A Statistical Model for Improved Surface Detection

Samuel Smith

Birmingham City University

samuel.smith@mail.bcu.ac.uk

January 14, 2016

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.

Software

- MIMICS

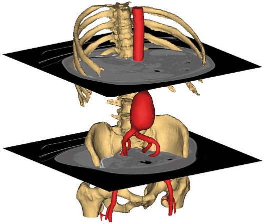


Figure: http://biomedical.materialise.com/mimics

- Simpleware

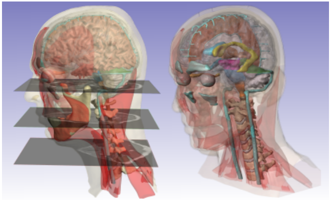
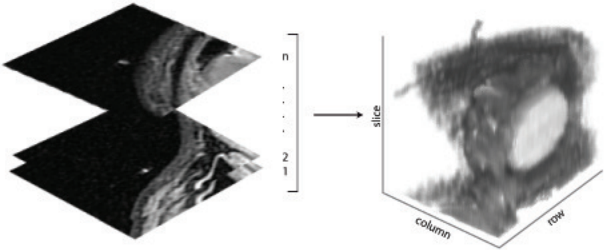


Figure: www.Simpleware.com

Three Dimensional Data



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

Example

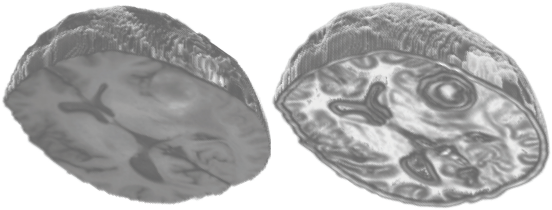


Figure: MR Volume (left). Surface Map (right)

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?

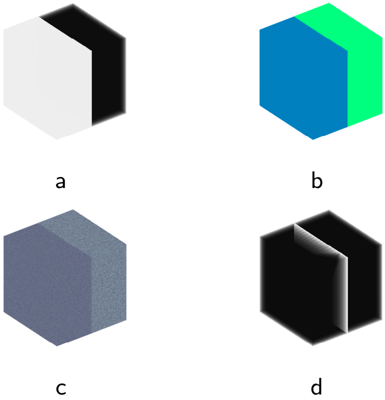
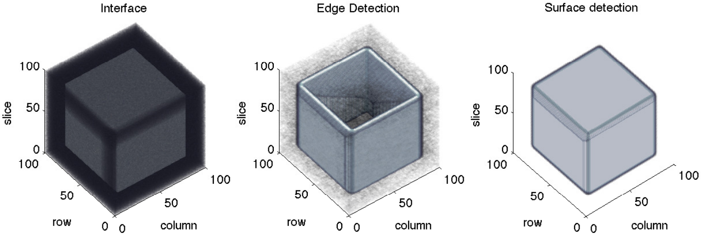


Figure: a) Intensity Interface. b) Colour Interface. c) Texture Interface. d) Interface Location

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.

Problems

- True 3D methods are not being utilised over multi-slice 2D methods
- Gradient methods fail to locate boundaries between regions of different texture when average intensity is uniform
- Gradient methods require a high degree of smoothing to prevent over detection, at the expense of correct responses and accurate localisation

Published Work

- Ian Williams, Nicholas Bowring, David Svoboda, A performance evaluation of statistical tests for edge detection in textured images, Computer Vision and Image Understanding, Volume 122, May 2014, Pages 115-130,
- Smith, S.; Williams, I., "A Statistical Method for Improved 3D Surface Detection," in Signal Processing Letters, IEEE , vol.22, no.8, pp.1045-1049, Aug. 2015

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.

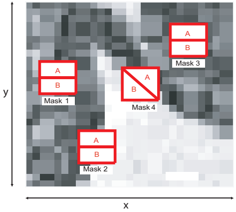


Figure: Mask 1,3, not located on a boundary (low output). Mask 2 located on boundary, incorrect orientation (low output). Mask 4 located on boundary, correct orientation (high output).

- 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.

Statistical Surface Detection Model

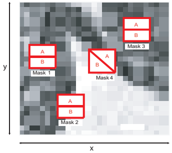
Extending upon the two dimensional statistical edge detection method, a model for detecting surfaces can be produced. The model should be able to:

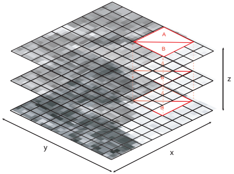
- Detect surfaces in three dimensional image volumes.
- Locate interfaces that lie in the plane of the surface operator.
- Determine interfaces at multiple intensity scales.
- Be robust to noise and image artefacts.

