Modeling and Analysis of Echo-Cardiogram Video

Jayanta Mukhopadhyay

Dept. of Computer Sc. & Engg.

IIT Kharagpur

Collaborators

- Dr. Biswajit Acharya
- Dr. Aditi Roy
- Prof. S. Sural
- Prof. A.K. Majumdar

1. **B. Acharya et. al.,(2006), Video models for dynamic objects, Information Sciences, 176, pp. 2567-2602.**
2. **A. Roy et. al. (2008), State based Modeling and Object Extraction from Echocardiogram Video , IEEE trans. on IT in Biomedicine, no. 3, vol. 12, pp. 366-376.**

Echo / Doppler Cardiography

- A study for finding information about
  - The structure and function of the heart
  - Cardiac hemodynamics of the heart
- Uses ultrasound imaging.

Different Phenomena in Imaging

- Shadow Formation
- Reverberation
- Side Lobes
- Foreshortening
- Harmonic Imaging

Video of a Dynamical System

- Video depicts external views of objects and their behavior
  - **External view depends on camera angles, focusing parameters, sharpness parameters, environmental parameters**, etc.
- Video presents temporal changes in spatial and topological relationships among the objects
  - **Changes in shape, size, color, location over time.**

Modeling Issues

- Video model is dictated by the objects present

- Defining relationships between objects
  - Spatial and temporal relationship
  - Behavioral relationship

Contextual Information

- Objects
  - Attributes
  - Temporal and spatial relationships with other objects.
  - States.
  - State transitions.
- Events

Event?

- A phenomenon or any observable occurrence
- Event recognition: A high-level interpretation of the phenomenon depicted in a video.

Detection

- Objects: Role / State

State Transition

Description

- Qualitative
- Quantitative

Duration

Dynamical Systems Model

- A system consists of several interacting objects.
- Objects can have static and dynamic properties.
- Dynamic behavior of an object is determined by its state.
- Objects can have finite number of states.
- Objects change states as a consequence of events.
- Interaction between objects need to be represented also (say by activity diagram).

Video of Dynamical Systems: Examples

- Echo-cardiogram Video
  - Objects: Heart Chambers, Valves
  - States: Diastole, Systole
  - Events: Mitral valve opening / closing
- Soccer Video
  - Objects: Players, Ball, Bill-boards, …
  - States: Play-on, break, interval, ….
  - Events: Kick-off, Ball-passing, Goal …

Examples (contd.)

- Video Surveillance
  - Objects: Human being, luggage, ..
  - States: Moving, Stationary, Gait cycle..
  - Events: Pull, Push, Draw, Throw, Keep
- Video for Pediatric Neurological Study
  - Objects: Infant, Dress, Appliances
  - States: Different Poses
  - Events: Pulling head, moving hand, ..

Example: Echocardiogram Video

- Depicts the physiological process of heart
- Depicts changes in objects of heart
  - **Objects** are:
    - Ventricles
    - Atria
    - Valves
    - Arteries
- Changes are
  - **Structural** (shape, size, wall movement, etc.)
  - **Behavioral** (expansion, contraction, opening, closing, etc.)

Echocardiogram Video

- Different measurements require different parts of heart to be seen on the screen

$\Rightarrow$ different parts can only be seen in different views (transducer position)

$\Rightarrow$ an echo video may contain different views

- Any component of heart may or may not be seen in a particular frame, but it's still undergoing the pre-defined changes

Echocardiogram Video Views

Short-axis view

Long-axis view

Apical view

Color Doppler

The Heart

- Heart changes state at the end of a cardiac cycle.
- Every object follows cardiac cycle.
- Behavior of different objects may not be same over a specified time interval (e.g. when atria contract, ventricles expand).
- Periodicity of events may not be constant over time (e.g. cardiac cycle duration may vary).

Dynamics of Heart Operation

- Heart in *systole*

  ⇒ LV and RV in *contracting*

  ⇒ LA and RA in *expanding*

  ⇒ AV in *open*, MV in *closed*

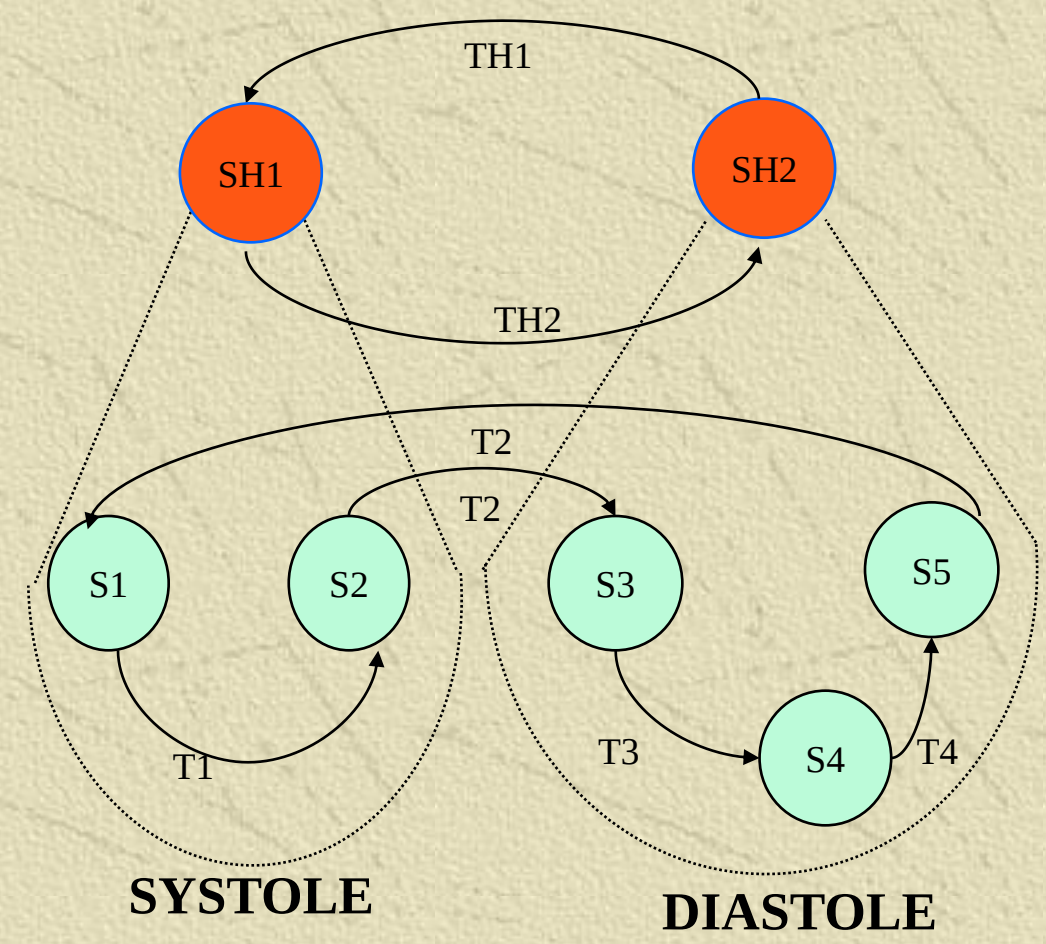
- Heart in *diastole*

  ⇒ LV and RV in *expanding*

  ⇒ LA and RA in *contracting*

  ⇒ AV in *closed*, MV in *open*

State Transition of Heart and Left Ventricle



STATES

SH1 SYSTOLE

SH2 DIASTOLE

S1 ISOVOLUMIC CONTRACTION

S2 EJECTION

S3 ISOVOLUMIC RELAXATION

S4 RAPID INFLOW

S5 FULLY EXPANDED

TRANSITION EVENTS

TH1 BLOOD INFLOW

TH2 BLOOD OUTFLOW

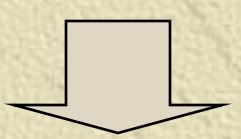
T1 VOLUME DECREASES

T2 VOLUME CONSTANT

T3 VOLUME INCREASES

T4 RAPID INCREASE OF VOLUME

The Model



State-based Video Model

Definitions of related concepts

- Video is a dense sequence of frames
  - Attributes
    - Vid (video id)
    - Pid (parent video id)
    - Name (e.g. news_video, echo_video, soccer_video)
    - Duration (length of video in #frames)
  - Methods
    - Start(vid) (first #frame w.r.t. parent video)
    - end(vid) (last #frame w.r.t. parent video)
    - Select_frame(frame_no) (extract i^{th} frame)

Definitions (contd.)

