

Proximity Priors for Variational Semantic Segmentation and Recognition

Julia Bergbauer¹, Claudia Nieuwenhuis², Mohamed Souiai¹ and Daniel Cremers¹

¹Technical University of Munich, Germany* ²UC Berkeley, ICSI, USA

Abstract

In this paper, we introduce the concept of proximity priors into semantic segmentation in order to discourage the presence of certain object classes (such as 'sheep' and 'wolf') 'in the vicinity' of each other. 'Vicinity' encompasses spatial distance as well as specific spatial directions simultaneously, e.g. 'plates' are found directly above 'tables', but do not fly over them. In this sense, our approach generalizes the co-occurrence prior by Ladicky et al. [3], which does not incorporate spatial information at all, and the non-metric label distance prior by Strelakovsky et al. [11], which only takes directly neighboring pixels into account and often hallucinates ghost regions. We formulate a convex energy minimization problem with an exact relaxation, which can be globally optimized. Results on the MSRC benchmark show that the proposed approach reduces the number of mislabeled objects compared to previous co-occurrence approaches.

1. Introduction

Image segmentation is an essential component in image content analysis and one of the most investigated problems in computer vision. The goal is to partition the image plane into 'meaningful' non-overlapping regions. Especially for complex real-world images, however, the definition of meaningful depends on the application or the user's intention. Typically, the desired segmentation consists of one region for each separate object or structure of the scene. Due to strongly varying texture and color models within and between different object classes, the segmentation task is very complex and requires additional prior information. For example animals such as horses, cows and sheep have similar color models and similarly textured fur. Since many segmentation algorithms only consider local color or texture information to assign each pixel to an object class, they often generate incorrect segmentations, where e.g. part of the sheep is assigned the label 'cow' as shown in Figure 1b).

*This work was supported by the ERC Starting Grant 'ConvexVision' and the German Academic Exchange Service (DAAD).

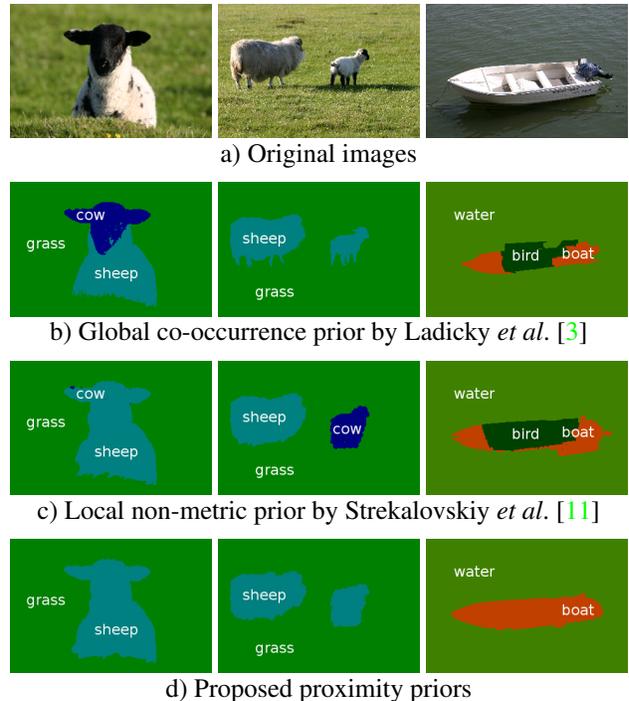


Figure 1: **Proximity priors** discourage the simultaneous occurrence of label pairs within specific directions and distances. Hence, they extend both global [3] co-occurrence priors, which altogether disregard spatial information, and local [11] co-occurrence priors, which only consider directly adjacent pixels as close and often create ghost regions (see Figure 4).

For humans the task of recognizing objects strongly relies on their context and inter-relations with other objects. Therefore, we introduce high-level proximity priors. The key idea is to encourage or discourage the simultaneous appearance of objects within a specified range (distance and direction). Respective penalties for the proximity of various label pairs (encourage 'vases' directly above the 'table' but not further above or below the 'table') can be learned statistically from a set of segmented images. Figure 1 shows three examples where previous co-occurrence priors fail but

proximity priors correctly propagate co-occurrence information yielding the correct segmentation result.

