Max-margin Regularization for Reducing Accidentalness in Chamfer Matching

Angela Eigenstetter*, Pradeep Yarlagadda* and Björn Ommer

Interdisciplinary Center for Scientific Computing, University of Heidelberg, Germany
{aeigenst,pyarlaga,bommer}@iwr.uni-heidelberg.de

Abstract. Standard chamfer matching techniques and their state-of-the-art extensions are utilizing object contours which only measure the mere sum of location and orientation differences of contour pixels. In our approach we are increasing the specificity of the model contour by learning the relative importance of all model points instead of treating them as independent. However, chamfer matching is still prone to accidental matches in dense clutter. To detect such accidental matches we learn the co-occurrence of generic background contours to further eliminate the number of false detections. Since, clutter only interferes with the foreground model contour we learn where to place the background contours with respect to the foreground object boundary. The co-occurrence of foreground model points and background contours are both integrated into a single max-margin framework. Thus our approach combines the advantages of accurately detecting objects or parts via chamfer matching and the robustness of a max-margin learning. Our results on standard benchmark datasets show that our method significantly outperforms current directional chamfer matching, thus redefining the state-of-the-art in this field.

1   Introduction

Chamfer matching is a widely used technique for object detection. Due to its simplicity and efficiency it has been employed in a variety of applications to match whole object boundaries as well as partial object contours. Despite these advantages chamfer matching has a serious drawback when contours are matched in cluttered image regions. Contour matches in cluttered regions have a high accidentalness which can not be distinguished from matches on the actual object. Recent research made some attempts to improve specificity by including orientation information [1, 2] in the distance function. The limitation of these approaches is that the presence of individual model points in a query image is measured independently. As demonstrated by Biederman [3], Attneave [4], and various experiments on illusionary contours, object boundary pixels are not

---

* Both authors contributed equally to this work.

all equally important due to their statistical interdependence. To address this issue we learn the relevance of model points and give higher weight to more important model points. Furthermore we learn a flexible co-placement of generic background contours to reduce the accidentalness of matches in dense background clutter. We integrated these two learning steps into a single max-margin framework based on the state-of-the-art directional chamfer matching approach of Chellappa et al. [2] and evaluate the gain in performance.

2   Related Work

Chamfer matching has been used in a large number of applications in computer vision. It was first introduced by Barrow et al. [5] to match two sets of contour fragments. Since then chamfer matching has been widely applied and has been a successful technique for detecting complete objects or their parts. In [6] hierarchical chamfer matching was suggested where edge points are matched in a coarse-to-fine-manner. Later, chamfer matching was used to build powerful detectors as proposed in [7–9]. Leibe et al. [7] combine local features with global shape cues obtained from chamfer matching to verify and refine hypotheses. In [8], Gavrila and Munder have applied chamfer matching for real-time pedestrian detection and tracking. Lin et al. [9] have proposed a hierarchical part-template matching approach for detection and segmentation which measures shape information in terms of chamfer matching scores.

In [10] Thayananthan et al. have compared shape context [11] and chamfer matching of templates for object detection in cluttered images. They reported that chamfer matching is more robust to clutter than shape context. Nevertheless, false positives in cluttered background were still found to be the major downside of chamfer matching.

More recent research has made attempts to address this problem. Shotton et al. [1] suggested an improved matching scheme called oriented chamfer matching (OCM) that takes into account the orientation mismatch between pixels. In [2] an alternative approach for incorporating edge orientation has been proposed which solves the matching problem in an augmented space. It was shown that the suggested directional chamfer matching (DCM) achieves a superior performance compared to oriented chamfer matching. Another improvement was proposed in [12] where manually selected tuples of contour fragments have been used as normalizers for oriented chamfer matching.

[1, 2] focus on adding orientation information to improve the matching quality of the foreground template. In both approaches, the score for an object hypothesis is obtained by summing over all the template pixels in the distance transform of the query image. However, an object is more than the mere sum of the distance transformation of each template pixel, considering the evidence from [3, 4] which indicates that not all boundary pixels are equally important.

