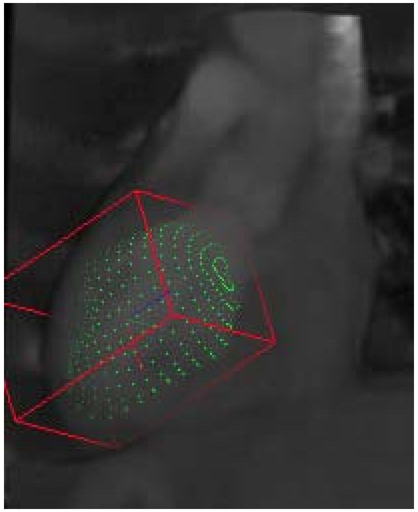
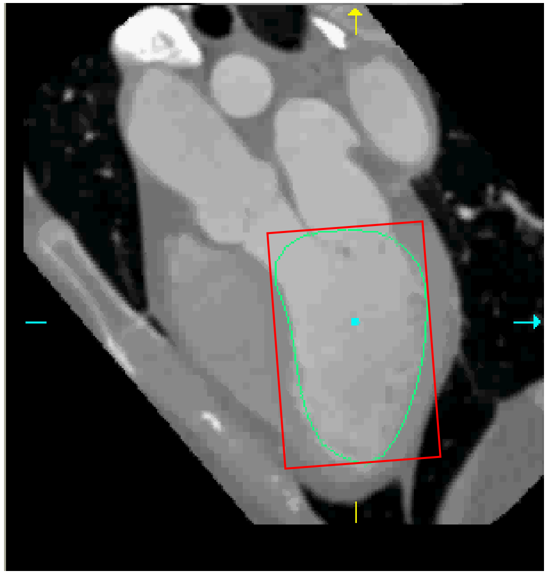
Marginal Space Learning

Adrian Barbu

3D Object Detection

- Find the 3D bounding box of the object of interest
- Supervised learning using PBT and features
- 9 parameter search space Ω:
  - Center $\mathbf{x}=(x,y,z)$
  - Scale $s=(s_1,s_2,s_3)$
  - Orientation $\boldsymbol{\theta}=(\theta_1,\theta_2,\theta_3)$





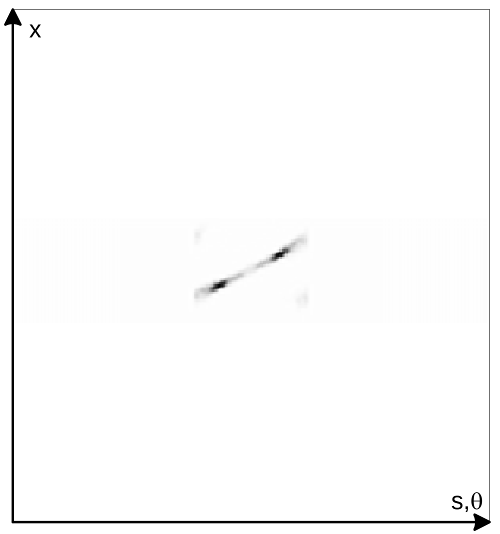
Computational Challenge

Problem:

- 9 parameter search space too large ~ 10^{13} locations
- Cascade or PBT not fast enough
  - Would take 1-3 months to find the solution!

But:

- What we want is a single location out of 10^{13}
  - Needle in the haystack
  - $p(x,s,\theta|I)$ is very peaked
  - Like a black dot on a large white board



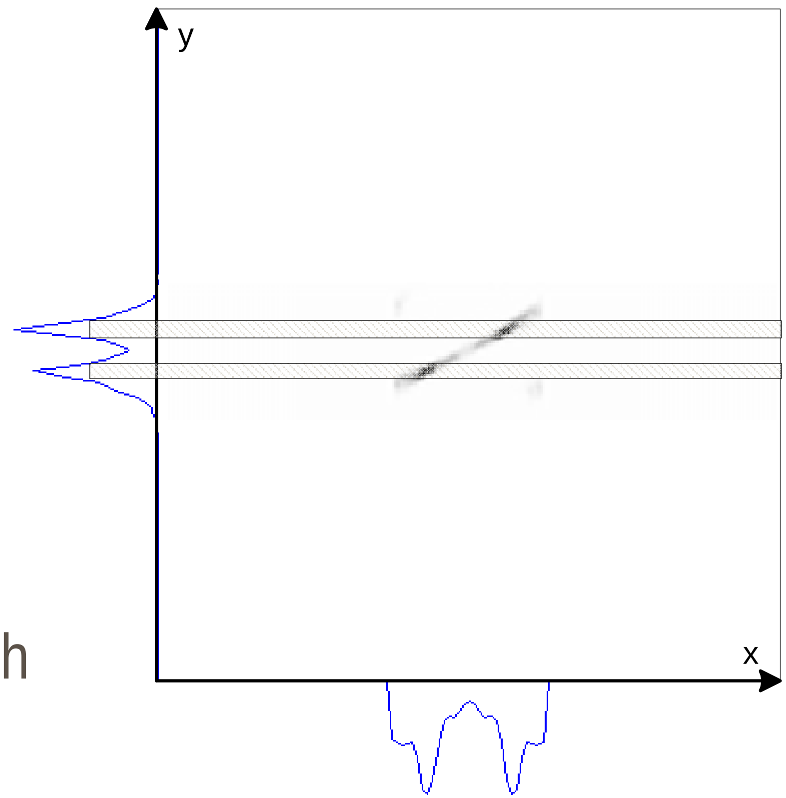
Marginal Space Learning

Intuition

- Probability mass $p(x, y|I), (x, y) \in \Omega$ is usually sparse (concentrated at very few locations)
- Can learn a marginal probability

$$P(y|I) = \int_{\Omega_x} P(x, y|I)dx$$

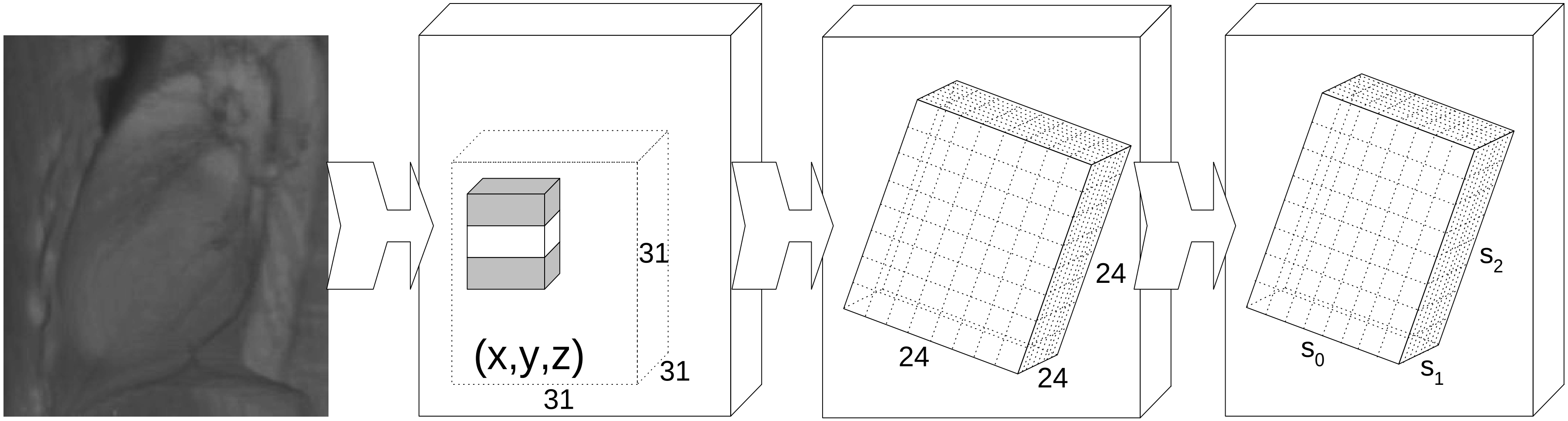
- Focus on locations with large $P(y|I)$



Advantages

- Search space is greatly reduced
- Global optimum typically still in search space

