Multi-Step Procedures for the Localization of 2D and 3D Point Landmarks and Automatic ROI Size Selection

Sönke Frantz, Karl Rohr, and H. Siegfried Stiehl

Universität Hamburg, Fachbereich Informatik
Arbeitsbereich Kognitive Systeme, D-22527 Hamburg, Germany
`frantz@informatik.uni-hamburg.de`

Abstract. In this contribution, we are concerned with the detection and refined localization of 3D point landmarks. We propose multi-step differential procedures for subvoxel localization of 3D point landmarks. Moreover, we address the problem of choosing an optimal size for a region-of-interest (ROI) around point landmarks. That is, to reliably localize the landmark position, on the one hand, as much as possible image information about the landmark should be incorporated. On the other hand, the ROI should be restricted such that other structures do not interfere. The multi-step procedures are generalizations of an existing two-step procedure for subpixel localization of 2D point landmarks. This two-step procedure combines landmark detection by applying a differential operator with refined localization through a differential edge intersection approach. In this paper, we investigate the localization performance of this two-step procedure for an analytical model of a Gaussian blurred L-corner. The results motivate the use of an analogous procedure for 3D point landmark localization. We generalize the edge intersection approach to 3D and combine it with 3D detection operators to obtain multi-step procedures for subvoxel localization of 3D point landmarks. The multi-step procedures are experimentally tested for 3D synthetic images and 3D MR images of the human head. We also propose an approach to automatically select an optimal ROI size. This approach exploits the uncertainty of the position estimate resulting from the edge intersection approach. We present first promising results for a real 2D image with different types of corners as well as for anatomical brain landmarks in 2D slices of a 3D MR image.

1 Introduction

The registration of medical images of the human head such as MR (Magnetic Resonance) and CT (X-Ray Computed Tomography) images is important, e.g., for the planning of neurosurgical interventions, radiotherapy, and therapy evaluation. One possibility to register two images is a point-based approach. In this case, prominent points, denoted also as point landmarks, have to be extracted from images. In general, the registration result heavily depends on the accurate localization of suitable point landmarks. Potential point landmarks of the

human head are salient tips, which can be found, for instance, on the ventricular system, the skull base, as well as on other anatomical structures. Usually, such 3D point landmarks are manually selected—a task which is difficult, time-consuming, and which often lacks accuracy. An alternative to manual selection is a semi-automatic approach to landmark localization. The advantage of such an approach is that the user has the possibility to interactively control the results, which is important in clinical applications. For example, first, an approximate position of a specific landmark is manually determined. Second, to extract potential landmark candidates, a computational approach is applied within a region-of-interest (ROI) around the approximate position. Third, the user selects the most promising candidate. The computational approach has to reliably and robustly detect as well as to accurately localize prominent points. Recently, 3D differential operators have been introduced [15, 14, 2] which are, however, only designed for the detection of 3D point landmarks. Since reliable and robust landmark detection generally requires operators with large size, the accuracy of the detected points often is not satisfactory and additional steps are necessary to obtain better position estimates.

In the first part of this paper, we propose differential-based multi-step procedures for refined localization of 3D point landmarks. Our multi-step procedures are based on a two-step procedure for localizing point landmarks in 2D images [5]. This two-step procedure combines landmark detection by applying a 2D differential operator with refined localization through a 2D differential edge intersection approach. This two-step procedure only employs first-order partial derivatives of the image function and therefore does not suffer from potential instabilities of computing high-order partial derivatives. In this contribution, first, we investigate the localization performance of the two-step procedure for a 2D analytical model of a Gaussian blurred L-corner. We then generalize the 2D edge intersection approach to 3D. By combining the 3D edge intersection approach with 3D differential operators for landmark detection we propose multi-step procedures for refined localization of 3D point landmarks. We present experimental results for 3D synthetic images as well as 3D MR images of the human head.

In the second part of the paper, we address the problem of choosing an optimal size of the ROI. Obviously, the chosen ROI should be adapted to the scale of a landmark. On the one hand, it is reasonable to use as much as possible information about the structure containing the landmark to reliably localize the position of the landmark. On the other hand, it is clear that the ROI should not include other structures. In this paper, we introduce a new approach to automatically select the optimal size of a ROI. Our approach is based on the uncertainty of the position estimate resulting from the edge intersection approach. A comparable approach has been proposed in [7]. However, there a different detection operator has been used and also the criterion for ROI size selection is different. We show for a real 2D image that our approach allows the extraction of ROIs with optimal size for various image structures such as L-, Y-, and Arrow-corners. Results for a number of anatomical landmarks in 2D slices of a 3D MR image also demonstrate the suitability of our approach for medical images.

2 Multi-Step Procedures for Point Landmark Localization Using an Edge Intersection Approach

2.1 Two-step procedure for localizing 2D point landmarks

Förstner and Gülch [5] proposed a two-step procedure for localizing 2D point landmarks (corners) which combines landmark detection by applying a 2D differential operator with refined localization through a differential edge intersection approach.

Landmark detection The used differential operator [4,8] exploits the matrix

$$\mathbf{N} = \begin{pmatrix} \sum_i g_{x_i}^2 & \sum_i g_{x_i} g_{y_i} \\ \sum_i g_{x_i} g_{y_i} & \sum_i g_{y_i}^2 \end{pmatrix}.$$

This matrix represents (up to a factor) the averaged dyadic product of the intensity gradient. The subscripts x and y of the intensity function $g(x,y)$ stand for the partial derivatives in the respective spatial direction and the sum index i denotes the spatial location. The detection operator reads

$$F(\mathbf{x}) = \frac{det(\mathbf{N})}{tr(\mathbf{N})}, \tag{1}$$

where $det(\cdot)$ denotes the determinant and $tr(\cdot)$ the trace of a matrix. Each point $\mathbf{x} = (x, y)$ in the image is being assigned the measure in (1), where $\mathbf{N}$ is usually computed in a symmetric and quadratic window of certain size around that point. Point landmarks are detected by searching for local maxima of $F(\mathbf{x})$.