[12] proposed to reduce the chamfer matches in clutter by normalizing template matches with manually combined normalizer contours. These normalizers are placed at the center of the template matches. However, to sufficiently model
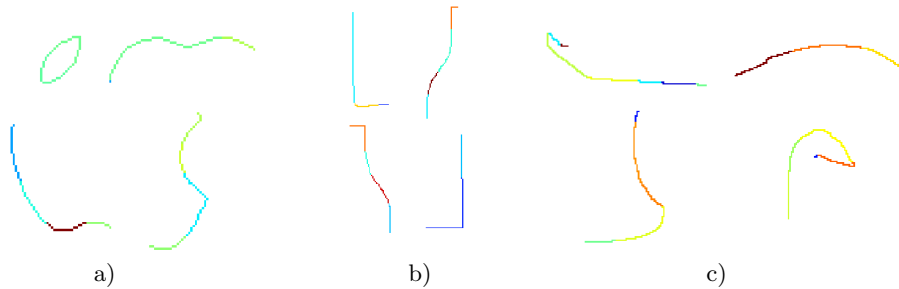
Fig. 1. a)-c) show the relative pixels weights for applelogos, bottles and swans learnt with a linear max-margin classifier. Red indicates high weight and blue low weight.

complex background, it is important to combine simple contours via flexible placement going beyond the manual combinations of normalizers. We measure the accidentalness of a match to clutter by learning the co-placement of background contours dependent on the foreground.

We integrate the i) relative importance of different pixels on a foreground template and ii) accidentalness of a template match to clutter into a single discriminative approach. Our final detection system improves the matching performance of a foreground template while suppressing spurious matches in cluttered background using the proposed background regularization.

3   Reducing Accidentalness in Chamfer Matching

Our approach is based on the publicly available fast directional chamfer matching approach of [2] which was shown to achieve state-of-the-art performance in chamfer-based matching. Let us now briefly review the fast directional chamfer matching [2] and introduce the required notation. Each object is represented by a collection of contours of its different parts. Let $P = \{\mathbf{p}_i\}$ and $Q = \{\mathbf{q}_j\}$ be the pixels of an object part and query edge maps respectively. Let $\phi(\mathbf{p}_i)$ be the edge orientation of the edge point $\mathbf{p}_i$.

For a given location $\mathbf{x}$ of the object part in the query image, directional chamfer matching aims to find the best $\mathbf{q}_j \in Q$ for each $\mathbf{p}_i \in P$ by minimizing the cost $|(\mathbf{p}_i + \mathbf{x}) - \mathbf{q}_j| + \lambda|\phi(\mathbf{p}_i + \mathbf{x}) - \phi(\mathbf{q}_j)|$. λ denotes the weighting factor between location and orientation terms. Thus the directional chamfer distance for placing an object part at location $\mathbf{x}$ is defined as

$$d_{DCM}^{(P,Q)}(\mathbf{x}) = \frac{1}{|P|} \sum_{\mathbf{p}_i \in P} \min_{\mathbf{q}_j \in Q} |(\mathbf{p}_i + \mathbf{x}) - \mathbf{q}_j| + \lambda|\phi(\mathbf{p}_i + \mathbf{x}) - \phi(\mathbf{q}_j)| \qquad (1)$$

3.1   Interdependence of Model Points

Not all the pixels on an object part are equally important for detecting objects. Consider for instance the famous Kanizsa triangle. Provided only contour

fragments around the corners, the whole triangle can be easily recognized. Similarly, Biederman [3] presents perceptual experiments with degraded contours that demonstrate the varying importance of different points on object contours. Another example is Attneave's cat [4], where for instance, points of high curvature are proposed as the most useful features for recognition. However, we want to automatically learn which parts of the model are important, rather than manually encoding a set of rules that define the importance of contour points. While there is related work, for instance, on saliency [13] and interest point detection [14], we seek a formulation that can be directly integrated into chamfer matching. Moreover, the interest points are detected based on each training image separately whereas we seek important points of an object part based on joint consideration of all the training images.

In chamfer matching, matching costs for an object part are obtained by summing over all the object part pixels in the distance transform of the query image as in Eq. (1). Thus, all the pixels are implicitly considered to be equally important when computing the matching costs. To take into account the fact that not all pixels are equally important, we learn discriminative weights for the co-occurrence of individual points of an object part, i.e., of their matching costs

$$t_i^{(P,Q)}(\mathbf{x}) = \min_{\mathbf{q}_j \in Q} |(\mathbf{p}_i + \mathbf{x}) - \mathbf{q}_j| + \lambda|\phi(\mathbf{p}_i + \mathbf{x}) - \phi(\mathbf{q}_j)| \qquad (2)$$

Since adjacent pixels of an object part are statistically dependent, we utilize the line representation for the templates from [2] and learn discriminative weights for each line of the object part. Thus, all the pixels which lie on the same line are assigned the same weight. Let $\bar{t}_l$ denote the matching cost of line l fitted to the object part.

$$\bar{t}_l = \sum_{i \in l} t_i^{(P,Q)}(\mathbf{x}) \qquad (3)$$

The discriminative learning algorithm that discovers the weights for the co-occurrences of lines is described in Sect. 3.3. For examples of relative pixel weights see Fig. 1.