Marginal Space Learning for 3D Detection

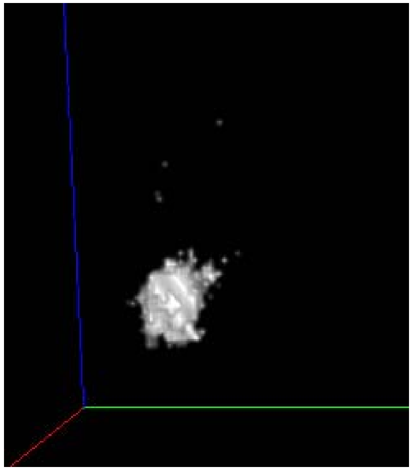


3D CT scan

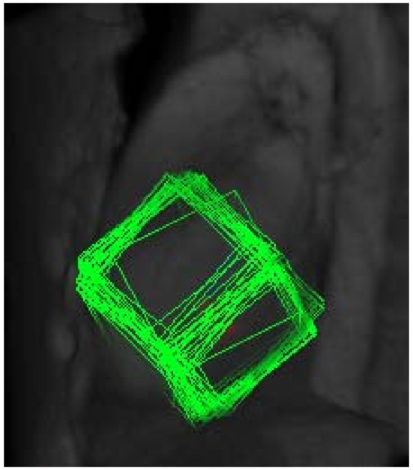
LV center detector

LV center and orientation

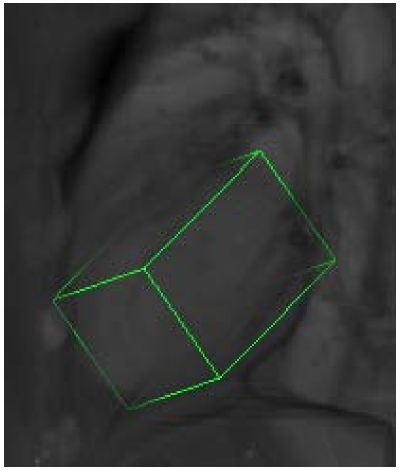
similarity transformation



Detected LV centers



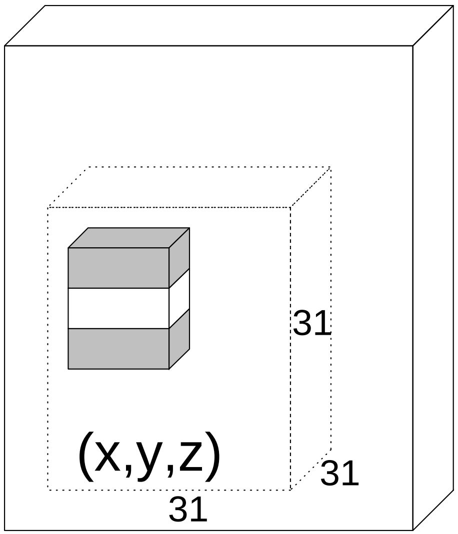
Detected LV centers and orientations



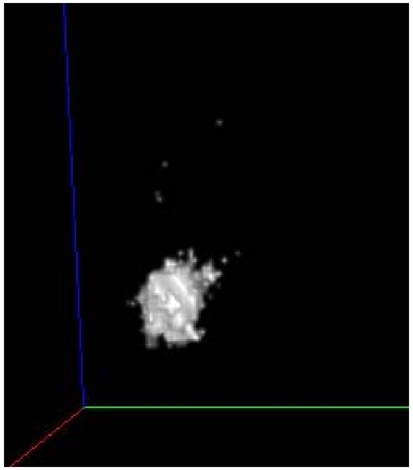
Detected LV

Training the MSL Classifiers for the LV

- Level 1: Location (x,y,z)
  - 323 volumes, 4-fold cross-validation
  - Positives: true LV center location in each volume
    - Plus perturbed $\pm$ 1 pixel
    - About 10k positives
  - Negatives: random locations (x,y,z) at distance at least 20 from LV center
    - About 500k negatives
  - 3D Haar features and 3D integral image
    - 10k Haar features
  - Cascade with 2 levels:
    - 99.2% detection rate
    - 0.17% false alarm rate



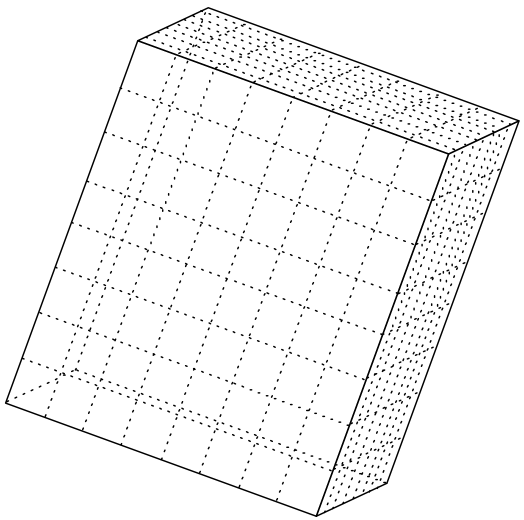
Training MSL



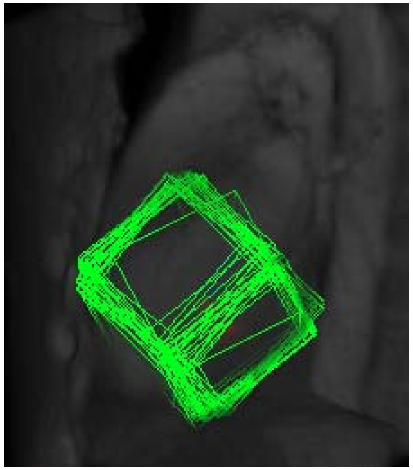
- From the location detector:
  - Best 100 candidates (x_i, y_i, z_i) are kept
- Level 2: Location and orientation $(x, y, z, \theta_1, \theta_2, \theta_3)$
  - Angle space discretized in 1000 combinations
  - Each of the 100 candidates
    - Is appended each of the 1000 angle combinations
    - Obtain 100k candidates $(x, y, z, \theta_1, \theta_2, \theta_3)$
  - Positives:
    - Candidates that are close to the LV parameters $(x^{LV}, y^{LV}, z^{LV}, \theta_1^{LV}, \theta_2^{LV}, \theta_3^{LV})$
  - Negatives
    - Candidates that are far from the LV parameters
  - PBT with 4 levels
    - Steerable features

Steerable Features

- Features at different locations on a 3D grid
  - E.g. 11x11x11 grid= 1331 locations
  - Rotated, scaled and translated by the object parameters
- At each location different feature types:
  - Gradient $g_x, g_y, g_z, ||g||$
  - Intensity
  - Combinations of the above
    - Sum, product, quotient, etc
  - Dot product of gradient direction and sample direction
  - Total 71 feature types
- Totally 91k features



Training MSL

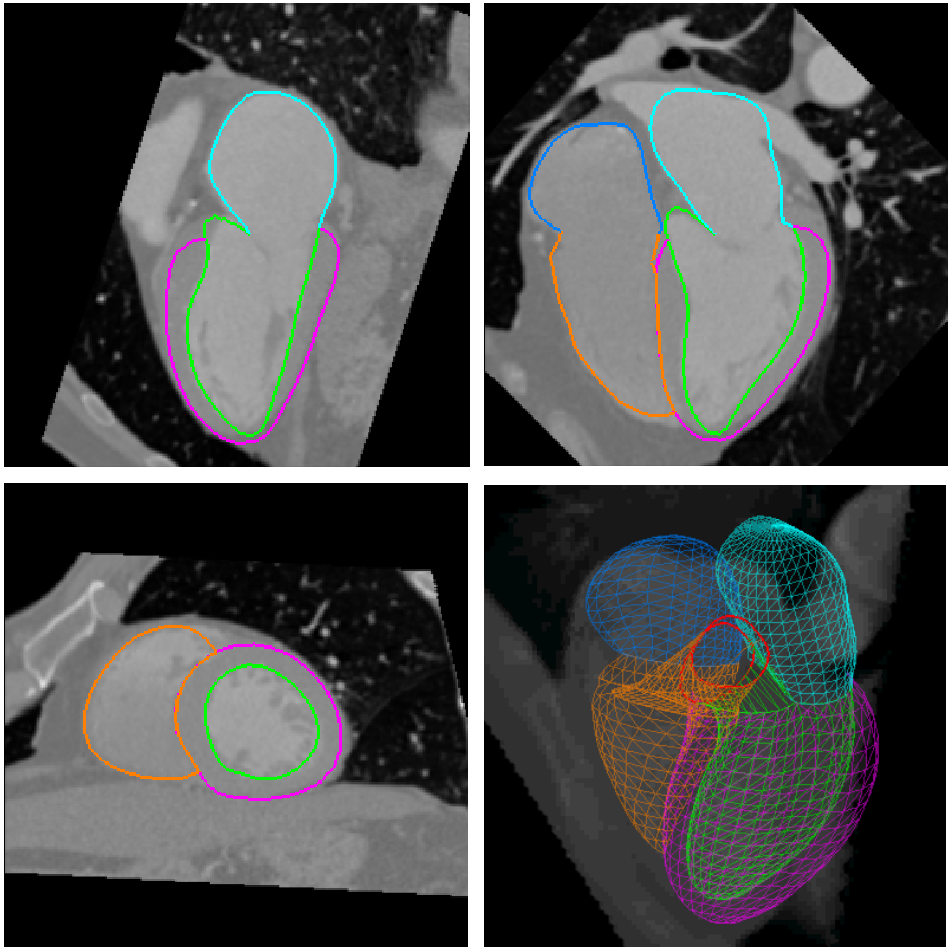


- From level 2:
  - Best 50 candidates $(x,y,z,\theta_1,\theta_2,\theta_3)$ are kept
- Level 3: position, orientation and scale $(x,y,z,\theta_1,\theta_2,\theta_3,s_1,s_2,s_3)$
  - Scale space discretized in 1000 combinations
  - Each of the 50 candidates $(x,y,z,\theta_1,\theta_2,\theta_3)$ from Level 2
    - Is appended each of the 1000 scale combinations
    - Obtain 50k candidates $(x,y,z,\theta_1,\theta_2,\theta_3,s_1,s_2,s_3)$
  - Positives:
    - Candidates that are close to the LV parameters
  - Negatives
    - Candidates that are far from the LV parameters
  - PBT with 5 levels
    - Steerable features

