Perception and Computer Vision with Alternate Sensors

Dave Tahmoush

tahmoush@cs.umd.edu

Computer Vision and Pattern Recognition

- The computer vision algorithms for visible range sensors in day and office-like environments are good.
    - Object detection, tracking, and classification
- Opportunities exist in infrared (IR), depth, x-ray, thermal, ladar, radar, and other non-visible imaging sensors
    - Historically high cost, low resolution, poor image quality, lack of widely available data sets
    - Potential advantages to the non-visible part of the spectrum
    - Adapt computer vision techniques to new sensors
- Sensory technology is advancing rapidly and the sensor cost is dropping dramatically.
    - Image sensing devices with high dynamic range and high sensitivity have started to appear
    - Used in medical, defense, and automotive domains as well as home and office security.

Perception and Recognition Examples

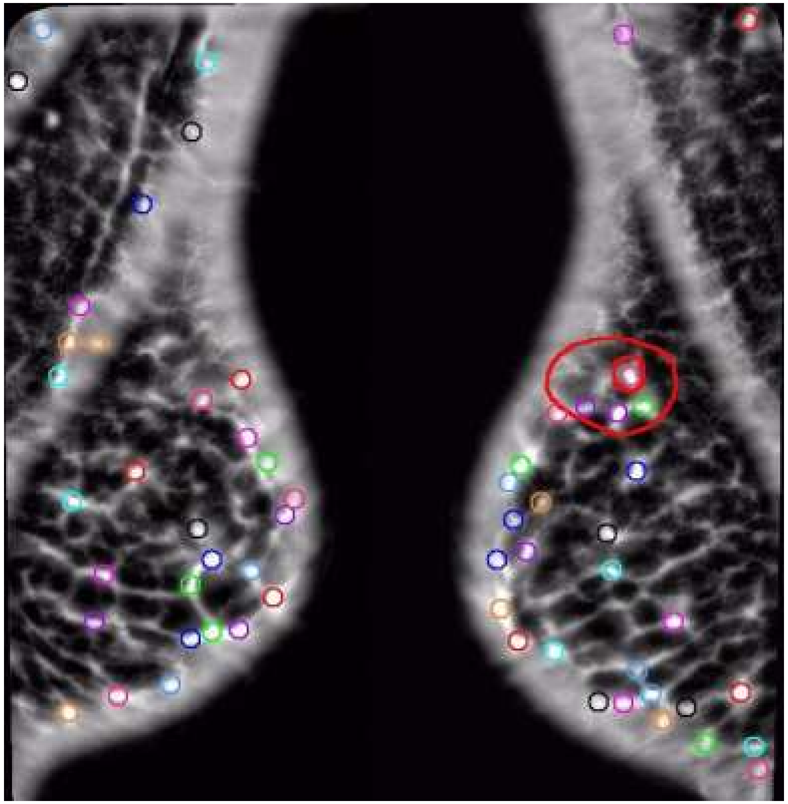
- In medical imaging, detect the existence and position of potential cancers in the body.
    - Noisy images that can be improved with computer vision techniques
    - Highly malignant cancer is isolated and recognized using custom features, clustering, and learning techniques.
- Computer vision is used for robotic perception
    - Primarily through ladar (lidar) but also through radar for robotic vehicles and video
- Video games use computer vision to determine the 3D motion of participants with Kinect-like sensors.
    - These sensors determine rudimentary point clouds
    - Motions are perceived and recognized
    - I will demonstrate how including ontological data can improve the feature set and the resulting perception
    - Will show results from online learning as well

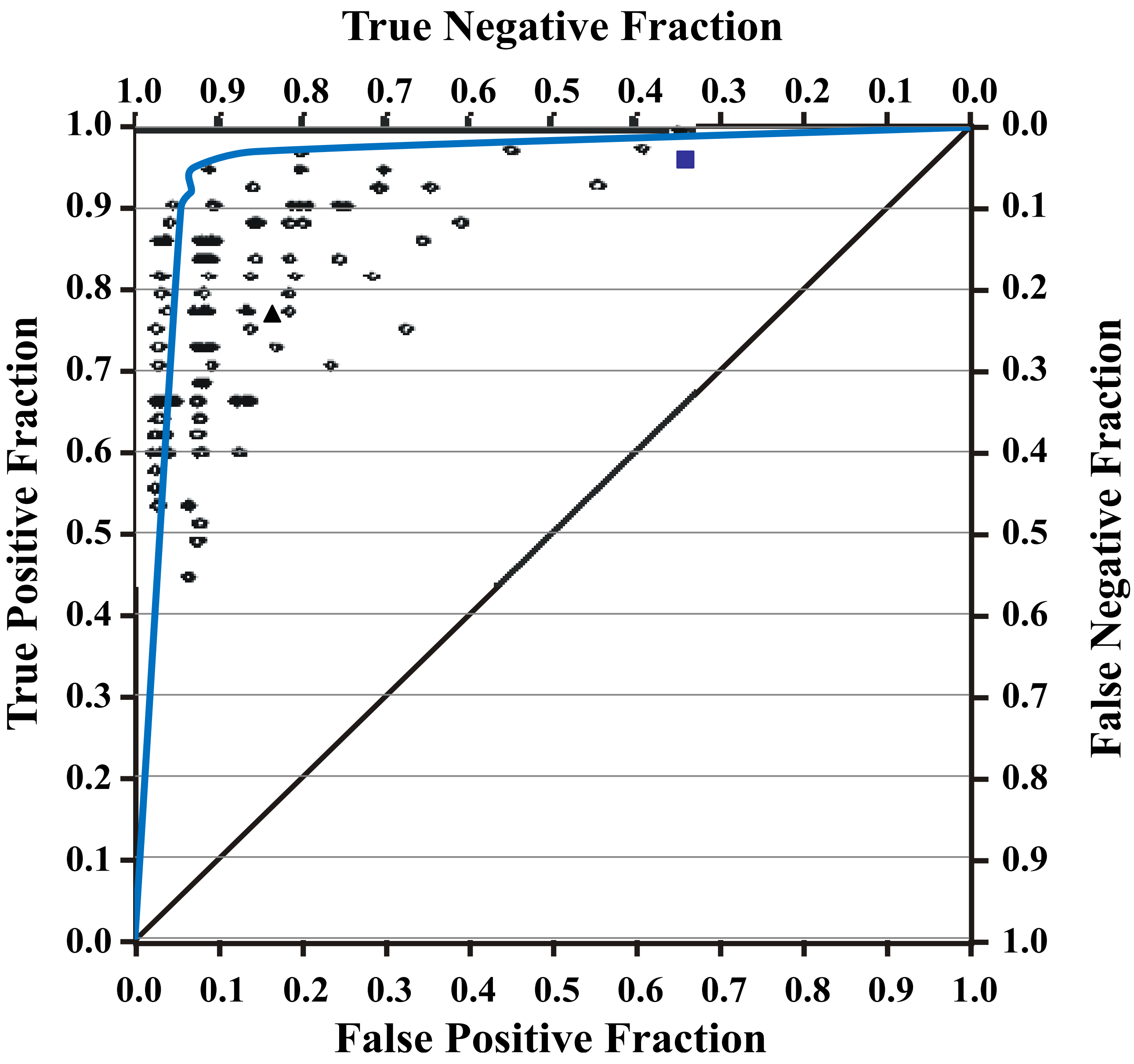
MEDICAL IMAGE ANALYSIS

Tahmoush, Dave. "Image similarity to improve the classification of breast cancer images." Algorithms 2, no. 4 (2009): 1503-1525.

MAMMOGRAM ANALYSIS

1. Tiny colored circles are extracted features
2. Compare the distribution of features from left to right breasts
   - Hypothesis: the cancer distorts the distribution of features
   - We create an analysis that effectively measures this small distortion
3. Registration of features from left to right images is possible but challenging --noisy, misaligned, features blocked or missing.
4. Must create a method that is robust at handling these difficulties.



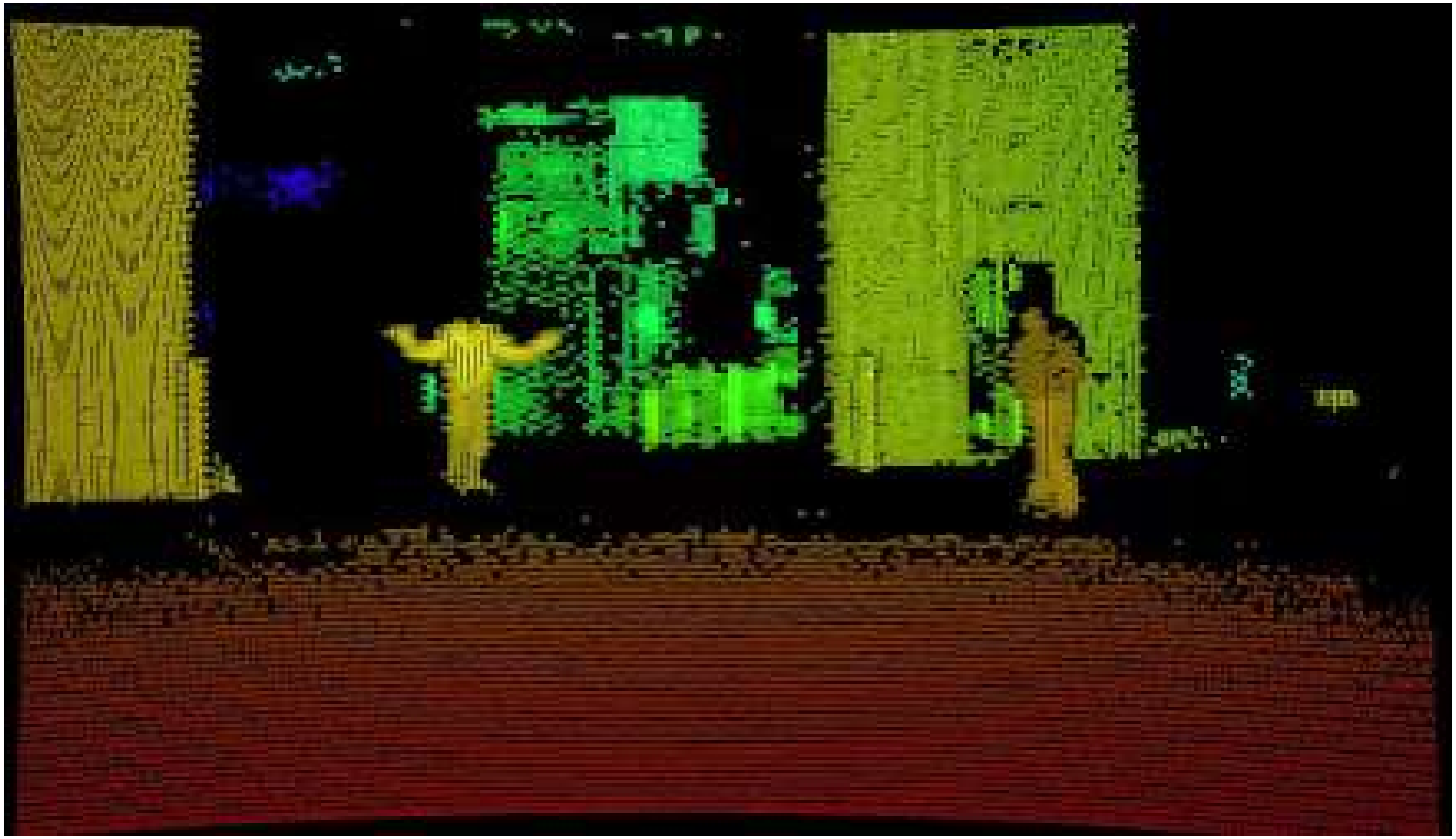


TPF *vs* FPF for 108 US radiologists in study [Beam et al], with our performance overlaid. Our technique compares favorably to radiologists.

3D VIDEO ANALYSIS AND ACTION RECOGNITION

Tahmoush, David. "Applying action attribute class validation to improve human activity recognition." In **CVPR** Workshop, pp. 15-21. 2015.

Ladar View



- Single frame of ladar data
- How we are viewed by robots

Results

-Found initial classification approach that worked well

-Estimated value of adding online capabilities

-5 of 6 of the improvable classes did improve

-No classes got worse

-Reduced misclassifications by 40%

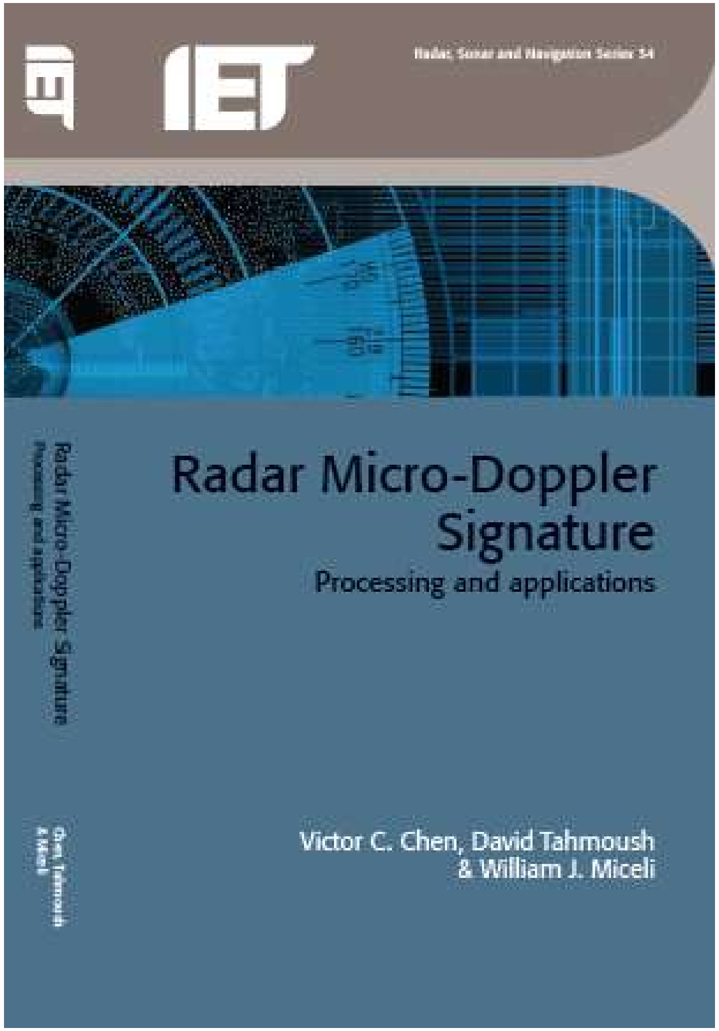
Method	Accuracy
Li et al	71.9%
Xia et al	85.5%
Yang et al	84.1%
Chen et al	83.3%
Initial Method	88.7%
Online Method	93.5%

More Perception and Recognition Examples

- In aerial imaging
  - Wide area motion imagery (WAMI)
  - Midwave Infrared
  - Full motion video
- In High-dimensional data
  - Curse of dimensionality
  - Cancer recognition
- If time permits

Book

- Look for my book
- 2014 publication
- Applications of computer vision with radar imaging and video


Radar, Sonar and Navigation Series 34
IET
Radar Micro-Doppler Signature
Processing and applications
Victor C. Chen, David Tahmoush & William J. Miceli
Radar Micro-Doppler Signature Processing and applications
Chen, Tahmoush & Miceli

MEDICAL IMAGE ANALYSIS

Tahmoush, Dave. "Image similarity to improve the classification of breast cancer images." **Algorithms** 2, no. 4 (2009): 1503-1525.

INTRODUCTION TO MEDICAL IMAGES

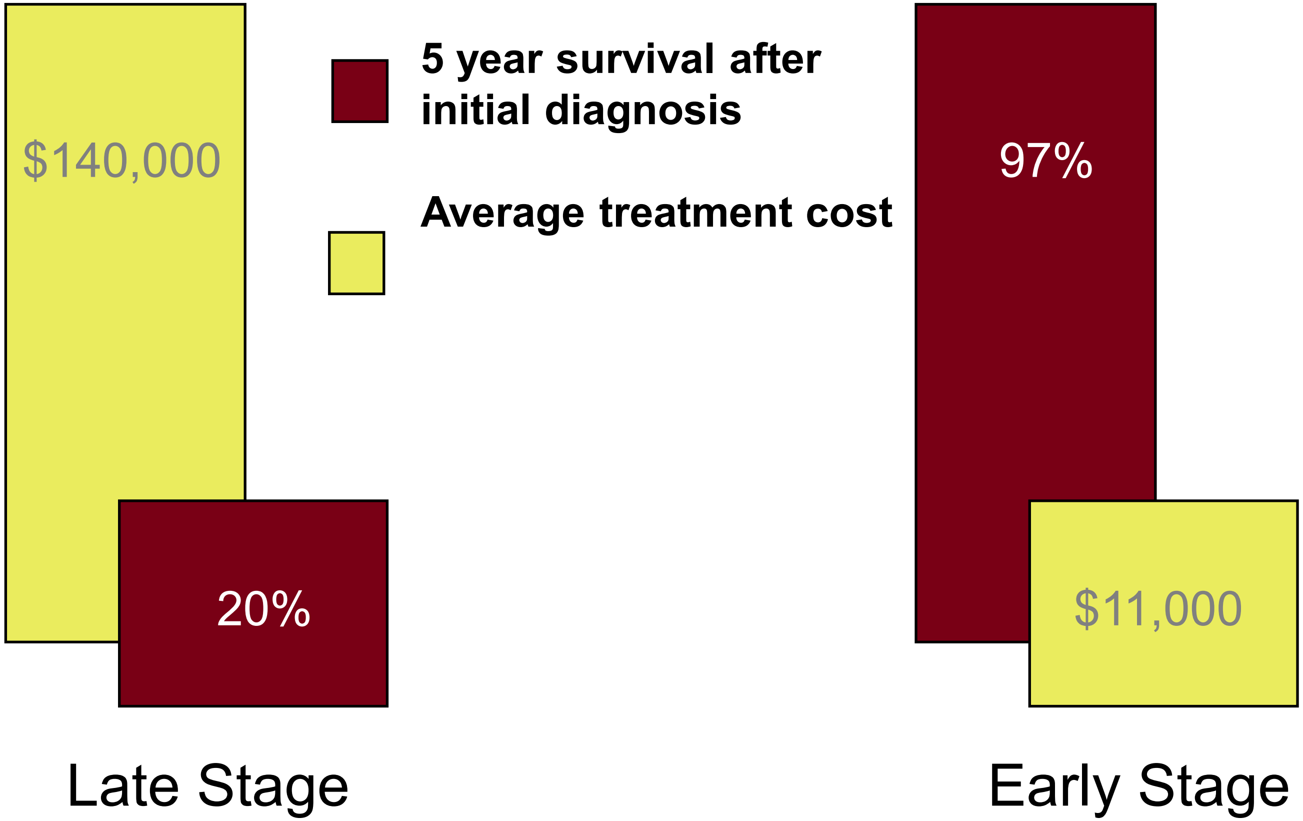
1. Several types of medical images
   - Focus is on mammograms
   - Also ultrasound, MRI
2. Medical images have been slow to digitize
   - Only 1.9% of mammograms are digital first
   - The rest are digitized from a film
3. High volume of images
4. Short viewing time for radiologists (difficult HCI problem)
5. Extremely low incidence of cancer found (3-10/1,000)
6. Sensitivity* of human screening mammography: ~ 80%
7. Specificity** of human screening mammography: 90 - 95%
8. Better detection means earlier detection and higher survival rate

