Journal of Computing and Information Technology, 6(4):435–447, 1998.

Multi-Step Differential Approaches for the Localization of 3D Point Landmarks in Medical Images

Sönke Frantz, Karl Rohr, and H. Siegfried Stiehl

Universität Hamburg, Fachbereich Informatik, Arbeitsbereich Kognitive Systeme,
Vogt-Kölln-Str. 30, D-22527 Hamburg, Germany
E-mail: {frantz,rohr,stiehl}@informatik.uni-hamburg.de

Abstract

In this contribution, we are concerned with the detection and refined subvoxel localization of 3D point landmarks. We propose multi-step differential approaches which are generalizations of an existing two-step approach for subpixel localization of 2D point landmarks. This two-step approach 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 approach for an analytical model of a Gaussian blurred L-corner as well as a Gaussian blurred ellipse. By varying the model parameters, differently tapered and curved structures are represented. The results motivate the use of an analogous approach to 3D point landmark localization. We generalize the edge intersection approach to 3D and, by combining it with 3D differential operators for landmark detection, we propose multi-step approaches for subvoxel localization of 3D point landmarks. The multi-step approaches are experimentally tested on 3D synthetic images and 3D MR images of the human head. We show that the multi-step approaches significantly improve the localization accuracy in comparison to applying a 3D detection operator alone.

Keywords: 3D anatomical point landmarks, 3D differential operators, 3D differential edge intersection approach, subvoxel localization

Introduction

The registration of 3D medical images such as Magnetic Resonance (MR) and X-Ray Computer Tomographic (CT) images with the aim to combine complementary information from different data sets is important, for example, for the planning of neurosurgical interventions, radiotherapy planning, and therapy evaluation. One possibility to register two images is a point-based approach, where corresponding anatomical points, denoted also as point landmarks, are localized in the images and then are used as features to compute a geometric

transformation for mapping one image onto another (e.g., Bookstein, 1989; Rohr *et al.*, 1996). Potential 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 points have been localized manually, which is time-consuming and often lacks accuracy. An alternative to this is a semi-automatic procedure, which has the advantage that the user can interactively control the results: First, the user determines an approximate position of a specific landmark. 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 detect and to accurately localize prominent points. Recently, 3D differential operators have been introduced which are, however, only designed for the detection of 3D point landmarks (Thirion, 1994; Rohr, 1997; Beil *et al.*, 1997).

In this contribution, we propose multi-step approaches for 3D point landmark localization, combining landmark detection with additional steps for refined localization. The multi-step approaches are based on an existing two-step approach of Förstner and Gülch (1987) for subpixel localization of 2D point landmarks. This approach combines landmark detection by applying a 2D differential operator with refined localization through a 2D differential edge intersection approach. In this paper, first, we investigate the localization performance of the two-step approach for a 2D analytical model of a Gaussian blurred L-corner and a Gaussian blurred ellipse. By varying the model parameters, differently tapered and curved structures are represented. The superior localization performance of the two-step approach in comparison to applying a 2D detection operator alone motivate the use of an analogous approach for 3D point landmark localization. We then generalize the differential edge intersection approach to 3D and combine it with existing 3D differential operators for landmark detection to obtain multi-step approaches for subvoxel localization of 3D point landmarks. The multi-step approaches are experimentally tested on 3D synthetic images and 3D MR images of the human head.

Two-step Approach for Localizing 2D Point Landmarks

Förstner and Gülch (1987) proposed a two-step approach for subpixel localization of 2D point landmarks. First, we summarize this two-step approach. Then, we investigate the localization performance of the two-step approach in comparison to applying a 2D detection operator alone. We study a 2D analytical model of a Gaussian blurred L-corner as well as a Gaussian blurred ellipse.

Description of the Approach

Landmark detection The used differential operator (see Förstner, 1986; Noble, 1987) exploits the matrix (up to a factor)

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

The subscripts x and y of the image function $g(x, y)$ stand for the partial derivatives in the respective spatial direction and the sum index i denotes the spatial location. The operator reads

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

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

Refined localization Suppose an L-corner has been detected and an observation window is placed around the detected position capturing sufficient edge information of the structure (see Fig. 1, where $\mathbf{x}_d$ denotes the detected position and the observation window is drawn with dashed style; $\mathbf{x}^*$ denotes the position of the tip, where the two edges meet).

[Figure 1 about here.]

For simplicity, a local coordinate system with the detected point as origin is chosen. At each edge point in the observation window a tangent is defined to locally approximate its corresponding edge direction. For this the image gradients are taken as normals to the tangents. The tangents are represented in the Hessian normal form. For example, for a point $\mathbf{x}_i$ with gradient ∇g_i the Hessian normal form reads $< \nabla g_i, \mathbf{x} >=< \nabla g_i, \mathbf{x}_i >$, where $< \cdot, \cdot >$ denotes the inner product. This is also illustrated in Fig. 1, where the tangent at $\mathbf{x}_i$ is drawn with bold style. Rewriting the tangent equation 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. An estimate for the position of the tip can be obtained through intersection of all tangents using the least-squares method, that is, through minimization of the residual error function

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

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 $\nabla \mathrm{E}(\mathbf{x}) = \mathbf{0}$ yields a system of normal equations

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

where $\mathbf{x}^*$ denotes the estimated subpixel position of the tip, $\mathbf{N}$ is the matrix from above, and

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

Localization Performance for 2D Analytical Models

The localization properties of the first step of the two-step approach of Förstner and Gülch (1987), that is, applying a detection operator alone, have been analyzed in Rohr (1994) for an analytical model of a Gaussian blurred L-corner (see Fig. 2a for the unblurred structure, where the position of the tip is marked through a black dot). It has been shown that the 2D differential operator in (1) yields systematic localization errors w.r.t. the correct position of the tip. In the following, we analyze the localization capabilities of the two-step approach for the analytical model of the Gaussian blurred L-corner and compare the achieved accuracy with that resulting from the 2D detection operator alone. We also analyze the localization capabilities of the two-step approach for a Gaussian blurred ellipse (see Fig. 2b for the unblurred structure, where the position of the considered tip is marked through a black dot).

