Whole Body Image Parsing Using Machine Learning

S. Kevin Zhou, Ph.D.
Image Analytics and Informatics
Siemens Corporate Research
Princeton, NJ

with contributions from Siemens colleagues and clinical collaborators

Long Term Research Goal

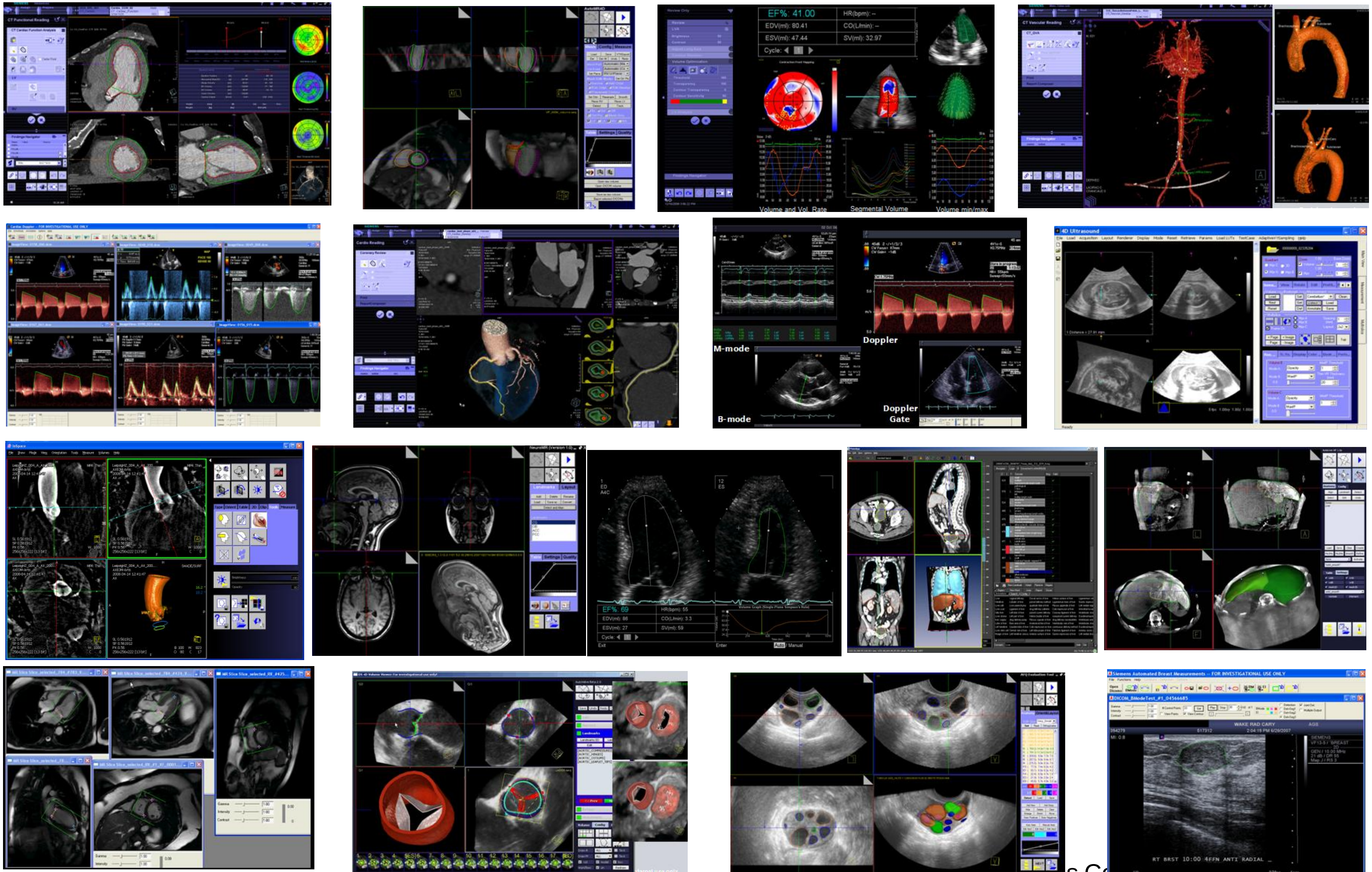
Image

Parsing

Anatomies

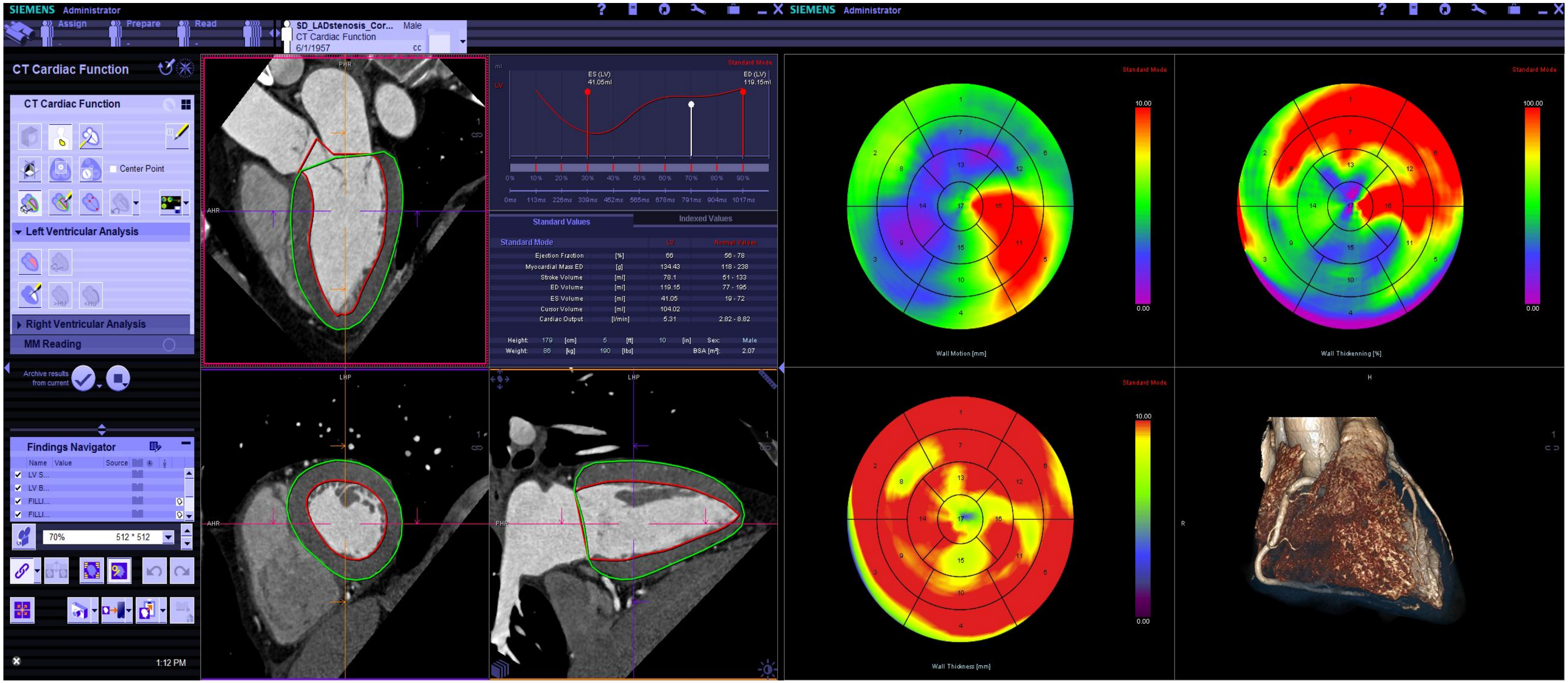
SIEMENS

SCR - Comprehensive Research on Biomedical Imaging



SIEMENS

Cardiac Function – Computed Tomography



• Presented in RSNA 2009

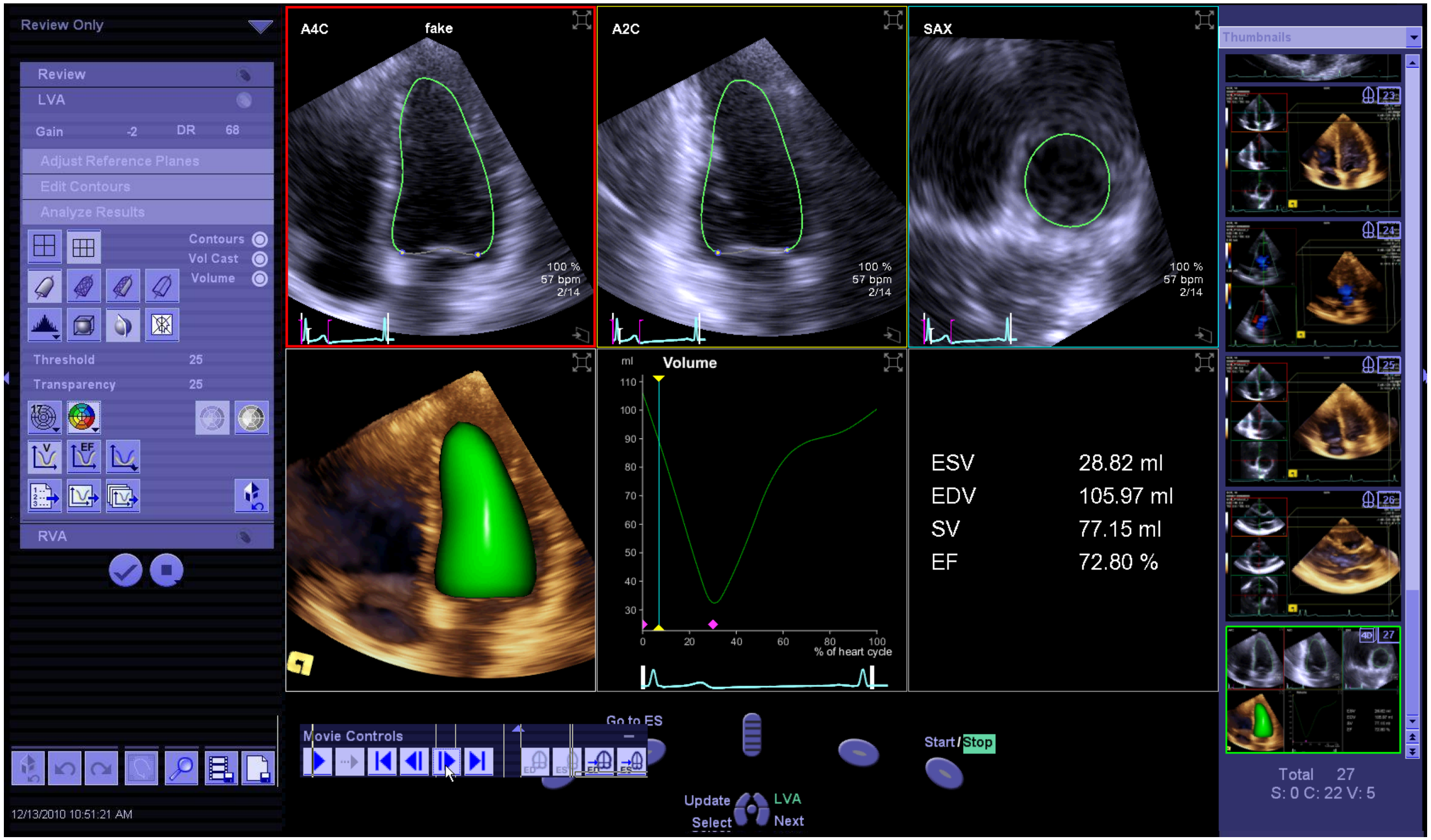
- Isolate the heart from the chest wall
- Quantify left and right ventricular ejection fraction and left ventricular mass

• Y. Zheng et al, Four-Chamber Heart Modeling and Automatic Segmentation for 3D Cardiac CT Volumes using Marginal Space Learning and Steerable Features, IEEE TMI, 2008

SIEMENS

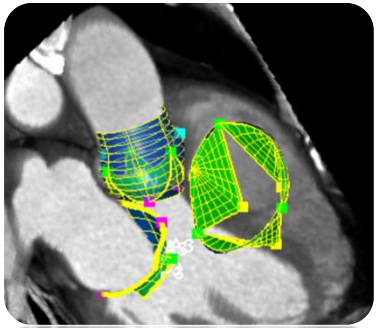
eSie Left Ventricle Analysis on SC2000

- Automatic navigation, detection, tracking and quantification of the left ventricle in 3D+T ultrasound imaging

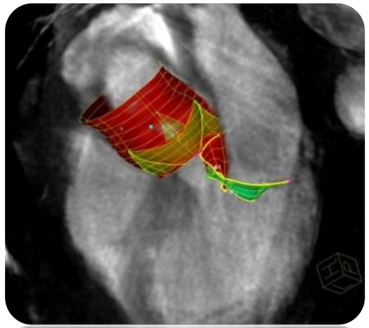


Work based on eSie LVA compared with MRI is one of the 5 Young Investigator's Award finalists presented at the American College of Cardiography (ACC) Scientific Sessions 2011