Refined localization We consider the neighborhood around a detected point. Suppose an L-corner has been detected and the observation window captures sufficient edge information of the structure. For simplicity, a local coordinate system with the detected point as origin is chosen. For each edge point in the observation window we define a tangent locally approximating the edge direction. For this, the intensity gradients are taken as normals to the tangents. The tangents are represented in the Hessian normal form, that is, for a point $\mathbf{x}_i$ with intensity gradient ∇g_i we have $< \nabla g_i, \mathbf{x} > = < \nabla g_i, \mathbf{x}_i >$, where $< \cdot, \cdot >$ denotes the inner product. Rewriting this as $\varepsilon_i(\mathbf{x}) = < \nabla g_i, \mathbf{x} - \mathbf{x}_i >$ yields the perpendicular distance from $\mathbf{x}$ to the tangent at $\mathbf{x}_i$. Note that the distance is implicitly multiplied with the gradient magnitude since ∇g_i generally is not a unit vector. The position of the tip can be estimated through intersection of all tangents using the least-squares method, that is, through minimization of the residual error function

$$E(\mathbf{x}) = \sum_i \varepsilon_i(\mathbf{x})^2. \tag{2}$$

So far, solely edge points have been considered. However, due to implicitly weighting the residual errors with the gradient magnitude, it is possible to give

up this restriction and to include in the sum in (2) all points within the observation window. Points in rather homogeneous regions with low gradient magnitude should hardly contribute to the sum anyway. On the other hand, edge points with generally high gradient magnitude should actually force a small distance from the position estimate to their corresponding tangent. Finally, the condition $\nabla E(\mathbf{x}) = \mathbf{0}$ yields the system of normal equations

$$\mathbf{N}\mathbf{x}^* = \mathbf{y}, \tag{3}$$

where $\mathbf{x}^*$ denotes the position of the tip w.r.t. the local coordinate system, $\mathbf{N}$ is the matrix from above, and

$$\mathbf{y} = \begin{pmatrix} \sum_i g_{x_i}^2 x_i + \sum_i g_{x_i} g_{y_i} y_i \\ \sum_i g_{x_i} g_{y_i} x_i + \sum_i g_{y_i}^2 y_i \end{pmatrix}.$$

2.2 Analytical study for an L-corner

It is known that for an L-corner the detection operator in (1) yields systematic localization errors w.r.t. the correct corner position. The error generally depends on the blur, the aperture angle of the structure, and the size of the operator (see [13] for an analytical study). Hence, in general, additional steps are necessary to improve the localization accuracy. In the following, we analyze the localization performance of the two-step procedure for an analytical model of a Gaussian blurred L-corner [11] and compare the achieved accuracy with that resulting from the detection operator alone.

In [13], the positions for the detection operator alone have been determined. We have centered the observation window for the edge intersection approach at the position determined with the detection operator. Then, we have computed the matrix $\mathbf{N}$ and the vector $\mathbf{y}$ and have solved the system of normal equations in (3) to obtain the position estimate.

In Fig. 1a, the distances e from the localized positions to the correct corner position are depicted in dependence on the aperture angle β. The standard deviation of the Gaussian blur function has been set to $\sigma = 1$. The solid line results from the detection operator alone and the dashed line results from the two-step procedure using the edge intersection approach. The size of the observation window for the edge intersection approach has been the same as that for the detection operator, that is, 3×3 pixels. It can be seen that the edge intersection approach yields an improvement of about $1pix$ (pix denotes spatial unity) for a large range of aperture angles. Moreover, we can obtain even better position estimates if we further enlarge the observation window for the edge intersection approach. The result for a window size of 15×15 pixels is shown in Fig. 1b. The localization error is thus reduced to nearly zero. We conclude that a reasonable strategy for improving the localization accuracy is to capture with the observation window as much as possible information of the intensity structure. On the other hand, it should be guaranteed that no other structures interfere. As a second main topic of this paper, we will report on optimal size selection for a region-of-interest (ROI) around point landmarks in Section 3.
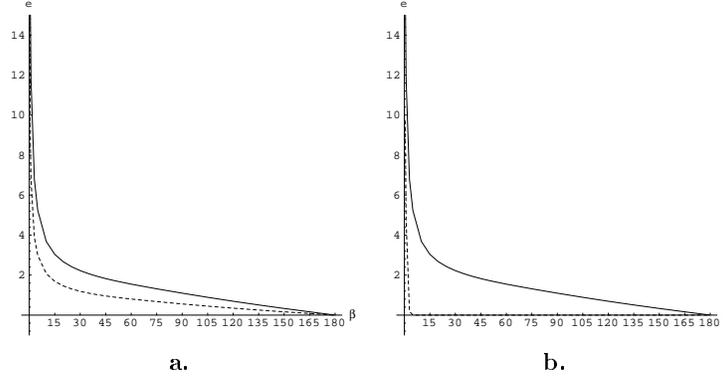
Fig. 1a and b. Localization accuracy for a Gaussian blurred L-corner with $\sigma = 1$. The distances e from the localized positions to the correct corner position are depicted in dependence on the aperture angle β. The solid lines result from the detection operator alone and the dashed lines result from the two-step procedure. Observation windows of sizes 3×3 pixels (a) and 15×15 pixels (b) have been used.