[Figure 2 about here.]

First, we have considered the detection operator in (1) and have determined the locus of the maximal operator response along the symmetric axis of the structures, that is, the x-axis. For the computation of the operator responses we have used the approximation

$$\mathbf{N} \approx \nabla g(\mathbf{x}) \nabla g(\mathbf{x})^T + c\mathbf{H}_g(\mathbf{x})^2,$$

where $\nabla g(\mathbf{x})$ is the gradient of the respective Gaussian blurred function (see below) and $\mathbf{H}_g(\mathbf{x})$ is the Hessian matrix. The parameter c is a measure for the size of the observation window of the operator. We have used $c = 2/3$, which stands for a window width of $3pix$ (pix denotes spatial unity). Then, for the 2D differential edge intersection approach 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.

Results for the L-corner The model function of a Gaussian blurred L-corner with unit height and aperture angle β can be obtained through convolution of the ideal function shown in Fig. 2a as

$$L(\mathbf{x}, \beta) = \begin{cases} 1 & \text{if } x \geq 0 \ \wedge \ |y| \leq x \tan \frac{\beta}{2}, \\ 0 & \text{otherwise} \end{cases}$$

with a Gaussian blur function $G(\mathbf{x},\sigma) = 1/(2\pi\sigma^2)\exp\Big(-(x^2+y^2)/(2\sigma^2)\Big)$, where σ denotes the standard deviation (see Rohr (1992) for details).

In Fig. 3a, the distances e from the localized positions to the correct corner position at $(0,0)$ are depicted in dependence on 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 approach 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. The localization error for the detection operator alone heavily depends on the aperture angle β. The localization error is very small for values near $\beta = 180°$ and high for small values of β. The position estimates resulting from the edge intersection approach are significantly more accurate. In fact, the additional second step improves the accuracy by about $1pix$ for a large range of values of β. Moreover, we can obtain even better position estimates if we further enlarge the observation window for the edge intersection approach and thus take more edge information into account. The result for a window size of 15×15 pixels is shown in Fig. 3b. It can be seen that the localization error is reduced to nearly zero.

[Figure 3 about here.]

Results for the Ellipse The ellipse with semi-axes a and b in Fig. 2b can be described by

$$E_2(\mathbf{x},a,b) = \begin{cases} 1 & \text{if } \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

The Gaussian blurred ellipse can be obtained through convolution with $G(\mathbf{x},\sigma)$. The considered landmark is the tip at $(a,0)$. We have investigated ellipses with different shapes. For this we have varied the ratio a/b. For instance, a large value of this ratio gives a more tapered shape and a value near 1 gives a more rounded shape. We have permanently set $b = 4$ and have varied a in the range of $a = 6 \ldots 128$.

In Fig. 4a, the distances e from the localized positions to the position of the tip are depicted in dependence on the ratio a/b. The standard deviation of the Gaussian blur function has been set to $\sigma = 1$. The localization error resulting from the detection operator alone amounts on average to a bit more than $1pix$. Again, the second step of the approach significantly improves the localization accuracy. However, the localization error increases for the edge intersection approach if we further enlarge the observation window. This is because the ellipse deviates from the assumed polygonal model and this deviation increases if we enlarge the observation window. Nevertheless, in Fig. 4b it is shown that even for an observation window of size 7×7 pixels the position estimates resulting from the edge intersection approach are significantly more accurate than those resulting from the detection operator alone.

[Figure 4 about here.]

Our investigation has demonstrated the superior localization performance of the two-step approach in comparison to applying a detection operator alone

and hence motivates the use of an analogous approach for 3D point landmark localization.

Extension of the Edge Intersection Approach to 3D

Let us now generalize the edge intersection approach of Förstner and Gülch (1987) to 3D. In the 2D case, we have dealt with an L-corner to motivate the approach. A 3D generalization of an L-corner is a tetrahedron, which is shown in Fig. 5.

[Figure 5 about here.]

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 which captures sufficient edge 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 at a landmark through tangent planes. The image 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 as in the 2D case 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,

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

and

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

Detection of 3D Point Landmarks

The 3D edge intersection approach described above requires the detection of point landmarks. Recently, 3D differential operators for point landmark detection have been introduced in Rohr (1997) which are 3D extensions of existing 2D corner detectors (Förstner, 1986; Noble, 1987; Rohr, 1987, 1992). The operators only employ first order partial derivatives and therefore do not suffer from potential instabilities of computing high order partial derivatives. By contrast, other approaches employ partial derivatives of up to order two or

three (Thirion, 1994; Beil *et al.*, 1997). The 3D operators essentially exploit the matrix $\mathbf{N}$ in (4) and read $Op3(\mathbf{x}) = det(\mathbf{N})/tr(\mathbf{N})$, $Op3'(\mathbf{x}) = 1/tr(\mathbf{N}^{-1})$, as well as $Op4(\mathbf{x}) = det(\mathbf{N})$. The operators $Op3$ and $Op3'$ are 3D extensions of the operator in (1) (see Förstner, 1986), whereas $Op4$ is a 3D extension of the operator of Rohr (1987, 1992). Experiments with 3D MR and CT images of the human head have yielded similar promising results for the different operators (see also Hartkens *et al.*, 1998). The operators can also be related to the localization uncertainty of the position estimate resulting from the 3D edge intersection approach. The localization uncertainty is given through the covariance matrix $\mathbf{\Sigma} = \sigma_\varepsilon^2 \mathbf{N}^{-1}$, supposed the residuals ε_i are independently and normally distributed with zero mean and variance σ_ε^2. Hence, the responses of the differential operators can also be understood as scalar measures for the localization (un-)certainty of the position estimate resulting from the 3D edge intersection approach.