Four Chamber Heart Segmentation

Heart Segmentation

- LV Endocardium
- LV Epicardium
- Left Atrium
- Aortic Trunk
- Right Ventricle
- Right Atrium



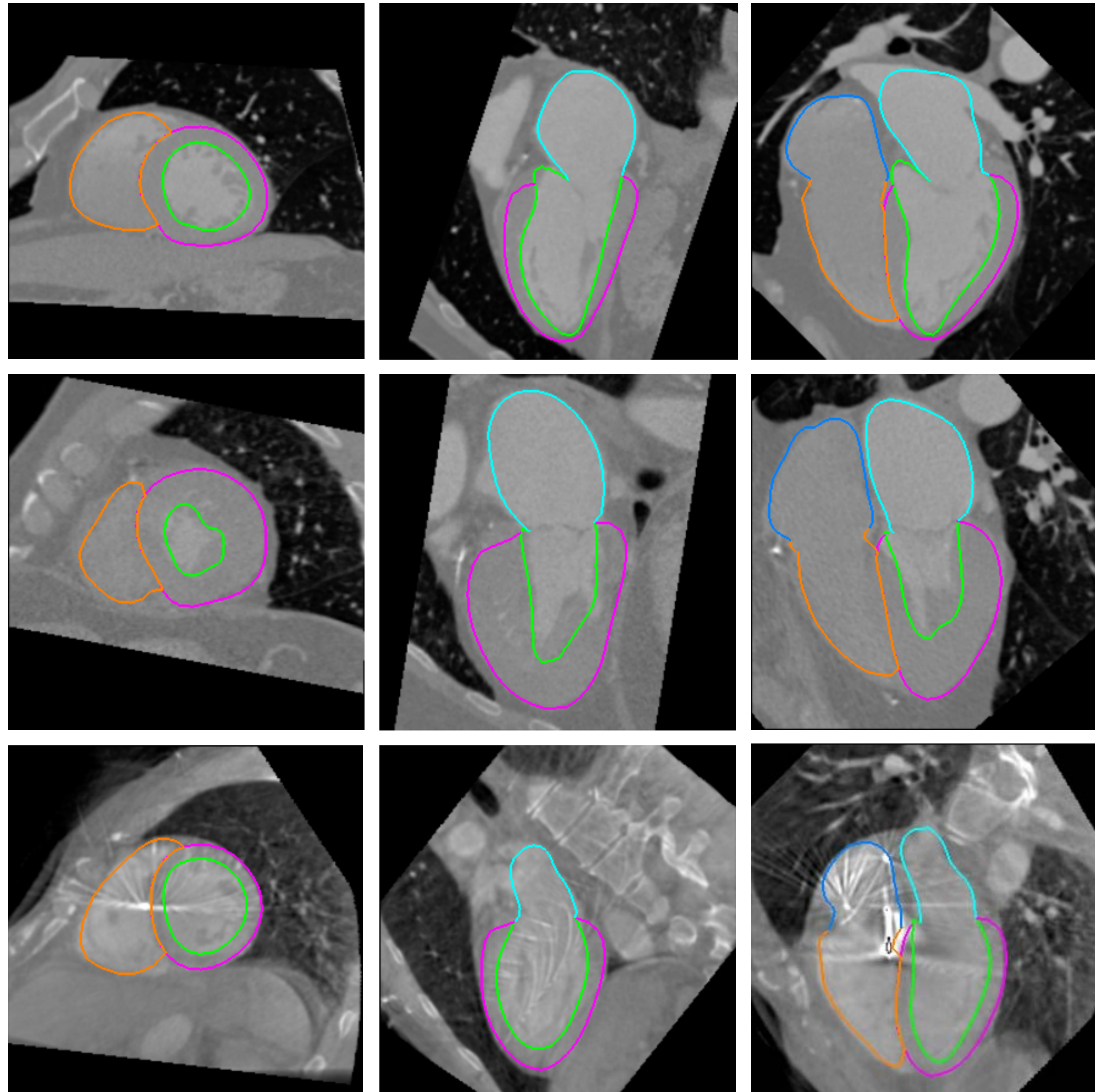
Heart Segmentation

More details:

- MSL to detect the location of each chamber
- Learning based boundary model
- PCA shape model

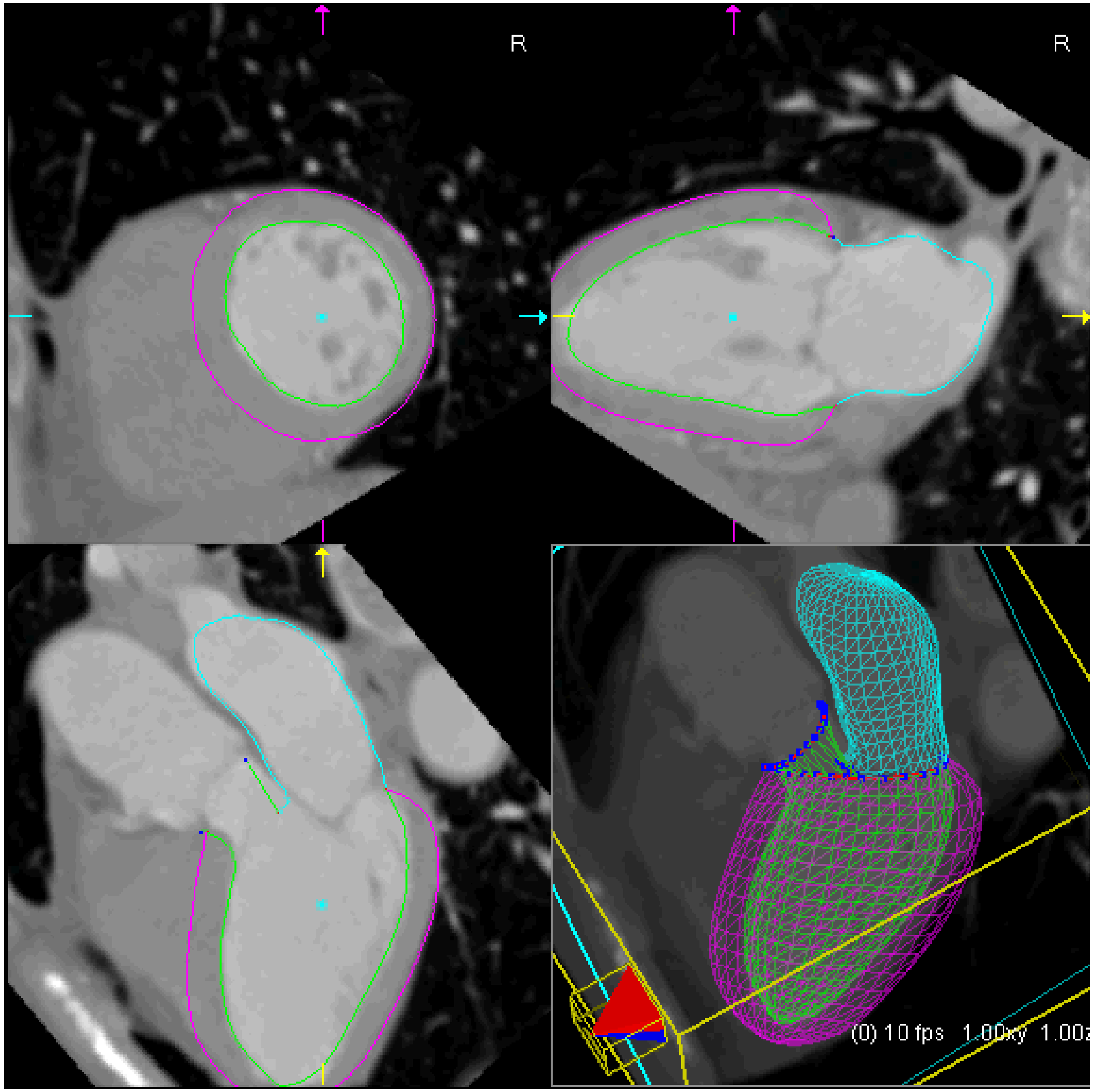
Performance:

- Working on 323 volumes with 4-fold cross-validation
- Mean error 1.3mm (1mm voxel size)
- 1 sec/volume
- 6 orders of magnitude faster than brute force search



