A Statistical Model for Improved Surface Detection

Samuel Smith

Birmingham City University

samuel.smith@mail.bcu.ac.uk

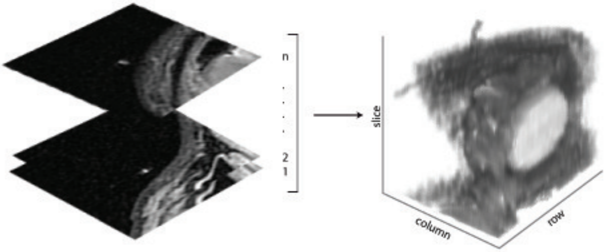
February 20, 2015

Overview

Rationale

- Three dimensional image data is becoming the common modality for many non-destructive testing, image analysis, visualisation and biomedical imaging systems. Typically:
  - Computed Tomography
  - Magnetic Resonance Imaging / functional MRI
- These high level processes require low level image processing techniques.
- Improvements offered in low level techniques, should offer improvements to higher level applications.

Three Dimensional Data



Three dimensional image data is stored as 3 dimensional array, consisting of a stack of two dimensional slices.

What is a Surface?

- A surface is an interface which exists in 2 or 3 dimensional data, it describes boundary, or plane, which separates two or more different regions.
- This boundary could be between two or more areas of different:
  - Voxel Intensities
  - Colour
  - Texture

What is a Surface?

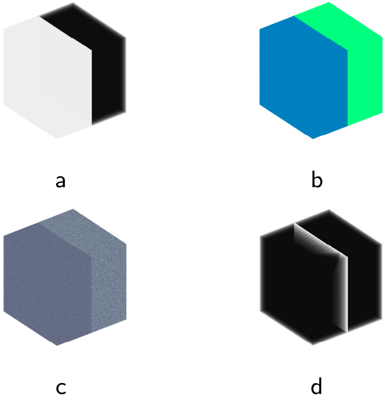
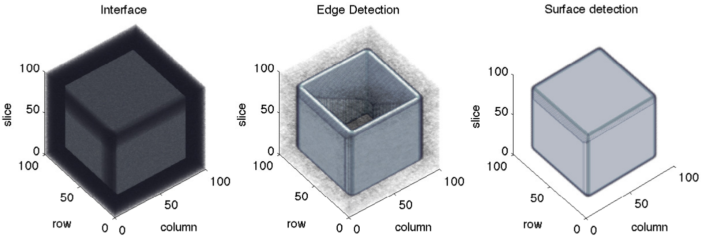


Figure: a) Intensity Interface. b) Colour Interface. c) Texture Interface. d) Interface Location

3D Surfaces

- Optimal plane of application is not known a priori
- Surfaces which lie in the plane of edge detection are not detected. Instead an outer edge of the surface is identified



Criteria for edge/surface detection

When designing a surface detection filter, there are certain criteria that needs to be met. Canny defined the following criteria for edges, but the following holds true for surfaces.

- Criterion 1: Good detection.
  Minimising the number of number of missed surface points, as well as minimising the number of spurious responses.
- Criterion 2: Good localisation.
  The points marked as surface points by the operator should be as close as possible to the center of the true surface.
- Criterion 3: Single Response.
  There should be only one response to a single surface. Duplicate responses for surfaces should be eliminated.

Methods for segmenting surface information from 3D image datasets.

- Gradient Based Operators.
- Steerable filters.
- Statistical Surface Detection.

Gradient based methods

Gradient methods can be broadly described as methods which use convolution to approximate the first and second derivative of images with regard to their intensity values. Where sharp changes in pixel intensities exist, these positions usually correspond to the location of a boundary.

- Sobel.
  An efficient convolution based method for determining the first derivative of an image.
- Laplacian.
  Takes the second derivative of an image.
- Canny.
  Applies Gaussian smoothing, non maximum suppression and hysteresis thresholding. Giving improved response in noisy images, and produces a single response for a single interface. For a long time considered the optimal method.

- State of the art method for surface detection.
- Developed by Aguet et al, extending upon the work of Jacob and Unser.
- The design of the filters are based on "Canny-like criteria", giving an optimal response for curves, edges and surfaces, while suppressing noise and unwanted image texture.
- This method utilises the concept of oriented filter masks to determine the maximum response in surface detection, and adaptive filtering.

2D Statistical Edge Detection

- Pixel intensity values are extracted using a simple 2D neighbourhood mask.
- The mask is divided into two sample regions, to which a statistical test is applied measuring dissimilarity.

