Proc. *25th Pattern Recognition Symposium (DAGM'03)*, Magdeburg, Germany,
10th-12th September 2003, In *Lecture Notes in Computer Science*,
Springer Verlag Berlin Heidelberg 2003, 220-227,

3D Parametric Intensity Models for the Localization of Different Types of 3D Anatomical Point Landmarks in Tomographic Images

Stefan Wörz, Karl Rohr

School of Information Technology, Computer Vision & Graphics Group
International University in Germany, 76646 Bruchsal
Email: {woerz,rohr}@i-u.de

Abstract. We introduce a new approach for the localization of 3D anatomical point landmarks based on 3D parametric intensity models which are directly fit to an image. We propose different analytic intensity models based on the Gaussian error function in conjunction with 3D rigid transformations as well as deformations to efficiently model tip-like, saddle-like, and sphere-like structures. The approach has been successfully applied to accurately localize anatomical landmarks in 3D MR and 3D CT image data. We have also compared the experimental results with the results of a previously proposed 3D differential operator. It turns out that the new approach significantly improves the localization accuracy.

1 Introduction

The localization of 3D anatomical point landmarks is an important task in medical image analysis. Landmarks are useful image features in a variety of applications, for example, for the registration of 3D brain images of different modalities or the registration of images with digital atlases. The current standard procedure, however, is to localize 3D anatomical point landmarks manually which is difficult, time consuming, and error-prone. To improve the current situation it is therefore important to develop automated methods.

In previous work on the localization of 3D anatomical point landmarks, 3D differential operators have been proposed (e.g., Thirion [12], Rohr [10]). Recently, an evaluation study of nine different 3D differential operators has been performed by Hartkens *et al.* [7]. 2D differential approaches for extracting point landmarks in 2D medical images have been described in Briquer *et al.* [3] and Hartkens *et al.* [6]. While being computationally efficient, differential operators incorporate only small local neighbourhoods of an image and are therefore relatively sensitive to noise, which leads to false detections and also affects the localization accuracy. Recently, an approach based on deformable models was introduced (Frantz *et al.* [4], Alker *et al.* [1]). With this approach tip-like anatomical structures are modeled by *surface* models, which are fit to the image data using an edge-based fitting measure. However, the approach requires the detection of 3D image edges as well as the formulation of a relatively complicated fitting measure.

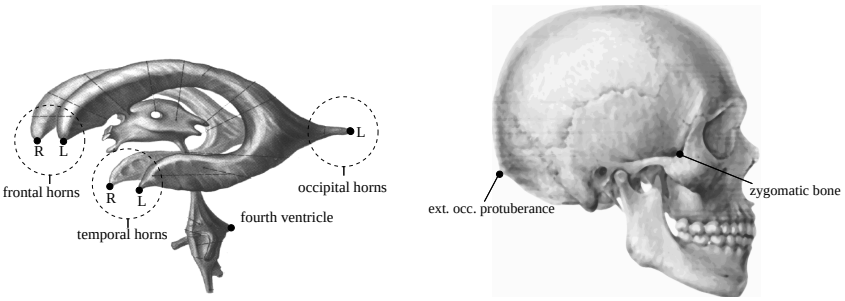

Fig. 1. Ventricular horns of the human brain (from [11]) and the human skull (from [2]). Examples of 3D point landmarks are indicated by black dots.

We have developed a new approach for the localization of 3D anatomical point landmarks. In contrast to previous approaches the central idea is to use 3D parametric *intensity* models of anatomical structures. In comparison to differential approaches, larger image regions and thus semi-global image information is taken into account. In comparison to approaches based on surface models, we directly exploit the intensity information of anatomical structures. Therefore, more a priori knowledge and much more image information is taken into account in our approach to improve the robustness against noise and to increase the localization accuracy. Also, a much simpler fitting measure can be used.

