Joint 3D Scene Reconstruction and Class Segmentation
Supplementary Material

Christian Häne[1], Christopher Zach[2], Andrea Cohen[1], Roland Angst[1*], Marc Pollefeys[1]

[1]ETH Zürich, Switzerland
{chaene, acohen, rangst, pomarc}@inf.ethz.ch

[2]Microsoft Research Cambridge, UK
chzach@microsoft.com

1 Illustration of the Wulff Shapes

In the main text the two Wulff shapes "line segment" and "half-sphere plus spherical cap" are described. In Fig. 1 the Wulff shapes and the corresponding smoothness terms ψ^{ij} are depicted.

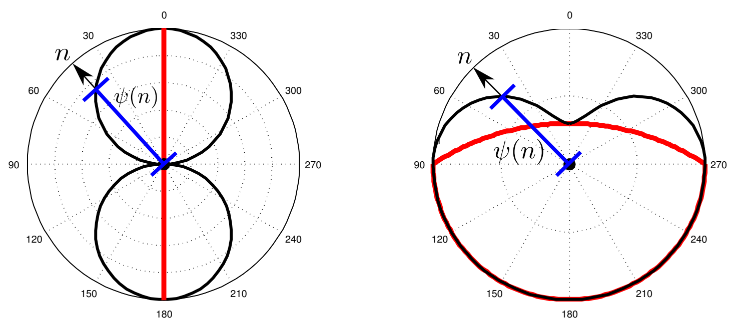


Figure 1: Visualization of the 2D version of the used Wulff shapes (line segment and half-sphere plus spherical cap). The red lines depict the Wulff shapes. The black lines are polar plots of the function ψ. The distance of a point on the black curve to the origin is the value that the function $\psi(n)$ attains for a normal vector n in the direction of the point (visualized in blue).

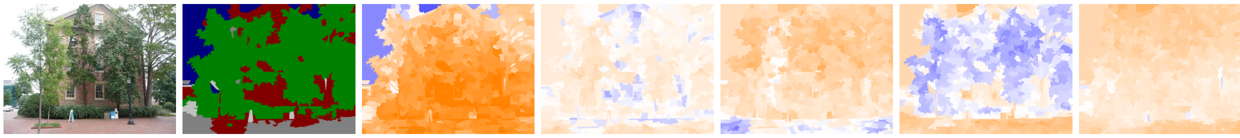
2 Output of the Boosted Decision Tree Classifier

This short section gives some additional information about the output of the boosted decision tree classifier used in this work.

The classifier trains a boosted decision tree for each of the labels separately in a one against all fashion. The output of the classifier is a score for each label which can be either positive or negative. In a probabilistic interpretation, these scores correspond to half the log likelihood-ratios [2]. This means that a positive score "votes" for a specific label and a negative one against it. In Fig. 2 the raw output of the classifier is visualized.

As explained in [2] in a multi-class setting the classifier scores can also be interpreted as unnormalized class log likelihoods.

*Now at Stanford University



(a) Input image (b) Best label (c) Sky (d) Building (e) Ground (f) Vegetation (g) Clutter

Figure 2: Output of the boosted decision tree classifier. Positive scores are visualized with blue colors and negative scores in orange. The intensity of the colors indicates the magnitude.

3 Inference

This section contains a description of our inference procedure. Therefore we first restate our objective

$$E_{\text{discr}}(\mathbf{x}) = \sum_{s\in\Omega}\left(\sum_i \rho_s^i x_s^i + \sum_{i,j:i<j} \phi^{ij}(x_s^{ij} - x_s^{ji})\right), \tag{1}$$

with the linear constraints

$$\begin{aligned} & x_s^i = \sum_j (x_s^{ij})_k, \quad x_s^i = \sum_j (x_{s-e_k}^{ji})_k \quad (k \in \{1,2,3\}) \\ & x_s \in \Delta, \qquad x_s^{ij} \geq 0. \end{aligned} \tag{2}$$

