



HAL
open science

Local Surf-Based Keypoint Transfer Segmentation

Antoine Bralet, Razmig Kéchichian, Sebastien Valette

► **To cite this version:**

Antoine Bralet, Razmig Kéchichian, Sebastien Valette. Local Surf-Based Keypoint Transfer Segmentation. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Apr 2021, Nice, France. pp.390-393, 10.1109/ISBI48211.2021.9434106 . hal-03244747

HAL Id: hal-03244747

<https://hal.science/hal-03244747>

Submitted on 1 Jun 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

LOCAL SURF-BASED KEYPOINT TRANSFER SEGMENTATION

Antoine Bralet Razmig K echichian S ebastien Valette

Univ Lyon, INSA-Lyon, Universit e Claude Bernard Lyon 1, UJM-Saint Etienne, CNRS, Inserm
CREATIS UMR 5220, U1206, F-69100, VILLEURBANNE, France

ABSTRACT

This paper presents an improvement of the keypoint transfer method for the segmentation of 3D medical images. Our approach is based on 3D SURF keypoint extraction, instead of 3D SIFT in the original algorithm. This yields a significantly higher number of keypoints, which allows to use a local segmentation transfer approach. The resulting segmentation accuracy is significantly increased, and smaller organs can be segmented correctly. We also propose a keypoint selection step which provides a good balance between speed and accuracy. We illustrate the efficiency of our approach with comparisons against state of the art methods.

Index Terms— Whole-body, Probabilistic and statistical models & methods, Image segmentation

1. INTRODUCTION AND RELATED WORK

Whole-body segmentation is still a difficult problem in medical image analysis. Recent advances in machine learning, and particularly in deep learning methods, have brought improvements to segmentation algorithms [1]. However applications of deep learning to whole body segmentation still face important challenges, such as the lack of large training datasets, and the large dimensions of 3D medical images (often about 512^3 voxels) which make the application of 3D deep models infeasible due to very high memory requirements. Indeed, most deep learning methods applied in 3D either resort to resampling or use 2D projections of 3D data [2].

We choose a keypoint transfer method for 3D segmentation, originally proposed in [3], as a starting point. Similarly to multi-atlas segmentation methods [4], this method segments organs in a target image by transferring organ segmentations (annotations) of reference images onto the target and combining them. However unlike most multi-atlas methods, it segments a whole-body image in less than a minute with a relatively small memory footprint. This is due to the use of keypoints instead of image matrices and the lack of complex and time consuming registration steps necessary in multi-atlas methods. We first describe the reference method, illustrated in Fig. 1, prior to presenting our contributions.

The keypoint transfer algorithm proceeds in 4 steps:

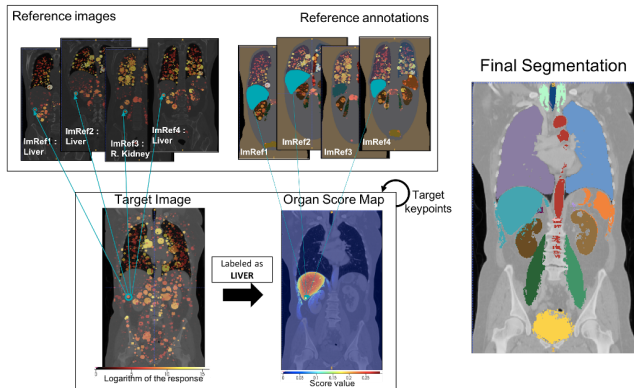


Fig. 1: Illustration of the keypoint transfer algorithm following a single target keypoint, circled in blue. Matched reference keypoints, circled in blue, vote for their labels: 3 votes for liver and 1 for R. kidney. Liver annotations of voting images are transferred to the target, adding up to its liver score map. Once all target keypoints are studied, organ maps are merged to obtain the final segmentation.

Preprocessing: Applied to reference images to extract keypoints (3D SIFT [5] in the reference method) and removing keypoint located outside organs of interest.