1. ***OBJECT***:

   A distinguishable component of the physical system depicted in the video, e.g. ventricles, atria, valves, arteries etc.

2. ***STATE***:

   Dynamic behavior of the object: captures temporal and spatial semantic of the underlying process.

3. ***TRANSITION EVENT***:

   Causes of state transitions, e.g. blood inflow, blood outflow, opening of valve, etc.

Video and States (Mapping)

Let $\mathcal{V}$ be the set of video frames of a domain

Let $\mathcal{O}$ the set of objects present in the schema

Then, the association of frames to objects is defined by

$$\alpha: \mathcal{V} \rightarrow \wp(\mathcal{O})$$

And, the association of objects to frames is defined by

$$\beta: \mathcal{O} \rightarrow \wp(\mathcal{V})$$

Object Properties

- Intrinsic properties
  - Name, object attributes, object components, object states
- Presentation properties
  - Visual (e.g. color, texture, size)
  - Perspective (e.g. top-view, front-view)
  - Camera-related (e.g. close-shot, long-shot,pan, zoom)
  - Positional and topological (e.g. left-of, inside, occluded)

Definitions

- Video Frame
  - Atomic part of a *simple* video or *composite* video
- Video Segment $v_s\ [i, j]$
  - Dense subsequence of a video from i^{th} frame to j^{th} frame
- Composite Video
  - "synchronized" collection of a set of video, called "component" video objects

Operations on Segments

- Spatial
  - Projection
  - Concatenation
  - Contained-in
  - Contains
  - Overlap
- Temporal
  - Follows
  - Follows-immediate
  - Followed-by
  - Followed-by-immediate

Video Calculus

- Σ **set of constant symbols consisting of**
  - **Atomic values (*int, float, duration etc.)***
  - **Set of object instances (defined in schema)**
  - **Set of video segments**
  - **Set of constant functions**
- $\hat{V}$ **set of variables**
- **connectives** $\wedge$(and), $\vee$(or), $\neg$(not)
- **operators** $:=$ (assignment), [ ] (interval), $\Rightarrow$(implication), $\equiv$(equivalent)
- **Quantifiers** $\forall$(universal quantifier), $\exists$ (existential quantifier)
- **Parenthesis** ( )
- **Predicate Symbol** p^1, p^2
- **Function Symbol** $\tilde{f}, \tilde{g}$

Video Calculus (contd.)

Let X, Y object variables, A, B atomic values, θ binary predicate. Then Atomic Formula can be written as

X.A θ c, X.A θ Y.B or θ(X.A, Y.B)

Well-formed formula:

- Atomic formula
- $F \vee G, F \wedge G, \neg F$
- $\forall xF, \exists xF$

Semantics

- "Meaningful" scene is defined in terms of occurrence of states of objects of interest.
- Dynamic behaviour is defined by state transition of objects present.
- Semantics of a video segment dictated by a well-formed formula, called guard condition.

Formula

- Relational

  $p(v_s)$: v_s . *duration* = 30 $\wedge$ v_s .*name* = '*soccer*'

- Compositional

  $p(v_s)$: $\forall$ *k* $\in$ [v_s . start , v_s .end] $\exists$ O $\in \alpha$ (v_s . *select_frame* (*k*)) (*O.player_name* = '*Pele*')

- State-based

  $p(v_s$, *LV.systole*) : $\forall$ *k* $\in$ [v_s . start , v_s .end] (*LV* $\in \alpha$ (v_s . *select_frame* (*k*)) $\wedge$ *LV. state* = '*systole*')

States & Reachability

reachable ($O.s_1, O.s_2$), written as $O.s_1 \sim> O.s_2$, defined recursively as:

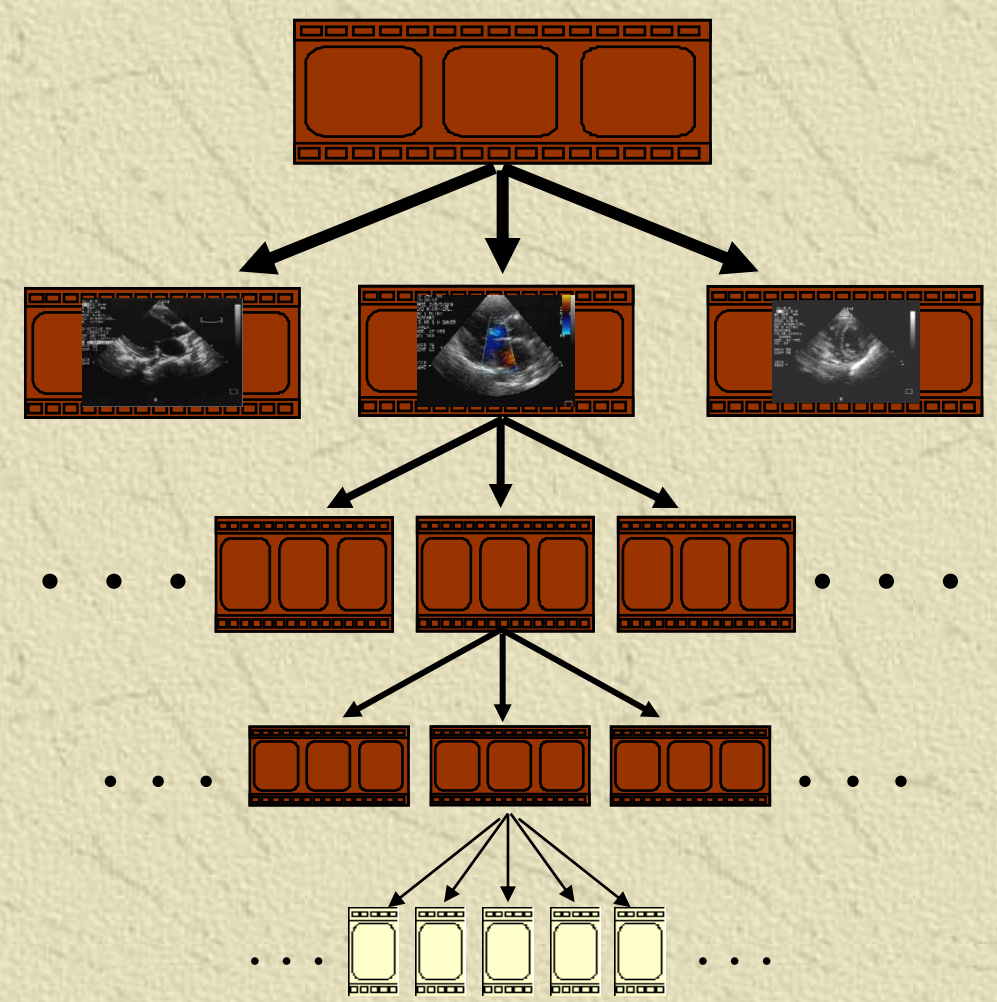
1. *next_state* ($O.s_1$, $O.s_2$) $\Rightarrow$ *reachable* ($O.s_1, O.s_2$)
2. *next_state* ($O.s_1$, $O.s_3$) $\wedge$ *reachable* ($O.s_3, O.s_2$)

   $\Rightarrow$ *reachable* ($O.s_1, O.s_2$)

**** The EDB predicate ***next_state*** ($O.s_1$, $O.s_2$) corresponds to the state transition table of O, where the state s_2 is reached as a one-step transition from the state s_1.