2.3   Extension of the edge intersection approach to 3D

In the 2D case, we have dealt with an L-corner to motivate the edge intersection approach. A 3D generalization of an L-corner is a tetrahedron. Suppose we have obtained an approximate position of the tip, for instance, by applying a 3D detection operator, and have placed there an observation window capturing sufficient image information. The position of the tip of the tetrahedron is the intersection point of three plane surfaces which correspond to 3D edges. Analogously to the 2D case, we locally approximate the surfaces of a landmark through tangent planes. The intensity gradients are taken as normals to the tangent planes. The Hessian normal form of the tangent plane at a point $\mathbf{x}_i$ reads $< \nabla g_i, \mathbf{x} > = < \nabla g_i, \mathbf{x}_i >$. The position of the tip can be estimated through intersection of all tangent planes using the least-squares method, that is, through minimization of a residual error function which formally agrees with that in (2) of the 2D edge intersection approach. Note that also here the gradient magnitudes implicitly weight the residual error function. We thus obtain a system of normal equations

$$\mathbf{N}\mathbf{x}^* = \mathbf{y}, \tag{4}$$

where $\mathbf{x}^*$ denotes the estimated subvoxel position of the tip and

$$\mathbf{N} = \begin{pmatrix} \sum_i g_{x_i}^2 & \sum_i g_{x_i} g_{y_i} & \sum_i g_{x_i} g_{z_i} \\ \sum_i g_{x_i} g_{y_i} & \sum_i g_{y_i}^2 & \sum_i g_{y_i} g_{z_i} \\ \sum_i g_{x_i} g_{z_i} & \sum_i g_{y_i} g_{z_i} & \sum_i g_{z_i}^2 \end{pmatrix}$$

and

$$\mathbf{y} = \begin{pmatrix} \sum_i g_{x_i}^2 x_i + \sum_i g_{x_i} g_{y_i} y_i + \sum_i g_{x_i} g_{z_i} z_i \\ \sum_i g_{x_i} g_{y_i} x_i + \sum_i g_{y_i}^2 y_i + \sum_i g_{y_i} g_{z_i} z_i \\ \sum_i g_{x_i} g_{z_i} x_i + \sum_i g_{y_i} g_{z_i} y_i + \sum_i g_{z_i}^2 z_i \end{pmatrix} .$$

2.4 Multi-step procedures for localizing 3D point landmarks using the 3D edge intersection approach

In [14], 3D differential operators for detecting point landmarks have been proposed. The operators exploit the matrix $\mathbf{N}$ in (4) (up to a factor) and read $Op3(\mathbf{x}) = det(\mathbf{N})/tr(\mathbf{N})$, $Op3'(\mathbf{x}) = 1/tr(\mathbf{N}^{-1})$, and $Op4(\mathbf{x}) = det(\mathbf{N})$. The operators are 3D extensions of existing 2D corner detectors in [4, 8, 10, 11]. In this contribution, we propose three multi-step procedures for refined point landmark localization in 3D images, combining the 3D detection operators with the 3D edge intersection approach.

i) **Two-step procedure**

First, points are detected with either $Op3$, $Op3'$, or $Op4$, where a large operator size is chosen for reasons of robustness w.r.t. noise in images. Second, to refine the positions, a small operator size is chosen and the respective differential operator is applied within a small neighborhood around the detected points. A similar procedure for 2D point landmark localization was proposed earlier [3, 10].

ii) **Two-step procedure with subvoxel localization**

First, points are detected with either $Op3$, $Op3'$, or $Op4$. Second, the positions are refined through the 3D edge intersection approach. This scheme essentially is the 3D extension of the two-step procedure of Förstner and Gülch [5].

iii) **Three-step procedure with subvoxel localization**

This procedure is a combination of the procedures i) and ii) and is therefore a three-step procedure. The first two steps correspond to the two-step procedure i). In the third step, the position estimates resulting from i) are further refined through the 3D edge intersection approach.

2.5 Experimental results

We report on experiments for 3D synthetic images as well as 3D MR images of the human head. The partial derivatives have been estimated with 3D extensions of the 2D filters of Beaudet [1]. For the detection step we have used a filter size of $5 \times 5 \times 5$ voxels. For the refinement steps of the procedures i) and iii) filters of size $3 \times 3 \times 3$ voxels have been used. The components of the matrix $\mathbf{N}$ and the vector $\mathbf{y}$ in (4) are the averaged values of products of the partial derivatives within an observation window. To study the localization accuracy of the different procedures in dependence on the width w of the observation window, we have investigated various window widths, starting with $w = 3$ voxels. The size of the observation window for the 3D edge intersection approach has been the same as that for the detection operator. Maxima of the detection operator responses have been determined by local maximum search in neighborhoods of $3 \times 3 \times 3$ voxels. In the case of several maxima within the chosen ROI of the image, we have selected the candidate with the largest operator response. To alleviate subjectivity in our experiments we have not used any thresholds on the operator responses. However, a few cases in the experiments on medical images have required special attention.

3D synthetic images We have investigated the localization accuracy of the procedures i), ii), and iii) for the tip of a 3D tetrahedron. This structure is a 3D generalization of an L-corner with aperture angle β. For the construction of the 3D object in Fig. 2a, where the tip is marked through a black dot, the symmetric axis of the L-corner has been spread into the direction of the z-axis such that it also encloses the angle β with the x-axis. The shape of the so constructed tetrahedron is determined by the aperture angle β.

