CS 378: Autonomous Intelligent Robotics

Instructor: Jivko Sinapov

http://www.cs.utexas.edu/~jsinapov/teaching/cs378/
Computer Vision: 2D Images
Announcements

• Homework 5 deadline extended to this Thursday
Announcements

Volunteers needed for another study!

As before, there will be extra credit

To sign up, email:
– Rodolfo Rodriguez <rcorona@utexas.edu>
– Jesse Thomason <thomason.jesse@gmail.com>
Final Project Timeline

• Project Proposal due: Mar. 29th  Apr. 1st

• Project Presentations / Demos: Last Week of Class (May 3rd and 5th)

• Final Report due: May 11th
Project Proposal Guidelines

• Work in groups of 2-3 (it's OK to work on your own if you really want to)

• Preferably, team up with people with different skills than yours

• Purpose of the proposal is to give you an outline / roadmap
Project Proposal Guidelines

• Each proposal should be about 2-3 pages

• Each proposal should include:
  – What is the application / task / problem?
  – Any previous experience you may have in that area
  – What do you expect to achieve by the end of the semester?
  – How do you plan to evaluate whether it works or not?
  – A timeline / schedule of progress and milestones
Project Proposal Guidelines

• Organization: your proposal should have sections and headings (don't just submit one long essay)

• For example:
  – Introduction / problem formulation
  – Proposed approach / software
  – Proposed evaluation
  – Summary of anticipated end result
Project Ideas

Vending Machine

Sonar Sensor
Project Ideas

Write ROS code to allow the robot to use an LED light strip
Project Ideas

Help the robot “see” something it currently cannot

Help the robot “hear” something (e.g., the elevator sound)

Help the robot “do” something (e.g., follow a person)
Project Ideas
Project Ideas
Project Ideas
Project Ideas
Final Project Timeline

The most important thing is to start early, and discuss your ideas with the TA, mentors and myself. We'll point you to a starting point, describe functionality that already exists, and help refine your ideas.
Final Project Timeline

• Project Proposal due: Mar. 29th  Apr. 1st

• Project Presentations / Demos: Last Week of Class (May 3rd and 5th)

• Final Report due: May 11th
Computer Vision: 2D Images
Readings


Readings (con't)

What is an image?
A grayscale image

<table>
<thead>
<tr>
<th>Index</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>32</td>
<td>33</td>
<td>34</td>
<td>35</td>
<td>36</td>
<td>37</td>
<td>38</td>
<td>39</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>41</td>
<td>42</td>
<td>43</td>
<td>44</td>
<td>45</td>
<td>46</td>
<td>47</td>
<td>48</td>
<td>49</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>51</td>
<td>52</td>
<td>53</td>
<td>54</td>
<td>55</td>
<td>56</td>
<td><strong>57</strong></td>
<td><strong>58</strong></td>
<td>59</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>61</td>
<td>62</td>
<td>63</td>
<td>64</td>
<td>65</td>
<td>66</td>
<td>67</td>
<td>68</td>
<td>69</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>71</td>
<td>72</td>
<td>73</td>
<td>74</td>
<td>75</td>
<td>76</td>
<td>77</td>
<td>78</td>
<td>79</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>81</td>
<td>82</td>
<td>83</td>
<td>84</td>
<td>85</td>
<td>86</td>
<td>87</td>
<td>88</td>
<td>89</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>91</td>
<td>92</td>
<td>93</td>
<td>94</td>
<td>95</td>
<td>96</td>
<td>97</td>
<td>98</td>
<td>99</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
An RGB image
How did computer vision start?

In 1966, Marvin Minsky at MIT asked his undergraduate student Gerald Jay Sussman to “spend the summer linking a camera to a computer and getting the computer to describe what it saw”. We now know that the problem is slightly more difficult than that!
Computer vision vs human vision

<table>
<thead>
<tr>
<th>What we see</th>
<th>What a computer sees</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>0</th>
<th>3</th>
<th>2</th>
<th>5</th>
<th>4</th>
<th>7</th>
<th>6</th>
<th>9</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Intensity Levels

- 2
- 32
- 64
- 128
- 256 (8 bits)
- 512
- ...
- 4096 (12 bits)
Intensity Levels

- 2
- 32
- 64
- 128
- 256 (8 bits)
- 512
- 1024
- ...
- 4096 (12 bits)
Image Plane v.s. Image Array