The specific challenge we face in this paper is to find an efficient and convex optimization approach for multi-label segmentation with proximity priors. Improved results based on proximity priors in comparison to related co-occurrence based approaches are shown in Figures 1, 4, 5.

1.1. Related Work

There have been several previous approaches on the integration of co-occurrence priors into semantic segmentation. The most closely related approaches are the global co-occurrence prior by Ladicky *et al.* [3] and the non-metric distance prior by Strekalovskiy *et al.* [11], which can be understood as a local co-occurrence prior.

Ladicky *et al.* [3] globally penalize label sets which occur together in the same image. Yet, this prior is entirely indifferent about *where* in the image respective labels emerge. Moreover, the penalty proposed in [3] is independent of the size of the labeled regions. As a consequence, if more pixels vote for a certain label then they may easily overrule penalties imposed by the co-occurrence term – leading to the segmentations in Figure 1b) with large adjacent regions despite large co-occurrence cost for ‘sheep’ and ‘cow’.

In contrast, Strekalovskiy *et al.* [11] introduced a local co-occurrence prior, which operates only on directly neighboring pixels. The authors formulate a variational approach, which allows for the introduction of non-metric label distances in order to handle learned arbitrary co-occurrence penalties, which often violate the triangle inequality. While labels ‘wolf’ and ‘grass’, for example, are common within an image and labels ‘sheep’ and ‘grass’ as well, sheep are rarely found next to wolves. The drawback of this approach is that the algorithm can avoid costly label transitions simply by introducing infinitesimal ‘ghost labels’ – see Figure 4. Furthermore, due to the strong locality the prior allows for regions to appear close to each other despite high co-occurrence penalties (see the labels ‘sheep’ and ‘cow’ in Figure 1c).

Considering more complex spatial label relationships will avoid ghost labels due to stronger penalization and will allow to propagate the co-occurrence penalty to more distant pixels of the second sheep. Therefore, we generalize these priors to a prior for arbitrary relative spatial relations. Figure 1d) shows examples where proximity priors successfully propagated the co-occurrence penalty to neighboring objects.

In the context of learning, relative spatial label distances have been successfully applied in [1, 2, 9].

1.2. Contributions

In this paper, we propose *proximity priors* for variational semantic segmentation and recognition. Specifically, we

make the following contributions:

- We integrate learned spatial relationships between different objects into a variational multi-label segmentation approach.
- We generalize global co-occurrence priors [3] and local co-occurrence priors [11] to co-occurrence priors with arbitrary spatial relationships.
- We give a convex relaxation which can be solved with fast primal-dual algorithms [8] in parallel on graphics hardware (GPUs).
- We avoid the emergence of artificial ‘ghost labels’.
- We do not rely on prior superpixel partitions but directly work on the pixel level.

2. Variational Multi-Label Segmentation

Let $I : \Omega \rightarrow \mathbb{R}^d$ denote the input image defined on the image domain $\Omega \subset \mathbb{R}^2$. The general multi-label image segmentation problem with $n \geq 1$ labels consists of the partitioning of the image domain Ω into n regions $\{\Omega_1, \dots, \Omega_n\}$. This task can be solved by computing binary labeling functions $u_i : \Omega \rightarrow \{0, 1\}$ in the space of functions of bounded variation (BV) such that $\Omega_i = \{x \mid u_i(x) = 1\}$. We compute a segmentation of the image by minimizing the following energy [13] (see [6] for a detailed survey and code)

$$E(\Omega_1, \dots, \Omega_n) = \frac{\lambda}{2} \sum_{i=1}^n \text{Per}_g(\Omega_i) + \sum_{i=1}^n \int_{\Omega_i} f_i(x) dx. \quad (1)$$

For comparability, we use the same appearance model f_i as in [3, 11]. $\text{Per}_g(\Omega_i)$ denotes the perimeter of each set Ω_i , which is minimized in order to favor segments of shorter boundary. These boundaries are measured with either an edge-dependent or an Euclidean metric defined by the non-negative function $g : \Omega \rightarrow \mathbb{R}^+$. For example,

$$g(x) = \exp\left(-\frac{|\nabla I(x)|^2}{2\sigma^2}\right), \quad \sigma^2 = \frac{1}{|\Omega|} \int_{\Omega} |\nabla I(x)|^2 dx$$

favors the coincidence of object and image edges.