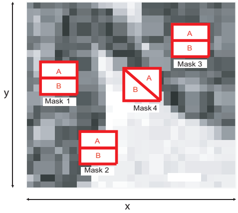
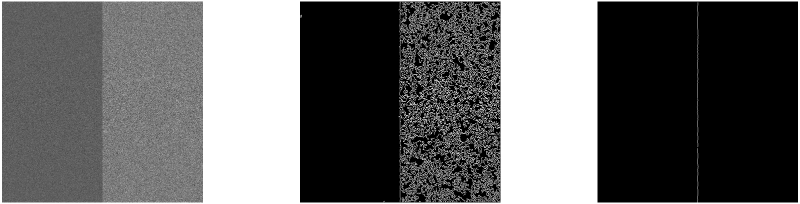


Figure: Mask 1,3, not located on a boundary (low output). Mask 2 located on boundary, incorrect orientation (low output). Mask 4 located on boundary, correct orientation (high output).

- Through a procedure of shifting the orientation of the mask, the position correlating to maximum dissimilarity then provides the output magnitude for that pixel.
- This process is repeated for each pixel in the image, evaluating the location, strength, and orientation of edges present in an image.

The problem with Gradient and Rotational methods is that their performance on texture based interfaces is less than optimal.
In 2D data alone, it has been shown that 2D statistical based methods outperform traditional gradient based detection where excessive texture is evident in the image, or when the intensity profile of an edge is weaker.



a b c

Figure: a) Noisy/texture interface. b) Output from 2D Canny method. c) Output from 2D Statistical edge detection

Statistical Surface Detection Model

Extending upon the two dimensional statistical edge detection method, a model for detecting surfaces can be produced.
The model should be able to:

- Detect surfaces in three dimensional image volumes.
- Locate interfaces that lie in the plane of the surface operator.
- Determine interfaces at multiple intensity scales.
- Be robust to noise and image artefacts.

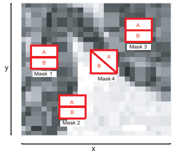
For this work, 2 methods of surface detection were developed.

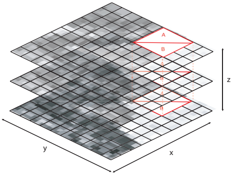
- Maximum Response method.
- Vector Magnitude method

Maximum Response Method

This method is a direct extension of 2D statistical edge detection into 3D.

- Here, voxel intensity values are extracted using a simple 3D neighbourhood mask applied across several 2D image slices.
- The mask is divided into two sample regions, to which a statistical test is applied measuring dissimilarity.





- Through a procedure of shifting the orientation of the mask, the position correlating to maximum dissimilarity then provides the output magnitude for that voxel.
- This process is repeated for each voxel in the image, evaluating the location, strength, and orientation of surfaces present in a 3D image.

- This method selects orientations based on a 26-connectivity neighbourhood.
- If we define a single orientation as the direction from the central neighbourhood voxel to a neighbouring voxel, due to the rotational symmetry of the mask, half of the positions would be redundant as the same output would be provided.
- Therefore this method computes 13 orientations.

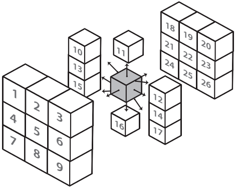


Figure: 26 connectivity Neighbourhood mask

- Alternatively, a statistical test can be applied in 3 orientations, across the x,y and z Cartesian planes.

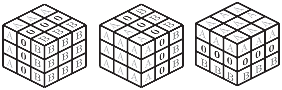


Figure: Vector Magnitude Method

Vector Magnitude

The outputs of the statistical tests in each orientation provide a vector of dissimilarity.

$$V = T_x \hat{i} + T_y \hat{j} + T_z \hat{k} \tag{1}$$

Where T_x, T_y, T_z are the magnitude components of the vector V, provided by the output of a statistical test measure of dissimilarity between neighbourhood regions The overall magnitude can then be computed by finding the Euclidean norm, defined as:

$$\|x\| = \sqrt{{T_x}^2 + {T_y}^2 + {T_z}^2} \tag{2}$$

By computing the the Ecludian norm over 3 orientations, instead of determining the maximum response of 13 orientations, it becomes significantly less computationally expensive.

