Matching Folded Garments to Unfolded Templates Using Robust Shape Analysis Techniques

Ioannis Mariolis and Sotiris Malassiotis

Information Techologies Institute, Centre for Research & Technology Hellas,
Thessaloniki, Greece
{ymariolis,malasiot}@iti.gr

Abstract. This work presents a novel method performing shape matching of folded garments to unfolded templates, aiming to facilitate unfolding by robotic manipulators. The proposed method incorporates robust shape analysis techniques, estimating point correspondences between contours of folded garments and unfolded templates. The analysis results are also used for estimating the location of the folding axis on the templates and discriminating between different types of garments. The method has been experimentally evaluated using both synthetic and real datasets of folded garments and the produced results indicate the usefulness of the proposed approach.

Keywords: fold detection, partial matching, shape analysis, unfolding.

1 Introduction

Garment unfolding by robotic manipulators is becoming a very active research topic. Perhaps the most challenging aspect of such manipulation is the related vision tasks regarding modeling, reconstruction and recognition of the folded garments. A limited number of existing studies address such tasks, employing different approaches to their realization.

A template matching approach has been adopted in Osawa et al. [1], which focuses on the unfolding of massive laundry. The degree of similarity between templates is calculated by evaluating covariance between images. In Hamajima et al. [2], unfolding is facilitated by the detection of hemlines, which was based on the appearance of shadows. Maitin et al. [3] address robotic handling of towels using stereo cameras for searching geometric cues. A system handling different types of garments by comparing observed contours to simulated ones is presented in Cusumano et al. [4]. In Willimon et al. [5], features such as corners, peak region and continuity of the cloth are used to determine a location and orientation in order to interact with the cloth and unfold or flatten it. The presented studies are aiming to unfold hung garments by locating suitable regrasping points. In a different direction, in Triantafyllou et al. [6], edges and corners of a piece of fabric lying on a table are extracted and classified according to an estimated

topology. The classified corners are used to detect folds and facilitate unfolding by a single manipulator. However, that study addresses folds on convex pieces of fabric instead of real garments.

In this work, in order to facilitate unfolding of real garments lying on a table, matching their contours to unfolded templates is proposed. To the authors' knowledge, this is the first study addressing the problem of matching folded shapes to unfolded templates in general. However, additional challenges are presented to the matching task in case garments are considered, since non-rigid deformations can also be introduced to the folded shapes. The proposed approach considers zero-curvature folds yielding approximately planar configurations for the folded garments. The key assumption of the approach is that the folding axis becomes part of the contour of the folded garment. This assumption is always valid when only a single zero-curvature fold exists. Although the presented method mainly addresses single folds, it can be recursively applied in case of multiple folds.

This paper is organized as follows. In Section 2 the proposed method performing matching of folded shapes is presented, whereas in Section 3 the experimental results are demonstrated. The paper concludes in Section 4.

2   Matching Folded Shapes to Unfolded Templates

According to the proposed method, an image of a folded garment is acquired and its contour is extracted and approximated by a polygon [7]. Then, each side of the simplified polygon is examined as potential folding axis. The examined side is removed from the folded contour and partial matches between the resulted open contour and an unfolded template are detected. Each partial match defines an affine transformation, which is applied to the examined side generating a hypothesis about the folding axis location on the template. In order to test the hypothesis, the template is virtually folded and the resulted polygon is matched to the folded garment using Inner Distance Shape Contexts (IDSC) algorithm [8]. Based on the hypothesis producing the best match, point correspondences between the matched polygons are established, whereas the location of the folding axis on the template is estimated. An overview of the method is depicted in the block diagram of Figure 1.