To rewrite the perimeter of the regions in terms of the indicator functions we make use of the total variation:

$$\text{Per}_g(\Omega_i) = \int_{\Omega} g(x) |Du_i| = \sup_{\xi_i : |\xi_i(x)| \leq g(x)} - \int_{\Omega} u_i \text{div } \xi_i dx.$$

Since the binary functions u_i are not differentiable Du_i denotes their distributional derivative. Furthermore, $\xi_i \in \mathcal{C}_c^1(\Omega; \mathbb{R}^2)$ are the dual variables and \mathcal{C}_c^1 denotes the space of smooth functions with compact support. We can

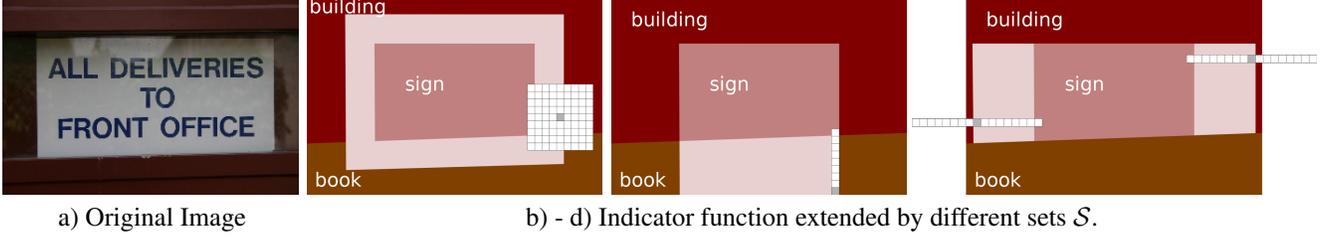


Figure 2: **Impact of \mathcal{S}** . Different sets \mathcal{S} in (3) convey different proximity priors. b) Symmetric sets \mathcal{S} only consider object distances, but are indifferent to directional relations. c) If \mathcal{S} is chosen as a vertical line centered at the bottom, the indicator function of the region 'sign' is extended to the bottom of the object, *e.g.* penalizing 'book' appearing closely below 'sign'. d) Horizontal lines penalize labels to the left and right.

rewrite the energy in (1) in terms of the indicator functions $u_i : \Omega \rightarrow \{0, 1\}$ [6, 13]:

$$E(u_1, \dots, u_n) = \sup_{\xi \in \mathcal{K}} \sum_{i=1}^n \int_{\Omega} (f_i - \text{div } \xi_i) u_i dx, \quad (2)$$

where $\mathcal{K} = \left\{ \xi \in \mathcal{C}_c^1(\Omega; \mathbb{R}^{2 \times n}) \mid |\xi_i(x)| \leq \frac{\lambda g(x)}{2} \right\}$.

2.1. The Novel Proximity Prior

To introduce the proximity prior into the optimization problem in (2), we define the proximity matrix $A \in \mathbb{R}_{\geq 0}^{n \times n}$. Each entry $A(i, j)$, $i \neq j$ indicates the penalty for the occurrence of label j in the proximity of label i , which we denote by $i \sim j$. For $i = j$ we set $A(i, i) := 0$. The penalties can be computed from co-occurrence probabilities of training segmentations, *e.g.* by $A(i, j) = -\log P(i \sim j)$. An example for a learned proximity matrix A is illustrated in Figure 3.

To compute the proximity of two labels, we first introduce the notion of an extended indicator function u_i denoted by $d_i : \Omega \rightarrow \{0, 1\}$, which 'enlarges' the indicator function



Figure 3: **Proximity matrix**. Learned penalty matrix for the MSRC benchmark (objects are color coded corresponding to benchmark convention in first row and column). The lighter the color the more likely is the occurrence of the corresponding labels within the relative spatial context, and the lower is the corresponding penalty.

in a specific direction and distance (see Figure 2):

$$d_i(x) = \sup_{y \in \Omega} u_i(y) + s(x - y) = \sup_{z \in \mathcal{S}} u_i(x + z), \quad (3)$$

where $s(x) = \begin{cases} 0, & x \in \mathcal{S}, \\ -\infty, & \text{otherwise.} \end{cases}$

The set $\mathcal{S} \subseteq \Omega$ determines the type of geometric spatial relationship we want to penalize, *i.e.* distance and direction, for example 'less than 20 pixels above'. Symmetric sets of specific sizes consider the proximity of two labels without preference of a specific direction. If \mathcal{S} is for example a line we can penalize the proximity of specific labels in specific directions, *e.g.* the occurrence of a book below a sign (compare Figure 2c). The larger \mathcal{S} the more pixels are considered adjacent to x . Runtime can be minimized here by choosing sparse sets \mathcal{S} .

