Autonomous Mobile Robot Design

Topic: Stereo

Dr. Kostas Alexis (CSE)
Shape from X

- Recovery of 3D shape from one or two images.
Shape from X

- Shape from...
  - Stereo
  - Motion
  - Shading
  - Photometric stereo
  - Texture
  - Contours
  - ...

Shape from X
Shape from Stereo
Stereo

B=Baseline
f=focal length
C1 and C2=Camera centers
X1, X2=Image location in left and right cameras

\[
\frac{Z + f}{Z} = \frac{x_1 + x_2 + B}{B}, \quad Z = \frac{fB}{x_1 + x_2}
\]

\[x_1 + x_2 = \text{disparity} = d\]
Stereo pairs and Depth maps
Rectification
Correspondence using Search

Criterion function:

\[ Z = \frac{fB}{d} \]
Correlation Based Stereo Methods

- Disparity map can be constructed based on a correlation measure

- Sum of Square Differences (SSD): \[ SSD = \sum \sum (I_{left} - I_{right})^2 \]

- Absolute Difference (SSD): \[ AD = \sum \sum |(I_{left} - I_{right})| \]

- Cross Correlation (CC): \[ CC = \sum \sum I_{left}I_{right} \]

- Normalized Correlation (NC): \[ NC = \frac{\sum \sum (I_{left}I_{right})}{\sqrt{\sum \sum I_{left}I_{right}}} \]

- Mutual Correlation (MC): \[ MC = \frac{1}{64\sigma_{left}\sigma_{right}} \sum \sum (I_{left} - \mu_{left})(I_{right} - \mu_{right}) \]
Correlation

- Similarity/Dissimilarity Measures
  - Sum of Squares Difference (SSD)
  - Normalized Correlation (NC)
  - Mutual Correlation (MC)
  - Mutual Information (MI) - $I(x, y) = \sum \sum p(x, y) \log \frac{p(x, y)}{p_1(x)p_2(y)}$

- Use
  - Gray levels
  - Laplacian of Gaussian
  - Gradient magnitude
Block Matching

- Can be used for
  - Computing MPEG motion vectors
  - Optical flow
  - Stereo (displacement limited to only x-axis)
  - Image matching
Block Matching

- For each 8x8 block, centered around pixel \((x, y)\) in right image, \(B_k\)
  - Obtain 16x16 block in left, centered around \((x, y)\), \(B_{k-1}\)
  - Computer Sum of Squares Differences (SSD) between an 8x8 block, \(B_k\), and all possible 8x8 blocks in \(B_{k-1}\)
  - The 8x8 block in \(B_{k-1}\) centered around \((x', y')\), which gives the least SSD is the match
  - The displacement vector (disparity, optical flow) is given by \(u = x - x'\), \(v = y - y'\)
Sum of Squares Differences (SSD)

\[
(u(x, y), v(x, y)) = \arg \min_{u,v=-4...4} \sum_{i=0}^{7} \sum_{j=0}^{7} \left( f_k(x+i, y+j) - f_{k-1}(x+i+u, y+j+v) \right)^2
\]
Maximum Absolute Difference (MAD)

\[(u(x, y), v(x, y)) = \arg\min_{u,v=-4...4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} \left| f_k(x + i, y + j) - f_{k-1}(x + i + u, y + j + v) \right| \]
Maximum Matching Pixel Count (MMPC)

\[ T(x, y; u, v) = \begin{cases} 
1 & \text{if } |f_k(x, y) - f_{k-1}(x+u, y+v)| \leq t \\
0 & \text{otherwise}
\end{cases} \]

\[ (u(x, y), v(x, y)) = \arg \max_{u, v = -4 \ldots 4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} T(x+i, y+j; u, v) \]
Cross Correlation

\[
(u(x, y), v(x, y)) = \arg \max_{u,v=-4\ldots 4} \sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_k(x + i, y + j) \cdot f_{k-1}(x + i + u, y + j + v))
\]
Normalized Correlation

\[ (u,v) = \arg \max_{u,v=-A...A} \frac{\sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_k(x+i,y+j) - \mu_1)(f_{k-1}(x+i+u,y+j+v) - \mu_2)}{\sqrt{\sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_k(x+i,y+j) - \mu_1)^2} \sqrt{\sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_{k-1}(x+i+u,y+j+v) - \mu_2)^2}} \]