2.1   Partial Matching

Matching the folded contour to the template is extremely challenging, since the original contour can be fragmented over several pieces after folding. Moreover, some fragments can be missing due to overlaps, whereas fragments belonging to different parts of the original contour can be wrongly connected after folding. The existence of fragmented correspondences indicates that only partial contour matching can be achieved. In order to tackle the above challenges, the approach proposed in Riemenschneider et al. [9], has been adopted.
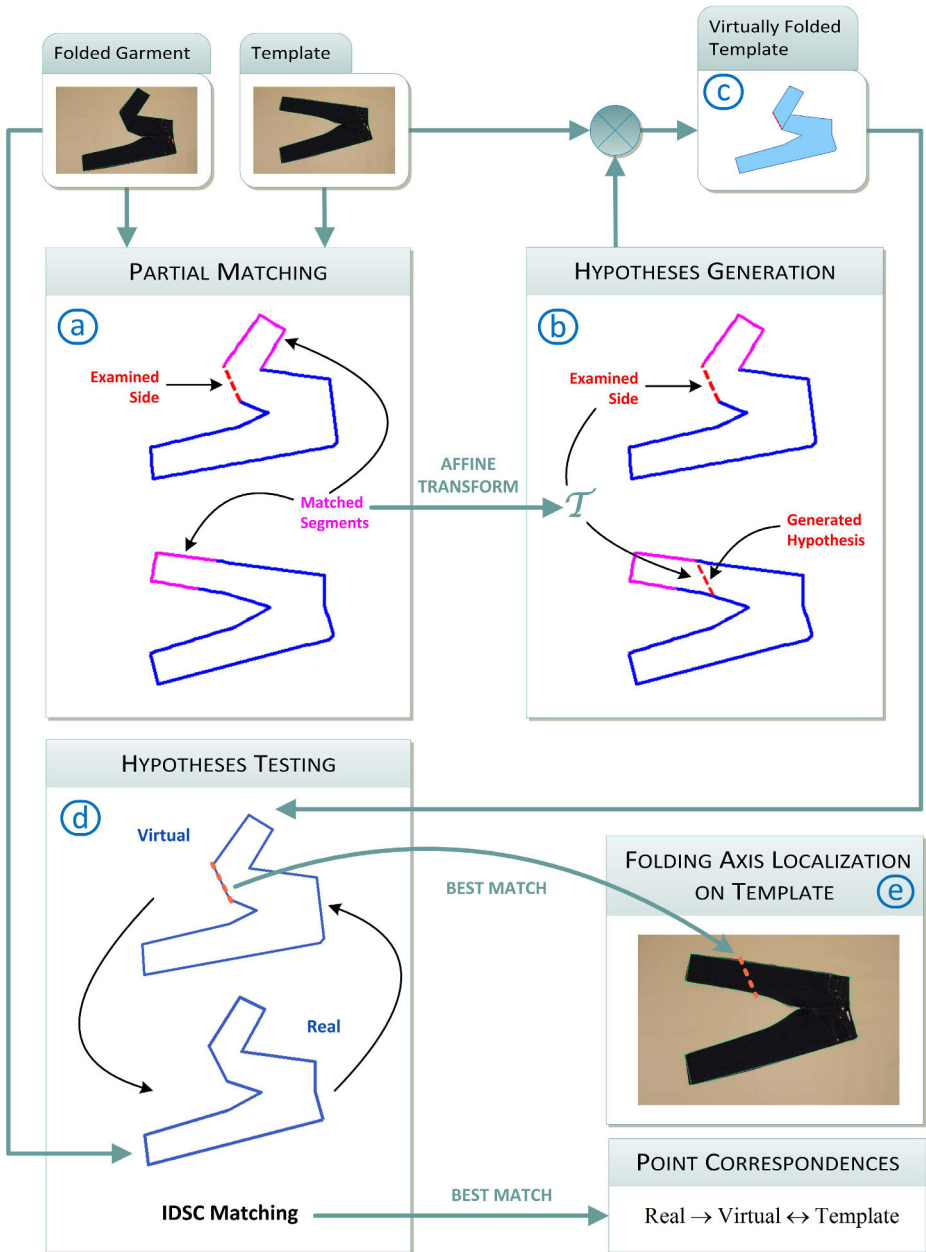
Fig. 1. Block diagram of the proposed method performing shape matching of folded garments