Model driven Video Processing

Video Segmentation (state-based)



ECHO-VIDEO

ECHO-SHOTS
(Camera Views)

PHASES
(State based)

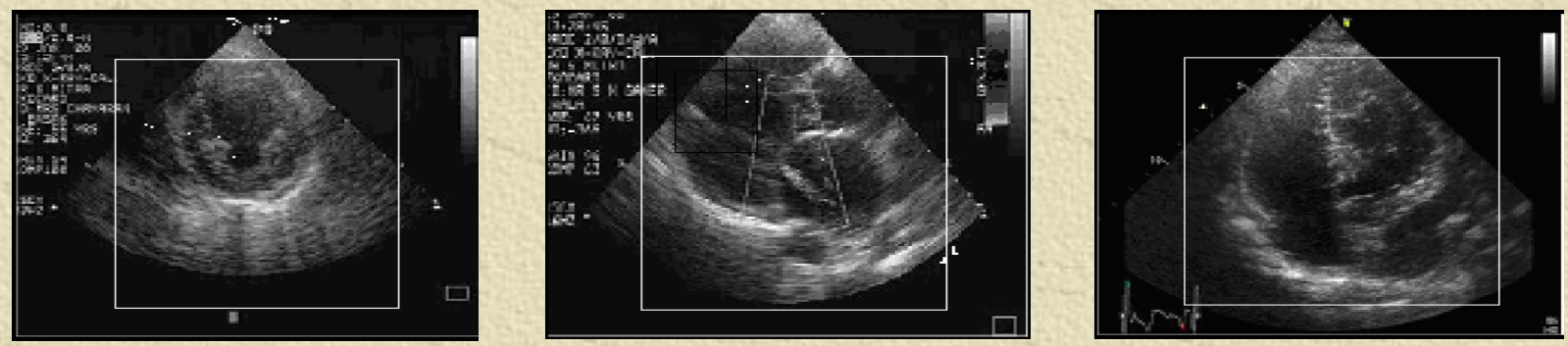
EPISODES
(Sub-state based)

FRAMES

View Detection and Identification

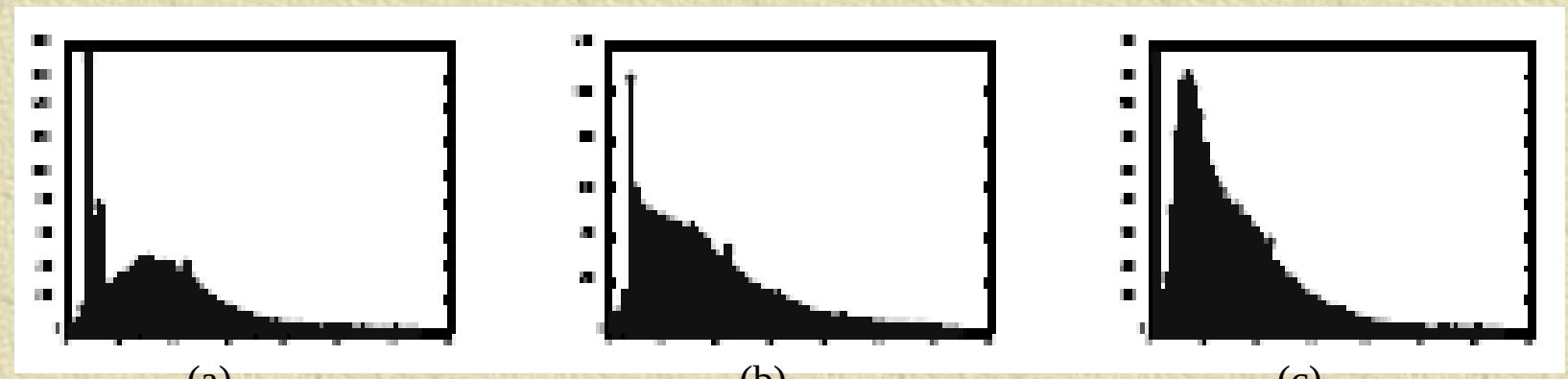
- Histogram based comparison
- Edge change ratio
- The number of chambers present, their orientation and the presence of heart muscles in each view, gives different histogram pattern.

View Detection and Identification (contd.)



(a) (b) (c)

Frames with ROI : (a) Short Axis View; (b) Long Axis View; (c) Apical View



(a) (b) (c)

Histogram of (a) Short Axis View, (b) Long Axis View, (c) Apical View for above figure

The Algortihm

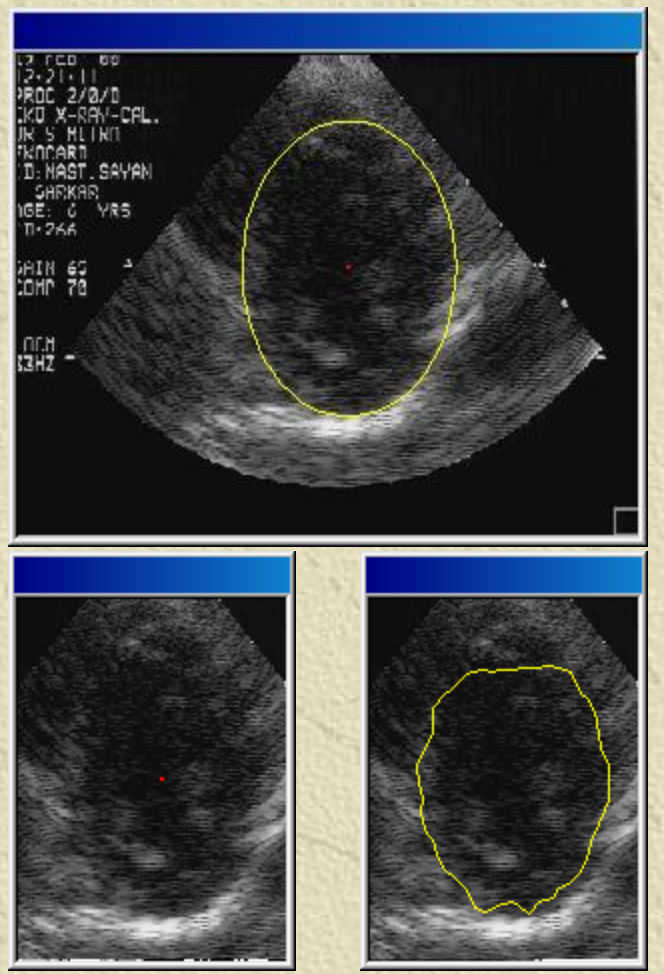
- *Step 1 :* Define a ‘region of interest’ for each frame.
- *Step 2 :* Generate gray scale histogram for this ROI using 64 bins.
- *Step 3 :* Normalize the histograms.
- *Step 4 :* Create 64-element input vector for each frame based on the normalized bin count value.
- *Step 5 :* Train a neural network with one input layer (having 64 units), one hidden layer (having 80 units) and one output layer (having 3 units one for each view).