2 Parametric Intensity Models for Anatomical Structures

Our approach uses 3D parametric intensity models which are fit directly to the intensities of the image data. These models describe the image intensities of anatomical structures in a semi-global region as a function of a certain number of parameters. The main characteristic in comparison to general deformable models is that they exhibit a prominent point which defines the position of the landmark. By fitting a parametric intensity model to the image intensities we obtain a *subvoxel* estimate of the position as well as estimates of the other parameters, e.g., the image contrast. As important classes of 3D anatomical point landmarks we here consider tip-like, saddle-like, and sphere-like structures.

3D Intensity Model of Tip-Like Structures Tip-like structures can be found, for example, within the human head at the ventricular system (e.g., the tips of the frontal, occipital, or temporal horns, see Fig. 1) and at the skull (e.g., the tip of the external occipital protuberance). The shape of these anatomical structures is ellipsoidal. Therefore, to model them we use a (half-)ellipsoid defined by three semi-axes (r_x, r_y, r_z) and the intensity levels a_0 (outside) and a_1 (inside). We also introduce Gaussian smoothing specified by a parameter σ to incorporate image blurring effects. The exact model of a Gaussian smoothed ellipsoid cannot be expressed in analytic form and thus is computationally expensive. To efficiently represent the resulting 3D intensity structure we developed an analytic model as an approximation. This model is based on the Gaussian error function $\Phi(x) = \int_{-\infty}^{x} (2\pi)^{-1/2} \, e^{-\xi^2/2} \, d\xi$ and can be written as

$$g_{Ell.}(\mathbf{x}) = a_0 + (a_1 - a_0)\,\Phi\left(\frac{\sqrt[3]{r_x r_y r_z}}{\sigma}\left(1 - \sqrt{\frac{x^2}{r_x^2} + \frac{y^2}{r_y^2} + \frac{(z+r_z)^2}{r_z^2}}\right)\right) \quad (1)$$

where $\mathbf{x} = (x, y, z)$. We define the tip of the ellipsoid w.r.t. the semi-axis r_z as the position of the landmark, which also is the center of the local coordinate system. In addition, we include a 3D rigid transform $\mathcal{R}$ with rotation parameters (α, β, γ) and translation parameters (x_0, y_0, z_0). Moreover, we extend our model to a more general class of tip-like structures by applying a tapering deformation $\mathcal{T}$ with the parameters ρ_x and ρ_y, and a bending deformation $\mathcal{B}$ with the parameters δ (strength) and ν (direction), which are defined by

$$\mathcal{T}(\mathbf{x}) = \begin{pmatrix} x\,(1 + z\,\rho_x/r_z) \\ y\,(1 + z\,\rho_y/r_z) \\ z \end{pmatrix} \quad \text{and} \quad \mathcal{B}(\mathbf{x}) = \begin{pmatrix} x - z^2\delta\cos\nu \\ y - z^2\delta\sin\nu \\ z \end{pmatrix} \quad (2)$$

This results in the parametric intensity model with a total of 16 parameters:

$$g_M(\mathbf{x}, \mathbf{p}) = g_{Ell.}(\mathcal{T}(\mathcal{B}(\mathcal{R}(\boldsymbol{x})))) \quad (3)$$

$$\mathbf{p}_{Ell.} = (r_x, r_y, r_z, a_0, a_1, \sigma, \rho_x, \rho_y, \delta, \nu, \alpha, \beta, \gamma, x_0, y_0, z_0) \quad (4)$$

