Diego Martin Crespo

A20432558

Fall 2018

Computer Vision : $\boxed{\text{Assignment 5}}$

[1] ROBUST ESTIMATION AND SEGMENTATION

a) Outliers are pixels which values does not follow the ones from the rest of the pixels. It can be due to noise, occlusion or alignment error for example.

b) $E(\theta) = \sum_{i=1}^{\wedge} \rho_\sigma(d(x_{ij}, \theta))$. The standard least error equation is $\rho_\sigma(x) = x^2$ which is more sensitive to outliers.

c) $\rho_\sigma(x) = \frac{x^2}{x^2+\sigma^2}$. Its advantage is that it puts a limit of 1 to the error for outliers. If we choose a large σ we may include outliers but if it is too small we not include enough points so we can estimate σ as $\sigma = 1.5\, \mathrm{med}(d(x_{ij}, \theta))$

d) RANSAC is an iterative method to estimate the parameters of a mathematical model from a data set that contains outliers. It provides a resonable result only with a certain probability that increases with the number of iterations. It should be smaller to reduce de number of outliers (reduce de probability of outliers)

e) The parameters of the RANSAC are:

n = number of points at each iteration

d = minimum number of points needed

k = number of trials

t = distance to determine outliers

w = probability that a point is a inlier

e) Colors, texture and location.

- Merge approach: start with each pixel in separate cluster. Iteratively merge clusters.
- Split approach: start will all pixels in one cluster. Iteratively splits clusters.

g) K-means: select k (number of clusters) with an initial guess of k-means $\{m_i\}_{i}^{k}$

Assign $l_i = \underset{j \in [1,k]}{\operatorname{argmin}} ||f_i - m_j||^2$ for each pixel and assign the correspond cluster.

Recompute the mean: $m_j = \frac{\sum_{i \in S_j} f_i}{\# S_j}$; $S_j = \{i \mid l_i = j\}$

Stop when m_j does not change

- Mixture of Gaussians - it is like k-means replacing
  $d = ||f_i - m_j||^2$ with $d = (f_i - m_j)^T \Sigma_j^{-1} (f_i - m_j)$

$$m_j = \frac{\sum f_i}{\# S_j} \qquad \Sigma_j = \frac{\sum (f_i - m_j)(f_i - m_j)^T}{\# S_j}$$

h) It is similar to k-means to:

Give a weight to each sample $\quad m_j = \frac{\sum_{i \in S_i} w(f_i - m_j) f_i}{\sum_{i \in S_j} w(f_i - m_j)}$

The closest a sample is to the mean, the more it affects it.

[2] CAMERA CALIBRATION

a) • Forward Projection: given a 3D word point project into the image using the projection matrix M.
• Calibration: given P_i world points and corresponding p_i image points find M and the intrinsic and extrinsic parameters
• Reconstruction: given p_i points in image and M find corresponding P_i world points.
The easiest one is forward projection and the most dificult reconstruction

b) A set of 3D world points and its image 2D corresponding points

c) ① Find the projection of matrix M
② Find the parameters from M

d) $p_i = M P_i = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 3 & 4 \\ 1 & 1 & 1 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 18 \\ 14 \\ 7 \end{pmatrix} \to \begin{pmatrix} 18/7 \\ 2 \end{pmatrix}$

e) $p_i = M P_i$; $\begin{pmatrix} 100 \\ 200 \\ 1 \end{pmatrix} = \begin{pmatrix} 10 & 20 & 10 & 20 \\ 20 & 20 & 40 & 20 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \\ 1 \end{pmatrix}$

f) Since we get 2 equations from each point and we have 11 unknowns we need 6 points $\Rightarrow$ 12 eqs

g) Extract parameters using orthogonality of $r_1\ r_2\ r_3$:
$r_1 \cdot r_2 = 0$, $r_2 \cdot r_3 = 0$, $r_1 \cdot r_3 = 0$
$r_1 \times r_2 = r_3$, $r_2 \times r_3 = r_1$, $r_3 \times r_1 = r_2$

h) we need to compute the error:
$$E(K^*, R^*, T^*) = \sum_{i}^{n} \left(x_i - \frac{m_1^T P_i}{m_3^T P_i}\right)^2 + \left(y_i - \frac{m_2^T P_i}{m_3^T P_i}\right)^2$$

i) In planar Calibration we have a target and we want to know the coordinates in the calibration plane. The error is the euclidean distance.

j) $p_i = M P_i$ $\quad M = K^* [r_1\ r_2\ r_3\ T^*]$

$M = K^* [r_1\ r_2\ T^*]$ assumes $Z = 0$