State and Sub-state Identification

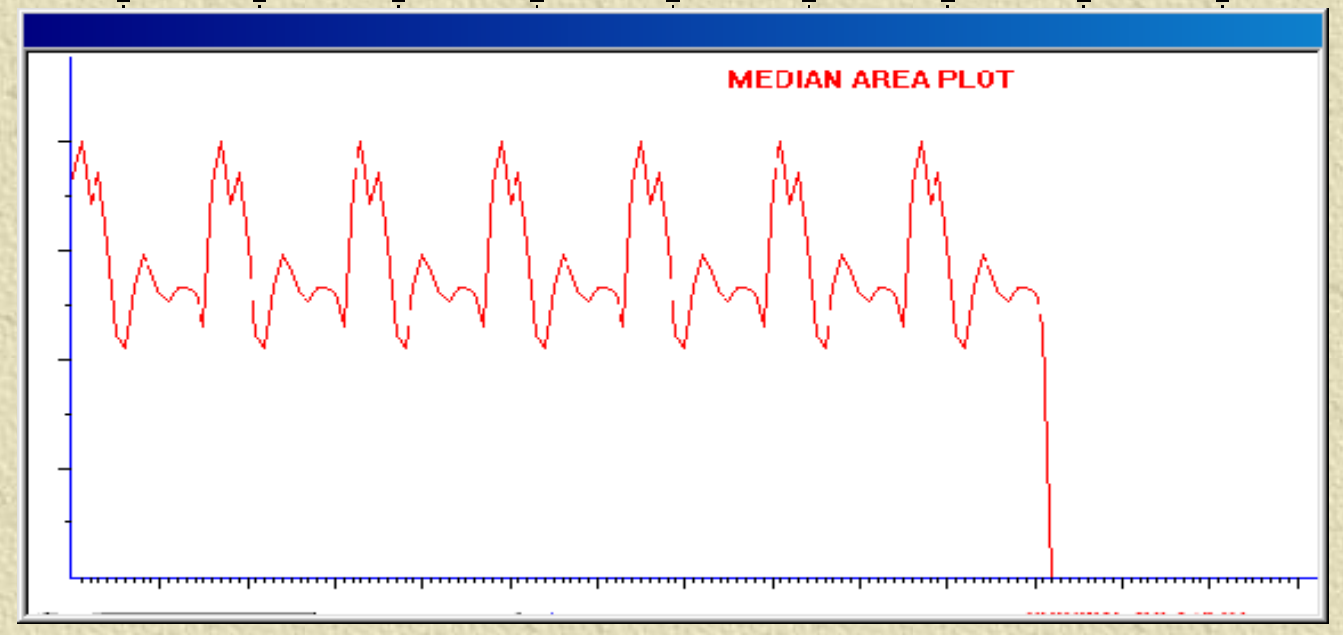
- Identification of states
  - Estimating dynamic properties of objects
  - Using Sweep M-mode
- Identification of sub-states using Radial Color M-mode

LV Contour Estimation

- **User's Input**
  - –Major axis
  - –Minor axis
- **Initial Contour**
  - –Ellipse Drawing
- **Active Contour**
  - –Relaxation Technique

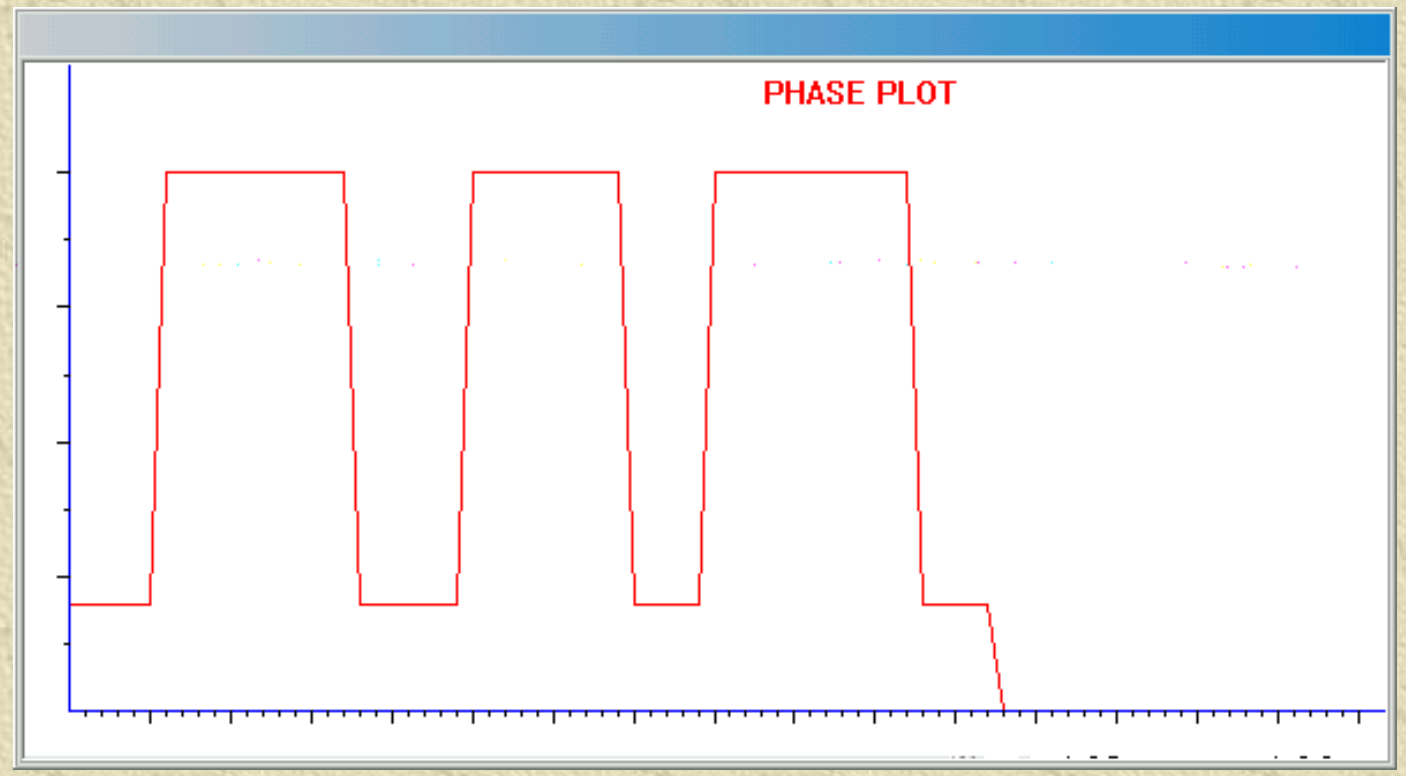


LV Relative Area Estimation



- Compute median cardiac cycle from correlation map
- Compute median area of corresponding frames from available cardiac cycles

