Deformable Wide Baseline Matching using Markov Random Fields

Selen Atasoy[1,2], Ben Glocker[1], Stamatia Giannarou[2], Diana Mateus[1], Alexander Meining[3] (MD), Guang-Zhong Yang[2] and Nassir Navab[1]

[1]Computer Aided Medical Procedures (CAMP), Technische Universitaet Muenchen
[2]Visual Information Processing Group, Imperial College London,
[3]Department of Gastroenterology, Klinikum Rechts der Isar, Technische Universitaet Muenchen

Introduction

Recent advances in biophotonics have enabled optical biopsy techniques for cancer surveillance. The non-invasive nature of these techniques, however, entails the difficulty of identifying previously visited biopsy locations. This work presents a novel wide baseline matching approach for endoscopic images of deformable soft tissue. The task of matching affine invariant image regions is modeled in a Markov Random Field (MRF) framework. The proposed model incorporates appearance based region similarities as well as spatial correlations of the regions. In particular, a geometric constraint is introduced that is robust to large scale change present in endoscopy and to a large degree of tissue deformation. The performance of the method as compared to the existing state-of-the-art is evaluated on both *in-vivo* and simulation datasets with varying levels of visual complexities.

Methods

1. **Region Detection and Description:** Affine covariant regions are detected independently on both images using the anisotopic region detector [1]. For viewpoint invariant description, each elliptical region is mapped onto a circular patch and described using the SIFT descriptor [2].
2. **Matching through Markov Random Fields:** Matching the region descriptors is modeled as the global optimization of a Markov Random Field (MRF). The regions in the first image are defined as the nodes and the regions in the second image as the labels of the MRF. Furthermore, a null-label is introduced to assign to regions without true correspondence in the second image. The matching is computed by finding the maximum a posteriori estimate for optimum labelling.
3. **Unary Costs:** Photometric similarities between the node and the label regions are evaluated by defining the cost of assigning the label l_p to the node p to be the angle between the SIFT descriptors normalized by the maximum possible angle. Furthermore, a ***null-cost function*** is introduced defining the cost of not matching a region (assigning the null label l_0). The motivation is that not matching a region with a strong matching candidate should have a higher cost than a region with a weak matching candidate. Thus, the unary costs are defined as:

$$V_p(l_p) = \begin{cases} \dfrac{\arccos(d(\bar{p},\sigma_{\bar{p}},\vartheta_{\bar{p}}) \cdot d(\bar{l_p},\sigma_{\bar{l_p}},\vartheta_{\bar{l_p}}))}{\arccos(0)} & \text{if } l_p = l_0 \quad \textit{Photometric Similarity} \\ \alpha \cdot (1 - \min(V_p(\cdot))) & \text{if } l_p \neq l_0 \quad \textit{Null-cost Function} \end{cases}$$

where α is the factor regulating the trade-off between the quality and the number of matches.

4. **Neighbourhood Systems:**
   - ***Global neighborhood system:*** In the context of the matching problem each region is allowed to have at most one correspondence in the second image. This ***uniqueness constraint*** is included into the energy function using a fully connected graph of nodes and defining the pair-wise cost of assigning same label to two nodes to be infinite.
   - ***Local neighborhood system:*** We further define a local neighborhood system for both, nodes and the labels to apply geometric constraints within a local region. Therefore, each region is connected to other regions within a distance threshold: $\mathcal{N}_{\text{local}}(p) = \{q \neq p \mid \; ||\, p - q \,|| < t\}$.
5. **Pair-wise Costs:** In this work we propose a flexible geometric constraint based on the assumption that neighboring regions move with similar transformations. If two neighboring regions p and q have corresponding regions l_p and l_q in the second image, then there exist two affine transformations $A_p = s_p \cdot R \cdot M_p$ and $A_q = s_q \cdot R \cdot M_q$, where R is the rotation of the local neighborhood, s_p and s_q the scale factors and A_p and A_q the local affine transformations mapping the elliptical regions to circles . Thus, if (p, l_p) and (q, l_q) are true correspondences, then the descriptors $d(\bar{q},\sigma_{\bar{q}},\vartheta_{\bar{p}})$ and $d(\bar{l_q},\sigma_{\bar{l_q}},\vartheta_{\bar{l_p}})$ should be similar. For all nodes within the local neighborhood system we define the ***geometric constraint*** to be:

$$\psi_{pq}(l_p,l_q) = \begin{cases} \infty & \text{if } (l_p \notin \mathcal{N}_{\text{local}}(l_q)) \\ \dfrac{\arccos(d(\bar{q},\sigma_{\bar{q}},\vartheta_{\bar{p}}) \cdot d(\bar{l_q},\sigma_{\bar{l_q}},\vartheta_{\bar{l_p}}))}{\arccos(0)} & \text{if } (l_p \in \mathcal{N}_{\text{local}}(l_q)) \end{cases}$$

The final pair-wise costs are defined as:

$$V_{pq}(l_p,l_q) = \begin{cases} \infty & \text{if } (l_p = l_q \neq l_0) \quad \textit{Uniqueness Constraint} \\ \alpha & \text{if } (l_p = l_0) \vee (l_q = l_0) \quad \textit{Null-cost Function} \\ \psi_{pq}(l_p,l_q) & \text{if } (q \in \mathcal{N}_{\text{local}}(p)) \quad \textit{Geometric Constraint} \\ 0 & \text{otherwise} \end{cases}$$