Keypoint matching: Keypoints are extracted from the target image, and for each keypoint the most similar keypoint in each reference image is found. The matching is done according to keypoint scale, descriptor and coordinates criteria after a simple translation-only registration of reference images onto the target.

Keypoint voting: For a given target keypoint, matched reference keypoints vote for their organ label, weighted by translational consistency and descriptor similarity criteria to take uncertainty into account. The label with the highest score is assigned to the target keypoint.

Keypoint segmentation: For each target keypoint, the annotations of all matched reference keypoints with the same label are transferred onto the target image, weighted by label score and voxel-wise intensity difference between reference and target images, thus creating a score map for each organ. The final target segmentation is created by thresholding organ score maps to reduce over-segmentation caused by outlier annotations, and by merging them keeping the label that has the highest score for a given voxel in case of overlap.

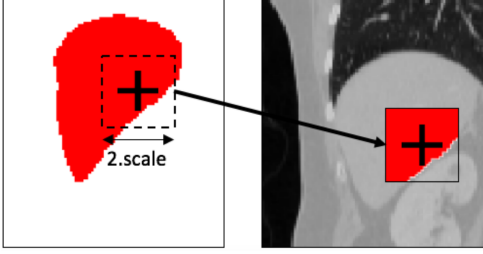


Fig. 2: Local segmentation: instead of transferring the full reference annotation (left), we extract a cubical region around the reference keypoint (the side of which depends on the keypoint scale) and transfer it to the target image (right).

The main shortcoming of the original keypoint transfer method is its relatively low segmentation accuracy, compared to multi-atlas methods, which also results in the loss of smaller organs. Therefore, contributions we propose aim at improving this drawback. Specifically, we increase the number of extracted keypoints greatly using SURF [6] instead of SIFT keypoints. Moreover we introduce a local transfer method which significantly improves segmentation accuracy. To further increase segmentation speed, we devise a keypoint selection algorithm to increase segmentation speed while maintaining accuracy levels.

2. METHODS

2.1. 3D SURF keypoint extraction

The keypoint extraction method we choose is the 3D SURF approach described in [7]. This choice is motivated by two criteria: its short computational time and the number of keypoints extracted, about 140,000 per image. Such a large set of keypoints helps to cover the majority of thoracic and abdominal organs, even the smallest ones such as the gallbladder and the adrenal glands. This enables us to apply the keypoint transfer algorithm to smaller organs, unlike [3].

2.2. Local segmentation

A major improvement of the keypoint-transfer segmentation algorithm comes from our introduction of *local* segmentation which consists in transferring only the region surrounding the reference keypoint to the target image as opposed to *global* segmentation which transfers the full organ annotation.

The motivation comes from observed target over-segmentations, which are due to the high inter-patient variability in organ shape and size of transferred reference annotations. A good match between two keypoints does not imply similar boundaries between reference and target organs: a keypoint carries information about its local neighborhood, it is oblivious to the larger context of the organ it lies in.

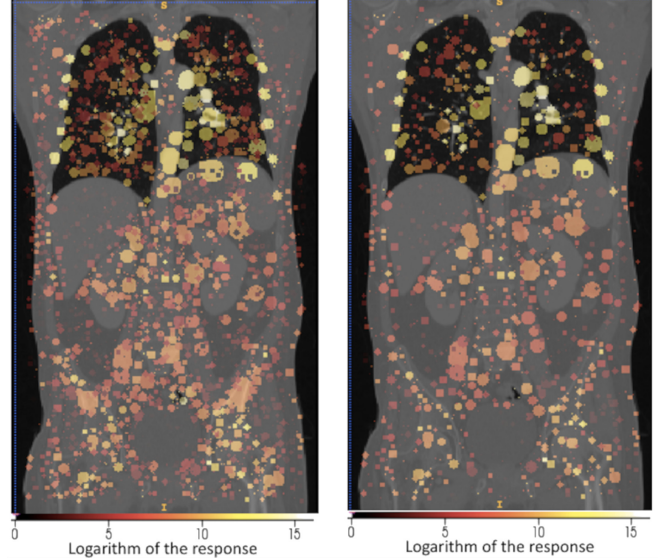


Fig. 3: Comparison between the number and location of keypoints before (left) and after (right) their selection which removes the majority of repetitive keypoints.

