PERGAMON

Pattern Recognition 32 (1999) 3–15

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

Extraction of 3d anatomical point landmarks based on invariance principles

Karl Rohr

Universität Hamburg, Fachbereich Informatik, Arbeitsbereich Kognitive Systeme, Vogt-Kölln-Str. 30, 22527 Hamburg, Germany

Received 28 October 1997, in revised form 20 April 1998; received for publication 12 February 1998

Abstract

We describe 3D operators for extracting anatomical landmarks which are based on only first-order partial derivatives of an image. To improve the predictability of the extraction results we analyze certain properties of the operators. First, we provide a statistical interpretation in terms of the Cramér–Rao bound representing the minimal localization uncertainty. Second, we show that the operators can be derived on the basis of invariance principles. It turns out that the operators form a complete set of principal invariants. Third, we analyze the detection performance using a certain type of performance visualization and a scalar performance measure. Experimental results are presented for 3D tomographic images of the human brain. © 1999 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

Keywords: Point-based registration; Medical image analysis; Landmark extraction; 3D differential operators; Invariance principles; Uncertainty lower bound; Detection performance

1. Introduction

The accuracy and reliability of point-based image registration in general strongly depends on the detection and accurate localization of landmarks. In particular, this holds true in the case of 3D medical image data, e.g. magnetic resonance (MR) and X-ray computed tomography (CT) images or digital atlases. In this context, the task is to accurately match multimodality images such that the complementary information of the different data sets can be combined. In either case, the goal is to increase the accuracy of localizing anatomical structures in 3D space which is of fundamental importance in computer-assisted neurosurgery. To match different data sets, often corresponding prominent points, also denoted as landmarks, are identified and used as features to compute the transformation between the data sets [1–4]. In comparison to fiducial markers placed outside the human head, anatomical landmarks have the advantage that they can be located within the relevant inner brain parts. Therefore, the registration accuracy is generally increased in these regions. However, usually 3D anatomical landmarks are localized manually which is time-consuming and often lacks accuracy. Instead, semi-automatic or automatic procedures promise to improve this situation.

Currently, there are only a few computational approaches for extracting point landmarks in 3D tomographic images [5,6]. Thirion [5], for example, has introduced 3D differential operators that consist of partial derivatives of an image up to the third order to detect certain points on 3D ridge lines. However, the computation of high-order partial derivatives generally is very sensitive to noise. Therefore, additional steps are necessary to diminish these instabilities [7]. Related operators based on isocontour curvature properties and

PII: S0031-3203(98)00088-0

using partial derivatives up to the second order can be found in Ref. [8] (see also Ref. [9]). In Ref. [6] we have introduced 3D differential operators which are generalizations of existing 2D corner detectors. These operators employ either only first-order partial derivatives or first- and second-order partial derivatives of an image. Therefore, these operators are computationally efficient and they do not suffer from instabilities of computing high-order partial derivatives. Recently, the 2D versions of these operators have been evaluated using 2D tomographic images of the human brain (see Ref. [10]). It turned out that the operators based on only first-order derivatives yield better results than those based on first- and second-order derivatives.

In this contribution, we consider 3D operators which employ only first-order partial derivatives of an image. Whereas in Ref. [6], these operators have been derived straightforwardly in accordance with the 2D case, we here provide a sound statistical interpretation of the operators. This interpretation is given in terms of the minimal localization uncertainty defined by the Cramér–Rao bound. Moreover, we show that the operators can be derived on the basis of invariance principles. It turns out that the operators form a complete set of principal invariants w.r.t. the covariance matrix of the position estimate. Application of the operators to 3D tomographic images reveals that a large number of important 3D anatomical landmarks can be detected with these operators; however, there are also a number of false detections. In this contribution, we therefore analyze the detection performance in more detail. We introduce and apply a certain type of performance visualization and also use a scalar quantity to measure the detection capability. Experimental results are presented for 3D tomographic images of the human brain.

2. Statistical interpretation

2D differential corner detectors have usually been designed to extract grey-value corners; however, in many applications they have been used to find general points of high-intensity variations. Therefore, we can summarize the general extraction criterion of these operators as finding "points with high intensity variations". This criterion motivated the 3D generalizations in Ref. [6] and their use for localizing anatomical point landmarks in 3D tomographic images. Whereas the criterion is rather general and unspecific, in this section we provide a statistical interpretation of the 3D operators. We consider the following three operators which are based only on first-order partial derivatives of an image and which represent 3D generalizations of the 2D corner detectors in Refs. [11] and [12]:

$$Op3 = \frac{det\,\mathbf{C}_g}{tr\,\mathbf{C}_g}, \qquad Op3' = \frac{1}{tr\,\mathbf{C}_g^{-1}}, \qquad Op4 = det\,\mathbf{C}_g, \tag{1}$$

where

$$\mathbf{C}_g = \overline{\nabla g\,(\nabla g)^{\mathrm{T}}} \tag{2}$$

is a symmetric 3×3 matrix which represents the averaged dyadic product of the 3D intensity gradient $\nabla g = (g_x, g_y, g_z)^{\mathrm{T}}$, $g(x, y, z)$ is the image function, and *det* and *tr* denote the determinant and the trace of a matrix, respectively. If $\mathbf{C}_g$ has full rank, then $\mathbf{C}_g$ is positive-definite and all eigenvalues of $\mathbf{C}_g$ as well as those of it's inverse are larger than zero. Therefore, the operators in Eq. (1) yield positive values, and landmark points can be found by computing local maxima of the expressions in Eq. (1).

In the following, let σ_n^2 denote the variance of additive white Gaussian image noise and m the number of voxels in a local 3D window. Then, we can relate the matrix $\mathbf{C}_g$, which captures the intensity variations inside the window, to the minimal localization uncertainty of the center of the window $\mathbf{x} = (x, y, z)$. The minimal localization uncertainty is given by the Cramér–Rao bound [13] and is represented by the covariance matrix

$$\mathbf{\Sigma}_g = \frac{\sigma_n^2}{m}\,\mathbf{C}_g^{-1}. \tag{3}$$

We see that $\mathbf{\Sigma}_g$ is proportional to the inverse of $\mathbf{C}_g$. From Eq. (3) we can derive the 3D error ellipsoid of the position estimate with semi-axes σ_x, σ_y, and σ_z. A quantitative measure for the localization uncertainty of a landmark is the volume of the 3D error ellipsoid which is defined as

$$V = \frac{4}{3}\pi\sqrt{det\,\mathbf{\Sigma}_g}, \tag{4}$$

where

$$det\,\mathbf{\Sigma}_g = \sigma_x^2\sigma_y^2\sigma_z^2. \tag{5}$$

The smaller the value of $det\,\mathbf{\Sigma}_g$ the smaller is the localization uncertainty. Thus, we can formulate the following criterion for localizing 3D point landmarks: Find those points with "minimal localization uncertainty", i.e. minimal volume of the 3D error ellipsoid. This requirement can be stated as

$$det\,\mathbf{\Sigma}_g \rightarrow min. \tag{6}$$

Since $det\,\mathbf{\Sigma}_g = 1/det\,\mathbf{\Sigma}_g^{-1}$ and $\mathbf{\Sigma}_g$ is the inverse of $\mathbf{C}_g$ (up to a factor) we see that Eq. (6) is equivalent to

$$det\,\mathbf{C}_g \rightarrow max, \tag{7}$$

which indeed is the operator $Op4$ in Eq. (1). Thus, this operator extracts 3D points with "minimal localization uncertainty", i.e. "highest possible localization precision". Note that $\mathbf{C}_g$ and thus $\mathbf{\Sigma}_g$ can directly be computed from the image data. Analogous, with Eqs. (1) and (3), we see that the operator $Op3'$ minimizes $tr\,\mathbf{\Sigma}_g$ which is the sum of the squared semi-axes of the 3D error ellipsoid (for the 2D case see also Ref. [11]),

$$tr\,\mathbf{\Sigma}_g = \sigma_x^2 + \sigma_y^2 + \sigma_z^2. \tag{8}$$

If the semi-axes are small, then generally also the error ellipsoid is small and *vice versa*. Therefore, this operator can be seen as an approximate measure for the volume of the ellipsoid. Finally, $Op3$ is equivalent to (see also Section 3 below)

$$\tfrac{1}{2}((tr\,\mathbf{\Sigma}_g)^2 - tr\,\mathbf{\Sigma}_g^2) = \sigma_x^2\sigma_y^2 + \sigma_x^2\sigma_z^2 + \sigma_y^2\sigma_z^2, \tag{9}$$

which can further be written as

$$\frac{1}{\pi^2}((\pi\sigma_x\sigma_y)^2 + (\pi\sigma_x\sigma_z)^2 + (\pi\sigma_y\sigma_z)^2). \tag{10}$$

