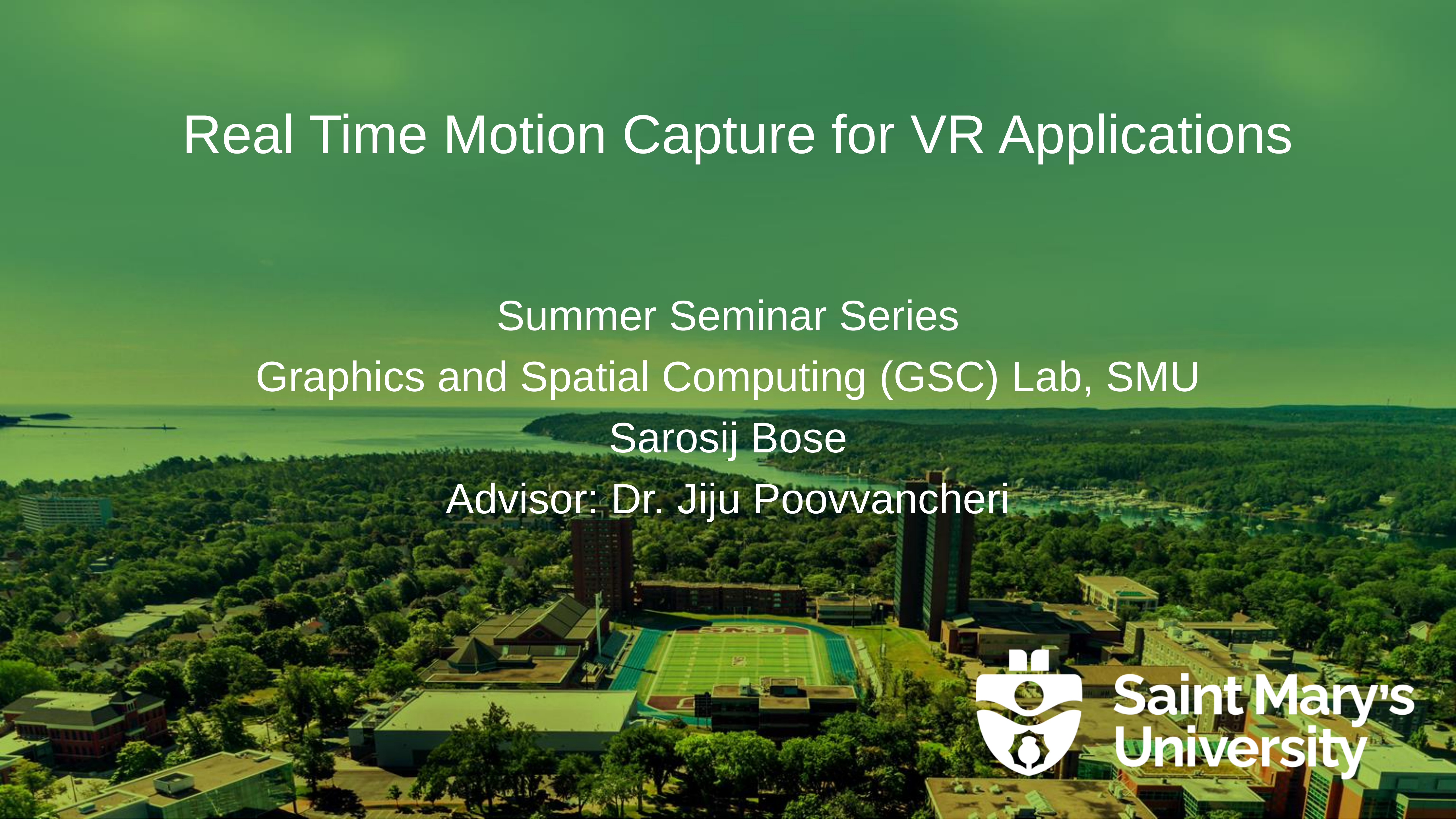


Real Time Motion Capture for VR Applications

Summer Seminar Series

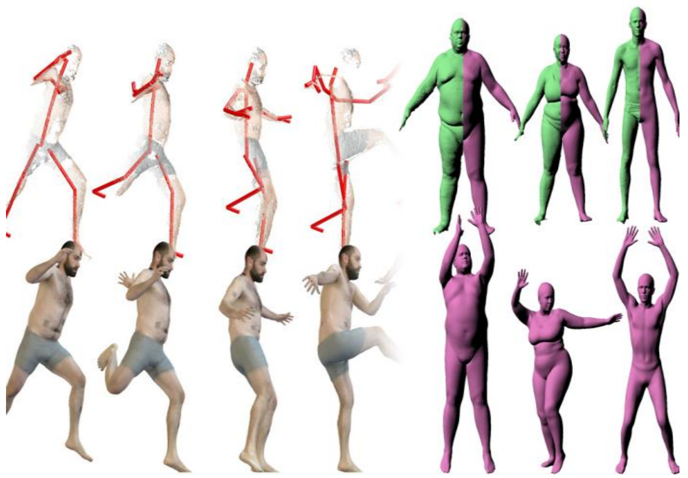
Graphics and Spatial Computing (GSC) Lab, SMU

Sarosij Bose

Advisor: Dr. Jiju Poovvancheri

Saint Mary's University

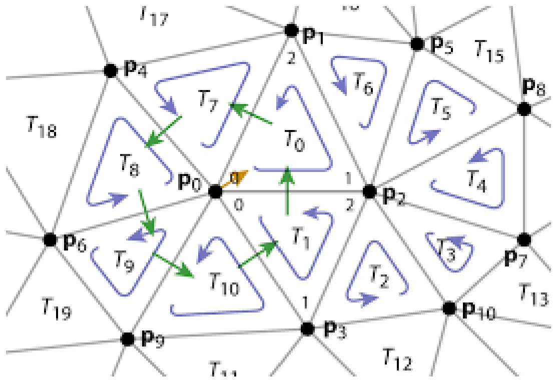
Introduction



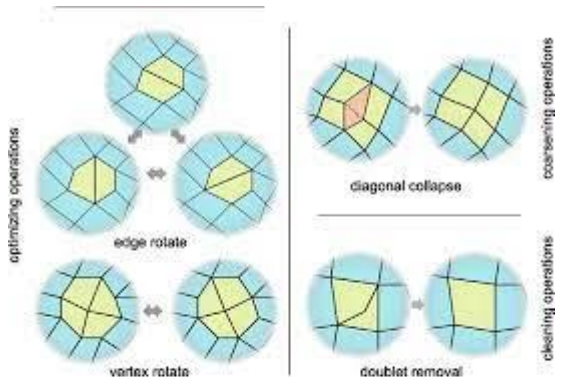
Modern devices such as Virtual Reality (VR) need robust real time body tracking systems to track the particular objects which may come in front of it. This means there is a need of an **appropriate representation**, which when combined with an **lightweight customized tracking algorithm** can be readily deployed with ease on an **edge device**.