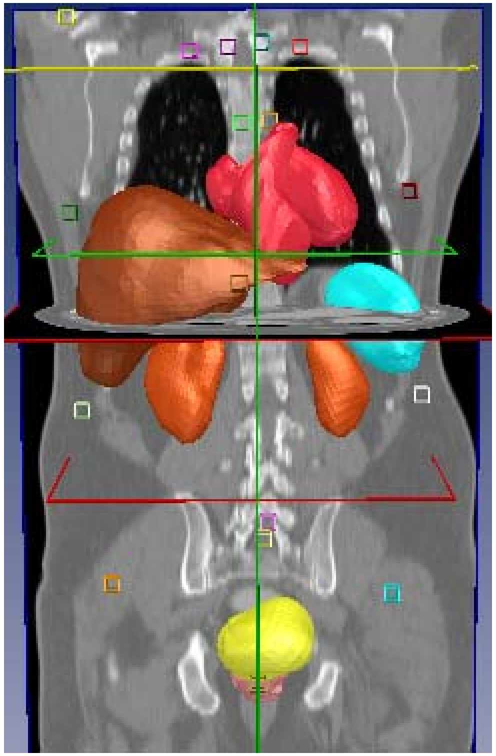
Heart Tracking

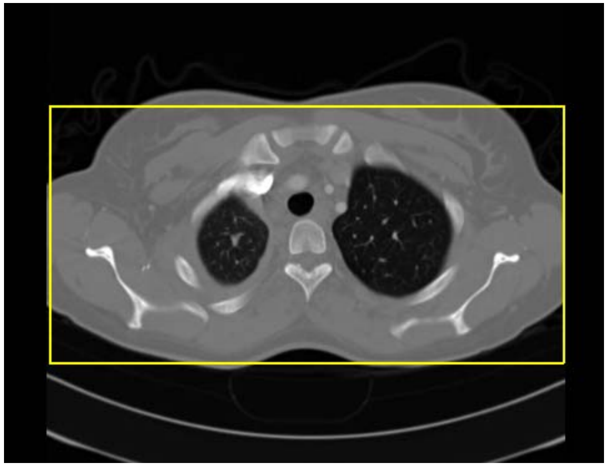
- Initialization by segmentation
- Appearance model using MSL
- Local search of parameters

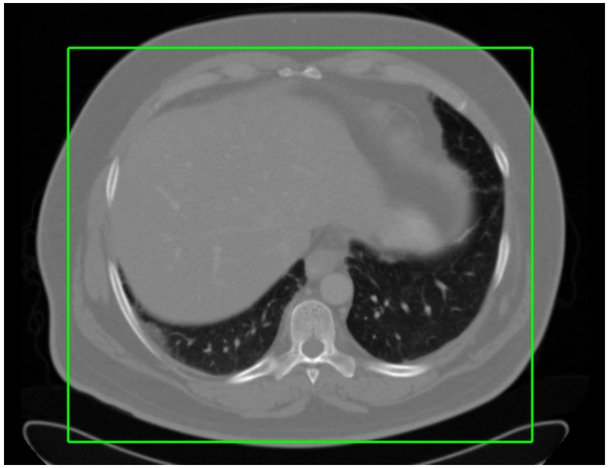


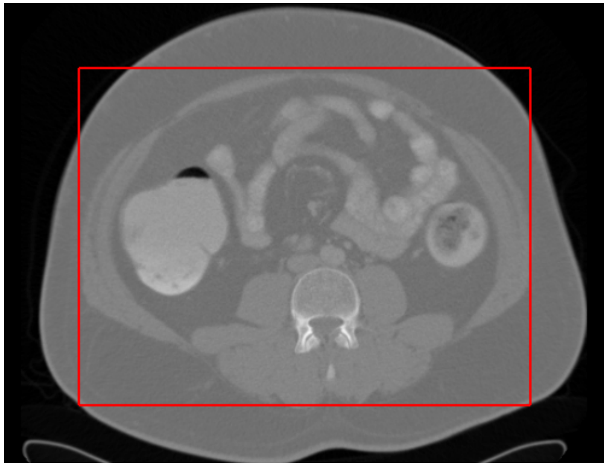
Full Body Parsing

- Marginal Space Learning
  - Cross-section detection
  - Landmark Detection
  - Organ Detection and Segmentation
- Cross-section detector:
  - Detect three salient cross-sections
  - Connected in a network for robustness



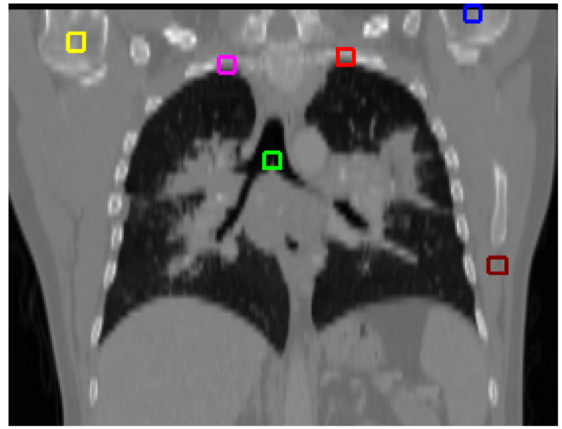


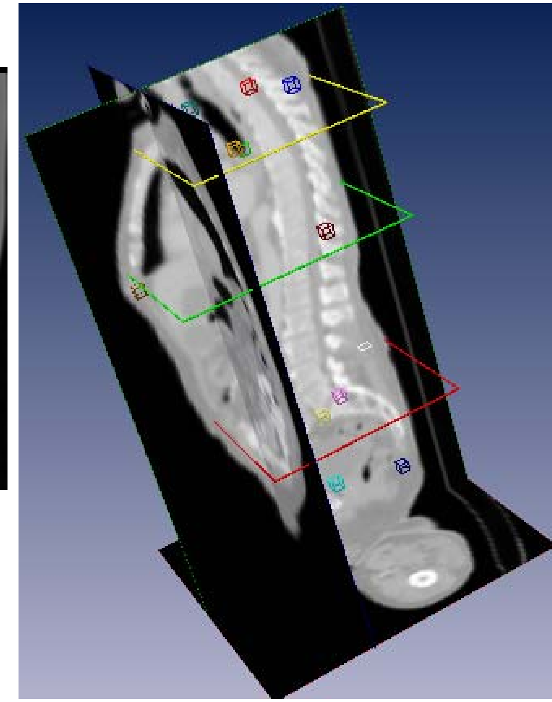




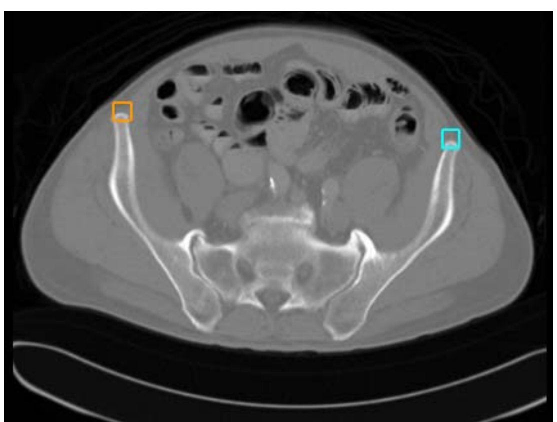
Landmark Detection

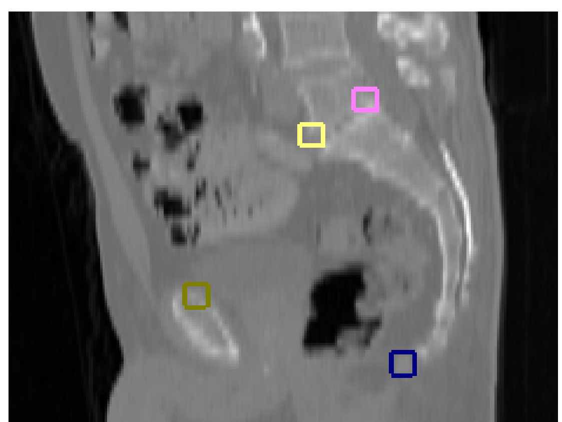
- Detect salient and important body landmarks
  - Aortic arch
  - Bronchial Bifurcation
  - Lung tips
  - Vessel bifurcations
  - Humerus heads
  - Lumbar vertebrae
  - Coccyx
  - Pubic symphysis tip
  - Front corners of the hip bone