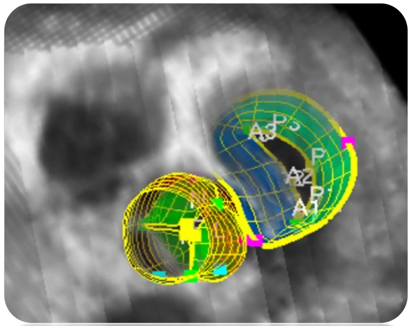
Patient-Specific Modeling of the Heart Valves



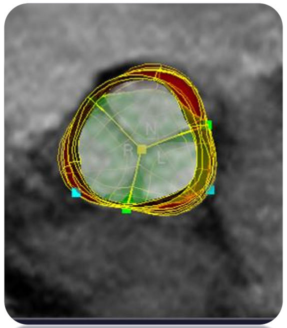
CT



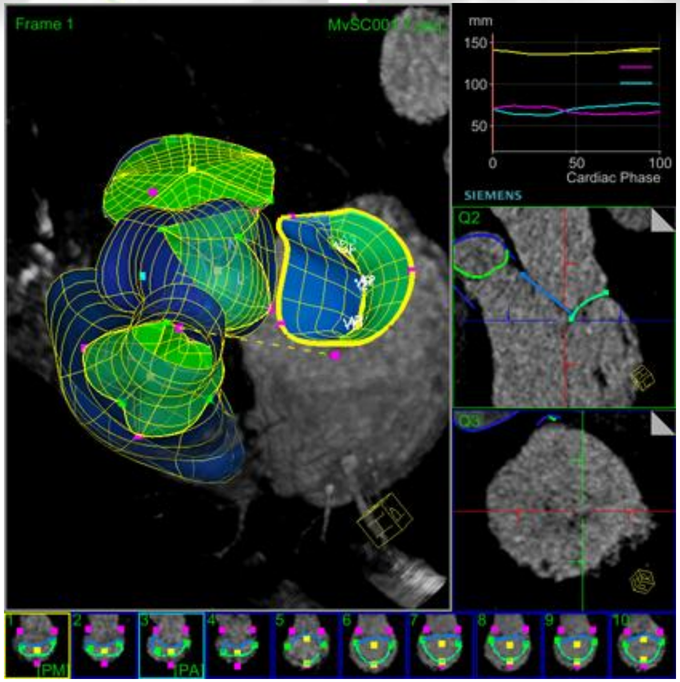
DynaCT

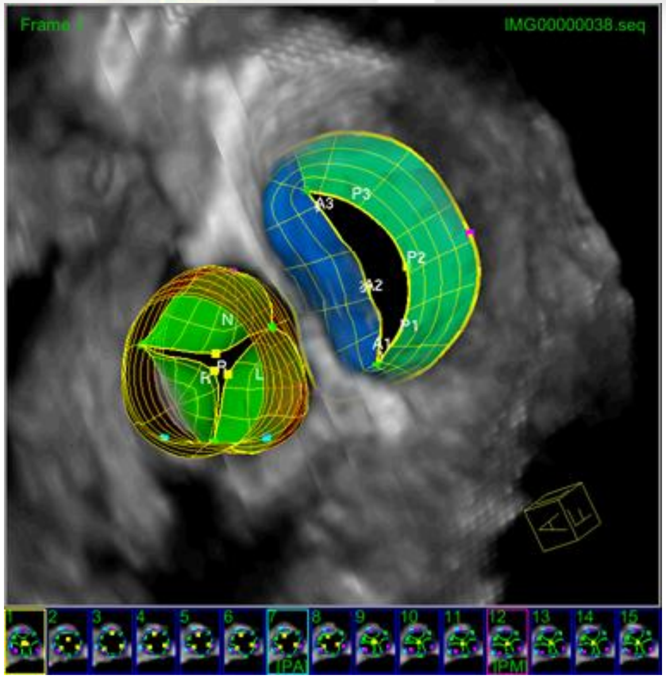


US (SC2000)



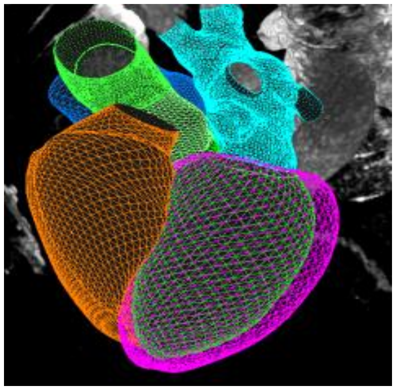
MRI

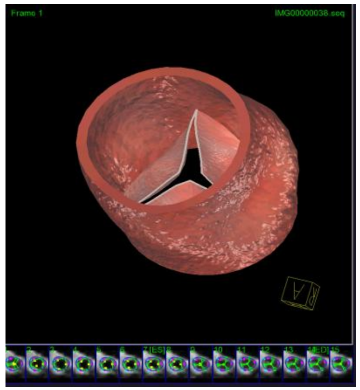




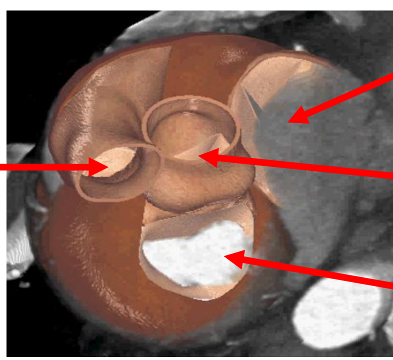
- Full valvular apparatus - aortic, mitral, pulmonary, tricuspid

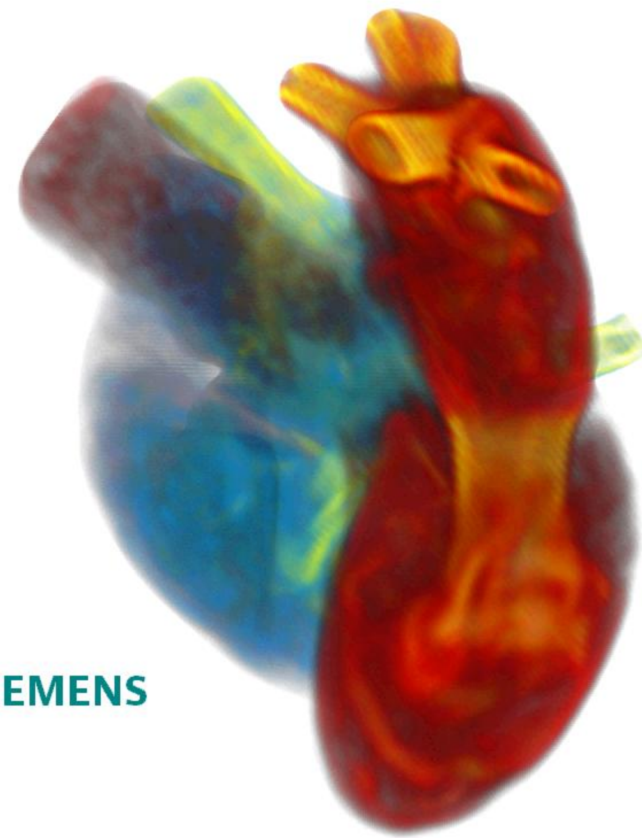
Heart Physiome: Computational Hemodynamics





4D CT

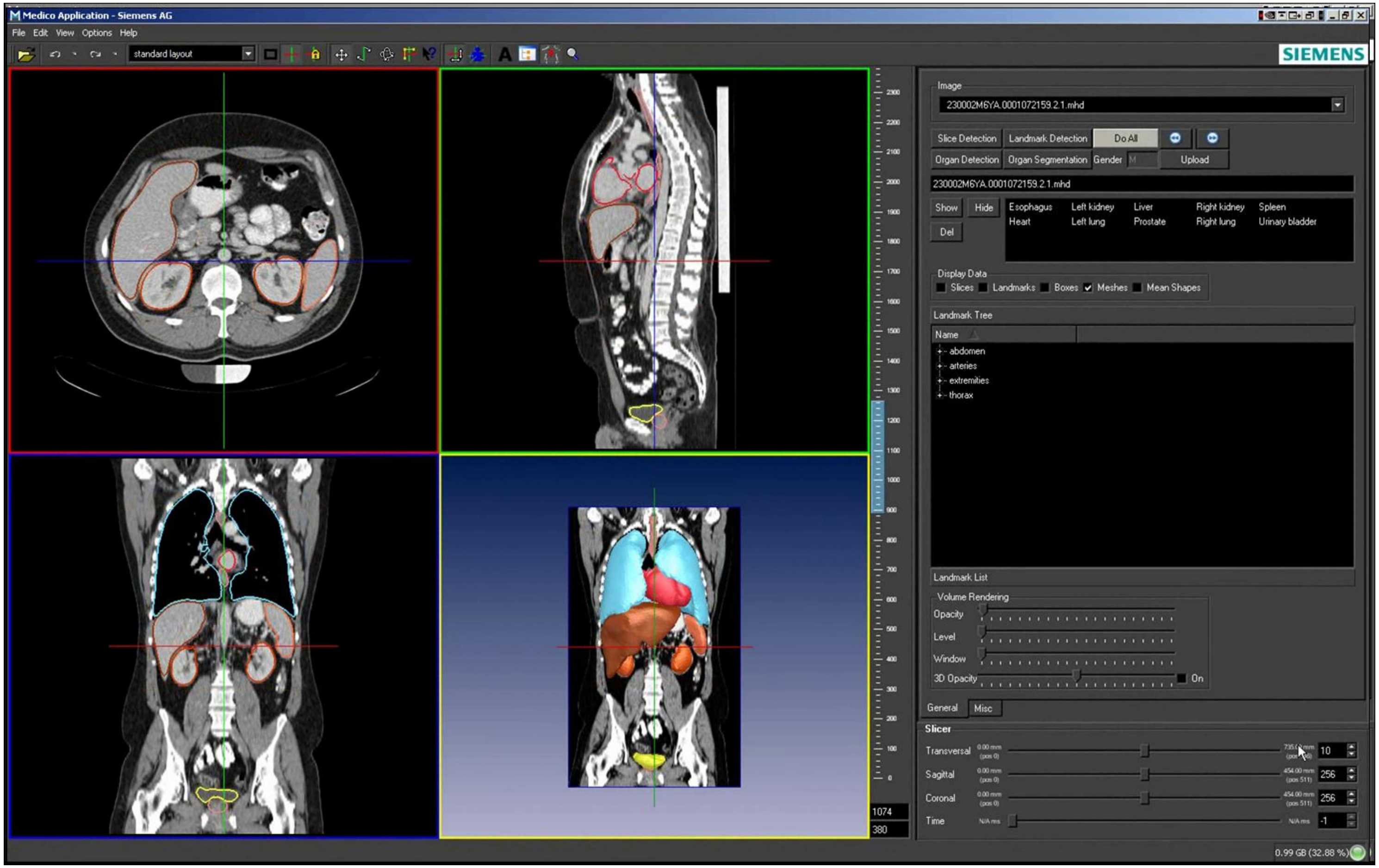




CFD Engine (level-set based)

Mihalef et al : Patient-Specific Modeling of Left Heart Anatomy, Dynamics and Hemodynamics from High Resolution 4D CT, Royal Society, 2011

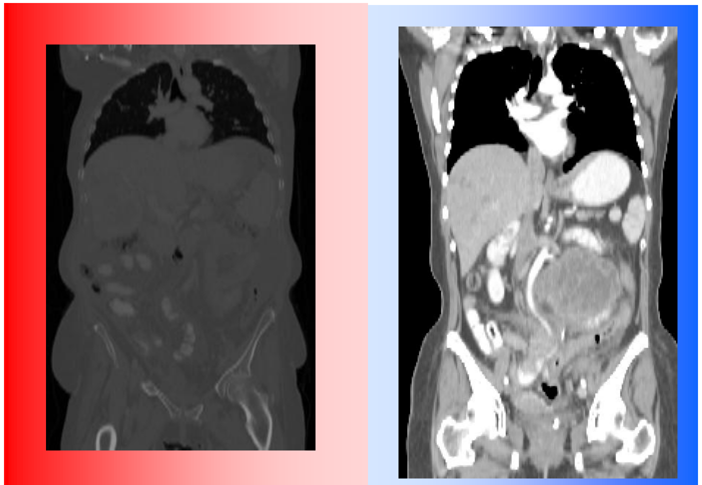
Automatic Volume Parsing and Metadata Indexing



• Volume parsing: Detects slides, 3D landmarks, organs, delineate organs

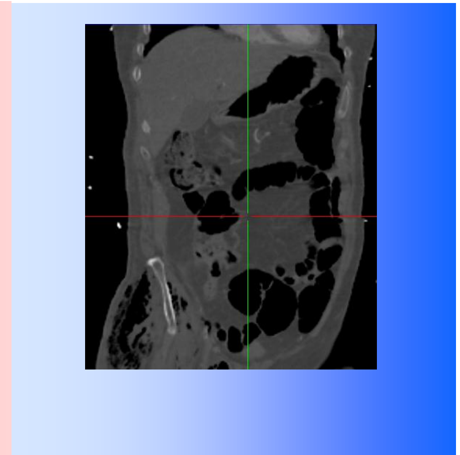
• S. Seifert et al: Semantic Annotation of Medical Images, Hierarchical Parsing and Semantic Navigation of Full Body CT Data, SPIE Medical Imaging, 2009-2010