Now we can see, that the three terms in the parentheses represent the areas of three (2D) ellipses with corresponding semi-axes, and the total expression is the sum of the squared areas of the ellipses. Actually, these ellipses are the three orthogonal sections of the 3D error ellipsoid determined through the directions of the eigenvectors of $\mathbf{\Sigma}_g$. Thus, also $Op3$ represents an approximation to the volume of the 3D error ellipsoid. In summary, all three operators $Op3$, $Op3'$, and $Op4$ can be interpreted as measures for the size of the 3D error ellipsoid, using either the sum of (squared) sectional ellipse areas, the sum of the (squared) lengths of the semi-axes, or the (squared) volume of the ellipsoid, respectively.

As an example, in Fig. 1, we show the estimated 3D error ellipsoid for the landmark 'genu of corpus callosum' within a 3D MR data set of the human brain. The 3D error ellipsoid has been displayed by three orthogonal views (sagittal, axial, coronal) and corresponds to the 1σ confidence regions. Note that the ellipsoid has been enlarged by a factor of 30 for visualization purposes. We have used a region-of-interest (ROI) of $21 \times 21 \times 21$ voxels centered around the detected landmark position, and $\sigma_n^2 = 25$ has been chosen for the variance of the image noise. In Fig. 1, we see within the sagittal plane that the localization uncertainty along the edge of the object is much larger than the localization uncertainty perpendicular to the edge. This is what we expect from the local intensity structure of this landmark. The axial and coronal views of the 3D error ellipsoid indicate an approximately isotropic localization uncertainty,

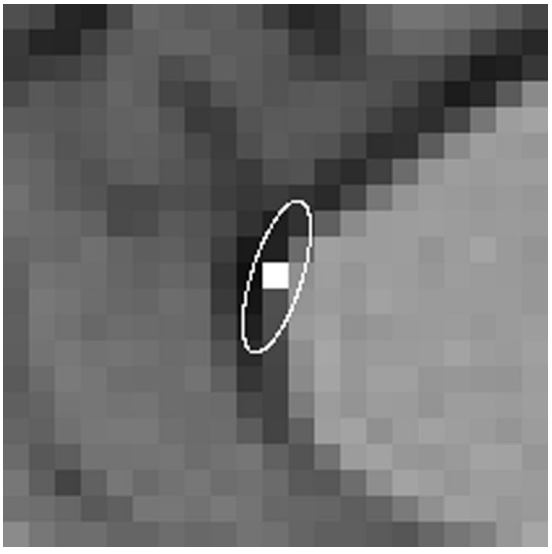

Fig. 1. Estimated 3D error ellipsoid (enlarged by a factor of 30) for the landmark genu of corpus callosum within a 3D MR data set of the human brain (orthogonal views: sagittal, axial, coronal).

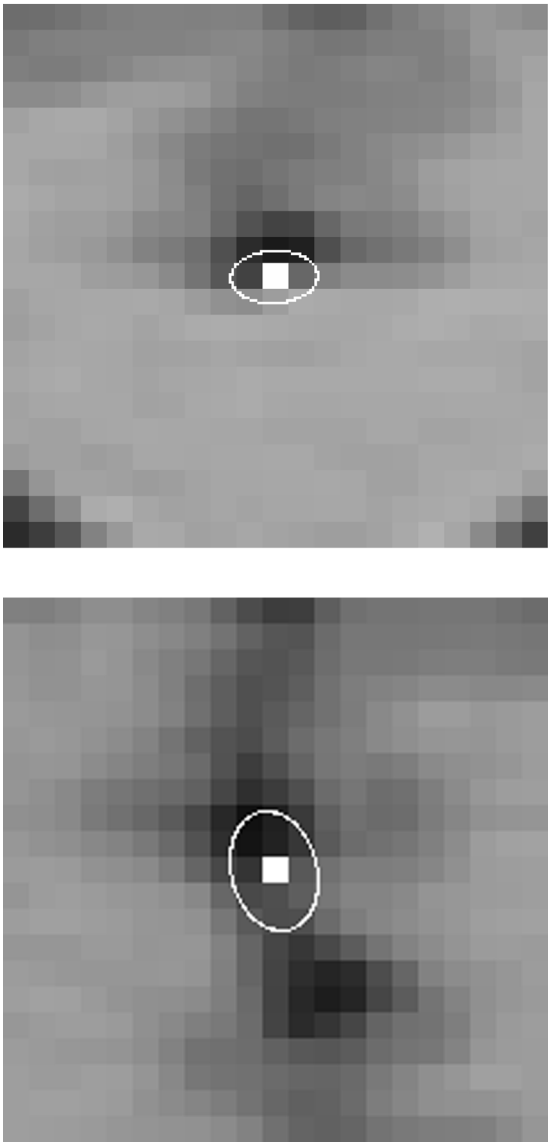

Fig. 1. (*Continued*).

where the uncertainty within the coronal plane is larger than in the axial plane. Another example is given in Fig. 2 which shows the estimated 3D error ellipsoid for the landmark "lower cusp between pons and medulla oblongata".

3. Relation to principal invariants

In this section, we show that the 3D differential operators described above can be derived by applying invariance principles. Invariance (absolute or relative invariance) is an important property which indicates preservation under certain types of transformations (e.g. rotations). Our analysis is based on the covariance matrix $\mathbf{\Sigma}_g$ as introduced above and which represents the minimal positional uncertainty (Cramér–Rao bound). We will consider the principal invariants of this matrix.

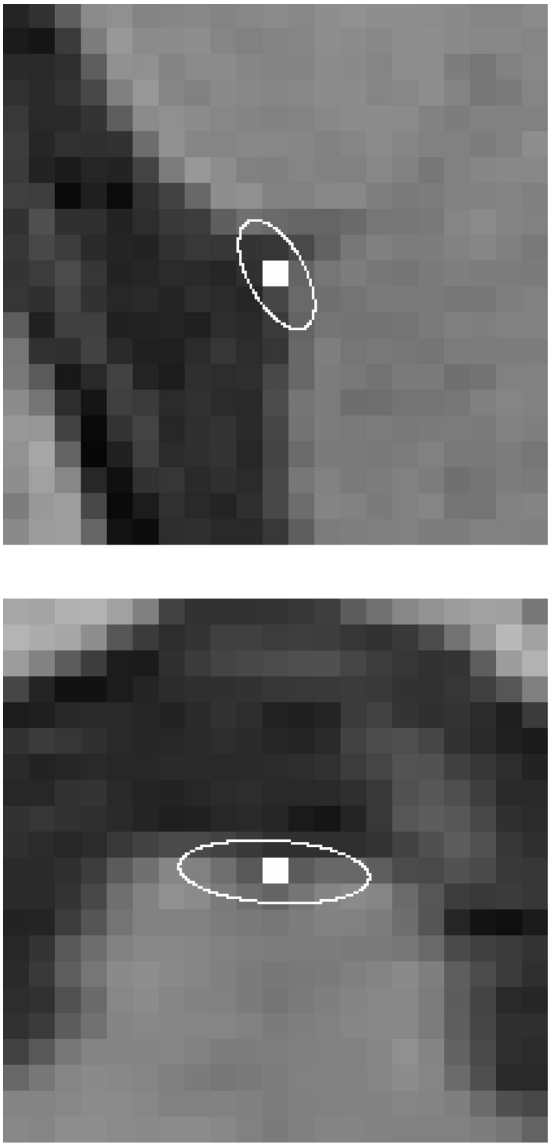

Fig. 2. Same as Fig. 1 but for the landmark lower cusp between pons and medulla oblongata.