* Sensitivity = TP / TP + FN ** Specificity = TN / TN + FP

TP = True Positives FN = False Negatives

TN = True Negatives FP = False Positives

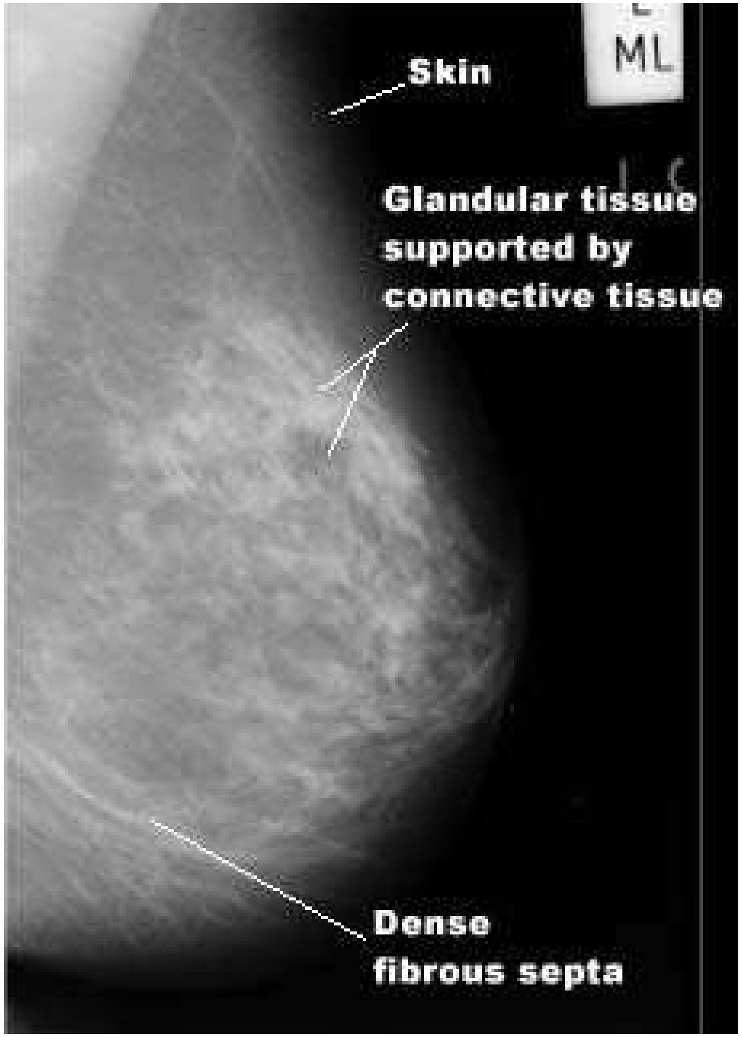
EARLY DETECTION BENEFITS



Source: American Cancer Society, 1999 Breast Cancer Facts and Figures

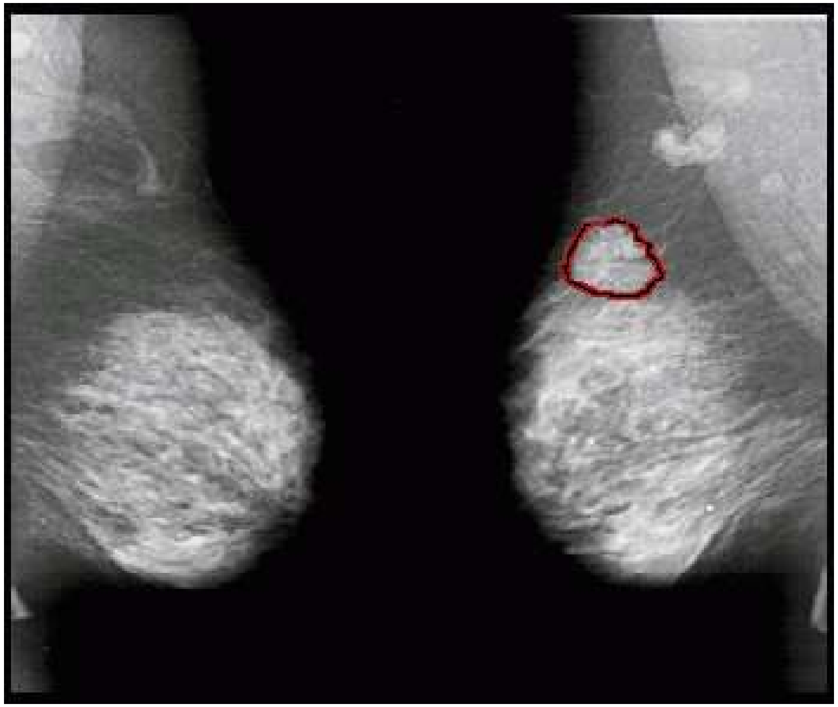
MAMMOGRAM BACKGROUND

1. A mammogram is an X-ray of the breast
2. There are many normal structures in the breast that absorb X-rays from the mammogram.
3. Absorption is shown in white
4. Note the chest wall in the upper left of the image
5. Mammogram is a projected image of superimposed breast structures
6. Normal structures can obscure cancerous structures
7. Septa are dense collagen fibers that provide structure in the body
8. A lot of non-cancerous structure



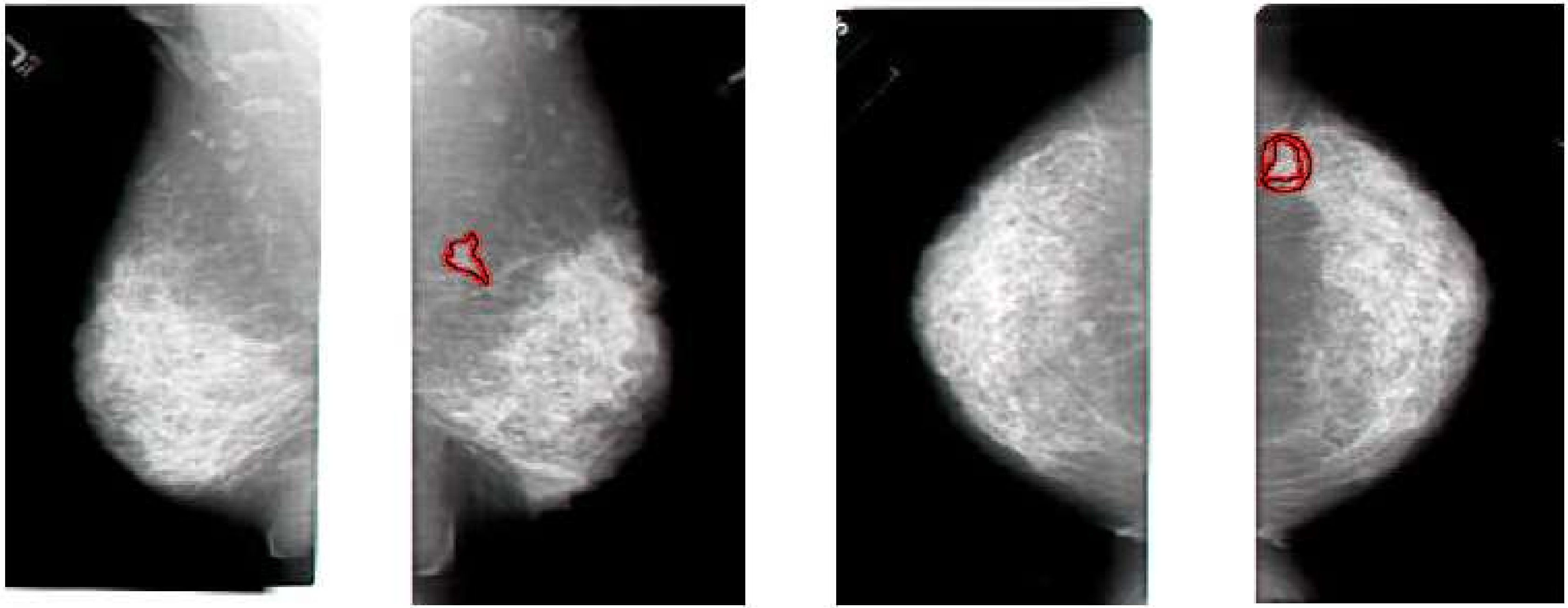
MAMMOGRAM BACKGROUND

1. Cancerous area is outlined in red/black
2. Texture of cancer can be similar to some of the normal tissue
4. Image comparison could be helpful in this case to recognize the cancer
5. Since the images come in sets, the non-cancerous cases are examples of similar images, while the cancerous cases are examples of dissimilar images, and these examples can be used to determine image classification.



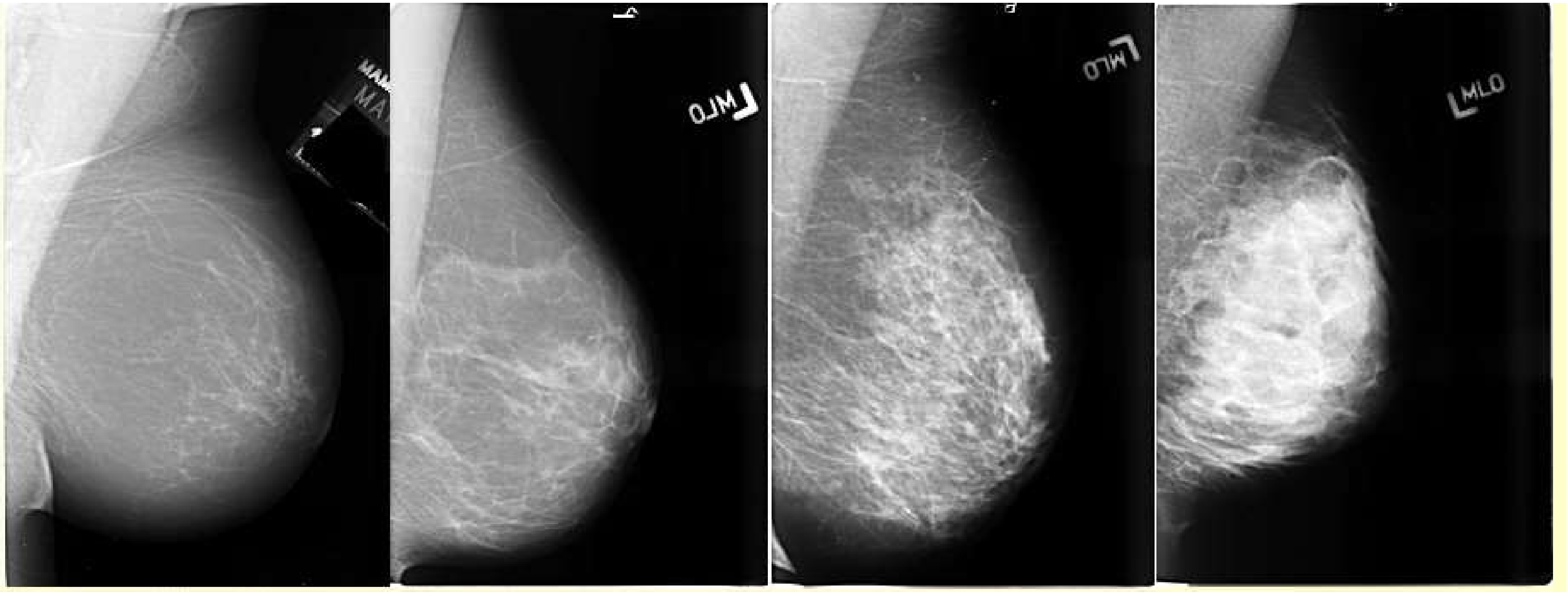
MAMMOGRAMS

1. The typical set of four images that make up a screening mammogram
   - The side view (MLO) of the left breast, the side view of the right breast
   - The top view (CC) of the left breast, the top view of the right breast
   - The cancerous areas are outlined in red
   - Note that the cancer is apparent in both images of the same breast, which provides additional information for the analysis
2. Breast tissue is compressed in the imaging process, making 3-D reconstruction challenging, and also only two images per breast



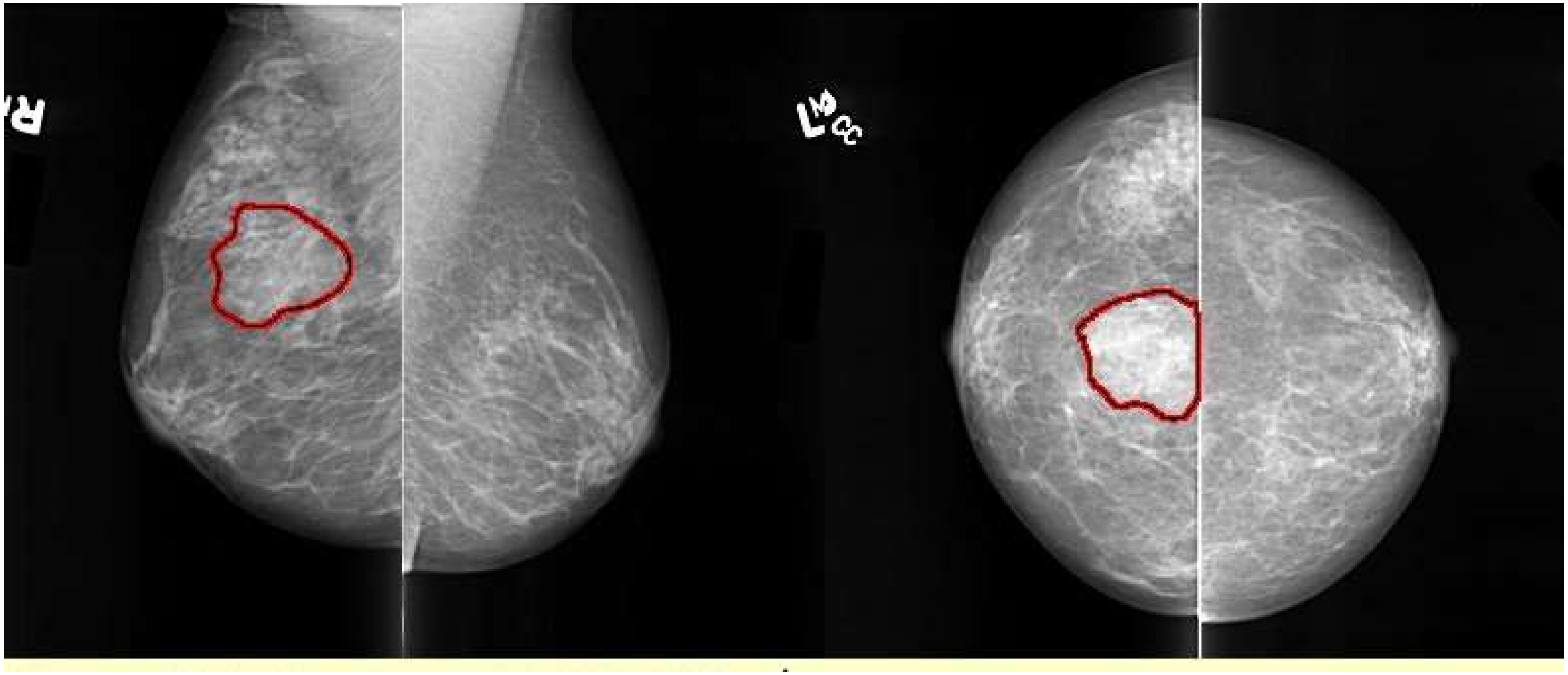
MAMMOGRAM VARIABILITY

1. Breast vary in density, and a range of 1-4 is shown from left to right.
2. From “almost entirely fat” to extremely dense
3. A dense breast pattern has not been proven to be a greater risk for breast cancer, but does limit the ability of a mammogram to detect breast cancer
4. Breasts usually develop symmetrically, but differences in the symmetry of breast tissue patterns or breast size are not necessarily abnormal or indicative of cancer



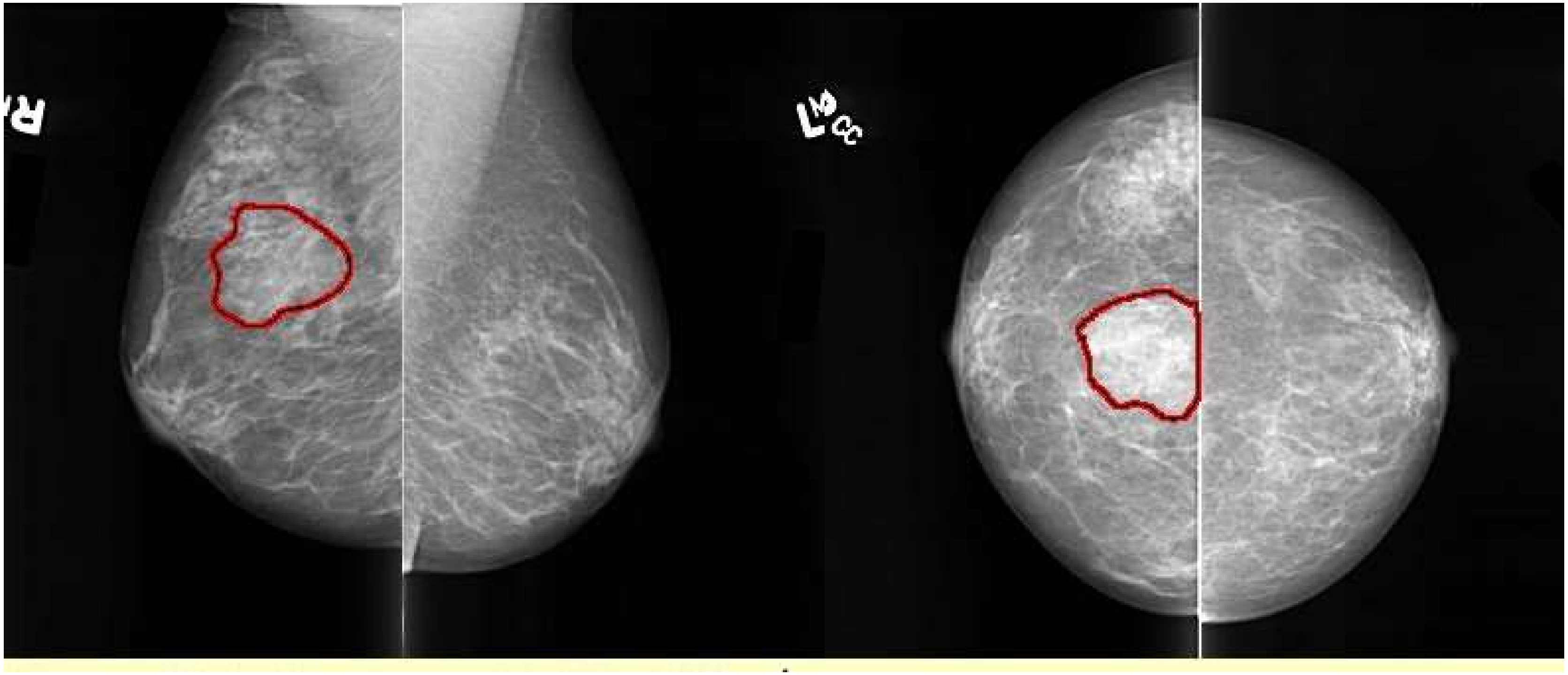
MAMMOGRAM ASYMMETRY

1. An asymmetric area may be indicative of
   - a developing mass
   - a variation of normal breast tissue
   - postoperative change from a previous biopsy
   - poor positioning and compression during imaging



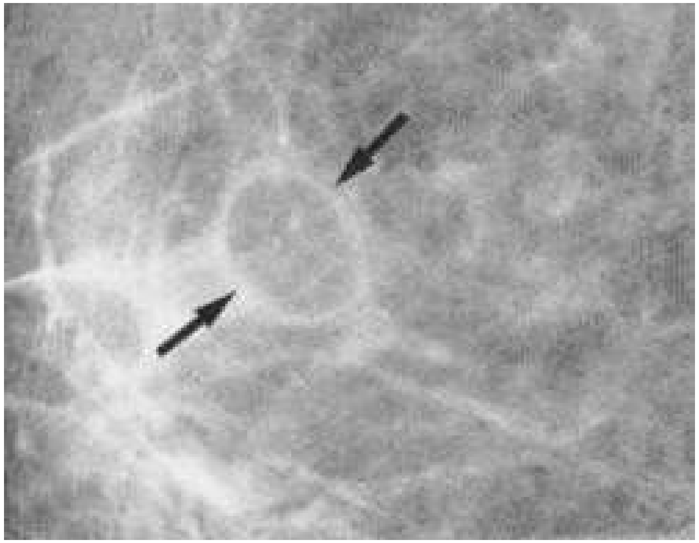
MAMMOGRAM ASYMMETRY

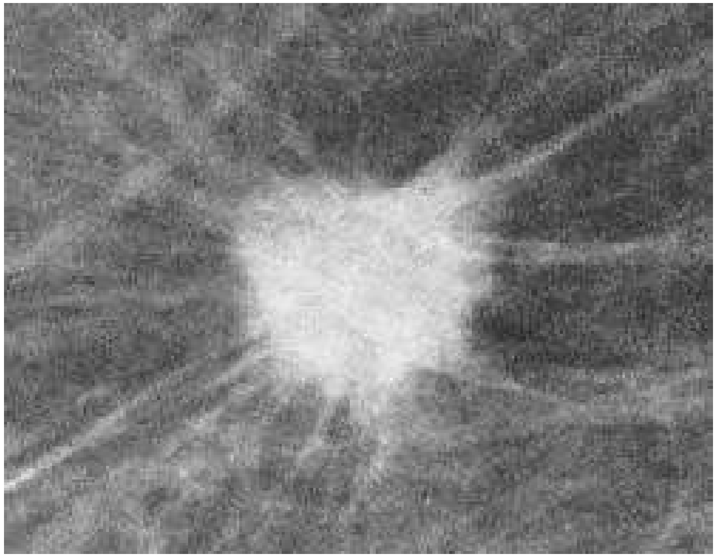
1. The appearance of asymmetries due to positioning and compression during imaging is often the result of superimposition of normal breast structures.
2. True breast asymmetry is three-dimensional and should be present on both MLO and CC views (side and top).
3. Asymmetry could be a benign variation of asymmetric breast tissue or a focal asymmetric density that may represent a significant mass.