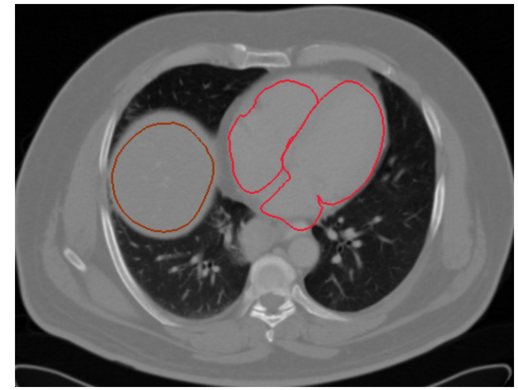
- Connected in a network for robustness

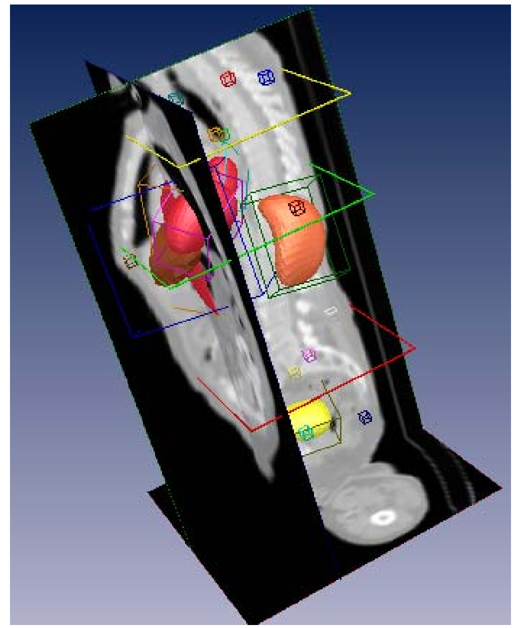


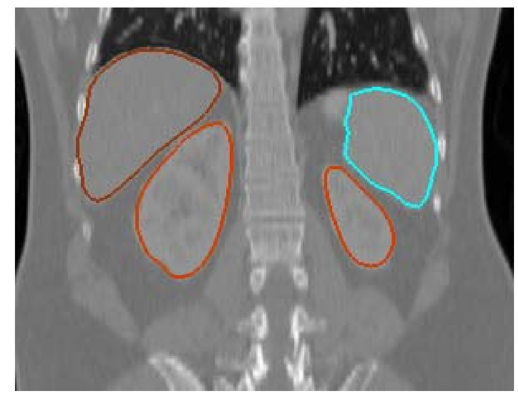


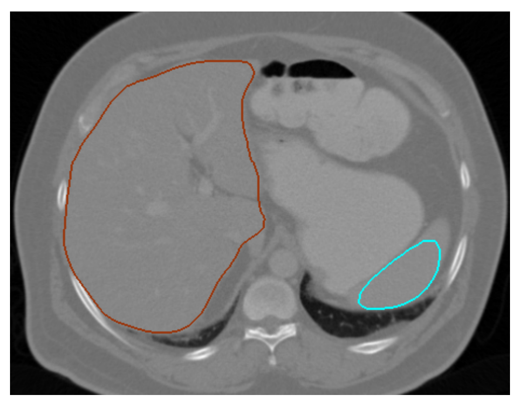
Organ Segmentation

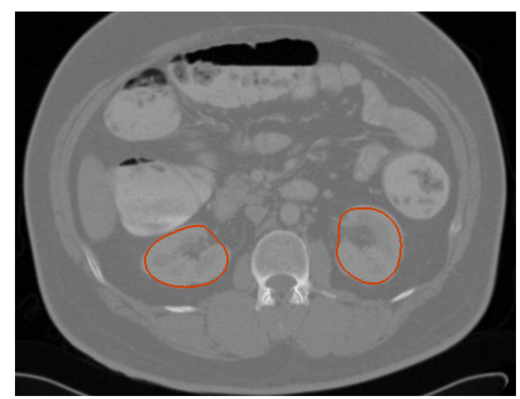
- Marginal Space Learning
- Organs:
  - Four heart chambers
  - Liver
  - Kidneys
  - Spleen
  - Prostate
  - Bladder





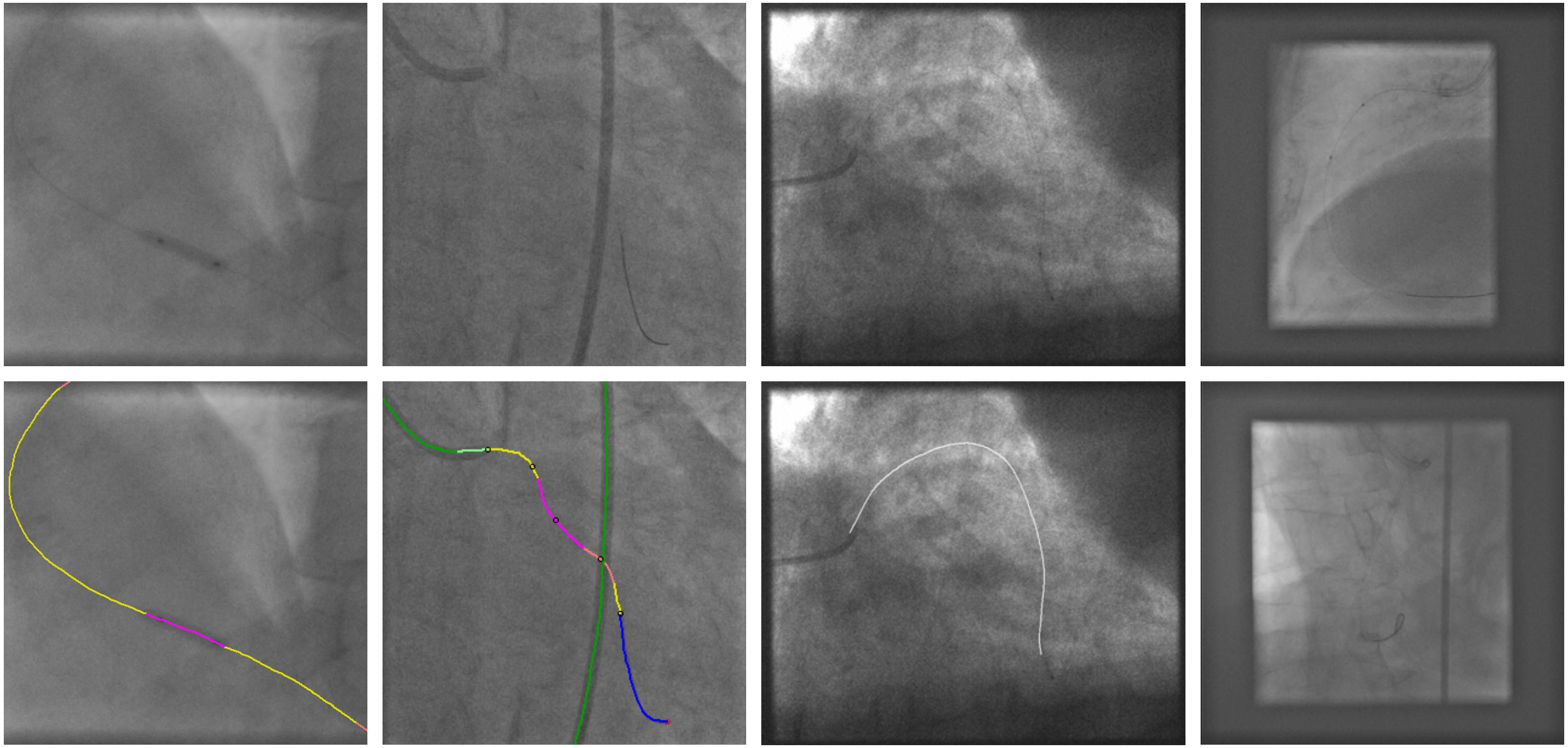






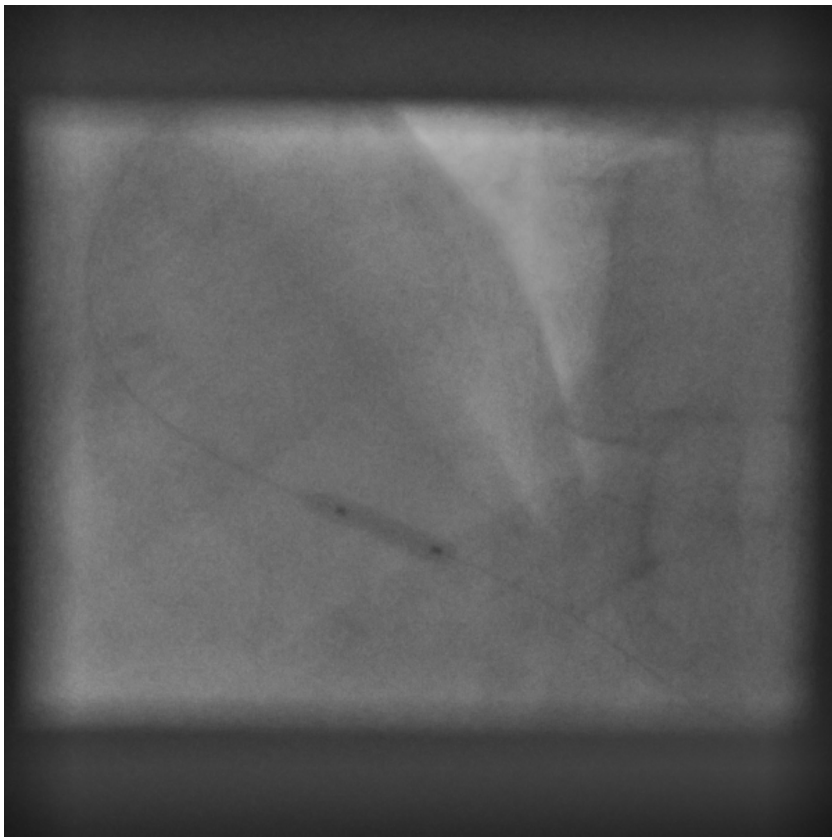
Guidewire Localization in Fluoroscopy

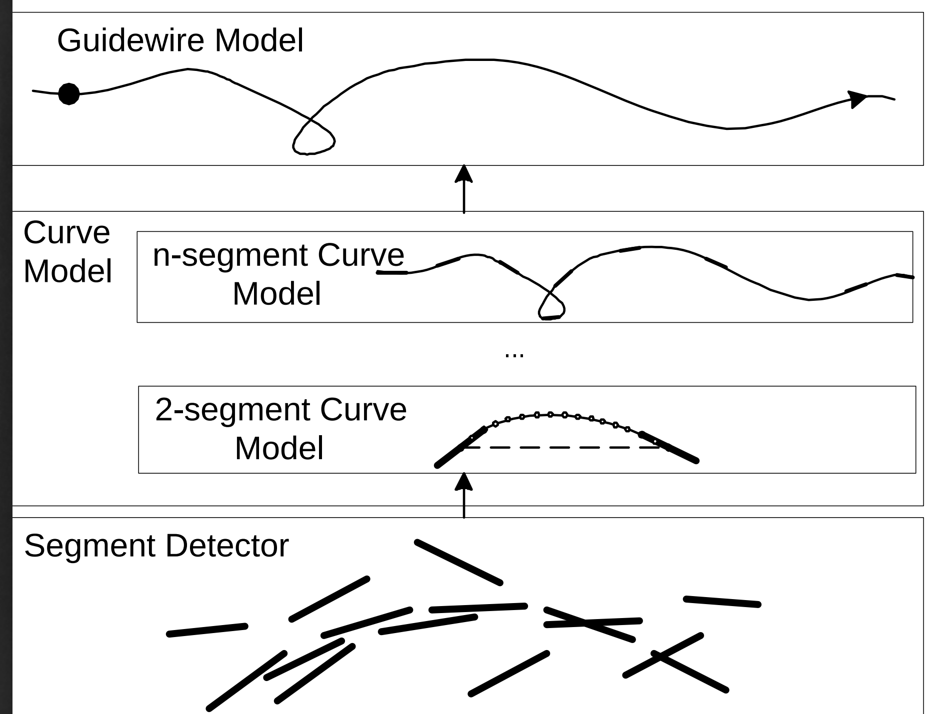
- 2D X-Ray
- Detect Flexible Guidewire
- Model both shape and appearance