Neighbourhood Mask

- The mask neighbourhood is of equal scale in 3 dimensions
- To centralise the neighbourhood mask around a voxel, the mask length in each direction must be an odd value.
- When choosing a suitable size for a neighbourhood mask, there is a trade off between reliability and accuracy

Mask Scale

As mask size increases

- Improved resolving power.
- Greater suppression of noise.
- Less susceptible to image artefacts.
- Image details smaller than the neighbourhood mask are not captured.
- More uncertainty in the location of the interface.
- More computationally expensive

Ideally we want to keep the mask as small as possible to locate finer details, while large enough to resolve texture based boundaries

A key aspect which determines the effective of statistical surface detection are the statistical tests which are applied. Different regions of an image contain different statistical properties such as mean, standard deviation, skewness and kurtosis, as well as feature descriptors such as those identified by Haralick. These features combine together to define the intensity and texture profile of a region.

Student's T test

In probability theory, the first and second moments of a probability density function are mean and variance, therefore these two properites play a fundamental role in describing image region characteristics. The Student T test is a popular mean based statistical test which tests the hypothesis that two distributions will have a similar mean value.

T-test

$$T = |\bar{x}_A - \bar{x}_B| \sqrt{\frac{N-1}{\sigma_A^2 + \sigma_B^2}} \quad (3)$$

Where $\bar{x}_A$, σ_A^2 and $\bar{x}_B$, σ_B^2 are respectively the mean and variance of mask regions A and B, and N is the number of pixels in each single region.

The Difference of Boxes test is the second mean based test used in this work. It's simply the absolute difference of the means of each region. This method provides similar outputs to gradient based methods.

Difference of Boxes -test

$$D = |\bar{x}_A - \bar{x}_B| \tag{4}$$

Where $\bar{x}_A$, and $\bar{x}_B$ are the mean of mask regions A and B,.

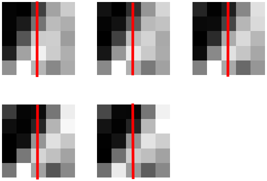
Theχ^2 test is a rank based test which checks for the independence of the two sorted datasets. It is a comparison measure that takes the relative difference in points at the same rank position for two binned data sets.

χ^2 Test

$$\chi^2 = \sum_i \frac{R_i - S_i}{R_i + S_i} \tag{5}$$

Here R_i is the number of values in *bin i* of region A, and S_i is the number of values in *bin i* of region B.

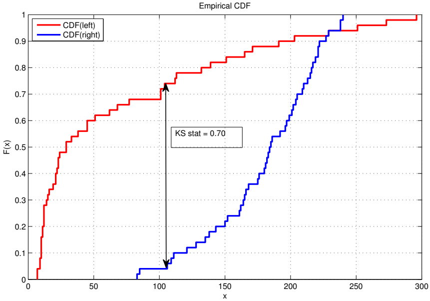
Two sample Kolmogorov-Smirnov test

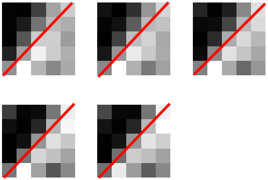


KS-Test

$$D_{n,n'} = \sup |F_{1,n}(x) - F_{2,n'}(y)|$$

Where $F_{1,n}$ and $F_{2,n}$ are the empirical distribution functions of the two samples

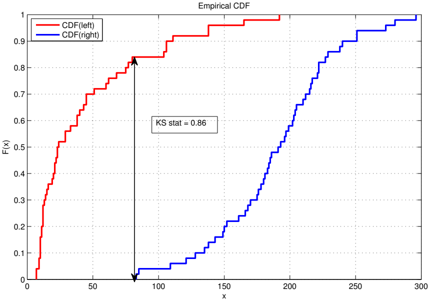




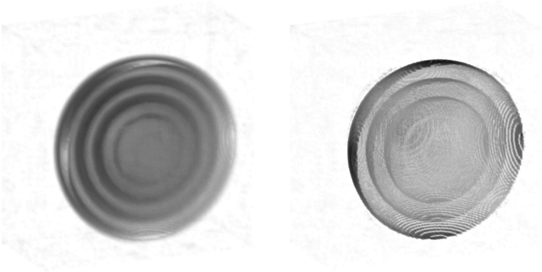
KS-Test

$$D_{n,n'} = \sup |F_{1,n}(x) - F_{2,n'}(y)|$$

Where $F_{1,n}$ and $F_{2,n}$ are the empirical distribution functions of the two samples