A significant advantage of the adopted matching technique is that it enables efficient aggregation of partial matches, whereas it is also translation and rotation invariant. In this work, a novel strategy for extracting the descriptors is proposed in order to achieve reflection invariance, as well. This property is very useful in case of folding, since some segments of the folded contour may represent reflections of segments belonging to the initial contour. According to the proposed strategy an initial contour orientation is selected for both the template and the folded contour and the descriptors are extracted, whereas a second set of descriptors is extracted by the folded contour after inverting its orientation. Then, partial matching is performed twice. The first time the initial set of descriptors is employed, matching unfolded segments on the contours. The second time the other set of descriptors is employed, matching segments on the template's contour to reflected segments on the folded contour. Each match is assigned with an affinity score As and after thresholding only strong matches remain. As described in [9] similar overlapping contours are merged and the remaining matches are aggregated.

The described matching procedure is repeated for every side of the folded contour. In each repetition, the template contour is matched to the open contour that results after removing the examined side. In Figure 1a the simplified contours of a folded pair of pants and an unfolded template are illustrated. The examined side is denoted by a red dashed line, whereas matched segments are denoted by magenta lines. The proposed method examines all sides and produces a large number of matches, which are not always correct. Each match is used to generate a hypothesis about the location of the folding axis on the template and the validity of the hypothesis is tested.

2.2 Hypothesis Generation and Testing

Each partial match is used to estimate a local affine transformation, which is then applied to the examined side. The new location of the transformed side generates a hypothesis about the position and orientation of the folding axis, with respect to the employed template. A significant advantage of using partial matches to generate the hypotheses is that the method presents insensitivity to local deformations. In Figure 1b an example of a generated hypothesis is illustrated. The red dashed line on the template corresponds to the generated hypothesis about the folding axis location. In this example the presented hypothesis is valid.

Each generated hypothesis is assigned with a confidence score Cs, which is computed as a product of the affinity score As of the match used for generating the hypothesis and Rs score, which is given by:

$$Rs = \begin{cases} e^{-\frac{Lout}{Lin}}, & \text{if } Lin \neq 0 \\ 0, & \text{if } Lin = 0 \end{cases} \tag{1}$$

where $Lout$ denotes the axis length that is predicted to be outside the template and Lin denotes the axis length that is predicted to be inside the template. Both As and Rs, take values in the $[0,1)$ interval. The higher their values, the greater

is the confidence on the validity of a certain hypothesis. In case $Rs = 0$, the hypothesis is automatically rejected.

Different partial matches can result in very similar hypotheses, in case the matches are correct. Therefore, an approach similar to the one used for merging the matched segments is proposed for aggregating the surviving hypotheses. Aggregation is based on a binary image, where each pixel represents a hypothesis. The horizontal image axis corresponds to angles, the vertical axis corresponds to distances and each pixel represents the folding axis of each hypothesis in polar form. Similar hypotheses are expected to create connected regions in this binary image, which can be extracted by connected component analysis. After extracting a component, the associated confidence scores are summed and the result is thresholded. The hypotheses of each component that survives thresholding are merged by calculating a weighted average of the corresponding folding axes. The assigned confidence scores are used as weights for the calculation. By thresholding and aggregating, only a small number of hypotheses proceed to the final testing stage, which is the most computationally intensive.

Each hypothesis determines a folding axis on the template contour. The predicted axis is used to virtually fold the template and a predicted folded contour is generated (Figure 1c). The predicted contour is matched to the folded contour using IDSC and a matching cost is estimated (Figure 1d). IDSC has been selected due to its insensitivity to articulations [8]. The hypothesis resulting to the minimum matching cost is selected as the most probable one. The selected hypothesis is accepted only if the associated matching cost is lower than the one produced when the unfolded template is employed. The IDSC correspondences between the folded garment and the virtual contour of the selected hypothesis are used for matching the folded garment to the unfolded template. The selected hypothesis provides also the location of the folding axis on the template.