To detect if two regions i and j are close to each other, we compute the overlap of the extended indicator function d_i and the indicator function u_j . For each two regions i and j we can now penalize their proximity by means of the following energy term:

$$E_{prox}(u) = \sum_{1 \leq i < j \leq n} A(i, j) \int_{\Omega} d_i(x) u_j(x) dx. \quad (4)$$

2.2. A Convex Relaxation

In the following we will propose a convex relaxation of the segmentation problem (2) combined with the proposed proximity prior in (4). To obtain a convex optimization problem, we require convex functions over convex domains.

Relaxation of the Binary Functions u_i The general multi-labeling problem is not convex due to the binary region indicator functions $u_i : \Omega \rightarrow \{0, 1\}$. To obtain a convex problem where each pixel is assigned to exactly one label, optimization is carried out over the convex set

$$\mathcal{U} = \left\{ u \in BV(\Omega; [0, 1]^n) \mid \sum_{j=1}^n u_j(x) = 1 \quad \forall x \in \Omega \right\}.$$

$$\begin{aligned}
\min_{\substack{u \in \mathcal{U} \\ d \in \mathcal{D} \\ \alpha \in \mathcal{A}}} \max_{\substack{\xi \in \mathcal{K} \\ \beta \in \mathcal{B} \\ q \in \mathcal{Q}}} & \sum_{i=1}^n \left\{ \int_{\Omega} (f_i - \operatorname{div} \xi_i) u_i dx + \sum_{z \in \mathcal{S}} \int_{\Omega} \beta_{i_z}(x) (d_i(x) - u_i(x+z)) dx \right. \\
& + \sum_{j=i+1}^n \int_{\Omega} q_{ij}^1 (1 - d_i) + q_{ij}^2 d_i + q_{ij}^3 (1 - u_j) + q_{ij}^4 u_j \\
& \left. + \alpha_{ij}^1 (q_{ij}^1 + q_{ij}^3) + \alpha_{ij}^2 (q_{ij}^1 + q_{ij}^4) + \alpha_{ij}^3 (q_{ij}^2 + q_{ij}^3) + \alpha_{ij}^4 (q_{ij}^2 + q_{ij}^4 - A(i, j)) dx \right\}. \tag{5}
\end{aligned}$$

Relaxation of the Dilation Constraints The dilation constraints in (3) are relaxed to

$$d_i(x) \geq u_i(x+z) \quad \forall x \in \Omega, z \in \mathcal{S}. \tag{6}$$

By simultaneously minimizing over the functions d_i we can assure that at the optimum d_i fulfills the constraints in (3) exactly. The inequality (6) can easily be included in the segmentation energy by introducing a set of Lagrange multipliers β_{i_z} and adding the following energy term:

$$\min_{d \in \mathcal{D}} \max_{\beta \in \mathcal{B}} \sum_{i=1}^n \sum_{z \in \mathcal{S}} \int_{\Omega} \beta_{i_z}(x) (d_i(x) - u_i(x+z)) dx, \tag{7}$$

$$\mathcal{B} = \{ \beta_{i_z} \mid \beta_{i_z} : \Omega \rightarrow [-\infty, 0] \quad \forall z \in \mathcal{S}, i = 1, \dots, n \},$$

$$\mathcal{D} = BV(\Omega; [0, 1]^n).$$

Relaxation of the Product in (4) The product of the dilation d_i and the indicator function u_j is not convex. A convex, tight relaxation of such energy terms was given by Strekalovskiy *et al.* [10]. To this end, we introduce additional dual variables q_{ij} and Lagrange multipliers α_{ij} :

$$\mathcal{Q} = \{ q_{ij} \mid q_{ij} : \Omega \rightarrow \mathbb{R}^4, 1 \leq i < j \leq n \}, \tag{8}$$

$$\mathcal{A} = \{ \alpha_{ij} \mid \alpha_{ij} : \Omega \rightarrow [-\infty, 0]^4, 1 \leq i < j \leq n \}.$$

Resulting Optimization Problem After carrying out these relaxations we finally obtain the convex energy minimization problem in (5).