Appearance variations



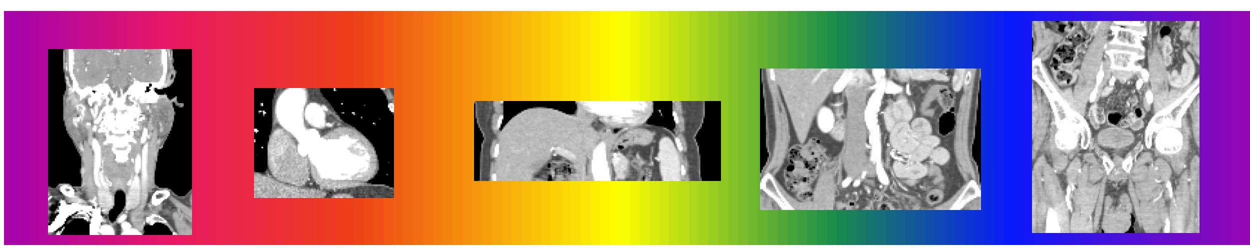
Contrast

Context



Pathologies

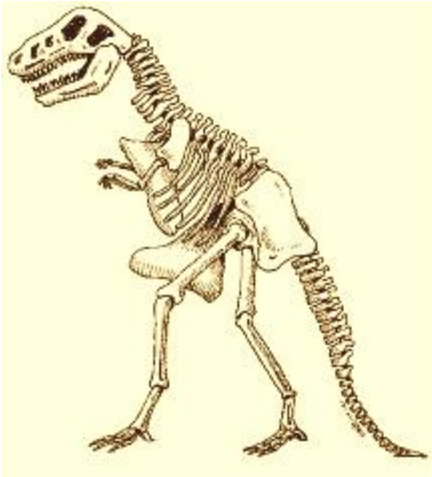
- Deformed
- Missing
- Misleading context



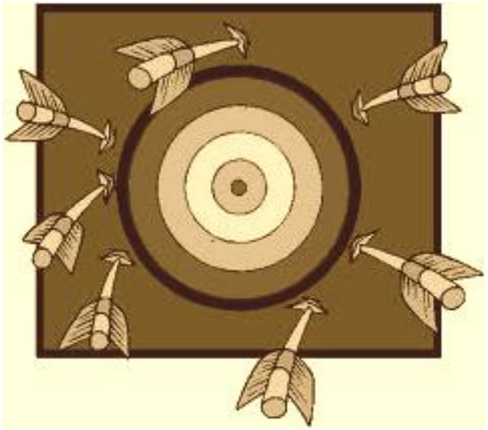
Body Portions

- Narrow FOV
- Occlusion

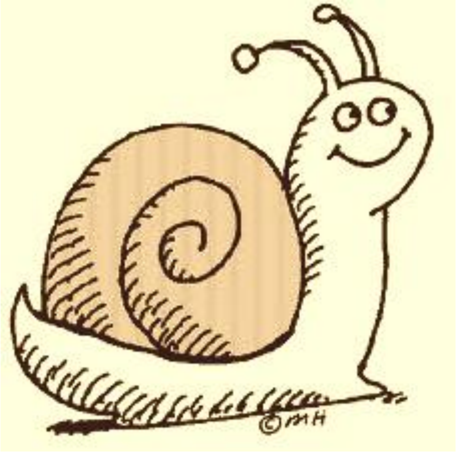
Challenges



Shape



Accuracy

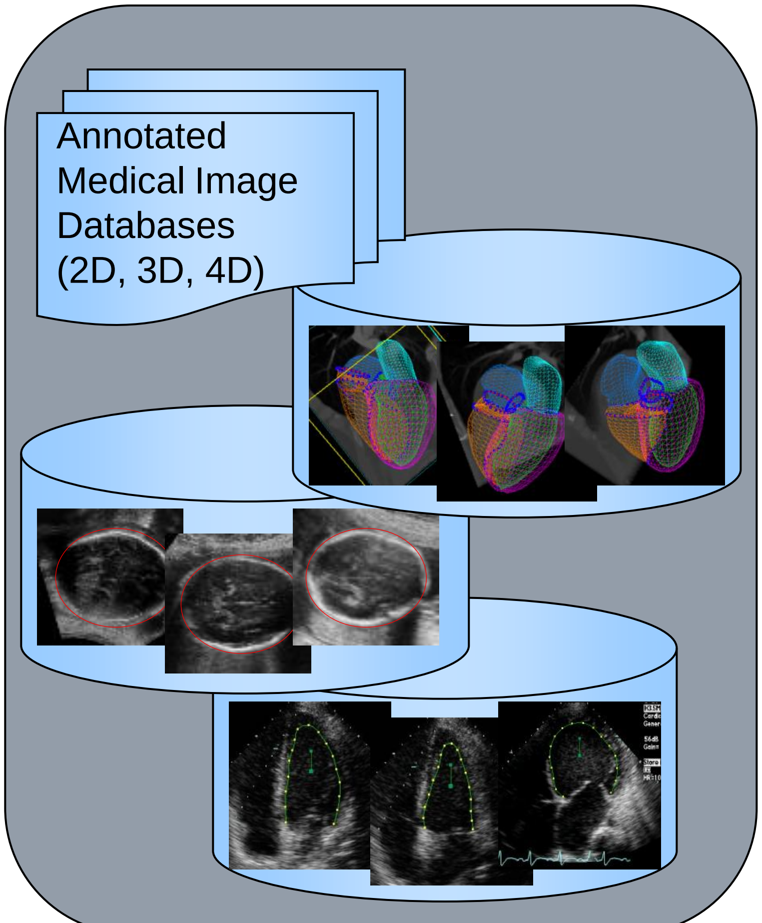


Speed

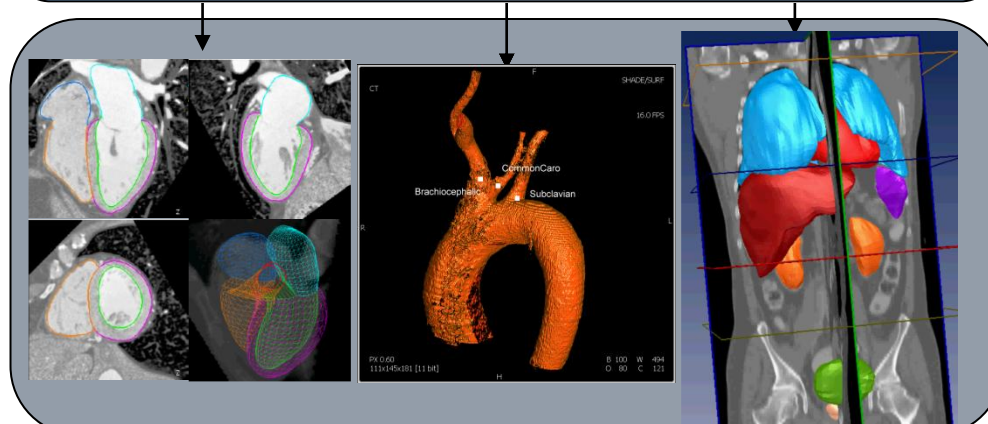
SIEMENS

Whole Body Analysis using Machine Learning

- Trainable solutions for fast, automatic landmark detection, organ labeling, segmentation, motion estimation and abnormally detection



- Discriminative Anatomical Network
- Probabilistic Boosting Tree, Random Forests
- Advanced Optimization – Random Walker
- Marginal Space Learning

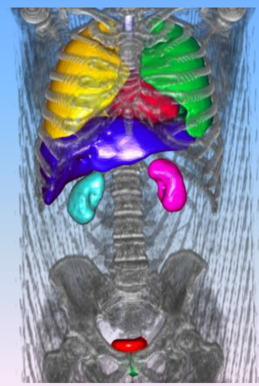


SIEMENS

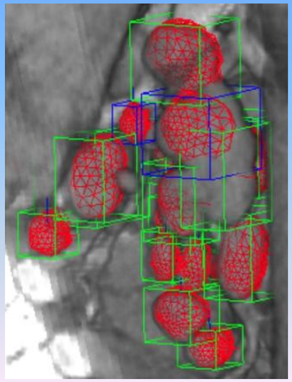
Outline

- WB Landmark

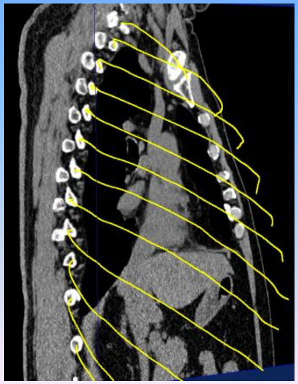
- WB Organ



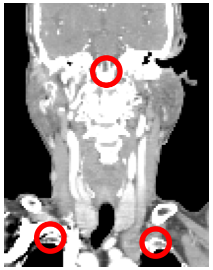
- WB Lesion



- WB Bone



Landmarks

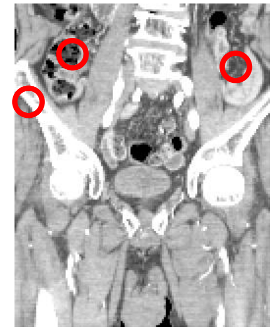


Skull Base
Lung Top

Trachea

Liver, Sternum

Liver, Hip



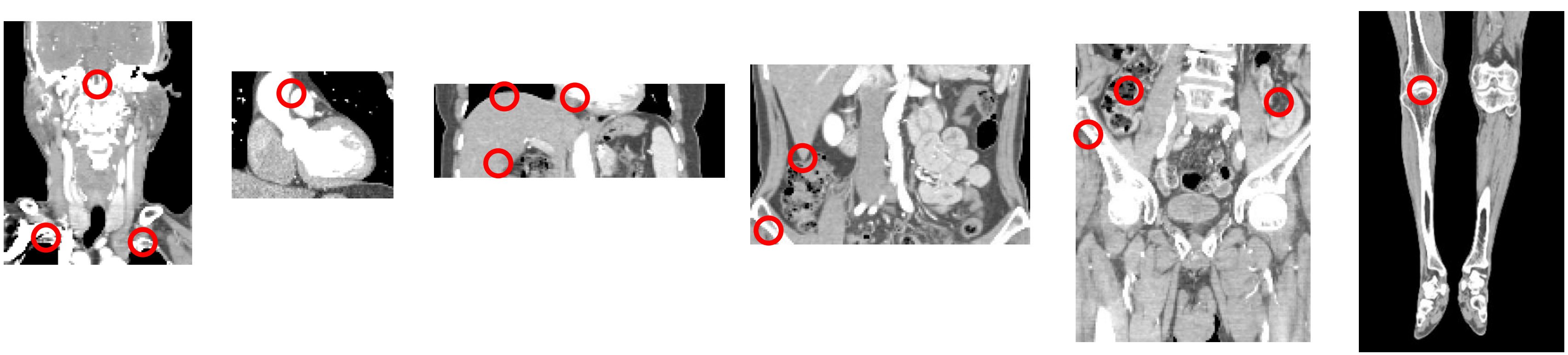
Hip, Kidney

Knee

238 landmarks

Characteristics

- **Used as pre-processing step to trigger other tasks in 3D**
- **Handles objects with large appearance variability**
- **Accuracy**
  - Near zero false negative and false positive rates
- **Speed**
  - Has to be fast as the first step.



Independent vs. Sequential Search

- **Independent Search**
  - Ignore the spatial relationship between landmarks
  - Computational complexity linearly depends on the volume size, the classifier complexity and the number of landmarks. SLOW

- **Sequential Search**
  - Leverage the spatial relationship between landmarks
  - Break down the linear dependency on volume size. FAST
  - Questions: What is the optimal search order?