3.2   Background Contours for Modelling Accidentalness

One of the main drawbacks of chamfer matching is the liability to false positive matches in background clutter. Increasing the specificity of the model contour matches by adding orientation information [2, 1] and learning the importance of foreground contour pixels can only partially solve the problems arising from background clutter Fig. 4. Consequently we need to measure the accidentalness of an object part matching in the background clutter. To measure such accidentalness, we introduce a set of simple, generic contour segments (see Fig. 2 a) that typically match equally well to background clutter and the correct part contour. We refer to this set of contours as background contours. Since each single background contour segment has a very low specificity we learn discriminative
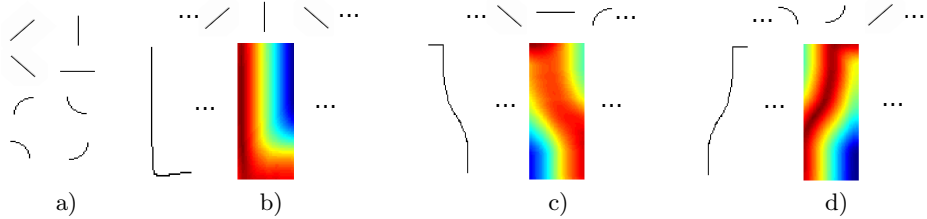
Fig. 2. a) shows a set of simple background contours. These background contours are used to regularize the chamfer response of a part p_i. b)-d) show the masks described in Eq. (4) obtained from placing the background contours at the top relative to the object part contour on the left.

co-occurrence patterns which have very low accidentalness. By going for flexible spatial arrangements of background contours, we avoid manually combining tuples of normalizers consisting of one or two contours to form hand designed complex background templates as in [12]. Furthermore, we measure the amount of clutter only in the neighborhood of model contours, where clutter actually interferes with the matching of the model contour while in [12] background contours are placed at a fixed single location (the center of the model contour). The importance of the second point is illustrated by the following example. Consider a U-shaped object part being matched to a query image. Clutter from the query image that is situated within the U does not interfere with the object part. Only clutter that is close to the contour of the U will have an impact. Thus, [12] miss out on measuring the susceptibility of the model contour to clutter and instead measure clutter simply at the center of the object.

To make sure that background contours T_{bg} are placed on the foreground contour P, where accidental matches typically occur, we create a mask for every combination of a foreground part and a background contour

$$M^{(T_{bg},P)}(\mathbf{x}) = 1 - d_{DCM}^{(T_{bg},P)}(\mathbf{x}) \tag{4}$$

These masks give high weight to regions where the background contour matches well on the part contour and low weight otherwise. Figure 2 shows the resultant masks for three different foreground bottle parts in combination with different background contours.

To describe the background matching costs for a hypothesis in a robust way we build weighted histograms over chamfer matching costs. Let $\bar{\mathbf{x}}$ be one specific placement of the foreground object part P on the query image Q. Furthermore we define $B(\bar{\mathbf{x}})$ to be the bounding box region of P centered at $\bar{\mathbf{x}}$. For each foreground hypothesis we build weighted histograms $h^{(T_{bg},Q)}$ over the directional chamfer matching costs $d_{DCM}^{(T_{bg},Q)}$ in the corresponding bounding box region. The weights introduced in Eq. (4) are used to weight the histogram votes according to their position relative to the foreground object part. Each histogram consists of K bins where $\mathcal{M}_k$ is the range of the kth bin and $k = 1, ..., K$. We define a

histogram bin $h_k^{(T_{bg},Q)}$ as

$$h_k^{(T_{bg},Q)} = \sum_{\substack{\mathbf{x} \in B(\bar{\mathbf{x}}) \\ d_{DCM}^{(T_{bg},Q)}(\mathbf{x}) \in \mathcal{M}_k}} M^{(T_{bg},P)}(\mathbf{x}), \tag{5}$$

for each background contour T_{bg} on a certain position of the foreground object part P in the query image Q.

