

Improving the Detection Performance in Semi-automatic Landmark Extraction

Sönke Frantz, Karl Rohr, and H. Siegfried Stiehl

Universität Hamburg, Fachbereich Informatik, Arbeitsbereich Kognitive Systeme
Vogt-Kölln-Str. 30, D-22527 Hamburg, Germany
{frantz,rohr,stiehl}@informatik.uni-hamburg.de
<http://kogs-www.informatik.uni-hamburg.de/PROJECTS/imagine/Imagine.html>

Abstract. Manually extracting 3D anatomical point landmarks from tomographic images is generally tedious and time-consuming. A semi-automatic procedure for landmark extraction, which allows for interactive control, offers the possibility to improve on this. The detection performance is decisive for the applicability of such a procedure. However, existing computational approaches to landmark detection often suffer from a larger number of false detections. A considerable number of false detections is caused by neighboring anatomical structures that are captured by the region-of-interest (ROI) at a landmark. In this paper, we present two different approaches to reducing false detections caused by neighboring structures. First, we present a statistical, differential approach to automatically selecting a suitable size for the 3D ROI. Second, we present a differential approach that incorporates additional prior knowledge of the intensity structure at a landmark. Combining both approaches with a robust 3D differential operator for landmark detection, we develop a new algorithm for landmark detection. In estimating the partial derivatives of the intensity function, we can cope with anisotropic voxel sizes using a scheme based on B-spline image interpolation. The new algorithm is applied within a semi-automatic procedure to extract anatomical point landmarks from 3D MR and CT images of the human head.

1 Introduction

Anatomical landmarks are useful features for a wide spectrum of applications in medical image analysis. If selected suitably, such landmarks may represent substantial image information very concisely, which is important in medical applications considering the vast amount of data one has to deal with. Among the different types of landmarks, we are particularly interested in 3D anatomical point landmarks of the human head. The driving application is landmark-based 3D image registration. One advantage of using point landmarks is efficiency, e.g., in establishing correspondences. Also, point landmarks are well suited for interactive control, which we consider crucial in clinical practice. However, manual landmark extraction from images is generally tedious and time-consuming. A (semi-)automatic procedure offers the possibility to improve on this.

Only a few computational approaches to extracting 3D anatomical point landmarks exist (e.g., [1],[2],[3]). In these approaches, 3D differential operators are used to detect different types of landmarks. The complexity of these operators ranges from those using only first order partial derivatives of the intensity function [2] to those using derivatives up to the second or even the third order [3],[1]. In our case, we apply operators within a semi-automatic procedure, which has the advantage that the user can control the results. Semi-automatic means that (i) the user interactively determines the position of a landmark coarsely, (ii) a differential operator is applied within a region-of-interest (ROI) to detect landmark candidates, and (iii) the user selects one candidate. However, a general problem with differential operators is that often a larger number of false detections occurs (see [4] for a study of the detection performance of 3D differential operators for landmark detection). A large number of false detections not only affects the selection of the correct candidate but also the confidence in the results, which in turn would reduce the acceptance of such a procedure. One reason for false detections is that only local intensity information is used, which makes the operators sensitive to noise and small intensity variations. Operators that use high order derivatives are particularly affected by this. Second, a considerable number of false detections is caused by neighboring anatomical structures that are captured by the ROI.

In this paper, we introduce two different approaches to reducing false detections caused by neighboring structures. First, we address the problem of selecting a suitable 3D ROI at a landmark. We present a statistical, differential approach to automatically selecting an optimal ROI size (Sect. 2). Second, we take advantage of prior knowledge of the intensity structure at a landmark to impose additional constraints on the detected candidates. Based on the curvature of the iso-intensity surfaces at the detected positions, we automatically reject those candidates where the present intensity structure is inconsistent with the prior knowledge of the landmark at hand (Sect. 3). Combining both approaches with a robust 3D differential operator for landmark detection, which uses only first order derivatives, we present a new algorithm for landmark detection. In image derivative estimation, we can cope with the typical case of anisotropic voxel sizes using a scheme based on B-spline image interpolation (Sect. 4). The new algorithm is applied within a semi-automatic procedure to extract anatomical point landmarks from 3D MR and CT images of the human head (Sect. 5).

