

# LOCAL WALL MOTION CLASSIFICATION OF STRESS ECHOCARDIOGRAPHY USING A HIDDEN MARKOV MODEL APPROACH

Sarina Mansor, J Alison Noble

Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, UK.

## ABSTRACT

In this paper, we present a new approach to regional heart functional analysis employing a Hidden Markov Model (HMM) approach for cardiac disease classification. We examine the use of an HMM for local wall motion classification based on stress echocardiography. A wall segment model is developed for a normal and an abnormal heart and the experiments are performed on rest and stress sequences. We achieve a good accuracy of classification, particularly for the normal data but with also promising results for the abnormal case.

**Index Terms**— Local Wall Motion Analysis, Stress Echocardiography, Hidden Markov Models.

## 1. INTRODUCTION

Analysis of left ventricular (LV) regional function is important for diagnosing heart disease, especially to detect myocardial ischemia (a disease of heart muscle). In current clinical practice, the analysis mostly relies on visual assessment by experienced cardiologists. This makes the diagnosis of regional heart function disease a highly subjective and operator-dependent problem. While much has been done to attempt to automate the task of wall motion analysis on rest image sequences, there is a very small literature on analysis of stress echocardiography. Stress sequences (where the heart has been stressed by exercise or use of a pharmacological drug e.g. dobutamine) are more difficult to automatically analyse as the non-rigid motion of the heart is more challenging to track. Further, to our knowledge there has been no prior attempt to automatically classify heart motion based on the *combined information* derived from a rest and a stress sequence. This is the subject of this paper.

The first step of automated regional heart function analysis is to detect the heart wall borders. Many techniques have been developed to detect and automatically track both the endocardial and epicardial borders of the left ventricle (LV). Current automated 2D echocardiography image tracking technology is now sufficiently well-developed for application on good-to-medium quality rest data [1]. However, further studies need to be done to show that this is

also true for stress echocardiography. In the work reported here, we used semi-automated derived contours as our initial focus has been on classification not automated tracking.

The second step of automated regional heart function analysis is classification of heart segmental function as either normal or abnormal based on the extracted contours (heart wall borders). This is the focus of this paper. In our prior work [2], we performed global wall motion classification using an HMM approach and promising results were obtained. In the literature we are only aware of local wall motion classification using the PCA approach and on rest sequences only [3].

This contribution investigates the use of a Hidden Markov Model (HMM) as a tool for regional stress echocardiography classification. Hidden Markov models are especially known for their application in temporal pattern recognition such as speech [4] and biosignals [5] because of their ability to successfully learn the time-varying characteristics of signals. Therefore, we would like to employ an HMM approach since the cardiac data inherits the time-varying and sequential properties. To the best of our knowledge, there is no other published work incorporating HMMs with regional heart function analysis. The most closely related work is the use of a HMM for 2D shape analysis [6]-[8], particularly [7] which implemented a HMM for spatio-temporal pattern recognition.

## 2. HIDDEN MARKOV MODELS

A detailed tutorial on HMMs can be found in [4] and a brief description based on that paper is given here. A hidden Markov model (HMM) is a probabilistic model which describes the statistical relationship between an observable sequence  $O$  and a “hidden” state sequence  $S$ . The hidden state is discrete, but the observation values may be either continuous or discrete in nature [5].

An HMM is characterized by the following parameters:

- 1) The number of states of the model,  $N$ .
- 2) The state transition matrix,  $A$
- 3) The observation probability distribution,  $B = \{b_N\}$  for each state  $N$
- 4) The initial state distribution,  $\pi$

However, for convenience, the HMM  $\lambda$  is parameterized by  $A$ ,  $B$  and  $\pi$ , with the notation:  $\lambda = (A, B, \pi)$ .

There are three basic problems associated with HMMs:

1. The classification/evaluation problem  
Given a model  $\lambda$  and a sequence of observations  $O$ , we would like to compute the probability that the observations are generated by the model,  $p(O|\lambda)$ . This problem can be solved by a ‘forward-backward’ procedure [4].
2. The decoding problem  
Given a model  $\lambda$  and a sequence of observations  $O$ , we would like to find the most likely sequence of hidden states that could have generated the observations. This problem is typically solved by the Viterbi algorithm [4].
3. The learning/training problem  
Given a set of observation sequences, we would like to adjust the model parameters,  $\lambda = (A, B, \pi)$  to maximize the probability of the given dataset. This problem is typically solved by the EM algorithm [4].