3D Intensity Model of Saddle-Like Structures Saddle-like structures can be found, for example, within the human head at the zygomatic bone (see Fig. 1). These structures can be modelled by a bended ellipsoid where the bending is symmetrical w.r.t. the center of the ellipsoid. Therefore, we modify (1) such that the center of the local coordinate system is localized at the tip of the ellipsoid w.r.t the semi-axis r_x. By restricting the direction of the bending deformation $\mathcal{B}$ towards the x-axis, i.e. setting $\nu = 0$ in (2), we achieve a saddle-like structure where the curvature of the bending is maximal at the center of the local coordinate system. This defines the position of the landmark. Besides the bending deformation we also apply a 3D rigid transform $\mathcal{R}$. Here, we do not use a tapering deformation. Applying the transformations, we obtain the parametric intensity model with a total of 13 parameters:

$$\mathbf{p}_{Saddle} = (r_x, r_y, r_z, a_0, a_1, \sigma, \delta, \alpha, \beta, \gamma, x_0, y_0, z_0) \quad (5)$$

3D Intensity Model of Sphere-Like Structures Sphere-like structures are, for example, human eyes. These structures can be modelled by a sphere with radius R. Fortunately, the exact model of a Gaussian smoothed sphere can be expressed in analytic form (see [8]) and is given by

$$g_{Sphere}(\mathbf{x}) = \Phi_\sigma(R - r) - \Phi_\sigma(-R - r) - \sigma^2 r^{-1}\left(G_\sigma(R - r) - G_\sigma(R + r)\right) \quad (6)$$

where $r = \sqrt{x^2 + y^2 + z^2}$, $\Phi_\sigma(x) = \Phi(x/\sigma)$, and $G_\sigma(x) = \left(\sqrt{2\pi}\sigma\right)^{-1} e^{-\frac{x^2}{2\sigma^2}}$. We define the center of the sphere as the position of the landmark. In addition, we include the intensity levels a_0 and a_1 as well as a 3D translation. This results in the parametric intensity model with a total of 7 parameters:

$$\mathbf{p}_{Sphere} = (R, a_0, a_1, \sigma, x_0, y_0, z_0) \quad (7)$$

Table 1. Size and resolution of the medical 3D images used in the experiments.

Image	Slices	Size (Voxels)	Resolution (mm^3)
Woho (MR)	sagittal	$256 \times 256 \times 256$	$1.0 \times 1.0 \times 1.0$
C06 (MR)	axial	$256 \times 256 \times 120$	$0.859 \times 0.859 \times 1.2$
C06 (CT)	axial	$320 \times 320 \times 87$	$0.625 \times 0.625 \times 1.0$

3 Model Fitting Approach

Estimates of the model parameters in (4), (5), and (7) are found by a least-squares fit of the model to the image intensities $g(\boldsymbol{x})$ within semi-global regions-of-interest (ROIs), thus minimizing the objective function

$$\sum_{\boldsymbol{x} \in \mathrm{ROI}} \left(g_M(\boldsymbol{x}, \boldsymbol{p}) - g(\boldsymbol{x}) \right)^2 \tag{8}$$

Note, the fitting measure does not include any derivatives. This is in contrast to previous fitting measures for surface models which incorporate the image gradient as well as 1st order derivatives of the model (e.g., [4]).

For the minimization we apply the method of Levenberg-Marquardt, incorporating 1st order partial derivatives of the intensity model w.r.t. the model parameters. In [9] such type of approach has been used for localizing 2D corner and edge features. The partial derivatives can be derived analytically using the generalized chain rule. Note, we do not need to compute the image gradient as is the case with surface models. We need 1st order derivatives of the intensity model only for the minimization process, whereas the surface model approach requires 2nd order derivatives for the minimization.

4 Experimental Results: 3D Synthetic Data

We applied our approach to synthetic 3D image data generated by the three intensity models itself with added Gaussian noise. For the ellipsoidal model, we carried out about 3000 experiments with different parameter settings and achieved a very high localization accuracy with a maximum error in the estimated position of less than 0.15 voxels (except in one case we got 0.35 voxels). We also found that the approach is robust w.r.t. the choice of initial parameters. Additionally, for about 2000 experiments with similar settings but very intense Gaussian noise down to a signal-to-noise ratio of ca. 1, the maximum localization error turned out to be less than 0.53 voxels. For the saddle model, we carried out about 5000 similar experiments. The resulting localization errors turned out to be about twice as large as for the ellipsoidal model. For the spherical model, the results of 5000 similar experiments are much better than the results of the other two models, i.e. the localization error is less than 0.17 voxels even with very intense Gaussian noise down to a signal-to-noise ratio of ca. 0.1.