The projections onto the respective convex sets of ξ , d , β and α are done by simple clipping while that of the primal variable u is a projection onto the simplex in \mathbb{R}^n [5].

3. Implementation

In order to find the globally optimal solution to this relaxed convex optimization problem, we employ the primal-dual algorithm published in [8]. Optimization is done by alternating a gradient descent with respect to the functions u , d and α and a gradient ascent for the dual variables ξ , β and q interlaced with an over-relaxation step on the primal variables. The step sizes are chosen optimally according to [7]. We stopped the iterations when the average update

of the indicator function $u(x)$ per pixel was less than 10^{-5} . By allowing the primal variables u_i to take on intermediate values between 0 and 1 we may end up with non-binary solutions. In order to obtain a binary solution to the original optimization problem, we assign each pixel x to the label L with maximum value after optimizing the relaxed problem:

$$L(x) = \arg \max_i \{ u_i(x) \}, x \in \Omega. \tag{9}$$

We observed that the computed relaxed solutions u are binary almost everywhere.

Due to the inherent parallel structure of the optimization algorithm [8] the approach can be easily parallelized and implemented on graphics hardware. We used a parallel CUDA implementation on an NVIDIA GTX 680 GPU. To reduce the runtime of the approach we randomly subsampled only very few entries in \mathcal{S} and neglected the others yielding equivalent results in around 180 seconds on average (note that we do not work on super pixels). We can conclude that already very sparse sets \mathcal{S} containing around ten entries yield results very similar to the full set \mathcal{S} .

4. Experiments and Results

We have defined proximity priors within a variational multi-label approach in order to integrate spatial relations between object labels. One of the major advantages of the proposed algorithm is that we can utilize sets \mathcal{S} of different sizes and shapes which allow us to take into account larger neighborhoods of pixels in specific directions and to prevent 'ghost labels'. In the following we will show results on the MSRC database and compare our segmentations to state-of-the-art approaches for semantic labeling and co-occurrence priors.

4.1. Preventing Ghost Labels

'Ghost labels' denote thin artificial regions which are easily introduced if label distances are learned from training data, see for example [11].

If the distance function does not obey the triangle inequality 'ghost labels' can appear. They reduce costs of direct label transitions by taking a 'detour' over a third, unrelated but less expensive label. Examples are given in Figure 4b) with a closeup in Figure 4c). The segmentation result obtained by [11] *e.g.* contains very thin 'boat' regions at the

	Accur. per pixel	Accur. per class	Building	Grass	Tree	Cow	Sheep	Sky	Plane	Water	Face	Car	Bicycle	Flower	Sign	Bird	Book	Chair	Road	Cat	Dog	Body	Boat
Gould <i>et al.</i> [1] CRF + rel. loc.	76.5	64.38	72	95	81	66	71	93	74	70	70	69	72	68	55	23	82	40	77	60	50	50	14
Ladicky <i>et al.</i> [3] co-oc.	80	67.76	77	96	80	69	82	98	69	82	79	75	75	81	85	35	76	17	89	25	61	50	22
Lucchi <i>et al.</i> [4], DPG local	75	68.62	54	88	83	79	82	95	87	70	85	81	97	69	72	27	88	46	60	74	27	49	28
Lucchi <i>et al.</i> [4], DPG loc.+glob.	80	74.62	65	87	87	84	75	93	94	78	83	72	93	86	70	50	93	80	86	78	28	58	27
Vezhnevets <i>et al.</i> [12], weak sup.	67	66.52	12	83	70	81	93	84	91	55	97	87	92	82	69	51	61	59	66	53	44	9	58
Vezhnevets <i>et al.</i> [12], full sup.	72	71.71	21	93	77	86	93	96	92	61	79	89	89	89	68	50	74	54	76	68	47	49	55
Stekalovskiy <i>et al.</i> [11]	84.85	77.52	70	97	92	89	85	96	81	83	90	82	92	83	66	45	92	63	86	80	51	73	32
Proposed Proximity Priors	84.97	78.19	69	97	92	87	87	97	87	82	91	83	94	84	62	44	93	67	86	83	57	74	26
Ladicky <i>et al.</i> [3] co-oc. + hier.	87	76.76	82	95	88	73	88	100	83	92	88	87	88	96	96	27	85	37	93	49	80	65	20