3.3   Learning Co-occurrences for Foreground and Background

We combine the twin problems of i) modeling the co-occurrence of points on a foreground object part and ii) modeling the accidentalness of a match by means of co-occurrence patterns of background contours into a single discriminative approach. We regularize directional chamfer matching by learning the characteristic co-occurrence of foreground object part pixels and the joint placement of background contours.

   As training data this learning algorithm utilizes the object hypotheses obtained from running the directional chamfer matching code [2] on the training images. A hypothesis j having an overlap greater than 80% with the groundtruth is labeled as positive $y_j = 1$, while a hypothesis with an overlap smaller than 40% is labeled as negative $y_j = -1$. For each object hypothesis we build a feature vector f_j as constructed in Eq. (6).

$$f_j = [\bar{t_1} \, ... \, \bar{t_L} \; h_1 \, ... \, h_K] \tag{6}$$

L denotes the number of line segments fitted to the object part and K denotes the number of background contours.

   Let $\mathcal{K}$ be a kernel such that $\mathcal{K}(f_i, f_j)$ represents the similarity between feature vectors f_i, f_j. Subsequently, we use the radial basis kernel $\mathcal{K}(f_i, f_j) = \exp\left(-\gamma \|f_i - f_j\|^2\right)$. It is common practice in the field of kernel machines, to interpret the kernel $\mathcal{K}(f_i, f_j)$ as a dot product of transformed features $\psi(f_i), \psi(f_j)$. Here ψ represents the mapping of the feature vector into a higher dimensional space. Due to the seminal 'kernel trick' [15] it is sufficient to define the kernel $\mathcal{K}$ without explicitly representing the mapping ψ. We then seek weights w to be applied on $\psi(f_i)$ so that the margin between positive and negative hypotheses in the transformed space is maximized. Since we employ the non-linear RBF kernel, the resulting classifier will learn non-linear relationships between the features and model the joint co-occurrences of foreground and background contours. We need to optimize the following max-margin classification problem to learn the weights w.

$$\min_{w,b,\xi} \frac{1}{2}\|w\|_2^2 + C \sum_{j=1}^{N} \xi_j \tag{7}$$

subject to :

$$y_j(w^T \psi(f_j) + b) \geq 1 - \xi_j \qquad \wedge \qquad \xi_j \geq 0$$

a)   b)   c)
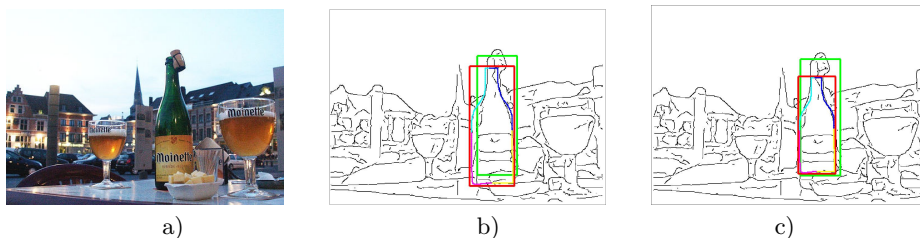
Fig. 3. Learning discriminative weights for the co-occurrences of $t_i^{(P,Q)}(\mathbf{x})$ improves the matching score of shape template as shown in the example here. The original image, the result obtained from directional chamfer matching and the result obtained from foreground reweighting are shown in panels a,b and c respectively. The groundtruth bounding box is shown in green and the top scoring object hypotheses are shown in red.

where N is the number of training samples, b is the offset, C is the penalty and ξ_j are slack variables allowing for margin violations. Commonly Eq. (7) is converted into its dual form and solved for the dual SVM parameters, the support vectors S_i, their coefficients α_i and the offset b.

3.4   Object Detection using Regularized Chamfer Matching

In the previous section we have described how the relevance of model points and the accidentalness measured using background contours can be jointly learned. Let us now utilize the combined model of foreground relevance and background accidentalness from Eq. (7) to improve upon the directional chamfer matching cost function Eq. (1). This improved, regularized chamfer distance $d_{RDCM}^{(P,Q)}(\mathbf{x})$ again measures the distortion cost of an object part f_j. Let the j-th object hypothesis f_j, which is described by the feature vector from Eq. (6), be the placement of object part P at location $\mathbf{x}$ in the query image Q. Since a non-linear radial basis kernel is employed, the regularized chamfer distance is obtained using the dual SVM parameters, obtained by solving the SVM optimization problem from Eq. (7) in its dual form,

$$d_{RDCM}^{(P,Q)}(\mathbf{x}) = 1 - \left(\sum_i \alpha_i \mathcal{K}(f_j, S_i) + b \right). \tag{8}$$

Each object part matched to a query image casts a vote with weight d_{RDCM} as computed in Eq. (8) for different placements of the part in the query image. The votes from various parts are collected in a Hough accumulator and non-max suppression is performed to obtain final candidate hypotheses for objects.