Two-Step and Three-Step Approaches for the Localization of 3D Point Landmarks

Based on the 3D detection operators and the 3D edge intersection approach, we here propose three multi-step approaches for the refined localization of 3D point landmarks.

i) Two-step approach
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 approach for 2D point landmark localization was proposed earlier (Dreschler, 1981; Rohr, 1987).

ii) Two-step approach 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 approach of Förstner and Gülch (1987).

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

Experimental Results

We present results from experiments with 3D synthetic images and 3D MR images of the human head. The experiments have been carried out in the KHOROS environment. The partial derivatives have been estimated with

3D extensions of the 2D filters of Beaudet (1978). For the detection step we have used a filter size of $5 \times 5 \times 5$ voxels. For the refinement steps of the approaches 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 the products of the partial derivatives within an observation window of a certain spatial extent. To study the localization accuracy of the different approaches in dependence on the width w of the observation window, we have investigated various 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 a local maximum search in neighborhoods of $3 \times 3 \times 3$ voxels. In the case of several maxima within the image, we have selected the candidate with the largest operator response.

Synthetic Images

We have investigated the localization accuracy of the approaches i), ii), and iii) as well as of a 3D detection operator alone for synthetically generated tetrahedrons and ellipsoids. In the experiments, we have used $Op3'$ for both landmark detection and the refinement steps of the approaches i) and iii).

Results for the Tetrahedron This structure is a 3D generalization of an L-corner with aperture angle β. To construct the 3D object in Fig. 5, the symmetric axis of the L-corner is spread into the direction of the z-axis such that it also encloses the angle β with the x-axis. This gives

$$T(\mathbf{x}, \beta) = \begin{cases} 1 & \text{if } x \geq 0 \ \wedge \ 0 \leq z \leq xs \ \wedge \ |y| \leq xt - z\frac{1-t^2}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

where $s = \tan\beta$, $t = \tan(\beta/2)$, and $0° \leq \beta < 90°$. The shape of the such constructed binary tetrahedron is determined by the aperture angle β. For example, the choice $\beta = 90°$ gives the corner of a cube, while a small value of β gives a very tapered structure.

In Fig. 6a, 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. We see that applying the detection operator alone yields the worst results. Approach i) is only for $w = 5$ better than DET. The position estimates resulting from the approaches ii) and iii) are significantly more accurate than those resulting from DET and approach i). The accuracy gets worse for DET and approach i) if w increases. By contrast, the accuracy for the approaches ii) and iii) gets better if w increases. In Fig. 6b, 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 approaches ii) and iii) the localization error slightly increases if $w \geq 11$ and $w \geq 7$, resp. We suspect that discretization errors give rise to this effect.

[Figure 6 about here.]

Results for the Ellipsoid A binary ellipsoid can be defined through

$$E_3(\mathbf{x}, a, b, c) = \begin{cases} 1 & \text{if } \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

where a, b, and c are the lengths of the semi-axes in x-, y-, and z-direction, resp. We consider as point landmark the tip at $(0, 0, c)$.

In Fig. 7a, for a rotationally symmetric ellipsoid with $a = 8, b = 8, c = 40$ 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. In Fig. 7b, the results for an ellipsoid with $a = 16, b = 8, c = 40$ are shown. It can be seen that for the first ellipsoid the position estimates resulting from DET can be improved with approach i). However, for the second ellipsoid the localization error for approach i) partly is higher in comparison to DET. The position estimates resulting from the approaches ii) and iii) generally are significantly more accurate than those resulting from DET and i). Approach ii) improves the accuracy w.r.t. DET by about 2-3vox (vox denotes spatial unity). Approach iii) yields for the first ellipsoid comparable results as ii). For the second ellipsoid the position estimates resulting from approach iii) generally are worse than those resulting from ii). In general, for DET and approach i) the localization error increases if the width w of the observation window increases. We also note for the approaches ii) and iii) a larger localization error for small and large values of w. For small values of w too little information for the 3D edge intersection approach is incorporated and for large values of w the approximation of the surfaces through tangent planes gets worse.

[Figure 7 about here.]

3D MR Images of the Human Head

We consider as 3D anatomical point landmarks the tips of the frontal, occipital, and temporal horns of the ventricular system of the human head, abbreviated by MC6, MC7 and MC13, resp. The tips are indicated in Fig. 8a through black dots within dashed circles. The letters 'l' and 'r' denote the respective hemispheric part. In Fig. 8b, axial slices of a 3D MR image with the horns of the ventricular system are shown. The positions of the tips in the depicted slices are marked through white crosses.

[Figure 8 about here.]