And \( \mu_1, \mu_2 \) are the mans of patch1, patch2 respectively.
Mutual Correlation

\[
(u(x, y), v(x, y)) = \arg \max_{u, v = -4 \ldots 4} \frac{1}{64 \sigma_1 \sigma_2} \sum_{i=0}^{-7} \sum_{j=0}^{-7} (f_k(x + i, y + j) - \mu_1) \cdot f_{k-1}(x + i + u, y + j + v) - \mu_2
\]

And \(\sigma, \mu\) are the standard deviation and mean (for patch1 and pathc2) respectively.
Bernard’s Stereo Method

- Similar intensity
- Similar to brightness constraint
- Smoothness of disparity

\[
E = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \left\| I_{\text{left}}(x+i, y+j) - I_{\text{right}}(x+i + D_x(x, y), y+j) \right\| + \lambda \left\| \nabla D(x, y) \right\|
\]

\[
\nabla D(x, y) = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \left| D(x+i, y+j) - D(x, y) \right|
\]
Bernard’s Stereo Method

- Energy can be minimized using brute force search
  - Let max allowed disparity is 10 pixels
  - For 128x128 image for 10 possible levels of disparity, there can be $10^{16384}$ possible disparity values.
- We can select any minimization technique
  - A critical point is to find a fast method
  - Bernard’s method employed Simulated Annealing
Simulated Annealing

- Sample a random state $S$ // disparities
- Select a high temperature $T$
  - Sample random $S'$
  - $\Delta E = E(S') - E(S)$
  - **If** $\Delta E < 0$
    - $S = S'$
  - **Else**
    - $P = \exp(-\Delta E / T)$
    - $X = \text{random}(0,1)$
    - **If** $X < P$
      - $S = S'$
    - **If** no decrease in several iterations
      - $T = T'$ // $T' < T$
Simulated Annealing

- Sample a random state $S$ // disparities
- Select a high temperature $T$
  - Sample random $S'$
  - $\Delta E = E(S') - E(S)$
  - If $\Delta E < 0$
    - $S = S'$
  - Else
    - $P = \exp(-\Delta E/T)$
    - $X = \text{random}(0,1)$
    - If $X < P$
      - $S = S'$
    - If no decrease in several iterations
      - $T = T'$ // $T' < T$
Stereo results

Scene

Ground truth
Stereo results

Scene

Ground truth
Stereo results (Data from Tsukuba University)

Window-based matching  Ground truth
Stereo results

State-of-the-art method

Ground truth

Boykov et al., “Fast Approximate Energy Minimization via Graph Cuts”, ICCV, 199
Stereo systems in Robotics

Uncertainty-aware Receding Horizon Exploration and Mapping using Aerial Robots
Christos Papachristos, Shehryar Khattak, Kostas Alexis

Real-Time Visual-Inertial Mapping, Re-localization and Planning Onboard MAVs in Unknown Environments
Michael Burri, Helen Oleynikova, Markus W. Achtelik and Roland Siegwart
Code Examples and Tasks


How does this apply to my project?

- To be able to perform stereo depth map and structure from motion.
Find out more

- “Fundamentals of computer vision”, Mubarak Shah
- Luong, Quan-Tuan, and Olivier D. Faugeras. "The fundamental matrix: Theory, algorithms, and stability analysis." International journal of computer vision 17.1 (1996): 43-75. APA
Seminar Lecture

- **Robotics Short Seminars**
- Next seminar talk:
  - **What:** Autonomous Exploration and Inspection via Active Perception
  - **When:** September 30, 2016, 4pm
  - **Where:** SEM201
  - **Who:** Kostas Alexis, University of Nevada, Reno
Thank you!
Please ask your question!