A Three-Layered Approach to Facade Parsing - Supplementary Material

Anđelo Martinović[1], Markus Mathias[1], Julien Weissenberg[2], and Luc Van Gool[1,2]

[1] ESAT-PSI/VISICS, KU Leuven
[2] Computer Vision Laboratory, ETH Zurich

1 Results - Confusion Matrices

First, we show the full confusion matrices for both datasets used in the paper. The results are shown for outputs of each layer.

Table 1: Confusion matrices for the **ECP** dataset. Left: Bottom layer output. Center: Middle layer output. Right: Top layer output.

$$
\begin{pmatrix}
62 & 24 & 10 & 0 & 3 & 0 & 1\\
3 & 91 & 3 & 0 & 0 & 0 & 3\\
8 & 17 & \mathbf{74} & 0 & 1 & 0 & 0\\
2 & 9 & 0 & 43 & 0 & 0 & 45\\
6 & 13 & 3 & 0 & 70 & 5 & 0\\
0 & 1 & 0 & 0 & 7 & \mathbf{91} & 0\\
1 & 16 & 1 & 3 & 0 & 0 & 79
\end{pmatrix}
\quad
\begin{pmatrix}
69 & 22 & 5 & 0 & 2 & 0 & 1\\
2 & \mathbf{93} & 1 & 0 & 0 & 0 & 3\\
9 & 18 & 71 & 0 & 1 & 0 & 0\\
2 & 7 & 0 & 60 & 0 & 0 & 31\\
6 & 13 & 1 & 0 & \mathbf{73} & 5 & 0\\
1 & 1 & 0 & 0 & 7 & \mathbf{91} & 0\\
1 & 11 & 0 & 1 & 0 & 0 & 86
\end{pmatrix}
\quad
\begin{pmatrix}
\mathbf{75} & 14 & 6 & 0 & 2 & 1 & 2\\
4 & \mathbf{88} & 3 & 0 & 0 & 0 & 6\\
7 & 22 & 70 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & \mathbf{67} & 0 & 0 & 32\\
9 & 9 & 2 & 0 & 74 & 5 & 0\\
0 & 0 & 0 & 0 & 2 & \mathbf{97} & 0\\
0 & 5 & 0 & 1 & 0 & 0 & \mathbf{93}
\end{pmatrix}
\begin{array}{l}
\textit{window}\\
\textit{wall}\\
\textit{balcony}\\
\textit{door}\\
\textit{roof}\\
\textit{sky}\\
\textit{shop}
\end{array}
$$

Table 2: Confusion matrices for the **eTrims** dataset. Left: Bottom layer output. Center: Middle layer output. Right: Top layer output.

$$
\begin{pmatrix}
89 & 1 & 0 & 0 & 0 & 0 & 2 & 7\\
17 & \mathbf{67} & 0 & 3 & 3 & 0 & 4 & 4\\
43 & 4 & \mathbf{24} & 0 & 0 & 0 & 6 & 23\\
23 & 12 & 0 & \mathbf{35} & 28 & 0 & 1 & 1\\
21 & 6 & 0 & 26 & \mathbf{47} & 0 & 0 & 0\\
8 & 0 & 0 & 0 & 0 & \mathbf{91} & 0 & 0\\
13 & 1 & 1 & 0 & 0 & 1 & 82 & 1\\
25 & 0 & 0 & 0 & 0 & 0 & 1 & 72
\end{pmatrix}
\quad
\begin{pmatrix}
\mathbf{90} & 1 & 0 & 0 & 0 & 0 & 2 & 6\\
21 & 66 & 1 & 2 & 2 & 0 & 4 & 4\\
48 & 3 & 20 & 0 & 0 & 0 & 4 & 25\\
23 & 12 & 0 & \mathbf{35} & 28 & 0 & 1 & 0\\
21 & 5 & 0 & 26 & \mathbf{47} & 0 & 0 & 0\\
10 & 0 & 0 & 0 & 0 & 90 & 0 & 0\\
13 & 1 & 0 & 0 & 0 & 1 & \mathbf{83} & 1\\
23 & 0 & 0 & 0 & 0 & 0 & 1 & 75
\end{pmatrix}
\quad
\begin{pmatrix}
86 & 1 & 0 & 0 & 0 & 0 & 1 & 10\\
19 & \mathbf{67} & 0 & 2 & 3 & 0 & 4 & 5\\
44 & 3 & 18 & 0 & 0 & 0 & 4 & 32\\
23 & 12 & 0 & \mathbf{35} & 28 & 0 & 1 & 1\\
21 & 6 & 0 & 26 & \mathbf{47} & 0 & 0 & 0\\
9 & 0 & 0 & 0 & 0 & \mathbf{91} & 0 & 0\\
13 & 1 & 1 & 0 & 0 & 1 & 81 & 3\\
19 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{80}
\end{pmatrix}
\begin{array}{l}
\textit{building}\\
\textit{car}\\
\textit{door}\\
\textit{pavement}\\
\textit{road}\\
\textit{sky}\\
\textit{vegetation}\\
\textit{window}
\end{array}
$$