4   Experimental Evaluation

To demonstrate the utility of the proposed discriminative chamfer regularization, we evaluate our approach on benchmark datasets for chamfer matching. Since we

integrated our regularization into the publicly available code of [2], the results reported in [2] have been used as the baseline. To demonstrate the advantage of our regularization over learning the normalization for chamfer distances [12], a comparison is made with the results documented in [12]. We also compare with a sophisticated learning and inference approach applied on object contours [16]. Furthermore, an analysis of the running time overhead caused by discriminative chamfer regularization compared to the running time of the chamfer matching approach of [2] is presented.

To extract the edge maps from input RGB images, we utilize the probabilistic boundary detector of [17]. The weights w in Eq. (7) are learnt using the support vector machine implementation of [18]. To measure the performance of our detection system, we employ the standard PASCAL overlap criterion according to which a detection is correct if the ratio of intersection and overlap between groundtruth and the detection is larger than 50%.

The individual contributions of the proposed foreground and background regularization are presented in Sect. 4.3. We analyze the contribution of reweighting the foreground template pixels as in Eq. (2) and then the performance obtained by the combined foreground and background regularization as in Eq. (7). The gain obtained by these contributions is compared with the baseline obtained by [2]. Sect. 4.4 compares the proposed regularization with other state-of-the-art extensions to chamfer matching such as [12, 16].

4.1 Datasets

For our experimental evaluation, we use the TUD Pedestrians, TUD Cows and the ETHZ Shape datasets. These are the benchmark datasets for chamfer matching and approaches such as [1, 2, 12] report their results on one or more of these datasets. The TUD Pedestrian dataset provides two training sets with 210 and 400 side-view pedestrians. Following the protocol of [12], we use 400 images for training, 5 masks from the training images as model shape templates and 250 images for testing. The Cow dataset from the PASCAL Object Recognition Database Collection [19] consists of 111 images in which cows appear with quite different articulation. The protocol from [12] is used to divide the dataset into training and testing sets. Five masks from the training images are obtained as the shape templates. For the ETHZ shape dataset, according to the standard protocol one hand-drawn example provided along with each category is used as a shape template. For the object categories applelogos, bottles and swans,the template is decomposed into four parts while for the categories giraffes and mugs the full template was utilized.

4.2 Running time

To obtain the initial matches for the templates, we run the publicly available directional chamfer matching code of [2] using the default parameters for all the datasets. In our experimental evaluations, we have observed that computing the distance transformation of a query image for each angular quantization is the

Table 1. Comparison of **average precision** for three datasets namely, TUD Pedestrians, Cows and the ETHZ Shape classes. We compare DCM [2] which constitutes the basis of our approach with the extension from Sec. 3.1 and our final learning of regularized chamfer matching. All the detections are evaluated based on PASCAL overlap criterion with the groundtruth object annotations.

	Pedestrians	Cows	Applelogos	Bottles	Giraffes	Mugs	Swans
DCM [2]	3.0	88.1	60.8	85.5	27.0	10.1	33.1
FG Regularization	6.8	89.2	62.0	86.9	36.3	**27.3**	33.8
Combined Regularization	**11.2**	**91.9**	**81.8**	**90.4**	**43.0**	**27.3**	**47.3**

most time consuming part in the code of [2]. The proposed chamfer regularization added only a marginal overhead to the computation time. For instance, only 2 second overhead is observed per image from TUD Cow dataset. On the other hand, computations for the baseline performance [2] took about 15 seconds per image. Thus, our approach turns out to be easily integrable into a state-of-the-art chamfer matching approach, without adding significant overhead in terms of running time.