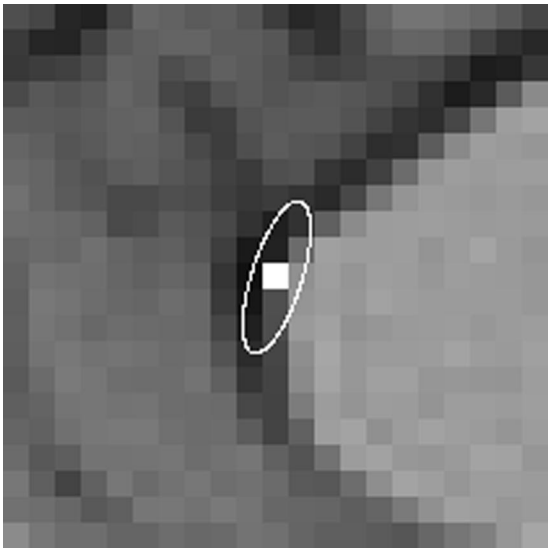

Fig. 2. (*Continued*).

3.1. Principal invariants of a matrix

Generally, the principal invariants of a $d \times d$ matrix $\mathbf{B} = (b_{ij})$ are the coefficients $\iota_1, \ldots, \iota_d$, also denoted $\iota_1(\mathbf{B}), ..., \iota_d(\mathbf{B})$ if we wish to make the dependence on $\mathbf{B}$ explicit. These coefficients appear in the characteristic polynomial of $\mathbf{B}$:

$$det(\mathbf{B} - \lambda\mathbf{I}) = (-1)^d\lambda^d + (-1)^{d-1}\iota_1(\mathbf{B})\lambda^{d-1} + \cdots - \iota_{d-1}(\mathbf{B})\lambda + \iota_d(\mathbf{B}), \quad (11)$$

where $\mathbf{I}$ is a $d \times d$ unity matrix, such that $\iota_1(\mathbf{B}) = tr\,\mathbf{B}$, $\iota_{d-1}(\mathbf{B}) = tr(cof\mathbf{B})$ ($= det\,\mathbf{B}\,tr\mathbf{B}^{-1}$ if $\mathbf{B}$ is invertible), and $\iota_d(\mathbf{B}) = det\,\mathbf{B}$ [14]. The complete set of the principal invariants can be abbreviated as $\iota_\mathbf{B} = (\iota_1(\mathbf{B}), ..., \iota_d(\mathbf{B}))$. The $\iota_i(\mathbf{B})$ are invariant under similarity transformations (translation, rotation, scale) and they are independent of each other.

In the case of a 3×3 matrix we have the coefficients $\iota_\mathbf{B} = (\iota_1(\mathbf{B}), \iota_2(\mathbf{B}), \iota_3(\mathbf{B}))$ which appear in the characteristic polynomial

$$det(\mathbf{B} - \lambda\mathbf{I}) = -\lambda^3 + \iota_1(\mathbf{B})\lambda^2 - \iota_2(\mathbf{B})\lambda + \iota_3(\mathbf{B}). \quad (12)$$

If we denote the eigenvalues of $\mathbf{B}$ by $\lambda_1, \lambda_2, \lambda_3$ then the following relations can be deduced:

$$\iota_1(\mathbf{B}) = tr\,\mathbf{B} = \lambda_1 + \lambda_2 + \lambda_3,$$

$$\iota_2(\mathbf{B}) = \tfrac{1}{2}\left((tr\,\mathbf{B})^2 - tr\,\mathbf{B}^2\right) = \lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_1\lambda_3,$$

$$\iota_3(\mathbf{B}) = det\,\mathbf{B} = \lambda_1\lambda_2\lambda_3. \tag{13}$$

Other examples of invariants are $tr\,\mathbf{B}^2$ or $tr\,\mathbf{B}^3$; however, these invariants are not principal invariants. In the 2D case, i.e. for a 2×2 matrix, we have $det(\mathbf{B} - \lambda\mathbf{I}) = \lambda^2 - \iota_1(\mathbf{B})\lambda + \iota_2(\mathbf{B})$ with the two principal invariants $\iota_1(\mathbf{B}) = tr\,\mathbf{B}$ and $\iota_2(\mathbf{B}) = det\,\mathbf{B}$.

3.2. Interpretation of the 3D differential operators

The covariance matrix $\mathbf{\Sigma}_g$ given in Eq. (3) transforms under similarity transformations $\tilde{\mathbf{x}} = \mathbf{A}\mathbf{x}$ as $\tilde{\mathbf{\Sigma}}_g = \mathbf{A}\mathbf{\Sigma}_g\mathbf{A}^T$. In the 3D case, according to Eq. (13), the three principal invariants are

$$\iota_1(\mathbf{\Sigma}_g) = tr\,\mathbf{\Sigma}_g, \qquad \iota_2(\mathbf{\Sigma}_g) = \tfrac{1}{2}\left((tr\,\mathbf{\Sigma}_g)^2 - tr\,\mathbf{\Sigma}_g^2\right),$$

$$\iota_3(\mathbf{\Sigma}_g) = det\,\mathbf{\Sigma}_g. \tag{14}$$

Regarding the first principal invariant and knowing from Eq. (3) that $\mathbf{C}_g$ is the inverse of $\mathbf{\Sigma}_g$ (up to a factor) it follows that $\iota_1(\mathbf{\Sigma}_g)$ is equivalent to $Op3' = 1/tr\,\mathbf{C}_g^{-1}$ in Eq. (1). Note that for localizing landmarks, we compute local maxima of the operator values. Thus, constant factors of the operators are not relevant. Also note, that maximizing an expression is equivalent to minimizing the inverse of it. Considering $\iota_2(\mathbf{\Sigma}_g)$ in order to derive a relation to our 3D operators, we first rewrite the second principal invariant as $\iota_2(\mathbf{\Sigma}_g) = det\,\mathbf{\Sigma}_g tr\,\mathbf{\Sigma}_g^{-1}$. Since $det\,\mathbf{\Sigma}_g = 1/det\,\mathbf{\Sigma}_g^{-1}$ we have $\iota_2(\mathbf{\Sigma}_g) = tr\,\mathbf{\Sigma}_g^{-1}/det\,\mathbf{\Sigma}_g^{-1}$. Thus, $\iota_2(\mathbf{\Sigma}_g)$ is equivalent to $Op3 = det\,\mathbf{C}_g/tr\,\mathbf{C}_g$. Finally, again since $det\,\mathbf{\Sigma}_g = 1/det\,\mathbf{\Sigma}_g^{-1}$ and due to Eq. (3), it follows that the third principal invariant $\iota_3(\mathbf{\Sigma}_g)$ is equivalent to $Op4 = det\,\mathbf{C}_g$. In summary, the 3D operators $Op3$, $Op3'$, and $Op4$ in Eq. (1) are equivalent to the principal invariants of the covariance matrix $\mathbf{\Sigma}_g$. The operators represent the complete set of principal invariants of this matrix. With $\propto$ denoting proportional or inverse proportional, we can write

$$\iota_1(\mathbf{\Sigma}_g) \propto \frac{1}{tr\,\mathbf{C}_g^{-1}}, \qquad \iota_2(\mathbf{\Sigma}_g) \propto \frac{det\,\mathbf{C}_g}{tr\,\mathbf{C}_g}, \qquad \iota_3(\mathbf{\Sigma}_g) \propto det\,\mathbf{C}_g. \tag{15}$$

Note, that $\iota_1(\mathbf{\Sigma}_g)$ can also be expressed exclusively in terms of the matrix $\mathbf{C}_g$ as $\iota_1(\mathbf{\Sigma}_g) \propto 2\,det\,\mathbf{C}_g/((tr\,\mathbf{C}_g)^2 - tr\,\mathbf{C}_g^2)$.