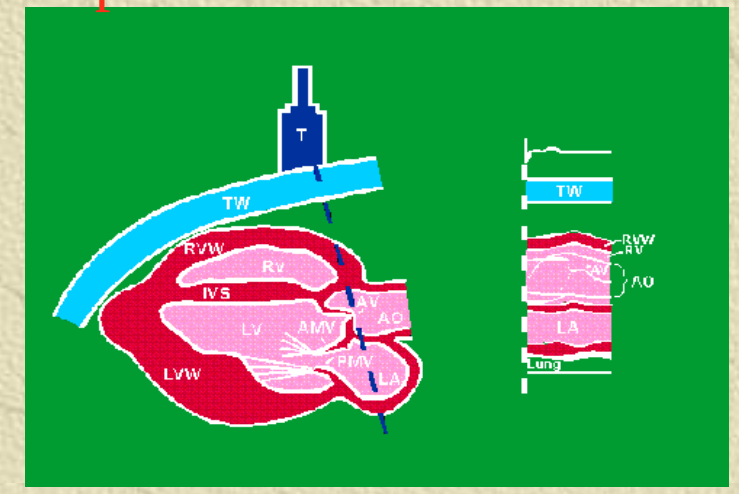
Heart State Detection



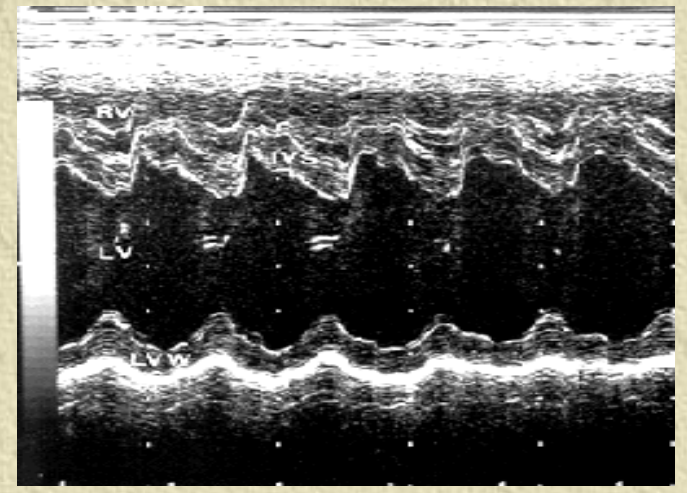
What is M-mode ?

- M-mode is represented as an image in which abscissa and ordinate represent time and depth (or distance), respectively.

Temporal variation of intensity profile



Transducer position for obtaining M-mode



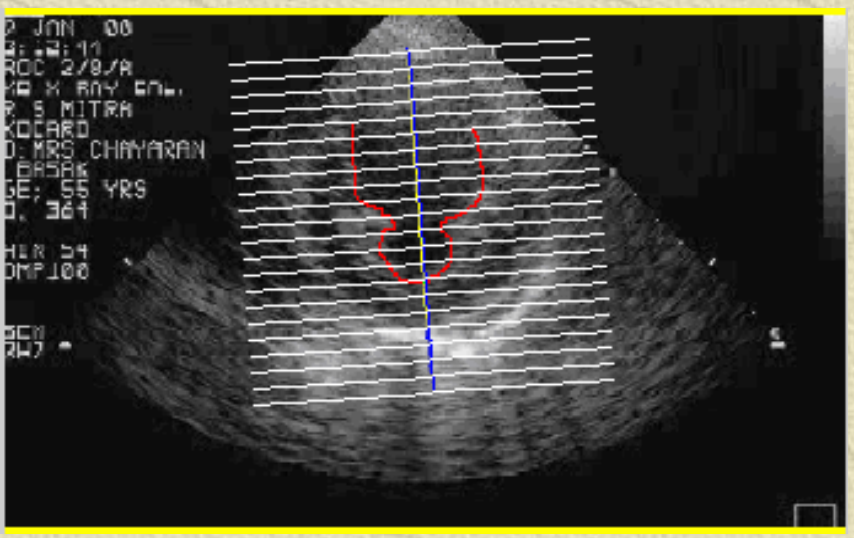
M-mode image

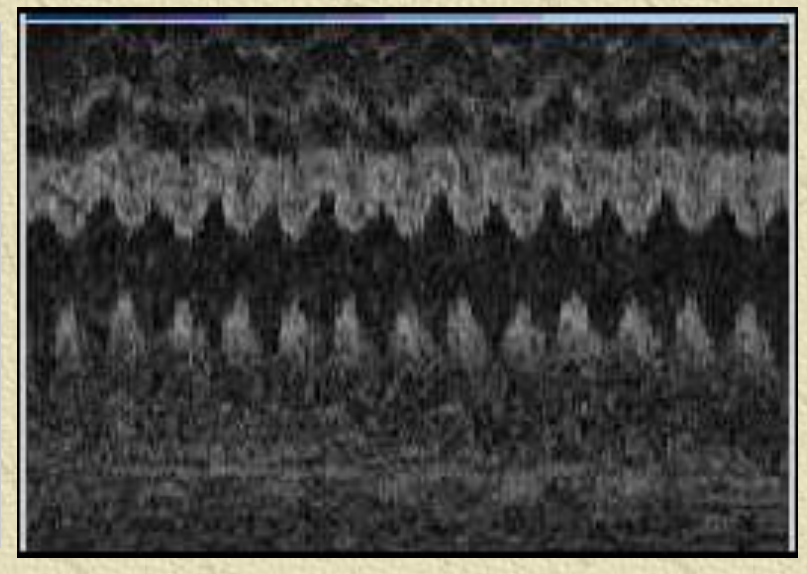
Identification of States using Sweep M-mode

- Synthetic sweep M-Mode generation
- Pre-processing
- Border extraction
- State identification

Step 1 : Synthetic Sweep M-Mode Generation (SAX)

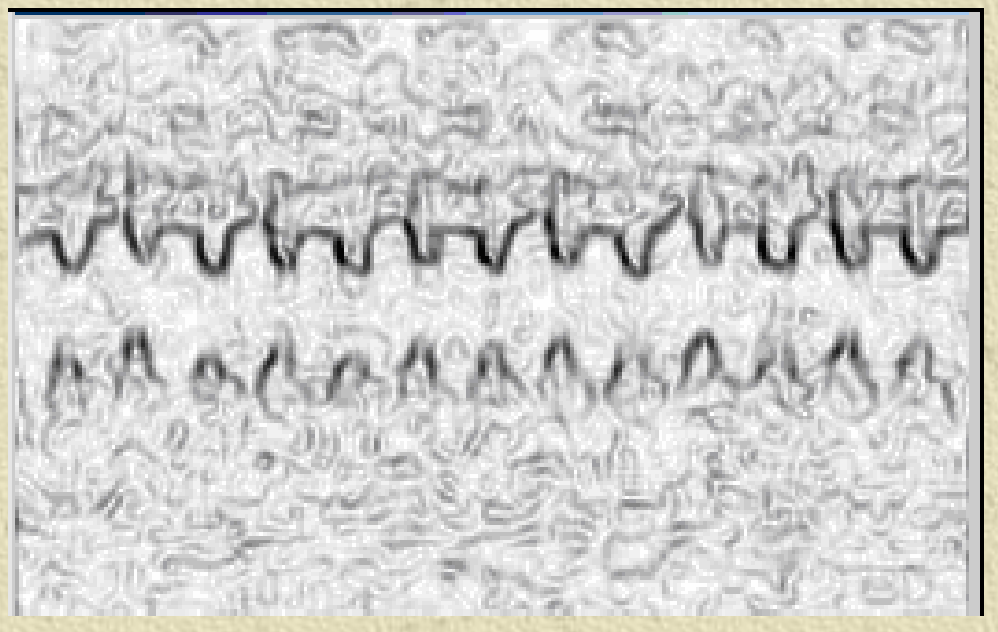
- Multiple M-modes along horizontal lines.