Determining the Search Order

- **Exhaustively evaluating the search order in 1 volume**
  - 12 landmarks: 12! x 1sec /60/60/24/365 > 15 years
  - 43 landmarks: 43! x 1sec /60/60/24/365 > 1045 years

- **Difficult to find the best search order even offline**

- **The landmarks could be missing.**

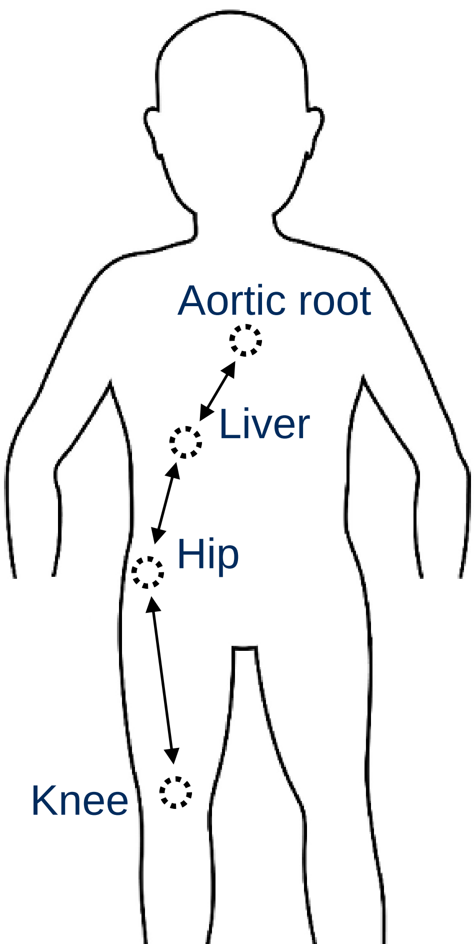
SIEMENS

"Greedy Search" for Fast Detection [Liu et al. CVPR 2010]

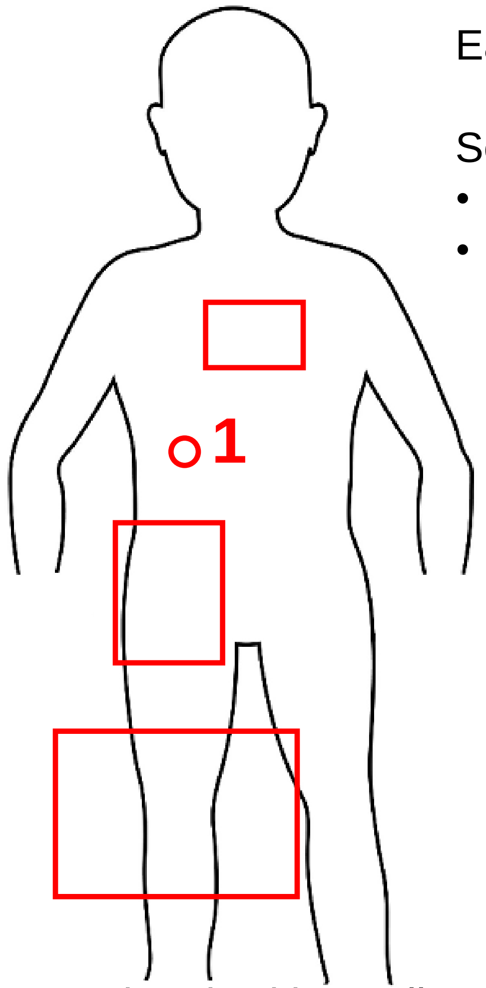
Each box indicates a search window.

Search window could incorporate

- Training data statistics
- Landmark classifier complexity

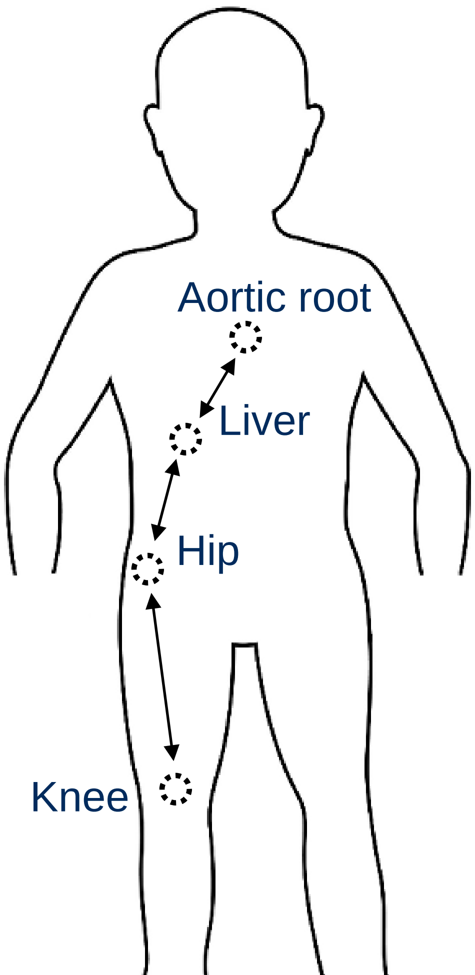


Spatial relationship between landmarks provided by training data

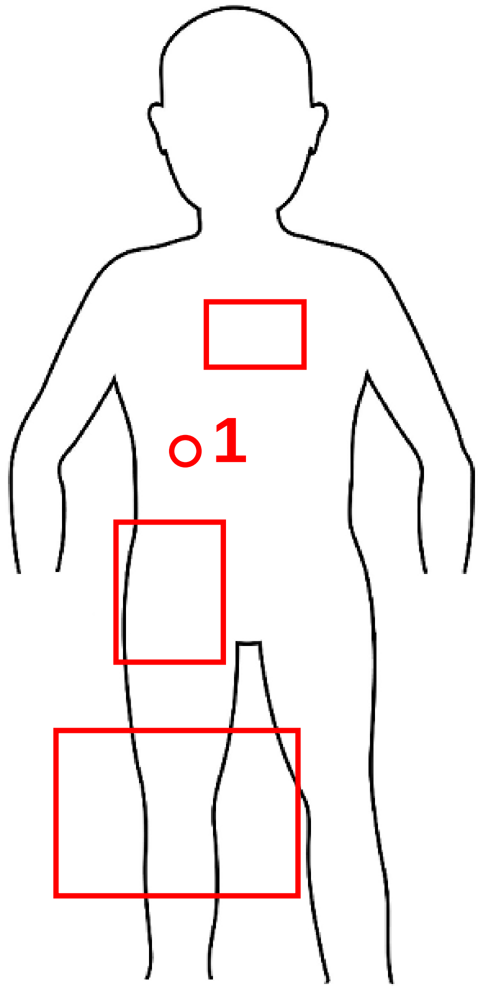


Landmark with smallest search range: Aortic root

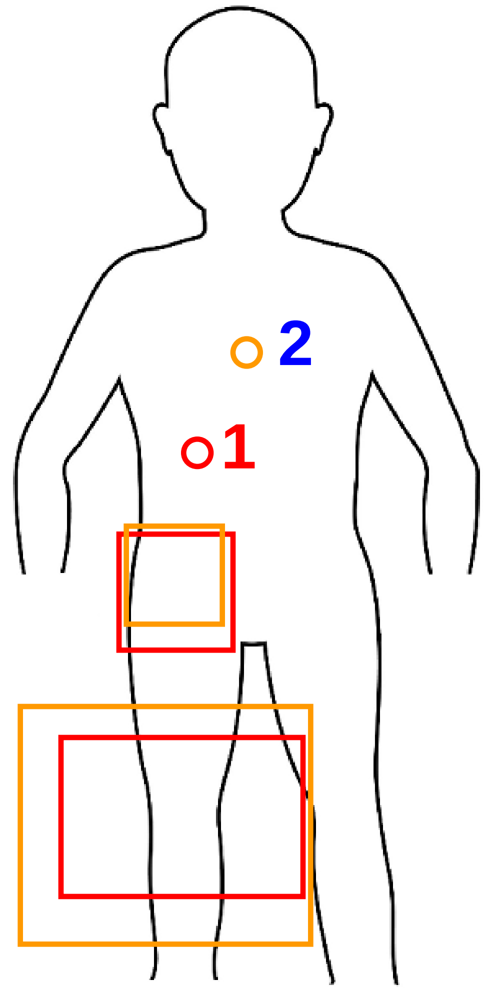
“Greedy Search” for Fast Detection



Spatial relationship between landmarks provided by training data



Landmark with smallest search range: Aortic root



Landmark with smallest search range: Liver

SIEMENS

Algorithm

- **In each round of the greedy algorithm, each detected landmark d provides a search space V_{ud} for each undetected landmark u**

$$d \in S_{(1):(k)} = \{l_{(1)} \prec l_{(2)} \prec \ldots \prec l_{(k)}\}$$

- **Each un-detected landmark selects the smallest search space**

$$\forall u \,, \quad V_u(S_{(1):(k)}) = \min_{d \in S_{(1):(k)}} V_{ud}$$

- **The un-detected landmark that has the smallest search space is chosen, and the cost is**

$$C_{k+1}(S_{(1):(k)}) = \min_u V_u(S_{(1):(k)})$$

- **This algorithm approximately solves**

$$\min \sum_{k=2}^{N} C_k(S_{(1):(k-1)})$$

SIEMENS

Submodular Maximization

Define $F_k(S) = C_k(\phi) - C_k(S)$

$$\min \sum_{k=2}^{N} C_k(S_{(1):(k-1)}) \rightarrow \max \sum_{k=2}^{N} F_k(S_{(1):(k-1)})$$

$F_k(.)$ **is a submodular function**

$$F_k(S \cup \{l\}) - F_k(S) \geq F_k(T \cup \{l\}) - F_k(T) \qquad \forall S \subseteq T$$

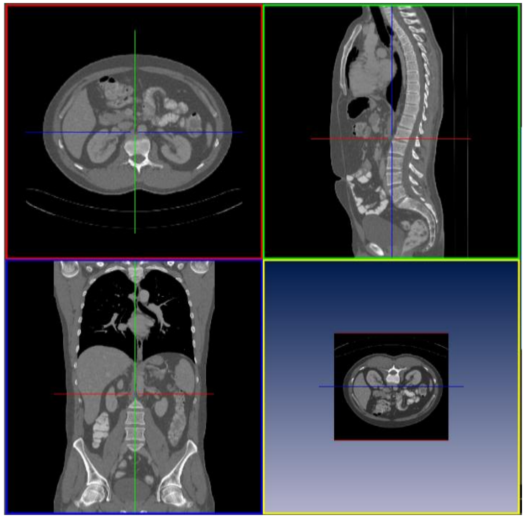
- **Theorem: If** F **is a submodular, nondecreasing function and** $F(\phi) = 0$**, then the greedy algorithm finds a set** S' **such that**

$$F(S') \geq (1 - 1/e) \max F(S)$$

- **Approximation reaches at least 63% of optimal solution (off-line bound)**

“Greedy Search” is adaptive

SIEMENS



23000GpSKp.11.1.detRaw

1 LiverTop

→ Skull

→ 2 FemurHeadR (1)

→ 3 HipR (2)

→ 4 HipL (2)

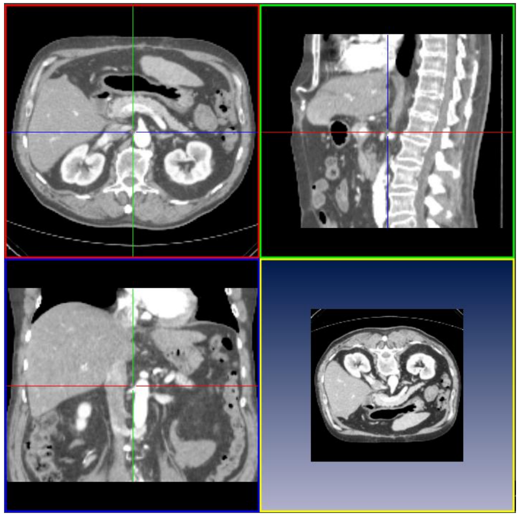
→ 5 KidneyR (2)

→ 6 KidneyL (2)

→ 7 LungTopL (4)

→ 8 LungTopL (7)

…



MZ-143330-CTSV_3042.3874.5.1.detRaw

1 LiverTop

→ Skull

→ FemurHeadR

→ 2 AortaRoot (1)

→ TracheaBif (1)

→ LungTopL (1)

→ 3 SternumBot (1)

→ LiverCent (1)

→ 4 KidneyL (1)

…

Time vs. Volume Size

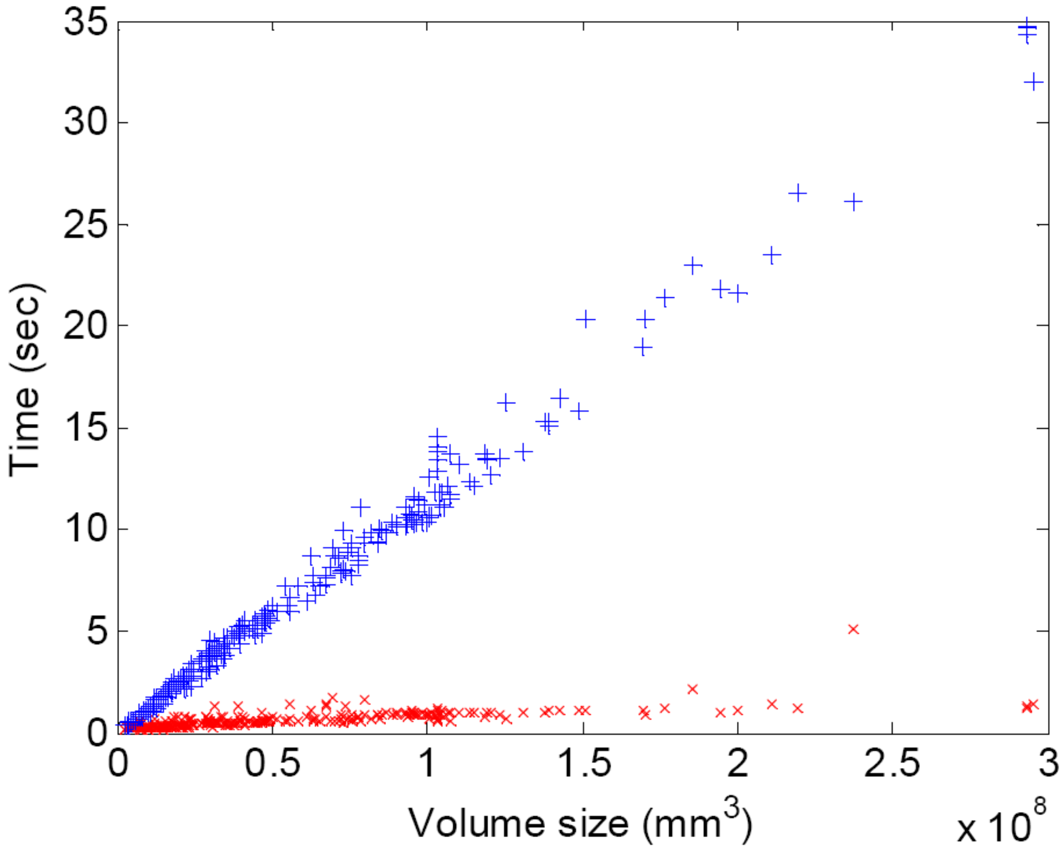


Figure 4. Detection time as a function of volume size. Blue (+): independent landmark detectors. Red (x): Greedy search.

