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
SCR - Comprehensive Research on Biomedical Imaging
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
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

- 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
Appearance variations

Contrast

Pathologies
- Deformed
- Missing
- Misleading context

Context

Body Portions
- Narrow FOV
- Occlusion
Challenges

- **Shape**
- **Accuracy**
- **Speed**
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
Outline

- WB Landmark
- WB Organ
- WB Lesion
- WB Bone
Landmarks

Skull Base
Liver, Hip
Lung Top
Liver, Sternum
Trachea
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: $12! \times 1\text{sec} /60/60/24/365 > 15\text{ years}$
  - 43 landmarks: $43! \times 1\text{sec} /60/60/24/365 > 1045\text{ years}$

- Difficult to find the best search order even offline

- The landmarks could be missing.
“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
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

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

\[ F_k(S \cup \{l\}) - F_k(S) \geq F_k(T \cup \{l\}) - F_k(T) \quad \forall S \subseteq T \]

- **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

1 LiverTop
→ Skull
→ 2 FemurHeadR (1)
→ 3 HipR (2)
→ 4 HipL (2)
→ 5 KidneyR (2)
→ 6 KidneyL (2)
→ 7 LungTopL (4)
→ 8 LungTopL (7)

…

1 LiverTop
→ Skull 
→ 2 AortaRoot (1)
→ 3 SternumBot (1)
→ 4 KidneyL (1)
→ 5 KidneyR (2)
→ 6 HipL (2)
→ 7 HipR (2)
→ 8 LungTopL (4)

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

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>std</th>
<th>Q95</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent $D_{8mm}$ $N = 63$</td>
<td>17.30</td>
<td>6.16</td>
<td>46.24</td>
<td>84.51</td>
</tr>
<tr>
<td>Greedy $D_{8mm}$ $N = 63$</td>
<td>1.14</td>
<td>0.47</td>
<td>1.92</td>
<td>2.44</td>
</tr>
<tr>
<td>Independent $D_{8mm}$ $N = 25$</td>
<td>6.72</td>
<td>6.40</td>
<td>17.73</td>
<td>35.00</td>
</tr>
<tr>
<td>Greedy $D_{8mm}$ $N = 25$</td>
<td>0.65</td>
<td>0.43</td>
<td>1.26</td>
<td>5.08</td>
</tr>
</tbody>
</table>
## FPR and FNR

<table>
<thead>
<tr>
<th></th>
<th>Greedy</th>
<th>Independent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$FP_A$</td>
<td>$FN_A$</td>
</tr>
<tr>
<td>SkullBase</td>
<td>0 (193)</td>
<td>0 [50]</td>
</tr>
<tr>
<td>R.LungTop</td>
<td>0 (84)</td>
<td>1 [114]</td>
</tr>
<tr>
<td>LiverDome</td>
<td>0 (86)</td>
<td>2 [65]</td>
</tr>
<tr>
<td>R.HipTip</td>
<td>0 (131)</td>
<td>0 [94]</td>
</tr>
<tr>
<td>R.Knee</td>
<td>0 (265)</td>
<td>0 [12]</td>
</tr>
<tr>
<td>LiverBott.</td>
<td>2 (33)</td>
<td>1 [33]</td>
</tr>
<tr>
<td>TracheaBif.</td>
<td>0 (44)</td>
<td>0 [41]</td>
</tr>
<tr>
<td>LiverCent.</td>
<td>0 (90)</td>
<td>1 [136]</td>
</tr>
<tr>
<td>L.HumerusHead</td>
<td>0 (96)</td>
<td>1 [12]</td>
</tr>
<tr>
<td>R.HumerusHead</td>
<td>1 (80)</td>
<td>2 [7]</td>
</tr>
<tr>
<td>L.LungTop</td>
<td>0 (61)</td>
<td>1 [21]</td>
</tr>
<tr>
<td>L.HipTip</td>
<td>0 (94)</td>
<td>1 [46]</td>
</tr>
<tr>
<td>L.FemurHead</td>
<td>0 (124)</td>
<td>0 [16]</td>
</tr>
<tr>
<td>R.FemurHead</td>
<td>0 (120)</td>
<td>0 [16]</td>
</tr>
<tr>
<td>CoccyxTip</td>
<td>0 (118)</td>
<td>0 [16]</td>
</tr>
<tr>
<td>PubicSymph.Top</td>
<td>0 (133)</td>
<td>0 [23]</td>
</tr>
<tr>
<td>SternumTip</td>
<td>3 (51)</td>
<td>1 [22]</td>
</tr>
<tr>
<td>AortaBend</td>
<td>0 (31)</td>
<td>1 [53]</td>
</tr>
<tr>
<td>Brachioceph.</td>
<td>1 (35)</td>
<td>3 [132]</td>
</tr>
<tr>
<td>R.Kidney</td>
<td>2 (59)</td>
<td>5 [61]</td>
</tr>
<tr>
<td>L.Kidney</td>
<td>0 (71)</td>
<td>0 [76]</td>
</tr>
</tbody>
</table>
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