4.3   Evaluating Foreground and Background Regularization

Table 1 compares the baseline directional chamfer matching, which constitutes the basis of our approach, with the different components of our discriminative chamfer regularization. In particular, the performance of our foreground regularization method and the performance of our combined detector using foreground and background regularization are evaluated. Figure 1 shows the relative importance of various pixels of the foreground template learnt by using a linear SVM. The experiments show that foreground regularization alone already improves performance in terms of average precision on all of these object categories. Additionally applying the background regularization suppresses even more false positives in cluttered background and, thus, yields a significant further gain.

For the TUD Pedestrian dataset the images in the testing set are provided at a very high resolution which yields very low average precision for the directional chamfer matching which is around 3%. The low baseline can be attributed to the high resolution of the test images, since it is known that chamfer matching is sensitive to all the fine details in the edge map. Our foreground regularization more than doubled the average precision obtained from the baseline. Adding the background regularization brought a further gain of 4.4% in average precision. For the Cow dataset directional chamfer matching yields very good performance around 88% average precision. Nevertheless, our combined detector still improves the performance about 4% by exploiting the advantages of foreground and background regularization. In Fig. 3 one can see how foreground reweighting improves the alignment with the groundtruth. The background normalization becomes particularly useful for categories such as applelogos, bottles and swans where most of the performance gain can be attributed to background regularization.

<div align="center">a)   b)   c)</div>
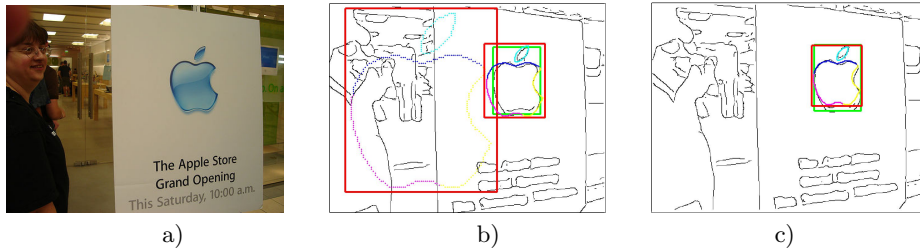
Fig. 4. This example shows how combined foreground and background regularization Eq. (7) can remove false positive detections which could not be eliminated by foreground reweighting alone. Panel a) shows the original image, b) the result obtained by using foreground reweighting and c) the results from the combined foreground and background regularization. Best viewed in color.

The example in Fig. 4 shows that foreground regularization is not always able to suppress false positives in cluttered background and how background regularization can handle such cases. For a challenging category like Giraffes with articulations and background clutter, both foreground and background regularization are found to be equally helpful. For the category of mugs we observed that explaining the foreground more accurately is more important than suppressing false detections in cluttered background. We observe 17.3% improvement in average precision by learning the co-occurrence of template pixels while our combined detector yields results in the same range. All in all our combined detector using foreground and background regularization achieves significant gain on all of the seven categories compared to directional chamfer matching. Additional detection results comparing the regularized chamfer matching to directional chamfer matching are provided in Fig. 5.

4.4   Comparison with State-of-the-Art Extensions to Chamfer Matching

We compare our combined foreground and background regularization with other state-of-the-art extensions to chamfer matching such as the normalized oriented chamfer matching by Ma et al. [12] (NOCM) and the hierarchical deformable template model (HDT) by Zhu et al. [16].

[12] have reported results on two datasets: the TUD Pedestrian dataset [20] and the Cow dataset [19]. [16] have evaluated their method on the Cow dataset. Both the approaches report their results in terms of detection rate at 10% precision. In the previous section, we have reported the gain obtained by our regularization in terms of average precision, since it is taking into account the area under the precision recall curve instead of just one point on the performance curve and therefore is a more robust measure. Nevertheless, to compare ourselves with [12, 16], we need to report results in terms of detection rate at 10% precision.
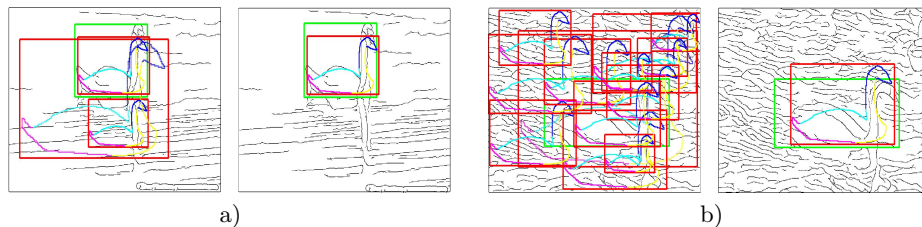
Fig. 5. Panel a) and b) show detection results for two examples. The left image of each panel shows results obtained by directional chamfer matching. The right image of each panel shows the improved detection result after applying chamfer regularization. The groundtruth bounding box is shown in green and the top scoring object hypotheses are shown in red. Best viewed in color.