SIEMENS

Detection Time

	mean	std	Q95	max
Independent D_{8mm} $N = 63$	17.30	6.16	46.24	84.51
Greedy D_{8mm} $N = 63$	$\boxed{1.14}$	0.47	1.92	2.44
Independent D_{8mm} $N = 25$	6.72	6.40	17.73	35.00
Greedy D_{8mm} $N = 25$	$\boxed{0.65}$	0.43	1.26	5.08

FPR and FNR

	Greedy		Independent	
	FP_A	FN_A	FP_B	FN_B
SkullBase	**0 (193)**	0 [50]	1 (192)	0 [50]
R.LungTop	**0 (84)**	1 [114]	1 (83)	1 [114]
LiverDome	0 (86)	2 [65]	0 [86]	2 [65]
R.HipTip	**0 (131)**	0 [94]	1 (130)	0 [94]
R.Knee	0 (265)	0 [12]	0 (265)	0 [12]
LiverBott.	2 (33)	1 [33]	2 (33)	1 [33]
TracheaBif.	0 (44)	0 [41]	0 (44)	0 [41]
LiverCent.	**0 (90)**	1 [136]	2 (88)	1 [136]
L.HumerusHead	0 (96)	1 [12]	0 (96)	1 [12]
R.HumerusHead	1 (80)	2 [7]	1 (80)	2 [7]
L.LungTop	**0** (61)	1 [21]	1 (61)	1 [20]
L.HipTip	**0 (94)**	**1 [46]**	2 (92)	2 [45]
L.FemurHead	0 (124)	0 [16]	0 (124)	0 [16]
R.FemurHead	0 (120)	0 [16]	0 (120)	0 [16]
CoccyxTip	0 (118)	0 [16]	0 (118)	0 [16]
PubicSymph.Top	0 (133)	0 [23]	0 (133)	0 [23]
SternumTip	3 (51)	1 [22]	3 (51)	1 [22]
AortaBend	**0 (31)**	1 [53]	1 (30)	1 [53]
Brachioceph.	1 (35)	3 [132]	1 (35)	3 [132]
R.Kidney	2 (59)	5 [61]	2 (59)	5 [61]
L.Kidney	0 (71)	0 [76]	0 (71)	0 [76]

MMBIA 2012

IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)

January 9th, 2012 **Breckenridge, Colorado**

http://www.mmbia.org/mmbia2012

Important Dates

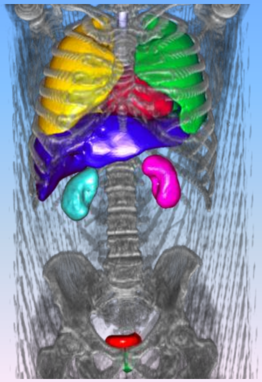
Paper Submission Deadline: **September 12th, 2011**
Paper Decisions: October 31st, 2011
Final Papers Due: December 1st, 2011

Outline

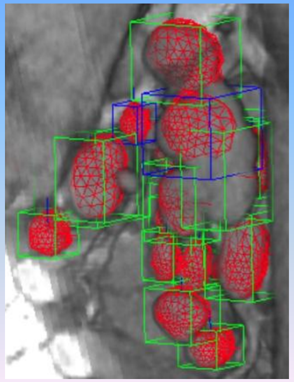
SIEMENS

- WB Landmark

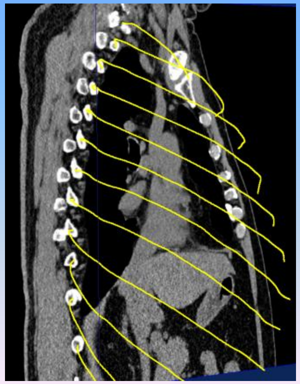
- WB Organ



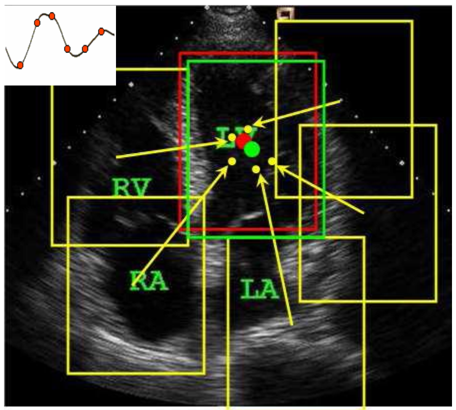
- WB Lesion



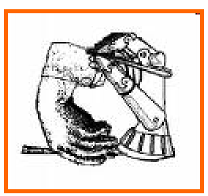
- WB Bone



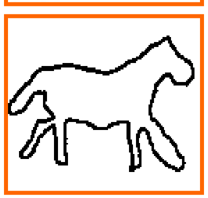
Shape Regression Machine (SRM) [Zhou MIA 2010]



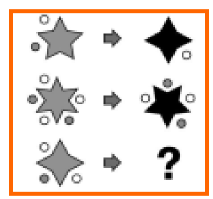
Efficient deformable shape segmentation



Learning: Regression



Annotation: Full shape

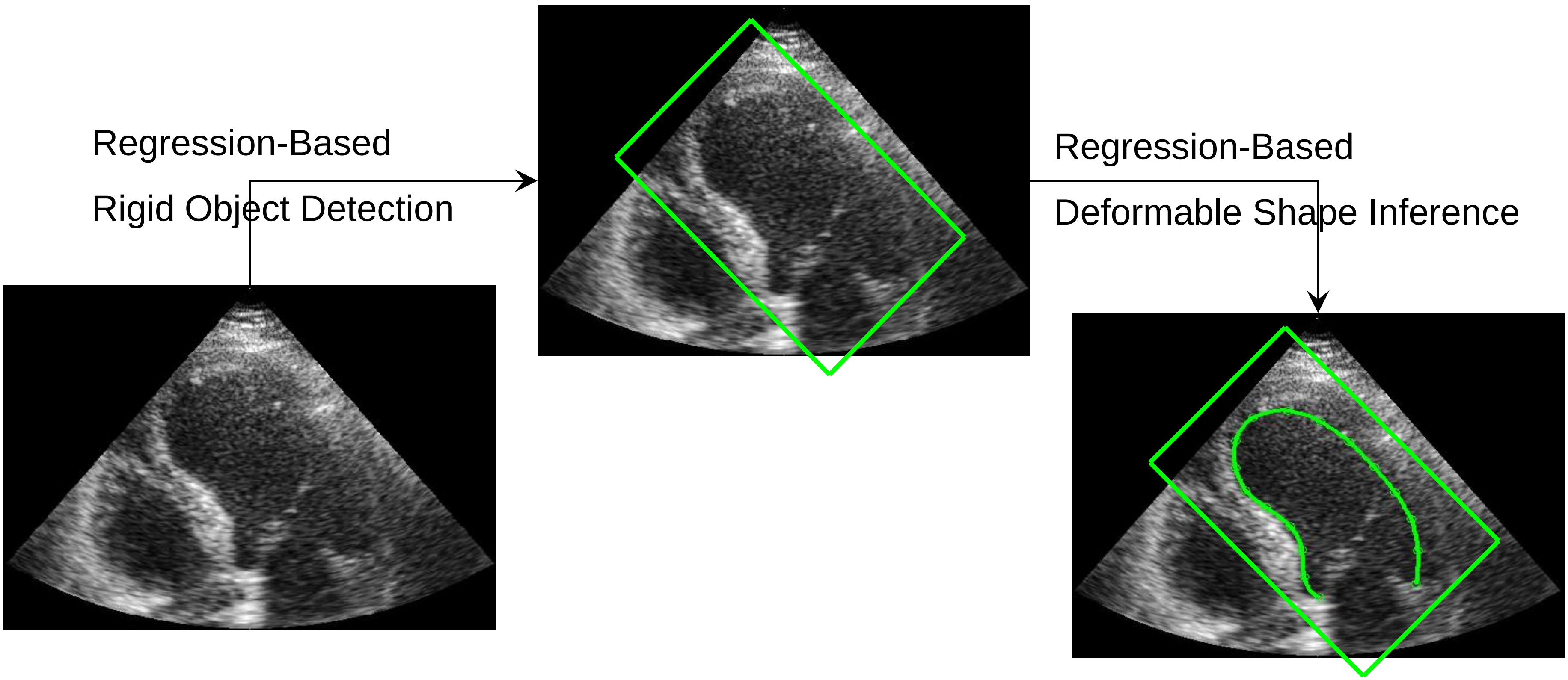


Inference: Sample averaging

Context: Shape, anatomy, appearance

Shape Representation & Two-Stage Approach

- **Shape *C* = rigid θ + deformable *S***
  - For LV endocardium, $\theta = (t_x, t_y, log(s_x), log(s_y), \alpha)$
  - *S* consists of a cohort of landmarks $(x_1, y_1, x_2, y_2, \ldots, x_N, y_N)$



Object Detection and Context

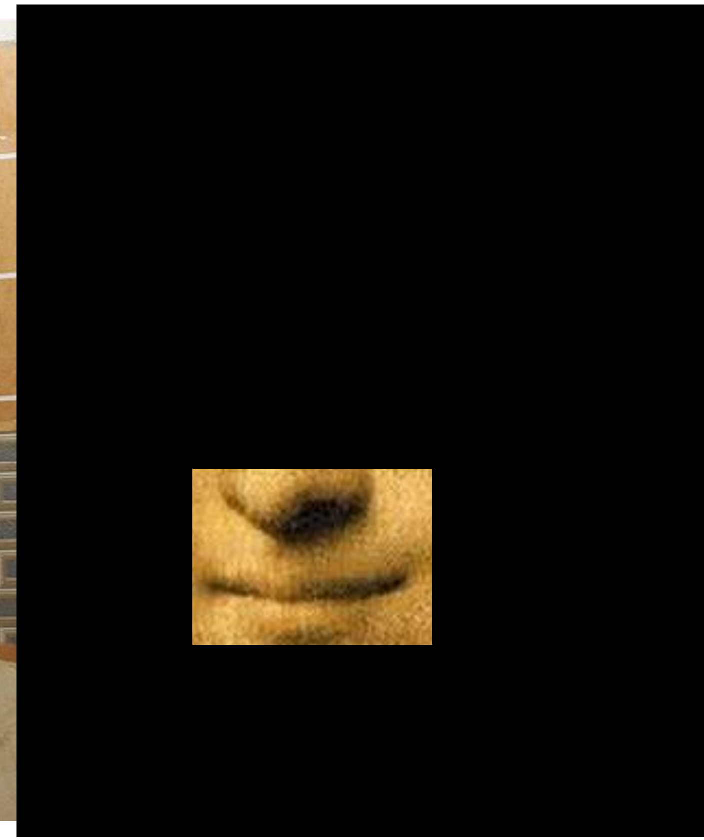
“I Spy”

no context

Painting by Picasso

weak context

Mona Lisa



strong context

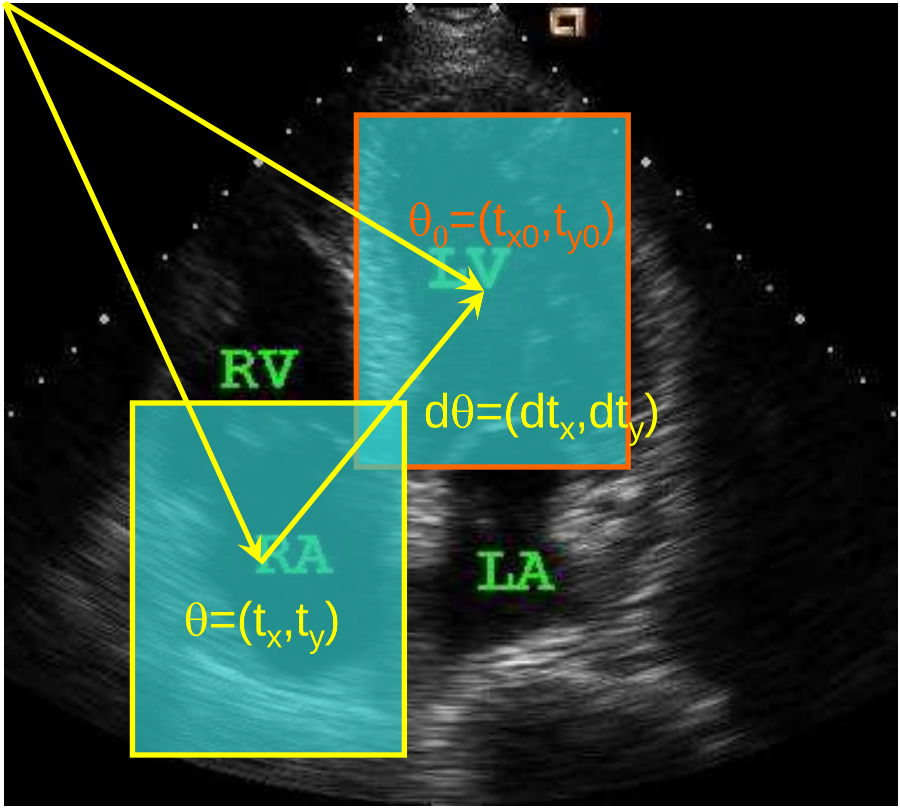
Regression-Based Object Detection: Basic Idea