5 Experimental Results: 3D Medical Images

We also applied the new approach to three real 3D tomographic images of the human head (datasets Woho and C06, see Table 1).

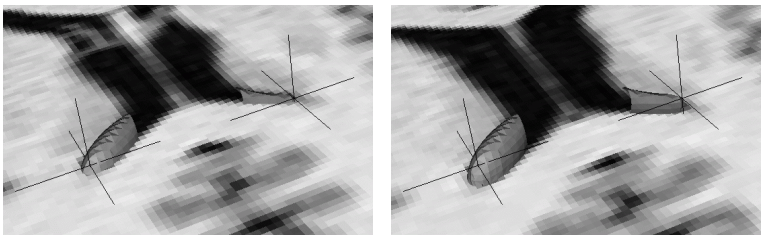

Fig. 2. 3D contour plots of the fitted intensity models for the left and right frontal horn within an MR image (Woho). The result is shown for two different slices of the original data. The marked axes indicate the estimated landmark positions.

We considered seven tip-like landmarks, i.e. the frontal, occipital, and temporal horns (left and right) as well as the external occipital protuberance, and two saddle-like landmarks, i.e. the zygomatic bone (left and right). For these landmarks in all three images we used as ground truth, positions that were manually determined in agreement with up to four persons. For the CT image, we did not consider the temporal horns since either the ground truth position was missing due to very low signal-to-noise ratio (left horn) or it was not possible to successfully fit the intensity model (right horn). Particularly with this landmark the image quality was very bad. In addition, we considered two sphere-like landmarks in the MR images, i.e. the left and right eye.

Parameter Settings The fitting procedure described above requires the determination of suitable initial parameter values. The specification of these values is not a trivial task. Often all values are initialized manually, which is time-consuming. In case of the ellipsoidal model, we automatically initialize half of the model parameters. Values for the most important parameters, namely, the translation parameters (x_0, y_0, z_0) defining the position of the landmark were obtained by a 3D differential operator. Here we used the operator $Op3 = det\boldsymbol{C}_g / trace\boldsymbol{C}_g$, where $\boldsymbol{C}_g$ is the averaged dyadic product of the image gradient ([10]). This initialization was successfull for model fitting except in two cases, where the positions of Op3 were relatively far away from the ground truth positions (see Table 4). In these two cases, we initialized the translation parameters manually. The smoothing parameter σ was always initialized with 1.0 and the deformation parameters ρ_x, ρ_y, δ, and ν were all initialized with 0.0, thus, the intensity model was always initialized as an ellipsoid without deformation. The remaining parameters and the size of the ROI were initialized manually. For the saddle and spherical model, all parameters were so far initialized manually.

Results Tables 2, 3, and 4 show the fitting results for the considered landmarks. In case of the ellipsoidal model, model fitting needed 75 iterations on average. We have visualized the fitting results of the left and right frontal horn within an MR image in Figure 2 and of the external occipital protuberance for the C06 image pair in Figure 3 using 3D Slicer ([5]). The average distance between the estimated landmark positions and the ground truth positions for all 19 tip-like landmarks computes to $\bar{e} = 1.14mm$. In comparison, using the 3D differential

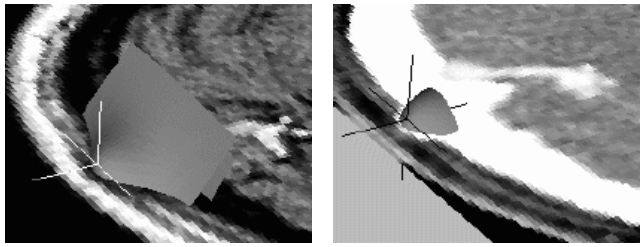

Fig. 3. 3D contour plots of the fitted intensity model for the external occipital protuberance within the original image pair C06 (left MR and right CT). Note, the size of the ROI and the used deformations are different.