Table 2. Comparison in terms of **detection rate** at 10% precision (in %) on the Cow dataset and the TUD Pedestrian dataset with OCM, NOCM and HDT.

	Cows	Pedestrians
OCM [1]	73.9	35.2
NOCM [12]	91.0	70.0
HDT [16]	88.2	-
Regularized Chamfer Matching	**98.3**	**80.0**

Table 2 shows the results for the Cow dataset and the TUD Pedestrian dataset. The results indicate that chamfer regularization significantly improves performance on the Cow dataset compared to HDT and NOCM. For TUD Pedestrians we gain 10% in detection rate compared to NOCM. All in all our results confirm that the regularized chamfer matching method significantly improves over state-of-the-art extensions to chamfer matching.

5   Conclusion

This contribution extends the well established and widely used chamfer matching technique, particularly by overcoming its susceptibility to clutter. Our results confirm, that learning the co-occurrence of model points is increasing the specificity of the template by supporting the differing relevance of model points instead of treating them as independent and equally important. Furthermore, accidental matches in background clutter can be suppressed by placing our generic contours on the model contour and learning to distinguish the typical co-occurrence of these contours on cluttered background compared to actual objects. These two contributions are integrated in a max-margin learning framework which is based on state-of-the art directional chamfer matching. [1]

References

1. Shotton, J., Blake, A., Cipolla, R.: Multi-scale categorical object recognition using contour fragments. PAMI **30** (2008) 1270–1281
2. Liu, M., Tuzel, O., A.Veeraraghavan, Chellappa, R.: Fast directional chamfer matching. In: CVPR. (2010)
3. Biederman, I.: Recognition-by-components: A theory of human image understanding. Psychological review **4** (1987) 115–147
4. Attneave, F.: Some informational aspects of visual perception. Psychological review **61** (1954)
5. Barrow, H.G., Tenenbaum, J.M., Bolles, R.C., Wolf, H.C.: Parametric correspondence and chamfer matching: Two new techiques for image matching. Int. Joint Conf. Artifical Intelligence (1977) 659–663
6. Borgefors, G.: Hierarchical chamfer matching: A parametric edge matching algorithm. PAMI **10** (1988) 849–865
7. Leibe, B., Seemann, E., Schiele, B.: Pedestrian detection in crowded scenes. CVPR (2005)
8. Gavrila, D.M., Munder, S.: Multi-cue pedestrian detection and tracking from a moving vehicle. International Journal of Computer Vision **73** (2007) 41–49
9. Lin, Z., Davis, L.S., Doermann, D., DeMenthon, D.: Hierarchical part template matching for human detection and segmentation. ICCV (2007)
10. Thayananthan, A., Stenger, B., Torr, P., Cipolla, R.: Shape context and chamfer matching in cluttered scenes. CVPR (2003)
11. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. PAMI (2002)
12. Ma, T., Yang, X., L.Latecki: Boosting chamfer matching by learning chamfer distance normalization. In: ECCV. (2010)
13. Kadir, T., Brady, M.: Saliency, scale and image description. IJCV **45** (2001)
14. Berg, A.C., Malik, J.: Geometric blur for template matching. In: CVPR. (2001) 607–614
15. Boser, B.E., Guyon, I.M., Vapnik, V.N.: A training algorithm for optimal margin classifiers. 5th Annual ACM Workshop on COLT (1992) 144–152
16. Zhu, L., Chen, Y., Yuille, A.: Learning a hierarchical deformable template for rapid deformable object parsing. PAMI **99** (2009)
17. Martin, D., Fowlkes, C., Malik, C.: Learning to detect natural image boundaries using local brightness, color and texture cues. PAMI **26** (2004) 530–549
18. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology **2** (2011) 27:1–27:27
19. Leibe, B., Leaonardis, A., Schiele., B.: Combined object categroization and segmentation with an implicit shape model. ECCV'04 Workshop on Statistical Learning in Computer Vision. (2004)
20. Andriluka, M., Roth, S., Schiele, B.: People-tracking-by-detection and people-detection-by-tracking. CVPR (2008)