If instead, we take the matrix $\mathbf{C}_g$ as the basis of our analysis and consider the principal invariants of this matrix, then we obtain

$$\iota_1(\mathbf{C}_g) = tr\,\mathbf{C}_g, \qquad \iota_2(\mathbf{C}_g) = \tfrac{1}{2}\left((tr\,\mathbf{C}_g)^2 - tr\,\mathbf{C}_g^2\right),$$

$$\iota_3(\mathbf{C}_g) = det\,\mathbf{C}_g. \tag{16}$$

We immediately see that $\iota_3(\mathbf{C}_g)$ is equivalent to $\iota_3(\mathbf{\Sigma}_g)$, while the other two invariants are different from the invariants of $\mathbf{\Sigma}_g$. Thus, we have two additional invariants and in total there are five different principal invariants. All invariants are given in terms of the matrix $\mathbf{C}_g$ which is easier to implement in comparison to using the matrix $\mathbf{\Sigma}_g$. However, since only $\mathbf{\Sigma}_g$ has a sound statistical interpretation in terms of the minimal uncertainty of the position estimate, we favor those operators which are based on this matrix, i.e. $\iota_1(\mathbf{\Sigma}_g)$, $\iota_2(\mathbf{\Sigma}_g)$, and $\iota_3(\mathbf{\Sigma}_g)$.

In the 2D case, we have the principal invariants $\iota_1(\mathbf{\Sigma}_g) = tr\,\mathbf{\Sigma}_g$ and $\iota_2(\mathbf{\Sigma}_g) = det\,\mathbf{\Sigma}_g$ which are indeed equivalent to the 2D corner detectors in Refs. [11] and [12]:

$$\iota_1(\mathbf{\Sigma}_g) \propto \frac{det\,\mathbf{C}_g}{tr\,\mathbf{C}_g},\ \iota_2(\mathbf{\Sigma}_g) \propto det\,\mathbf{C}_g. \tag{17}$$

Considering the matrix $\mathbf{C}_g$ we have

$$\iota_1(\mathbf{C}_g) = tr\,\mathbf{C}_g, \qquad \iota_2(\mathbf{C}_g) = det\,\mathbf{C}_g, \tag{18}$$

where $\iota_2(\mathbf{C}_g)$ is equivalent to $\iota_2(\mathbf{\Sigma}_g)$. Thus, overall we have three different principal invariants of $\mathbf{\Sigma}_g$ and $\mathbf{C}_g$ in the 2D case.

As an example for the 3D case, we show an application of the invariant $\iota_2(\mathbf{\Sigma}_g)$ which is equivalent to the operator $Op3$. We have used a 3D MR image of the human brain consisting of 68 slices of 165×210 resolution. The total number of detected points is about 1000. The result for a number of slices is shown in Fig. 3 (slices 7, 20, 36, 37, 38, and 48). It can be seen that significant structures have been detected. For example, in slices 7, 20, and 48 tips of the ventricular horns (see the dark elongated structure in the middle of the brain) and certain characteristic points on the sulci have been found. Additional detected landmarks are, for example, in slice 36 the anterior commissure and the genu of the corpus callosum, in slice 37 the tips between the midbrain and the pons, the posterior commissure, and the splenium of the corpus callosum, and in slice 38 the inferior colliculus (for the names of these landmarks see a standard book on human anatomy, e.g., Ref. [15]). Note, that each of these landmarks have been detected only in single slices which indicates that we have actually determined 3D points (to clearly see this behavior was the reason why we selected the three adjacent slices 36, 37, and 38). On the other hand, we also obtain some points which we cannot directly relate to points of anatomical relevance, e.g. in slices 37 and 38 in the upper and right part of the brain. Here, the problem may be the gap between the two hemispheres of the brain which is partly visible as dark areas in these slices. Anyway, in general the detection performance is very important, since the reliability and accuracy of the registration result strongly depends on this criterion. Therefore, we analyze the detection performance of the operators in more detail below.