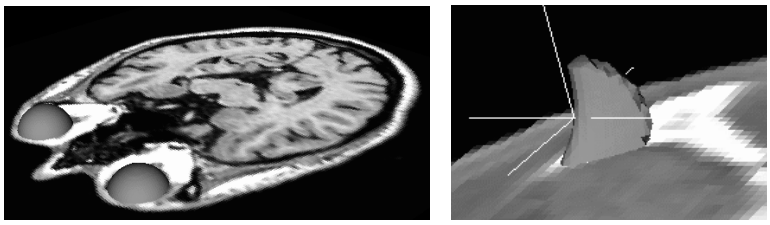

Fig. 4. 3D contour plots of the fitted intensity model for the eyes within an MR image (Woho) and for the left zygomatic bone within an CT image.

operator Op3, we obtain an average distance of $\bar{e}_{Op3} = 2.18mm$. Thus, the localization accuracy with our new approach turns out to be much better.

The results for the saddle model are worse in comparison to the ellipsoidal model. The average distance between the estimated landmark positions and the ground truth positions for all 6 saddle-like landmarks computes to $\bar{e} = 1.84mm$. In comparison, using the 3D differential operator Op3, we obtain an average distance of $\bar{e}_{Op3} = 1.63mm$. In addition, it turned out that the saddle model depends more on the initial parameter values than the other models. The results for the spherical model are very good. The fitted model describes the image intensities fairly well and also the model fitting is very robust w.r.t. the initial parameters. Figure 4 shows the fitting result for both eyes within an MR image.

The execution time of our algorithm is mainly dependent on the size of the ROI, the chosen variant of the deformation, and the quality of the initial parameters. As a typical example, the fitting time for the right temporal horn in the Woho image including tapering and bending deformations and a diameter of the ROI of 19 voxels is ca. 1s (on a AMD Athlon, 1.7GHz, running Linux).

6 Discussion

The experiments verify the applicability of our new approach, which yields subvoxel positions of 3D anatomical landmarks. The intensity models describe the anatomical structures fairly well as can be seen from the 3D contour plots. Also, the figures demonstrate that the spectrum of possible shapes of our intensity

Table 2. Fitting results for the ventricular horns and the external occipital protuberance (ellipsoidal model), for the zygomatic bone (saddle model), and for the eyes (spherical model) for the C06 image (MR). The estimated landmark position, intensity levels, and the distance e to the ground truth position are given. For comparison, the distance e_{Op3} of the differential operator $Op3$ to the ground truth position is listed.

C06 (MR)	$\hat{x}_0$	$\hat{y}_0$	$\hat{z}_0$	$\hat{a}_0$	$\hat{a}_1$	e	e_{Op3}
Left frontal horn	150.65	79.58	68.14	91.6	22.3	1.27mm	1.92mm
Right frontal horn	112.34	76.85	69.02	93.9	18.8	0.58mm	1.72mm
Left occipital horn	143.91	200.85	53.01	84.9	15.2	0.15mm	3.32mm
Right occipital horn	107.82	195.98	56.04	86.6	20.0	0.70mm	1.72mm
Left temporal horn	164.01	117.26	45.38	82.4	12.8	1.20mm	1.71mm
Right temporal horn	98.98	112.23	40.63	80.0	18.8	0.97mm	2.10mm
Ext. occipital protub.	130.05	230.94	32.97	61.6	8.7	0.06mm	1.21mm
					Mean	0.70mm	1.96mm
Left zygomatic bone	192.29	62.81	34.46	121.7	20.8	1.42mm	1.21mm
Right zygomatic bone	70.94	60.95	31.66	128.2	14.9	0.99mm	1.48mm
Left eye	167.16	41.69	44.42	90.4	24.7		
Right eye	98.00	40.68	43.50	97.5	25.4		

Table 3. Same as Table 2 but for the C06 image (CT).

