Anyone here? Smart embedded low-resolution omnidirectional video sensor to measure room occupancy

Timothy Callemein, Kristof Van Beeck, Toon Goedemé
Current solution - PIR sensors

- No privacy issues
- Easy to install
- Cheap

- Requires a level of movement
- Binary output (YES / NO)
- No people count
Research Overview

Camera

Performance

High → Low

Privacy
Camera
Camera
Camera

- Large field-of-view
- Single camera
- Image distortion
- Limited number of annotated datasets

Can we count people using these images?
Approach – Generating data

Redmond et al. “YOLO9000: Better, Faster, Stronger” 2017 (CVPR)
Approach – Lowering network resolution

Decrease network resolution

(448; 160; 96)

Yolov2

Redmond et al. “YOLO9000: Better, Faster, Stronger” 2017 (CVPR)
Approach – Lowering image resolution

Decrease image resolution

(64; 48; 32)
Approach – Using temporal data

People move vs. Static background

Can we improve our system by integrating temporal information?
## Performance

<table>
<thead>
<tr>
<th>Device</th>
<th>Resolution</th>
<th>Seconds per Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raspberry Pi 2</td>
<td>448</td>
<td>18.60</td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>3.60</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>2.17</td>
</tr>
<tr>
<td>Raspberry Pi 3B</td>
<td>448</td>
<td>16.60</td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>2.96</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>1.83</td>
</tr>
<tr>
<td>Raspberry Pi 3B+</td>
<td>448</td>
<td>11.72</td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>2.07</td>
</tr>
<tr>
<td></td>
<td>96</td>
<td>1.30</td>
</tr>
</tbody>
</table>
Results

Test on model trained with generated labels
- Different network resolutions
- Different image resolutions

PIROPO

PRIVATE OFFICE

MIRROR A

MIRROR B

PIROPO - https://www.gti.ssr.upm.es/research/gti-data/databases
Results

**PIROPO**

- 160_64: 74.60%
- 160_48: 74.50%
- 160_32: 75.90%
- 96_48: 35.10%
- 96_32: 27.10%

**PRIVATE**

- 160_64: 86.60%
- 160_48: 75.10%
- 160_32: 71.00%
- 96_48: 64.10%
- 96_32: 66.00%

**MIRROR A**

- 160_64: 57.50%
- 160_48: 63.60%
- 160_32: 58.50%
- 96_48: 43.50%
- 96_32: 55.70%

**MIRROR B**

- 160_64: 93.90%
- 160_48: 92.60%
- 160_32: 91.60%
- 96_48: 87.70%
- 96_32: 87.20%
Results

PIROPO

High movement level

PRIVATE

Low movement level

MIRROR A

Medium movement level

MIRROR B

Medium movement level

\[
\begin{align*}
\begin{bmatrix}
0 & -1 \\
-1 & 0
\end{bmatrix}
\begin{bmatrix}
0 & 1 \\
-1 & 0
\end{bmatrix}
\begin{bmatrix}
0 & -1 \\
-2 & -3
\end{bmatrix}
\begin{bmatrix}
t & 0 \\
-1 & 0
\end{bmatrix}
\end{align*}
\]
Results

PRIVATE

<table>
<thead>
<tr>
<th>k_1</th>
<th>k_2</th>
<th>k_3</th>
<th>k^t_2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>t</td>
</tr>
<tr>
<td>-1</td>
<td>0</td>
<td>-1</td>
<td>-2</td>
</tr>
<tr>
<td>0</td>
<td>-1</td>
<td>-3</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>-t</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Bar chart showing results for 160_64, 160_48, and 160_32 with different K values.

- K1
- K2 (t=1)
- K3
- K2 (t=2)
- K2 (t=3)
Conclusion

• Training possible with automatically generated annotations
• Can run on embedded hardware
• Already good performance
• Image resolution of 32 pixels

Future Work

• Improving label generation
• Influence of large room changes?
• What when the room gets bigger?
Thank you for your attention! Questions?

timothy.callemein@kuleuven.be
phd.callemein.be – www.eavise.be