Guidewire Localization in Fluoroscopy

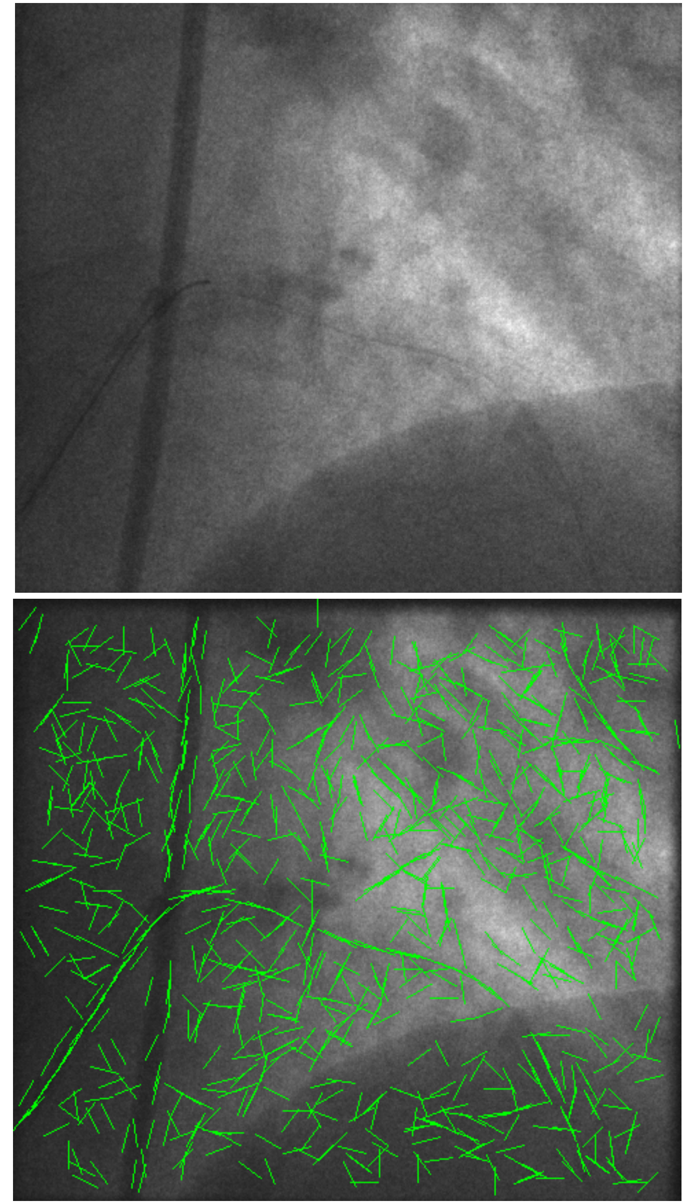
- Marginal Space Learning – Model gradually longer curves
- Hierarchical, part based model
- Joint model for shape and appearance



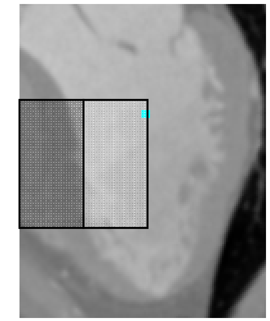


Guidewire Localization in Fluoroscopy

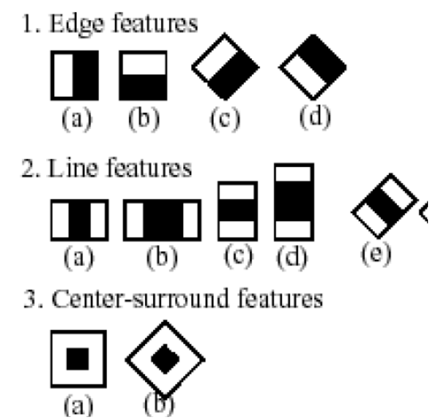
- Marginal Space Learning
  - Hierarchical model
  - Detect curves with more and more parameters
  - Curves are constructed from simpler curves from previous levels
- Level 1: Oriented Points
  - Parameters (x,y,θ)
  - Probabilistic Boosting Tree with Haar features



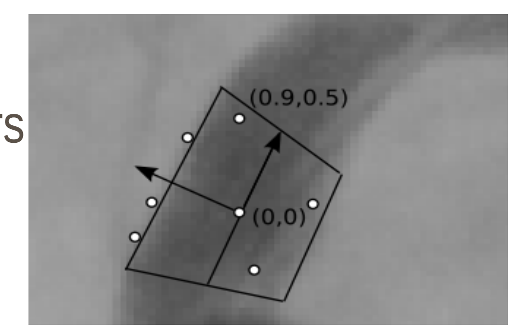
Features

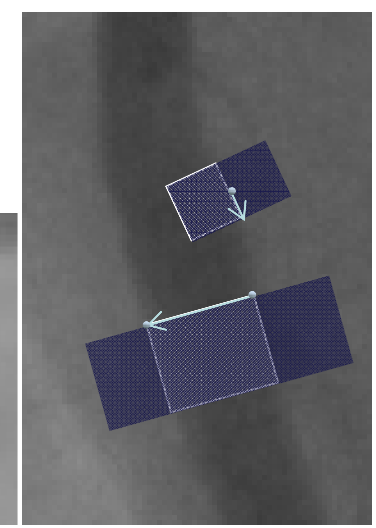


- Feature = a number obtained from the image through a predefined procedure
- Haar features
  - Sum of pixels in a rectangle
  - Can be computed efficiently using the integral image
  - Parameters: type, location (x,y) and size (dx,dy)
  - Are chosen to be invariant to brightness
  - E.g. feature: (1,10,20,15,30) type 1, location (10,20), size (15,30)



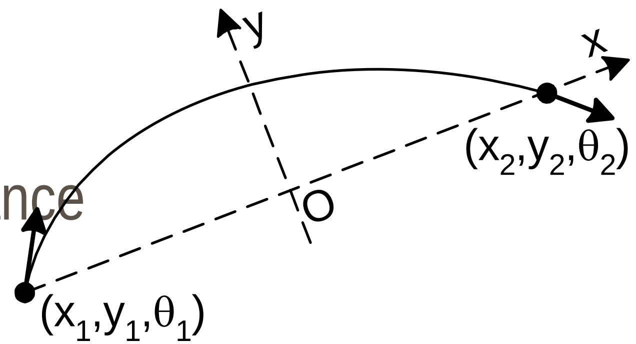
- Can easily obtain 100k features
- Other types of features exist:
  - Steerable features
  - Features obtained from different filters