M-mode image

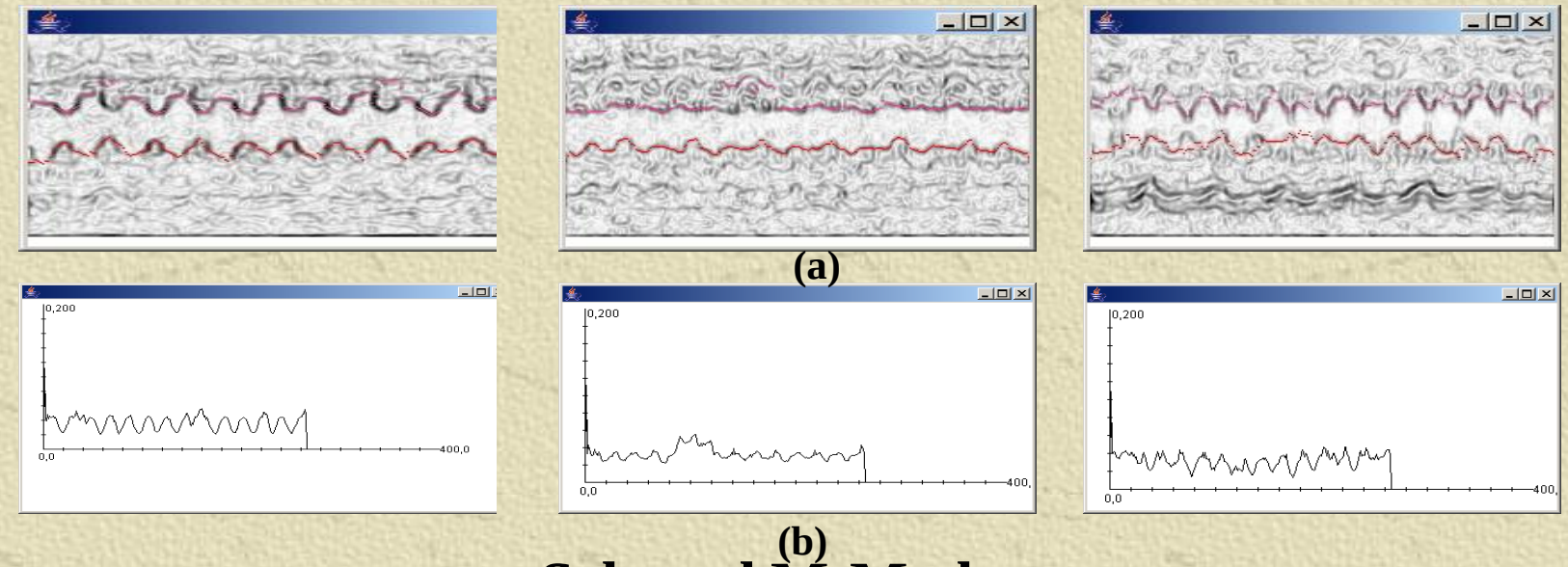
Step 2 : Pre-processing



Edge detected smoothed M-mode image

Step 3 : State Identification through tracking of homogeneous pixels along horizontal lines

- Distance between corresponding structures
- Selection of M-modes based on observed semi-periodicity of distance values.

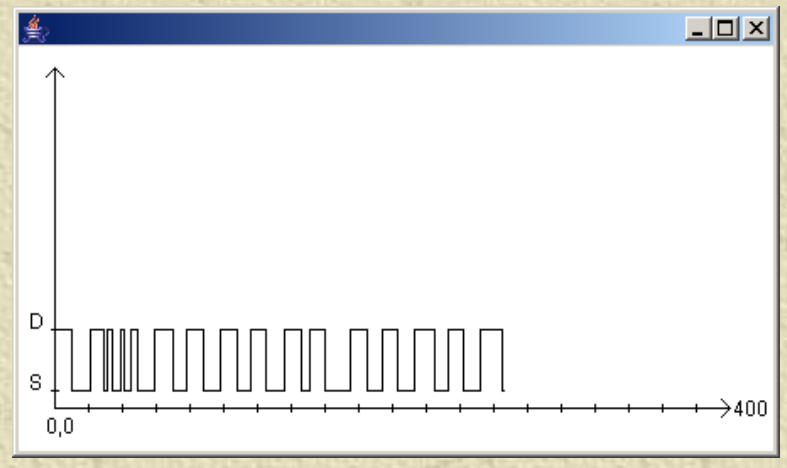


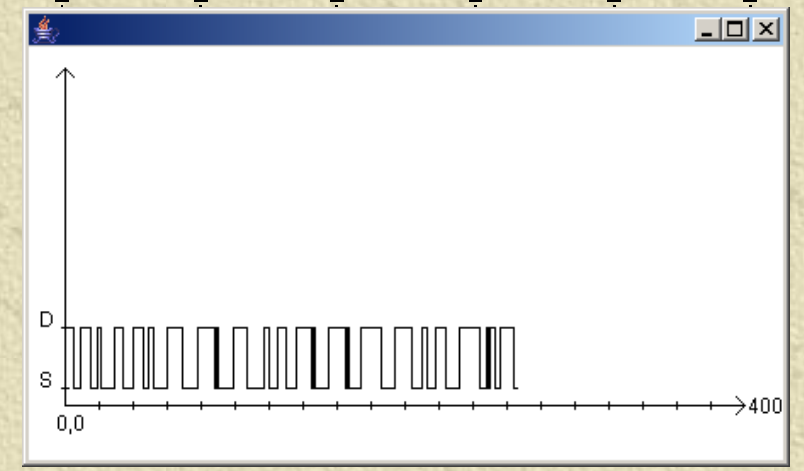
Selected M-Modes

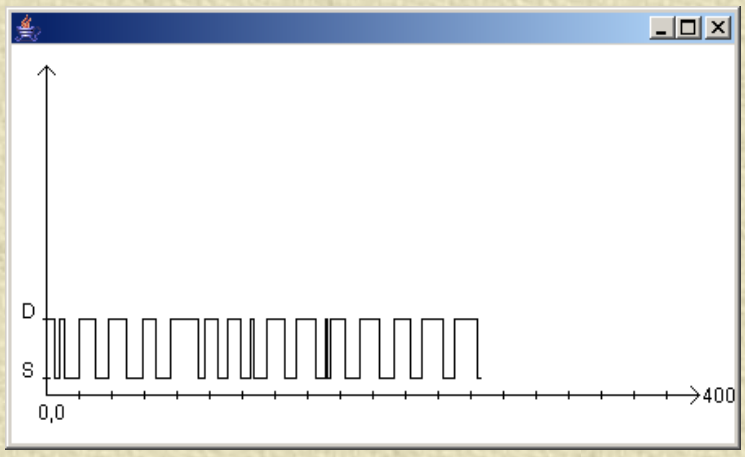
Identification of Diastole and Systole from SAX

- LV: Diastole: Expanding → Slope of distance function is +ve.
- LV: Systole: Contracting → Slope of distance function is -ve.
- Frames: +ve slope: Diastole
- Frames: -ve slope: Systole

