \in \mathbb{R}_+^{D \times 2} \quad (4)$$

$$\mathbf{H}^{(k+1)} := \arg \min \frac{1}{2} \|\mathbf{W}^{(k+1)} \mathbf{H}^T - \mathbf{X}\|_F^2, \quad \mathbf{H} \in \mathbb{R}_+^{N \times 2} \quad (5)$$

We then apply the algorithm in Ref. 9. Both (4) and (5) are generalized as:

$$\min_{\mathbf{g}_i} \frac{1}{2} \sum_{i=1}^N \|\mathbf{B}\mathbf{g}_i - \mathbf{y}_i\|^2, \quad \mathbf{g}_i \in \mathbb{R}_+^2, \mathbf{y}_i \in \mathbb{R}_+^N \quad (6)$$

Obviously, (6) is separable with respect to all the \mathbf{g}_i . Then, for each \mathbf{g}_i we consider the possible optimization problems associated with its entries g_{i1} and g_{i2} :

$$\min_{g_{i1} \geq 0, g_{i2} \geq 0} \|\mathbf{b}_1 g_{i1} + \mathbf{b}_2 g_{i2} - \mathbf{y}_i\|^2 \quad (7)$$

$$\Rightarrow \begin{cases} \min_{g_{i1} \geq 0} \|\mathbf{b}_1 g_{i1} - \mathbf{y}_i\|^2 \\ \min_{g_{i2} \geq 0} \|\mathbf{b}_2 g_{i2} - \mathbf{y}_i\|^2 \end{cases} \quad (8)$$

where $[\mathbf{b}_1 \ \mathbf{b}_2] = \mathbf{B}$. Because $\mathbf{B} \in \mathbb{R}_+^{N \times 2}$, $\mathbf{g}_i \in \mathbb{R}_+^2$, $\mathbf{y}_i \in \mathbb{R}_+^N$, solutions to the decomposed (8) are always feasible, and can be used as a good approximation when (7) is not feasible. Following the procedure above we solve (4) and (5) by obtaining \mathbf{g}_i for each i and concatenating them.

After solving the problem, we obtain the matrix $\mathbf{W} = [\mathbf{w}_1 \ \mathbf{w}_2]$. The end-members \mathbf{w}_1 and \mathbf{w}_2 are then reshaped into images J_1 and J_2 . Coefficients from end-members were subsequently used in order to select frames in the original videos corresponding to end-systole and end-diastole.

2.3 Data acquisition and description of dataset

A total of 99 echocardiographic ultrasound samples were collected randomly from a database at the Keck Medical Center at the University of Southern California, and properly de-identified for the analysis. The ultrasound images were originally obtained from the GE Vivid E9 ultrasound system under the same settings prior to be transferred to the data base. The samples include a diverse group of abnormal conditions and a few normal studies to review the anatomical structures pertinent to each condition and be subject of the analysis. Each sample involved the four standard two-dimensional videos including the apical four-chamber view, the apical two-chamber view, the parasternal long-axis view and the parasternal short-axis view to evaluate all left ventricular wall segments. In addition, each video was composed of two-beat regular rhythm cine-loop including complete diastolic and systolic cycles to evaluate the end-diastolic phase and end-systolic phase respectively. No distinction was made between two-dimensional echocardiographic samples of high and poor image quality for the purpose of this study; all echocardiograms were originally collected for diagnostic purposes.

The echocardiographic images were obtained from pathologic conditions that have different effects in the anatomical structures to be analyzed. These conditions included ischemic cardiomyopathy, hypertrophic cardiomyopathy, heart failure, aortic valve stenosis, myocardial infarction, pulmonary hypertension, systemic hypertension, myocardial toxicity, tamponade and congenital heart disease.