MAMMOGRAM MASSES

1. A mammographic image of a circumscribed lesion is on the left.
   - The ring structure is one of the key features that can be picked out of a mammogram.
   - Usually benign
2. A mammographic image of a spiculated lesion is on the right.
   - The bright center or core is one feature of these lesions, as well as the radiating lines which are called spiculations.
   - Usually malignant.
   - Some of the most difficult to detect [Lui 2001]



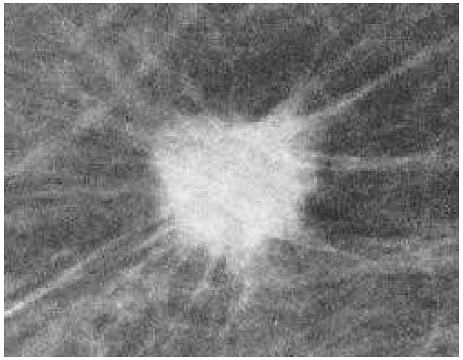


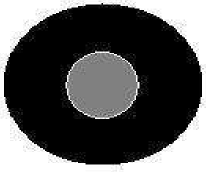
TYPICAL IMAGE CLASSIFICATION

1. Examples from the Caltech 101 image set
2. Non-structured, with arbitrary camera angles
3. Not image sets but single images
4. In these examples, simple approaches like color histograms can classify well
5. We know we are looking at mammograms, want to know whether there is cancer
6. Like asking whether the alligator has any cavities

DATA AND FEATURES

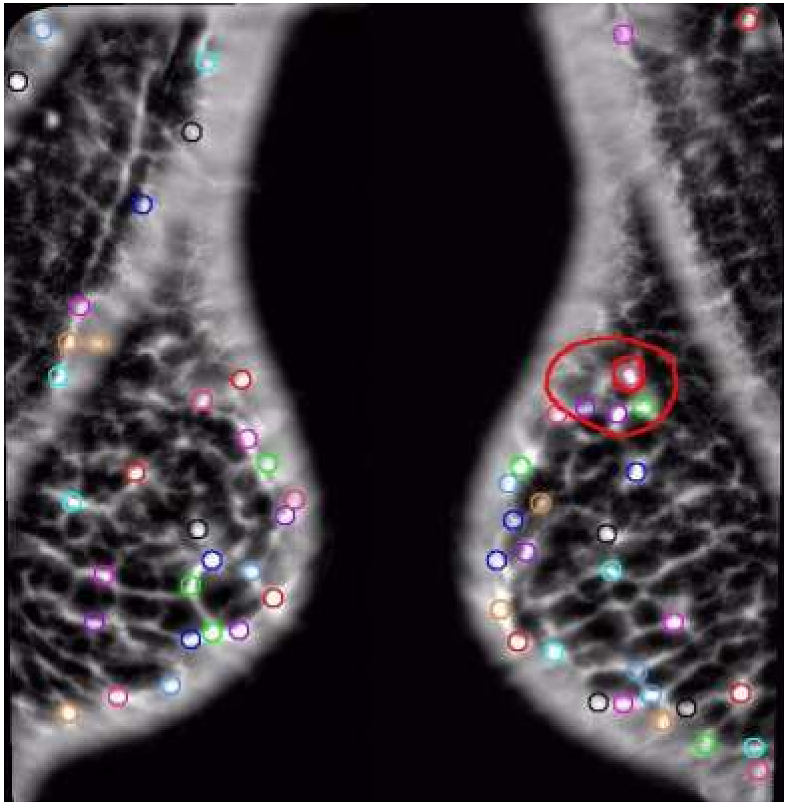
1. Used cases from the Digital Database for Screening Mammography.
2. 77 Normal cases and 77 cases with spiculated lesions. Focused on difficult to detect spiculated lesions.
3. The feature used was a multiscaled AFUM filter that detects the bright central core of the spiculated lesion
4. The CAD suspiciousness calculation is performed at each pixel location *(x,y)* in the images.
5. The minimum intensity I_{min} within r_1 is found, and then the fraction of pixels between r_1 and r_2 with intensities less than I_{min} is calculated.
6. This gives the fraction under the minimum (FUM) for one set of r_1 and r_2.
7. Averaging the FUM over a range of r_1 determines the average fraction under the minimum (AFUM)





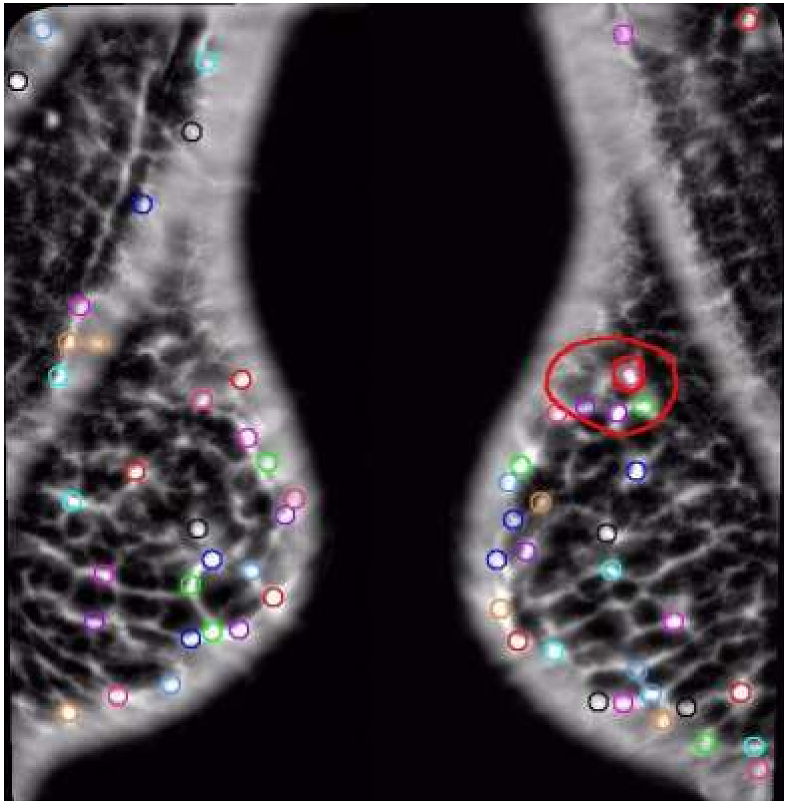
MAMMOGRAM FEATURES

1. AFUM features shown over a pair of mammogram images
2. Tiny colored circles are features
3. Thicker red lines are hand-drawn radiologist annotations of cancer
4. Cancerous area is detected by this feature, with tiny colored circle inside the thicker red lines
5. Features cluster around the cancer, multiple tiny circles inside the boundary of the thicker red lines
6. Many false positives with this feature
7. Need to determine how to classify images using this noisy feature
8. Like finding many teeth of the alligator, need to find the cavity

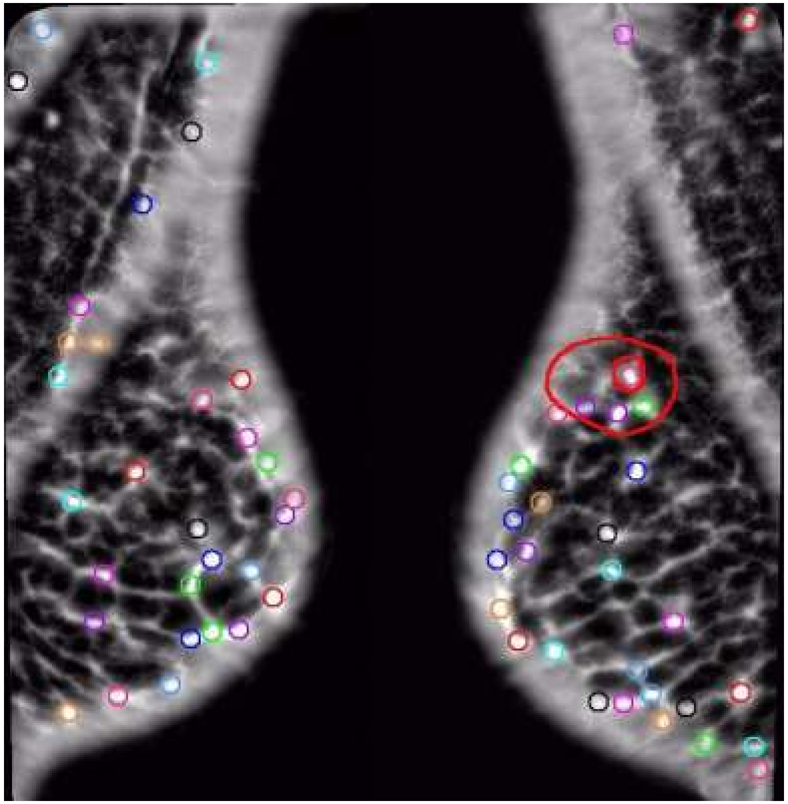


MAMMOGRAM ANALYSIS

1. Tiny colored circles are extracted features
2. Compare the distribution of features from left to right breasts
   - Hypothesis: the cancer distorts the distribution of features
   - We create an analysis that effectively measures this small distortion
3. Registration of features from left to right images is possible but challenging --noisy, misaligned, features blocked or missing.
4. Must create a method that is robust at handling these difficulties.



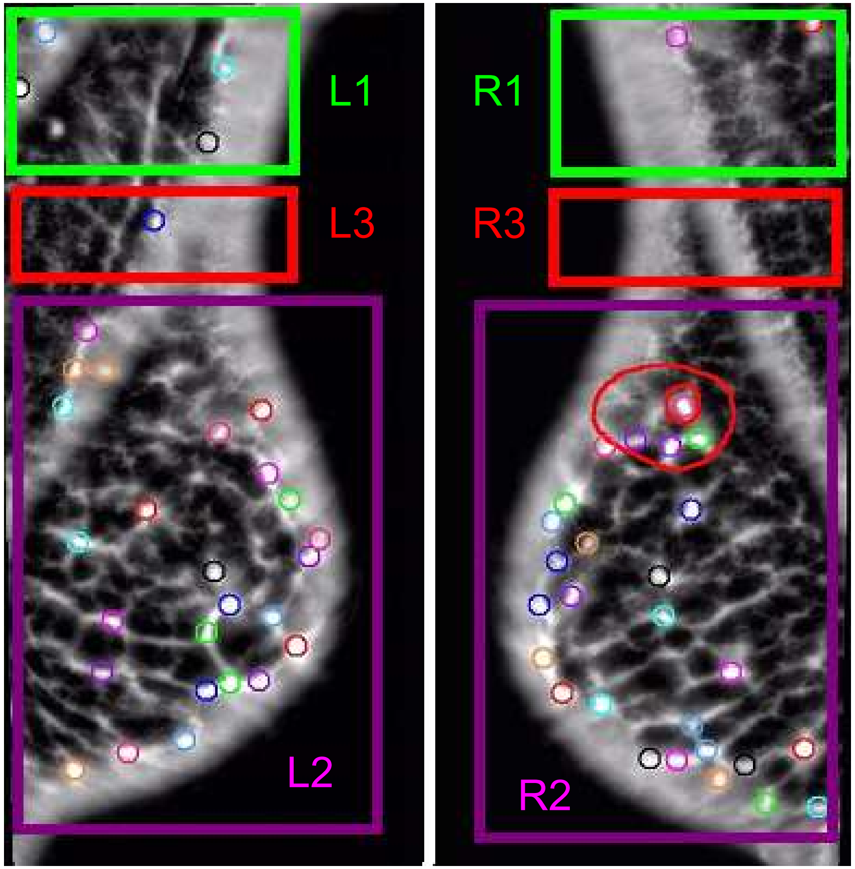
SIMPLIFIED ANALYSIS



1. Create a model for the comparison and learn the parameters
2. Must choose performance measure
   - The number of correct classifications
   - Weight the value of each correct classification to vary performance characteristics. For example, weight cancerous cases higher than normal cases
3. Data space is x,y coordinates of the features and the feature values f(x,y) to give 3D feature points
4. Use a fixed number of feature points for each breast, ranked by their feature value f(x,y)
5. Simplest model -- try to break up the space into clusters, learn the best parameters on a training set, and compare using a distance function.

SIMPLE ANALYSIS

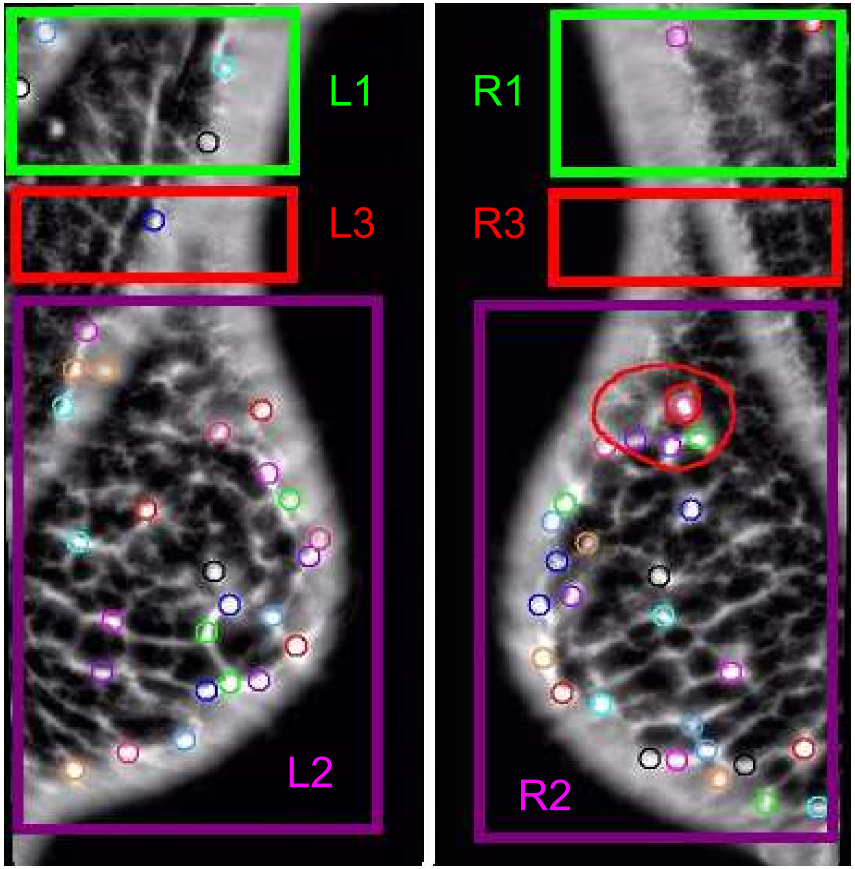
1. Learn the parameters of the clusters along with the threshold of the distance function *D*1.
2. Create clusters L1, L2, and L3, also R1, R2, and R3.
3. Clusters and their volumes are fed into distance function *D*1.
4. Image difference distance compared against learned threshold
5. if small distance then similar and no cancer.



$$D1 = \sum_{Clusters} |(\iiint_{\substack{Volume\\ of\\ Cluster}} df \sum_{i} \delta(f - \bar{a}_i) - \delta(f - \bar{b}_i))|$$

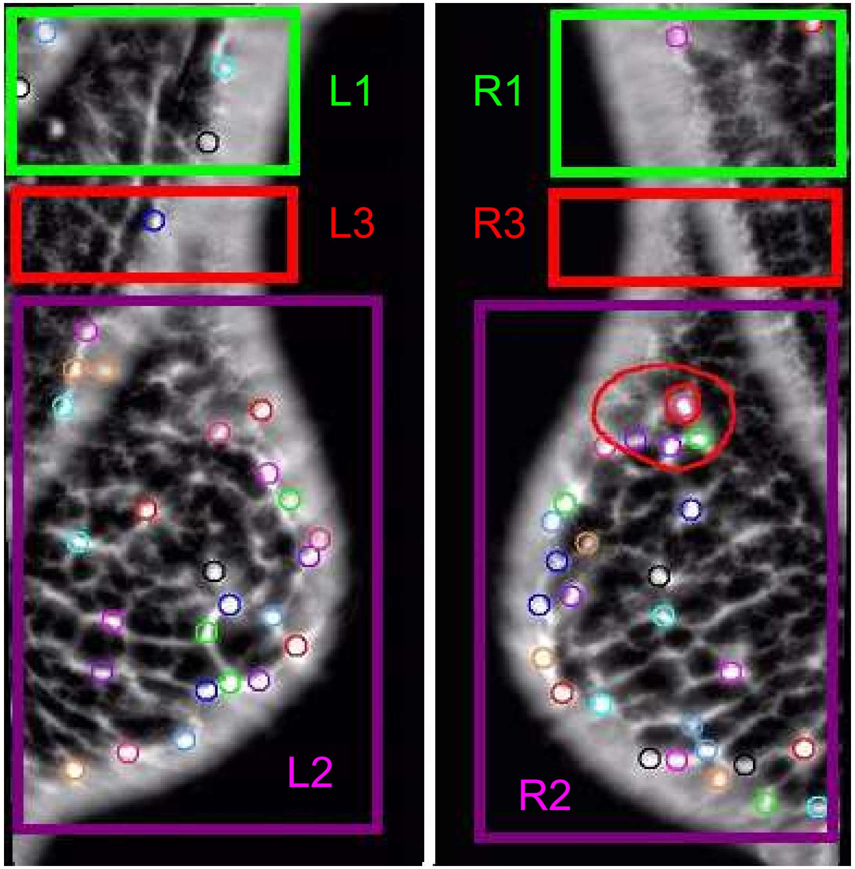
WHY WOULD THIS WORK?