In our work we deal with the first and third problems. A regional heart function model needs to be learned from a training set by the EM algorithm before using the forward-backward procedure for classification of a new dataset.

### 3. METHODOLOGY

#### 3.1. Patient Data

The available database consisted of 30 studies of Contrast DSE (Dobutamine Stress Echocardiography) data acquired by two cardiologists as standard B-mode ultrasound image sequences. Each study contained:

- 4 planes: 2-chamber (2C), 3-chamber (3C), 4-chamber (4C) and short-axis (SAX) views.
- 3 stress stages: rest (no dobutamine), intermediate (low-dose dobutamine) and peak (the maximum dose that a patient can take).

In this paper, we utilized the 2C data (which has 6 segments) at the resting phase and at the peak level of stress. In a standard stress examination, only the systolic (heart contracts) phase is recorded. Therefore, each data consists of systole only starting from the end of diastolic (ED) phase to the end-of-systolic (ES).

For regional functional analysis, the left ventricle is divided into 17 segments but mostly only 16 are measurable (especially for abnormal hearts) [9]. Clinical wall functional assessment evaluates systolic thickening<sup>1</sup> and endocardial wall motion for each segment. A normal (healthy) myocardial segment shows systolic thickening and also endocardial movement towards the centre of the cavity. Each of the 16 segments were labelled 1 - 4 (1 = normal, 2 = hypokinesis, 3 = akinetic, 4 = dyskinetic). For simplicity, we grouped them into 2 classes only (1 = normal, 2-4 = abnormal).

<sup>1</sup> The myocardial wall thickening during systole (the contraction of the heart) phase

#### 3.2. Wall Motion Classification

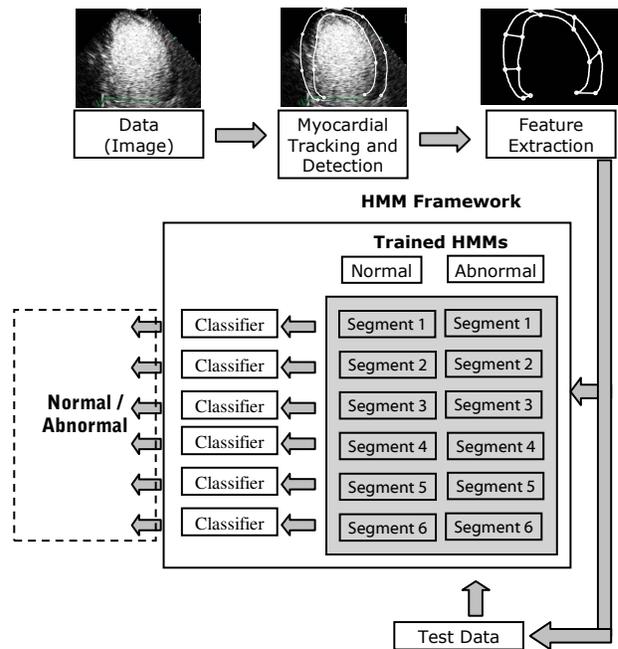


Figure 1: Diagram of Wall Motion Classification process

Figure 1 shows the general process of our HMM-based wall motion classification method, which can be summarized as follows:

1. Firstly, the myocardial contours were extracted by Quamus<sup>®2</sup> software. These were then validated by two experienced cardiologists.
2. Each contour was then divided into six segments (for 2C data) and a number of features estimated to characterize cardiac wall motion (e.g.: wall thickening, cavity area, endocardial motion).
3. The cardiac data was separated into two groups based on wall motion scoring: Normal (score = 1), Abnormal (score = 2-4).
4. Two HMMs were then developed for each segment: one for normals and one for abnormal. Thus, there were 12 trained models in total.
5. Finally, the trained HMMs (models) were then be used to classify a new (unseen) dataset.

#### 3.3. Feature Vector / Object Representation

As mentioned earlier, several features could be used to detect segmental wall motion abnormalities. In our previous paper [2], we used myocardium area as the feature (observation) vector, which gave a good accuracy of

<sup>2</sup> A semi-automatic boundary detection software developed by Mirada Solutions Ltd, Oxford, UK.

classification. In this paper, we chose the ‘cavity area’ (area under endocardial contour) for each of six segments, as shown in Figure 2. This feature vector was suggested by our cardiologist collaborators as it is the simplest and easiest observation representation that can be viewed from all echocardiographic data. Furthermore, we need only endocardial contours (which are easier to track/segment than epicardial contours) in order to compute for the cavity area. The cavity area (CA) is then normalized by the following equation (1):

$$CA(i) = \frac{CA_{ED} - CA(i)}{CA_{ED}} \quad (1)$$

where  $CA_{ED}$  = Cavity Area at End of Diastolic.