2 Automatically Selecting a Suitable 3D ROI Size

Within our semi-automatic procedure, the user interactively determines the position of the landmark at hand coarsely. At this position, a 3D ROI is placed in which a differential operator is applied to detect landmark candidates. In previous work [2],[5], we used a fixed ROI size. Because of this, however, neighboring anatomical structures were often captured by the ROI, which gave rise to additional detections. Here, we present a statistical approach to automatically selecting a suitable ROI size based on a 3D differential edge intersection

approach, which was previously used for refined landmark localization [6]. In the edge intersection approach, tangent planes are computed to locally approximate the surface at a landmark. The landmark position is estimated by intersecting the tangent planes using the least-squares method. This approach can be used for ROI size selection by taking the statistical uncertainty of the position estimate as a criterion for isolating an anatomical landmark (e.g., a tip) within the ROI. First results based on a 2D version of this approach were presented in [7].

Consider a cubic ROI of width w centered at the interactively determined position. Let $\hat{\mathbf{x}}_w$ denote the position estimate resulting from the 3D edge intersection approach [6]. The statistical uncertainty of $\hat{\mathbf{x}}_w$ is given by the covariance matrix,

$$\mathbf{\Sigma}_w = \sigma_\epsilon^2 \left(\sum_i \nabla g(\mathbf{x}_i) \nabla g(\mathbf{x}_i)^T \right)^{-1},$$

where σ_ϵ^2 is a data-dependent noise term and $\nabla g(\mathbf{x}_i)$ is the intensity gradient at \mathbf{x}_i ; the sum index i addresses all voxels within the ROI. The matrix $\mathbf{\Sigma}_w$ reflects the consistency of the observed data with the assumed polyhedral model of the surface at the landmark. A scalar measure for the uncertainty is the determinant of the covariance matrix, $U_w = \det(\mathbf{\Sigma}_w)$ (generalized variance). The idea behind ROI size selection is to vary the ROI size and then to select the optimal ROI size based on minimal uncertainty.

The procedure for ROI size selection is as follows: We start with a user-specified minimal ROI size (e.g., $w_{min} = 7$ voxels). When the ROI does not capture sufficient intensity information to reliably estimate the landmark position, U_w is large. Hence, taking more image information into account by enlarging the ROI, U_w can be expected to decrease. However, when neighboring structures are captured by the ROI, U_w significantly increases, which suggests that further enlargement of the ROI is not useful. In our implementation, we detect such a change of U_w at $w_{increase}$, say, by requiring that (a) U_w increases and (b) the relative spatial variation of the position estimate exceeds a threshold t_V . Additionally, a maximal ROI size w_{max} is prescribed. Finally, the optimal ROI size w_{opt} is selected in between, i.e., $w_{opt} = \arg \min U_w, w_{min} \leq w \leq \min\{w_{increase}, w_{max}\}$.

3 Incorporating Prior Knowledge of the Landmark

We take advantage of prior knowledge of the intensity structure at a landmark to impose additional constraints on the detected candidates. In general, the user knows the landmark type (e.g., a tip or a saddle point) as well as the used modality and the imaging parameters. Here, we distinguish between tips and saddle points. Additionally, we distinguish between dark and bright tips w.r.t. the background. To classify these structures, we exploit curvature properties of the iso-intensity surfaces at the detected positions.

Suppose we have detected a point \mathbf{x}_d on the surface of an anatomical structure. Let K denote the Gaussian curvature and H the mean curvature of the

isointensity surface at \mathbf{x}_d , which is implicitly defined by $g(\mathbf{x}) - g(\mathbf{x}_d) = 0$ (for details on computing differential measures of isointensity surfaces, see, e.g., [8],[1]). Using the sign of K , we distinguish between tips ($K > 0$) and saddle points ($K < 0$). Using the sign of H , we further distinguish between dark ($H < 0$) and bright ($H > 0$) tips w.r.t. the background. The candidate is rejected if the classification of the present intensity structure is inconsistent with the expected intensity structure.

4 Improved Algorithm for Landmark Detection

In sum, our new algorithm for landmark detection comprises three constituents:

- (1) A cubic ROI is centered at the interactively determined position, and then an optimal size for the ROI is automatically selected (Sect. 2).
- (2) Landmark candidates are detected by applying a computationally efficient 3D differential operator [2]. The operator is applicable to different types of landmarks and is relatively robust w.r.t. noise since only first order derivatives are used (cf. [4]). The operator reads $Op3 = \det(\mathbf{C}) / \text{trace}(\mathbf{C})$, where \mathbf{C} denotes the averaged dyadic product of the intensity gradient, $\mathbf{C} = \overline{\nabla g \nabla g^T}$.
- (3) Detected candidates with an intensity structure being inconsistent with the prior knowledge of the landmark at hand are automatically rejected (Sect. 3).