State Transition from M-Modes





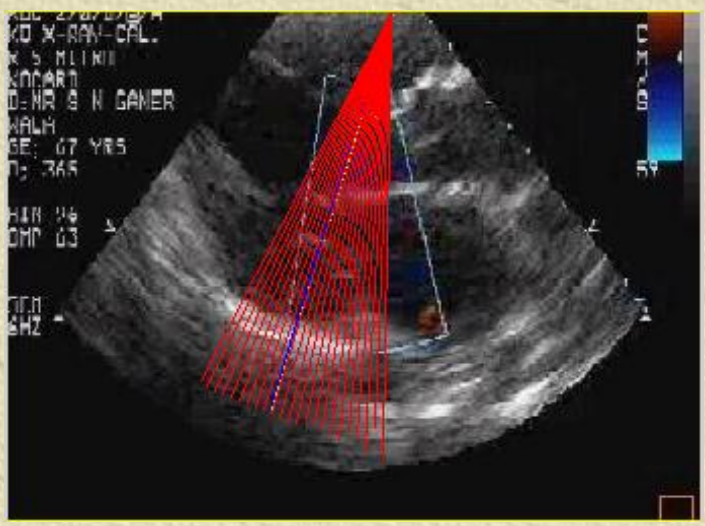


Identification of Sub-states using Radial Color M-mode

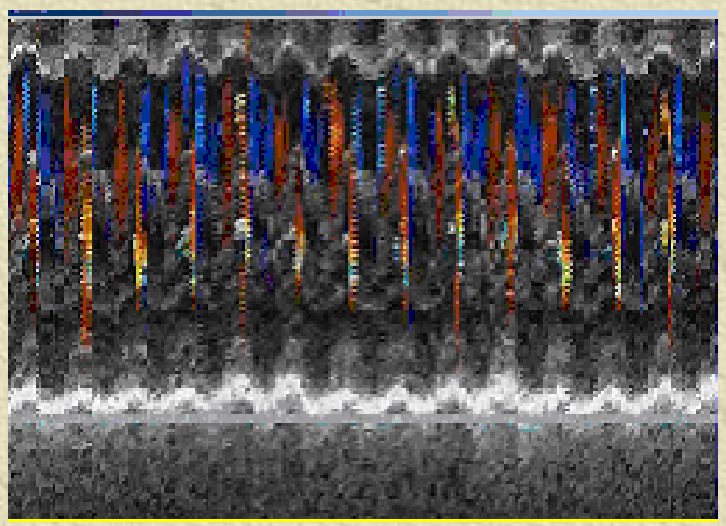
- Radial Color M-Mode generation
- Border extraction
- Color transition detection
- Sub-state identification

Radial Color M-mode (LAX)

The straight line is continuously swept by a specified angle keeping one end fixed.

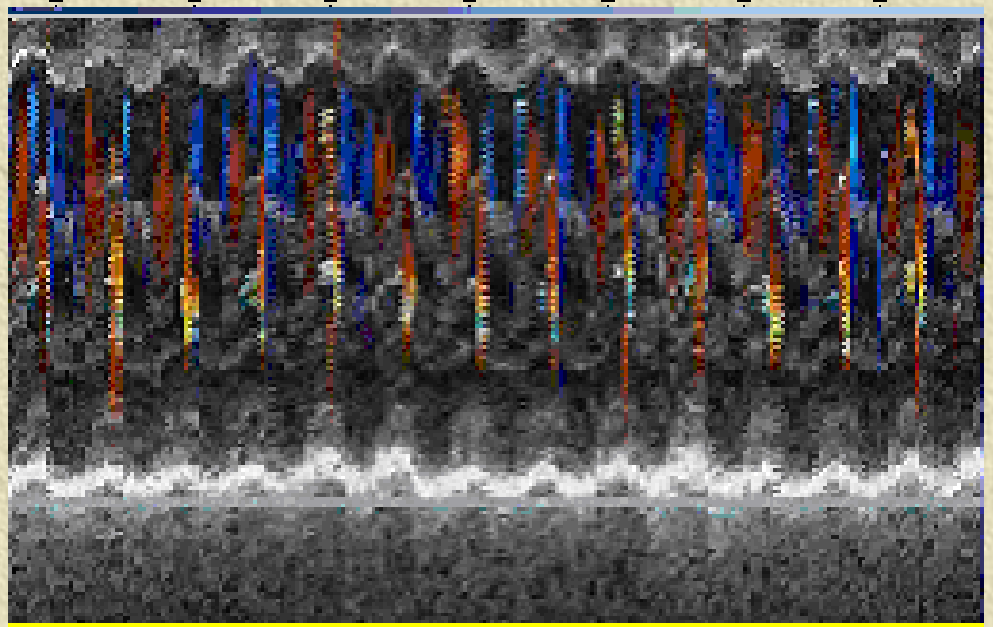


Radial M-modes are generated along the red straight lines.



Color M-mode image

Observations in Color M-mode (At mitral valve)

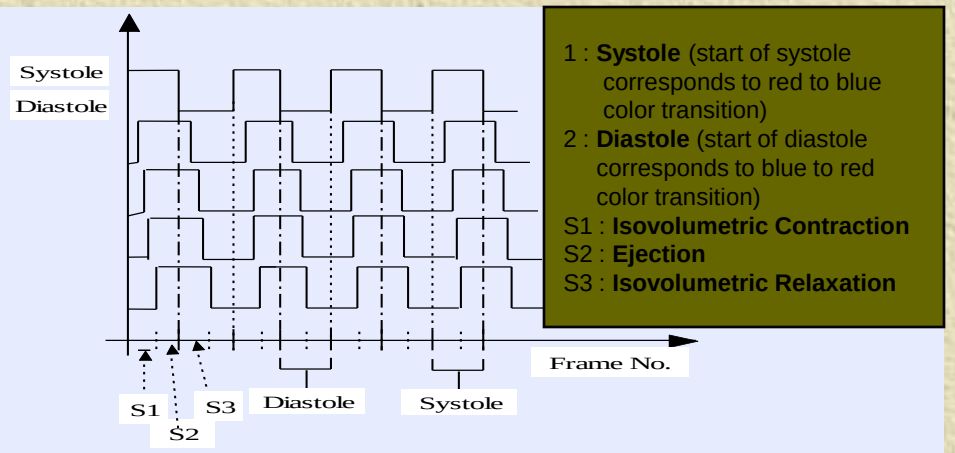


Blood flow in LV
(Upper portion)

Blood flow in LA
(Lower Portion)

- BLUE (-ve Flow, Away from probe)→ Systole
- RED (+ve Flow, Towards probe) → Diastole
- Color transition: State Transition

State Transition Diagram



State transition diagram of left ventricle

- As the blood flows from one end to another in the left ventricular cavity, there is a delay in color transition as we move from the aortic valve tip to mitral valve tip.
- The delay in diastole to systole transition is related to the isovolumetric contraction sub-state.
- Similarly, the delay in systole to diastole transition is related to isovolumetric relaxation sub-state.

Performance of Algorithms:

View Recognition

True Class	Predicted Class		
	Short Axis View (no. of frames)	Long Axis View (no. of frames)	Apical View (no. of frames)
Short Axis View	128	2	1
Long Axis View	1	80	1
Apical View	3	12	140