3D Statistical Detection

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.

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

Statistical test

Statistical surface detection can employ a range of different statistical comparison tests to identify different types of texture boundaries.

- Difference of Boxes
- Student t-Test
- Likelihood Test
- Fisher Test
- Kolmogorov-Smirnov Test
- χ^2 Test
- Robust Rank Order test
- Mann-Whitney U-Test

Student's T test

In probability theory, the first and second moments of a probability density function are mean and variance, therefore these two properties 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

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

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.

Difference of Boxes

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

$$D = |\bar{x}_A - \bar{x}_B| \tag{2}$$

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

χ^2 Test

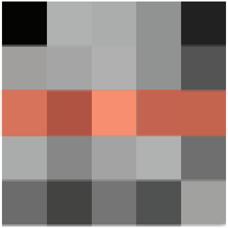
The χ^2 test is a rank based test which checks for the independence of the two sorted datasets. It is a comparison measure that takes the relative difference in points at the same rank position for two binned data sets.

χ^2 Test

$$\chi^2 = \sum_i \frac{R_i - S_i}{R_i + S_i} \tag{3}$$

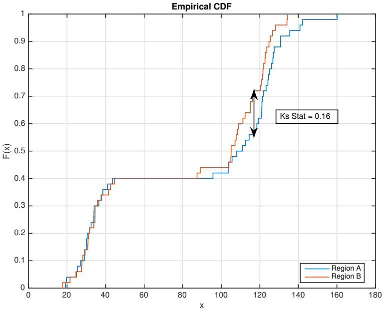
Here R_i is the number of values in *bin* i of region A, and S_i is the number of values in *bin* i of region B.

Two sample Kolmogorov-Smirnov test



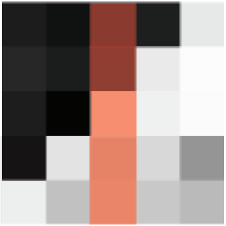
KS-Test

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



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

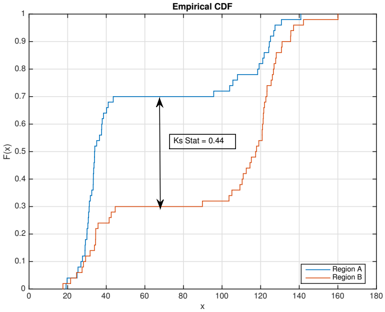
Two sample Kolmogorov-Smirnov test



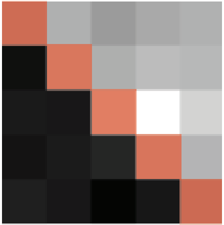
Where $F_{1,n}$ and $F_{2,n}$ are the empirical distribution functions of the two samples