We here report on experiments for these landmarks in three T1-weighted MR images. MR Image 1 consists of 235 sagittal slices of 256×256 voxels (voxel resolution is $1 \times 1 \times 1\text{mm}^3$), MR image 2 consists of 192 axial slices of 150×200 voxels (voxel resolution is $1.075 \times 1.075 \times 1.075\text{mm}^3$), and MR image 3 consists of 120 axial slices of 256×256 voxels (voxel resolution is $0.86 \times 0.86 \times 1.2\text{mm}^3$). 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 also that we have only determined

voxel positions, while the multi-step approaches 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. The size of the ROI has been set to $21 \times 21 \times 21$ voxels. For landmark detection and the refinement steps of the approaches i) and iii) we have used the operator $Op3$. We have used the same filter widths for estimating the partial derivatives as in the experiments with 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 procedure described in the introduction. Note, however, that some landmarks have required special attention.

In Tabs. 1–3, the results for the ventricle landmarks are shown. The first column gives the considered landmark (LM) and the width w (in voxels) of the observation window for the detection operator and the 3D edge intersection approach. Then follow the respectively localized positions resulting from the detection operator alone (DET) and the approaches i), ii), and iii). All positions are relative w.r.t. the manually specified landmark positions (top row of a box). The values below are the Euclidean distances to the manually specified landmark positions. It can be seen that, in general, the approaches ii) and iii) yield the most accurate positions. Also, approach i) generally yields better positions than DET, although not as good as the approaches ii) and iii). 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 poorly pronounced. Therefore, also the positions of these landmarks resulting from manual localization are rather uncertain. All detected points for MC13l,r in MR image 2 and MC13l in MR image 3 have extremely low operator responses indicating high localization uncertainties. Therefore, we have labeled the respective entries in the tables by a question-mark '?' meaning that 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. To indicate the different selection criteria in this case, we have labeled the corresponding entries in Tab. 3 by an asterisk '$*$'. For the left occipital horn MC7l in MR image 1 and the observation window width $w = 3$ we note a relatively large localization error. Visual inspection has actually revealed a 'double horn' giving rise to two candidates with extremely high operator responses. For $w = 3$ the candidate with the highest operator response is farer away from the manually specified position.

[Table 1 about here.]

[Table 2 about here.]

[Table 3 about here.]

In Figs. 9a and b, 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 w.r.t. each landmark and w.r.t. each MR image, resp. The landmarks which are labeled by a question-mark in Tabs. 2 and 3 have not been taken into account for the computation of the mean values. The graphs exhibit the superior localization capabilities of the approaches ii) and iii) in comparison to approach i) and the detection operator alone.

[Figure 9 about here.]

All in all, for DET the mean of the Euclidean distances from the localized positions to the manually specified positions amounts to $3.27 vox$ (vox denotes spatial unity). The two-step approach 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 approach iii) improves the accuracy w.r.t. DET by $1.52 vox$. The two-step approach ii) improves the accuracy w.r.t. DET by $1.14 vox$.

To visualize the localization capabilities of the different approaches, we also show in Fig. 10b orthogonal image views at the respectively localized positions for the tip of the left frontal 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. 10a. Note that, due to technical reasons, the subvoxel coordinates resulting from the approaches ii) and iii) have been rounded to voxel coordinates.

[Figure 10 about here.]

Conclusion

We have investigated multi-step differential approaches for the detection and refined localization of 3D point landmarks. The promising results due to our analytical investigation of the 2D edge intersection approach of Förstner and Gülch (1987) have motivated a generalization to 3D. Based on the 3D extension of this approach, we have proposed two two-step approaches as well as one three-step approach for subvoxel landmark localization. These approaches combine landmark detection by applying a 3D differential operator with refined localization through the 3D edge intersection approach. We have experimentally tested these approaches on 3D synthetic images and 3D MR images of the human head. The multi-step approaches with the 3D edge intersection approach have yielded the most accurate position estimates.

Acknowledgement

Support of Philips Research Laboratories Hamburg, project IMAGINE (IMage- and Atlas-Guided Interventions in NEurosurgery), is gratefully acknowledged.

The MR images have been provided by the Philips Research Laboratories Hamburg, the IMDM (Prof. K.H. Höhne) at the University Hospital Hamburg-Eppendorf, as well as the AIM project COVIRA (COmputer VIsion in RAdiology), resp. We thank the anonymous referees for their constructive comments on the manuscript.

References

P.R. Beaudet (1978). Rotationally invariant image operators. In *Proc. International Conference on Pattern Recognition*, pp. 579–583, Kyoto, JP.

W. Beil, K. Rohr, and H.S. Stiehl (1997). 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 (CAR'97)*, pp. 265–270, Berlin, FRG. Elsevier Science.

F.L. Bookstein (1989). Principal Warps: Thin-Plate Splines and the Decomposition of Deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(6), pp. 567–585.

L. Dreschler (1981). *Ermittlung markanter Punkte auf den Bildern bewegter Objekte und Berechnung einer 3D-Beschreibung auf dieser Grundlage*. Dissertation, Fachbereich Informatik, Universität Hamburg, FRG.

W. Förstner (1986). A feature based algorithm for image matching. *International Archives of Photogrammetry and Remote Sensing*, **26**(3), pp. 150–166.

W. Förstner and E. Gülch (1987). A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. In *Proc. Intercommission Conference on Fast Processing of Photogrammetric Data*, pp. 281–305, Interlaken, CH.

T. Hartkens, K. Rohr, and H.S. Stiehl (1998). Evaluierung der Detektionsleistung von 3D-Operatoren zur Ermittlung anatomischer Landmarken in tomographischen Bildern. In T. Lehmann, V. Metzler, K. Spitzer, and T. Tolxdorff, editors, *Proc. Bildverarbeitung für die Medizin*, pp. 93–97, Aachen, FRG. Springer.