In Fig. 2b, for a tetrahedron with $\beta = 90°$ the Euclidean distances e from the localized positions to the position of the tip are depicted in dependence on the width w (in voxels) of the observation window. DET denotes the detection operator $Op3'$. We see that applying the detection operator alone yields the worst results. Procedure i) is only for $w = 5$ better than DET. The position estimates resulting from the procedures ii) and iii) are significantly more accurate than those resulting from DET and procedure i). The accuracy gets worse for DET and procedure i) if w increases. By contrast, the accuracy for the procedures ii) and iii) gets better if w increases. In Fig. 2c, the results for a much more tapered tetrahedron with $\beta = 45°$ are shown. The results are comparable with those for the tetrahdron with $\beta = 90°$. However, for the procedures ii) and iii) the localization error slightly increases for $w \geq 11$ and $w \geq 7$, resp. We suspect that discretization errors give rise to this effect.
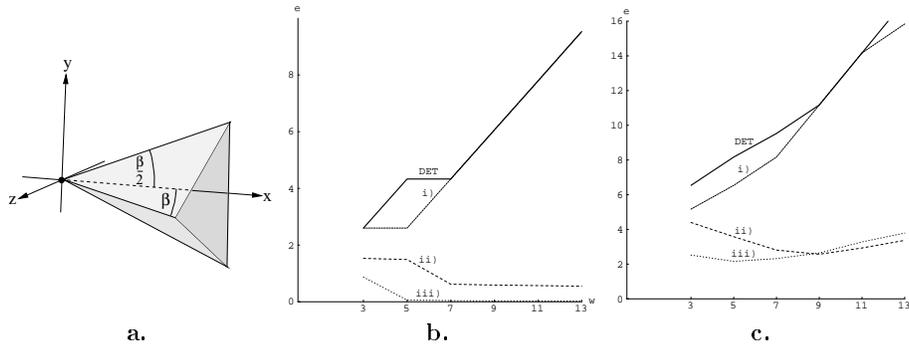
Fig. 2a. Tetrahedron with aperture angle β. **b and c.** Localization accuracy for tetrahedrons with $\beta = 90°$ and $\beta = 45°$, resp. The Euclidean distances e from the localized positions to the position of the tip are depicted in dependence on the width w of the observation window.

3D MR images Here, we present results for three 3D MR images of the human head. We consider as anatomical landmarks the tips of the frontal, occipital, and temporal horns of the ventricular system, abbreviated by MC6, MC7, and MC13, resp. The tips are indicated in Fig. 3a through black dots within dashed circles. The letters 'l' and 'r' denote the respective hemispheric part. We have manually specified the positions of these landmarks in the investigated data sets and have taken them as 'ground truth' positions, although we know that manual localization of 3D landmarks generally is difficult and may be prone to error. Note
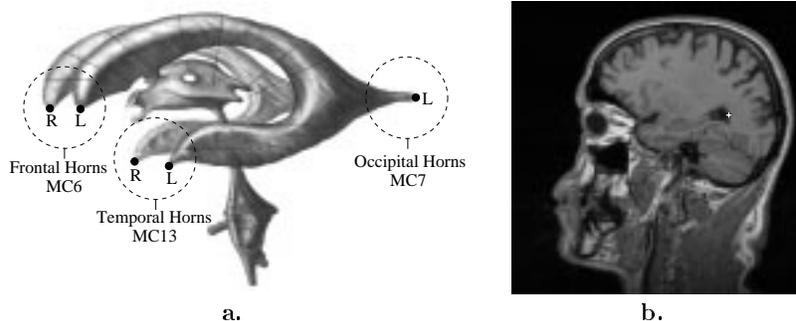
Fig. 3a. Anatomical preparate of the ventricular system of the human head (adapted from Sobotta, 1988) **b.** Off-sagittal slice of MR image 1. The manually specified position of the tip of the left occipital horn is marked through a white cross.

also that manually we have only determined voxel positions, while the multi-step procedures ii) and iii) use the 3D edge intersection approach and hence yield subvoxel positions. We have respectively chosen a region-of-interest (ROI) around the 'ground truth' positions and then have applied the computational approaches described above. For landmark detection and the refinement steps of the procedures i) and iii) we here have used the operator $Op3$. We have used the same filter widths to estimate the partial derivatives as in the experiments for the 3D synthetic images. The size of the observation window for the 3D edge intersection approach has been the same as that for the detection operator. To alleviate subjectivity, we have not used any thresholds on the detection operator responses. In the case of several detected points within the ROI, we have selected that point with the largest operator response. The thus selected candidates have been visually inspected for validity, according to the semi-automatic approach described in Section 1.

In Figs. 4a and b, the mean values $\bar{e}$ for both $w = 3$ and $w = 5$ of the Euclidean distances from the localized positions to the manually specified positions are depicted according to each landmark and according to each MR image, resp. In general, the procedures ii) and iii) yield the most accurate positions. Also procedure i) generally yields better positions than the detection operator alone, although not as good as ii) and iii). Basically, we have selected the candidates with the largest operator responses. The visual inspection of the detected candidates for the left and right temporal horns MC13l,r in MR image 2 and the left temporal horn in MR image 3 has been very difficult since in these images the temporal horns are extremely poorly pronounced. Also, the positions of these landmarks resulting from manual localization are rather uncertain. Therefore, for these images the respective landmarks have not been considered. For the right temporal horn MC13r in MR image 3 two points with extremely high operator responses have been detected. Visual inspection has revealed that the respective candidates with the highest operator responses are false detections and has caused us instead to select the candidates with the second-highest operator responses.

All in all, for the detection operator alone the mean of the Euclidean dis-
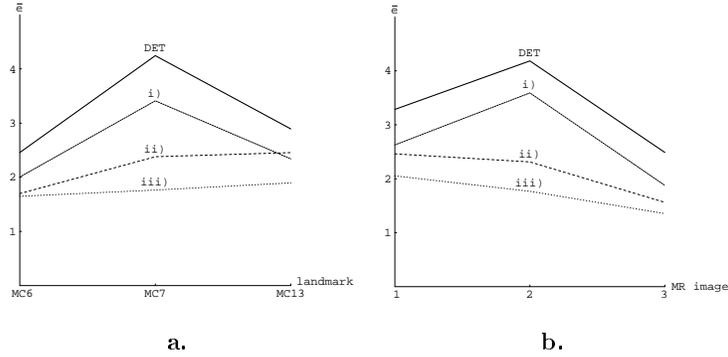
Fig. 4a and b. Localization accuracy of the detection operator alone (DET) and the multi-step procedures i), ii), and iii) for the ventricle landmarks in the investigated three MR images. The mean values $\overline{e}$ for both $w = 3$ and $w = 5$ of the Euclidean distances from the localized positions to the manually specified positions are depicted according to each landmark (a) and according to each MR image (b).

tances from the localized positions to the manually specified positions amounts to $3.27\,vox$ (vox denotes spatial unity). The two-step procedure i) improves the accuracy w.r.t. DET by $0.59\,vox$. Additionally using the 3D edge intersection approach further improves the accuracy w.r.t. i) by $0.93\,vox$. Thus, the three-step procedure iii) improves the accuracy w.r.t. DET by $1.52\,vox$. The two-step procedure ii) improves the accuracy w.r.t. DET by $1.14\,vox$.