Saint Mary's University

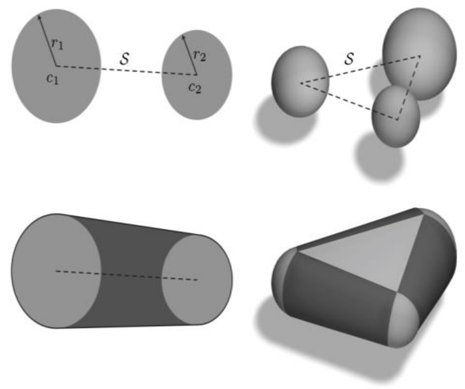
Geometric Representations



Traingle Mesh



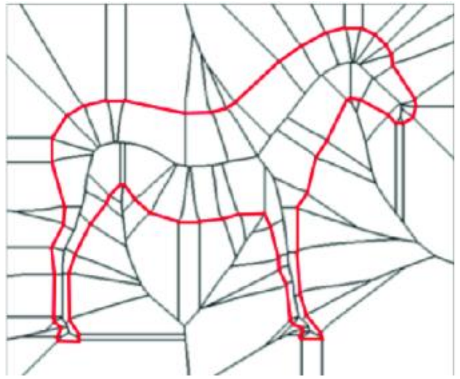
Quad Mesh

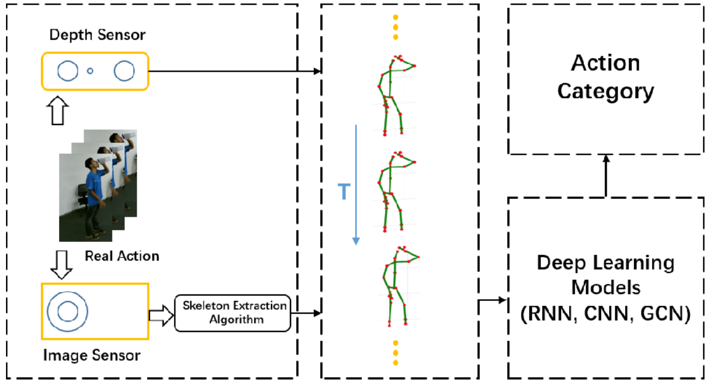


Sphere Mesh

Saint Mary's University

Existing Approaches




Depth Sensor
Real Action
Image Sensor
Skeleton Extraction Algorithm
T
Action Category
Deep Learning Models (RNN, CNN, GCN)

Voronoi diagram/Ray tracing based methods

The idea is to construct an approximate represenatation using a basic unit (covered below). Such methods are less prone to discretization errors leading to higher precision but they require heavy computational requiements.