J.A. Noble (1987). Finding corners. In *Proc. Third Alvey Vision Conference*, pp. 267–274, Cambridge, UK.

K. Rohr (1987). *Untersuchung von grauwertabhängigen Transformationen zur Ermittlung des optischen Flusses in Bildfolgen*. Diplomarbeit, Institut für Nachrichtentechnik, Universität Karlsruhe, FRG.

K. Rohr (1992). Modelling and identification of characteristic intensity variations. *Image and Vision Computing*, **10**(3), pp. 66–76.

K. Rohr (1994). Localization Properties of Direct Corner Detectors. *Journal of Mathematical Imaging and Vision*, **4**(2), pp. 139–150.

K. Rohr (1997). On 3D differential operators for detecting point landmarks. *Image and Vision Computing*, **15**(3), pp. 219–233.

K. Rohr, H.S. Stiehl, R. Sprengel, W. Beil, T.M. Buzug, J. Weese, and M.H. Kuhn (1996). Point-Based Elastic Registration of Medical Image Data Using Approximating Thin-Plate Splines. In K.H. Höhne and R. Kikinis, editors, *Proc. Visualization in Biomedical Computing (VBC'96)*, pp. 297–306, Hamburg, FRG. Springer.

J. Sobotta (1988). *Atlas der Anatomie des Menschen*. Urban & Schwarzenberg, 19 edition.

J.-P. Thirion (1994). Extremal Points: Definition and Application to 3D Image Registration. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pp. 587–592, Seattle, WA, USA.

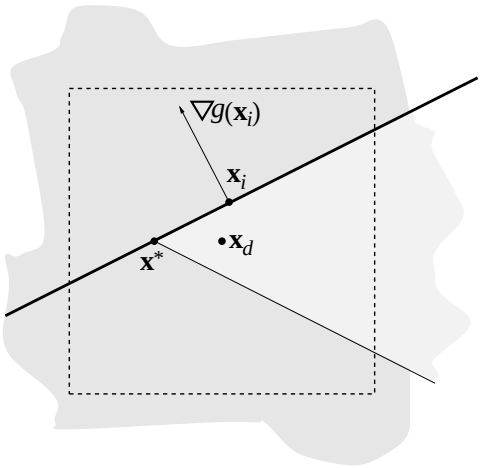


Figure 1 Sketch of the 2D edge intersection approach, where $\mathbf{x}_d$ denotes the detected position and $\mathbf{x}^*$ denotes the correct position of the tip, where the two edges meet. The observation window around $\mathbf{x}_d$ is drawn with dashed style.

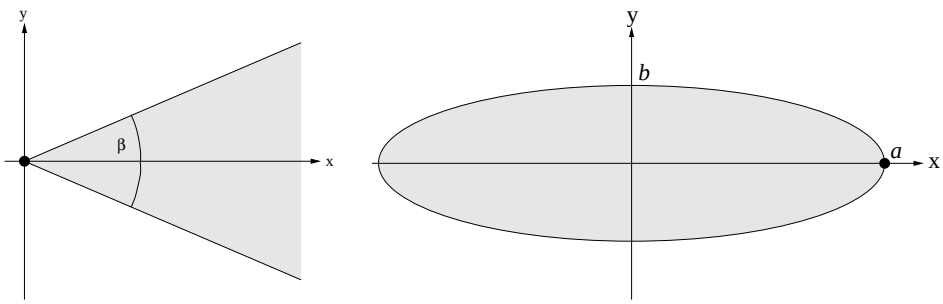
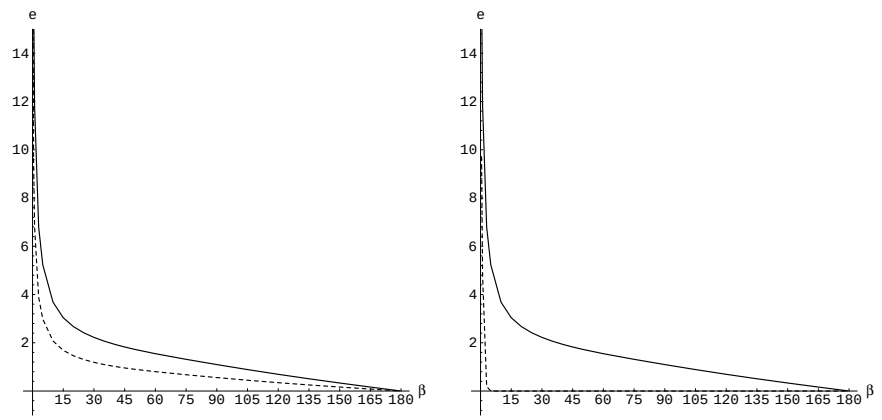
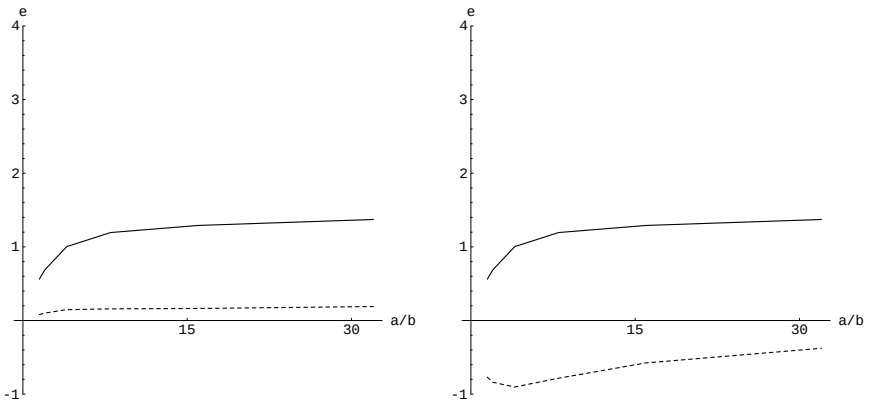