To give a visual impression of the localization capabilities of the different procedures, we also show in Fig. 5 orthogonal image cuts at the respectively localized positions for the tip of the left occipital horn in MR image 1 ($w = 5$). The location of the landmark within the human head is marked through a white cross in a sagittal view in Fig. 3b. Note that, due to technical reasons, the subvoxel coordinates resulting from the multi-step procedures ii) and iii) have been rounded to voxel coordinates. Nevertheless, it can be seen that the multi-step procedures ii) and iii) yield the best results.

3   Automatic ROI Size Selection for Point Landmarks

The performance of the semi-automatic approach for landmark localization described in Section 1 substantially depends on the size of the region-of-interest which generally should be adapted to the scale of a specific landmark. On the one hand, the ROI should capture as much as possible information about the landmark. On the other hand, the ROI should not include neighboring structures. In the following, we introduce a novel approach for automatically determining an optimal size for the ROI. The approach is based on the uncertainty of the position estimate resulting from the edge intersection approach.
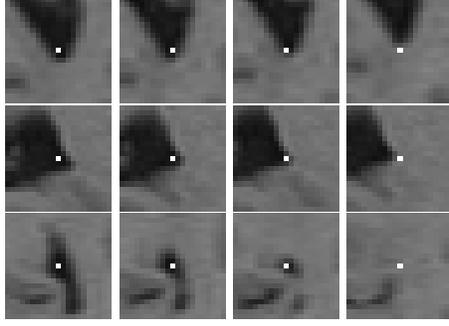
Fig. 5. Localized positions at the tip of the left occipital horn in MR image 1 ($w = 5$) for the detection operator alone and the multi-step procedures i), ii), and iii) (from left to right) in axial, sagittal, and coronal views (from top to bottom).

3.1 Localization uncertainty of the edge intersection position estimate

The covariance matrix of the position estimate $\mathbf{x}^*$ in (3) (resp. (4) in the 3D case) is given by

$$\boldsymbol{\Sigma} = \sigma_\varepsilon^2 \mathbf{N}^{-1}, \tag{5}$$

under the assumption that the residuals ε_i are independently and normally distributed random errors with zero mean and variance σ_ε^2. We estimate the variance through $\sigma_\varepsilon^2 = \mathrm{E}(\mathbf{x}^*)/(n-2)$ (2D) and $\sigma_\varepsilon^2 = \mathrm{E}(\mathbf{x}^*)/(n-3)$ (3D), where $\mathrm{E}(\cdot)$ is the respective residual error function and n stands for the number of points in the chosen observation window (see also [5]). In the 2D case, the localization uncertainty of the position estimate is represented by the covariance ellipse (covariance ellipsoid in the 3D case). Evidently, the size of the covariance ellipse (ellipsoid) should be small, indicating low localization uncertainty of the position estimate. In the 2D case, the Eigenvalues of the covariance matrix, λ_1 and λ_2, are the squared lengths of the the semi-axes of the covariance ellipse. A scalar measure for the localization uncertainty of the position estimate is the product of the Eigenvalues $U = \lambda_1 \lambda_2$, which is actually the squared area of the covariance ellipse. Analogously, a scalar measure for the localization uncertainty in 3D is $U = \lambda_1 \lambda_2 \lambda_3$, where λ_i denote the Eigenvalues of the 3D covariance matrix and represent the squared lengths of the semi-axes of the ellipsoid. That is, in the 3D case, U is the squared volume of the covariance ellipsoid. Equivalently, when using the determinant of the covariance matrix, we can write in both cases

$$U = det(\sigma_\varepsilon^2 \mathbf{N}^{-1}). \tag{6}$$

3.2 Localization uncertainty in dependence on the ROI size

Suppose we have an initial position for a landmark around which we place a symmetric observation window. Let us define an uncertainty function $U(w)$ depending on the width w of the ROI. Then, for example, $U(5)$ gives the localization uncertainty of the position estimate for an observation window of size

5×5 pixels ($5 \times 5 \times 5$ voxels in the 3D case). Assume that further enlarging the observation window gives us more edge information about the landmark but other structures are still excluded. In this case, we do not expect a significant signal change in $U(w)$. On the other hand, additionally capturing other structures by the ROI gives rise to a more or less strong increase of $U(w)$ which tells us that further enlargement of the ROI is not useful. The respective value of w then gives us the width of a suitable ROI.

Let us now see how this works in practice for the 2D case. We consider an L- and a Y-corner in a real image containing a polyhedral object as shown in Fig. 6. In accordance to the semi-approach procedure for landmark localization, we have manually determined the corner positions. The manually selected corner positions are marked by a small square for the L-corner in Fig. 6a and the Y-corner in Fig. 6b. We have placed symmetric observation windows of width $w = 5$ pixels around that points. The widths of the windows have been successively enlarged by 2 pixels, that is, 1 pixel on each side. In Fig. 7a, for the L-corner the localization uncertainty in (6) in dependence on the width of the ROI is depicted. In Fig. 7b, for the Y-corner the uncertainty measure in dependence on the width of the ROI is depicted (please ignore for a moment the vertical lines which have been added to the graphs). We see that for the L-corner the uncertainty function

a.   b.
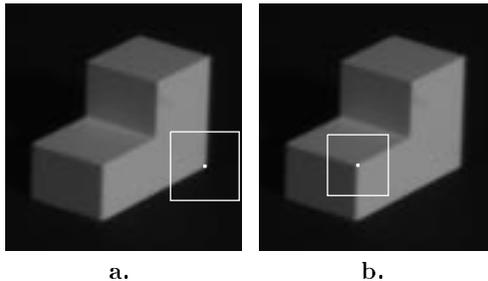
Fig. 6a and b. ROI size selection for two differently complex structures in a real image of a polyhedral object. The, based on the uncertainty measure U, manually selected optimal ROIs are drawn for an L-corner and a Y-corner, resp. The also manually specified corner positions (centers of the ROIs) are marked by small squares (see text).