Table 1: **MSRC benchmark scores.** We compare the segmentation accuracy to state-of-the-art segmentation algorithms with co-occurrence priors on the MSRC benchmark. The approach by Ladicky *et al.* in the last row is added for completeness but is not comparable since it includes hierarchical label priors and uses potentials of the highest order $|\Omega|$ instead of order two as in our approach.

edge of the 'grass' label, because the transition between the labels 'water' and 'boat' and 'boat' and 'grass' is in sum less costly than the direct transition between 'water' and 'grass'.

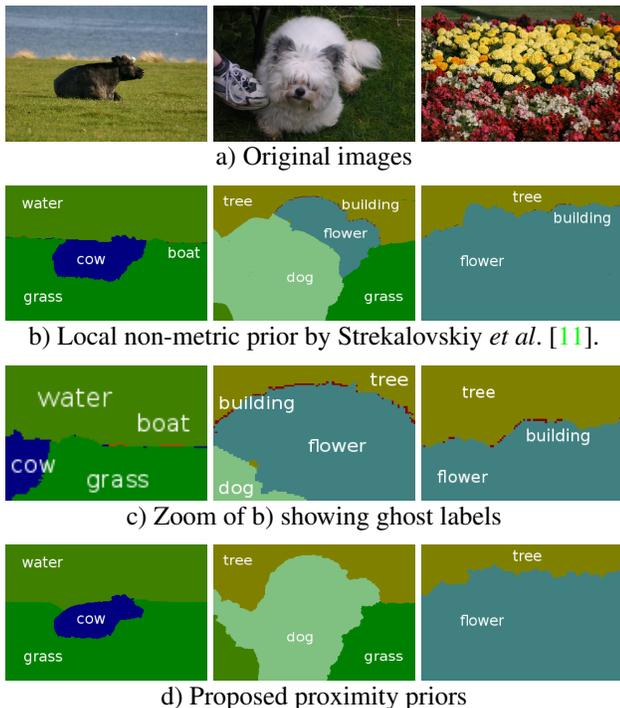


Figure 4: **Proximity priors prevent ghost labels.** If the transition of two labels is cheaper via a third label artificial labels will be introduced as shown in b) and as closeup in c). The proposed proximity priors consider regions with more than one-pixel distance still as adjacent and thus avoid ghost labels.

Proximity priors prevent such ghost labels by considering more than a single pixel wide margin as close to the object, see for example Figure 4d).

4.2. MSRC Segmentation Benchmark

To evaluate the proposed segmentation algorithm we apply it to the task of object segmentation and recognition on the MSRC benchmark. This benchmark comprises 591 images which contain 21 different labels such as 'cow', 'book', 'building' or 'grass'. For the benchmark experiments we chose a symmetric set \mathcal{S} of size 9×9 and set $\lambda = 0.3$.

The penalty matrix A defined in Section 2.1 is learned from training data based on the relative frequencies of label occurrences within the local range defined by \mathcal{S} . For a symmetric set \mathcal{S} of size 9×9 we obtain the penalty matrix in Figure 3.

To evaluate the segmentation accuracy of the proposed method, in Table 1 we compare the benchmark scores of our method to the approaches by Gould *et al.* [1] with relative location priors, Ladicky *et al.* [3] with co-occurrence prior (with and without hierarchical prior), Lucchi *et al.* [4] for the data pairwise global and local models, Vezhnevets *et al.* [12] for the weakly and fully supervised approach and to Stekalovskiy *et al.* [11] with the non-metric distance functions for multi-label problems. The scores denote the average accuracy on the benchmark given as $\frac{\text{True Positives} \cdot 100}{\text{True Positives} + \text{False Negatives}}$ per pixel and per class. The results indicate that we outperform the other co-occurrence based methods in average class and pixel accuracy.