In case multiple templates belonging to different garment types are employed, the selected hypothesis can be also used to predict the actual type of the folded garment. Thus, the garments can be recognized while they are folded and the estimated configuration can be exploited for planning the unfolding strategy.

In order to deal with garments presenting multiple folds the presented method can be applied recursively. The method is initially applied and the folding axis resulting to the most severe deformation is detected. Then, the corresponding virtually folded template is employed as new template and the method is reapplied. This procedure can be repeated until the configuration that presents the lowest matching cost is reached.

3   Experimental Evaluation

A series of experiments has been designed in order to assess the effectiveness of the proposed approach. In the first part of the experiments a synthetic dataset is created consisting of virtually folded templates, whereas for the second part of the experiments images of folded garments are employed.

3.1   Synthetic Dataset

The synthetic dataset consists of virtually folded polygons that were extracted from images of real garments. Five different types of garment shapes have been considered: shirts, t-shirts, pants, shorts and skirts. The extracted polygons were sampled by roughly 120 points and for each polygon 60 different folding axes have been randomly selected using pairs of contour points. Then, the polygons have been virtually folded using the selected axes. The original polygons before folding were used as templates.

   Matching results are strongly dependent on the accuracy of the folding axis localization. Therefore, the main measure used for the evaluation of the method is the difference between the predicted location of the folding axis and the ground truth. A summary of the results is presented in Table 1, where the mean differences in the polar form parameters $(d\rho, d\theta)$ of the axis location over the 60 folds and the corresponding standard deviations are provided (Table 1, columns 2 and 3). A 0.1 difference in the normalized distances[1] and a 5 degrees difference in the angles have been selected as reasonable thresholds for measuring the rate of correct axis localization (Table 1, column 4). However, there are a few cases where localization failure is not a matter of accuracy, but is caused by accepting erroneous hypotheses that are based on wrong matches. However, as shown in the last column of Table 1, in the total of 300 folding scenarios only 4 such cases occur.

Table 1. Folding axis localization results for 5 garments, using 60 folding scenarios

Type	$d\rho$	$d\theta^o$	Localization Rate	Mismatches
Pants	0.019 ± 0.020	1.75 ± 1.82	96.7%	2
Shorts	0.014 ± 0.019	0.96 ± 1.01	96.7%	1
Shirt	0.017 ± 0.018	1.17 ± 1.04	95%	1
T-shirt	0.016 ± 0.022	0.90 ± 1.12	98.3%	0
Skirt	0.014 ± 0.015	1.04 ± 0.91	98.3%	0

3.2   Real Dataset

In the second part of the experiments the method's performance in case of actual folds has been evaluated. Apart from the garment types used in the first part of the experiment, towels were also considered. Two different garments were employed for each type and an image database of 12 garments folded by 4 or 5 different axes has been acquired. In Figure 2 example images of the created database are presented. At first, the garments used for folding were also used as templates and a correct localization rate of 96.7%, with no mismatches, has been achieved for the entire database. Then, different templates have been employed

---

[1] The distance is normalized by dividing with the length of the major axis of the ellipse presenting the same second central moments with the template.

and the correct localization rate dropped to 88.5%, with two mismatches. However, in this case due to the lack of ground truth for the folding axis location on the template, localization failure has been determined subjectively through manual inspection.

