COMP 50: Autonomous Intelligent Robotics

Instructor: Jivko Sinapov

http://www.cs.tufts.edu/comp/50AIR/
Computer Vision: 2D Images
Semester Schedule

- Learning Robot Operating System (ROS)
- Learning to use our robots
- Computational Perception
- Developmental Robotics
- Human-Robot Interaction

You are here

Time
Announcements

• Homework 5 is due Friday 3/9
Announcements
Reading for this week


• Chapter 6, “Exploration, navigation, and localization” of Introduction to Autonomous Robots
Project Proposal Write-ups

• **Due** Friday, March 16
• **Length:** 4 to 5 pages
• **Sections:** Introduction, Related Work, Problem Formulation / Technical Approach, Evaluation Plan / Criteria for Success
• **Format:** default Google doc style or default overleaf.com LaTeX template
Computer Vision: 2D Images
Some material adapted from:

• Links are posted on the class website
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>57</td>
<td>58</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

<table>
<thead>
<tr>
<th>Array RGB</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.689</td>
</tr>
<tr>
<td>0.535</td>
</tr>
<tr>
<td>0.314</td>
</tr>
<tr>
<td>0.553</td>
</tr>
<tr>
<td>0.441</td>
</tr>
<tr>
<td>0.398</td>
</tr>
<tr>
<td>0.912</td>
</tr>
<tr>
<td>0.219</td>
</tr>
<tr>
<td>0.128</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Page 3 - blue intensity values</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.706</td>
</tr>
<tr>
<td>0.633</td>
</tr>
<tr>
<td>0.512</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Page 2 - green intensity values</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.515</td>
</tr>
<tr>
<td>0.712</td>
</tr>
<tr>
<td>0.121</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Page 1 - red intensity values</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.234</td>
</tr>
<tr>
<td>0.986</td>
</tr>
</tbody>
</table>
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

What we see

What a computer sees
Intensity Levels

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

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

Point Operations

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

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

Local Operations

$f_A(x, y)$  

$f_B(x, y)$  

Global Operations

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

Thresholding an Image

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

\((x, y)\)

\((x, y)\)

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

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

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
   Interior pixels
   Surrounds pixels

An Image (a) and Its Connected Components (b)

Color Perception
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?
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
6) Use rosbag to collect image data
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.shtml
THE END