1. Contextual similarity of the features
2. Spatial similarity of the distance function
3. Avoids direct image or feature registration
   - Very difficult
   - Inaccurate
   - Bilateral Subtraction, only works well on breasts with low density
   - Too much obscuring structure
4. Supervised learning
   - Makes method adaptable to many problem types
5. Analyzing small variations in the distribution



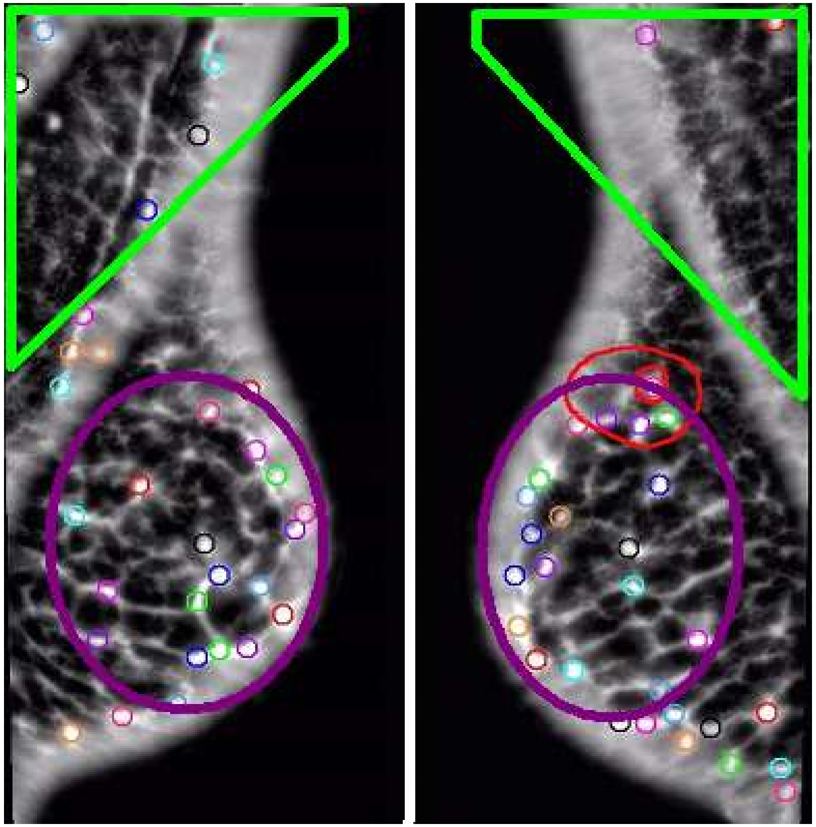
WHY WOULD THIS WORK?

1. Clustering provides flexibility to the spatial part of the analysis, not just fixed grid or pyramid
2. Flexible clustering allows machine learning to find the best structure
3. Utilize CAD prompts as a feature set for the classification of breast cancer images
   - Contextually very significant features
   - Difficult data set to work with since noisy
4. Clustering is separate from the classification
   - Can be used to find interesting areas of the images
   - Distance function gives classification of images



MORE COMPLICATED ANALYSIS

1. Flexible number of clusters
   - Cluster analysis can be non-space-filling and also non-disjoint
   - Greater range of shapes and flexibility of analysis
   - Cost is more parameters, greater risk of overfitting
2. Performance was better with more features
   - even though cancer is often in top eight features
   - The cancer distorts the distribution more than just by moving one feature.
3. More complicated distance functions are more effective



MORE COMPLICATED DISTANCE FUNCTIONS

Delta function δ

a_i and b_i are the feature vectors of the *i*th feature of the two images

The integral is over feature space, with the variable *f*

The sum over *i* is over the features

Equation D2 breaks the distance up into individual cluster distances

Equation D3 uses a probability of cancer density function Φ and a variable number of parameters per image

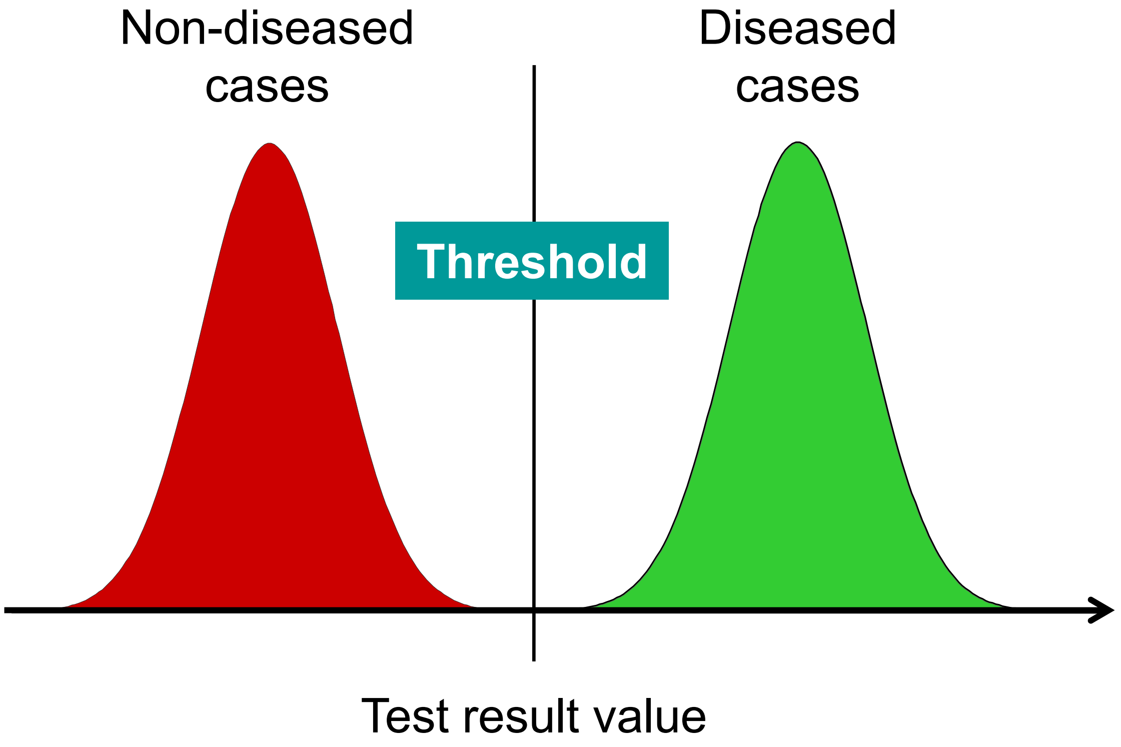
$$D1 = \sum_{Clusters} |(\iiint_{\substack{Volume\\ of\\ Cluster}} df \sum_i \delta(f - \overline{a}_i) - \delta(f - \overline{b}_i)) |$$

$$D2_{Cluster} = \iiint_{\substack{Volume\\ of\\ Cluster}} df \sum_i \delta(f - \overline{a}_i) - \delta(f - \overline{b}_i)$$

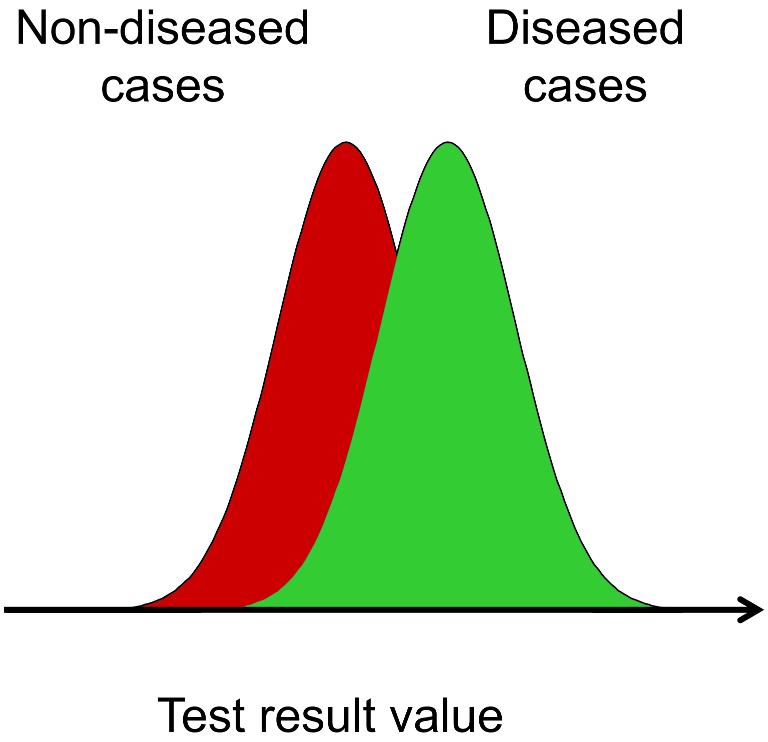
$$D3 = \sum_{Clusters} |(\iiint_{\substack{Volume\\ of\\ Cluster}} df \sum_i \Phi(f, \overline{a}_i) - \sum_j \Phi(f, \overline{b}_j)) |$$

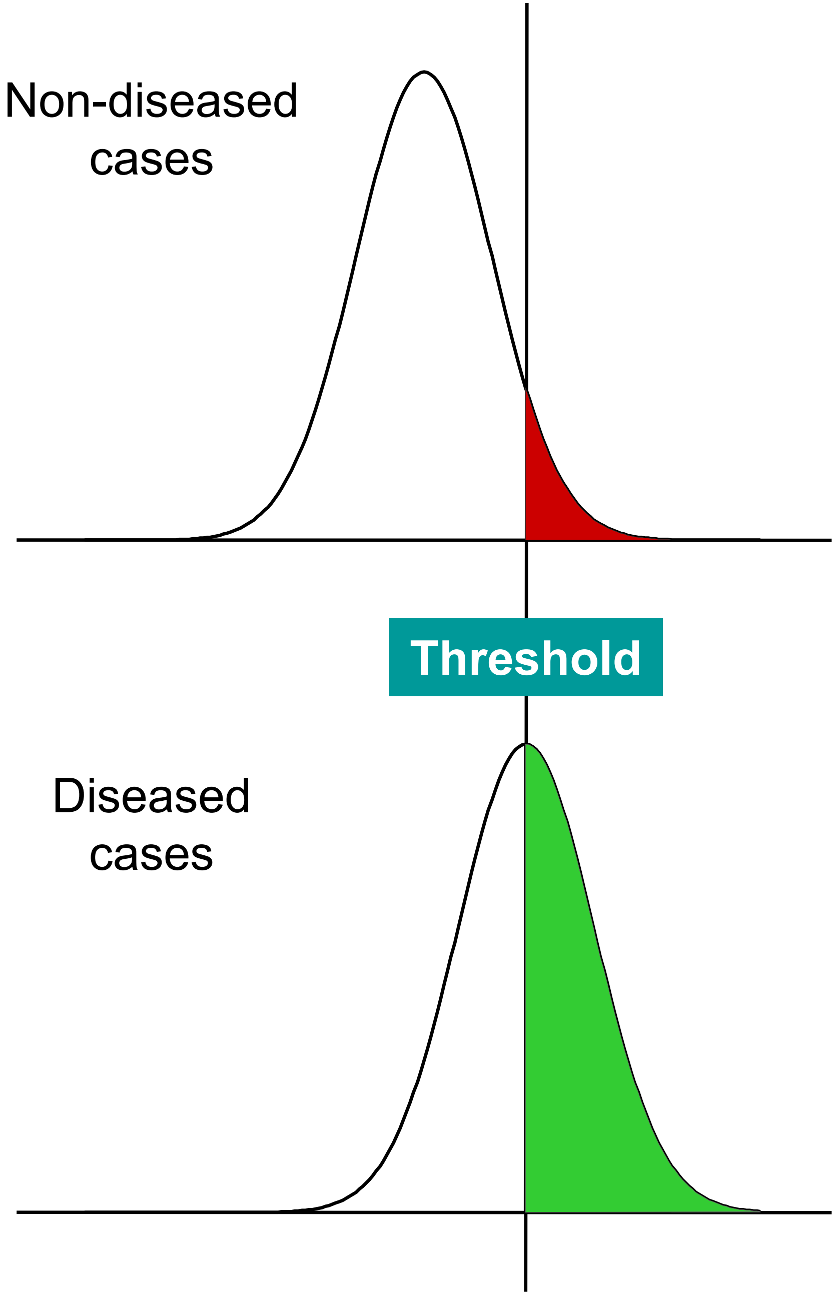
RESULTS

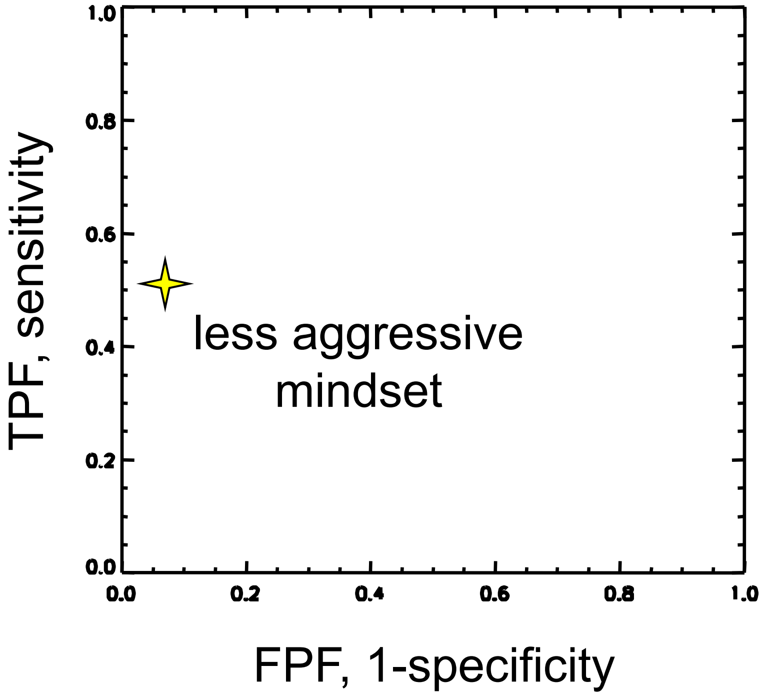
1. Use a Receiver Operating Characteristic (ROC) curve to demonstrate performance
   - Will give a short intro to ROC curves
   - Used in many two class systems
2. Approach compares well against commercial and wavelet methods
   - Outperforms each in overall performance, sensitivity, and specificity
3. Approach also compares well against humans
4. First method to outperform radiologists

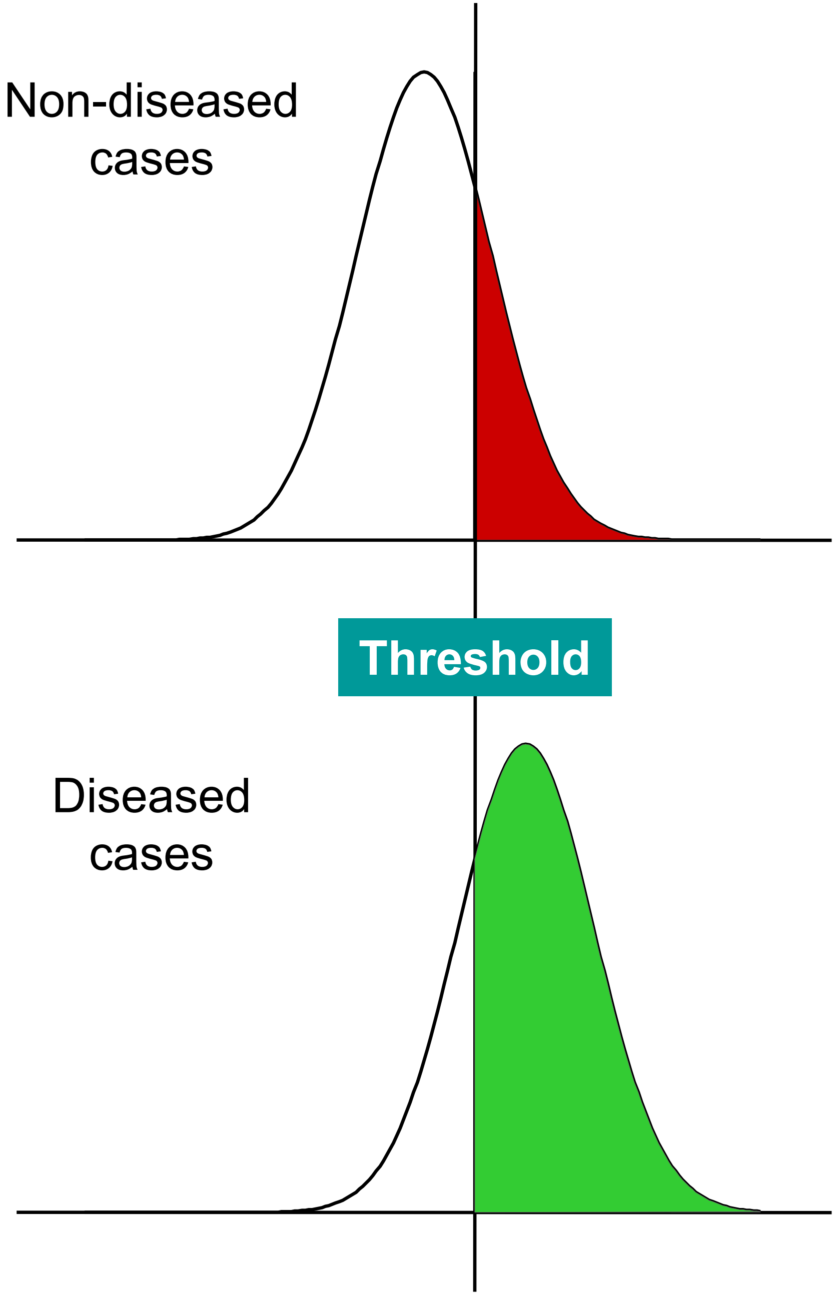

Non-diseased cases
Diseased cases
Threshold
Test result value

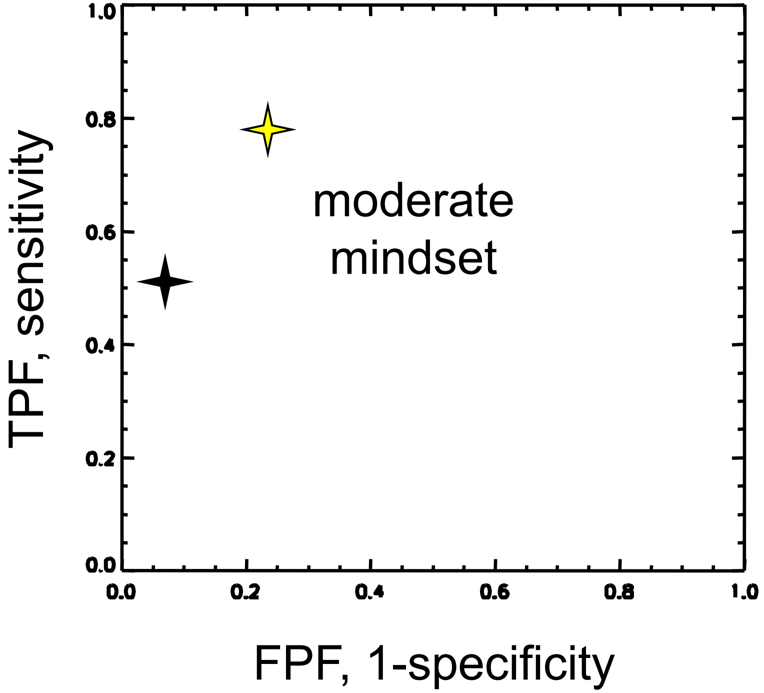
more typically:

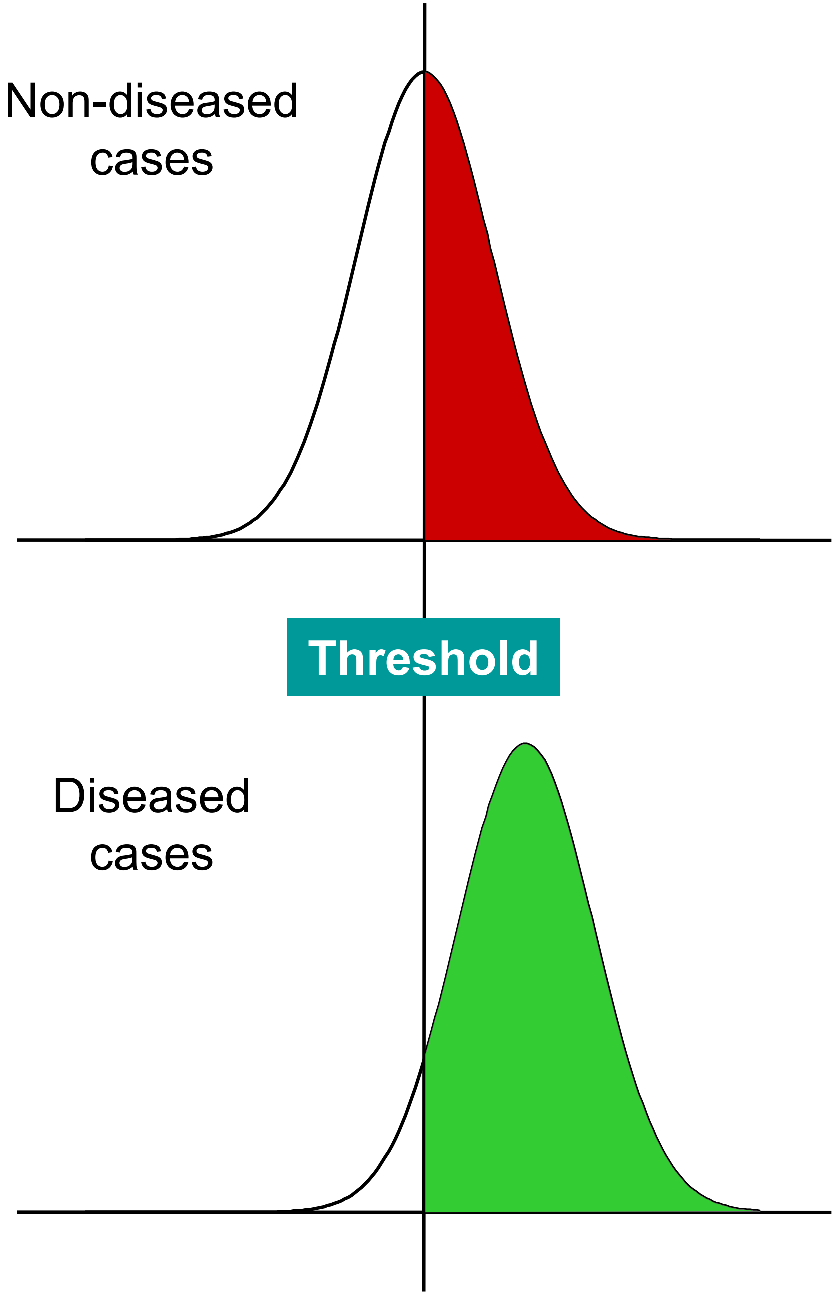


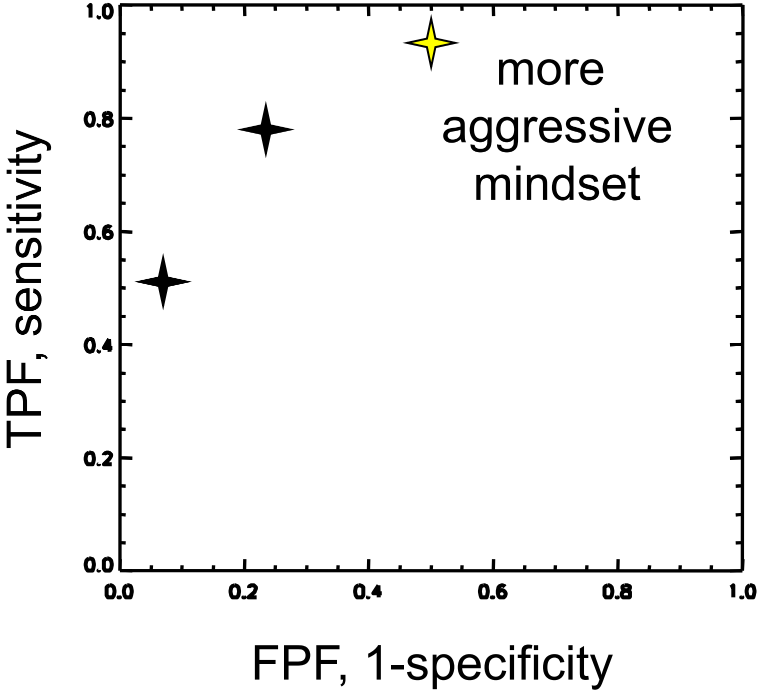

Non-diseased cases
Threshold
Diseased cases

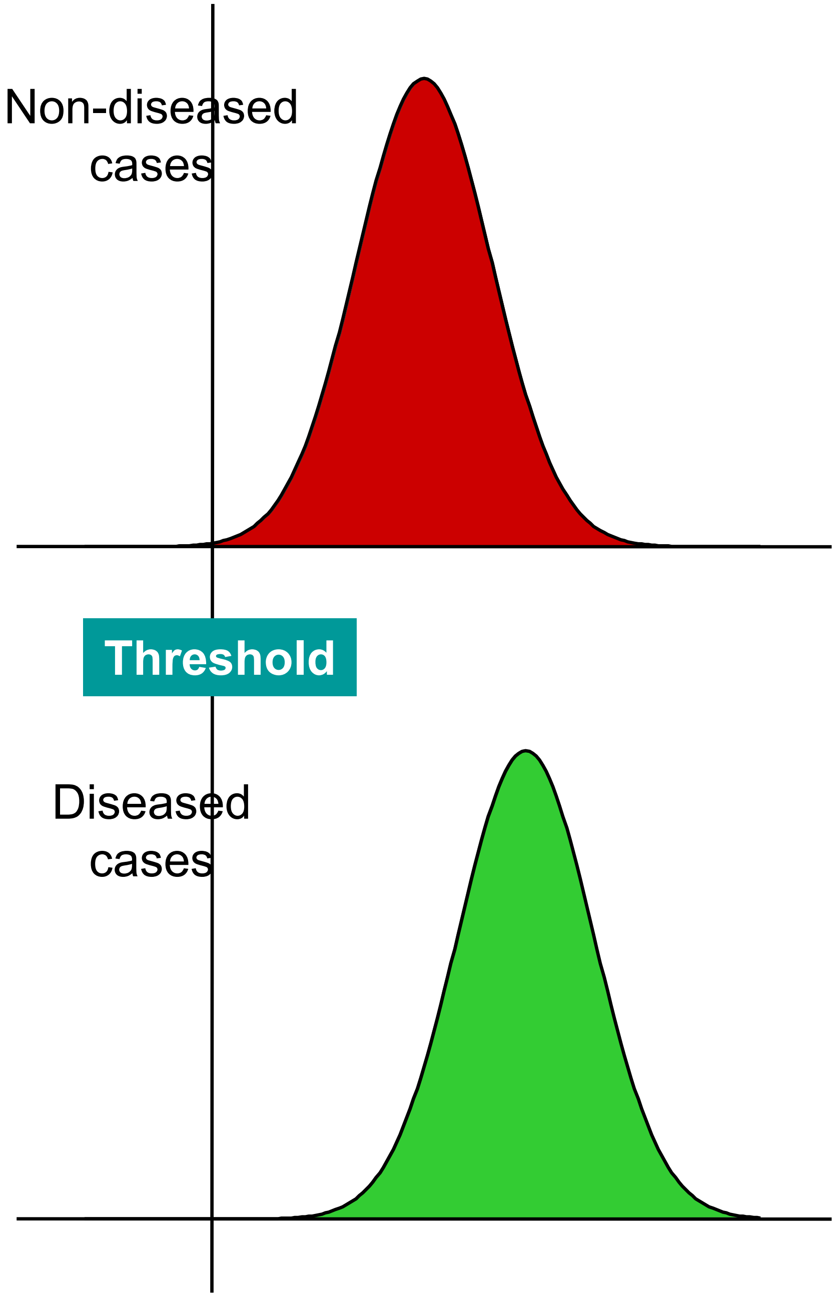

1.0
0.8
0.6
0.4
0.2
0.0
TPF, sensitivity
less aggressive mindset
0.0
0.2
0.4
0.6
0.8
1.0
FPF, 1-specificity


Non-diseased cases
Threshold
Diseased cases

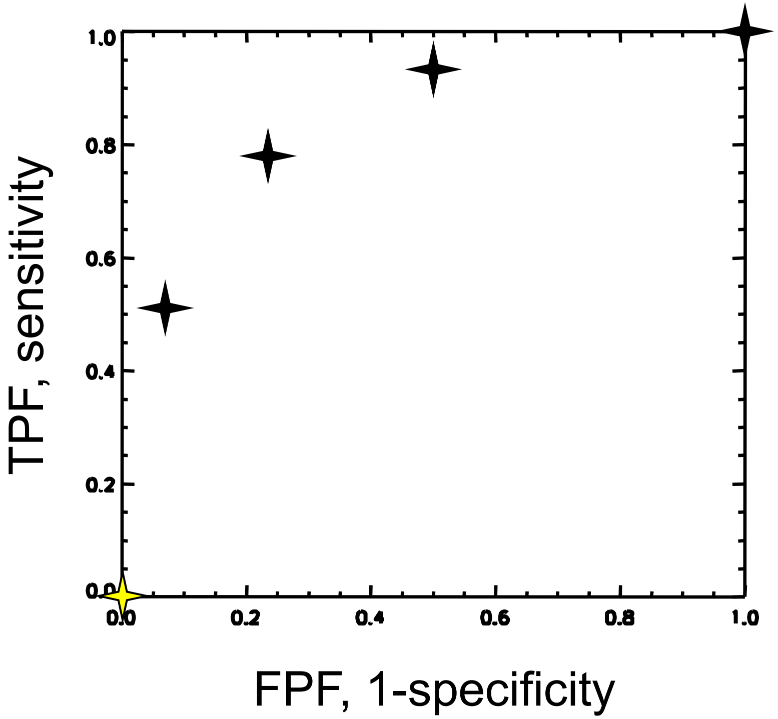

1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
TPF, sensitivity
FPF, 1-specificity
moderate mindset

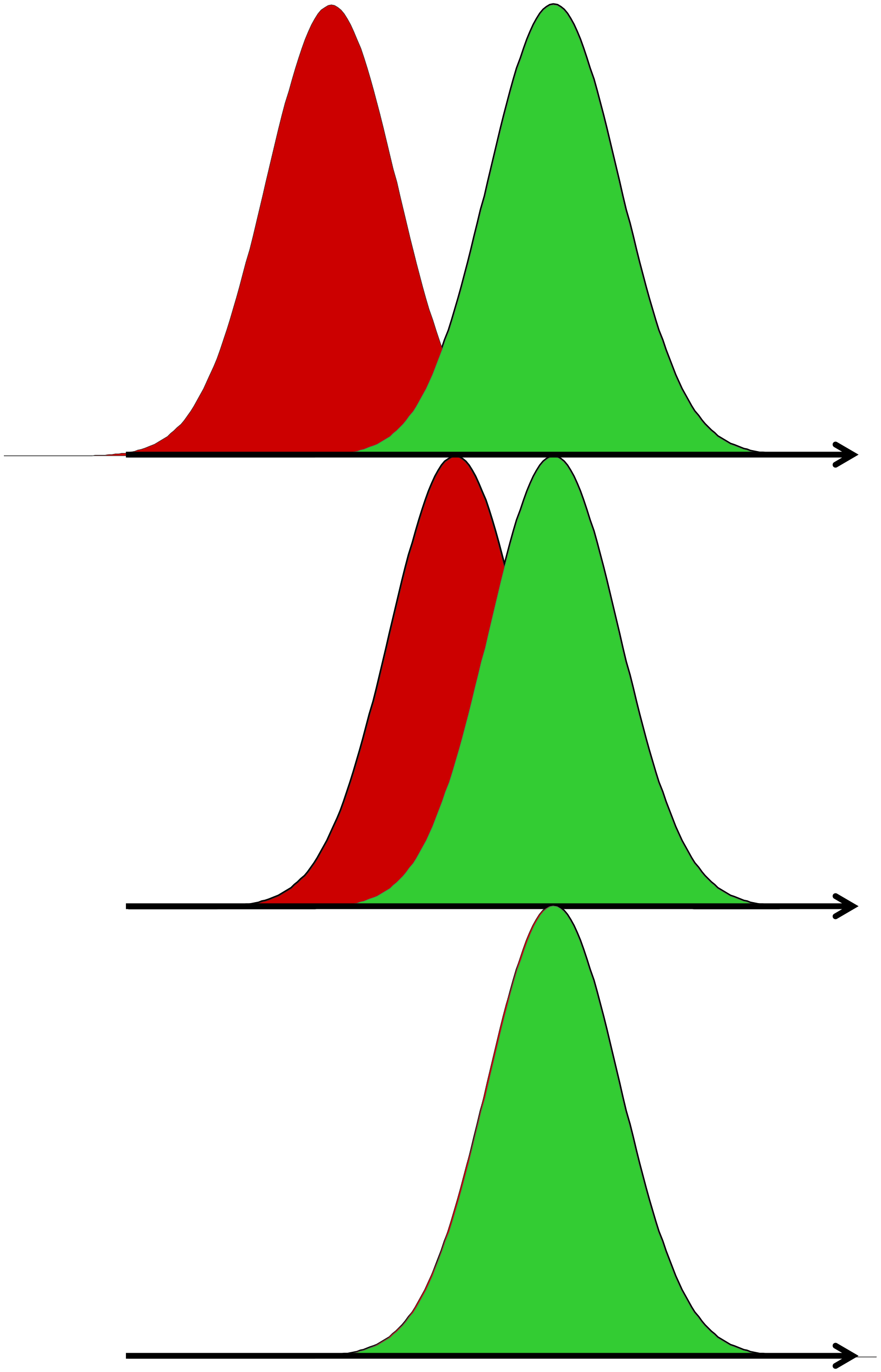

Non-diseased cases
Threshold
Diseased cases


1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
TPF, sensitivity
FPF, 1-specificity
more aggressive mindset