In case multiple templates of different garment types were employed, the method was always able to discriminate between them based on the template that generates the hypothesis with the minimum matching cost.
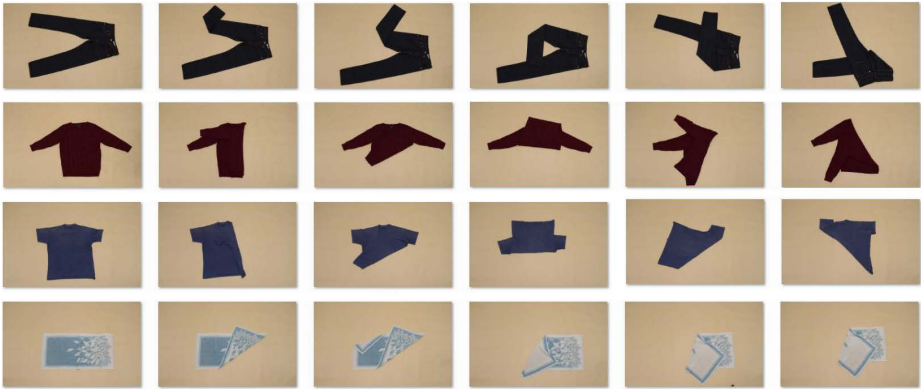

Fig. 2. Example images of the folded garments in the real dataset

4   Conclusion

In this work a method for matching contours of folded garments to unfolded templates in order to facilitate unfolding has been proposed. The method is based on the simple but reasonable assumption that the folding axis becomes part of the outer boundary of the folded garment. Hence, each side of the folded contour is considered to be potential folding axis. Using unfolded templates, hypotheses can be generated about the position and orientation of the axis with respect to these templates. The hypotheses are generated based on partial matching between the template's contour and the folded contour. Using only contour information is an important feature of the proposed method, since when it is applied to real garments it is not affected by their great variation in texture and colour. Another important feature of the proposed method is that matching is performed in both local and global scale using robust techniques that present insensitivity to deformations and articulations, respectively.

The presented experimental evaluation documented high accuracy in folding axis localization. Moreover, the correct localization rate for both virtual and real folds has been about 97%. As expected, this rate is decreased in case different garments are used as templates. Hence, using the estimated correspondences for refining the templates should be investigated. In future work, using the matching results of the proposed method for robotic planning of unfolding will also be investigated.

References

1. Osawa, F., Seki, H., Kamiya, Y.: Unfolding of Massive Laundry and Classification Types. J. of Adv. Comp. Intell. and Intell. Inf. 11, 457–463 (2007)
2. Hamajima, K., Kakikura, M.: Planning strategy for task of unfolding clothes. Rob. and Aut. Syst. 32, 145–152 (2000)
3. Maitin-Shepard, J., Cusumano-Towner, M., Lei, J., Abbeel, P.: Cloth Grasp Point Detection based on Multiple-View Geometric Cues with Application to Robotic Towel Folding. In: ICRA, pp. 2308–2315 (2010)
4. Cusumano-Towner, M., Singh, A., Miller, S., O'Brien, J.F., Abbeel, P.: Bringing Clothing into Desired Configurations with Limited Perception. In: ICRA, pp. 3893–3900 (2011)
5. Willimon, B., Birchfield, S., Walker, I.D.: Model for unfolding laundry using interactive perception. In: IROS, pp. 4871–4876. IEEE Press (2011)
6. Triantafyllou, D., Aspragathos, N.A.: A Vision System for the Unfolding of Highly Non-rigid Objects on a Table by One Manipulator. In: Jeschke, S., Liu, H., Schilberg, D. (eds.) ICIRA 2011, Part I. LNCS, vol. 7101, pp. 509–519. Springer, Heidelberg (2011)
7. Hershberger, J., Snoeyink, J.: An o(nlogn) implementation of the douglas-peucker algorithm for line simplification. In: Proceedings of the 10th Annual Symposium on Computational Geometry, SCG 1994, pp. 383–384. ACM, New York (1994)
8. Ling, H., Jacobs, D.W.: Shape classification using the inner-distance. IEEE Trans. Pattern Anal. Mach. Intell. 29, 286–299 (2007)
9. Riemenschneider, H., Donoser, M., Bischof, H.: Using partial edge contour matches for efficient object category localization. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part V. LNCS, vol. 6315, pp. 29–42. Springer, Heidelberg (2010)