In addition, in estimating the partial derivatives of the intensity function, we can cope with anisotropic voxel sizes. Based on [9], we implemented a scheme using cubic B-spline image interpolation and Gaussian smoothing. The derivatives are calculated based on the reconstructed continuous signal, taking anisotropic image resolution into account. In contrast, in the previous algorithm for landmark detection (e.g., [2],[5]) the detection operator $Op3$ was applied within a ROI of fixed size. Reduction of false detections by incorporating additional prior knowledge of the landmark was not considered.

5 Experimental Results Using 3D MR and CT Images

5.1 Experimental setting

Data. We used five T1-weighted MR/CT image pairs from different patients: one (*C06*) acquired at Utrecht University Hospital and four (*V101*, *V104*, *V107*, and *V109*) acquired at Vanderbilt University. The voxel sizes of the *C06* data are $0.86 \times 0.86 \times 1.2 \text{mm}^3$ (MR) and $0.63 \times 0.63 \times 1.0 \text{mm}^3$ (CT). The voxel sizes of the remaining original data are $0.85 \times 0.85 \times 3.0 \text{mm}^3$ (MR) and $0.42 \times 0.42 \times 3.0 \text{mm}^3$ (CT). Instead of the original data in the latter case, we used up-sampled images with a slice thickness of 1.0mm based on cubic B-spline interpolation [9].

Landmarks. We considered visually salient features located on the skull and within the brain: the saddle points at the zygomatic bones (MC15), the tip of the external occipital protuberance (MC5e), the topmost concavity of the fourth

ventricle (MC2), the junction at the upper end of pons (MC18), and the tips of the frontal (MC6) and occipital ventricular horns (MC7) (see Fig. 1). A suffix added to the landmark symbols indicates the respective hemisphere, e.g., MC15l refers to the saddle point at the left zygomatic bone.

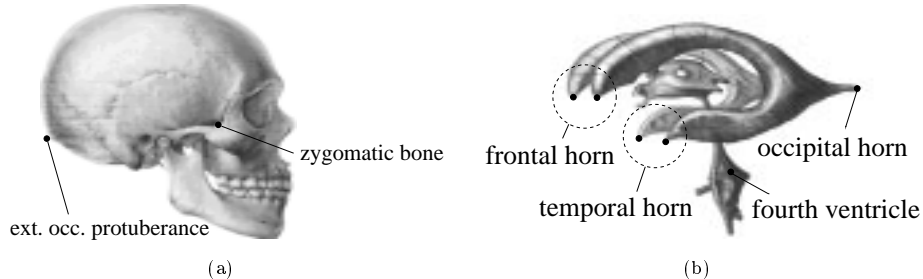


Fig. 1. Anatomical point landmarks located (a) on the skull (adapted from [10]) and (b) on the ventricular system (adapted from [11]).

Parameters. The minimal and maximal width of the cubic ROI was set to $w_{min} = 7$ voxels and $w_{max} = 21$ voxels, resp. The threshold for the spatial variation of the position estimate was set to $t_V = 0.5\text{mm}$. In the experiments with the previous algorithm for landmark detection, we used a fixed ROI width of $w = w_{max} = 21$ voxels. In estimating the derivatives, we used two different scales for the Gaussian filters, depending on the scale of the respective landmark: $\sigma = 1.0\text{mm}$ for MC15 as well as MC18 and $\sigma = 1.5\text{mm}$ for MC5e, MC2, MC6, and MC7 (note that these scales were automatically adapted according to the actual voxel size). Averaging the intensity gradient (matrix $\mathbf{C} = \overline{\nabla g \nabla g^T}$) was done within a $5 \times 5 \times 5$ window. Local maxima of the operator responses were determined in $3 \times 3 \times 3$ neighborhoods.