3. RESULTS

3.1 Biologically Meaningful End-Members Generation

Rank-2 NMF is used to generate end-members in each video; these are biologically meaningful images corresponding to end-systole and end-diastole. This method works extremely well on the apical 4 view and NMF finds quite accurate end-members for almost all 99 cases. To compare, we manually select “ground truth” frames based on a coregistered ECG dataset using a standard method from the medical literature.⁴ We select end-diastole as the frame following mitral valve closure or the frame in the cardiac cycle in which the cardiac dimension is largest. End-systole is selected as the frame preceding the mitral valve opening or the time in the cardiac cycle in which the cardiac dimension is smallest in a normal heart.¹⁰ Results for the apical 4 view, the apical 2 view, the short axis view, and the long axis view are presented respectively in Figure 1, 2, 3 and 4. It’s clear that end-members generated by NMF are quite similar to the manually selected ground truth. Moreover, the NMF images actually appear smoother than the ground truth, with noise reduced to some extent, illustrating that the NMF end-members are good candidates for further analysis.

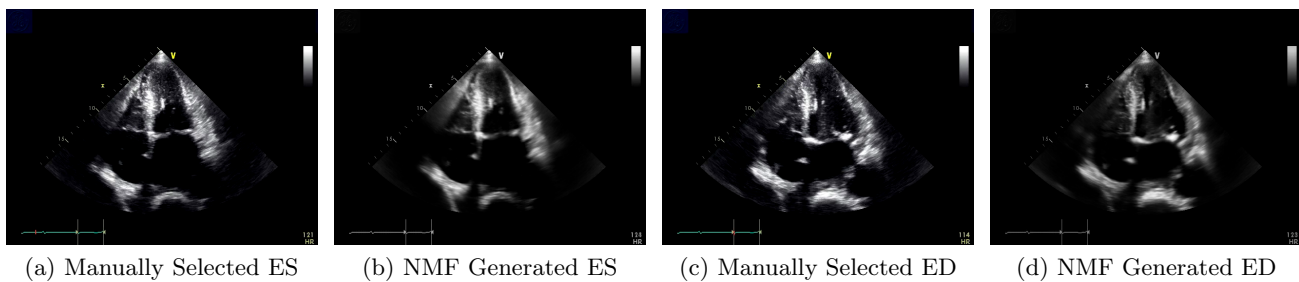


Figure 1: AP4 view: Comparison between manually selected end-members and NMF generated end-members.

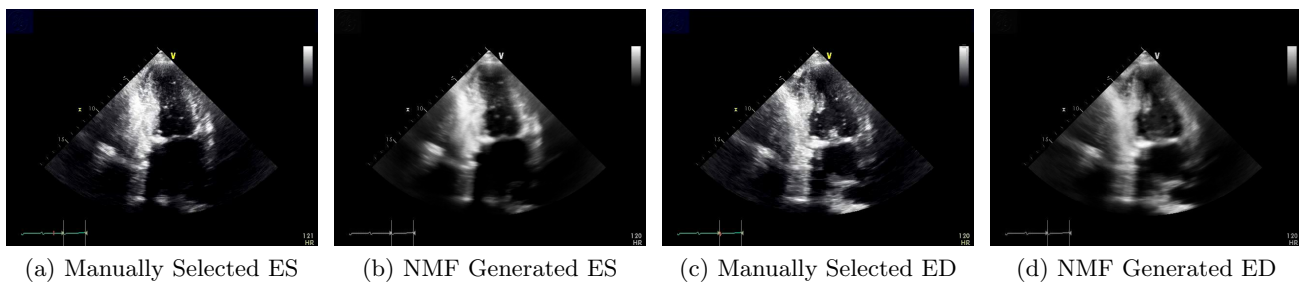


Figure 2: Apical 2 View: Comparison between manually selected end-members and NMF generated end-members.