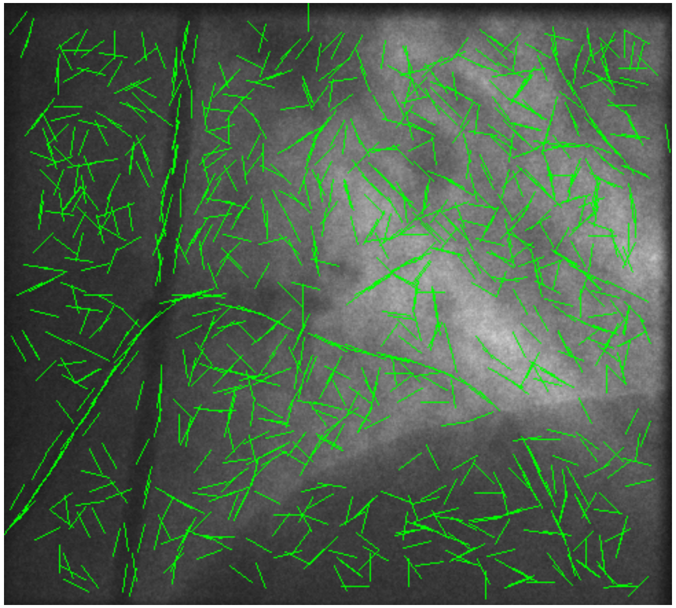


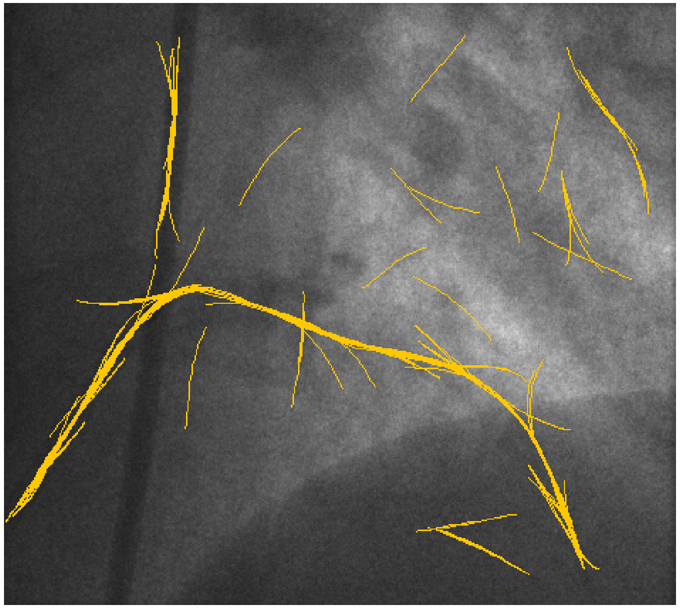


Guidewire Localization in Fluoroscopy

- Level 2: Short Curves
  - Length 32-80
  - Polynomial model from 2 segments
  - PBT with many types of shape/appearance features
    - Curvature
    - Statistics of gaps

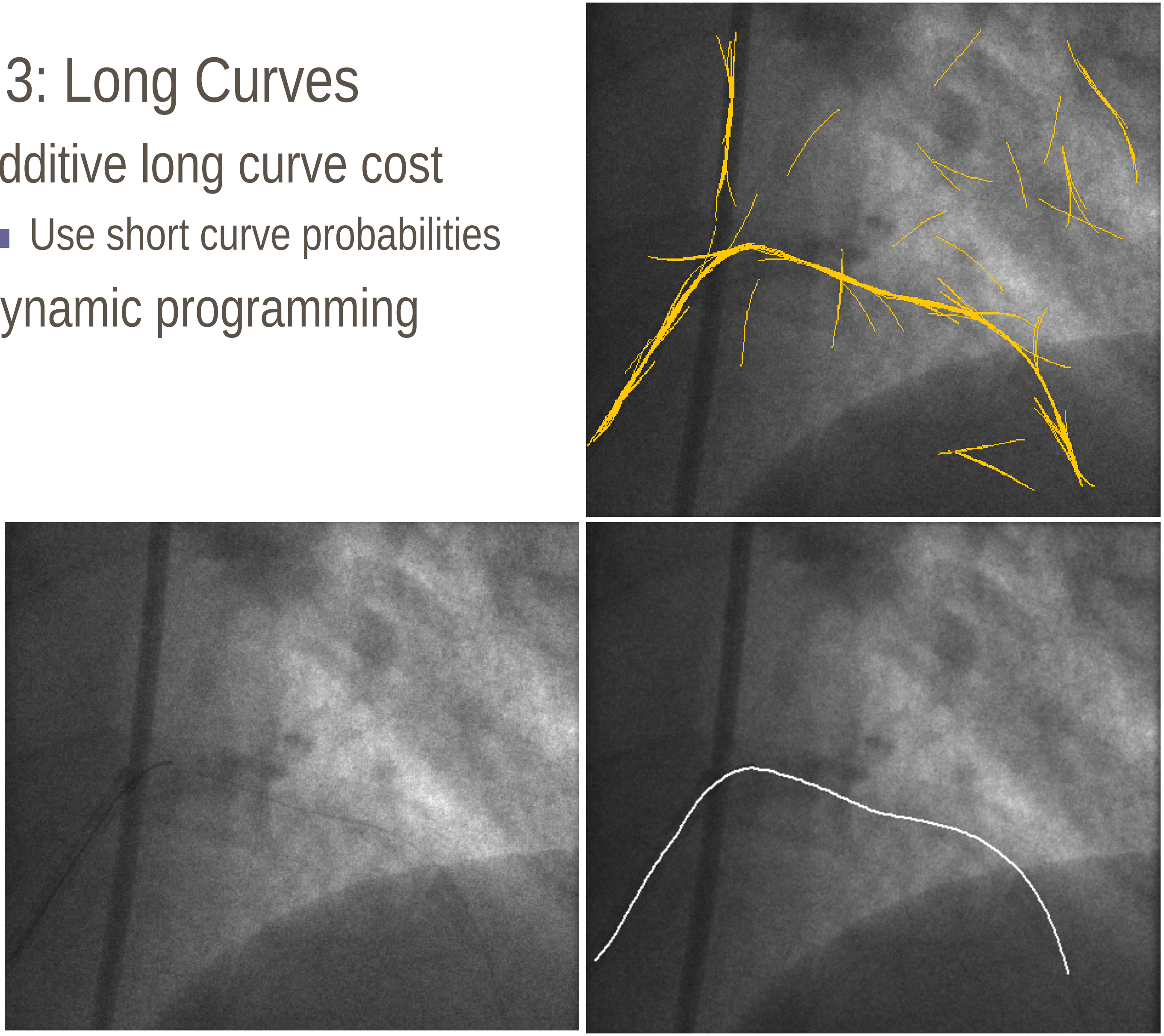






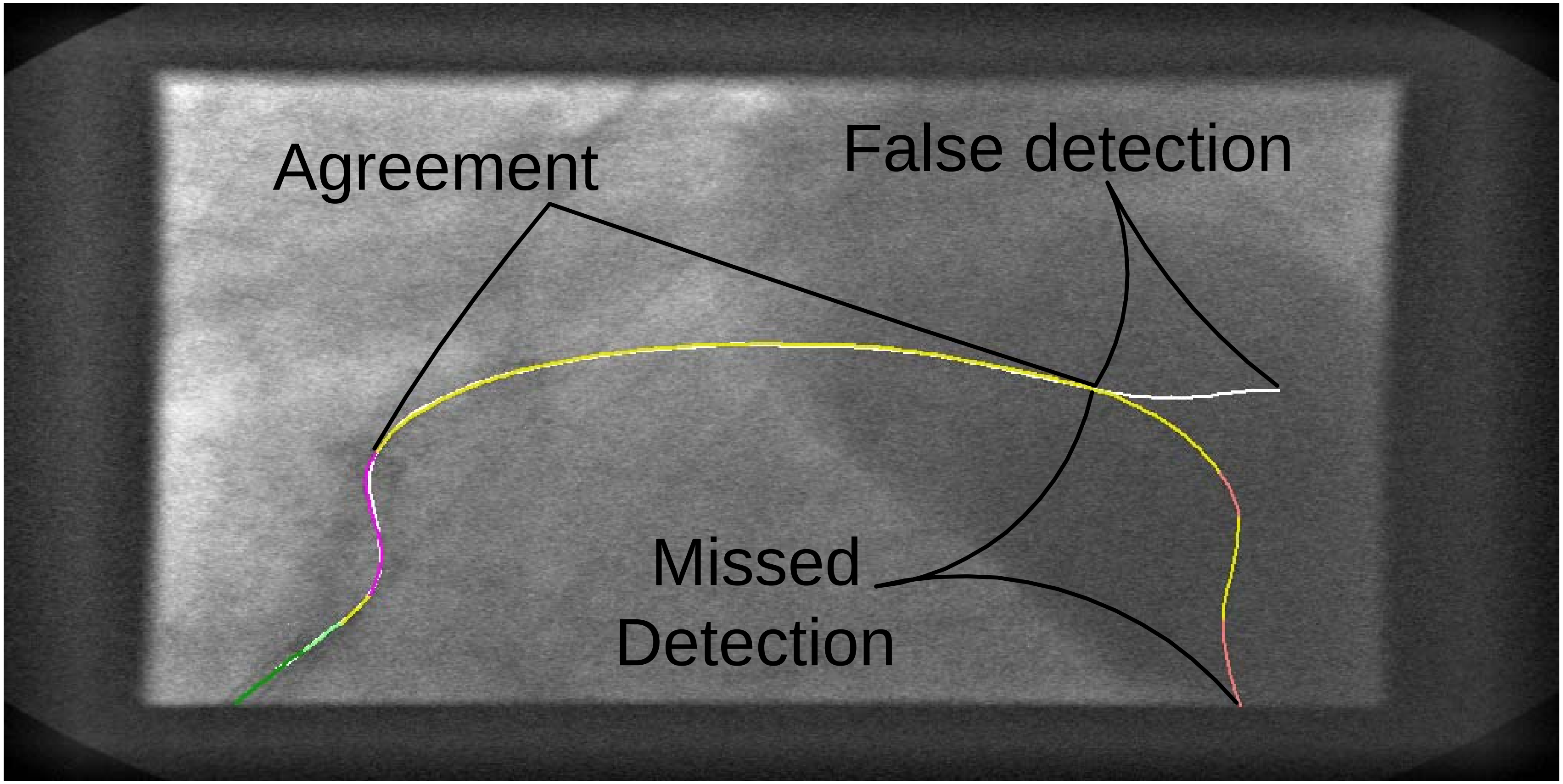
Guidewire Localization in Fluoroscopy

- Level 3: Long Curves
  - Additive long curve cost
    - Use short curve probabilities
  - Dynamic programming