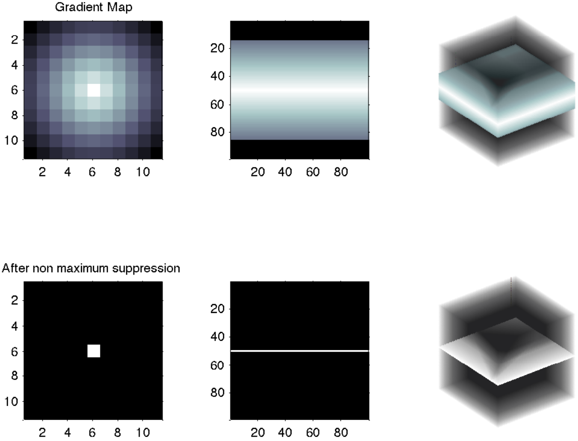
Non Maximum Suppression



a b

Figure: a) Surface map. b) Non-maximally suppressed Surface Map

Non Maximum Suppression



- Depending on the application, not all surfaces may be required.
- Sometimes erroneous surfaces caused by excessive noise may be present.
- A simple process often used for removing these edge or surface points is Hysteresis thresholding.
- Hysteresis results in a logical array or binary map where edge or surface points are given the value 1, while non-edge and non-surface points are given the value 0.

Hysteresis Thresholding

- We apply Hysteresis after non maximum suppression.
- It requires two threshold values.
- Any surface point with an intensity greater than the upper threshold is declared an edge or surface element.
- The 26 connected voxel neighbourhood surrounding a surface element is checked for intensities above the lower threshold
  - Those above the lower threshold are declared surface elements and their neighbourhoods are then checked too.
  - Those below the lower threshold are removed

Hysteresis Thresholding

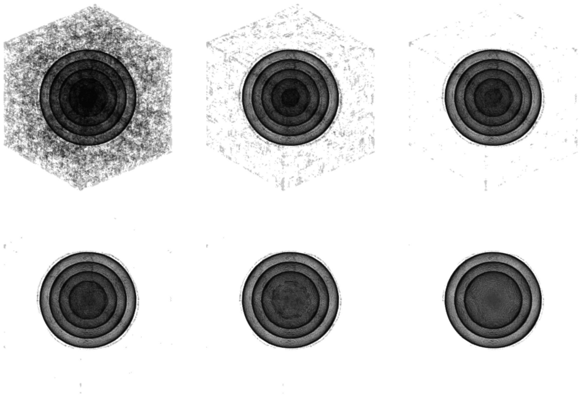


Figure: Hysteresis Thresholding. Upper thresholds a) 50. b) 75. c) 100. d) 125. e) 150. f) 200. Lower threshold set to 40% of upper

Performance Measures

- Qualitative
- Quantitative
- Hybrid

Quantitative Performance Measures

- Pratt Figure of Merit (PFOM)
- Receiver operating characteristic curves (ROC)
- Precision Recall
- Pixel Correspondence Metric (PCM)

The 3D *PFOM*

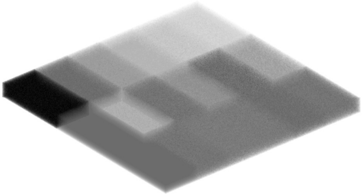
$$PFOM_{\eta} = \frac{1}{\max(N_I, N_B)} \sum_{i=1}^{N_B} \frac{1}{1 + \alpha \times d_i^2} \tag{6}$$

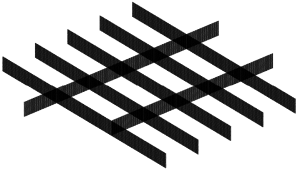
Where η is the maximum ideal tolerable error allowed between voxels, and was set to allow for a displacement of 1 voxel either side of the ideal. N_I and N_B are the surface points in the image volume and ground truth volume, respectively, d_i is the Euclidean distance between a detected surface voxel and the nearest voxel of the ideal, and α is a calibration constant set at $\alpha = 1/9$, a value established by Pratt (1979)

Synthetic Data Creation

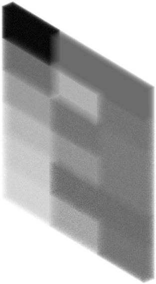
- Assessing the performance of surface detection methods is non-trivial if reliant on real world image data. This is due to the fact defining a ground truth data for real imagery is near impossible, and completely dependent on the scale which which one defines the existence of a boundary.
- By creating synthetic images, we can accurately determine the ground truth solutions.
- There are numerous considerations to be made when creating synthetic image volumes
  - Topology of interface
  - Type of interface
  - Number of interfaces
  - Bias of interfaces (major / minor lines)
  - Scale
  - Corners
  - Data acquisition