Note that the high score of the approach by Stekalovskiy *et al.* [11] does not reflect the ghost label problem since these regions contain only very few pixels. However, the introduction of entirely unrelated objects, albeit small ones, is

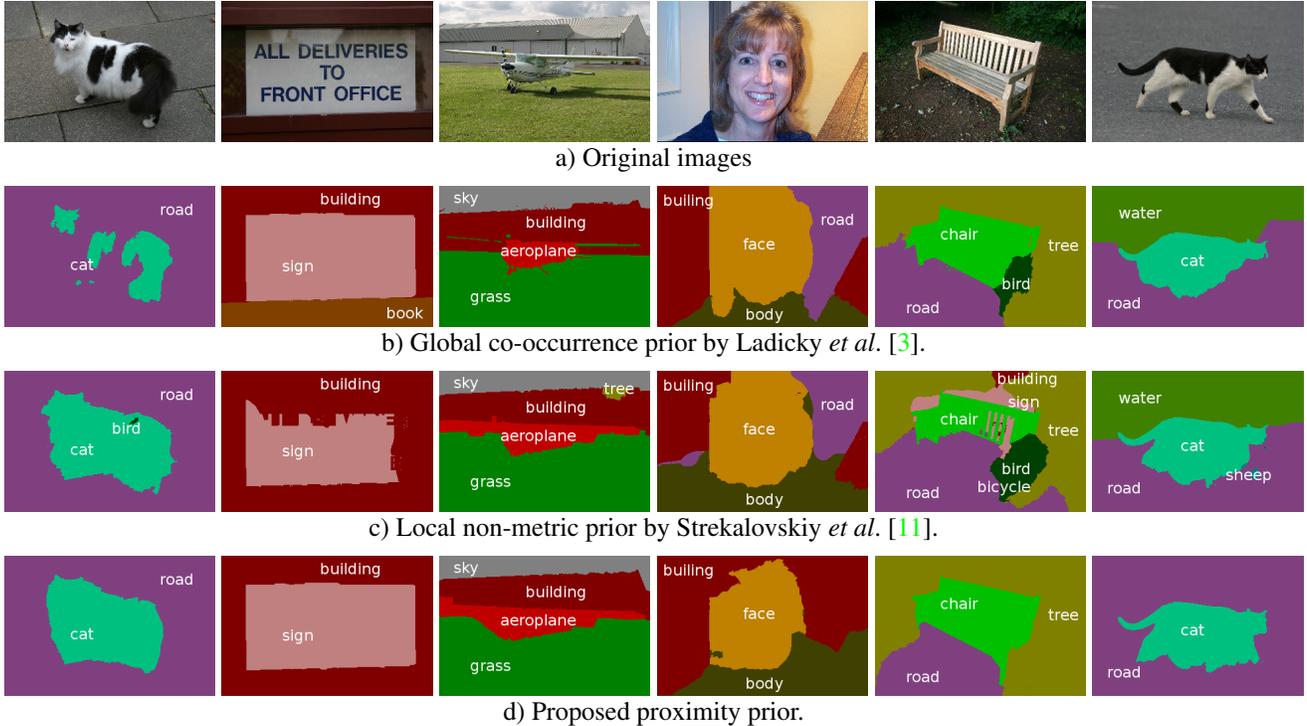


Figure 5: **Improved results on the MSRC benchmark.** Proximity Priors capture richer semantic information on spatial object inter-relations such as distances, direction and relative location than previous approaches such as global co-occurrence [3] or local co-occurrence [11].

often problematic for applications.

The benchmark results in general suggest rather small improvements for the integration of geometric spatial priors. This is somewhat surprising since the images show strong improvements and the prior corresponds to typical human thinking. As already mentioned by Lucchi *et al.* [4] who stated similar findings this is probably due to the rather crude ground truth of the benchmark with large unlabeled regions especially at object boundaries. These regions are not counted in the score, but nevertheless leave a lot of room for misclassification or improvements. Therefore, we think that the benchmark score should not be overstressed here.

Qualitative comparisons with the two best scoring of the above mentioned methods by Ladicky *et al.* [3] with hierarchical prior and by Strelakovsky *et al.* [11] on the MSRC database are given in Figures 1 and 5. The results show that the proposed method reduces the number of mislabeled objects. For example, our approach is the only one which correctly detects the boat in Figure 1 without assigning part of it to the label 'bird'. Another example is the head of the sheep in the first column of Figure 1 which is correctly labeled without any 'cow' pixels. The result of the cat in the first column of Figure 5 shows that we can avoid problems which appear due to prior superpixel segmentations.