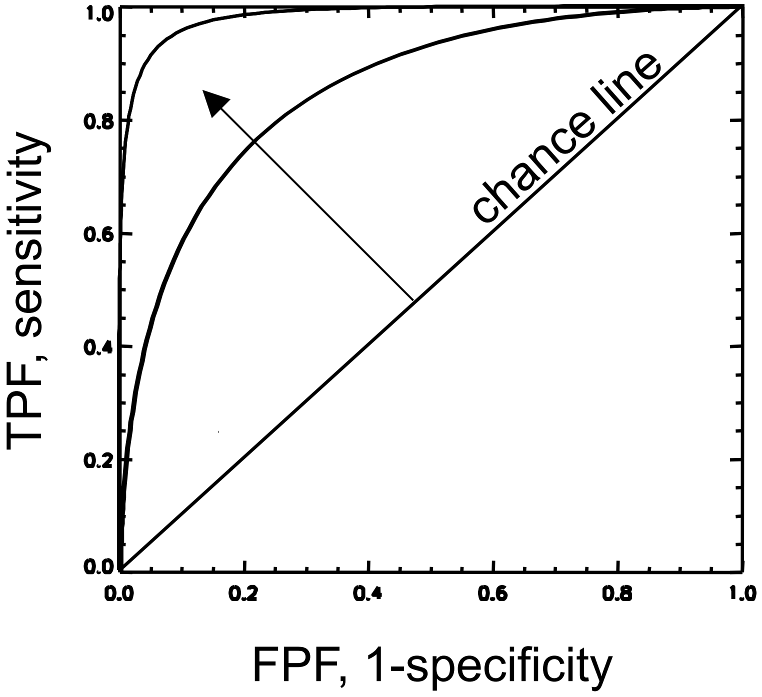


Entire ROC curve



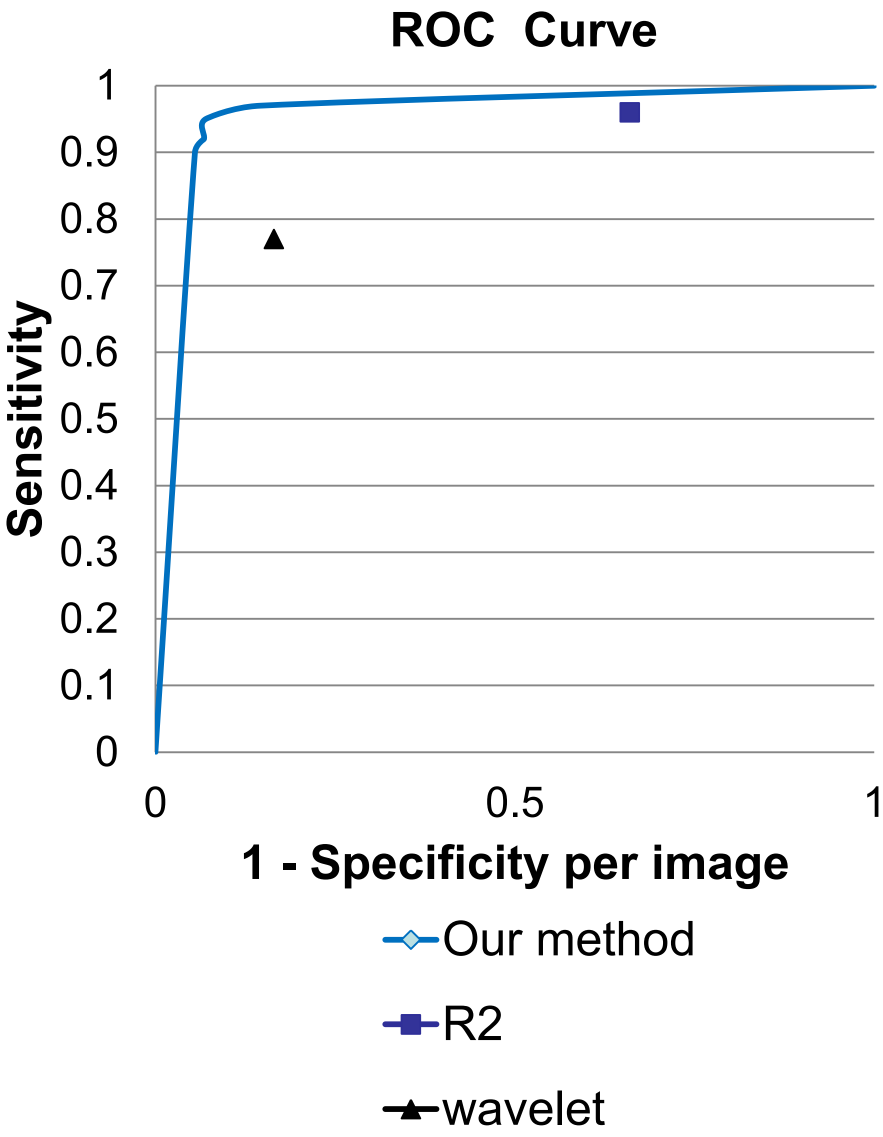


Entire ROC curve



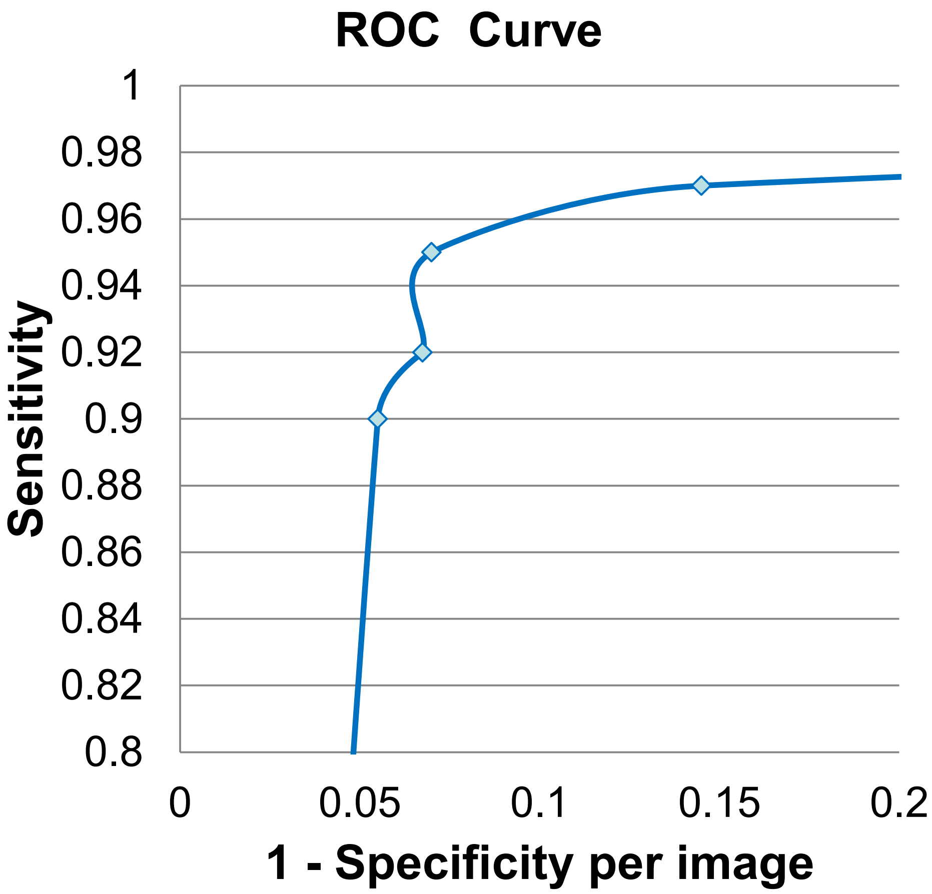
RESULTS

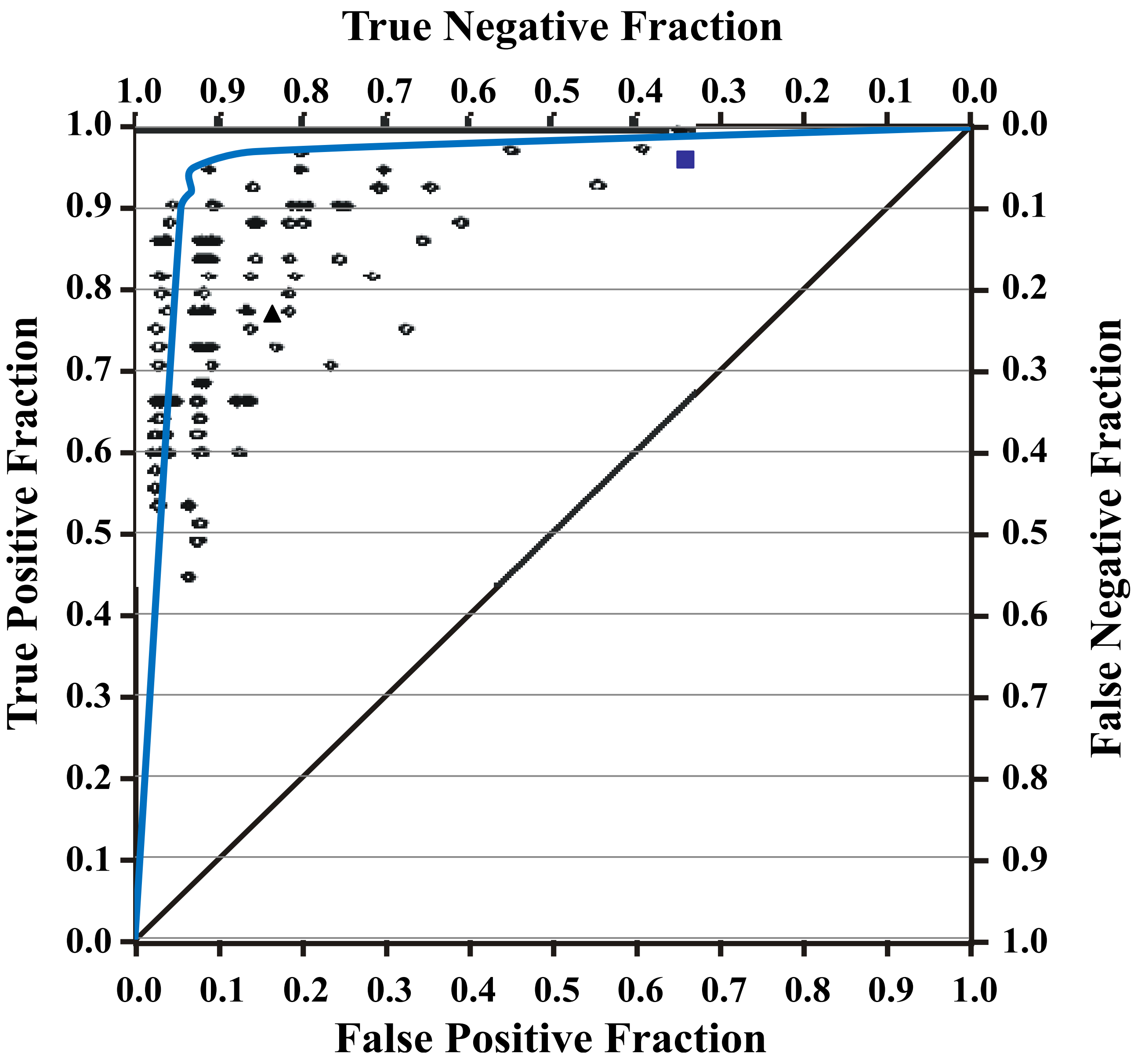
1. When used alone, we can achieve high true positive fraction (97%) at the expense of high false positives (15%)
2. More balanced at 95%-7%
3. Little bend in ROC curve due to fitting program



RESULTS

1. When used alone, we can achieve high true positive fraction (97%) at the expense of high false positives (15%)
2. More balanced at 95%-7%



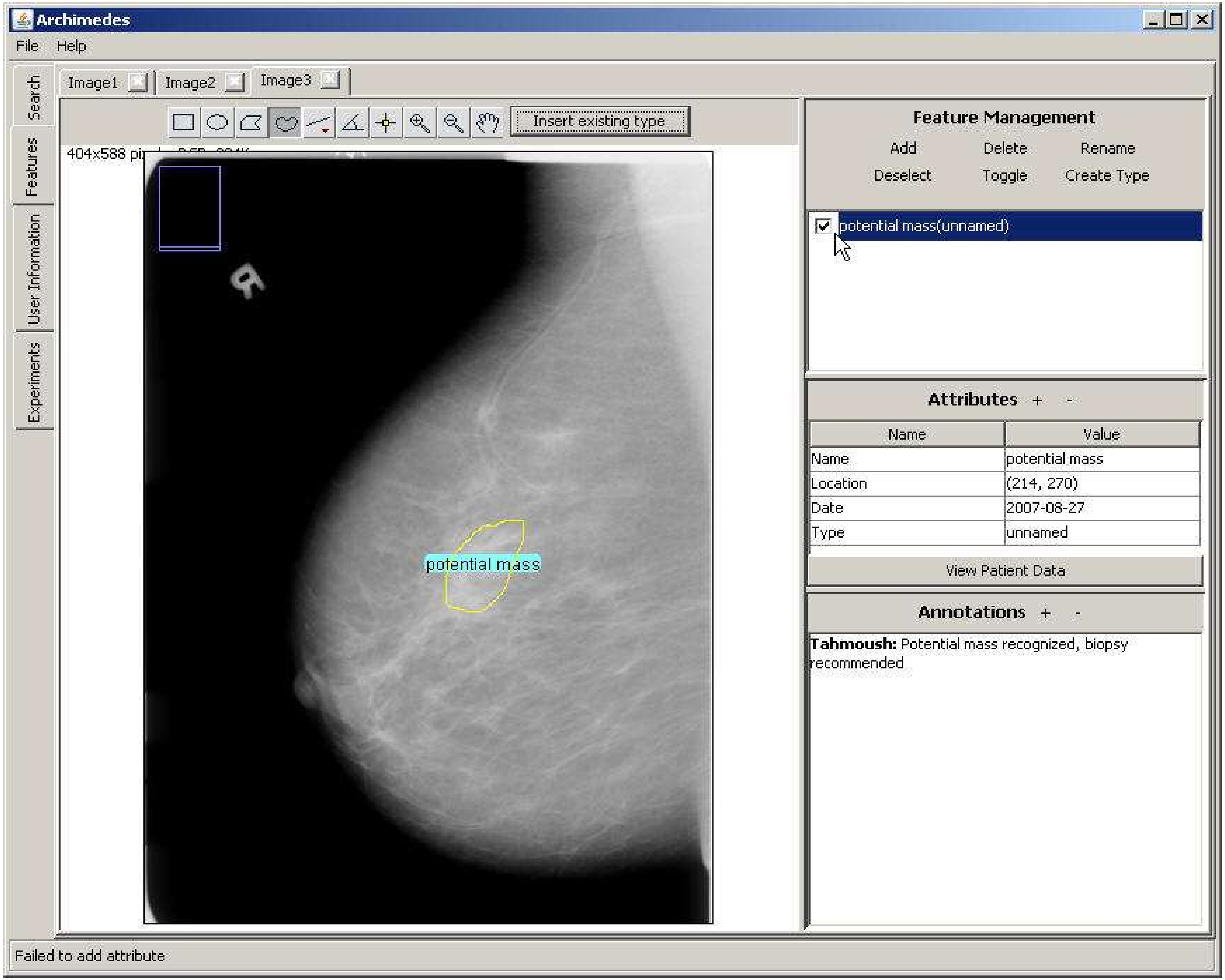


TPF *vs* FPF for 108 US radiologists in study [Beam et al], with our performance overlaid. Our technique compares favorably to radiologists.

PARTIAL CONCLUSIONS

1. Computer analysis can rival human diagnosis
    - Need large amounts of data to find enough of a particular cancer
    - NIH changing rules, forcing data sets to be shared
2. Humans can learn from a few examples – will always be ahead of computers
    - Imaging technology changes, need to verify that computer analysis still works well, so will always lag behind
    - But can get tired
    - Radiologists require intense training, and still large discrepancies in performance from one radiologist to another
    - HCI could improve radiologist performance
3. Computers can be more accurate
    - Have the processing power and techniques
    - Need more data, images, digitized radiologist annotations, notes
    - Created a database to collect this type of data

MEDICAL IMAGE DATABASE



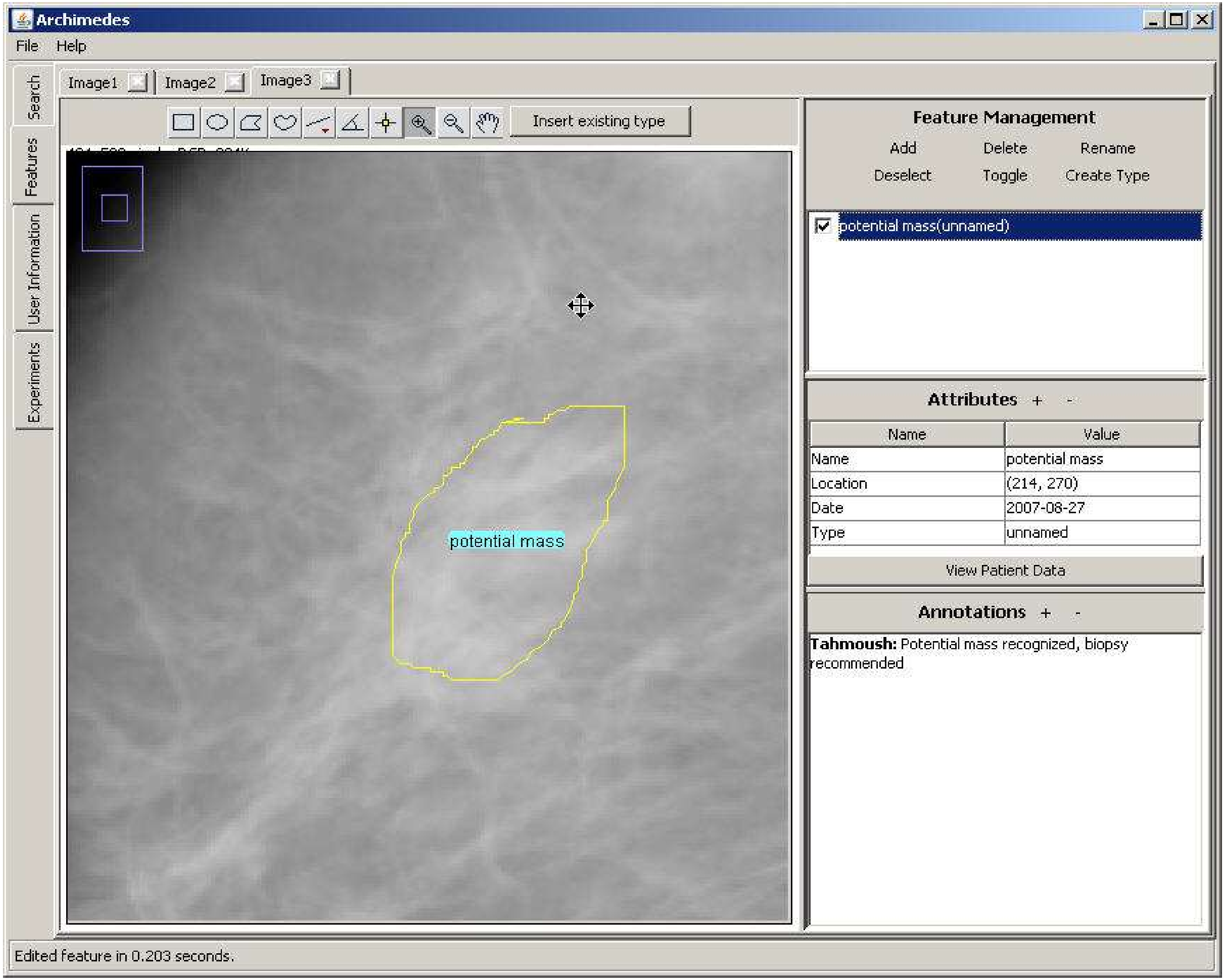
☑ potential mass(unnamed)

Name	Value
Name	potential mass
Location	(214, 270)
Date	2007-08-27
Type	unnamed

Tahmoush: Potential mass recognized, biopsy recommended

Failed to add attribute

MEDICAL IMAGE DATABASE



☑ potential mass(unnamed)

Name	Value
Name	potential mass
Location	(214, 270)
Date	2007-08-27
Type	unnamed

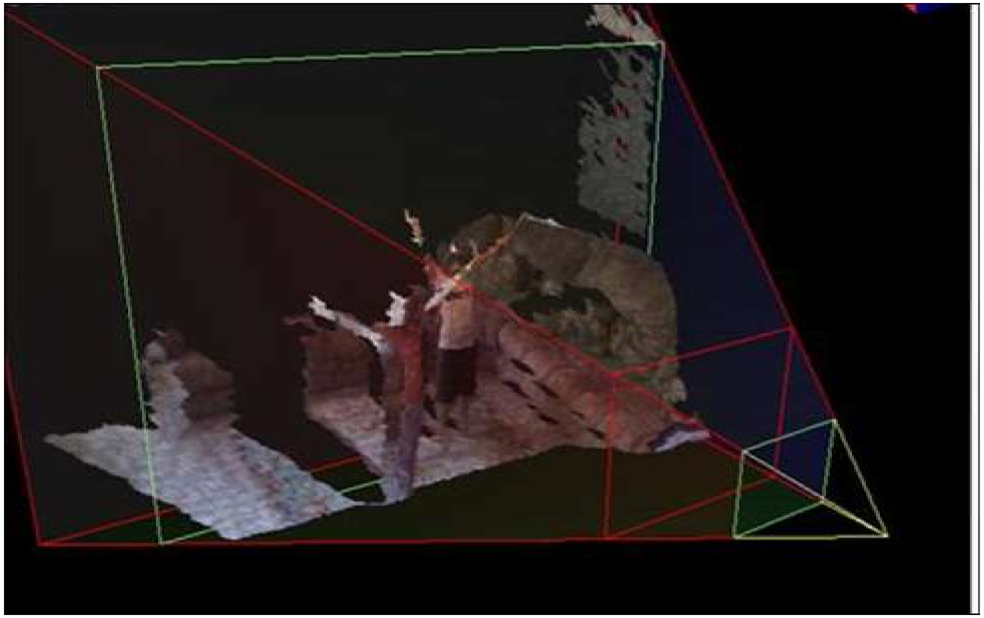
Tahmoush: Potential mass recognized, biopsy recommended

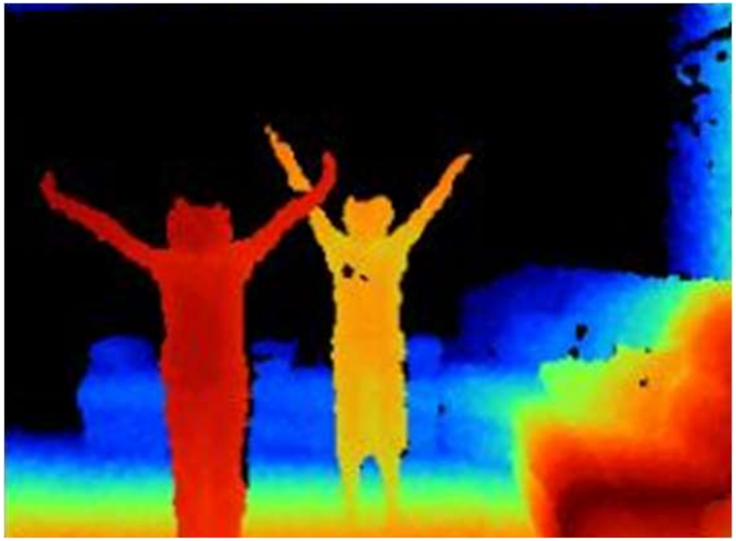
Edited feature in 0.203 seconds.

3D VIDEO ANALYSIS AND ACTION RECOGNITION

Tahmoush, David. "Applying action attribute class validation to improve human activity recognition." In **CVPR** workshop, pp. 15-21. 2015.