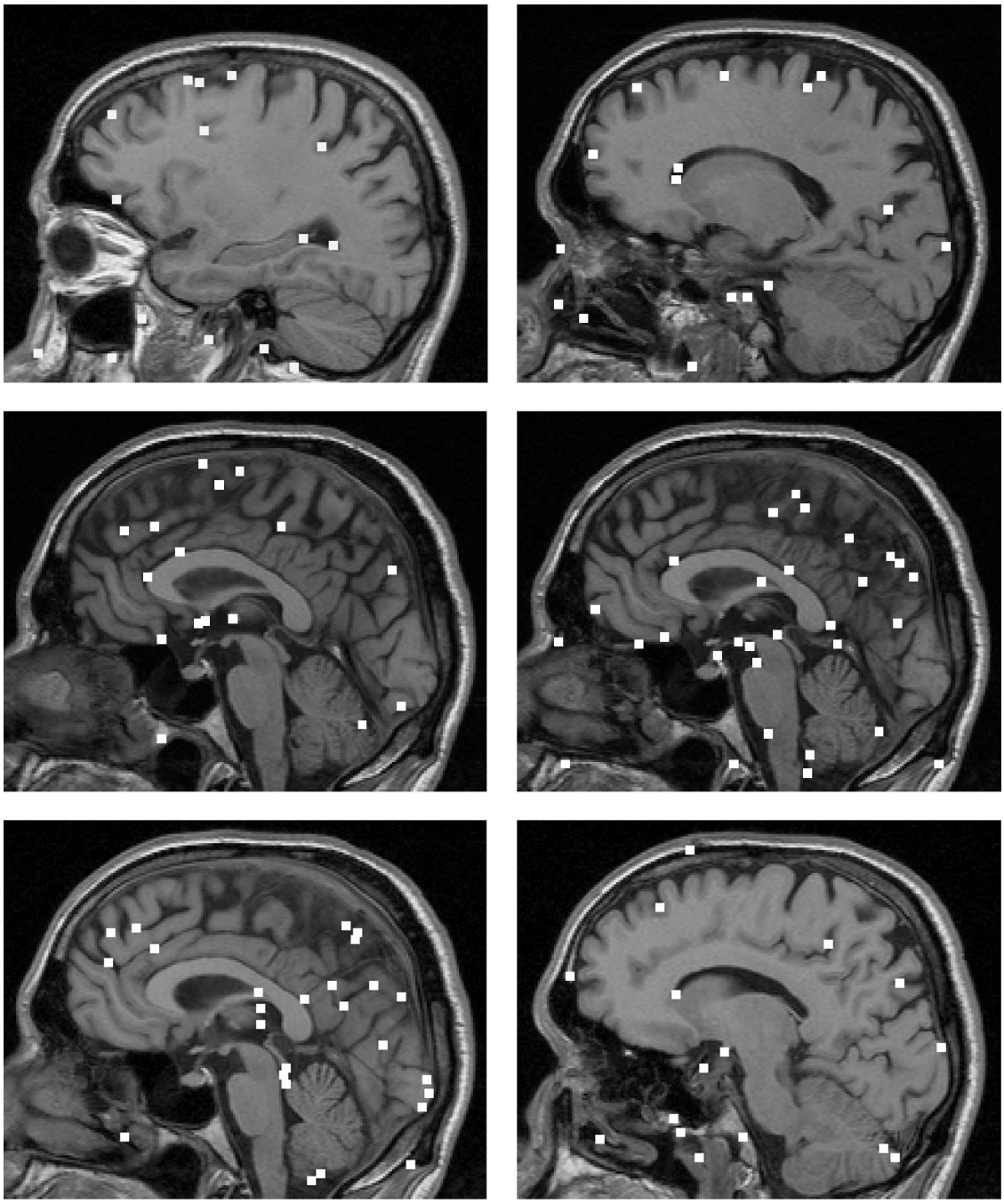

Fig. 3. Detected point landmarks for a 3D MR image of a human brain using the 3D operator $Op3$ (slices 7, 20, 36, 37, 38, and 48).

4. Detection performance

To localize anatomical landmarks in 3D tomographic images, we use a semi-automatic procedure on the basis of the 3D differential operators described above. In comparison to a fully automatic scheme such a semi-automatic approach has the advantage that the user has the possibility to control the results, which is important in clinical scenaria. In our case, we rely on the following semi-automatic procedure for localizing a certain

landmark (see also Ref. [16]. First, the user specifies a region-of-interest (ROI) together with an approximate position (e.g. the center of the ROI), and second, a 3D operator is applied yielding landmark candidates within the selected ROI. Ideally, there should be only one (correct) landmark candidate. However, due to noise and interaction effects with neighboring anatomical structures there are generally also false detections. In either case, the user should have the possiblility to accept or reject detected candidates.

4.1. Performance visualization

The described semi-automatic procedure can be applied most reliably and efficiently if the number of false detections is small and moreover if the operator response of the correct detection is much larger than the operator responses corresponding to the false detections. Note that landmark candidates with very low operator responses can be excluded by applying a low threshold. However, since still there would be false detections left such a procedure would not principally solve the problem. Also note that since we deal with 3D images which are displayed on a slice-by-slice basis in practice, it is generally very difficult to distinguish false detections from correct detections. In the 2D case, the selection is generally no problem since, for example, we can mark all detections in the original image and display the whole result on the screen. To improve the situation for 3D images, we suggest to use a certain performance visualization which gives an indication of the detection performance. With this visualization we first compute the Euclidean distances $\|\mathbf{x}_i - \mathbf{x}_{ROI}\|_2$ between the positions $\mathbf{x}_i$ of the detected candidates and the specified center of the ROI abbreviated by $\mathbf{x}_{ROI}$. Then, the computed distances are represented together with the operator responses (cf. Figs. 4–6). Thus, this representation combines the information of the number of false detections with the significance of the detections. Also, we obtain an indication of how the locations of the detections are distributed w.r.t. $\mathbf{x}_{ROI}$. However, a disadvantage is that we lose the full information of the distribution in 3D space.

4.2. A performance measure

Whereas the performance visualization described above gives a visual impression, we are also interested in quantitative measures for the detection performance. One measure is the number of false detections; however, with this measure the strength of the operator responses is not taken into account. Instead, we suggest using another measure. Suppose we have a number of n detections within a ROI and the operator responses are denoted by $R_i = Op(x_i)$, $i = 1, \ldots, n$, where the maximal operator response is $R_{max} = Op(x_{max})$. The candidates may or may not be the result after applying a threshold

$$\Omega = \{x_i \,|\, R_i \geqslant \varepsilon R_{max}\}, \tag{19}$$

where ε is user-defined, e.g. $\varepsilon = 0.01$. For this set we can compute the following measure:

$$\psi = \begin{cases} 0, & n = 0, \\ \sum_{i=1}^{n} \dfrac{R_i}{R_{max}} & n \geqslant 1. \end{cases} \tag{20}$$

If there are no detections at all then we have $\psi = 0$. Alternatively, if there is only one detection we obtain $\psi = 1$, and additional false detections with low operator responses yield a value of $\psi \approx 1$. In this case, the correct detection can clearly be distinguished from the false detections. On the other hand, if there are operator responses with similar values as the maximal operator response, then ψ is much larger than 1. Thus, in summary we here have a scalar quantity which gives an indication of the detection performance. Additionally, we can employ the mean value $\bar{\psi} = \psi/n$. In the case of several exclusively similar operator values, where we have $\psi \gg 1$, we obtain $\bar{\psi} \approx 1$. In the case of one strong response and a large number of low responses which also sum up to $\psi \gg 1$, we instead obtain $\bar{\psi} \ll 1$. Note, that for our operators the responses R_i are always larger or equal to zero. Thus, when computing ψ for other operators which possibly yield positive as well as negative responses, then absolute values of R_i should be used.

5. Experimental results

In this section, we report on experimental results concerning the detection performance and the registration accuracy.

5.1. Detection performance

The plots on the left side of Fig. 4 show an application of the detection performance visualization described in Section 4.1 above. We have used three different anatomical landmarks within a 3D MR image of the human brain (on the right side are the results for another 3D MR data set). The landmarks are located on the ventricular system and on the skull base. We have used the operator $Op3$ and Gaussian derivative filters of $\sigma = 1.5$ to compute the partial derivatives of the image. The size of the ROI has been chosen to $21 \times 21 \times 21$ voxels and the center of the ROI is the position due to best manual localization. Thus, we take the manually selected position as "ground truth", although we know that this position may be prone to error. In our case, we have alleviated this problem by using landmark positions