When using 3D SURF extraction, around 15,000 keypoints lie inside reference organs, which is significantly higher than the number of 3D SIFT keypoints in [3]. This allows us to use a local segmentation approach whereby only a local region of interest around a reference keypoint is transferred from the reference annotation to the target image, as shown in Fig. 2. Following the extraction parameters in [7], the transferred region is a cubical region the side of which is twice the scale of the keypoint.

2.3. Keypoint selection

As previously mentioned, we have a very large number of keypoints which facilitate local segmentation. One advantage of the original keypoint-transfer segmentation algorithm [3] is its very short computation time, which is partially due to the very small number of keypoints (no more than 1,000 per image). Hence the motivation to reduce the number of 3D SURF keypoints. The left image of Fig. 3 shows that the majority of keypoints overlap. To reduce the number of keypoints, we order them in decreasing order according to the detector response, then for each keypoint k we examine the spherical image area it covers (according to its coordinates and scale). If k covers a region that is already 50% covered by one or more keypoints with a higher detector response, k is removed from the set of keypoints. The set of keypoints which results from this algorithm is illustrated on the right image of Fig. 3. We can notice that the number of keypoints is greatly reduced, and the remaining keypoints are less cluttered.

	All keypoints (~ 140 000)		Selected keypoints (~ 50 000)	
	Global	Local	Global	Local
Execution Time	56:11	43:24	38:47	21:19
Dice	0.6363	0.6906	0.6305	0.6817
Average Distance (mm)	3.46	3.06	3.55	3.12
Hausdorff Distance (mm)	41.58	39.09	42.25	39.19

Table 1: Mean performance metrics and execution times for cross-validated experiments. Two strategies are explored: global and local transfer using all and selected keypoints. Dark green cells report best results for the given strategy and light green ones report the second best.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1. Data and compliance with ethical standards

We evaluate our method on 20 thoracic-abdominal contrast-enhanced 3D CT images with 20 organ annotations, manually segmented by specialists. These images come from the VISCERAL Challenge training dataset [8], obtained in compliance with ethical standards. We have resampled images to a mean dimension of $200 \times 200 \times 349$ for comparison with [3].

The experimental protocol follows a leave-one-out strategy, whereby 19 images are used as references and 1 as target, for all images. Segmentation quality has been evaluated using the Dice similarity index, the average surface distance and the Hausdorff distance, averaged over experiments to obtain overall and organ-wise performances of the algorithm.

3.2. Evaluation of contributions

Table 1 gives mean performance metrics over all experiments, in addition to mean execution times as it is one of the important aspects of the keypoint transfer algorithm. The only settings that change from an experiment to another are the segmentation strategy (global or local transfer, Section 2.2) and the number of keypoints (all keypoints or selected ones, Section 2.3). The table shows that when using all keypoints, local transfer is better than global transfer. This is confirmed by a 0.05 increase of mean Dice and a 0.4 mm reduction of mean average distance values, in addition to a 23% decrease of execution time. The one-sided paired t-test comparing mean Dice performances of global and local segmentations gives a p-value of 2.8×10^{-5} , which confirms the improvement of the local method. We can also see that selecting about one third of original keypoints has little influence on global segmentation accuracy, however it reduces execution time by about 31%. Finally, local segmentation using selected keypoints yields the most interesting results as it reduces execution time by 62% while maintaining segmentation quality. Indeed, a one-sided paired t-test on mean Dice measures between this experiment and local transfer with all keypoints gives a p-value of 0.08, which means that their organ-wise performances are quite similar.