5.2 Evaluation of the Detection Performance

To evaluate the detection performance, we use a certain type of performance visualization as well as a scalar quantity measuring the detection capability [12]. Additionally, we consider the number of detections. The operator responses at the detected candidates are plotted as a function of the distance to the ROI center. Thus, in these plots the spatial scatter of the detected positions within the ROI, the number of detections, as well as the significance of the different detections (in terms of the strength of the operator response) are reflected. To quantitatively evaluate the detection performance, we use a measure that takes into account both the number of detections as well as the significance of the different detections. Suppose we have obtained n detections for a landmark. Let $R_i > 0$, $i = 1, \dots, n$, denote the operator responses at the detected positions

and let R_{max} denote the maximum of these values. We consider the measure

$$\psi = \begin{cases} 0 & n = 0, \\ \sum_{i=1}^n \frac{R_i}{R_{max}} & n \geq 1. \end{cases}$$

If we obtain only one correct detection, we have $\psi = 1$. Additional false detections with small operator responses give $\psi \approx 1$. In this case, the correct detection can clearly be distinguished. On the other hand, if there are detections with operator responses similar to the maximal operator response, ψ is much larger than 1. Thus, the closer ψ is to 1, the better the detection performance.

5.3 Analysis of the Efficacy in Reducing False Detections

Exemplarily, we consider six landmarks located within the mid-sagittal plane and within the left hemisphere. We compare the detection performance of the new algorithm for landmark detection with that of the previous algorithm (see Sect. 4). Here, we present in detail the results obtained for the *V109* MR/CT image pair. In the experiments, the ROI centers are the positions resulting from best manual landmark localization (in agreement of two persons). Using this ‘ground truth’, we can also study the impact of our approaches on the localization performance. No thresholds were applied to the operator responses. In Fig. 2 (MR) and Fig. 3 (CT), the detection performance is visualized as described in Sect. 5.2. Thick bars indicate the detections obtained with the new algorithm for landmark detection. Narrow bars indicate those detections that would additionally be obtained with the previous algorithm. Note that the new algorithm always yields a subset of those candidates obtained with the previous algorithm. Additionally, the number of detections n obtained with both algorithms as well as the detection performance measure ψ are given.

For both modalities, the new algorithm showed a significantly better detection performance than the previous algorithm. A larger number of false detections with significant operator responses were suppressed, e.g., in MR in the case of the occipital protuberance and the fourth ventricle (Figs. 2b,c) and in CT in the case of the fourth ventricle, the top of pons, and the left occipital horn (Figs. 3c,d,f). Note that in MR in the case of the fourth ventricle (Fig. 2c) and in CT in the case of the fourth ventricle and the top of pons (Figs. 3c,d) the previous algorithm yielded a number of false detections with larger operator responses than those candidates with minimal distance to the manual positions. These detections were rejected with the new algorithm. In almost all cases, the localization performance was not affected, i.e., the detections with minimal distance to the manual positions were retained. The fact that the detections obtained with the new algorithm are better distinguishable w.r.t. the operator response is also reflected by smaller values of ψ . Both approaches to reducing false detections complementarily improve the performance. For example, in MR in the case of the fourth ventricle (Fig. 2c) ROI size selection alone yielded two detections, while incorporation of prior knowledge of the landmark yielded four detections. The combination of both approaches yielded only one detection. For the other four MR/CT image pairs, we obtained similar results.