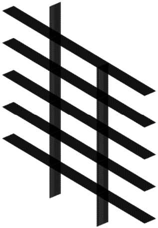
Synthetically created data - Multiple scale



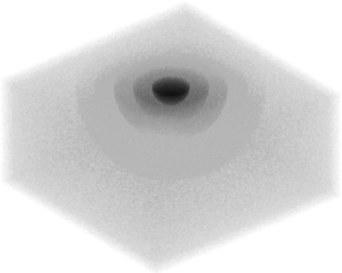


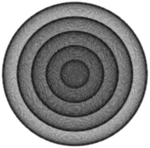
Synthetically Created Data - Multiple scale rotated



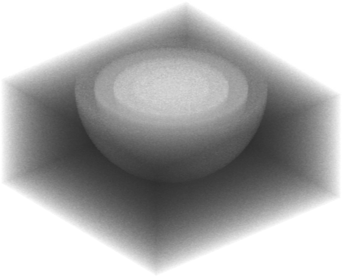


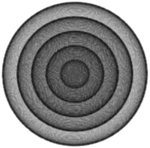
Synthetically Created Data - Multiple sphere



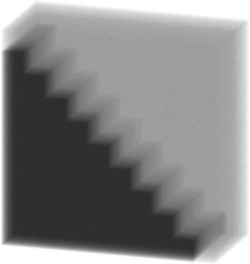


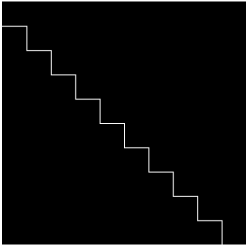
Synthetically Created Data - Reverse Multiple sphere

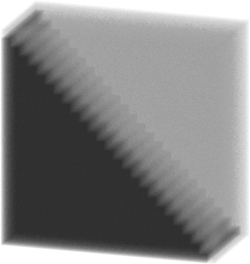


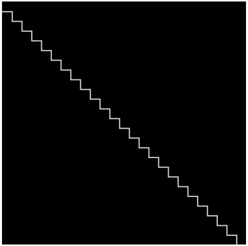


Synthetically Created Data - Staircase

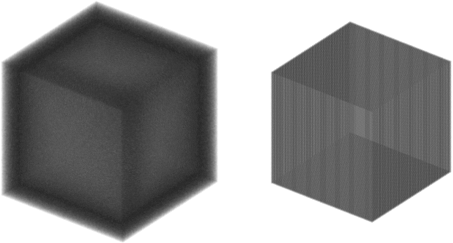




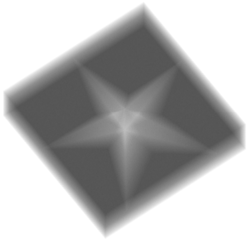


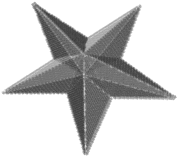


Synthetically Created Data - Cube



Synthetically Created Data - Star





In this study we used a Monte Carlo Methodology of testing.

- Three examples of each image
- Pseudo-random values containing the same statistical properties for each image
- Mean and standard deviation of results of each image type determines the score.

PFOM Results

Sphere $\bar{x}$ Scores							
σ	2D Canny	3D Canny	3D Steerable	Mask Size	2D stat	3D stat $\|\|x\|\|$	3D stat (max)
1	86.28	92.28	89.60	5	78.46	89.88	91.24
2	87.21	91.93	96.92	7	84.83	96.90	97.73
3	87.37	92.31	91.20	9	89.12	98.79	99.22
4	87.44	92.43	81.34	11	92.26	98.46	99.18
Multiple interface							
σ	2D Canny	3D Canny	3D Steerable	Mask Size	2D stat	3D stat $\|\|x\|\|$	3D stat (max)
1	64.86	70.13	65.07	5	72.01	80.64	80.89
2	67.62	72.55	82.98	7	84.16	90.23	92.33
3	67.46	72.60	81.71	9	93.87	95.99	96.17
4	67.35	72.60	n/a	11	94.16	96.76	95.59
Multiple interface rotated							
σ	2D Canny	3D Canny	3D Steerable	Mask Size	2D stat	3D stat $\|\|x\|\|$	3D stat (max)
1	42.72	70.48	65.22	5	42.05	80.82	82.57
2	43.68	73.36	83.71	7	44.68	92.47	92.50
3	43.81	73.43	82.92	9	50.02	97.02	96.46
4	43.82	73.39	n/a	11	52.16	96.10	95.90

Table: *PFOM* results on each test image compared to the ideal. Highlights: Poor(Red), Adequate(Yellow), Good(Green).