KS-Test

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



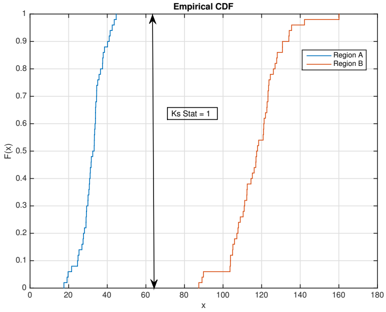
Two sample Kolmogorov-Smirnov test



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

KS-Test

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



Matlab

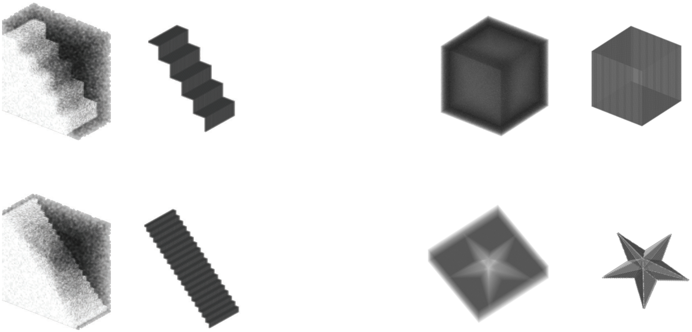
Problems

Solutions

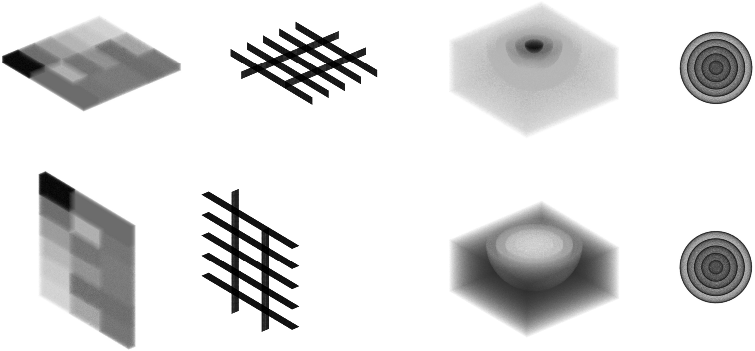
Synthetic Data Creation

- 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 - Single Scale



Synthetically Created data - Multi Scale

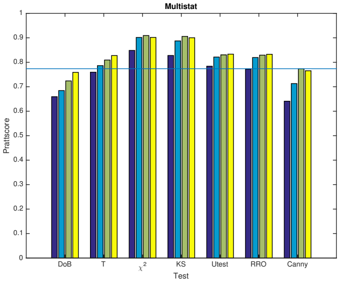


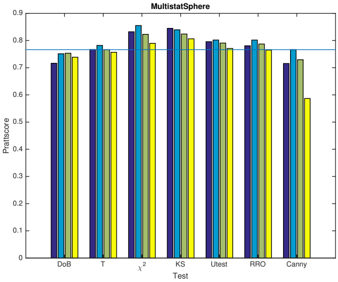
Monte Carlo Analysis

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

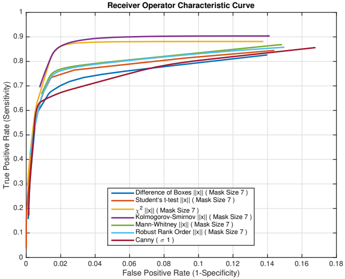
- Three examples of each image
- Pseudo-random values containing the same statistical properties for each image
- Mean and standard deviation of results of each image type determines the score.

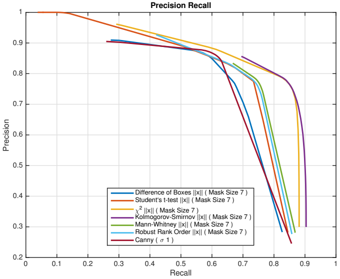
Results


Multistat
Prattscore
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
DoB
T
χ^2
KS
Utest
RRO
Canny
Test


MultistatSphere
Prattscore
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
DoB
T
χ^2
KS
Utest
RRO
Canny
Test

Results


Receiver Operator Characteristic Curve
True Positive Rate (Sensitivity)
False Positive Rate (1-Specificity)
0
0.02
0.04
0.06
0.08
0.1
0.12
0.14
0.16
0.18
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Difference of Boxes ||x|| (Mask Size 7)
Student's t-test ||x|| (Mask Size 7)
χ^2 ||x|| (Mask Size 7)
Kolmogorov-Smirnov ||x|| (Mask Size 7)
Mann-Whitney ||x|| (Mask Size 7)
Robust Rank Order ||x|| (Mask Size 7)
Canny (σ 1)


Precision Recall
Precision
Recall
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Difference of Boxes ||x|| (Mask Size 7)
Student's t-test ||x|| (Mask Size 7)
χ^2 ||x|| (Mask Size 7)
Kolmogorov-Smirnov ||x|| (Mask Size 7)
Mann-Whitney ||x|| (Mask Size 7)
Robust Rank Order ||x|| (Mask Size 7)
Canny (σ 1)

Visual Results - Sphere

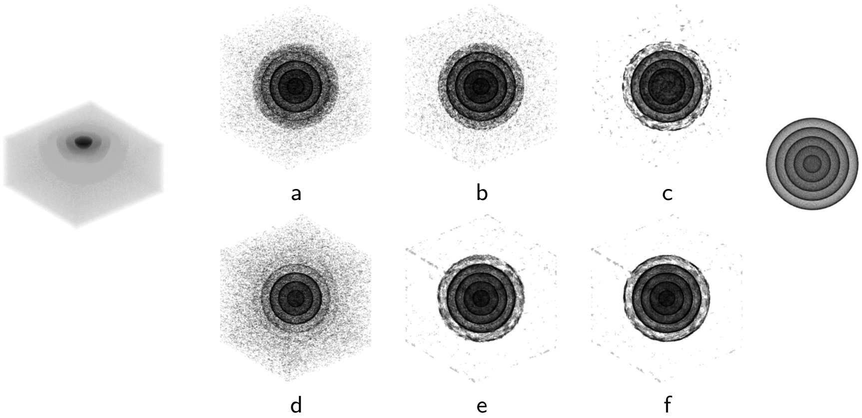


Figure: a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical f) 3D Statistical.