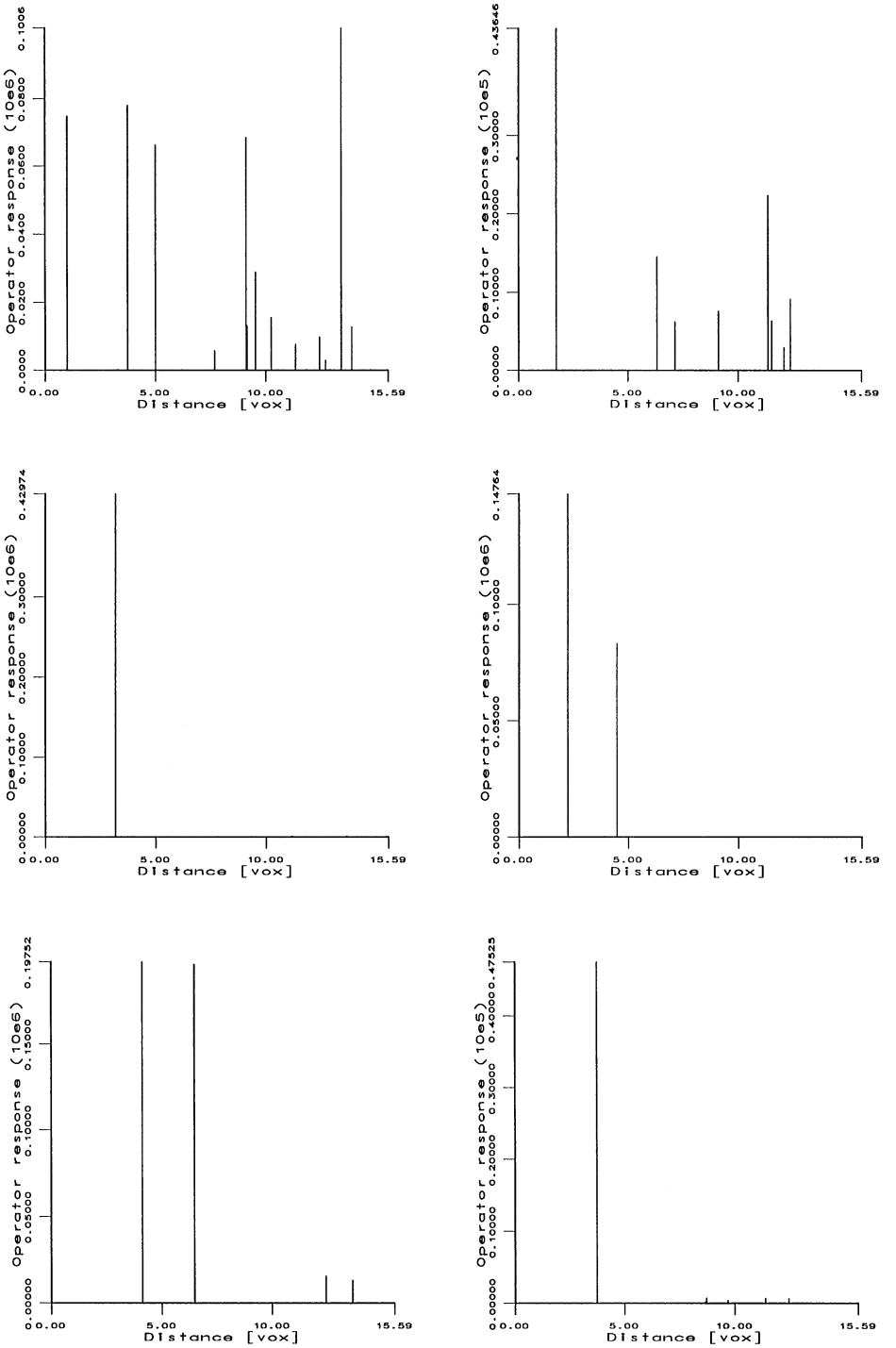

Fig. 4. Operator responses of *Op*3 for landmarks external occipital protuberance (top), left frontal ventricular horn (middle), and left occipital ventricular horn (bottom) in two different 3D MR data sets (left and right).

which have been specified in agreement with the judgement of three to four persons. To further reduce the subjectivity in our experiments we have used no thresholds at all. For the landmark on the top left of Fig. 4 (external occipital protuberance) we see that we have a detected landmark candidate close to the manually specified position. However, we also have a number of false detections with relatively large responses. Thus, the

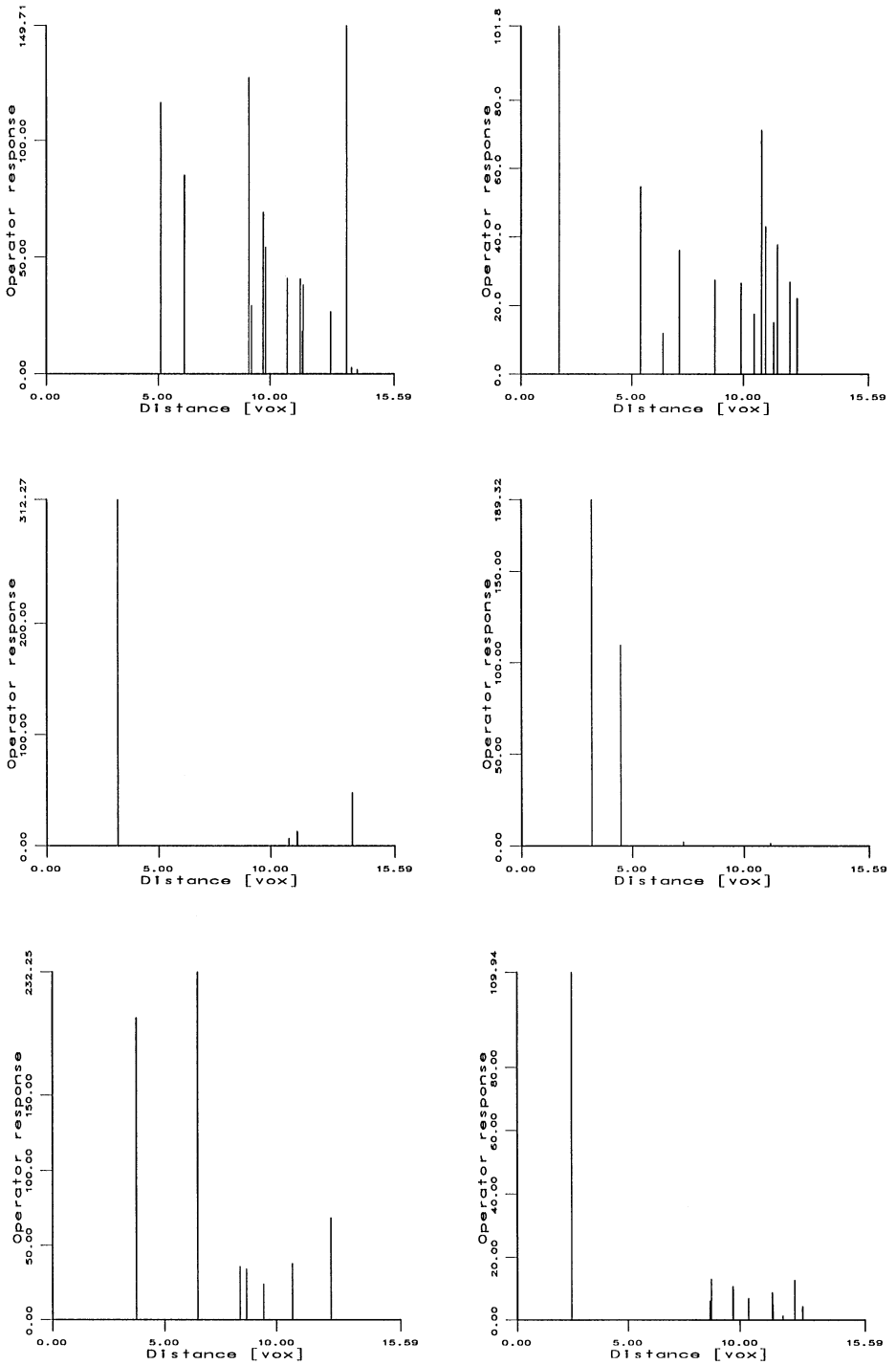

Fig. 5. Same as Fig. 4 but for operator *Op3′*.

detection performance for this landmark is relatively bad and it will be difficult and time-consuming for a user to select the correct candidate. Instead, for the landmark in the middle left of Fig. 4 (tip of left frontal ventricular horn) we have a perfect detection result with only one strong candidate within the whole 3D ROI. For the landmark on the bottom left of Fig. 4 (tip of left occipital ventricular horn), we essentially have two