Visual Results - Sphere

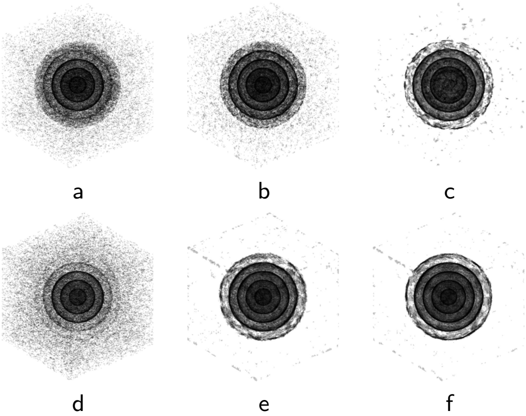


Figure: Results from a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical MR. f) 3D Statistical $\|x\|$.

Visual Results -Sphere Reversed

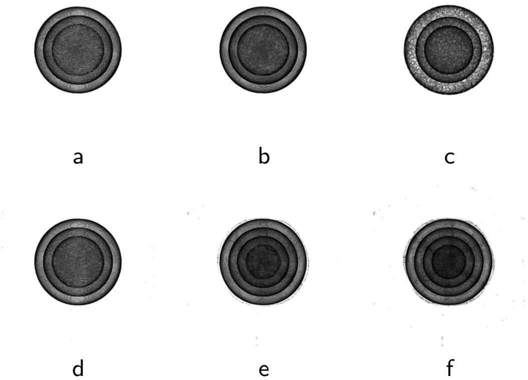


Figure: Results from a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical MR. f) 3D Statistical $\|x\|$.

Visual Results - Multiple Scale

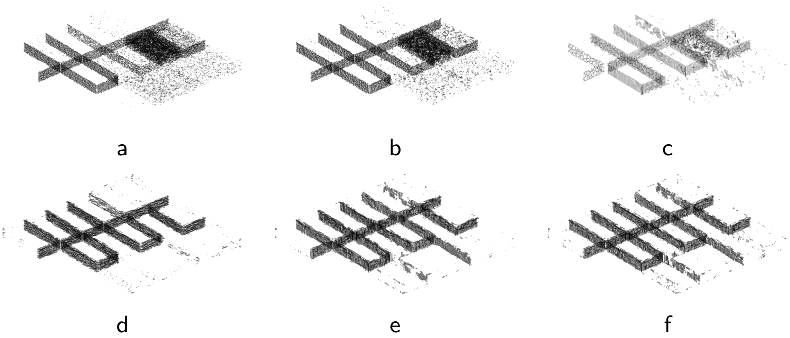


Figure: Results from a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical MR. f) 3D Statistical $\|x\|$.

Visual Results - Multiple Scale Rotated

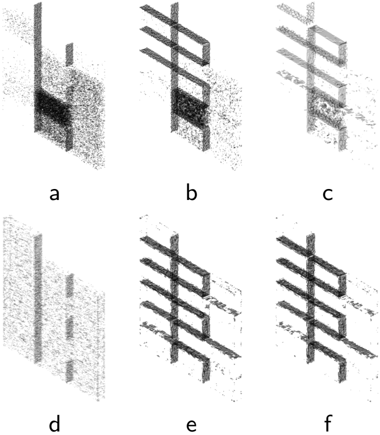
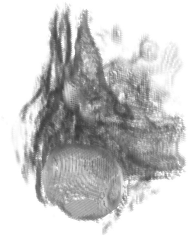
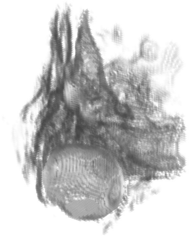


Figure: Results from a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical MR. f) 3D Statistical $\|x\|$.

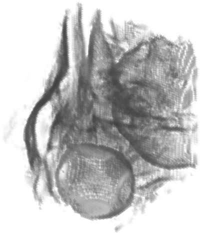
Real Image Results



Steerable



Canny



Statistical

Conclusions

- Outperforms 3D Canny and Steerable filters, improved response to texture and noise.
- Outperforms all 2D edge detection methods.
- When possible, 3D surface detection should always be used instead of 2D, and where texture defines image boundaries, Statistical methods should be employed.

Future Work

- Synthetic data creation.
- Statistical tests
- Mask shape
- Real World application testing. (Active Contours/surfaces, snakes GVFCs, etc)

End

pub ?