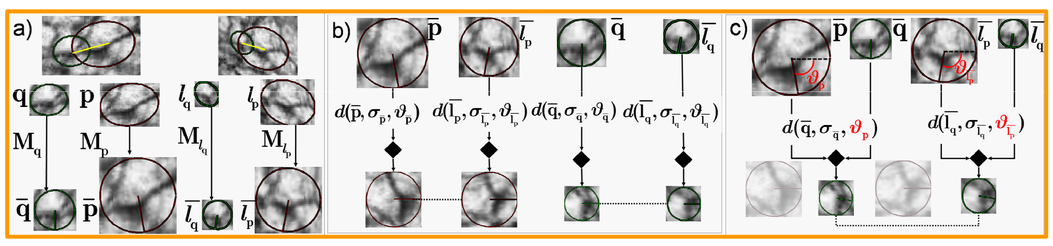

Fig. 1. a) Viewpoint invariant region description. b) Unary costs c) The proposed pair-wise costs.

6. **MAP Labelling:** Finally, the maximum a posteriori estimate of the MRF labelling is computed by minimizing the final energy function

$$l^* = \arg\min_l E_{\text{MRF}}(l) = \arg\min_l \sum_{p \in G} V_p(l_p) + \sum_{p \in G} \sum_{q \in N(p)} V_{pq}(l_p,l_q)$$

using Belief Propagation.

Results

The performance of the proposed method is evaluated on 4 *in-vivo* and 4 simulation datasets and compared to three descriptor matching strategies evaluated in [3]; the threshold-based (TB), nearest-neighbor (NN) and the nearest neighbor distance ratio matching (NNDR).

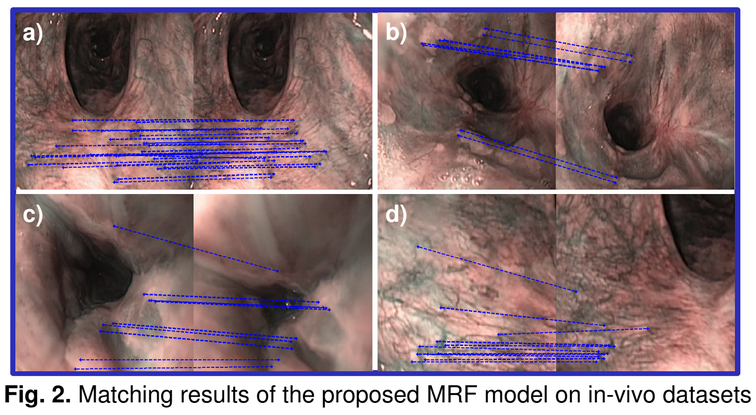

Fig. 2. Matching results of the proposed MRF model on in-vivo datasets

The matching result of our method on *in-vivo* datasets is illustrated in Fig.2.

For the quantitative analysis, the measures *recall* and *precision* of each matching method are evaluated. For all datasets (simulation and *in-vivo*) maximum recall values for the precision interval (80%-100% inliers) is demonstrated in Fig.3.

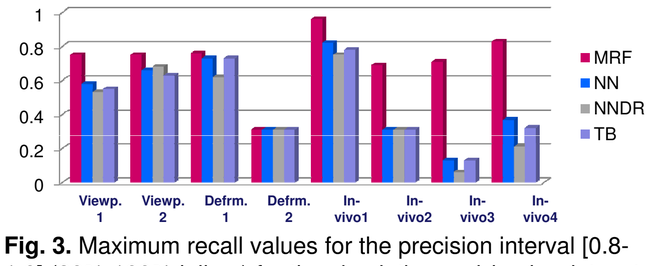

Fig. 3. Maximum recall values for the precision interval [0.8-1.0] (80%-100% inliers) for the simulation and *in-vivo* datasets.

Conclusion

In this work, the task of wide baseline matching in endoscopic images is investigated.

A specific MRF model is deigned for deformable soft tissue matching. In the presented model the appearance and geometric constraints are evaluated in the same space (photometry), allowing for their seamless integration into the MRF objective function.

A novel geometric constraint is introduced, which is invariant to large scale change present in endoscopy and to a large degree of tissue deformation.

Our results demonstrate the robustness of the proposed approach for deformable wide-baseline matching towards an image-based solution for targeted optical biopsy.

References

[1] S. Giannarou, M. Visentini-Scarzanella, G.Z. Yang: Affine-Invariant Anisotropic Detector For Soft Tissue Tracking in Minimally Invasive Surgery , to be published in IEEE International Symposium on Biomedical Imaging (ISBI), 2009.
[2] D. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision, 60(2):91-110, 2004.
[3] K. Mikolajczyk and C. Schmid: A performance evaluation of local descriptors, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(10):1615-1630, 2005.

* This work will be published and presented in *Medical Image Computing and Computer-Assisted Intervention, MICCAI, 2009, London, UK, September 20-24 2009*