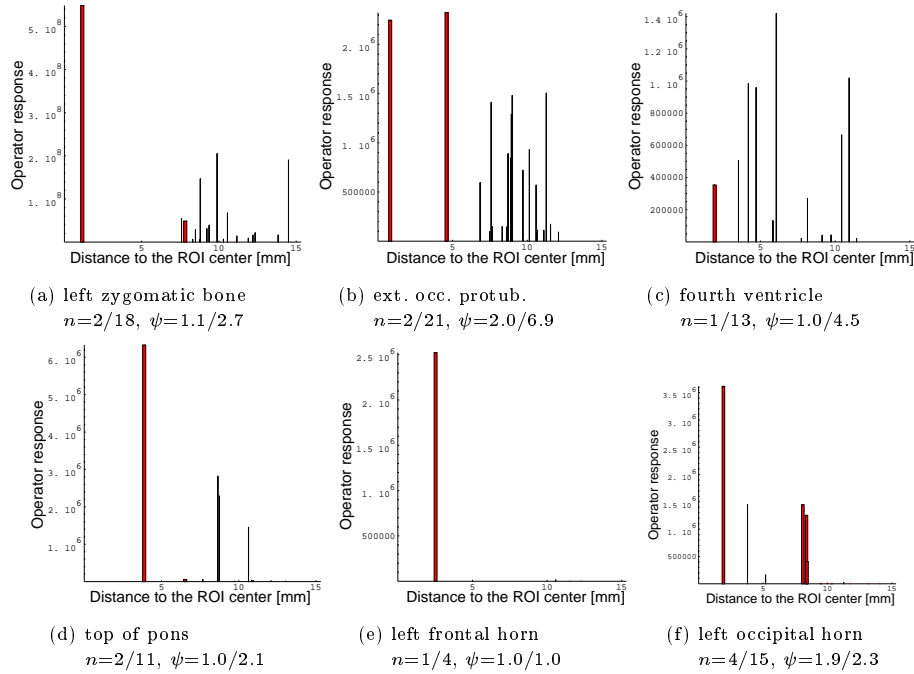


Fig. 2. Comparison of the detection performance of the new algorithm with that of the previous algorithm (*V109* MR image). The operator responses at the detected positions are drawn as a function of the distance to the ROI center. Thick bars indicate the remaining detections within the selected ROI obtained with the new algorithm. Narrow bars indicate those detections that would additionally be obtained with the previous algorithm. For both algorithms, the number of detections n as well as the detection performance measure ψ are given. For example, in (a) $n = 2/18$ means two detections for the new algorithm and 18 detections for the previous algorithm.

5.4 Application to Five MR/CT Image Pairs

We applied our new algorithm to extract landmarks from all five different MR/CT image pairs specified in Sect. 5.1. Depending on the field-of-view, the image quality, and lesions, we selected for each image pair 7–9 landmarks, which were simultaneously extracted from both modalities using the semi-automatic procedure. In the experiments, we applied a dynamic threshold to the operator responses to suppress insignificant detections (10% of the maximal operator response). The user inputs (e.g., the interactively determined positions and the selected candidates) were automatically recorded by the computer system. In Tables 1 and 2, the detection performance of the new algorithm is documented. Table 1 shows the results obtained for the five MR images, and Table 2 shows the results obtained for the five CT images (see Sect. 5.1 for the landmark symbols used for abbreviation). We compare the detection performance with that of the previous algorithm, which we applied at the same positions. For each landmark,

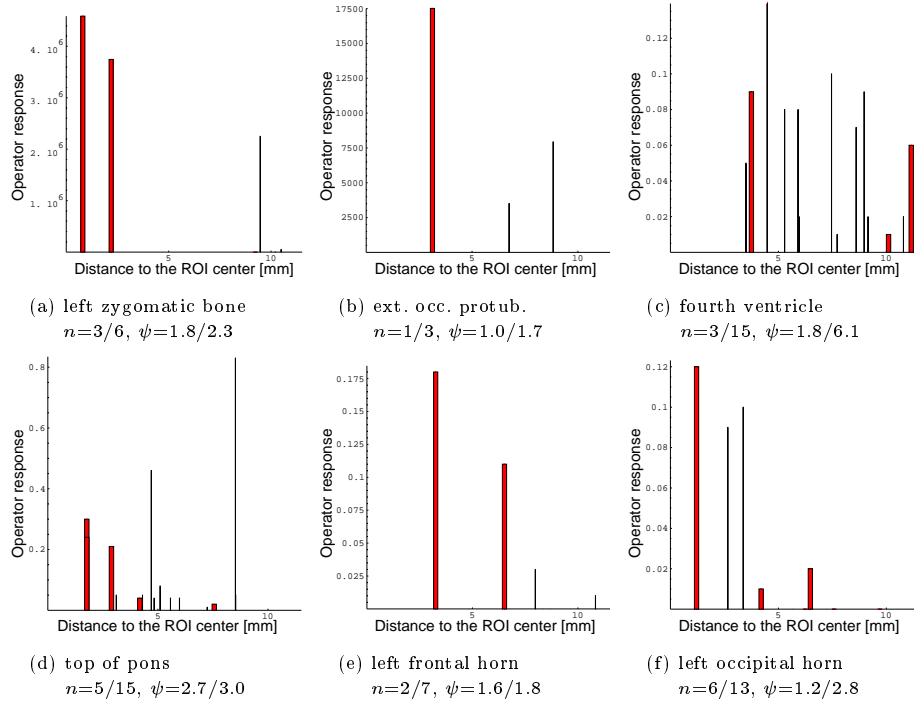


Fig. 3. Same as Fig. 2 but for the *V109* CT image.

we consider the mean \bar{n} and the maximal number n_{max} of detections, the relative difference $\Delta\bar{n}$ between these mean values, as well as the mean $\bar{\psi}$ and the maximum ψ_{max} of the detection performance measure ψ .