Deep Networks

Deep Learning based networks extracts the features (ex: skeleton points) from a certain Region of Interest (ROI) and then uses a GCN (or CNN in some cases) to obtain the frame level prediction. This is then extrapolated over the previous/next frames if necessary.

Saint Mary's University

Tracking Techniques

Tracking techniques can be broadly divided into 2 categories:-

- **Generative** Tracking
- **Discriminative** Tracking

Each of these approaches have their own respectively advantages and disadvantages:-

- Discriminative tracking procedures are usually **more precise** (i.e; prevent error propagation) and they need more data.
- Generative tracking procedures can be **re-initialized and they are based on templates**. However, final convergence and optimization depends significantly on the initial template chosen.

An ideal scenario would usually be a **mix of the two** methods: Develop an initial **novel representation** (generative) and use it for **realtime** (discriminative) hand tracking.

Saint Mary's University

Sphere Mesh

Why Sphere Mesh?

- In [1], it was shown that a sphere mesh based tracking template performed significantly better than other template based models. Some of it's advantages were:-

  * Complex motion tracking

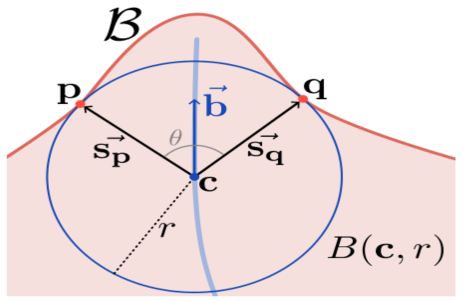
  * Tracking at **60 FPS**

  * No need of **per-frame re-initialization**.

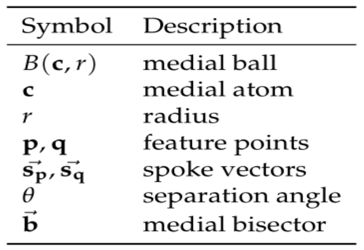
An additional advantage of using sphere meshes is that it's exceptionally stable – usually MAT based algorithms are known to be **notoriously hard to properly optimize** to a particular shape's boundary [2]. This in turn helped to design our proposed algorithm efficiently.

Saint Mary's University

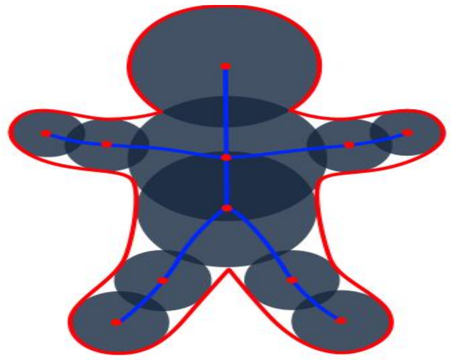
Medial Ball and Axis



Medial Ball B structure



Symbol	Description
$B(\mathbf{c}, r)$	medial ball
$\mathbf{c}$	medial atom
r	radius
$\mathbf{p}, \mathbf{q}$	feature points
$\vec{\mathbf{s_p}}, \vec{\mathbf{s_q}}$	spoke vectors
θ	separation angle
$\vec{\mathbf{b}}$	medial bisector



Medial Axis representation of a 2D human

Medial Balls

- Medial balls are the maximal empty balls with respect to a particular surface S.

Medial Axis (MA)

- Medial Axis of a closed surface is the set of centers of empty balls which touch the surface at more than one point. [3]

Saint Mary's University

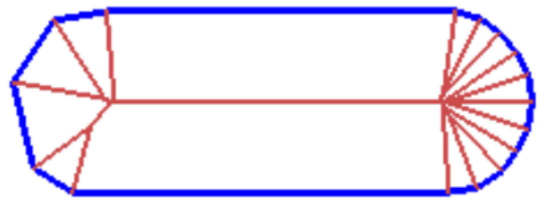
Medial Axis Transform (MAT)

What does Medial Axis Transform (MAT) actually mean?