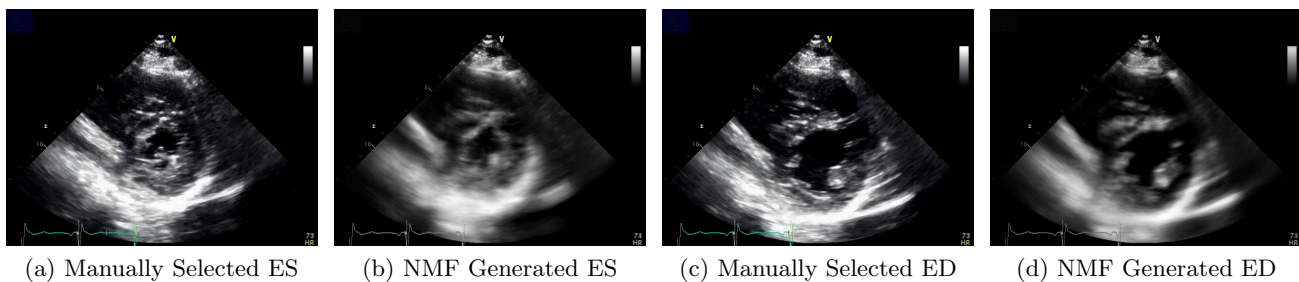


Figure 3: Short Axis View: Comparison between manually selected end-members and NMF generated end-members.

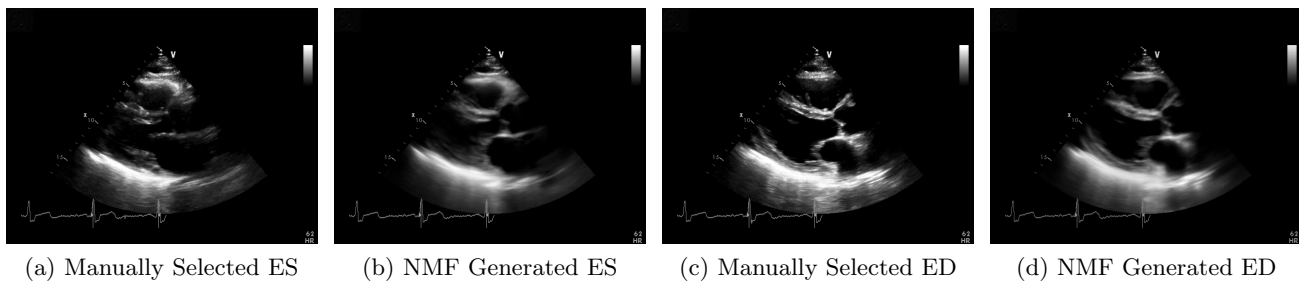


Figure 4: Long Axis View: Comparison between manually selected end-members and NMF generated end-members.

3.2 ES and ED Extraction based on NMF Coefficients

In this section we examine the two-dimensional projection of the echocardiogram onto the NMF end-members. Figure 5 shows an example for one patient with the NMF coefficients plotted against the frame number of the video (the time axis). One can see the periodic property of coefficient curves, consistent with the cyclic nature of heart movement. For each frame, we calculate the difference between the coefficients of the end-members. This produced a simple way to automate the identification of the frame number corresponding to end-systole and end-diastole. For comparison, the methods LLE and ISOMAP can also be used to select end-systole and end-diastole.^{3,4} Since the omit-third-phase method in Ref. 4 needs other priors or data to find the cycle start marker, it is not a good comparison for our method. Thus we focus on the similarity method³ to select end-systole and end-diastole, which identifies the pair with the smallest correlation on the manifold. Unlike previous work, we test LLE and ISOMAP on the whole image rather than a subregion of interest, in order to compare directly with our NMF-based method which likewise uses the entire image. All results are acquired from the apical 4 view. The performance is quantified by comparing the frame number difference between the extracted ED or ES with the ground truth. A comparison for the mean and variance of the error for these three methods is listed in the Table 1. We find that NMF has notably lower mean and variance, indicating its robustness and stability. That said, the LLE and ISOMAP methods were employed differently here by considering the entire image rather than subregions, and we have not studied the effect that a subregion vs. entire image has on those methods.