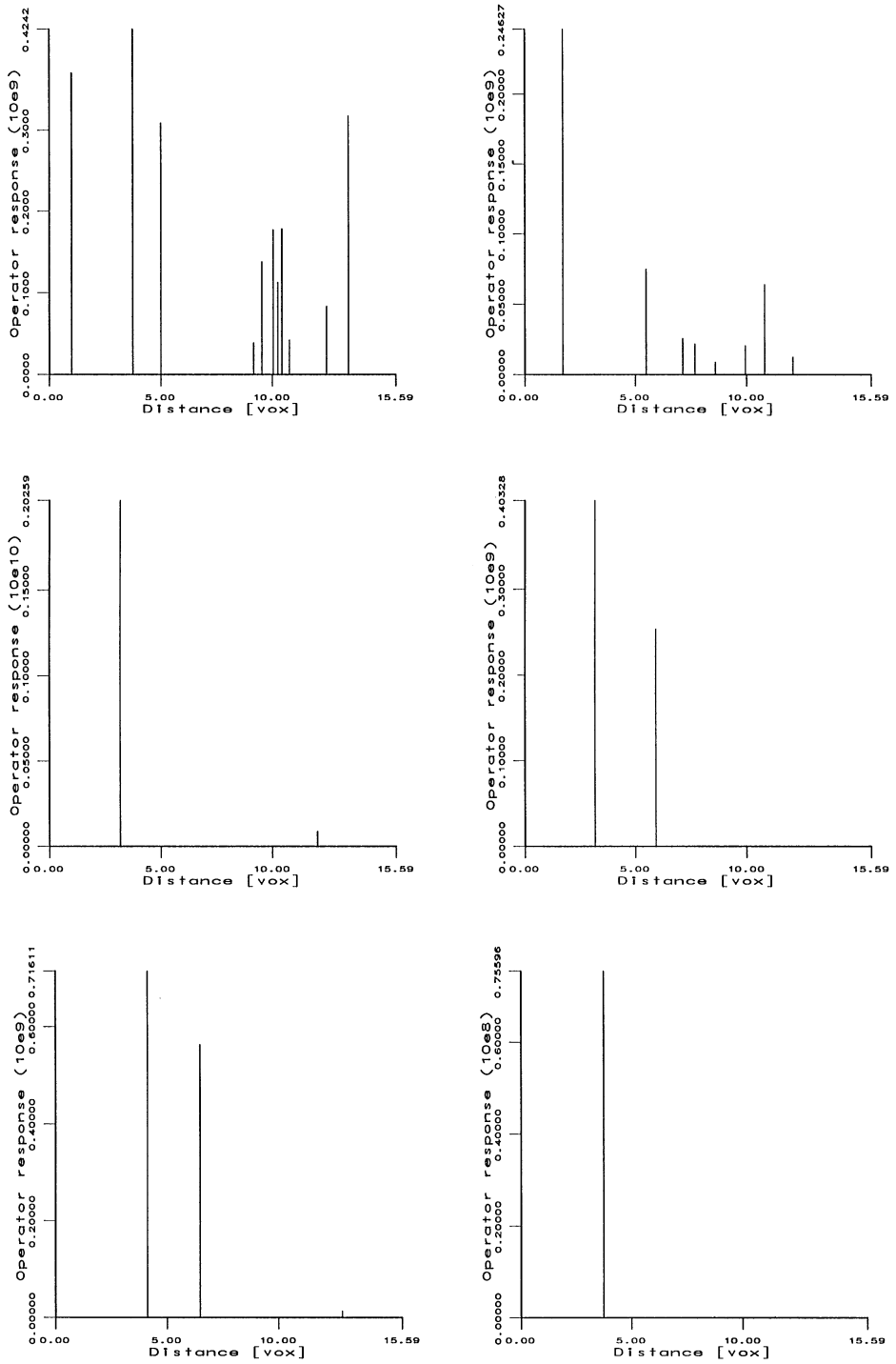

Fig. 6. Same as Fig. 4 but for operator *Op*4.

strong operator responses of similar values, where one of them is closer to the manually specified position. Inspection of the original image reveals that the double detection due to the individual anatomy of the imaged person. Here, the occipital ventricular horn of this person actually is a 'double horn' consisting of two tips. Therefore, in this case both landmarks are suited for use in registration and the user has to decide which of the

landmarks should be used. In Fig. 4 on the right are the results for the same landmarks as before but for a different 3D MR data set. Here, we see that for the first landmark (top right of Fig. 4) the detection result is much better in comparison to the previous data set. We have one strong detection close to the manually specified position. The additional false detections have lower operator responses and are farer away from the landmark position. For the second landmark (middle right) the detection result is good but it is worse in comparison to the previous data set, since here we have one false detection with a relatively large operator response. For the third landmark (bottom right), we have a perfect detection result up to some false detections with extremely low operator responses. Comparison with the previous data set shows that we here have a unique tip of the occipital ventricular horn corresponding to "normal" anatomy.

We have also applied the operators $Op3'$ and $Op4$ for the same landmarks and the same data sets as above (see Figs. 5 and 6, respectively). It turns out that the results of $Op3'$ are somewhat worse than those of $Op3$. For the landmark on the top left in Fig. 5 there is no strong detection close to the manually selected position, and for the landmark on the bottom left the detection closest to the manually selected position has a lower operator response than the one farer away. Additionally, for $Op3'$ the operator responses of the false detections with relatively low reponses are larger than those for $Op3$. Thus, the discrimination power of $Op3'$ is worse (e.g. compare the results on the bottom right of the Figs. 4 and 5). Considering the results of $Op4$ in Fig. 6 we see that this operator yields similar results as $Op3$. We have nearly the same strong detections and also the discrimination power between strong and weak detections is comparable.

For all examples above we have also computed the performance measure ψ as defined in Eq. (20). The results in Tables 1 and 2 correspond to the left and right sides of the Figs. 4–6, respectively. It turns out that the computed quantities are in well agreement with our observations in Figs. 4–6. First of all, we see that the detection performance for the first landmark is much worse in comparison to the other landmarks. Second, the operator $Op3'$ generally yields larger values of ψ than $Op3$ as well as $Op4$ and thus the detection performance of this operator is worse. Third, it can also be seen that $Op3$ and $Op4$ yield very similar results. We also note, that the measure ψ is a much better performance characterization than the number of false detections alone. For example, for $Op3$ in Table 2 we have for the first and third landmark exactly the same number of false detections (7 false detections). However, the detection performance for the third landmark is much better than that for the first one, since for the third landmark the false detections have much lower operator reponses. This fact is clearly quantified by the corresponding values of ψ which are in accordance with the performance visualizations on the top and bottom right of Fig. 4.

Table 1
Performance measure ψ for landmarks in a 3D MR data set

Landmark	$Op3$	$Op3'$	$Op4$
External occipital protuberance	4.82	6.38	5.17
Frontal ventricular horn	1.01	1.31	1.04
Occipital ventricular horn	2.14	2.93	1.81

Table 2
Same as Table 1 but for a different 3D MR data set

Landmark	$Op3$	$Op3'$	$Op4$
External occipital protuberance	2.58	4.84	1.93
Frontal ventricular horn	1.56	1.60	1.63
Occipital ventricular horn	1.06	1.62	1.01

5.2. Registration accuracy

We have also used the semi-automatically localized landmarks as features for the registration of 3D MR and CT images of the human brain. In the following experiment, we have applied the operator $Op3$ and Gaussian derivative filters of $\sigma = 1.5$. The investigated MR data set consists of 120 axial slices of 256×256 resolution, and the CT data set consists of 87 axial slices of 320×320 resolution. The images represent the brain of the same person. We have used the following anatomical point landmarks: topmost concavity of fourth ventricle roof, external occipital protuberance, tip of frontal ventricular horn, tip of occipital ventricular horn, tip of temporal ventricular horn, saddle point at zygomatic bone, as well as saddle point at processus mastoideus. The last five landmarks can be found in both hemispheres of the human brain. Note, that for the used CT data set the tip of the temporal ventricular horn was siginificantly worse pronounced in comparison to the other landmarks, therefore, we did not include this landmark. In total, we thus have a number of 11 homologous landmarks. To assess the registration accuracy of the localized landmarks we have applied an optimal affine transformation model. We hereby assume that an affine transformation model well approximates the true transformation. Note also that the registration accuracy only checks the global consistency of the landmarks according to the applied transformation model. Anyway, using 11 landmarks and an optimal affine transformation model we obtained a good registration result with a mean Euclidean error at the landmarks of $\bar{e} \approx 1.5$ voxels, where the individual errors at the landmarks varied between 0.4 and 2.6 voxels.