- A medial ball is a ball that fits completely inside B and does not contain any other ball that would fit inside B. The MAT is defined as the set of points that are the centres of all medial balls of B (see Figure in the previous slide). Each medial ball touches the medial ball B in at least two points, called its feature points.

Why is it useful?

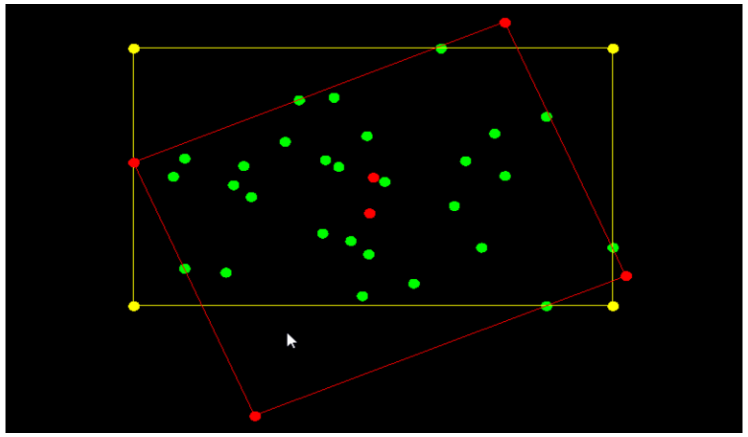
The MAT can be typically subdivided into open clusters that correspond to various features such as curves and sharp edges in the given shape. Further, such representations expose different properties of a shape which has their own unique use cases.



MAT for a random shape

Rotating Calipers (R.C.) Algorithm

Rotating Calipers Algorithm: This algorithm is used to find out the set of antipodal points in a convex 2D hull. This algorithm has a complexity of **O(NLogN)**. This is especially useful in implementing our algorithm as we are required to find out the furthest point in the surface from the initial chosen point.



Convex Hull for a set of points. Points in red are antipodal.

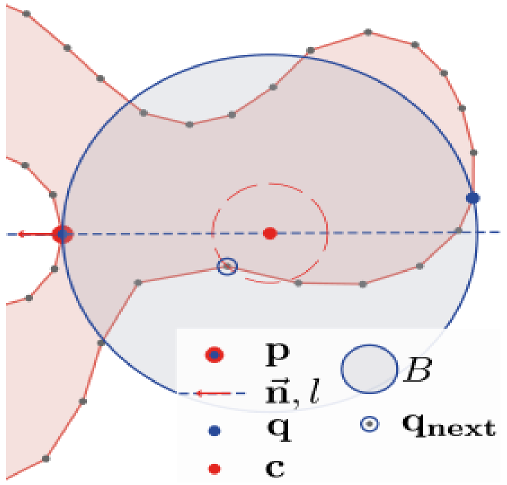
Initial Ball and Point Initialization

Initial Ball construction:-

We use the algorithm proposed in [4] to obtain the initial maximal ball. The initial ball radius is set which is iteratively made to be maximal by running a KDTree query to find out the nearest neighbours from that point and checking whether the radius equals the distance from the sample point to the center of the ball under consideration.

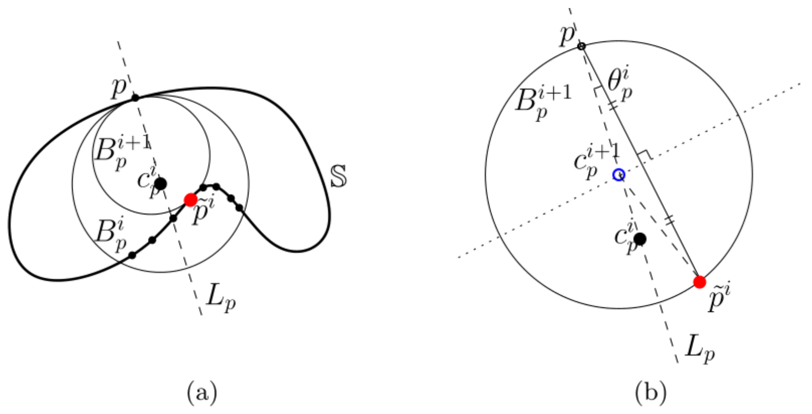
Initial point:-

This can be any random point from the given set of coodinates.