continuously decreases with increasing width of the ROI which is what we expect. Other structures do not interfere. In such a case, it is reasonable to choose the ROI as large as possible. The width is only restricted due to the image size. The thus manually determined ROI ($w = 71$) is drawn in Fig. 6a. For the Y-corner we note interferences of other edges when the observation window exceeds a certain size. This is reflected through a strong increase of the uncertainty function, beginning at $w = 63$. Therefore, a maximal ROI with $w = 63$ seems suitable. An observation window of this size is drawn in Fig. 6b.

3.3 Determining ROIs with optimal size

According to the above discussion, there are two principal optimality criteria w.r.t. the size of the ROI. We speak of 'criterion A' if we mean optimality of the ROI w.r.t. minimal localization uncertainty. We refer with 'criterion B' to optimality w.r.t. maximal size of the ROI. Thus, depending on the need of the user, we suggest the following procedure to select a ROI with optimal size: First, determine a ROI with maximal size (criterion B). Then, if desired, reduce the ROI by selecting the corresponding width of the minimal value of $U(w)$ (criterion A).

We have to detect a strong increase of the uncertainty function (criterion B). Signal changes in $U(w)$ are generally accompanied by a variation of the position estimate resulting from the edge intersection approach. We have found that the variation of the position estimate is a strong criterion to decide whether other structures interfere due to enlargement of the observation window. In our case, we compute the distance between the position estimates $\mathbf{x}_w^*$ (for the observation window width w) and $\mathbf{x}_{w-2}^*$ (for the observation window width $w-2$) and define a function

$$D(w) = |\mathbf{x}_w^* - \mathbf{x}_{w-2}^*| \tag{7}$$

which is compared with a threshold t_D prescribing to what extent variations of the position estimate are allowed. Additionally, we use a second criterion which measures the increase of uncertainty. In our case, we simply compare successive values of $U(w)$ which has given good results. Alternative measures are the 'derivative' or the 'curvature' of $U(w)$. Altogether, we propose to abort ROI enlargement if a) $U(w+2) > U(w)$ (increase of uncertainty) and b) $D(w+2) \geq t_D$. If these conditions are both fulfilled, then the value of w gives us the optimal size of the ROI according to criterion B.

3.4 Resulting algorithm for automatic ROI size selection

To sum up, our algorithm for ROI size selection consists of the following four steps:

1. Determine an initial position for the searched landmark. Place a symmetric and quadratic observation window of width $w = w_{min}$ around that point, for example, $w_{min} = 5$ pixels. Note that w_{min} determines the minimal width of the ROI. Compute the uncertainty measure $U(w)$ in (6) for the position estimate.

2. If further enlargement of the observation window is not possible, for instance, due to user restriction or image size, then either stop (criterion B with optimal width $w_{opt} = w$) or go to step 4 (criterion A, then let $w_B = w$ denote the maximal width of the ROI). Otherwise enlarge the observation window: $w = w + 2$.

3. Compute $U(w)$ and the variation of the position estimate $D(w)$ in (7). If a) $U(w) > U(w-2)$ and b) $D(w) \geq t_D$, then either stop (criterion B,

$w_{opt} = w - 2$) or goto step 4 (criterion A, then let $w_B = w - 2$ denote the maximal width of the ROI). Otherwise continue with step 2.

4. Select the corresponding width ($w_{min} \leq w \leq w_B$) for which $U(w)$ attains its minimum, that is, minimal localization uncertainty for the position estimate resulting from the edge intersection approach. Set $w_{opt} = w$.

3.5   Experimental results

We present results for real 2D images. In our experiments, we have used Gaussian filters for derivative estimation. If not stated otherwise, then the standard deviation $\sigma = 1.0$ has been chosen. The minimal width of the ROI has always been set to $w_{min} = 5$. The initial positions (ROI centers) have been selected manually. In all experiments the threshold for the variation of the position estimate has been fixed to $t_D = 0.5 pix$.

First, we have applied the algorithm on the image containing a polyhedral object. In Fig. 7, the uncertainty measure U in dependence on the width of the observation window is depicted for an L-corner (a), Y-corner (b), and Arrow-corner (c). In these graphs, the automatically determined optimal ROI widths are marked by a solid line for criterion A (minimal localization uncertainty of the position estimate) and by a dashed line for criterion B (maximal size of the ROI). The automatically determined ROIs are shown in Fig. 8a for criterion A and in Fig. 8b for criterion B. The loci of the position estimates resulting from the edge intersection approach are indicated through white crosses, although, due to technical reasons, the coordinates have been rounded to pixel coordinates. We see that for citerion B we obtain ROIs with more or less maximal size.

a.   b.   c.
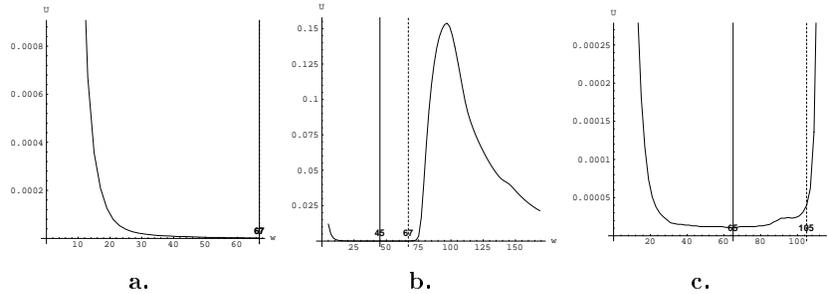
Fig. 7. Uncertainty measure U in dependence on the window width w for the L-corner (a), Y-corner (b), and Arrow-corner (c). The automatically selected ROI widths are marked by a solid vertical line for criterion A and by a dashed vertical line for criterion B (see text).