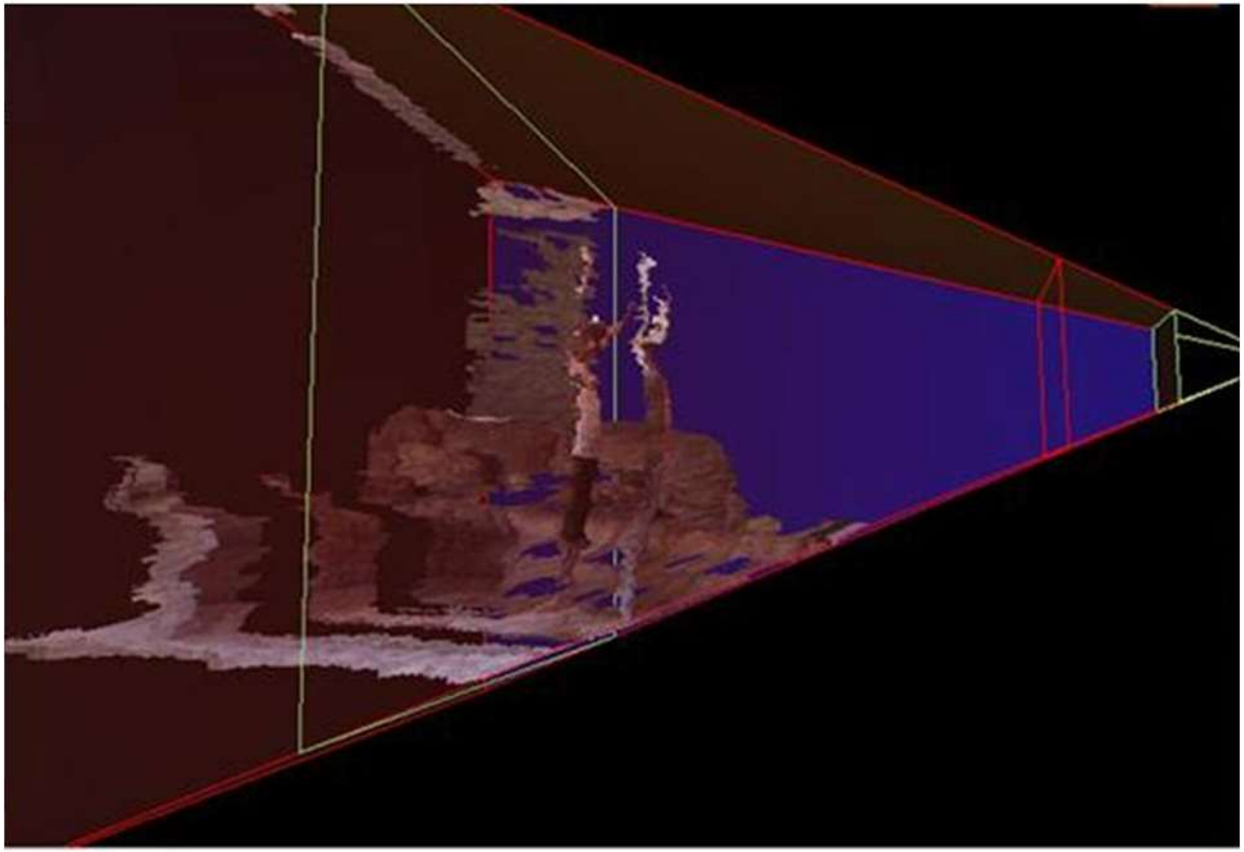
3-D Imagery





- 3D imaging technologies becoming commonplace
  - Ladar data
  - Microsoft Kinect
  - Time-of-flight cameras
- Costs driven down
- Video frame rates
- 3-D point cloud
- Distance often depicted in color
- Still occlusions

3-D Data, Single Source

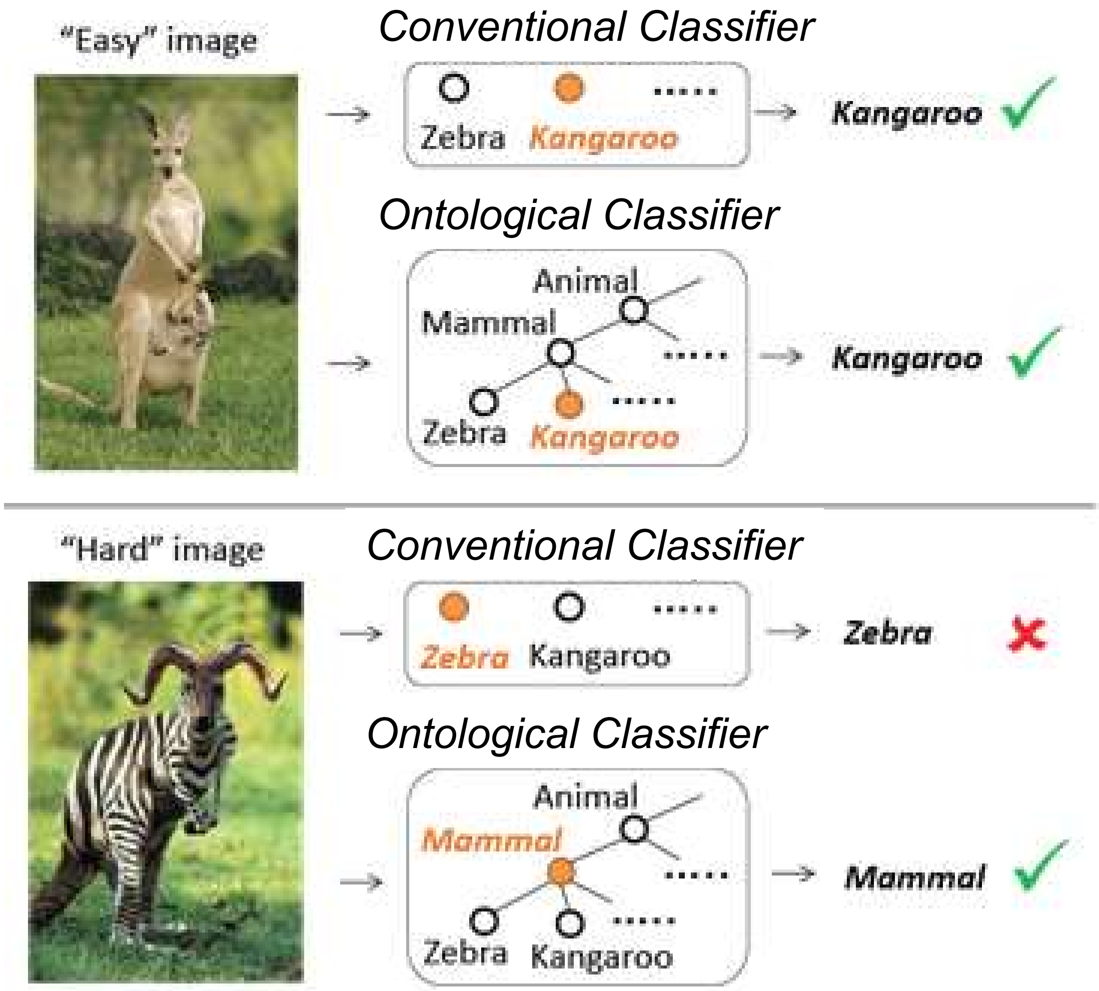


- Single source
- Occlusions
- Partial data
- Not great for perfect imaging
- Good enough for motion analysis

Activity Recognition

- Activity recognition has traditionally used low-level features
- **Can include human knowledge-driven associations**
    - Between and within actions and their attributes
    - Recognize lower-level attributes with their temporal relationships
        - **Could learn a much greater set of activities**
        - Reuse attributes to measure new activities
- In an ontology, actions can be decomposed into attributes with temporal and spatial relationships
- Throwing can be broken down
    - BodyPartsUsed = Hand
    - BodyPartArticulation-Arm = OneArmRaisedOverHead
- Build general attributes from video into an ontology
- Each activity or event classifier is composed of interacting attributes
    - Like sentences composed of interacting letters
    - Create a complete language
    - Easier to recognize letters and words than entire sentences

Ontological Versus Label Classification



- Label-based classification is a conventional approach
- New classifier for each instance
- Can have difficulty on harder images, rare cases
- Ontological classification is more conservative
    - Recognize mammal characteristics, but not particular label
- Example trained on ImageNet
- Uses expert information (kangaroo is a marsupial which is a type of mammal)
- Eventual goal

Approach

- Move to ontological classifier for understanding verbs in 3D video
- Building on human knowledge-driven associations
  - Between attributes and actions/activities
  - Recognizing the lower-level attributes
  - Recognizing their temporal relationships
  - Combine with video features
- For example, “bend down” which can be combined with “throw”
  - Combining in a temporal sequence gives “pick up and throw”
  - Implies stones or objects on the ground are the missiles
  - Infer hostility and anger with minimal threat
  - Throwing an object from a concealed area on the body creates a greater inference of threat with an unknown missile type

Approach

-On the textual analysis of throw in VerbNet, a typical frame is:

example	**"Steve tossed the ball from the corner to the garden."**
syntax	Agent V Theme {{+src}} Initial_Location {{+dest_dir}} Destination

where minimal physical information is encoded

-Initial location and destination included here, often inferred

-Not much informational content of verbs like throw

-Use related video to include some of the important physical articulation issues inherent in throw

-This may enable semantic understanding of video data

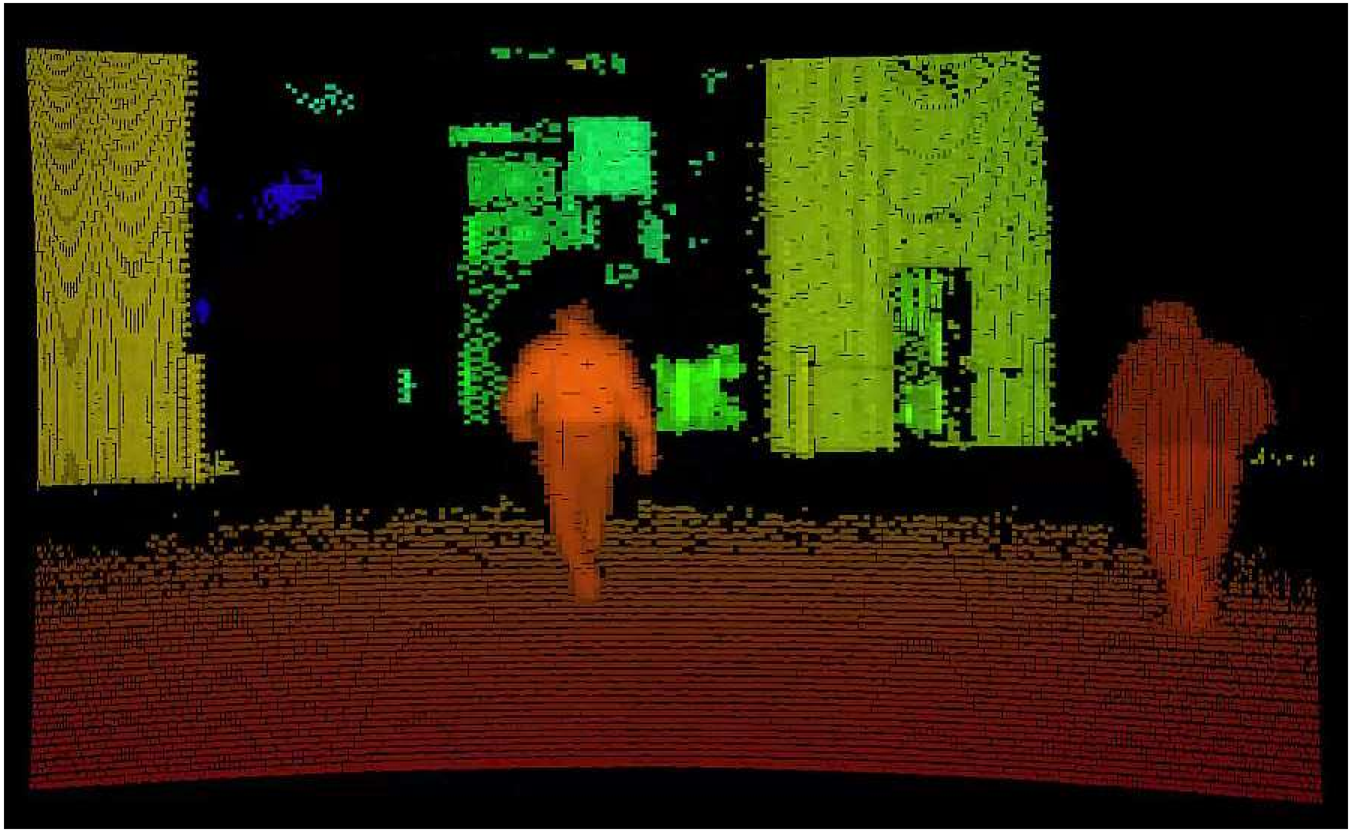
-Part of this was already done, compiling attributes for actions

Action Attributes

- Partial list of action attributes for "punch"
- Attributes include pose, motion of parts, and relations to the environment
- Reused from Action Classification Challenge ICCV 2013 [1]
- Extended
- Not all attributes currently used
- Convert to temporal measurements and link into ontology

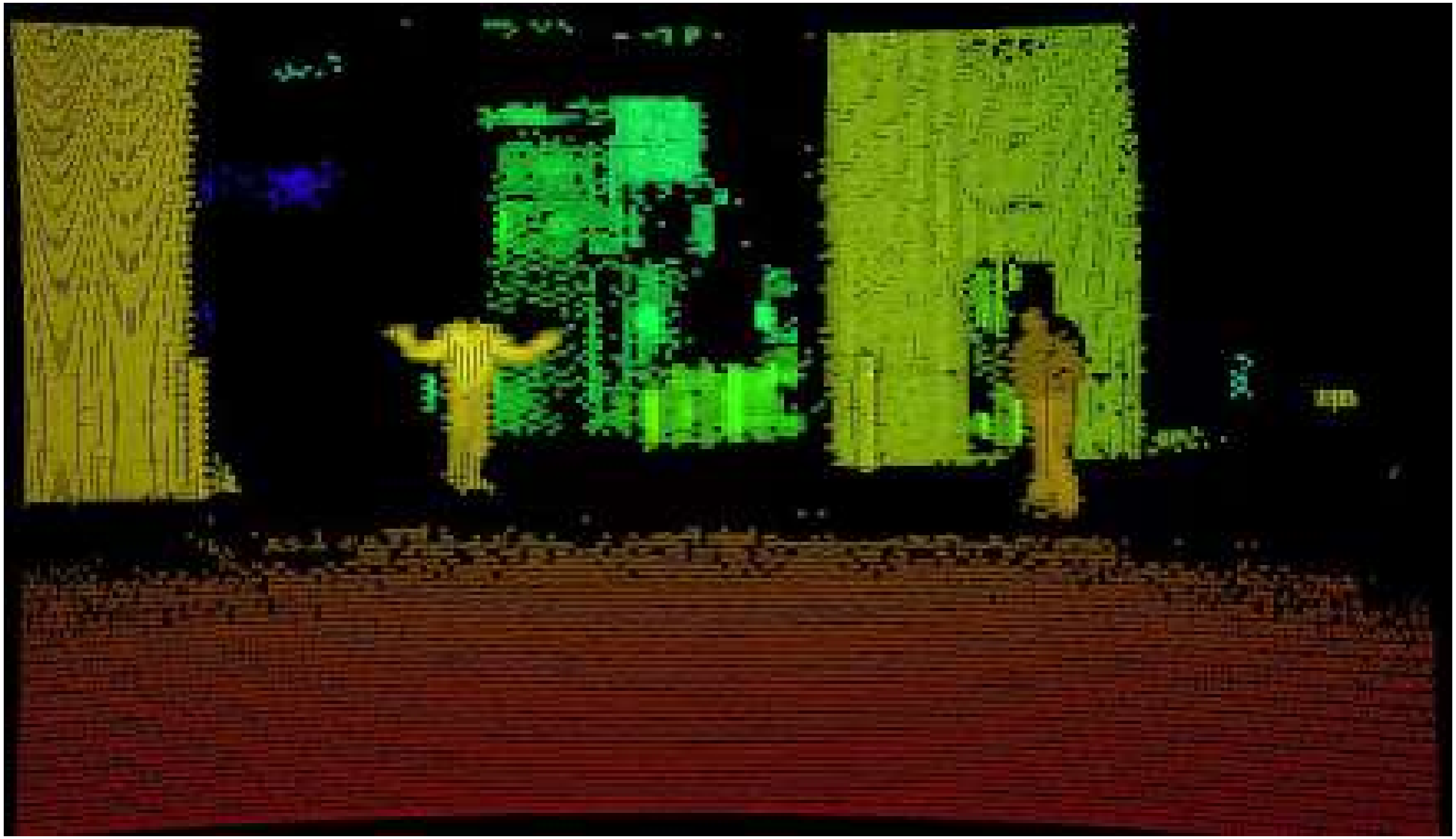
Action Attributes	**Punch**
Body Part Articulation-Arm = One_Arm_Motion	1
Body Part Articulation-Head = Straight_Position	1
Body Part Articulation-Torso = Down_Forward_Motion	0
Body Part Articulation-Torso = Straight_Up_Position	1
Body Part Articulation-Feet = Touching_Ground	1
Body Part Articulation-Feet = In_Air	0

Ladar View



- Note shadows cast on building
- Note ‘headless’ human doing two-handed wave

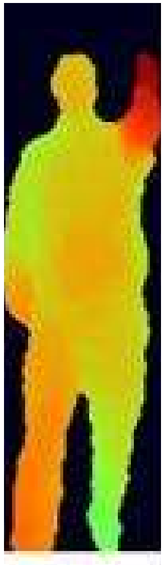
Ladar View



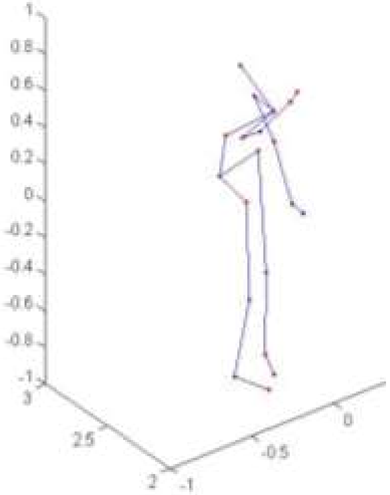
- Single frame of ladar data
- Note shadows cast on building
- Note ‘headless’ human doing two-handed wave