For all landmarks, the mean number of detections was reduced using the new algorithm. In most cases, we obtained significantly fewer detections. The values of ψ are significantly smaller, which indicates that the remaining detections are better distinguishable w.r.t. the operator responses. In MR, especially for the saddle points at the zygomatic bones (MC15), the occipital protuberance (MC5e), the fourth ventricle (MC2), and the top of pons (MC18) the results were significantly improved. In CT, for MC2 and MC18 we obtained significantly better results. In general, the detection performance was better for the CT images, i.e., the number of detections were smaller and ψ was smaller. In most cases, we obtained only 1–2 detections with the improved algorithm (in MR in 72% of all cases and in CT in 83% of all cases). In either case, we obtained not more than five detections for a landmark, while the previous algorithm yielded 15 detections in the worst case. Using the improved algorithm, in most cases the candidate with the maximal operator response was selected (in MR in 89% of all cases and in CT in 78% of all cases).

Recently, we also validated our improved algorithm for landmark detection within an application study in which five different observers participated [13].

Table 1. Comparison of the detection performance of the new algorithm with that of the previous algorithm (for five MR images). For each landmark, the mean \bar{n} and the maximal n_{max} number of detections, the relative difference $\Delta\bar{n}$ between these mean values, as well as the mean $\bar{\psi}$ and the maximum ψ_{max} of the detection performance measure ψ are given. For example, for MC15l $\bar{n} = 2.0/4.8$ denotes 2.0 detections in the mean for the new algorithm and 4.8 detections for the previous algorithm.

	MC15l	MC15r	MC5e	MC2	MC18	MC6l	MC6r	MC7l	MC7r
\bar{n}	2.0/4.8	3.3/8.3	1.8/8.3	1.0/12.0	1.0/5.0	2.5/3.0	2.4/3.4	1.8/2.8	2.0/3.4
n_{max}	3/8	5/12	2/13	1/15	1/7	3/4	3/5	3/4	3/5
$\Delta\bar{n}$	59%	61%	79%	92%	80%	17%	30%	36%	42%
$\bar{\psi}$	1.5/2.6	1.8/3.6	1.5/4.0	1.0/4.9	1.0/2.6	1.7/1.8	1.7/2.2	1.3/1.5	1.4/1.8
ψ_{max}	2.1/4.5	2.8/4.9	2.0/6.0	1.0/4.9	1.0/3.5	2.3/2.7	2.4/3.6	1.9/2.0	1.6/2.5

Table 2. Same as Table 1 but for five CT images.

	MC15l	MC15r	MC5e	MC2	MC18	MC6l	MC6r	MC7l	MC7r
\bar{n}	1.6/2.6	1.8/3.0	1.5/2.3	1.0/8.0	3.0/6.0	1.3/2.0	2.8/3.6	1.6/3.8	2.0/3.8
n_{max}	2/3	2/4	2/4	1/10	5/7	2/2	5/6	2/7	3/7
$\Delta\bar{n}$	39%	40%	35%	88%	50%	35%	23%	58%	48%
$\bar{\psi}$	1.3/1.9	1.6/2.1	1.2/1.4	1.0/3.5	1.7/2.7	1.0/1.2	1.9/2.2	1.3/2.1	1.5/2.0
ψ_{max}	1.8/2.3	2.0/2.8	1.5/2.0	1.0/3.9	2.6/3.3	1.2/1.2	3.6/3.9	1.9/2.6	1.8/2.9

The aim of this study was to evaluate the performance of the semi-automatic procedure in extracting landmarks from MR/CT images in comparison to a manual procedure (e.g., in terms of the elapsed time spent for landmark extraction as well as the registration results using a rigid transformation). The main findings were that (a) the elapsed time spent for landmark extraction can significantly be reduced with the semi-automatic procedure (the mean relative reduction of the elapsed time compared to a manual procedure was 38%) and (b) the registration results of both procedures generally showed similar quality.