3.3. Comparison with related methods

We compare our segmentation method with two related ones on the VISCERAL training dataset: [9], which is a multi-atlas algorithm using a keypoint-based registration approach, and the original keypoint transfer algorithm [3].

Fig. 4 gives per-organ mean Dice values for the three methods. Note that Dice measures are given only for 10 organs in [3]. We can see that our method outperforms [3] on all organs but the lungs. The mean Dice value for the 10 organs segmented in [3] with our local approach is 0.843 when using all keypoints and 0.844 with selected keypoints, compared to 0.796 in [3]. The one-sided paired t-test comparing our algorithm with [3] gives respectively a p-value of 0.023 when using all and 0.017 when using selected keypoints, confirming the significance of the improvement brought by our method. Investigation has shown that the relatively low Dice value for the lung segmentation is due to a slight under-segmentation caused by a relatively low number of keypoints at the basis of lungs. This could be corrected by using a specific threshold for the lungs prior to merging organ score maps. Fig. 4 shows that the multi-atlas method in [9] performs best in terms of segmentation accuracy. However our method comes quite close, especially on organs such as the psoas and abdominal muscles, the aorta and the adrenal glands. Moreover, our method outperforms [9] in terms of execution speed as the latter requires over 2 hours to terminate.

3.4. Timing and memory footprint

All experiments are conducted on a Linux system running on a 2.2 GHz CPU. The memory footprint of local segmentation is 17.7 GB with selected and 20.2 GB with all keypoints.

We should mention that, as of the current implementation of our method, it still is around 20 times slower than the original keypoint transfer algorithm [3]. This is due to two reasons: the much higher number of keypoints and the current Python implementation. A C++ re-implementation is underway and already shows perspectives of important speedup. Indeed, the step that consists in reading descriptors, calculating weights and voting which takes about 15 minutes on all keypoints in Python, terminates in about 38 seconds in the C++ re-implementation.

4. CONCLUSIONS AND FUTURE WORK

In this paper, we have proposed an improvement of the keypoint transfer segmentation algorithm by using 3D SURF keypoints, allowing us to use a local instead of a global segmentation transfer. The new approach improves segmentation accuracy and allows to segment smaller organs as well. As the number of extracted keypoints is much larger compared to the original method, our algorithm is slower. A good trade-off is to reduce the number of keypoints, which greatly reduces execution time while preserving segmentation quality.