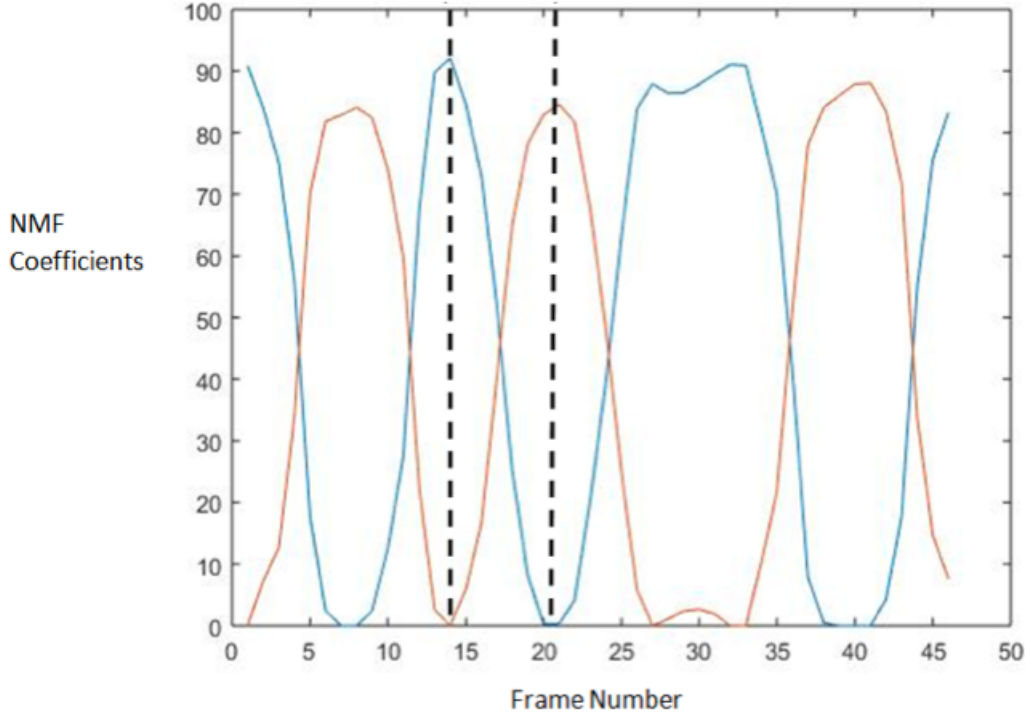


Figure 5: Rank-2-NMF end-member coefficients plotted against frame number for case 40. End-systole and end-diastole correspond to frames 14 and 21 respectively. Similar results are obtained for other cases in the data set.

Table 1: Comparison between NMF, LLE and ISOMAP results for all 99 cases in the apical 4 view.

	ES Difference ¹		ED Difference ²	
	mean	variance	mean	variance
NMF	0.93	1.72	0.93	1.17
LLE	2.76	5.92	2.22	8.79
ISOMAP	1.93	3.94	1.90	8.75

1. Difference between the frame number of extracted ES and Ground truth ES.
2. Difference between the frame number of extracted ED and Ground truth ED.

3.3 Cardiac Cycle Length and Heart Rate Estimation

End-systole and end-diastole NMF coefficients were used to estimate both the cardiac cycle length as well as the heart rate. The period of the NMF coefficient curve corresponded to the cardiac cycle length which, in combination with the frame rate, was used to determine the heart rate. The results from Rank-2-NMF were compared with those obtained from ISOMAP and LLE (Figure 6). In this study, only 65 cases from the data set were used since the cardiac cycle ground truth was only present for 65 cases. After eliminating 4 difficult cases, for which we could not find periods from the low-dimensional representation, the estimated heart rate was regressed on the ground truth and the results depicted in Figure 6. The number of cases with relatively reasonable results for ISOMAP, LLE and NMF is respectively 53, 57 and 61 out of the total 61 cases, which shows the robustness and universality of NMF methods. Furthermore, NMF’s results have a notably higher correlation with the ground truth with $R = 0.9196$, while the R of ISOMAP and LLE are 0.7404 and 0.8284, much smaller than that of NMF, indicating the effectiveness and accuracy of the NMF method.