Skeletonization

subject



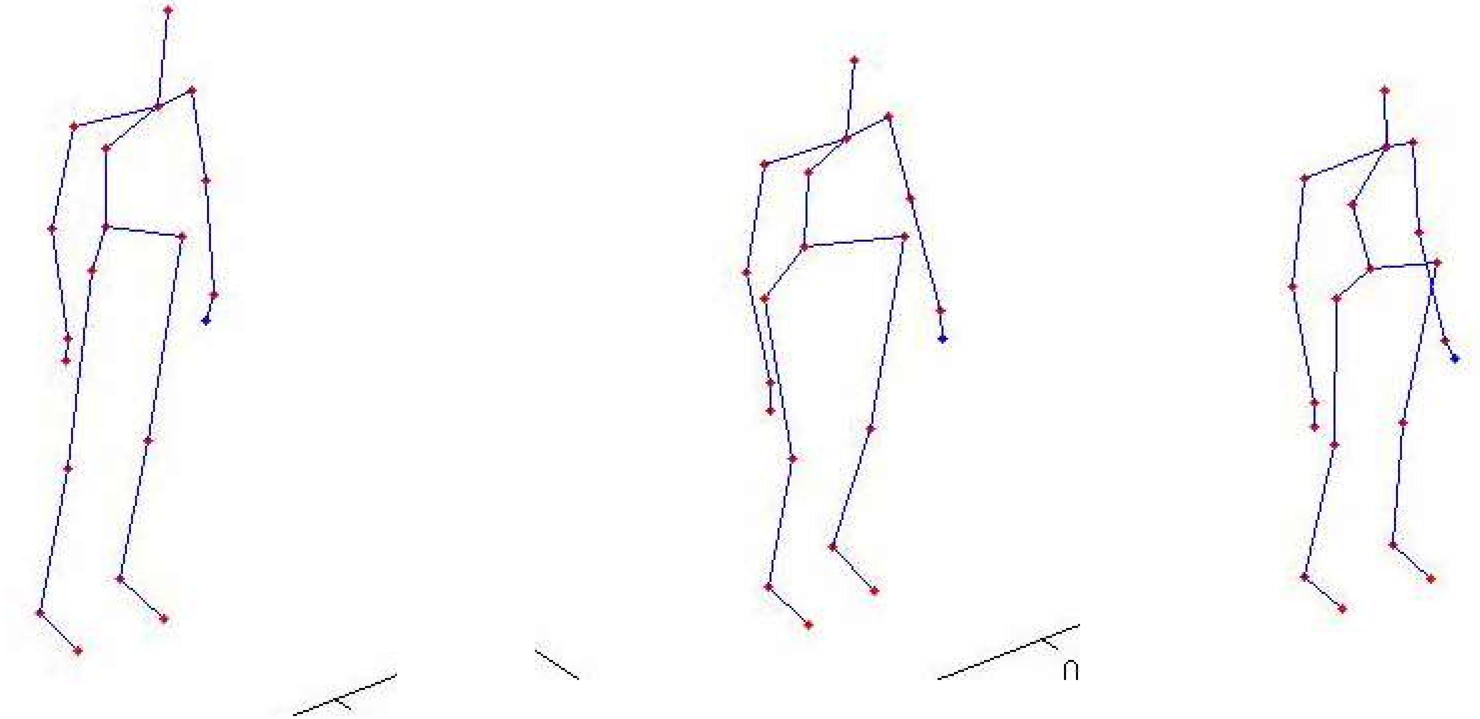
extracted point cloud



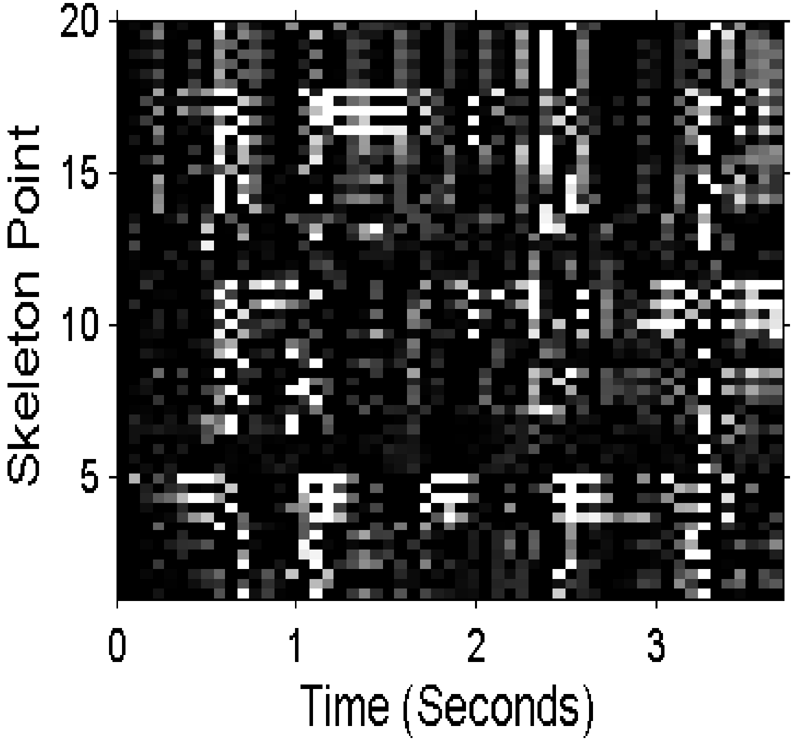
extracted skeleton

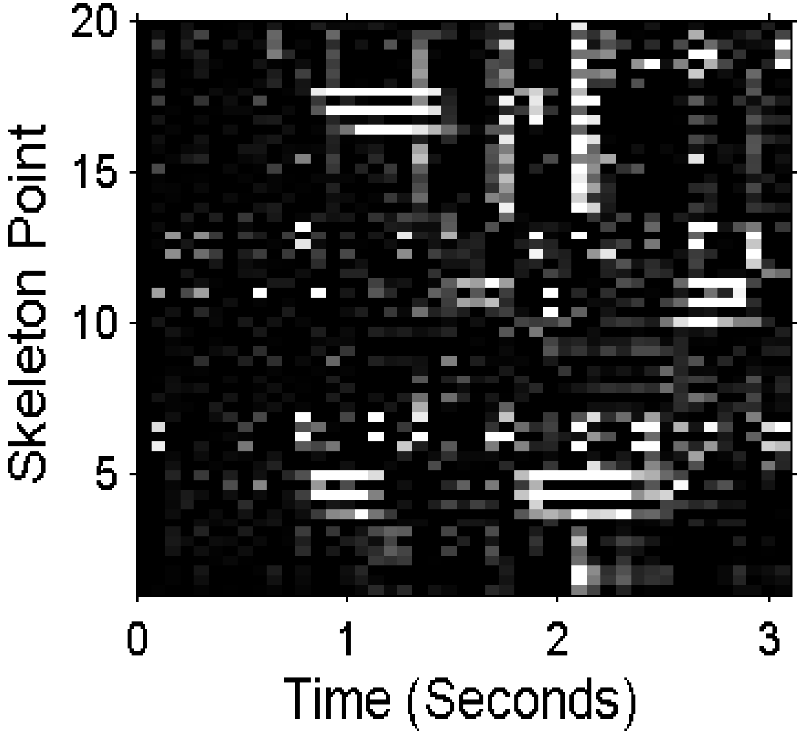
- Skeletonization significantly reduces data size with minimal information loss (sparse for fast analysis)
- Skeletonization introduces errors and can fail
- Action recognition should recognize skeletonization failure and be robust to errors

Two Hand Wave



Skeleton Motion



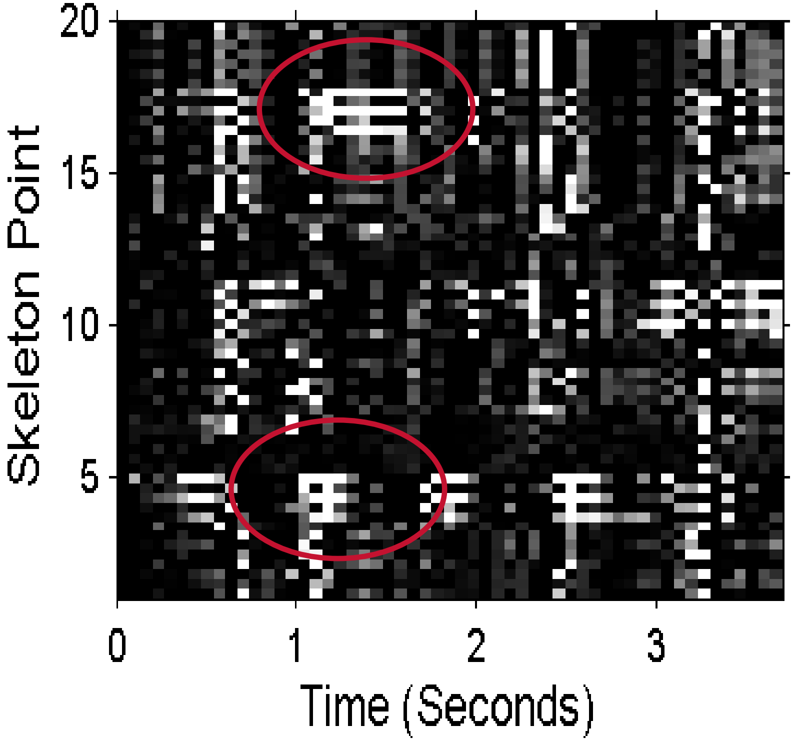


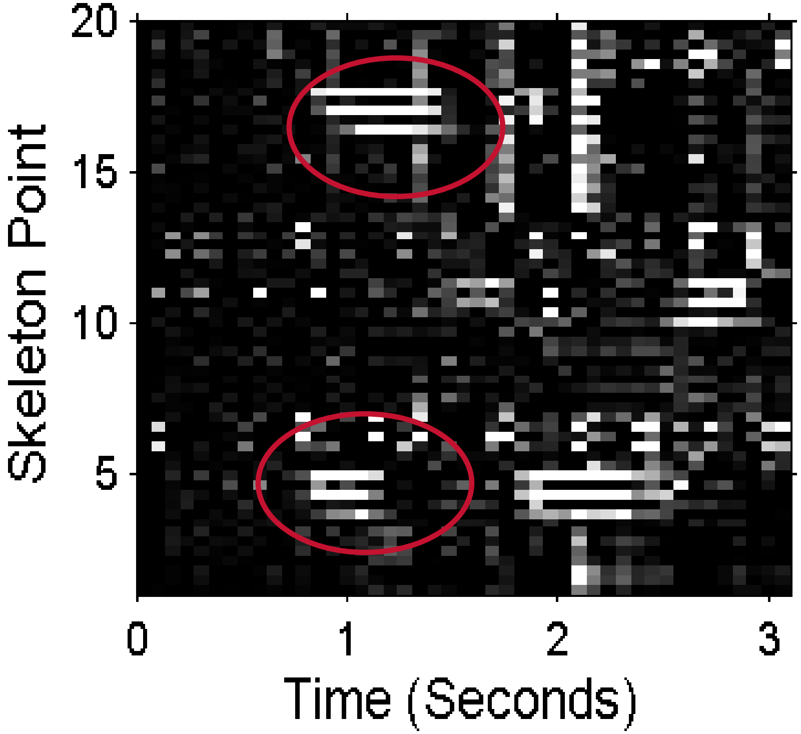
The joint-velocity magnitude heatmap for waving by 2 subjects

The log of the absolute value of the joint-velocity

Collection time for each signature is different

Skeleton Motion



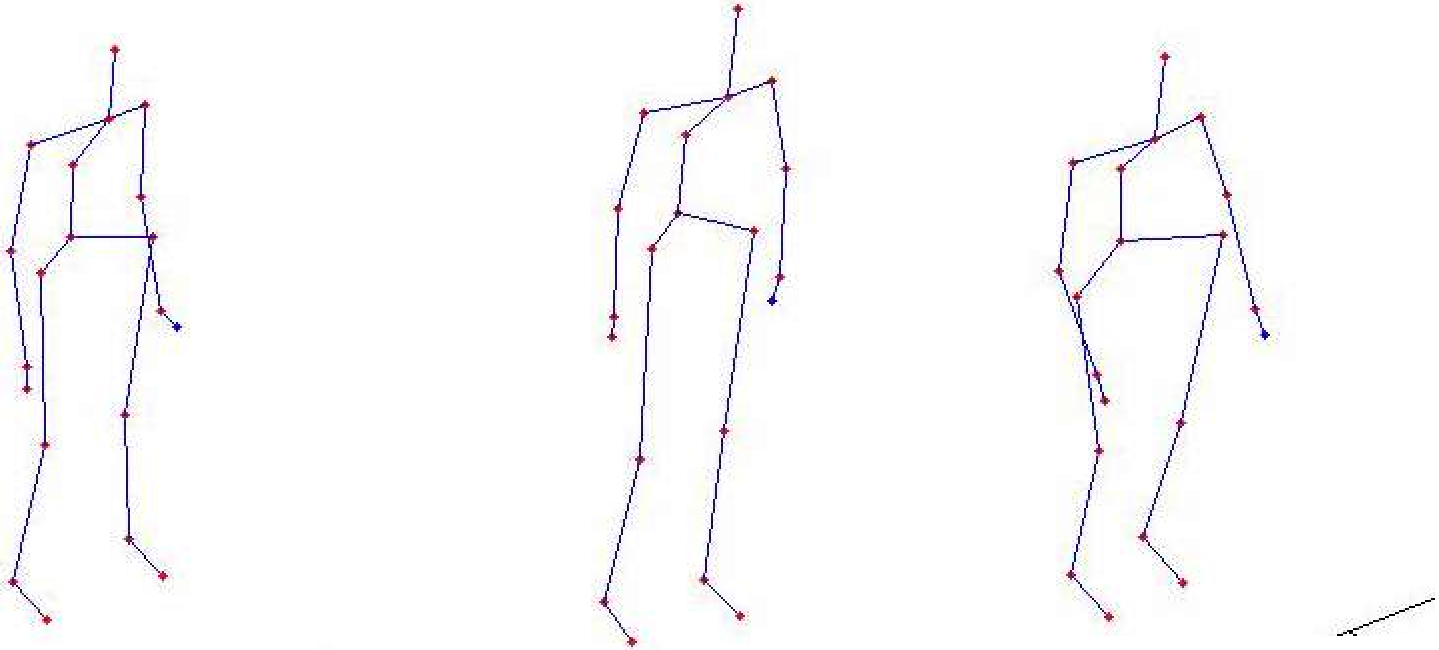


The joint-velocity magnitude heatmap for waving by 2 subjects

The log of the absolute value of the joint-velocity

Collection time for each signature is different

Waving



Attributes Utilized

- -Attributes that were useful in separating classes were used
  - -'Body Part Articulation-Arm = Two Arms Motion' which only is in ‘Two Hand Wave’ cases
  - -We extended the ontology to include 'Body Part Articulation-Arm = One Arm Raised Head Level' and 'Body Part Articulation-Arm = One Arm Extend Side' to help evaluate the ‘Side Boxing’ and ‘High Wave’ classes.
- -Fully implementing all attributes may do better
- -Attributes were easy to program on skeletal motions
  - -However, not clear how big motion should be (above neck or higher)
  - -Strictness not well explored
  - -Relative weight of measured attributes versus motion features

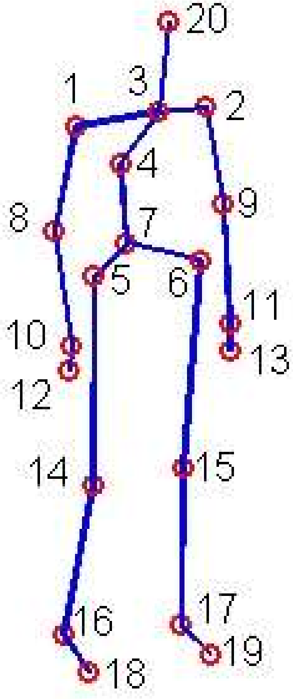
Attribute Programming

- Measuring attributes was simplified by utilizing the extracted skeletal joints
- For example, measuring the attribute 'Body Part Articulation-Arm = One Arm Raised Above Head'
    - a check on whether only one hand was above the location of the neck.
    - Then the calculation was

      IF $(P_{13}(3) > P_{12}(3))$ $(P_{13}(3) - P_{3}(3))$ * **H**$((P_{13}(3) - P_{3}(3))$ * **H**$((P_{3}(3) - P_{12}(3))$

      ELSE $(P_{12}(3) - P_{3}(3))$ * **H**$(P_{12}(3) - P_{3}(3))$ * **H**$((P_{3}(3) - P_{12}(3))$

      -Where **H** is the Heaviside function
- Note hand above neck was much more stable since head could be mis-identified (as arm!)
- Attribute is 0 when no hand is above neck (Heaviside) or if both hands are above the neck
- Converted attribute to video feature with a temporal extent



Skeleton Motion Similarity

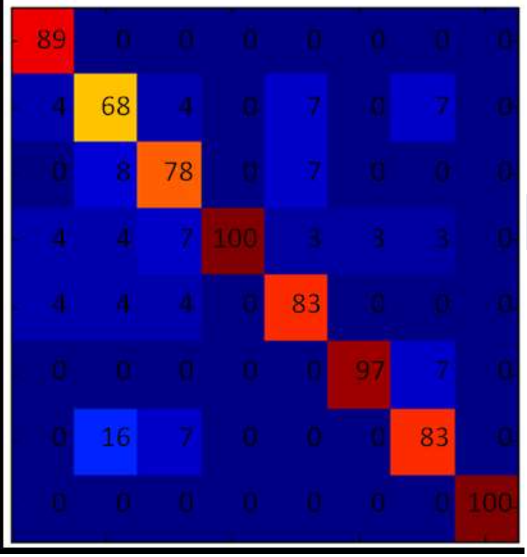
Normalized 3-dimensional cross-correlation

$$\frac{1}{n-1}\sum_{p,v,t}\frac{(f(p,v,t)-\overline{f})(g(p,v,t)-\overline{g})}{\sigma_f\sigma_g}$$

- n is number of samples, f is the test data, g is the case from the database. Data are pose points on the skeleton p, velocity v of that point, and time t
- Normalized cross-correlation is used as a distance function
- Used leave-one-actor-out (LOO) method for comparisons
- Nearest neighbor classification function
- Second neighbor can improve confidence in classification
    - Provides uncertainty if not support 1st nearest neighbor

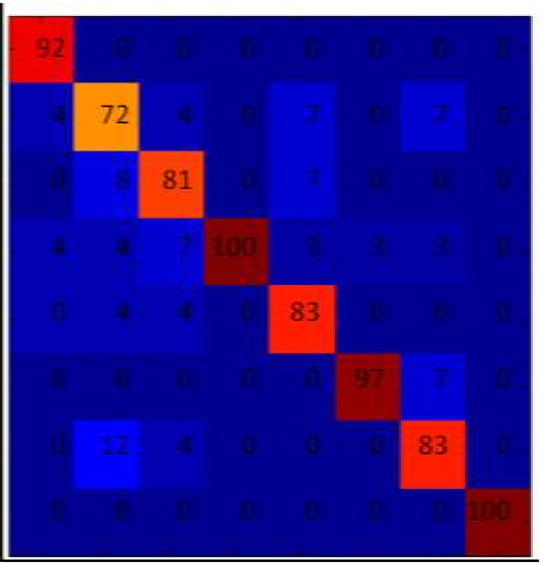
Classification Between Actions

[redacted]



89	0	0	0	0	0	0	0
4	68	4	0	7	0	7	0
0	8	78	0	7	0	0	0
4	4	7	100	3	3	3	0
4	4	4	0	83	0	0	0
0	0	0	0	0	97	7	0
0	16	7	0	0	0	83	0
0	0	0	0	0	0	0	100

[redacted]



92	0	0	0	0	0	0	0
4	72	4	0	7	0	7	0
0	8	81	0	7	0	0	0
4	4	7	100	3	3	3	0
0	4	4	0	83	0	0	0
0	0	0	0	0	97	7	0
0	12	4	0	0	0	83	0
0	0	0	0	0	0	0	100

Initial results including binary attributes with motion features showed small improvement

Reduced bad cases in training data, not test data

Results

- -Found several attributes that identified abnormal cases
- -Not too difficult to determine on skeleton data
- -Optimistic about improvement
- -Attributes did help 1.4%
- -50% of the improvable classes did improve
- -Damaged cases still in test set
- -Table is not all-inclusive
- -Surprisingly good performance

Method	Accuracy
Li et al [4]	71.9%
Lu et al [11]	85.5%
Yang et al [15]	84.1%
Chen et al [16]	83.3%
Initial Method	87.2%
Initial Method with Attributes	88.6%

Online Classifier

- An online classifier is harder to use but will take in new examples from its experience to add to its knowledge.
- In effect, an online classifier is constantly trying to learn from its environment.
    - Collects new data and adds to classifier
    - Retraining time is an issue, hence nearest neighbor
    - Motivation: people do things differently
    - Should improve performance if capable of online learning
- We estimate the improvement using an online approach by replacing the nearest neighbor classifier with an online nearest neighbor approach that takes in new data.
- Non-online case: treat one subject as “new” and classify using only other subjects’ data
- Online case: include learned data from new subject in classification

Results

- -Found initial classification approach that worked well
- -Estimated value of adding online capabilities
- -5 of 6 of the improvable classes did improve
- -No classes got worse
- -Table is not all-inclusive

Method	Accuracy
Li et al	71.9%
Xia et al	85.5%
Yang et al	84.1%
Chen et al	83.3%
Initial Method	88.7%
Online Method	93.5%

Partial Conclusions

- This work utilizes a single point-of-view 3D imaging system to approximate ladar captured data
- The classification of human activities was shown to be feasible with a motion-based classification approach on 3D ladar data
- The ladar approach performed at 87% on a set of activities, which was at or above the sate-of-the-art
- Including measurable text attributes did improve classification to 88%
- Utilizing an online approach improved the performance to 93% and effectively cut the number of misclassifications in half
- The online capability was added after the initial experiments by a careful choice of a classifier that could be replaced with an online classifier, implying that fielded robots could be upgraded as capabilities are proven out
- The incorporation of online capabilities is shown to be an important improvement to human activity recognition

Extra Slides