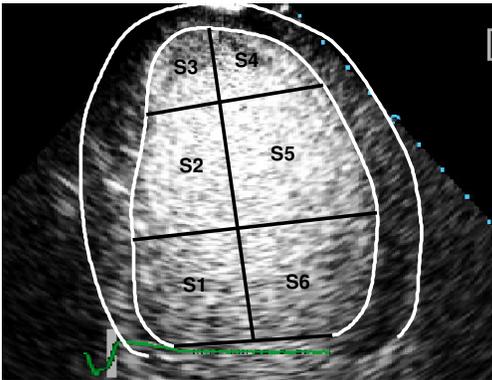


Figure 2: Cavity Area (CA) for all six segments of 2C data (S1 – S6 = CA for Segment 1 – Segment 6)

### 3.4. Training

The sequential nature of the heart phases can be well modelled by a left-right HMM. We developed two models for each segment: one for normals and another for abnormal. Each model is comprised of three hidden states. In our case, the states of the HMM do not have a true physical meaning – they simply reflect common statistical properties of the observation vectors in the feature space. We could have employed more states but that would require more training samples for an accurate estimation of the model parameters [6].

We trained the two HMMs (one for normals, one for abnormal) in an *unsupervised* manner. Prior to training, all dataset were resampled to have 30 frames per patient. We then calculated the normalised cavity area for each frame, as previously described.

In the first stage of training, we initialised the model parameters as follows. For each patient in the dataset, the first ten frames were assigned to state 1, the next ten to state 2 and the final ten to state 3. The parameters of the transition matrix  $a_{ij}$  were then initialised using the maximum likelihood estimates, based on the following equation (2):

$$a_{ij} = n_{ij} / \sum_k n_{ik} \quad (2)$$

where  $n_{ij}$  is the total number of transitions from state  $i$  to state  $j$  over all of the label sequences.

The observation probability densities  $b_i$  (which were Gaussian) were initialised by fitting a Gaussian to the set of features extracted from the frames assigned to each state  $i$ . Following initialisation of the HMM parameters, we then ran the EM algorithm to update the model parameters until the change in the log likelihood was less than 0.01% (indicating convergence).

### 3.5. Classification and Confidence Measure

For classification purposes, we compute the posterior probability,  $p(model|data)$  for all models and the maximum value gives the most probable model for the data. The posterior probability can be computed by using Bayes’ rule:

$$p(model | data) = \frac{p(data | model) p(model)}{\sum_{model} p(data, model)} \quad (3)$$

where  $p(data|model)$  is computed using the forwards-backwards algorithm and  $p(model)$  is the prior probability for each model.

A confidence measure is then needed to quantify the degree of confidence in the result (i.e. how well a model matches the data). The resulting posterior probability  $p(model|data)$  is a number in the range [0,1]. We can then assess the confidence in the resulting classification by assessing how close this value is to 1.

For example, we have 2 models: one for Normals (N) and one for Abnormals (A). Suppose that  $p(N | data) = 0.99$  and  $p(A | data) = 0.01$ , then we would be fairly sure that this person has a normal (healthy) heart. However, if instead we found that  $p(N | data) = 0.51$  and  $p(A | data) = 0.49$  then we would be much less sure since the difference in the posterior probabilities for the two models is very small. A suitable confidence measure is therefore:

$$confidence = |p(N|data) - p(A|data)| \quad (4)$$

which is zero if the two probabilities are 0.5 (maximal uncertainty) and 1.0 if either of the probabilities equals 1.0 (maximal certainty).

## 4. RESULTS AND DISCUSSION

The new approach was evaluated using two classification experiments: 1) a leave-one-out cross validation and 2) a new dataset testing experiment. Recall that we have a total of 30 sequences: 10 Normals and 10 Abnormals are used for training set and the remaining 10 are the testing set.

In a leave-one-out cross-validation one of the 10 sequences for the model being tested is held out and the remaining 9 sequences are used to derive the HMM. The

held-out sequence is then tested on this HMM and the other fully trained HMM. Tables 1 and 2 show the classification accuracy results for rest and stress data using both global and local classification approach. The global classification method can be found in our previous paper [2].