4.3. Direction Dependent Proximity

Some object pairs only appear in specific spatial constellations, for example cars do not appear above water or books on top of buildings. Such relations can be encoded by applying directional sets \mathcal{S} , *e.g.* horizontal or vertical lines (compare Figure 2). The corresponding penalties can either be defined or learned from training data.

For the portrait of the woman in Figure 5 we used a learning based approach for a horizontal set \mathcal{S} such as the one shown in Figure 2d). We derive the penalties A from the relative frequencies of objects appearing up to 40 pixels left and right of the label 'face' in the training images. For the bench and cat images (rightmost in Figure 5) we used a vertical set \mathcal{S} centered at the bottom/top (compare Figure 2c) together with a distance of 20 to penalize the label 'bird' below 'chair' and the label 'water' above 'street'.

5. Conclusion

We introduced proximity priors for semantic segmentation and recognition within a variational multi-label framework. Instead of introducing co-occurrence probabilities of label combinations, proximity priors define likelihoods for specific geometric spatial relationships of label pairs, *i.e.* their direction and distance. In this way, proximity priors

generalize both global co-occurrence priors [3], which take into account all labels irrespective of their spatial location, and local co-occurrence priors [11] which are only imposed on directly adjacent pixels.

The label cost penalty is proportional to the size of the labeled regions and also affects object labels at larger spatial distances. In addition, the proposed approach does not require the computation of superpixels and prevents the emergence of one pixel wide 'ghost labels'.

References

- [1] S. Gould, J. Rodgers, D. Cohen, G. Elidan, and D. Koller. Multi-class segmentation with relative location prior. *Int. J. of Computer Vision*, 2008. 2, 5
- [2] P. Kotschieder, P. Kohli, J. Shotton, and A. Criminisi. Geof: Geodesic forests for learning coupled predictors. In *Int. Conf. on Comp. Vision and Pattern Recognition*, 2013. 2
- [3] L. Ladicky, C. Russell, P. Kohli, and P. Torr. Graph cut based inference with co-occurrence statistics. In *Proceedings of ECCV*, 2010. 1, 2, 5, 6, 7
- [4] A. Lucchi, Y. Li, X. Boix, K. Smith, and P. Fua. Are spatial and global constraints really necessary for segmentation? In *IEEE Int. Conf. on Computer Vision*, 2011. 5, 6
- [5] C. Michelot. A finite algorithm for finding the projection of a point onto the canonical simplex of \mathbb{R}^n . *Journal of Optimization Theory and Applications*, 50(1):195–200, July 1986. 4
- [6] C. Nieuwenhuis, E. Toeppe, and D. Cremers. A survey and comparison of discrete and continuous multi-label optimization approaches for the potts model. *Int. J. of Computer Vision*, 2013. 2, 3
- [7] T. Pock and A. Chambolle. Diagonal preconditioning for first order primal-dual algorithms in convex optimization. In *IEEE Int. Conf. on Computer Vision*, 2011. 4
- [8] T. Pock, D. Cremers, H. Bischof, and A. Chambolle. An algorithm for minimizing the Mumford-Shah functional. In *IEEE Int. Conf. on Computer Vision*, Kyoto, 2009. 2, 4
- [9] S. Savarese, J. Winn, and A. Criminisi. Discriminative object class models of appearance and shape by correlatons. In *Int. Conf. on Computer Vision and Pattern Recognition*, 2006. 2
- [10] E. Strelakovsky, B. Goldluecke, and D. Cremers. Tight convex relaxations for vector-valued labeling problems. In *IEEE Int. Conf. on Computer Vision*, 2011. 4
- [11] E. Strelakovsky, C. Nieuwenhuis, and D. Cremers. Non-metric priors for continuous multilabel optimization. In *European Conference on Computer Vision*, 2012. 1, 2, 4, 5, 6, 7
- [12] A. Vezhnevets, V. Ferrari, and J. M. Buhmann. Weakly supervised semantic segmentation with a multi-image model. In *IEEE Int. Conf. on Computer Vision*, 2011. 5
- [13] C. Zach, D. Gallup, J.-M. Frahm, and M. Niethammer. Fast global labeling for real-time stereo using multiple plane sweeps. In *Vision, Modeling and Visualization Workshop VMV*, 2008. 2, 3