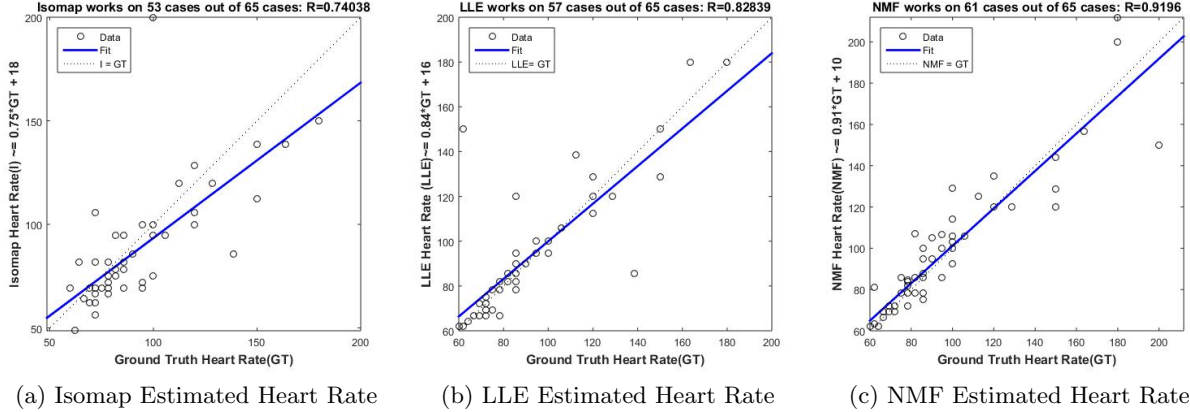


Figure 6: Regressions of Different Estimations on Ground-Truth

3.4 Video Decomposition

Recall that the ultrasound videos can be decomposed by NMF as follows:

$$W \times H + E = X,$$

where W , H , E and X respectively denote NMF generated end-members, corresponding coefficients, error term and original frames. Thus we can reconstruct the videos based on NMF end-members W and coefficients H .

Visualizations of the reconstructed NMF videos (the reshaped $W \times H$) and the difference videos ($E = X - W \times H$) are depicted in Figure 7, for a particular patient. It is clear that the reconstructed NMF videos capture wall motion and show the change of left ventricle volume, while the error term captures more information on muscle and valve movement, which leads to promising medical application.

