Mapping 3D objects using a single camera - Luca Mansari

[1]Jack Collins, [1]David Colson, [1]Cliona Donnelly, [1]Sam Geragthy, [1]Treacy Hanley, [2]Sam Kamperis, [1]Faik Mayah, [1]Suman Samantray

[1]NUI Galway, [2]Oxford Brookes

June 9, 2017

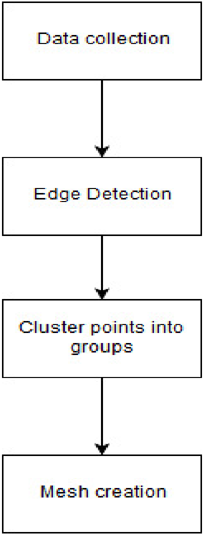
Problem description

- Extraction of depth information and 3D construction of an object.
- Constrained by single camera with one directional movement.

Assumptions:

- Distortion free camera.
- Negligible velocity change from different wavelengths of light.

Sequence of process



Method of Experimentation

- Optical rails used to confine movement of camera to one direction.
- Pictures taken of objects in increments of 5mm.
- Camera focal length set to 350mm.

Edge Detection - Sobel Filter

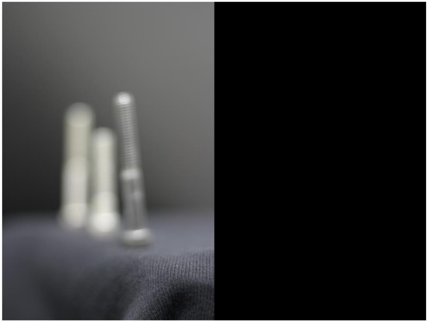
- Grayscale image represented by mxn matrix.
- Filter formulated by two 3x3 matrices.

$$G_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

- "Stamped" over each image matrix element to map horizontal and vertical changes.
- Calculate magnitude of G

$$G = \sqrt{G_x^2 + G_y^2}.$$

Animation of Process



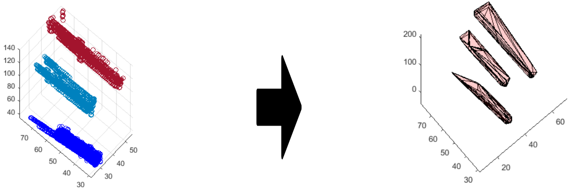
K-Means Clustering

1. Define k clusters.
2. Place k centroids into the object space.
3. Assign each element to closest centroid group.
4. Recalculate position of k centroids as barycentres of the clusters.
5. Repeat until results converge.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left|\left| x_i^{(j)} - c_j \right|\right|^2$$

Mesh creation - Delaunay Triangulation

- Tetrahedra constructed over set of P points such that no point in P is inside the circumsphere of any tetrahedra.
- Easily affected by outliers from noise.
- Image resolution reduction and edge threshold increase yielded best results.

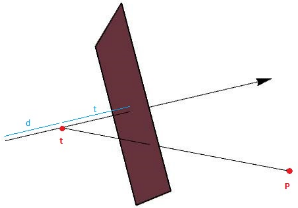


Calibration - Pixel to Metric Space

- Graph paper positioned 350mm from camera.
- Edges of area viewed by camera were manually identified and measured.
- This gave a 105mm x 70mm frame.

Projection Geometry

- We will look now at how a camera translates between pixel coordinates and real world coordinates.
- As the camera moves, the origin moves a distance t from the origin along the z axis. We consider a vector $\mathbf{t} = (0, 0, t)$.

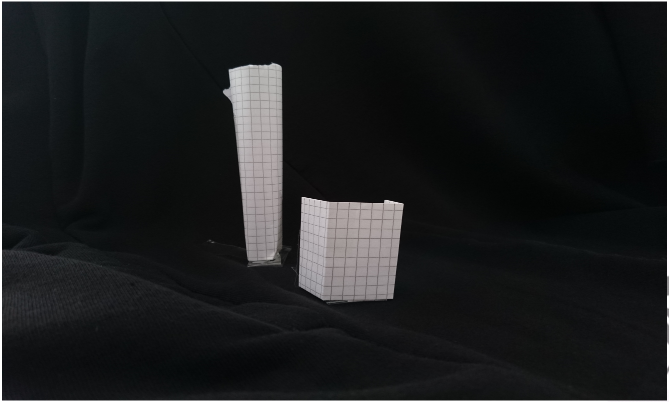


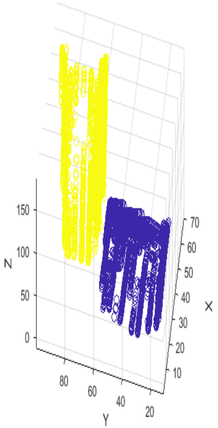
Projection Geometry

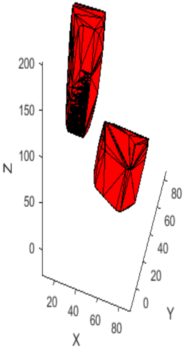
- The object at a point $\mathbf{P} = (P_x, P_y, P_z)$ is projected onto a plane $z = t + d$ to the point $\mathbf{R} = (R_x, R_y)$ which was solved to be $\frac{d}{P_z - t}\begin{pmatrix} P_x \\ P_y \end{pmatrix}$
- The equation of the line on which the camera moved was $L = k(\mathbf{P} - \mathbf{t})$ where k is a constant.
- We then solved this equation $\mid P - t \mid = \sqrt{P_x^2 + P_y^2 + (P_z - t)^2} = d + t$ for $P_z = \frac{d(d+t)}{\sqrt{R_x^2 + R_y^2 + d^2}}$

Extra data

- A 2nd data set was taken of some larger shapes.
- Graph paper yielded cleaner, more accurate results.




Z
150
100
50
0
80
60
40
20
Y
70
60
50
40
30
20
10
X


Z
200
150
100
50
0
20
40
60
80
X
80
60
40
20
0
Y

Further Research

- Relationship between camera characteristics, dimensions of objects and accuracy of results.
- Creating automatic threshold and cluster algorithms.
- Multiple mappings by rotating object to create more accurate 3D image.
- Segmentation of objects (multiple objects in one cluster).