- **Basic idea**
  - Regress the difference vector

$$d\theta = \mathcal{F}_1(\ I(\theta)\ )$$

  - Estimate the ground truth

$$\underline{\theta}_0 = \theta + d\theta = \theta + \mathcal{F}_1(\ I(\theta)\ )$$

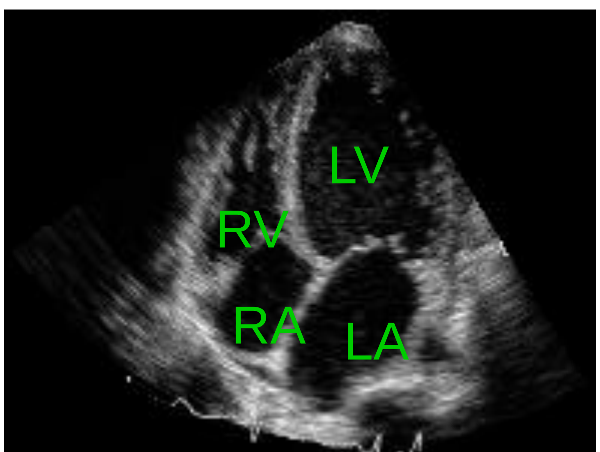


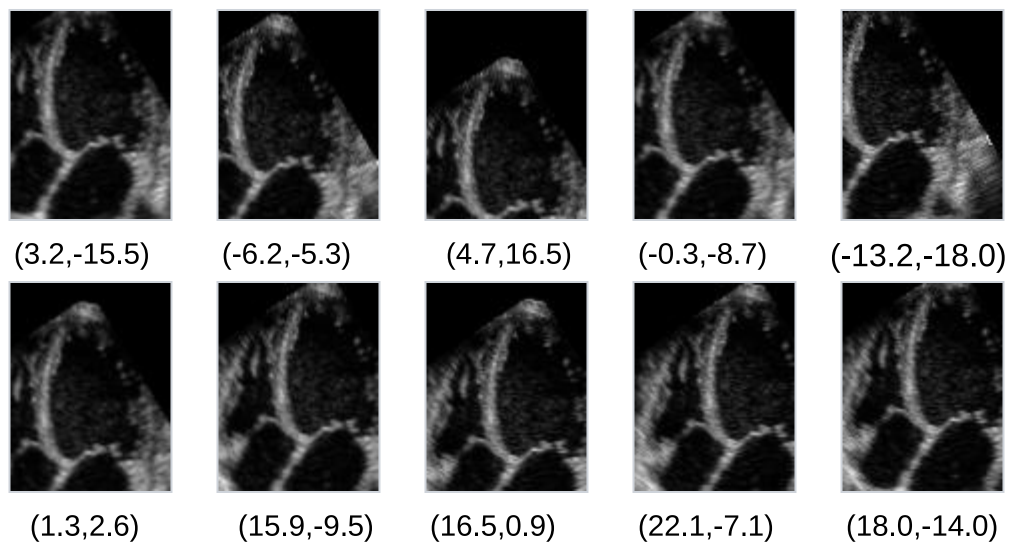
One scan solution!

SIEMENS

Two Questions?

- **Does such an oracle $\mathcal{F}_1$ exist?**
  - Context in anatomy and appearance at a *global* level
- **How to learn the oracle $\mathcal{F}_1$?**
  - Annotated database & machine learning





Robust Detection

- **Algorithm**
  - Sample

    $\{\theta^{<1>}, \theta^{<2>}, \ldots, \theta^{<M>}\}$

  - Estimate

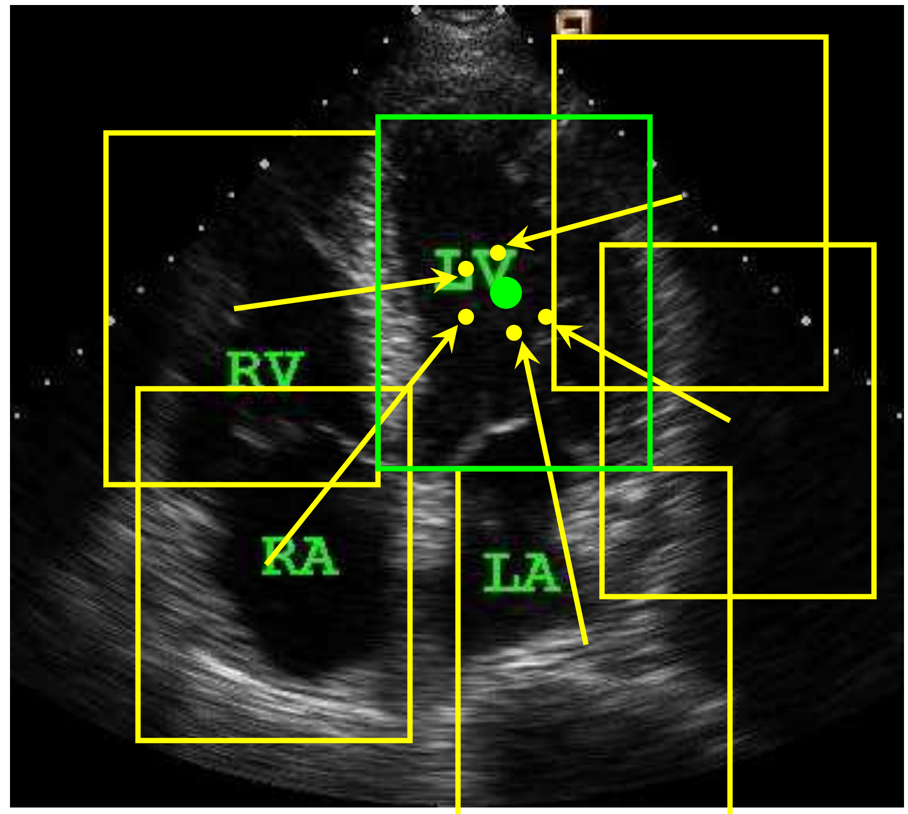
    $d\theta^{<m>} = \mathcal{F}_1(\, I(\theta^{<m>})\,)$

  - Predict

    $\underline{\theta}_0^{<m>} = \theta^{<m>} + d\theta^{<m>}$

  - Fuse by averaging

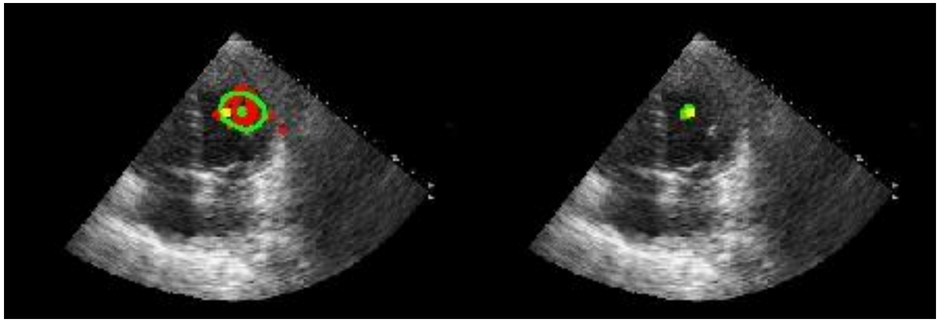
    $\underline{\theta}_0 = M^{-1}\, \Sigma_{m=1:M}\, \underline{\theta}_0^{<m>}$



SIEMENS

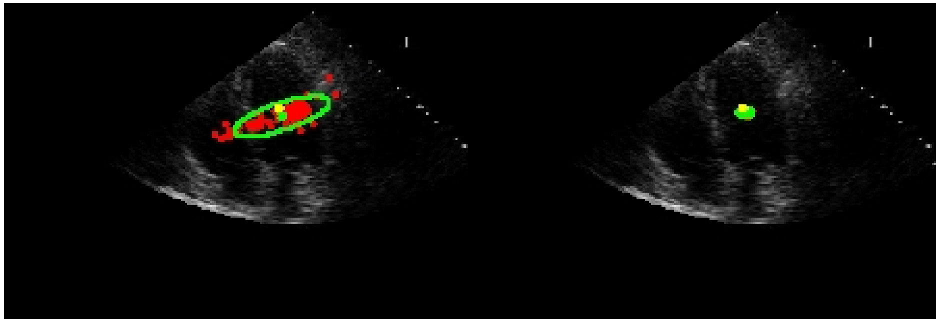
Improved Localization

- **Confidence score**
  - Train a binary classifier
  - p_d is the posterior prob. from the binary classifier

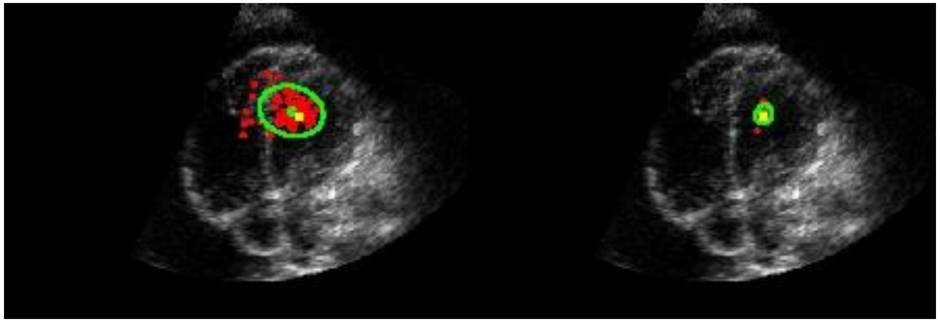


- **Weighted averaging**

$$\underline{\theta}_0 = \frac{\sum_j p_d^{<j>} \theta_0^{<j>}}{\sum_j p_d^{<j>}}$$

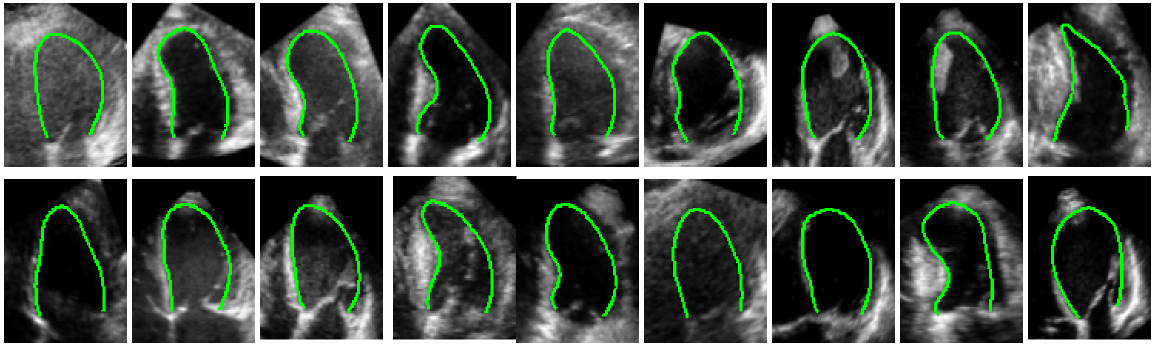


- **Faster computation**
  - Early stop



Regression-Based Deformable Shape Inference

- Basic idea $S = \mathcal{F}_2(I(\underline{\theta}_0))$. $I(\underline{\theta}_0)$: *Estimated ground truth patch*
- Does such an oracle $\mathcal{F}_2$ exist?
  - Context in shape and appearance at a *local* level