The ROIs capture no other structures. The algorithm stops nearby the obvious strong increases of $U(w)$ ($w_{opt} = 105$ for the Arrow-corner and $w_{opt} = 67$ for the Y-corner). However, for the L-corner no other edges interfere if the observation window is enlarged. Thus for the L-corner in both cases we have $w_{opt} = 67$. However, for criterion A we obtain for both the Arrow- and the Y-corner a significantly smaller ROI.
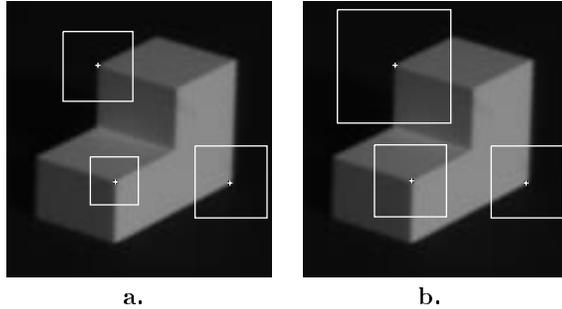
Fig. 8. Automatically determined ROIs for criterion A (a) and and criterion B (b). The position estimates resulting from the edge intersection approach are marked by white crosses (see text).

To analyze the localization accuracy of the position estimates resulting from the edge intersection approach in more detail, we have compared the here calculated positions with those obtained by fitting a parametric corner model to the data [12]. Using this 'ground truth' we have computed the respective distances to the position estimates. We have varied the scale σ of the Gaussian filters for estimating the partial derivatives to study the localization error in dependence on the scale. The distances in dependence on σ are listed for criterion A in Tab. 1a and for criterion B in Tab. 1b. To show the relation between the localization error and the size of the observation window, we have also listed the respective value w_{opt} (in parentheses) for each scale σ. From Tabs. 1a and b we see that the

σ	Arrow	L	Y
0.5	0.12 (31)	0.33 (61)	0.27 (33)
1.0	0.11 (65)	0.16 (67)	0.28 (45)
1.5	0.25 (79)	0.13 (65)	0.32 (51)

a.

σ	Arrow	L	Y
0.5	0.96 (105)	0.39 (71)	0.49 (67)
1.0	0.99 (105)	0.16 (67)	0.52 (67)
1.5	1.26 (105)	0.13 (65)	0.46 (67)

b.

Table 1. Localization error (in *pix*) of the position estimates resulting from the edge intersection approach for the Arrow-, L-, and Y-corner for criterion A (a) and criterion B (b). The values in parentheses denote the width of the determined ROI and σ denotes the respective scale of the Gaussian filters for estimating the partial derivatives (see text).

accuracy of the positions for criterion A is always better than or equal to that for criterion B. On the other hand, the selected ROIs are, with except for the L-corner, significantly smaller. A general statement concerning the best suited filter scale σ is not possible. However, we note that for criterion A we generally obtain a larger ROI with increasing scale. On the other hand, with increasing scale, for the Arrow- (criteria A and B) and Y-corner (criterion A), the localization error increases, whereas that for the L-corner decreases. We assume that the more complex structure of the Arrow- and Y-corner is the reason for this effect.

As a second example, we have investigated 2D slices of a 3D MR image of the

human brain. In Fig. 9a, for criterion A the automatically determined ROIs for several anatomical landmarks, namely the genu of corpus callosum (most left), top of pons (middle), and the tip of fourth ventricle (most right) are shown. In Fig. 9b, for criterion A the ROIs for the tip of the frontal horn (most left) and the tip of a gyrus (top) are shown. In Figs. 9c and d, for criterion B the ROIs are shown. The position estimates resulting from the edge intersection approach are marked by white crosses. The determined ROIs have nearly the same sizes for

a.   b.

c.   d.
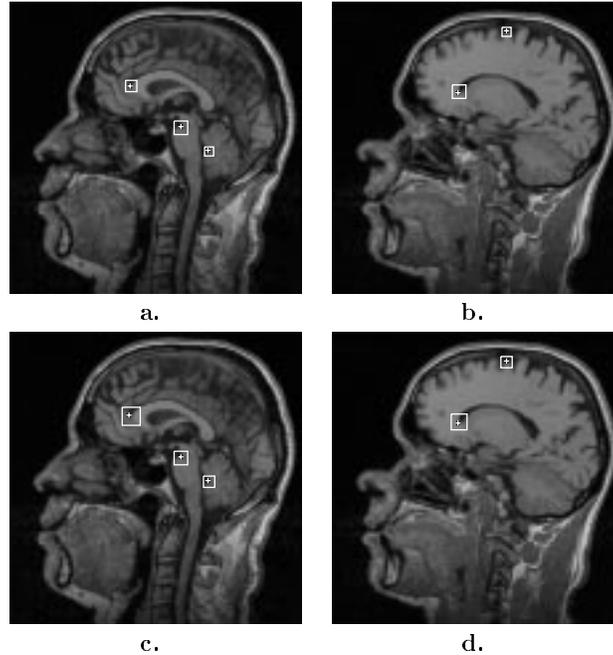
Fig. 9. Automatically selected ROIs for different anatomical landmarks in a 3D MR data set. For criterion A (a and b) and criterion B (c and d). The respective position estimates are marked by white crosses (see text).

criteria A and B, except for the genu of corpus callosum, where the widths are $w_{opt} = 11$ and $w_{opt} = 17$, resp. For this landmark, it appears that the accuracy of the position estimate for criterion B is much worse, possibly due to interfering brain convolutions in the neighborhood. One could expect that the maximal ROI for the frontal horn could be larger. However, note that there are several brain structures in the area below (see the lower border of the ROI) which causes our algorithm to stop with the selected width. All in all, it seems that the selected observation windows for both optimality criteria actually cover representative regions around all considered landmarks.