6 Conclusion

Existing computational approaches to extracting 3D anatomical point landmarks from medical images often suffer from false detections, which may seriously affect the applicability and the acceptance of such an approach. A considerable number of these false detections is caused by neighboring anatomical structures. In this paper, we introduced two different approaches to reducing false detections caused by neighboring structures. Combining both approaches with a robust 3D differential operator for landmark detection, we presented an improved algorithm for landmark detection. In estimating the partial derivatives of the intensity function, we can cope with the typical case of anisotropic voxel sizes. We applied our new algorithm within a semi-automatic procedure to extract anatomical point landmarks from 3D MR and CT images of the human head. Experimental results showed that both automatically selecting the ROI size as well as incorporating

additional prior knowledge of the landmark are very effective in reducing the number of false detections. Results of applying our novel semi-automatic procedure for landmark extraction to five MR/CT image pairs demonstrated the applicability of our procedure.

Acknowledgments

Support of Philips Research Hamburg, project IMAGINE (IMage- and Atlas-Guided Interventions in NEurosurgery), is gratefully acknowledged. Images were provided as part of the E.U. AIM project COVIRA and as part of the project “Evaluation of Retrospective Image Registration”, National Institutes of Health, Project Number 1 R01 NS33926-01, Principal Investigator, Dr. J. Michael Fitzpatrick, Vanderbilt University, Nashville, TN.

References

1. J.-P. Thirion. Extremal Points: Definition and Application to 3D Image Registration. In *Proc. CVPR'94*, pp. 587–592. IEEE Computer Society Press, 1994.
2. K. Rohr. On 3D differential operators for detecting point landmarks. *Image and Vision Computing*, 15(3):219–233, 1997.
3. W. Beil, K. Rohr, and H.S. Stiehl. Investigation of Approaches for the Localization of Anatomical Landmarks in 3D Medical Images. In *Proc. CAR'97*, pp. 265–270. Elsevier Science, 1997.
4. T. Hartkens, K. Rohr, and H.S. Stiehl. Performance of 3D differential operators for the detection of anatomical point landmarks in MR and CT images. In *Proc. SPIE's Medical Imaging 1999: Image Processing*. vol. 3661, pp. 32–43. SPIE, 1999.
5. K. Rohr. Image Registration Based on Thin-Plate Splines and Local Estimates of Anisotropic Landmark Localization Uncertainties. In *Proc. MICCAI'98*, LNCS 1496, pp. 1174–1183. Springer-Verlag, 1998.
6. S. Frantz, K. Rohr, and H.S. Stiehl. Refined Localization of Three-Dimensional Anatomical Point Landmarks Using Multi-Step Differential Approaches. In *Proc. SPIE's Medical Imaging 1998: Image Processing*, vol. 3338, pp. 28–38. SPIE, 1998.
7. S. Frantz, K. Rohr, and H.S. Stiehl. Multi-step Procedures for the Localization of 2D and 3D Point Landmarks and Automatic ROI Size Selection. In *Proc. ECCV'98*, vol. I, LNCS 1406, pp. 687–703. Springer-Verlag, 1998.
8. L.M.J. Florack, B.M. ter Romeny, J.J. Koenderink, and M.A. Viergever. General Intensity Transformations and Differential Invariants. *Journal of Mathematical Imaging and Vision*, 4:171–187, 1994.
9. M. Unser, A. Aldroubi, and M. Eden. B-Spline Signal Processing: Part I—Theory. *IEEE Trans. on Signal Processing*, 41(2):821–833, 1993.
10. R. Bertolini and G. Leutert. *Atlas der Anatomie des Menschen*, volume 3: Kopf und Hals, Gehirn, Rückenmark und Sinnesorgane. Springer-Verlag, 1982.
11. J. Sobotta. *Atlas der Anatomie des Menschen*, volume 1: Kopf, Hals, obere Extremität, Haut. Urban & Schwarzenberg, 19th edition, 1988.
12. K. Rohr. Extraction of 3d anatomical point landmarks based on invariance principles. *Pattern Recognition*, 32(1):3–15, 1999.
13. S. Frantz, K. Rohr, H.S. Stiehl, S.-I. Kim, and J. Weese. Validating Point-based MR/CT Registration Based on Semi-automatic Landmark Extraction. In *Proc. CARS'99*. Elsevier Science, to appear.