- How to learn the oracle $\mathcal{F}_2$?
  - Annotated database & machine learning
  - Perturb the rigid parameter to allow imperfect detection

Deformable Shape Inference Algorithm

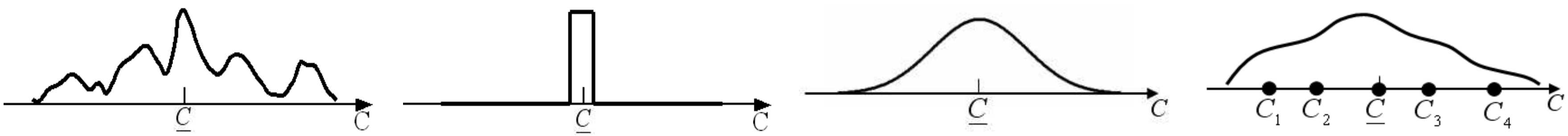
- **Algorithm**
  - Sample
    - Perturb the bounding box to generate K random samples $\{I^{<1>}, I^{<2>}, \ldots, I^{<K>}\}$
  - Estimate
    - $S^{<k>} = \mathcal{F}_2(I^{<k>})$
  - Fuse
    - Build a nonparametric kernel density $p_s(S)$
    - Weighted averaging

$$\underline{S} = \frac{\sum_k p_s^{<k>} S^{<k>}}{\sum_k p_s^{<k>}}$$

Discriminative Learning for Deformable Shape Segmentation [Zhang et al. ECCV 2008]

SIEMENS

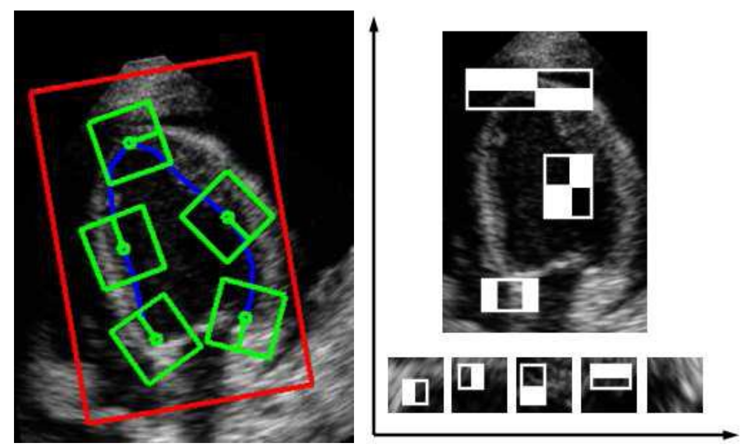
- **Learn a score function s(I, C) using classification, regression, or ranking.**



Energy minimization | Classification | Regression | Ranking

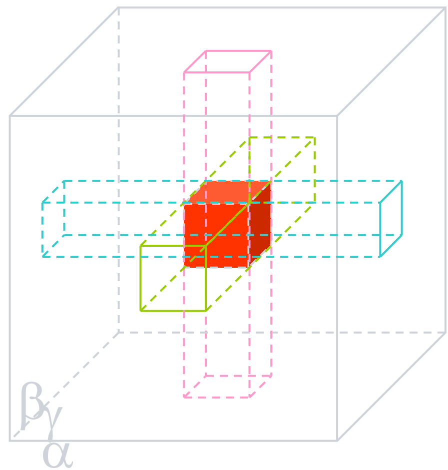
- **Maximize the score function using standard optimization methods (e.g., simplex).**

- **Better feature representation.**

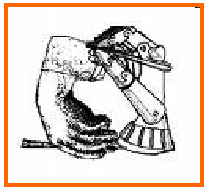


SIEMENS

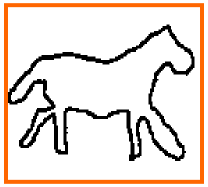
Marginal Space Learning (MSL) [Zheng et al. TMI 2008]



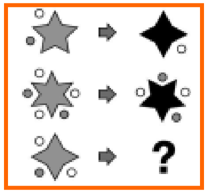
- **Efficient anatomy detection from 3D volumes**
- **Rigid parameterization (9D)**
  - 3 for translation α
  - 3 rotation β
  - 3 for anisotropic scale γ



Learning: Binary classification



Annotation: Bounding box



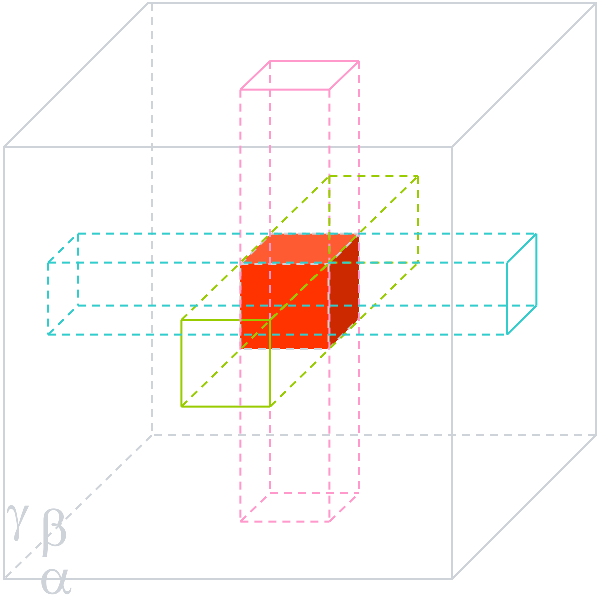
Inference: Exhaustive scanning

Context: Shape & appearance

Classification-based Object Detection [Voila & Jones]

- **Object detection: MAP in the search space Θ**

  $(\underline{\alpha},\underline{\beta},\underline{\gamma}) = \arg\max_{\{(\alpha,\beta,\gamma)\ \text{in}\ \Theta\}} \Pr(\alpha,\beta,\gamma|V)$



- **Offline learning**
  - Learn $\Pr(\alpha,\beta,\gamma|V)$ via binary classification

    $\Pr(+1|V[\alpha,\beta,\gamma]) = \Pr(\alpha,\beta,\gamma|V)$
  - High learning complexity: 1-vs-all
  - Computationally challenging

- **Online inference**
  - Exhaustive search in the full 9D space is prohibitive

38

Marginal Space Learning (MSL)

- **Offline learning**
  - Break down the learning complexity

    $\Pr(\alpha,\beta,\gamma|V) = \underline{\Pr(\alpha|V)} \times \underline{\Pr(\alpha,\beta|V)/\Pr(\alpha|V)} \times \underline{\Pr(\alpha,\beta,\gamma|V)/\Pr(\alpha,\beta|V)}$

    Translation detector: $\Pr(+1|V[\alpha]) = \Pr(\alpha|V)$

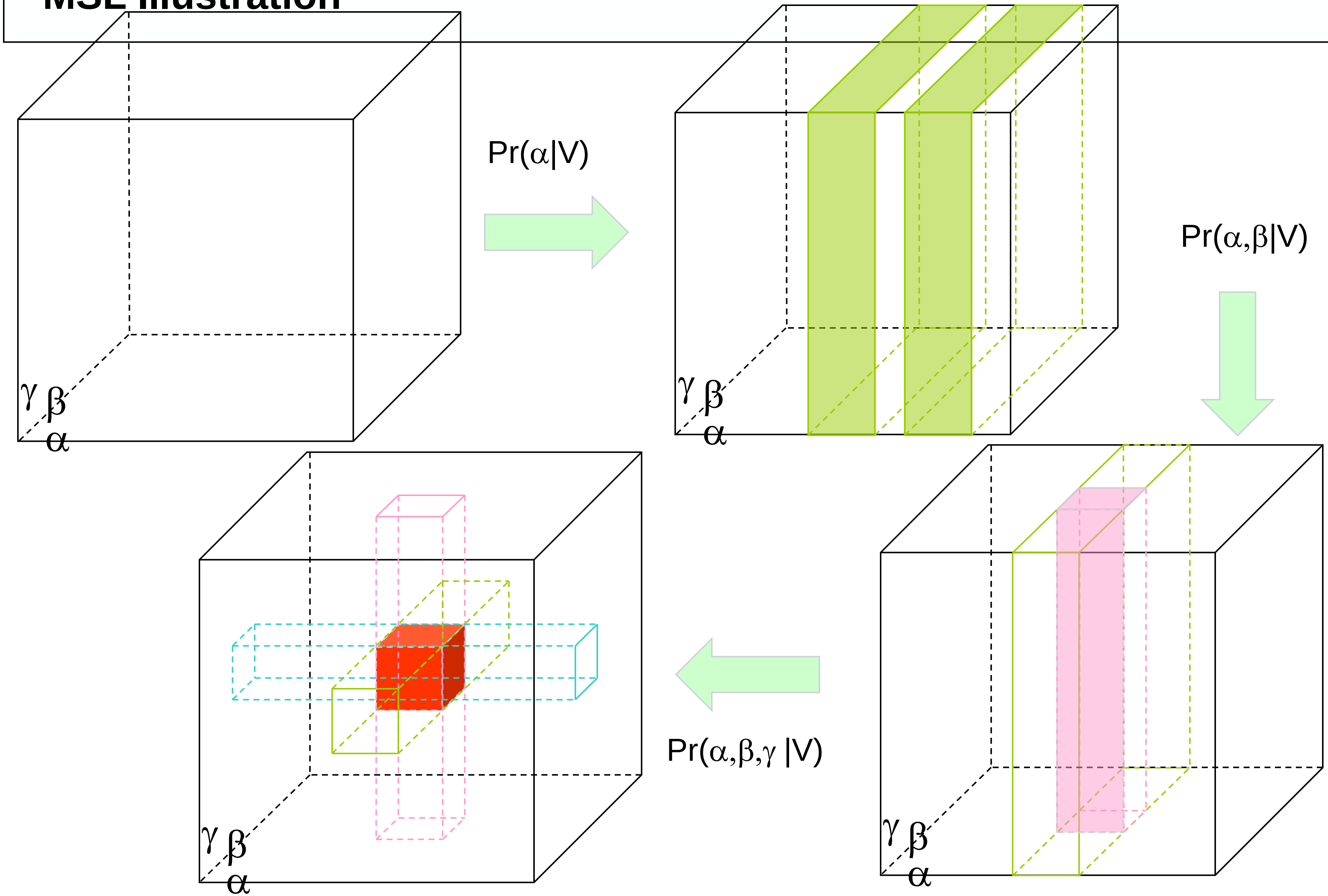
    Rotation detector: $\Pr(+1|V[\alpha,\beta]) = \Pr(\alpha,\beta|V)$

    Scale detector: $\Pr(+1|V[\alpha,\beta,\gamma]) = \Pr(\alpha,\beta,\gamma|V)$
  - Bootstrapping to reduce the number of negatives

- **Online inference**
  - Search in three spaces: {T}, {T,R}, {T,R,S}

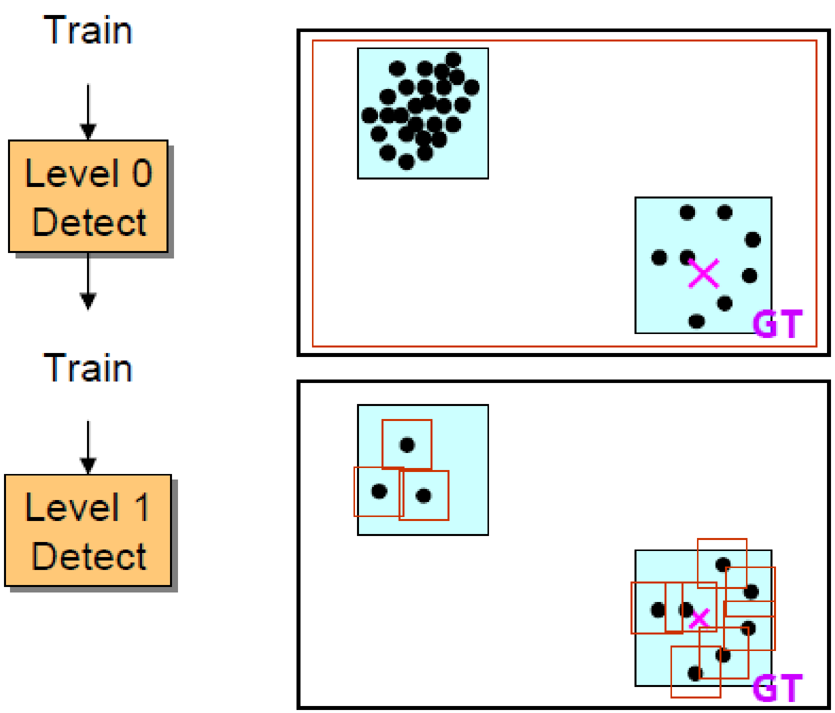
- **Extensible to cope with deformable shape space**

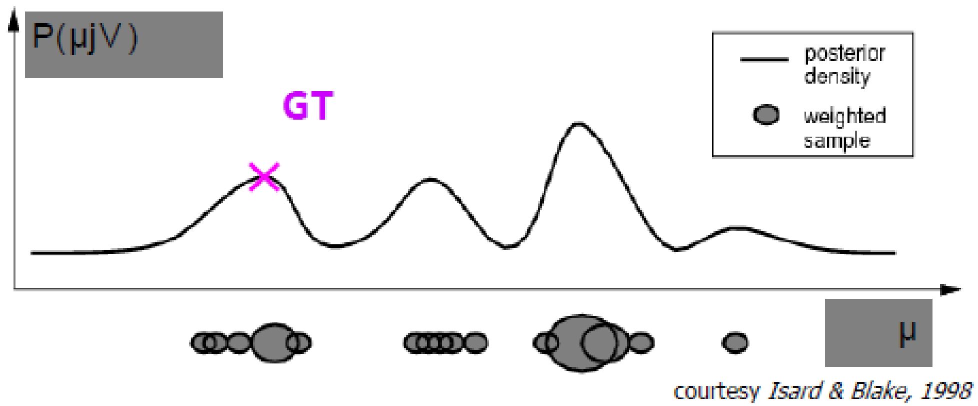
MSL Illustration



Hierarchical MSL – Image Pyramid to Improved Robustness [Sokfa et al, CVPR 2010]

SIEMENS





$$\hat{\mu} = \arg\max_{\mu} P(y = +1 | \mu; V)$$

Example: Liver segmentation [Ling et al. CVPR 2007]

SIEMENS

Mesh – $M(P,T)$
- P is the point set,
- T is the triangle index set.