Figure 2a L-corner with aperture angle β **b** Ellipse with semi-axes a and b. The positions of the tips are marked through black dots.



a **b**

Figure 3a 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 at $(0,0)$ are depicted in dependence on the aperture angle β. The solid line results from the detection operator alone and the dashed line results from the two-step approach. Observation windows of sizes 3×3 pixels (a) and 15×15 pixels (b) have been used.



a **b**

Figure 4a and b Localization accuracy for a Gaussian blurred ellipse with $\sigma = 1$. The distances e from the localized positions to the position of the tip at $(a, 0)$ are depicted in dependence on the ratio a/b, where we have permanently set $b = 4$ and have varied a in the range of $a = 6 \ldots 128$. The solid line results from the detection operator alone and the dashed line results from the two-step approach. Observation windows of sizes 3×3 pixels (a) and 7×7 pixels (b) have been used.

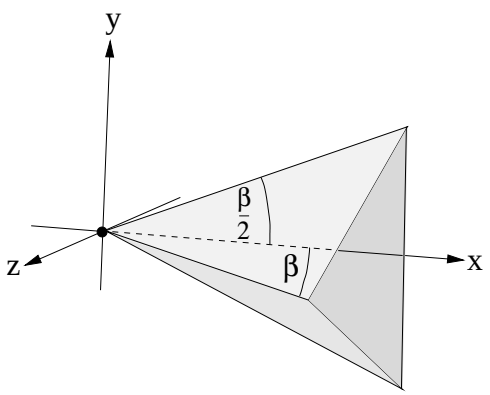


Figure 5 Tetrahedron with aperture angle β. The position of the tip is marked through a black dot.

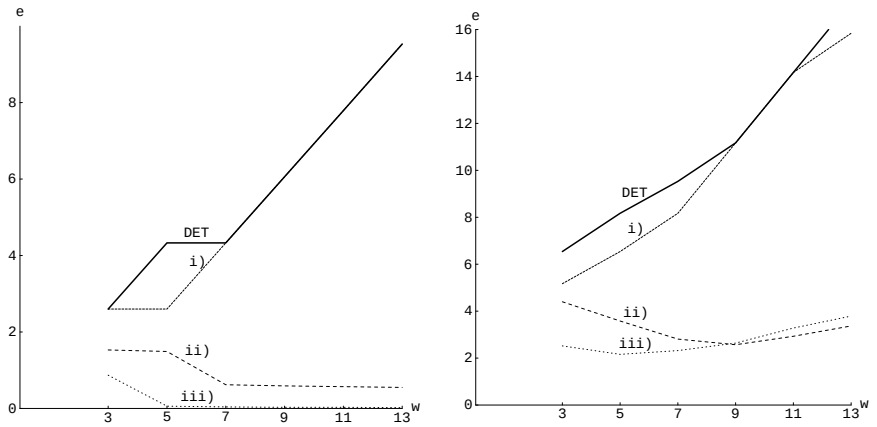


Figure 6a and b Localization accuracy for tetrahedrons with $\beta = 90°$ and $\beta = 45°$, resp. The Euclidean distances e from the localized positions to the position of the tip are depicted in dependence on the width w of the observation window.

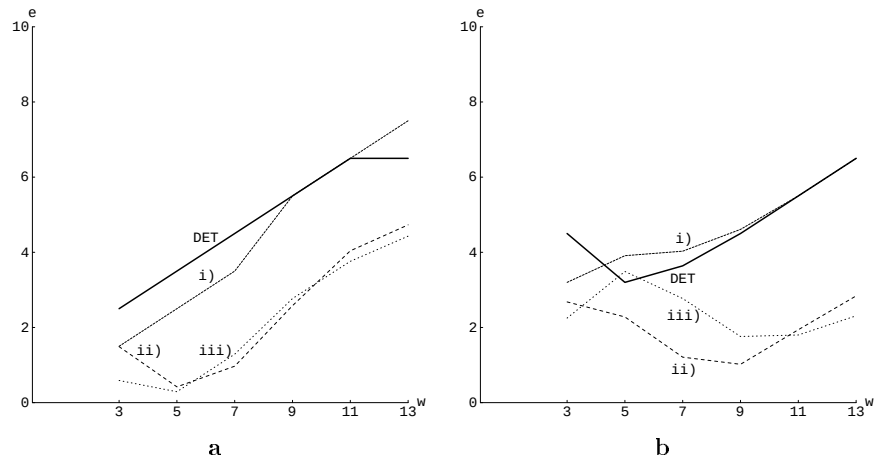
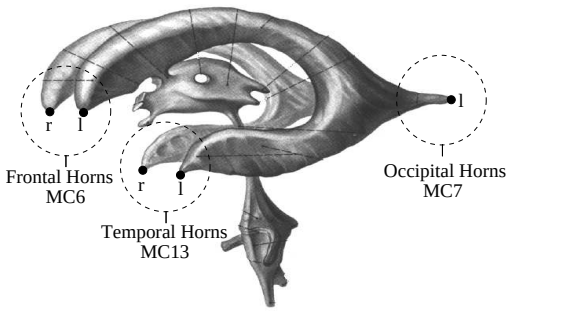
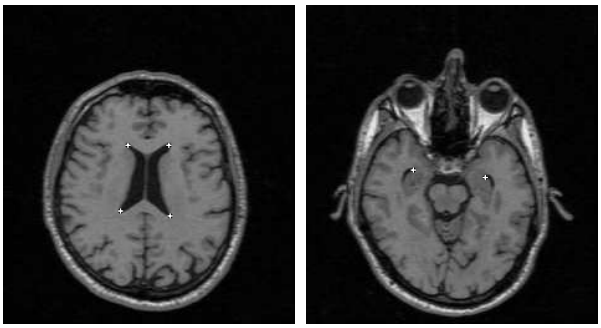


Figure 7a and b Localization accuracy for ellipsoids with semi-axes $a = 8, b = 8, c = 40$ and $a = 16, b = 8, c = 40$, resp. The Euclidean distances e from the localized positions to the position of the tip at $(0, 0, c)$ are depicted in dependence on the width w of the observation window.