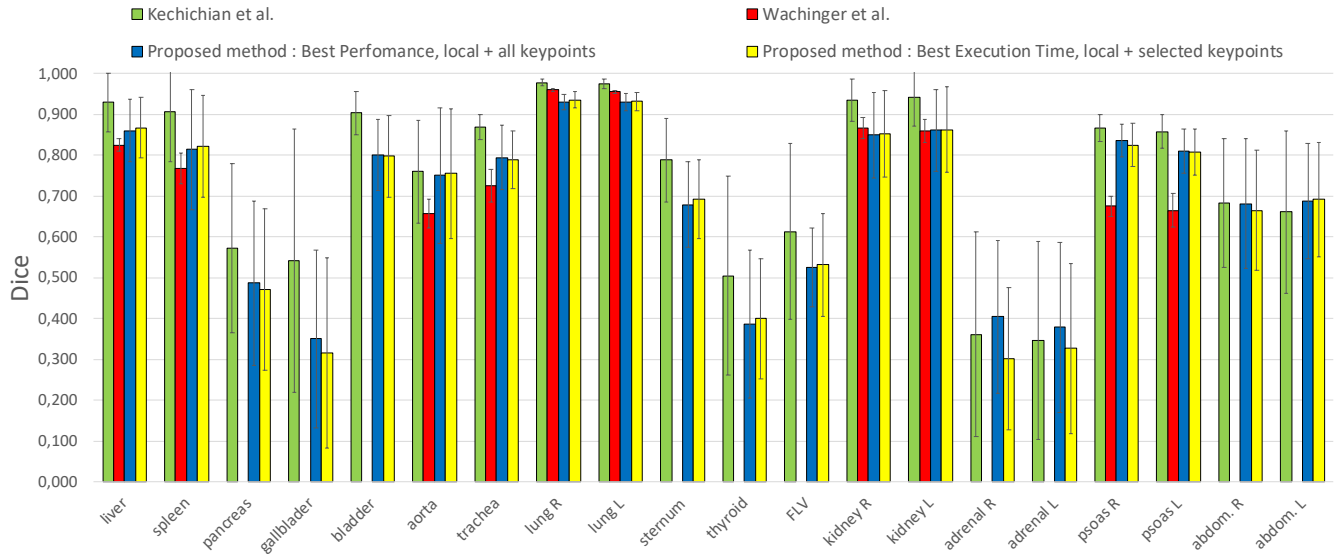


Fig. 4: Comparison of the local transfer segmentation method (using all and selected keypoints) with related state-of-the-art methods on per-organ mean Dice measures evaluated on the VISCERAL training dataset.

Although the performance of the proposed method is promising, there is much room for further improvement. Organ-specific threshold parameters could be defined and optimized. The segmentation of smaller organs (e.g. the gallbladder) could be improved by transferring annotations from nearby larger organs (e.g. the liver). The selection of keypoints, as described in Section 2.3, could be done according to their repeatability in other reference images, as in [10], instead of the detector response. A repeatability score could then be used for sorting keypoints prior to selection and as an additional weight for voting and segmentation.

5. ACKNOWLEDGEMENTS

This work is financed by the TOPACS ANR-19-CE45-0015 project of the French National Research Agency (ANR).

6. REFERENCES

- [1] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, “Image segmentation using deep learning: A survey,” *arXiv preprint arXiv:2001.05566*, 2020.
- [2] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghahfoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, “A survey on deep learning in medical image analysis,” *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.
- [3] C. Wachinger, M. Toews, G. Langs, W. Wells, and P. Golland, “Keypoint transfer for fast whole-body segmentation,” *IEEE Trans. Med. Imag.*, vol. 39, no. 2, pp. 273–282, 2020.
- [4] J. E. Iglesias and M. R. Sabuncu, “Multi-atlas segmentation of biomedical images: a survey,” *Med. Image Anal.*, vol. 24, no. 1, pp. 205–219, 2015.
- [5] M. Toews and W. Wells, “Efficient and robust model-to-image alignment using 3d scale-invariant features,” *Med. Image Anal.*, vol. 17, no. 3, pp. 271–282, 2013.
- [6] H. Bay, T. Tuytelaars, and L. Van Gool, “Surf: Speeded up robust features,” in *European Conf. on Computer Vision*. Springer, 2006, pp. 404–417.
- [7] R. Agier, S. Valette, R. Kéchichian, L. Fanton, and R. Prost, “Hubless keypoint-based 3d deformable groupwise registration,” *Med. Image Anal.*, vol. 59, p. 101564, 2020.
- [8] O. J. del Toro, H. Müller, M. Krenn, K. Gruenberg *et al.*, “Cloud-based evaluation of anatomical structure segmentation and landmark detection algorithms: Visceral anatomy benchmarks,” *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2459–2475, 2016.
- [9] R. Kéchichian, S. Valette, and M. Desvignes, “Automatic multiorgan segmentation via multiscale registration and graph cut,” *IEEE Trans. Med. Imag.*, vol. 37, no. 12, pp. 2739–2749, 2018.
- [10] F. Kahl, J. Alvé, O. Enqvist, F. Fejné, J. Ulén, J. Fredriksson, M. Landgren, and V. Larsson, “Good features for reliable registration in multi-atlas segmentation.” in *VISCERAL Challenge@ ISBI*, 2015, pp. 12–17.