Mesh Hierarchical
- Base mesh M^b
- Hierarchical meshes

$$\begin{cases} M^0 = \downarrow(M^b) \\ M^1 = \downarrow(M^0) \\ M^2 = \downarrow(M^1) \end{cases}$$

Statistical Shape Model

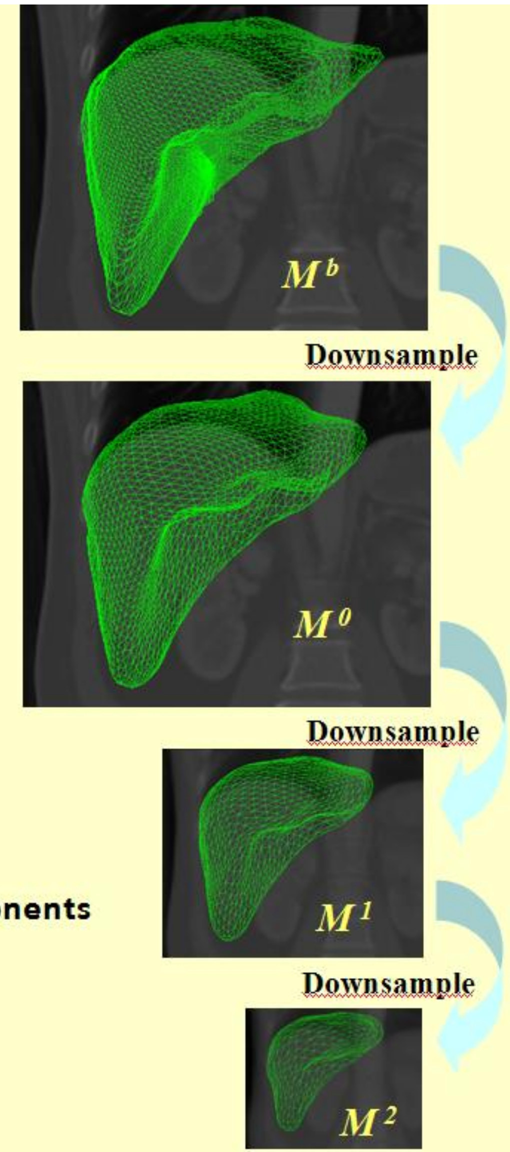
$$x = \mu + \sum_i c_i v_i$$

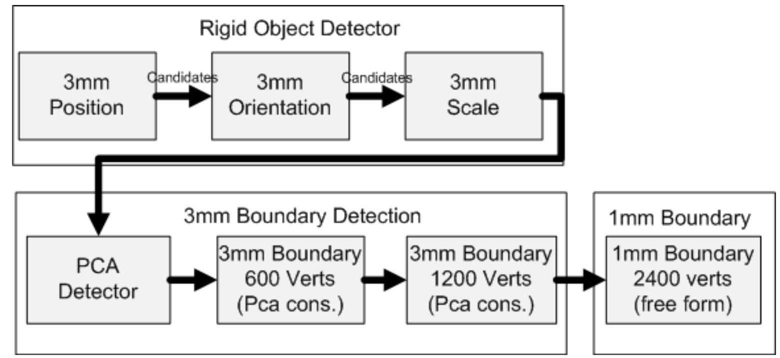
shape

mean shape

shape coefficients

components





Method: Learning based hierarchical segmentation

- Hierarchical mesh framework
- Subspace shape initialization
- Learning-based detection and boundary refinement

Combine Learning-based and PDE-based techniques

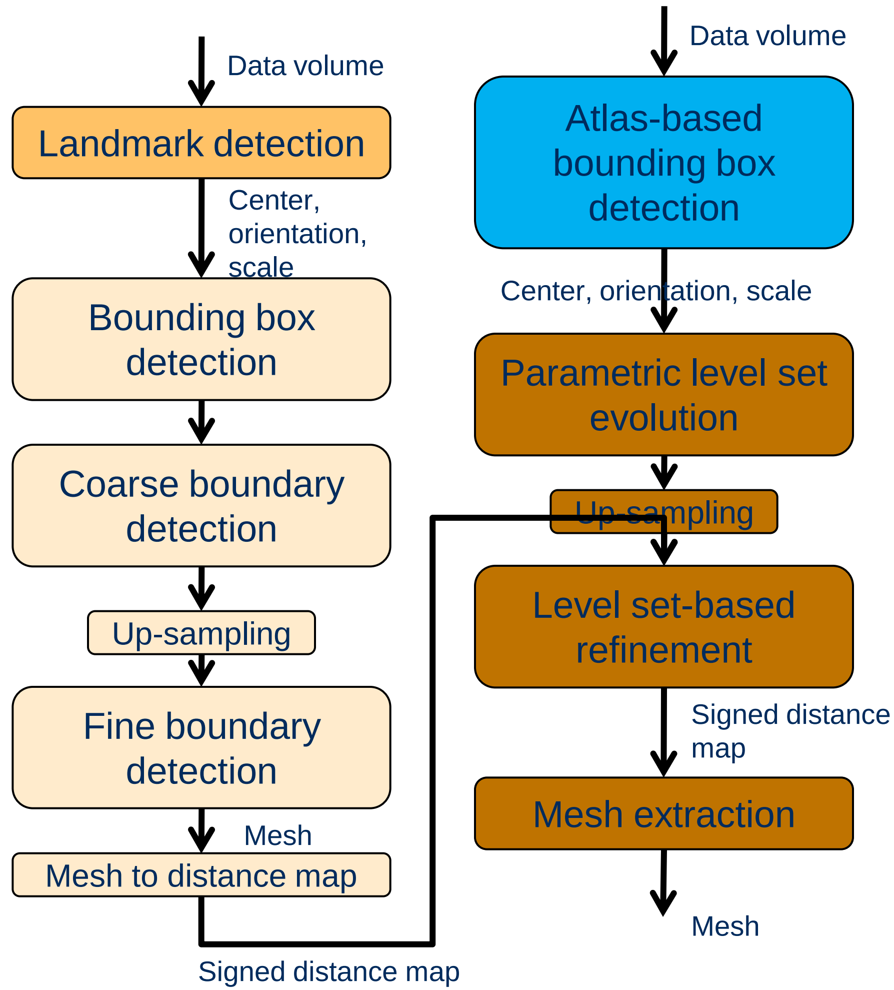
[Kohlberger et al. MICCAI 2011]

MSL + PBT + steerable features

Learning-based detection & segmentation + PDE-based refinement

Atlas-based initialization + level sets + shape model

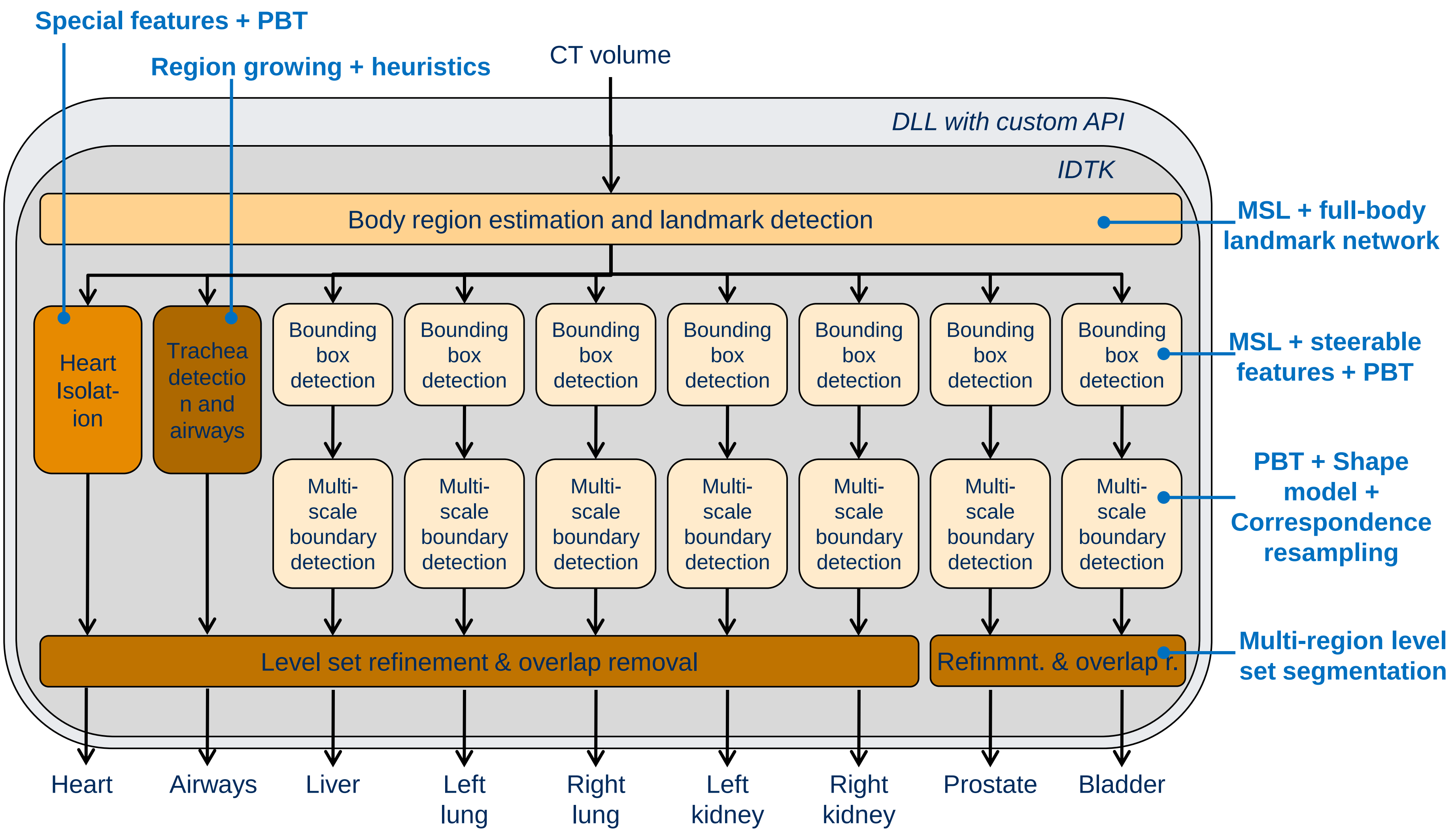
+ very robust if trained sufficiently
+ fast
- high number of training examples (~400)
- lack of detail due to point-based shape representation
- difficult to prevent overlaps



+ low number of training cases (20-100)
+ high details due to level set representation
+ overlaps easily preventable
- initialization not robust
- many parameters

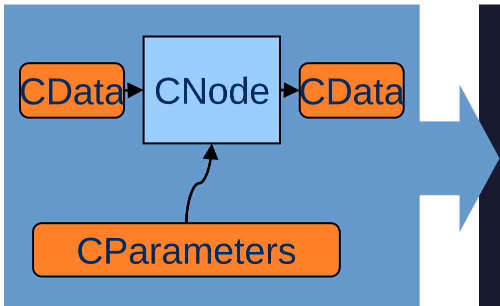
System Concept

SIEMENS



Integrated Detection Toolkit (IDTK) --- Building the Whole Body Parsing Project

SIEMENS



- Coding anatomical relationships by a network structure
- Flexible configuration
- Visual programming
- Scalable technology

MMBIA 2012

IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)

January 9th, 2012 **Breckenridge, Colorado**

http://www.mmbia.org/mmbia2012

Important Dates

Paper Submission Deadline: **September 12th, 2011**
Paper Decisions: October 31st, 2011
Final Papers Due: December 1st, 2011