Together they form a nonlinear and non-smooth convex program and can be minimized e.g. by proximal splitting methods. We use the primal-dual approach [1], and introduce Lagrange multipliers $\lambda_{st\to s}^{i,k}$ for the constraints $x_s^i = \sum_j x_{s,k}^{ij}$, $\lambda_{ts\to s}^{i,k}$ for $x_s^i = \sum_j x_{s-e_k,k}^{ij}$, and ν_s for $\sum_i x_s^i = 1$. We further partially dualize ψ^{ij} via

$$\psi^{ij}(x_s^{ij} - x_s^{ji}) = \max_{\mu_s^{ij}\in W_{\psi^{ij}}} \left(\mu_s^{ij}\right)^T \left(x_s^{ij} - x_s^{ji}\right),$$

thereby introducing additional dual variables μ_s^{ij}. Overall, the underlying saddle-point formulation reads as

$$\begin{aligned} E_{\text{S-P}}(x;\nu,\mu,\lambda) = & \sum_s \left(\sum\nolimits_i \rho_s^i x_s^i + \nu_s\left(\sum\nolimits_i x_s^i - 1\right)\right) \\ & + \sum_{s,i,j:i<j}\left(\left(\mu_s^{ij}\right)^T\left(x_s^{ij} - x_s^{ji}\right) + C^{ij}\left\|x_s^{ij} - x_s^{ji}\right\|_2\right) \\ & + \sum_{s,i}\left(\lambda_{st\to s}^i\right)^T\left(x_s^i\mathbf{1} - \sum\nolimits_j x_s^{ij}\right) \\ & + \sum_{s,i}\left(\lambda_{ts\to s}^i\right)^T\left(x_s^i\mathbf{1} - \sum_j \begin{pmatrix} x_{s-e_1,1}^{ij} \\ x_{s-e_2,2}^{ij} \\ x_{s-e_3,3}^{ij}\end{pmatrix}\right), \end{aligned} \tag{3}$$

subject to $x_s^i \geq 0$, $x_s^{ij} \geq 0$ and $\mu_s^{ij} \in \mathcal{W}_{\psi^{ij}}$. $\mathbf{1}$ denotes the vector $(1,1,1)^T$. The updates of the primal and dual variables are straightforward: gradient steps are followed either by projections to the respective feasible domain ($x_s^i \geq 0$, $\mu_s^{ij} \in \mathcal{W}_{\psi^{ij}}$) or the following proximity step,

$$\begin{aligned} \arg\min_{x_s^{ij},x_s^{ji}} & \frac{1}{2\tau}\left\|x_s^{ij}\right\|_2^2 + \frac{1}{2\tau}\left\|x_s^{ji}\right\|_2^2 \\ & + C^{ij}\left\|x_s^{ij} - x_s^{ji}\right\|_2 + \imath\{x_s^{ij} \geq 0, x_s^{ji} \geq 0\}. \end{aligned}$$

Instead of solving this proximity step, we slightly modify the objective $E_{\text{S-P}}$ as follows: since w.l.o.g. some minimizer of E_{discrete} (Eq. 1) will satisfy complementarity of x_s^{ij} and x_s^{ji} (i.e. $(x_s^{ij})^T x_s^{ji} = 0$, see [3] for a detailed explanation), we may replace $C^{ij}\|x_s^{ij} - x_s^{ji}\|_2$ in $E_{\text{S-P}}$ with

$$C^{ij} \left\| \begin{pmatrix} x_s^{ij} \\ x_s^{ji} \end{pmatrix} \right\|_2,$$

leading to a much simpler subproblem

$$\begin{aligned} \arg\min_{x_s^{ij}, x_s^{ji}} & \frac{1}{2\tau}\|x_s^{ij}\|_2^2 + \frac{1}{2\tau}\|x_s^{ji}\|_2^2 \\ & + C^{ij} \left\| \begin{pmatrix} x_s^{ij} \\ x_s^{ji} \end{pmatrix} \right\|_2 + \imath\{x_s^{ij} \geq 0, x_s^{ji} \geq 0\}, \end{aligned}$$

which corresponds essentially to a shrinkage step in $\mathbb{R}^6$.

References

[1] A. Chambolle and T. Pock. A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging. *J. Math. Imag. Vision*, pages 1–26, 2010.

[2] J. Friedman, T. Hastie, and R. Tibshirani. Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28:2000, 1998.

[3] C. Zach, C. Häne, and M. Pollefeys. What is optimized in tight convex relaxations for multi-label problems? In *Proc. CVPR*, 2012.