Point Operations

\[ f_A(x, y) \quad f_B(x, y) \]

\[(x, y) \quad (x, y)\]

Local Operations

\[ f_A(x, y) \]

\[ f_B(x, y) \]

(Jain, Kasturi, and Schunck (1995). Machine Vision, Ch. 1)

[Image of a jet plane flying over mountains]
Global Operations

\[ P = O_{\text{global}} \{ f[i,j] \} \]

Thresholding an Image

\[ f_A(x, y) \quad f_B(x, y) \]

\[(x, y) \quad (x, y)\]

Dark Image on a Light Background

\[ F_T[i, j] = \begin{cases} 
1 & \text{if } F[i, j] \leq T \\
0 & \text{otherwise.} 
\end{cases} \]

Selecting a range of intensity values

\[ F_T[i, j] = \begin{cases} 
1 & \text{if } T_1 \leq F[i, j] \leq T_2 \\
0 & \text{otherwise.} 
\end{cases} \]

Generalized Thresholding

A general thresholding scheme in which the intensity levels for an object may come from several disjoint intervals may be represented as

$$F_T[i, j] = \begin{cases} 
1 & \text{if } F[i, j] \in Z \\
0 & \text{otherwise}
\end{cases}$$

(2.4)

Thresholding Example (1)

Thresholding Example (2)

Original grayscale Image

\[ T = 48 \]
\[ T_1 = 2 | T_2 = 48 \]
Area of a Binary Image

\[ A = \sum_{i=1}^{n} \sum_{j=1}^{m} B[i, j]. \]

This figure now becomes important

Calculating the Position of an Object

\[
\bar{x} \sum_{i=1}^{n} \sum_{j=1}^{m} B[i, j] = \sum_{i=1}^{n} \sum_{j=1}^{m} jB[i, j]
\]

\[
\bar{y} \sum_{i=1}^{n} \sum_{j=1}^{m} B[i, j] = \sum_{i=1}^{n} \sum_{j=1}^{m} iB[i, j]
\]

The center is given by

\[
\bar{x} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} jB[i, j]}{A}
\]

\[
\bar{y} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{m} iB[i, j]}{A}
\]

Horizontal and Vertical Projections

Horizontal and Vertical Projections

Projection Formulas

\[ H[i] = \sum_{j=1}^{m} B[i, j] \]

\[ V[j] = \sum_{i=1}^{n} B[i, j]. \]

Diagonal Projection

The area and the position can be computed from the H and V projections

\[
A = \sum_{j=1}^{m} V[j] = \sum_{i=1}^{n} H[i]
\]

\[
\bar{y} = \frac{\sum_{i=1}^{n} iH[i]}{A}
\]

\[
\bar{x} = \frac{\sum_{j=1}^{m} jV[j]}{A}
\]

Neighbors and Connectivity
4-Connected

4-neighbors \([i + 1, j], [i - 1, j], [i, j + 1], [i, j - 1]\)
8-connected

8-neighbors \([i + 1, j + 1], [i + 1, j - 1], [i - 1, j + 1], [i - 1, j - 1]\) plus all of the 4-neighbors

Examples of Paths

(a) 4-path

(b) 8-path

Boundary, Interior, and Background

(a) Original image

(b) Boundary pixels
- Gray: Interior pixels
- White: Surrounds pixels
An Image (a) and Its Connected Components (b)

The RGB Color Space

[http://www.arcsoft.com/images/topics/darkroom/what-is-color-space-RGB.jpg]
The RGB Color Space

3D Scatter Plot for a patch of skin
The HSV Color Space
Color Detection and Segmentation
Color Detection and Segmentation
Discussion: how may we achieve this?
Example Hand Tracking using Color
Computer Vision in ROS
Computer Vision in ROS

1) Subscribing to an image topic
2) Converting a ROS image to an OpenCV image
3) Copy an image
4) Convert an image to grayscale
5) Access and set individual pixel values
Example Color Detection in ROS using OpenCV
Resources

• OpenCV in ROS:
  – http://wiki.ros.org/vision_opencv
  – http://wiki.ros.org/cv_bridge/Tutorials
  – http://docs.opencv.org/2.4/doc/tutorials/tutorials/tutorial.html
THE END