3.6 Related Work

Related work on the selection of suitable windows has been done in the context of matching stereo images by Okutomi and Kanade [9] and for junction localiza-

tion by Lindeberg [7]. Lindeberg presented an approach for junction localization with automatic scale selection. This approach differs in various aspects from our approach. First of all, he selects so-called scale-space maxima of the operator by Kitchen and Rosenfeld [6]. The respective scale at which a certain junction is detected is assumed to be an appropriate scale for an observation window around the structure. Note, however, that the used operator for structures of more complex form than an L-corner (e.g., Y- and Arrow-corners) lacks for a well-founded geometric interpretation. Also, it is not clear whether such a selected ROI is optimal w.r.t. localization uncertainty. That is, his criterion for selecting the size of the ROI differs from ours. Second, he defines an error function based on the residuals of the edge intersection approach. This error function is minimized by varying the scale of the filters for partial derivative estimation. The scale of the obtained minimum is taken as optimal scale for derivative estimation. However, with this approach the residual error is minimized, whereas we have considered the uncertainty of the position estimate which is generally not the same.

4 Conclusion and Further Work

We have investigated multi-step differential procedures for the refined localization of 3D point landmarks. The promising results due to our analytical study of the 2D edge intersection approach motivated a generalization to 3D. Based on this 3D extension, we have proposed two two-step and one three-step procedures for the localization of 3D point landmarks, combining the detection of points with refined localization through 3D edge intersection. We have experimentally studied these approaches for 3D synthetic images as well as for 3D MR images of the human head. For synthetic data, the combination of a two-step procedure of large-scale and small-scale 3D differential operators followed by the 3D edge intersection approach, that is, the three-step procedure, has yielded the most accurate results. For the MR images, the two-step and three-step procedures, both including 3D edge intersection, have yielded the best results. In comparison to applying a 3D detection operator alone, the localization accuracy has been significantly improved. We have also proposed a novel approach for automatically determining an optimal size for the ROI given a considered landmark. Our approach is based on the uncertainty of the position estimate resulting from the edge intersection approach. We have presented results for a real 2D image containing a polyhedral object as well as for 2D slices of a 3D MR image demonstrating the applicability of our approach.

Further experiments on medical images will be performed taking also other types of point landmarks into account. Our approach for ROI size selection has been formulated for use in 2D as well as 3D but so far has only been applied to 2D images. Thus, we will also study the performance for 3D images. Future work will also concern the registration of 3D medical images using the semi-automatically localized point landmarks.

Acknowledgement

References

1. P.R. Beaudet. Rotationally invariant image operators. In *Proc. International
   Conference on Pattern Recognition*, pages 579–583, Kyoto,Japan, 1978.
2. W. Beil, K. Rohr, and H.S. Stiehl. Investigation of Approaches for the Localization
   of Anatomical Landmarks in 3D Medical Images. In H.U. Lemke, M.W. Vannier,
   and K. Inamura, editors, *Proc. Computer Assisted Radiology and Surgery*, pages
   265–270, Berlin,Germany, 1997. Elsevier Amsterdam Lausanne.
3. L. Dreschler. *Ermittlung markanter Punkte auf den Bildern bewegter Objekte und
   Berechnung einer 3D-Beschreibung auf dieser Grundlage*. Dissertation, FB Infor-
   matik, Universität Hamburg, 1981.
4. W. Förstner. A feature based algorithm for image matching. *International Archives
   of Photogrammetry and Remote Sensing*, 26(3):150–166, 1986.
5. W. Förstner and E. Gülch. A Fast Operator for Detection and Precise Location
   of Distinct Points, Corners and Centres of Circular Features. In *Proc. Intercom-
   mission Conference on Fast Processing on Photogrammetric Data*, pages 281–305,
   Interlaken,Switzerland, 1987.
6. L. Kitchen and A. Rosenfeld. Gray-level corner detection. *Pattern Recognition
   Letters*, 1(2):95–102, 1982.
7. T. Lindeberg. Junction Detection with Automatic Selection of Detection Scales and
   Localization Scales. In *Proc. First International Conference on Image Processing*,
   volume 1, pages 924–928, Austin,TX,USA, 1994.
8. J.A. Noble. Finding corners. In *Proc. Third Alvey Vision Conference*, pages 267–
   274, Cambridge,UK, 1987.
9. M. Okutomi and T. Kanade. A Locally Adaptive Window for Signal Matching.
   International Journal of Computer Vision, 7(2):143–162, 1992.
10. K. Rohr. *Untersuchung von grauwertabhängigen Transformationen zur Ermittlung
    des optischen Flusses in Bildfolgen*. Diplomarbeit, Institut für Nachrichtentechnik,
    Universität Karlsruhe, 1987.
11. K. Rohr. Modelling and identification of characteristic intensity variations. *Image
    and Vision Computing*, 10(3):66–76, 1992.
12. K. Rohr. Recognizing Corners by Fitting Parametric Models. *International Journal
    of Computer Vision*, 9(3):213–230, 1992.
13. K. Rohr. Localization Properties of Direct Corner Detectors. *Journal of Mathe-
    matical Imaging and Vision*, 4(2):139–150, 1994.
14. K. Rohr. On 3D differential operators for detecting point landmarks. *Image and
    Vision Computing*, 15(3):219–233, 1997.
15. J.-P. Thirion. Extremal Points: Definition and Application to 3D Image Regis-
    tration. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*,
    pages 587–592, Seattle,WA,USA, 1994.