2 Results - Output Labelings

On the next couple of pages we present some example results on both datasets. We show the outputs from all three layers of our system.

Fig. 1: Results on the ECP dataset. (a) The original image. (b-d) Outputs from the bottom, middle and top layers, respectively. (e) Ground truth.

Fig. 2: Results on the ECP dataset. (a) The original image. (b-d) Outputs from the bottom, middle and top layers, respectively. (e) Ground truth.

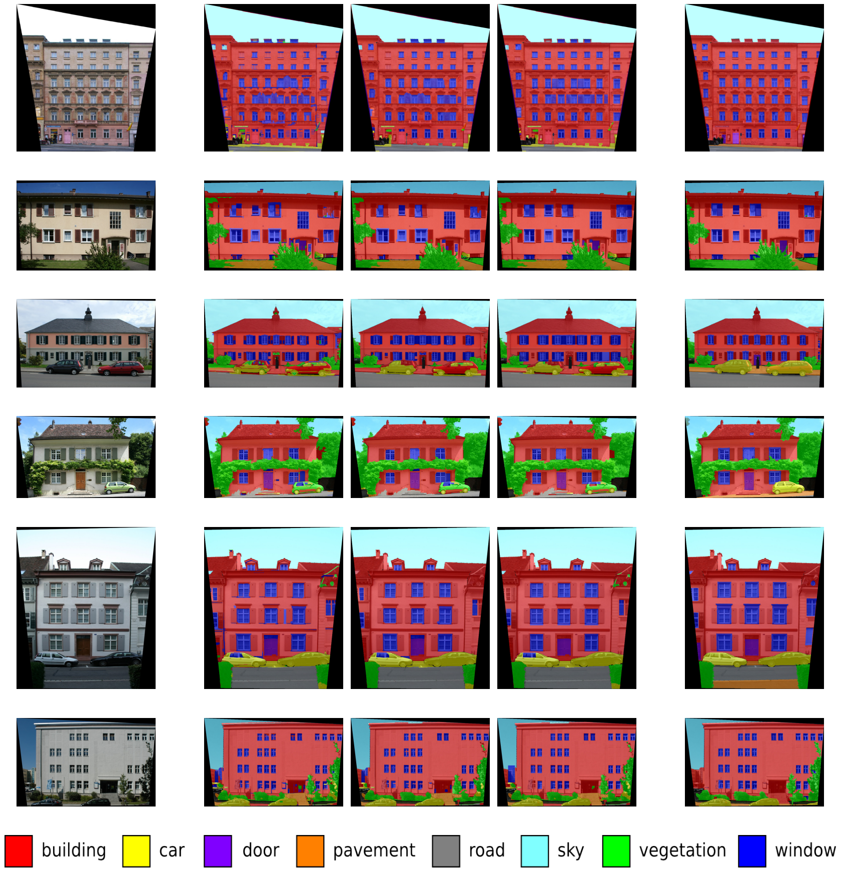

Fig. 3: Results on the eTrims dataset. (a) The original image. (b-d) Outputs from the bottom, middle and top layers, respectively. (e) Ground truth.