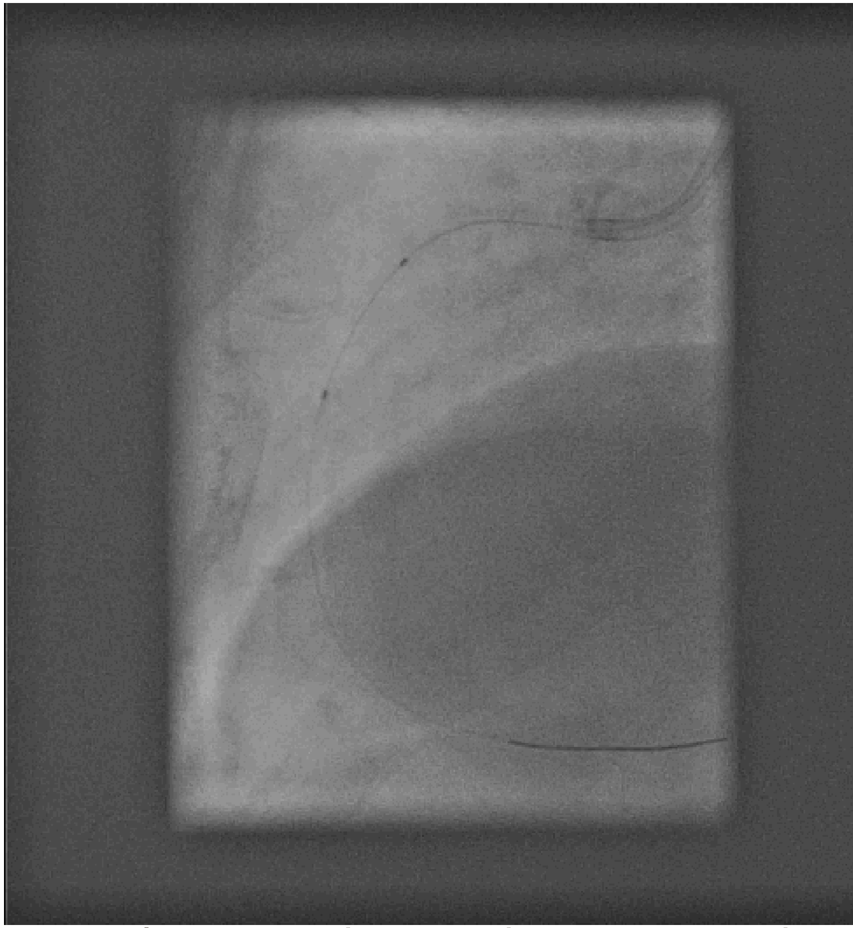


Evaluation

- Two error measures:
  - Missed detection – percentage of the guidewire that was not detected
  - False detection – percentage of the result that is not guidewire



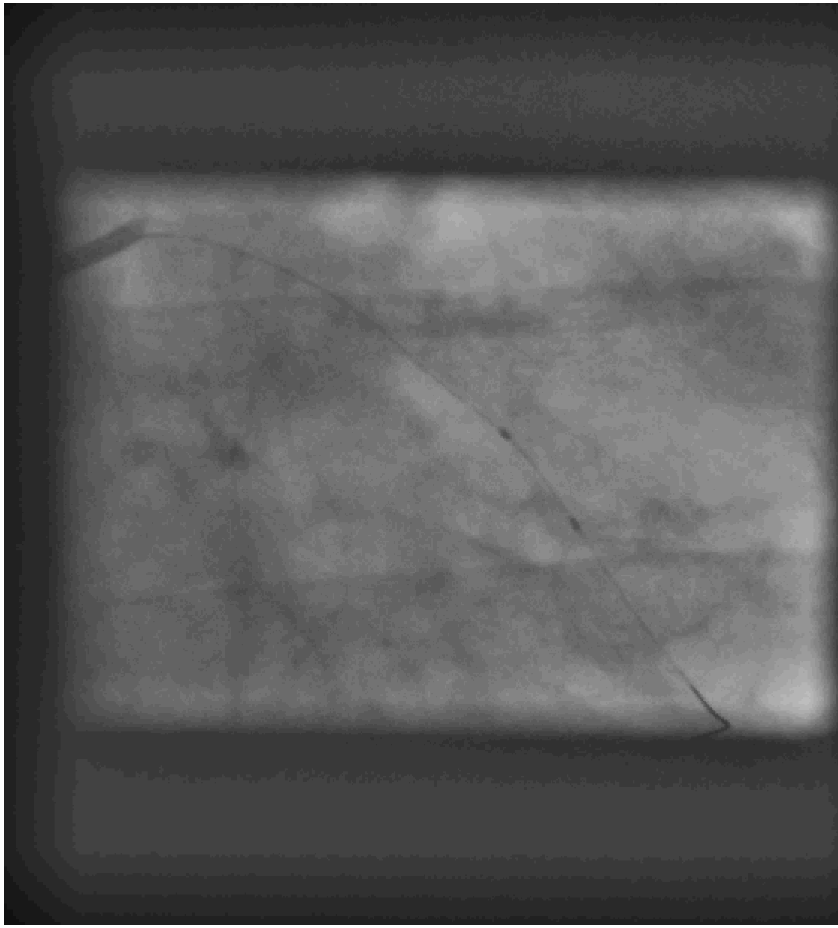
Results

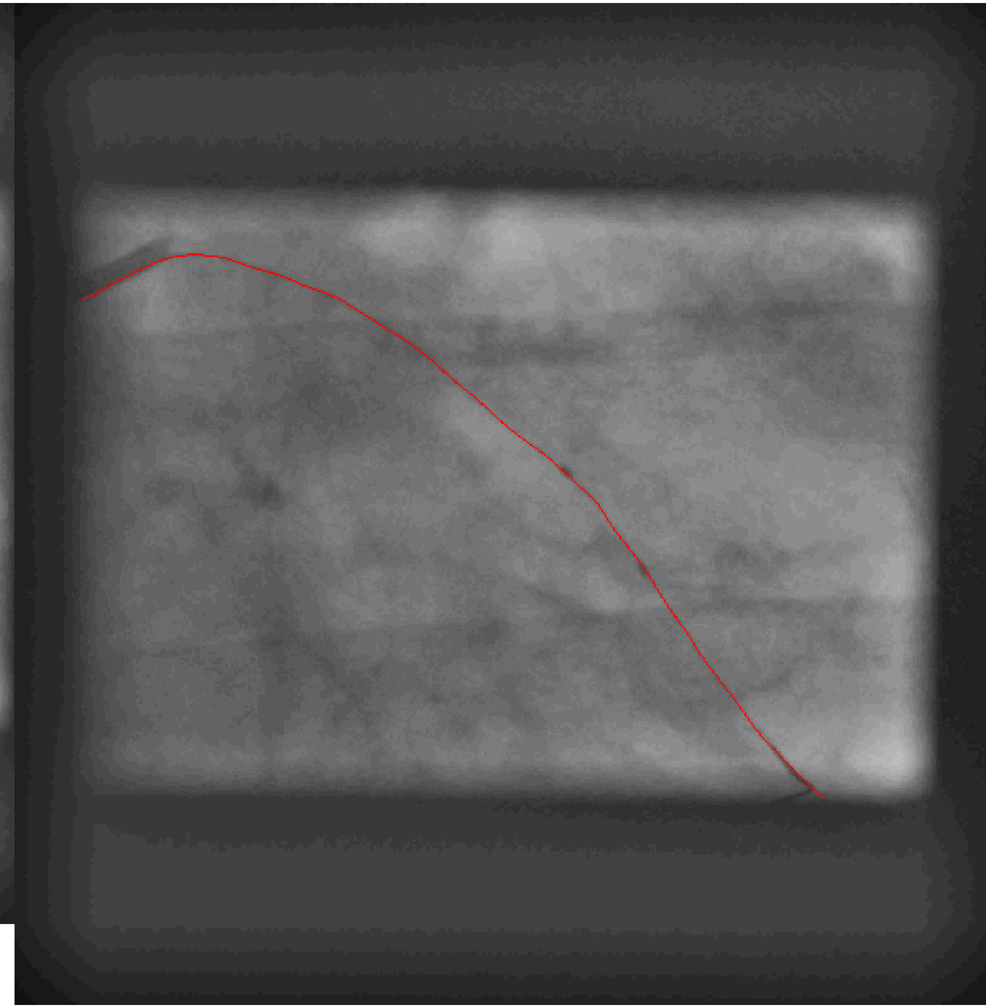


Results on 700 images (73 sequences):

- Missed detection 19.0%
- False detection 5.6%
- 1-2 seconds/frame

Results





Results

Conclusion

- Marginal Space Learning
  - A computational solution for detection in high dimensional parameter spaces
  - Detects in marginal spaces of increasing dimension
  - Good detection rate in marginal space:
    - True location usually not lost
    - Otherwise, still quite close
  - Small false alarm (e.g. 0.1%) in marginal space
    - Greatly reduce search in full space (e.g. 1000 times smaller space)
    - Great speedup (e.g. 1000 times faster)
  - Training
    - Each level is trained using candidates from previous level