Classification accuracy : 95.34%

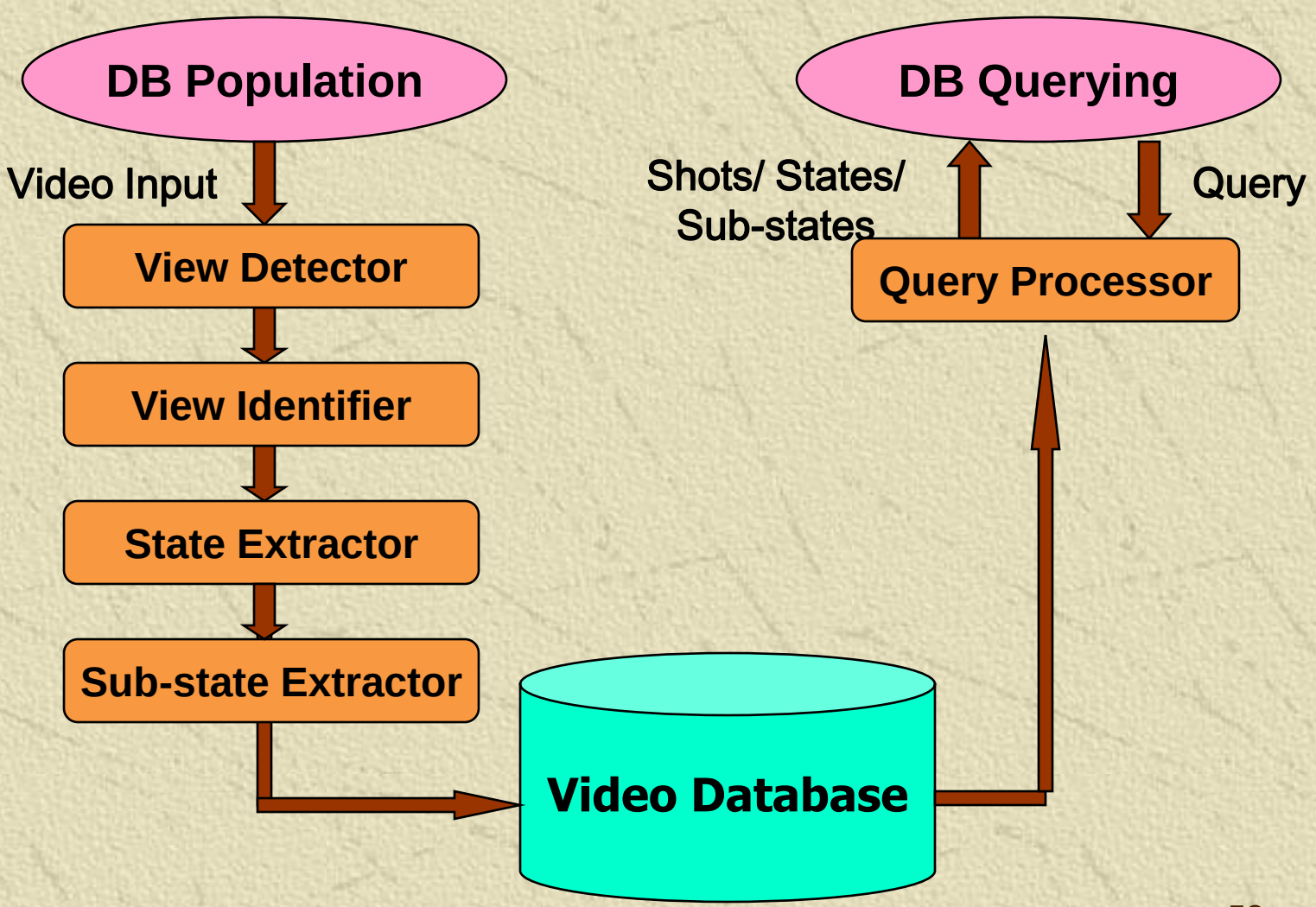
State Identification using M-mode

True Class	Predicted Class					
	Single M-mode			Sweep M-mode		
	Systole (no. of frames)	Diastole (no. of frames)	Undetected (no. of frames)	Systole (no. of frames)	Diastole (no. of frames)	Undetected (no. of frames)
Systole	96	5	21	111	10	2
Diastole	26	101	18	18	123	3

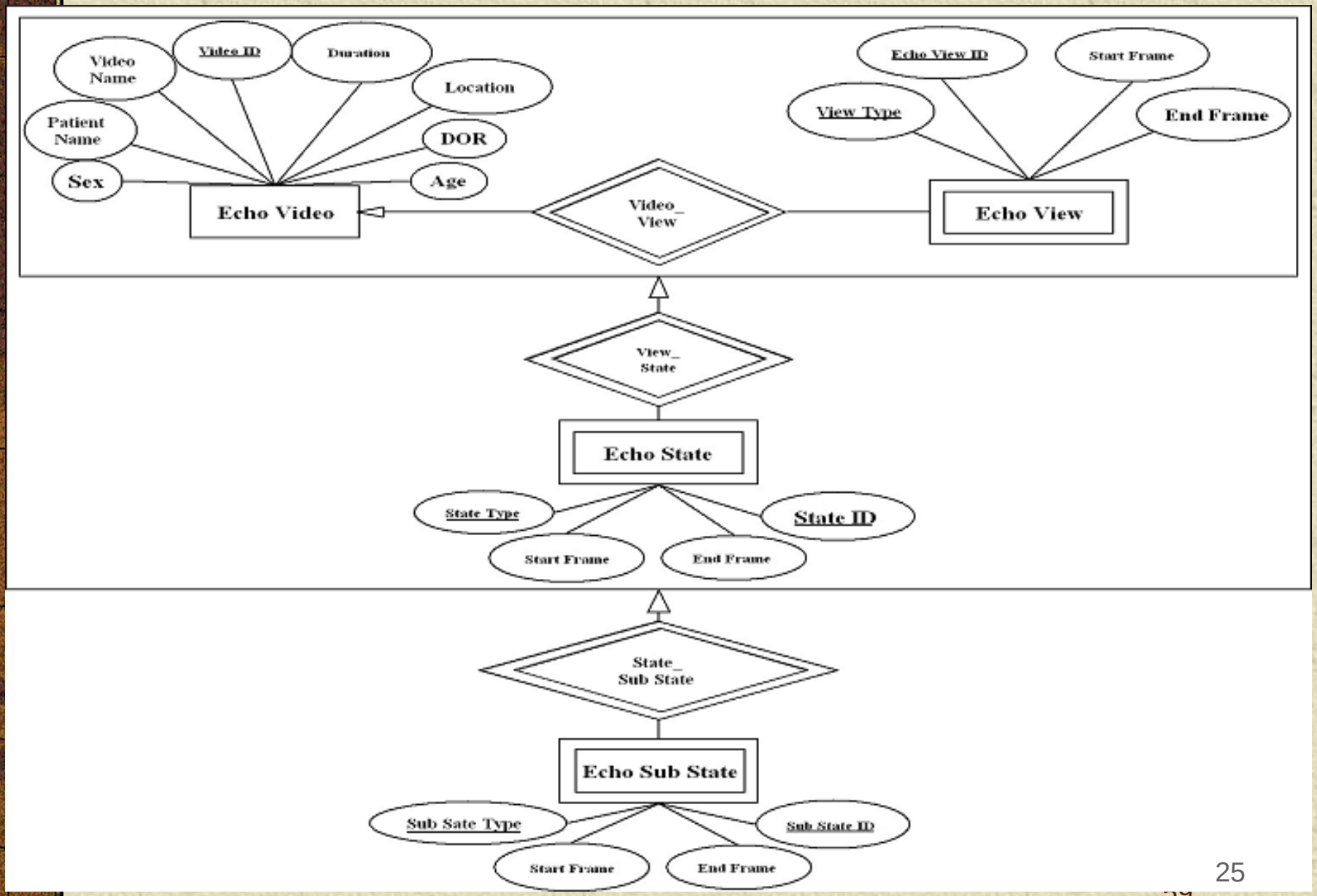
Misclassification error using single M-mode : 26.22%

Misclassification error using sweep M-mode : 12.36%

An Echo-cardiogram Video Archival System



ER Model


Video Name
Video ID
Duration
Location
Patient Name
DOR
Sex
Age
Echo Video
Video_ View
Echo View ID
Start Frame
View Type
End Frame
Echo View
View_ State
Echo State
State Type
State ID
Start Frame
End Frame
State_ Sub State
Echo Sub State
Sub Sate Type
Sub State ID
Start Frame
End Frame

State-based Retrieval (Example 1)

select *v* from *Vdb*

where

O in *v* and *O.name* = "left_ventricle"

and *O.state* = "expanding"

Meaning: Retrieve all segments *v* where the object left_ventricle is in state expanding

State-based Retrieval (Example 2)

select *v* from *Vdb*

where

O in *v* and *O.state* = " s_1"

and O^1 in (select v^1 from *Vdb*

where O^1.*state* = " s_2")

and O = O^1 and *reachable* ($O.s_1$, $O.s_2$)

Meaning: Retrieve all segments *v* that precede a given segment v^1 , both containing object O, such that state of O in v^1 reachable from state of O in *v*

Conclusion

- Use of a hierarchical state based model of dynamical cardiac system for extracting contextual information.
- It allows storage and indexing of the echo video.
- It helps browsing, searching and content based retrieval of video data.