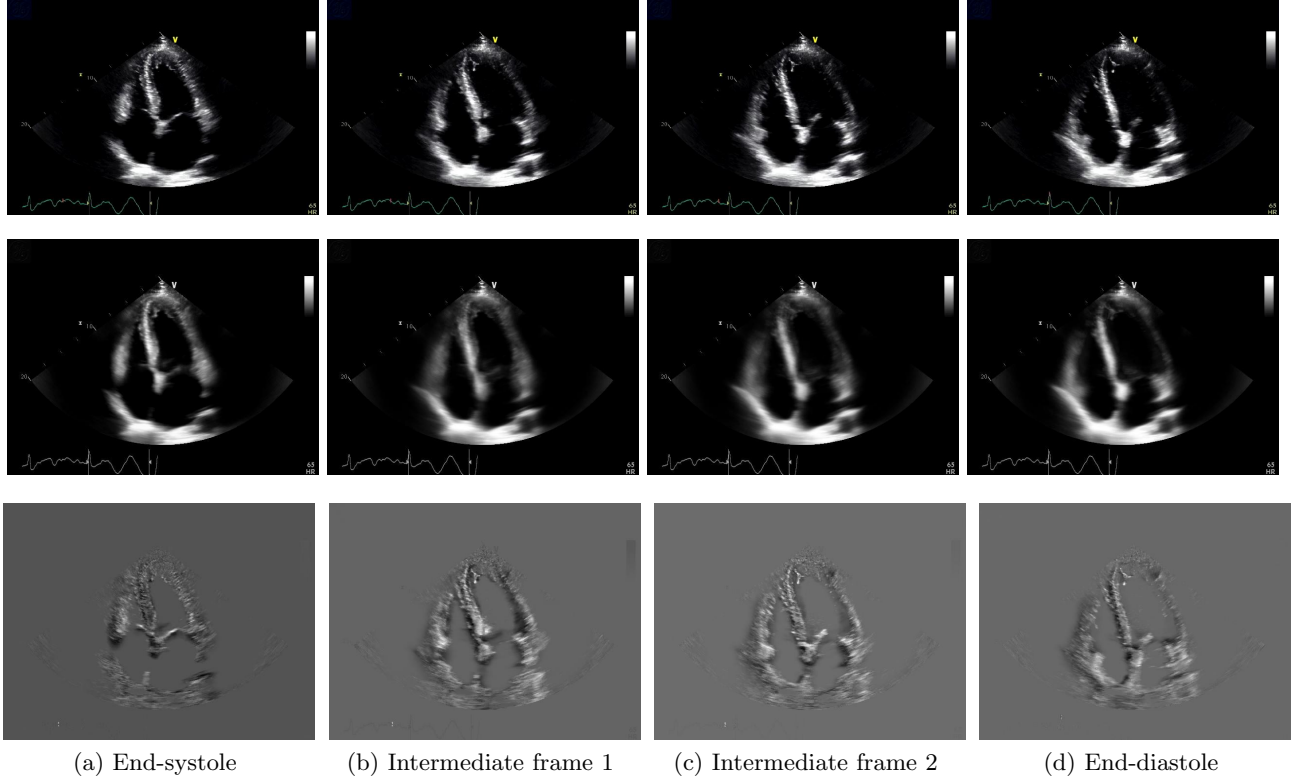


Figure 7: Frames from left to right show the heart between end-systole and end-diastole. Rows 1-3 show the original video, the NMF generated video ($W \times H$) and the difference video respectively.

4. DISCUSSION

In this paper, a popular machine learning method, Non-negative Matrix Factorization, has been applied to cardiac ultrasound videos, resulting in the observation that all frames in the video can be roughly written as a linear combination of end-systolic and end-diastolic frames. NMF tries to seek a linear lower dimensional representation, while the previous manifold learning methods LLE and ISOMAP attempt to achieve nonlinear dimension reduction. We find that NMF is more robust to noise, a major problem for ultrasound videos. While manifold learning techniques require the selection of a nearest neighbor parameter, the NMF method is entirely parameter-independent. Furthermore, our NMF technique requires no prior knowledge (such as ECG readings). Perhaps more importantly, NMF end-members are visually quite similar to the selected ground truth end-systole and end-diastole, justifying the NMF’s functionality to successfully find a two-dimensional representation of the full video sequence. Furthermore, the NMF generated images contain less noise, and also accurately retain the heart topology of the original ground truth. This property makes NMF images a good choice for further quantification of heart performance in terms of biologically important indices such as the ejection fraction.

It is also clear that NMF coefficients change periodically, reflecting the cyclic cardiac nature. By analyzing these coefficients and selecting the peaks and valleys of the coefficient curves, we can detect end-systolic and end-diastolic frame number robustly. Comparing NMF with LLE and ISOMAP (Figure 5), we conclude that the NMF method is far more robust to noise with more stable performance from its lower difference mean and variance (Table 1). Based on the selected end-systole and end-diastole, we directly calculate the heart rate and cardiac cycle length (Figure 6). The higher correlation with the ground truth on a greater number of cases illustrates NMF’s superiority to ISOMAP and LLE and its potential to capture information to describe cyclic heart movement.