Saint Mary's University

Radius Calculation



Iterative refinement of the radius at each step

Radius Calculation

The formula for the radius at next step can be given by:-

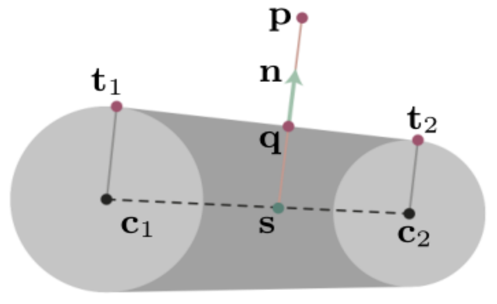
$$\rho_p^{i+1} = \frac{d(p, \tilde{p}^i)}{2\cos\theta_p^i},$$

Ex, p = [-0.543455 3.69187 -0.297564 -0.00031227 0.000179097 1]

Pill: Another Representation

Pill: The 'Pill' data structure can be thought of as a set of 2 balls which are connected by 2 direct common tangents (DCTs). It can be thought of as a **'primitive'** building unit.

Applications: This construct is particularly important for approximating surfaces – these can be thought of as a **fundamental representation** which can be further broken down as far as possible. The more these 'pills' are broken down, the better the obtained figure is for that shape.



A basic Pill structure

Saint Mary's University

Proposed Algorithm

Our proposed algorithm can hence be summarized down into the following steps:-

Step 1: *Set the initial point*

Step 2: *Construct the initial ball*

Step 3: *Find out the furthest point on the 2D surface using the R.C. algorithm*

Step 4: *Construct another ball obtained on Step 3.*

Step 5: *Using the balls from Step 2 and Step 4, construct the initial pill.*

Step 6: *Project normally from the obtained pill to find out the furthest point on the given surface.*

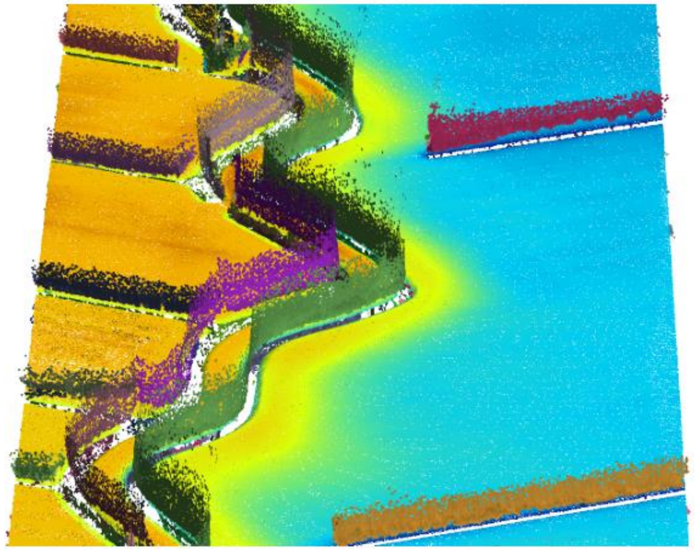
Step 7: *Break down the pill obtained on Step 5 by constructing two new pills using the ball constructed on the point in Step 6.*

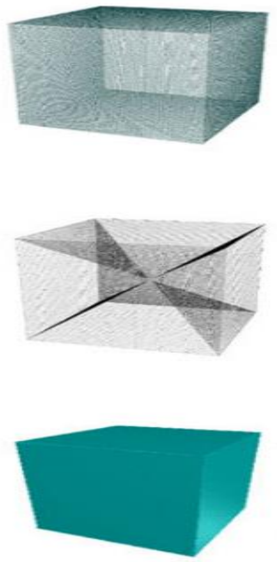
Step 8: *Iteratively repeat step 7 for the given budget.*

Saint Mary's University

Future Applications

Real time VR tracking


Watercourse Detection


Object Reconstruction

And many more….

Saint Mary's University

References

[1] Tkach A, Pauly M, Tagliasacchi A. Sphere-meshes for real-time hand modeling and tracking. ACM Transactions on Graphics (ToG). 2016 Nov 11;35(6):1-1.

[2] Katz RA, Pizer SM. Untangling the Blum medial axis transform. International Journal of Computer Vision. 2003 Nov;55(2):139-53.

[3]https://groups.csail.mit.edu/graphics/classes/6.838/F01/lectures/MedialAxisEtc/presentation/5.html

[4] Ma J, Bae SW, Choi S. 3D medial axis point approximation using nearest neighbors and the normal field. The Visual Computer. 2012 Jan;28(1):7-19.

Saint Mary's University

THANK YOU!

Saint Mary's University