Visual Results -Sphere Reversed

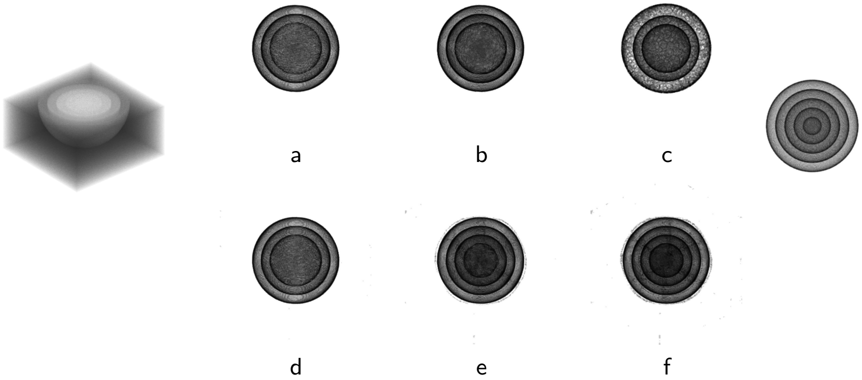


Figure: a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical f) 3D Statistical.

Visual Results - Multiple Scale

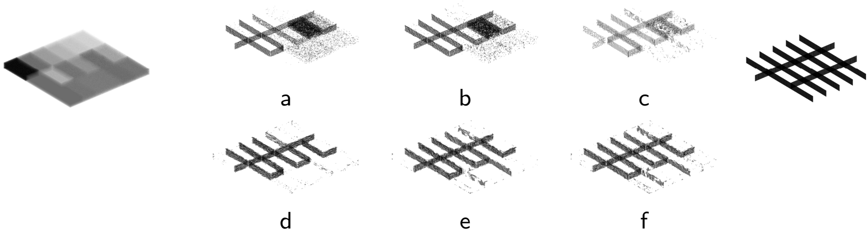


Figure: a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical f) 3D Statistical.

Visual Results - Multiple Scale Rotated

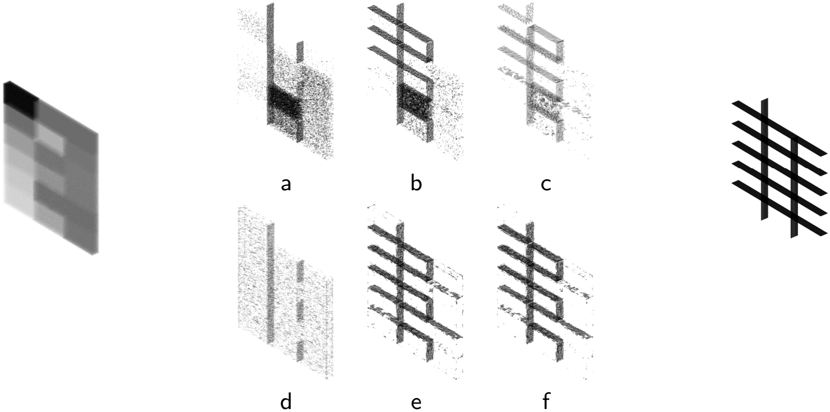
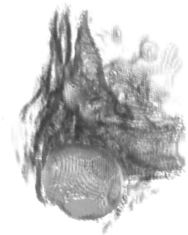
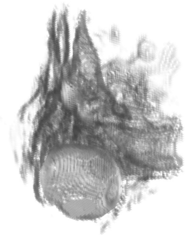


Figure: a) Canny 2D. b) Canny 3D. c) Steerable. d) 2D Statistical. e) 3D Statistical f) 3D Statistical.

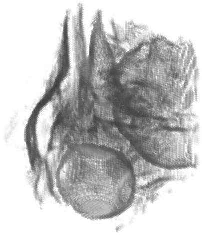
Real Image Results



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)