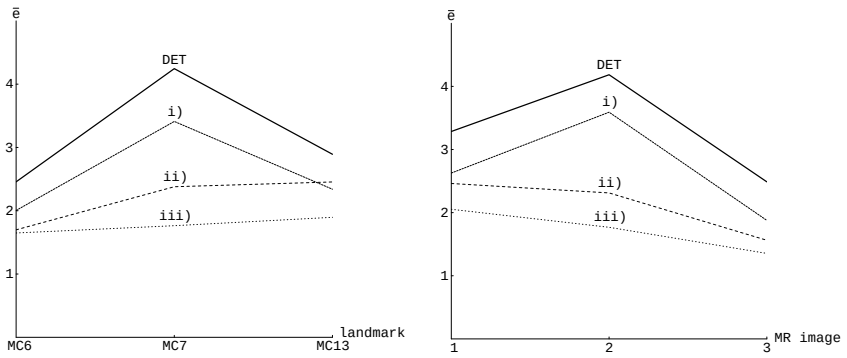


a



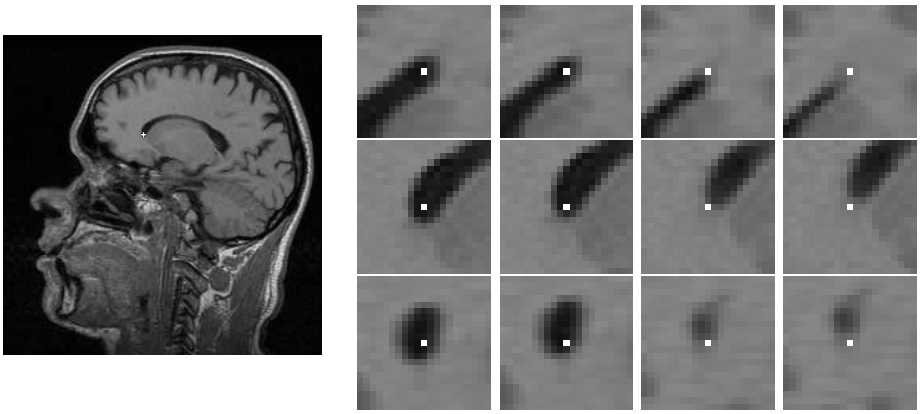
b

Figure 8a Anatomical preparate of the ventricular system of the human head (adapted from Sobotta, 1988) **b** Axial slices of a 3D MR image with the frontal, occipital, and temporal horns of the ventricular system. The tips of the frontal and occipital horns (left) and the tips of the temporal horns (right) in the depicted slices are marked through white crosses.



a **b**

Figure 9a and b Localization accuracy of detection operator alone (DET) and the approaches i), ii), and iii) for the ventricle landmarks in the investigated 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 w.r.t. each landmark (a) and w.r.t. each MR image (b).



a b

Figure 10a Off-sagittal slice of MR image 1. The manually specified position of the tip of the left frontal horn is marked through a white cross. **b** The localized positions at the tip of the left frontal horn in MR image 1 ($w = 5$) are shown for the detection operator alone and the multi-step approaches i), ii), and iii) (from left to right) in axial, sagittal, and coronal views (from top to bottom).

LM/w	DET	i)	ii)	iii)
MC6l/3	1, 2, 0 2.24	0, 1, 0 1.00	0.08, 1.50, 0.85 1.73	0.79, 0.27, 0.36 0.91
MC6l/5	1, 2, 0 2.24	1, 1, 0 1.41	0.91, 0.43, 1.16 1.54	0.91, 0.59, 1.00 1.47
MC7l/3	4, 5, 1 6.48	3, 6, 1 6.78	2.52, 6.19, 1.12 6.78	0.82, 7.46, 0.48 7.52
MC7l/5	3, 2, 2 4.12	2, 1, 1 2.45	0.53, 0.96, 1.51 1.86	0.52, 0.15, 0.86 1.02
MC13l/3	1, 2, 0 2.24	1, 1, 1 1.73	0.79, 1.88, 0.86 2.21	0.56, 0.06, 3.31 3.36
MC13l/5	2, 2, 1 3.00	2, 2, 1 3.00	0.92, 1.21, 0.88 1.76	0.59, 0.95, 0.55 1.25

LM/w	DET	i)	ii)	iii)
MC6r/3	1, 2, 0 2.24	1, 1, 0 1.41	0.03, 1.50, 0.13 1.51	0.30, 0.21, 0.52 0.63
MC6r/5	1, 2, 0 2.24	1, 2, 0 2.24	0.84, 0.90, 0.76 1.45	1.44, 0.23, 1.27 1.93
MC7r/3	3, 2, 0 3.61	2, 1, 0 2.24	1.80, 1.58, 0.14 2.40	0.62, 0.52, 0.15 0.82
MC7r/5	3, 2, 0 3.61	3, 2, 0 3.61	0.75, 0.61, 0.82 1.27	0.28, 0.03, 0.89 0.93
MC13r/3	0, 1, 2 2.24	0, 1, 1 1.41	0.91, 0.41, 0.87 1.33	0.72, 0.61, 1.31 1.61
MC13r/5	1, 1, 5 5.20	1, 1, 4 4.24	0.60, 0.30, 5.67 5.71	1.51, 0.75, 2.68 3.17

Table 1 Localized positions of the ventricle landmarks in MR image 1. LM denotes the landmark, w denotes the width of the observation window for the 3D detection operator and the 3D edge intersection approach. The following columns give the relative localized positions as well as the Euclidean distances to the manually specified landmark positions.