Fig. 4: Results on the eTrims dataset. (a) The original image. (b-d) Outputs from the bottom, middle and top layers, respectively. (e) Ground truth.

3 Procedural models

On the next page we show a few examples of procedural models created in CityEngine [1] directly from the output of our system, and rendered in Blender. The shop texture is projected from the original image, while all the other assets are imported from a database of basic architectural elements and textures.

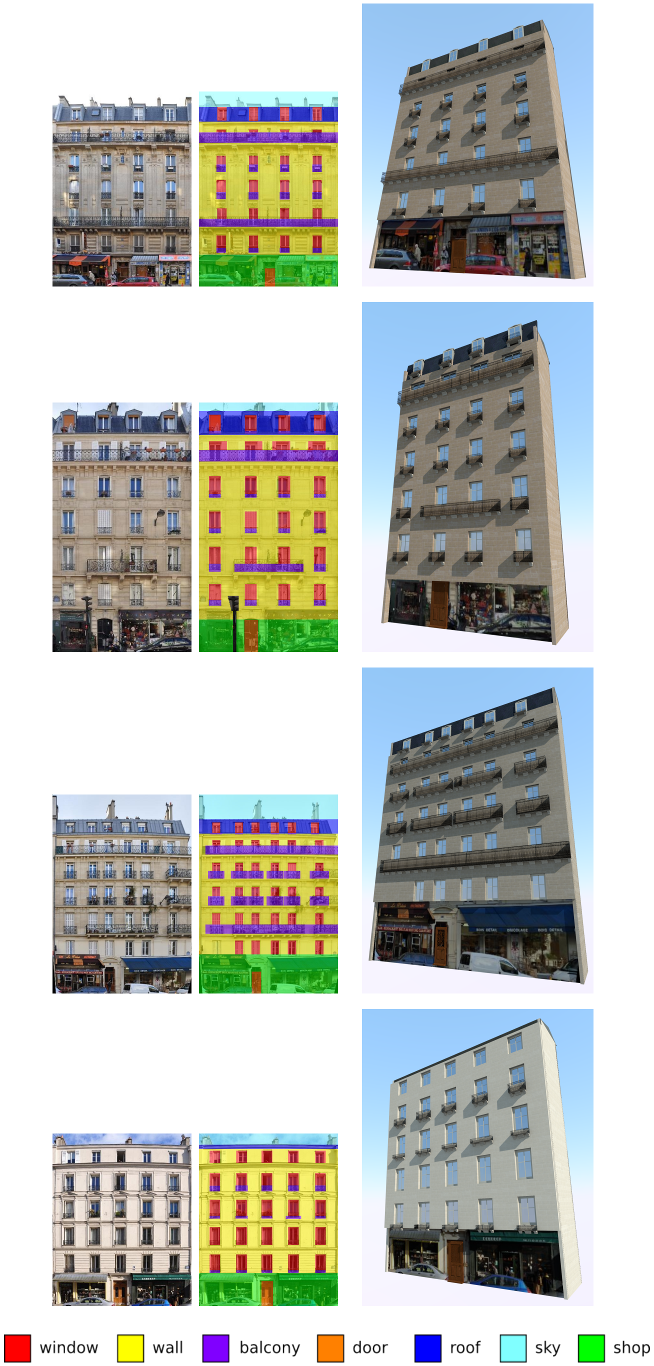

Fig. 5: Towards inverse procedural modeling. (a) Original image. (b) Top layer output of our system. (c) Procedural model.

4 Updated annotations for the ECP dataset

Fig. 6: Ground truth data. (a) Labels provided by [2]. (b) Original image. (c) Our updated ground truth.

References

1. Procedural: CityEngine. http://www.procedural.com/ (2010)
2. Teboul, O., Kokkinos, I., Simon, L., Koutsourakis, P., Paragios, N.: Shape grammar parsing via reinforcement learning. In: CVPR. (2011)