Another very promising application is the decomposition of the original videos. The NMF generated videos and the error term have captured different aspects of heart movement (Figure 7). Since the NMF video is a

linear combination of generated end-members, it shows the heart's change within the two dimensional NMF projection. In other words, one can clearly see the heart volume change between end-systole and end-diastole in the endmembers. On the other hand, the information of muscle change and valve movement has been well-identified in the difference video. One can clearly see the muscle contraction and expansion and the valve motion in the difference videos. Finally, the noise has been suppressed in both the NMF projection video and the difference video, laying a foundation for a more robust analysis.

These results show that NMF method is a powerful and promising tool in dealing with ultrasound videos. An interesting future line of inquiry is to investigate higher rank NMF to determine whether this can capture more clinically relevant information beyond end-systole and end-diastole. Another interesting line of inquiry is further exploration of the decomposed error term, since the heart dynamics is so clearly visible in this setting.

5. CONCLUSIONS

In this paper we provide a fully automatic, patient friendly method to determine end-systole and end-diastole robustly and accurately, which is then extended to calculate heart rate. The generated end-members were shown to closely represent end-systole and end-diastole and can potentially be used for further analysis. We are also able to deconstruct the original video into a sum of two features: volume change and muscle movement.

ACKNOWLEDGMENTS

We thank Richard Koffler of Viderics, Inc. for useful comments. We thank Da Kuang for help with the NMF code. This work was supported by UCLA through the Physical Sciences Division including the Entrepreneurship and Innovation Fund and the Department of Mathematics. GI was supported by NSF grant DMS-1045536. ALB was supported by NSF grant DMS-1417674 and ONR grant N00014-16-1-2119. XX and RZ were supported by the Cross-disciplinary Scholars in Science and Technology (CSST) program at UCLA.

REFERENCES

- [1] Lang, R. M., Badano, L. P., Mor-Avi, V., Afilalo, J., Armstrong, A., Ernande, L., Flachskampf, F. A., Foster, E., Goldstein, S. A., Kuznetsova, T., et al., "Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the american society of echocardiography and the european association of cardiovascular imaging," *Journal of the American Society of Echocardiography* **28**(1), 1–39 (2015).
- [2] Darvishi, S., Behnam, H., Pouladian, M., and Samiei, N., "Measuring left ventricular volumes in two-dimensional echocardiography image sequence using level-set method for automatic detection of end-diastole and end-systole frames," *Research in Cardiovascular Medicine* **2**(1), 39 (2013).
- [3] Gifani, P., Behnam, H., Shalbaf, A., and Sani, Z. A., "Automatic detection of end-diastole and end-systole from echocardiography images using manifold learning," *Physiological Measurement* **31**(9), 1091 (2010).
- [4] Shalbaf, A., AlizadehSani, Z., and Behnam, H., "Echocardiography without electrocardiogram using non-linear dimensionality reduction methods," *Journal of Medical Ultrasonics* **42**(2), 137–149 (2015).
- [5] Hall, J. E., [*Guyton and Hall textbook of medical physiology*], Elsevier Health Sciences (2015).
- [6] Lee, D. D. and Seung, H. S., "Learning the parts of objects by non-negative matrix factorization," *Nature* **401**(6755), 788–791 (1999).
- [7] Lai, E., Moyer, D., Yuan, B., Fox, E., Hunter, B., Bertozzi, A. L., and Brantingham, J., "Topic time series analysis of microblogs," *IMA Journal of Applied Mathematics* **81**(3), 409–431 (2016).
- [8] Lin, C.-J., "Projected gradient methods for nonnegative matrix factorization," *Neural computation* **19**(10), 2756–2779 (2007).
- [9] Kuang, D. and Park, H., "Fast rank-2 nonnegative matrix factorization for hierarchical document clustering," *Proc. SIGKDD*, 739–747, ACM (2013).
- [10] Lang, R. M., Bierig, M., Devereux, R. B., Flachskampf, F. A., Foster, E., Pellikka, P. A., Picard, M. H., Roman, M. J., Seward, J., Shanewise, J., et al., "Recommendations for chamber quantification," *European Heart Journal-Cardiovascular Imaging* **7**(2), 79–108 (2006).