LM/w	DET	i)	ii)	iii)
MC6l/3	2, 2, 0 2.83	2, 2, 0 2.83	1.61, 1.16, 1.04 2.24	1.58, 1.18, 1.74 2.63
MC6l/5	2, 2, 1 3.00	2, 2, 0 2.83	1.03, 0.17, 1.25 1.63	1.39, 0.23, 2.78 3.12
MC7l/3	0, 4, 1 4.12	0, 3, 0 3.00	0.10, 3.20, 0.08 3.21	0.05, 2.49, 0.97 2.67
MC7l/5	0, 4, 1 4.12	0, 4, 1 4.12	0.45, 1.81, 0.46 1.92	0.31, 1.00, 0.55 1.18
?MC13l/3	2, 1, 2 3.00	2, 2, 1 3.00	2.20, 1.38, 2.71 3.76	2.52, 4.47, 1.84 5.45
?MC13l/5	2, 1, 2 3.00	2, 1, 2 3.00	1.86, 1.86, 3.49 4.37	1.62, 2.66, 2.96 4.30

LM/w	DET	i)	ii)	iii)
MC6r/3	1, 2, 2 3.00	1, 2, 1 2.45	0.43, 0.88, 1.87 2.11	0.39, 0.92, 0.24 1.03
MC6r/5	1, 2, 2 3.00	1, 2, 2 3.00	0.24, 0.26, 1.64 1.68	0.71, 0.28, 1.59 1.76
MC7r/3	1, 5, 3 5.92	1, 4, 2 4.58	0.50, 3.06, 1.77 3.57	0.04, 1.07, 0.05 1.07
MC7r/5	2, 6, 4 7.48	1, 5, 3 5.92	1.39, 1.08, 1.22 2.14	0.09, 0.53, 0.42 0.68
?MC13r/3	6, 3, 2 7.00	6, 3, 1 6.78	6.22, 3.00, 1.63 7.10	5.82, 3.00, 1.08 6.64
?MC13r/5	6, 3, 1 6.78	5, 3, 1 5.92	6.50, 3.16, 0.87 7.28	5.94, 2.67, 1.09 6.60

Table 2 Localized positions of the ventricle landmarks in MR image 2. LM denotes the landmark, w denotes the width of the observation window for the 3D detection operator and the 3D edge intersection approach. The following columns give the relative localized positions as well as the Euclidean distances to the manually specified landmark positions.

LM/w	DET	i)	ii)	iii)
MC6l/3	1, 2, 0 2.24	0, 1, 0 1.00	0.52, 0.90, 0.64 1.23	0.59, 0.17, 0.06 0.62
MC6l/5	1, 2, 0 2.24	1, 2, 0 2.24	0.07, 0.14, 1.71 1.72	0.27, 0.57, 1.58 1.70
MC7l/3	2, 3, 1 3.74	1, 2, 0 2.24	1.09, 0.39, 0.77 1.39	0.14, 0.12, 0.75 0.77
MC7l/5	2, 3, 1 3.74	2, 2, 1 3.00	0.03, 1.76, 0.37 1.80	0.03, 1.85, 0.22 1.86
?MC13l/3	2, 5, 2 5.74	3, 6, 2 7.00	2.41, 2.16, 1.75 3.68	2.98, 3.75, 1.29 4.96
?MC13l/5	6, 6, 6 10.39	7, 6, 7 11.58	4.66, 7.63, 7.40 11.61	7.33, 6.59, 7.55 12.42

LM/w	DET	i)	ii)	iii)
MC6r/3	0, 2, 1 2.24	0, 1, 1 1.41	0.55, 1.13, 1.34 1.84	0.53, 0.09, 1.24 1.35
MC6r/5	0, 2, 0 2.00	0, 2, 1 2.24	0.80, 0.12, 1.52 1.72	0.54, 0.02, 2.56 2.62
MC7r/3	0, 2, 0 2.00	0, 1, 0 1.00	0.30, 0.38, 0.32 0.58	0.05, 0.44, 0.52 0.69
MC7r/5	0, 2, 0 2.00	0, 2, 0 2.00	0.69, 1.42, 0.42 1.63	0.44, 1.67, 0.88 1.94
*MC13r/3	1, 2, 1 2.45	0, 1, 1 1.41	2.33, 0.87, 0.29 2.50	0.31, 0.33, 0.52 0.69
*MC13r/5	0, 2, 1 2.24	0, 2, 1 2.24	1.14, 0.39, 0.14 1.21	1.27, 0, 0.27 1.30

Table 3 Localized positions of the ventricle landmarks in MR image 3. LM denotes the landmark, w denotes the width of the observation window for the 3D detection operator and the 3D edge intersection approach. The following columns give the relative localized positions as well as the Euclidean distances to the manually specified landmark positions.