Table 1: Classification accuracy of ‘rest’ sequences using a leave-one-out approach

Global Classification (%)		Local Classification (%)		
Normal	Abnormal		Normal	Abnormal
100	80	S1	90	100
		S2	90	80
		S3	80	80
		S4	80	60
		S5	100	90
		S6	100	80

Table 2: Classification accuracy of ‘stress’ sequences using a leave-one-out approach

Global Classification (%)		Local Classification (%)		
Normal	Abnormal		Normal	Abnormal
90	60	S1	80	70
		S2	80	50
		S3	70	100
		S4	90	70
		S5	80	70
		S6	70	80

In the testing experiment, we computed the log likelihood that each model (Normal and Abnormal) gives to the test sequences and classified them based on equation (3) (refer to Section 3.5). Table 3 gives the classification accuracy of the 10 testing sets for both rest and stress data.

Table 3: Classification accuracy for the testing dataset

Testing Set	Classification Accuracy (%)					
	S1	S2	S3	S4	S5	S6
2CV (Rest)	80	80	70	90	90	100
2CV (Stress)	80	90	60	70	70	80

As can be seen from the Tables 1 and 2, we can achieve more or less the same accuracy of local classification as the global ones. Local classification is harder because sometimes not all segments can be clearly seen in echocardiographic data, especially around apex area (segments 3 and 4). As for the testing experiment (refer to Table 3), a good accuracy of classification was achieved instead of a few number of data is used for trained model. This shows the potential use of HMM for classification.

In each experiment, the classification accuracy for the normal model is higher than for the abnormal case. This is because the wall motion of healthy hearts is better-defined than for the abnormal case which really covers subcategories of abnormal functional behaviour i.e. hypokinetic, akinetic,

dyskinetic. Since we had a limited number of patient datasets for the current study, we grouped all the abnormal cases into just one model. For a more accurate classification, they should be modelled according to each abnormal category.

Comparing experiments, the stress sequence has the lowest accuracy. This can be explained in part by the relatively large movement of the heart during stress. Furthermore, the image quality is often poor for this type of acquisition which affects the accuracy in tracing the myocardial borders.

## 5. CONCLUSIONS

In this paper, a new approach of local wall motion analysis of stress echocardiography data has been proposed based on HMMs. For each wall segment, two HMMs have been trained for the normal (healthy) and abnormal cases. A good accuracy of classification was achieved despite the limited number of data used. These encouraging results could be improved further by: 1) increasing the number of training data, for example at least 50 patients for each model, 2) developing different models for different types of abnormality and 3) investigating other possible observation vectors. The general approach would be equally applicable to 3D stress echocardiography which is receiving a lot of interest from the cardiology community at the current time. Finally, the methodology is not modality specific and could be equally applied to stress MRI.

## 6. REFERENCES

- [1] J.A. Noble “Cardiology meets Image Analysis: just another application or can image analysis usefully influence clinical practice?” *IEEE IWCVBIA*, China, 2005.
- [2] S. Mansor, J.A. Noble, “Regional Left Ventricular Wall Motion Analysis using Hidden Markov Models” *MIUA*, 2007.
- [3] K.Y Esther Leung and J.G. Bosch, “Local Wall-Motion Classification in Echocardiograms Using Shape Models and Orthomax Rotations,” *FIMH*, 2007.
- [4] L. Rabiner “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition”, *Proceedings of IEEE*, 77, 2, pp. 257-286, 1989.
- [5] N.P. Hughes et al “Markov Models for Automated ECG Interval Analysis”, *Advances in Neural Information Processing Systems*, 16, 2003.
- [6] Y. He & A. Kundu, “2-D Shape Classification Using Hidden Markov Model”. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13, 11, pp. 1172 - 1184, 1991.
- [7] K.H Fielding et al “Spatio-Temporal Pattern Recognition Using Hidden Markov Models”, *IEEE AES*, 31, 4, pp. 1292-1300, 1995.
- [8] M. Bicegon et al “Investigating Hidden Markov Models’ Capabilities in 2D Shape Classification” *IEEE Pattern Analysis Machine Intelligence*, 26, 2, pp. 281-286, 2004.
- [9] R. M Lang et al. “Recommendations for chamber quantification”, *European Journal Echocardiography*, 7, pp. 79-108, 2006.