6. Summary and future work

We have investigated 3D differential operators for extracting anatomical point landmarks in 3D tomographic images of the human brain. These operators only employ first-order partial derivatives of an image and are therefore computationally efficient and do not suffer from instabilities of computing high order partial derivatives. To improve the predictability of the operators we have analyzed certain properties of them. It has been shown that the operators can be interpreted in terms of the Cramér–Rao bound which represents the minimal localization uncertainty. All operators can be related to the geometry of the 3D error ellipsoid and they essentially represent the size of the ellipsoid. Moreover, we have shown that the operators can be derived on the basis of invariance principles. It turned out, that the operators form a complete set of principal invariants under similarity transformations. As a basis of our analysis we have used the covariance matrix of the position estimate due to the Cramér–Rao bound. Taking the inverse of that matrix as a basis, we additionally obtained two invariants in the 3D case and one in the 2D case.

We have also analyzed the detection performance of the operators. We have introduced and applied a certain type of performance visualization which combines the number of false detections with the significance of the detections. Also, we have suggested a scalar measure of the detection performance. It turned out, that this measure gives a good indication of the detection capability and that it is more appropriate than using the number of false detections alone as a detection performance criterion. Note that both the performance visualization as well as the performance measure can be applied within a semi-automatic procedure to guide the user in selecting the correct landmark candidates. Investigation of the different operators revealed that two operators, namely $Op3$ and $Op4$, yield superior results. We also saw that the detection performance is generally different for different landmarks.

Besides the detection performance, the localization accuracy is of primary interest. Recently, an extension of our operators has been reported in Ref. [17] which allows for subvoxel localization using multi-step differential approaches. The extracted landmarks serve as features for a point-based registration scheme. Note, that the information of the landmark error ellipsoids can also be incorporated as input for registration [18]. In future work, it will be important to analyze in detail the influence of landmark extraction on the registration result.

Acknowledgement

I thank S. Frantz and H.S. Stiehl for discussions and critical comments on the manuscript. This work has been supported by Philips Research Hamburg, project IMAGINE (IMage- and Atlas-Guided Interventions in NEurosurgery). The original images have kindly been provided by W. P. Th. M. Mali, L. Ramos, and C. W. M. van Veelen (Utrecht University Hospital) via ICS-AD of Philips Medical Systems Best and Philips Research Hamburg.

References

[1] G.T.Y. Chen, M. Kessler, S. Pitluck, Structure transfer between sets of three dimensional medical imaging data, Proc. 6th Annual Conf. on Computer Graphics, Vol. 6 (1985) pp. 171–177.

[2] F.L. Bookstein, Principal warps: thin-plate splines and the decomposition of deformations, IEEE Trans. Pattern Anal. Machine Intell. 11(6) (1989) 567–85.

[3] A.C. Evans, W. Dai, L. Collins, P. Neelin, S. Marrett, Warping of a computerized 3-D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis, in Medical Imaging V: Image Processing, San Jose, California/USA, Proc. SPIE 1445, ed. M.H. Loew, 1991, pp. 236–246.

[4] D.L.G. Hill, D.J. Hawkes, J.E. Crossman, M.J. Gleeson, T.C.S. Cox, E.E.C.M. Bracey, A.J. Strong, P. Graves, Registration of MR and CT images for skull base surgery using point-like anatomical features, Br. J. Radiol. 64 (767) (1991) 1030–1035.

[5] J.-P. Thirion, Extremal points: definition and application to 3D image registration, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Seattle/Washington, USA, 587–592, June 1994, 21–23.

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

[7] O. Monga, R. Lengagne, R. Deriche, Extraction of the zero-crossings of the curvature derivatives in volumic 3D medical images: a multi-scale approach, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Seattle/Washington, U.S.A., June 21–23, 1994, pp. 852–855.

[8] W. Beil, K. Rohr, H.S. Stiehl, Investigation of approaches for the localization of anatomical landmarks in 3D medical images, Proc. Computer Assisted Radiology and Surgery (CAR'97), Berlin, Germany, eds H.U. Lemke, M.W. Vannier, K. Inamura June 25–28 (1997), Elsevier, Amsterdam, 1997, pp. 265–270.

[9] J.B.A. Maintz, P.A. van den Elsen, M.A. Viergever, Comparison of feature-based matching of CT and MR brain images Proc. 1st Internat. Conf. Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed'95), Nice/France, April (1995), Lecture Notes in Computer Science 905, ed. N. Ayache, Springer, Berlin, 1995, pp. 219–228.

[10] T. Hartkens, K. Rohr, H.S. Stiehl, Evaluierung von Differentialoperatoren zur Detektion charakteristischer Punkte in tomographischen Bildern, 18. DAGM-Symposium Mustererkennung, 11–13. September 1996, Heidelberg, Germany, Informatik aktuell, eds B. Jähne, P. Geißler,

H. Haußecker, F. Hering, Springer, Berlin, 1996, pp. 637–644.
[11] W. Förstner, A feature based correspondence algorithm for image matching, Int. Arch. Photogrammetry Remote Sensing 26(3/3) (1986) 150–166.
[12] K. Rohr, Untersuchung von grauwertabhängigen Transformationen zur Ermittlung des optischen Flusses in Bildfolgen, Diplomarbeit, Institut für Nachrichtensysteme, Universität Karlsruhe, 1987.
[13] H.L. van Trees, Detection, Estimation, and Modulation Theory, Part I, Wiley, New York, 1968.
[14] P. G. Ciarlet, Mathematical Elasticity, Vol. I: Three-Dimensional Elasticity. North-Holland Amsterdam 1988.
[15] J. Sobotta, Atlas der Anatomie des Menschen, 1. Band: Kopf, Hals, Obere Extremintäten, eds H. Ferner and J. Staubesand. Urban und Scharzenberg, München, 1982.
[16] K. Rohr, H.S. Stiehl, R. Sprengel, W. Beil, T.M. Buzug, J. Weese, M.H. Kuhn, Point-based elastic registration of medical image data using approximating thin-plate splines, Proc. 4th Int. Conf. Visualization in Biomedical Computing (VBC'96), Hamburg, Germany, 22–25 September, 1996, Lecture Notes in Computer Science Vol. 1131, eds K.H. Höhne, R. Kikinis, Springer, Berlin, 1996, pp. 297–306.
[17] S. Frantz, K. Rohr, H.S. Stiehl, On the localization of 3D anatomical point landmarks in medical imagery using multi-step differential approaches, 19. DAGM-Symposium Mustererkennung, Braunschweig, Germany, 15–17 September (1997), Informatik aktuell, eds E. Paulus, F.M. Wahl, Springer, Berlin, 1997 pp. 340–347.
[18] K. Rohr, R. Sprengel, H.S. Stiehl, Incorporation of landmark error ellipsoids for image registration based on approximating thin-plate splines, Proc. Computer Assisted Radiology and Surgery (CAR'97), Berlin, Germany, 25–28 June (1997), eds H.U. Lemke, M.W. Vannier and K. Inamura, Elsevier Amsterdam, 1997, pp. 234–239.

About the Author—KARL ROHR received his Diploma degree in electrical engineering from the University of Karlsruhe, Germany, in 1987, and his Ph. D. degree in computer science from the University of Hamburg, Germany, in 1994. From 1988 to 1991 he was with the Department of Computer Science, University of Karlsruhe, and currently he is with the Department of Computer Science, University of Hamburg, Germany. In 1990 Dr. Rohr was awarded a DAGM prize for his paper on model-based recognition of grey-value corners. In 1995 he received the Springer Best Paper Award KI-95 for his joint work on model-based recognition and natural language description of human movements. His research interests include computer vision, medical image analysis, and computer graphics.