- 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** $C = \text{rigid } \theta + \text{deformable } S$
  - 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”  Painting by Picasso  Mona Lisa

no context  weak context  strong context
Regression-Based Object Detection: Basic Idea

- **Basic idea**
  - Regress the difference vector
    
    \[ d\theta = F_1(I(\theta)) \]
  - Estimate the ground truth
    
    \[ \theta_0 = \theta + d\theta = \theta + F_1(I(\theta)) \]

One scan solution!
Two Questions?

- Does such an oracle $F_1$ exist?
  - Context in anatomy and appearance at a *global* level

- How to learn the oracle $F_1$?
  - Annotated database & machine learning
Robust Detection

Algorithm

- Sample
  \( \{\theta^{<1>}, \theta^{<2>}, \ldots, \theta^{<M>}\} \)

- Estimate
  \( d\theta^{<m>} = F_i(I(\theta^{<m>})) \)

- Predict
  \( \theta_0^{<m>} = \theta^{<m>} + d\theta^{<m>} \)

- Fuse by averaging
  \( \theta_0 = M^{-1} \sum_{m=1:M} \theta_0^{<m>} \)
Improved Localization

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

- **Weighted averaging**
  \[
  \theta_0 = \frac{\sum_j p_d^{<j>} \theta_0^{<j>}}{\sum_j p_d^{<j>}}
  \]

- **Faster computation**
  - Early stop
Regression-Based Deformable Shape Inference

- Basic idea $S = F_2(I(\theta_0))$. $I(\theta_0)$: Estimated ground truth patch

- Does such an oracle $F_2$ exist?
  - Context in shape and appearance at a local level

- How to learn the oracle $F_2$?
  - 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**
    \[
    S^{<k>} = F_2(I^{<k>})
    \]
  – **Fuse**
    – Build a nonparametric kernel density $p_s(S)$
    – Weighted averaging
    \[
    S = \frac{\sum_k p_s^{<k>} S^{<k>}}{\sum_k p_s^{<k>}}
    \]
Learn a score function $s(l, C)$ using classification, regression, or ranking.

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

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

- Efficient anatomy detection from 3D volumes
- Rigid parameterization (9D)
  - 3 for translation $\alpha$
  - 3 rotation $\beta$
  - 3 for anisotropic scale $\gamma$

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 $\Theta$
  \[
  (\alpha, \beta, \gamma) = \arg \max \{ (\alpha, \beta, \gamma) \in \Theta \} \ Pr(\alpha, \beta, \gamma|V)
  \]

- **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
Marginal Space Learning (MSL)

- **Offline learning**
  - Break down the learning complexity
    \[
    \Pr(\alpha, \beta, \gamma | V) = \frac{\Pr(\alpha | V) \times \Pr(\alpha, \beta | V) / \Pr(\alpha | V)}{\Pr(\alpha, \beta | V) / \Pr(\alpha, \beta | V)}
    \]
    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
Hierarchical MSL – Image Pyramid to Improved Robustness [Sokfa et al, CVPR 2010]
**Example: Liver segmentation [Ling et al. CVPR 2007]**

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

**Mesh Hierarchical**
- Base mesh $M^0$
- Hierarchical meshes
  \[
  \begin{align*}
  M^0 &= \downarrow (M^3) \\
  M^1 &= \downarrow (M^0) \\
  M^2 &= \downarrow (M^1)
  \end{align*}
  \]

**Statistical Shape Model**
\[
x = \mu + \sum_i c_i V_i
\]
- shape
- mean shape
- shape components
- shape coefficients

**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]

- 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

Special features + PBT
Region growing + heuristics
CT volume

DLL with custom API
IDTK

Body region estimation and landmark detection

Heart Isolation
Trachea detection and airways
Bounding box detection
Bounding box detection
Bounding box detection
Bounding box detection
Bounding box detection
Bounding box detection
Bounding box detection

Multi-scale boundary detection
Multi-scale boundary detection
Multi-scale boundary detection
Multi-scale boundary detection
Multi-scale boundary detection
Multi-scale boundary detection
Multi-scale boundary detection

Level set refinement & overlap removal

Heart
Airways
Liver
Left lung
Right lung
Left kidney
Right kidney
Prostate
Bladder

MSL + full-body landmark network
MSL + steerable features + PBT
PBT + Shape model + Correspondence resampling
Multi-region level set segmentation
Integrated Detection Toolkit (IDTK) --- Building the Whole Body Parsing Project

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