C06 (CT)	$\hat{x}_0$	$\hat{y}_0$	$\hat{z}_0$	$\hat{a}_0$	$\hat{a}_1$	e	e_{Op3}
Left frontal horn	192.80	93.94	77.04	1043.5	996.8	1.33mm	0.63mm
Right frontal horn	135.31	90.46	78.14	1036.7	1001.8	1.26mm	2.10mm
Left occipital horn	184.07	260.57	69.21	1038.5	989.7	0.66mm	0.00mm
Right occipital horn	129.50	255.77	72.88	1045.0	994.0	0.94mm	1.33mm
Ext. occipital protub.	161.20	309.43	48.01	1007.9	2679.0	1.10mm	1.72mm
					Mean	1.06mm	1.16mm
Left zygomatic bone	245.91	77.37	49.73	976.1	2829.0	1.49mm	0.63mm
Right zygomatic bone	79.17	75.24	50.80	977.1	2918.0	1.81mm	1.00mm

models is relatively large. An issue for further work is the automatic initialization of all model parameters, e.g., based on differential properties of the image.

7 Acknowledgement

The original MR and CT images have kindly been provided by Philips Research Hamburg and W.P.Th.M. Mali, L. Ramos, and C.W.M. van Veelen (Utrecht University Hospital) via ICS-AD of Philips Medical Systems Best.

References

1. M. Alker, S. Frantz, K. Rohr, and H.S. Stiehl, "Improving the Robustness in Extracting 3D Point Landmarks from 3D Medical Images Using Parametric Deformable Models", *Proc. MICCAI'2001*, Utrecht, The Netherlands, Oct. 14-17, 2001, *Lecture Notes in Computer Science* 2208, W.J. Niessen and M.A. Viergever (Eds.), Springer-Verlag Berlin Heidelberg 2001, 582-590

Table 4. Same as Table 2 but for the Woho image (MR).

Woho	$\hat{x}_0$	$\hat{y}_0$	$\hat{z}_0$	$\hat{a}_0$	$\hat{a}_1$	e	e_{Op3}
Left frontal horn	111.26	78.26	101.84	124.0	23.8	2.22mm	3.16mm
Right frontal horn	111.49	77.54	132.27	117.3	20.1	1.44mm	2.24mm
Left occipital horn	189.38	101.53	91.62	107.3	23.3	2.31mm	4.12mm
Right occipital horn	182.63	97.42	150.02	112.7	15.9	0.68mm	3.61mm
Left temporal horn	134.90	111.86	88.81	95.1	44.3	1.80mm	2.83mm
Right temporal horn	129.24	114.36	150.16	109.6	35.8	1.46mm	4.58mm
Ext. occipital protub.	232.14	149.73	120.96	84.2	26.8	1.48mm	1.41mm
					Mean	1.63mm	3.14mm
Left zygomatic bone	96.46	125.64	58.69	167.2	47.1	2.22mm	3.00mm
Right zygomatic bone	89.59	128.50	179.83	197.8	12.1	3.12mm	2.45mm
Left eye	73.49	107.70	82.47	114.8	29.9		
Right eye	70.16	105.95	151.44	122.2	30.4		

2. R. Bertolini and G. Leutert, *Atlas der Anatomie des Menschen. Band 3:Kopf, Hals, Gehirn, Rückenmark und Sinnesorgane*, Springer-Verlag, Berlin, 1982
3. L. Le Briquer, F. Lachmann, and C. Barillot, "Using Local Extremum Curvatures to Extract Anatomical Landmarks from Medical Images", *Medical Imaging 1993: Image Processing,* 16-19 Febr. 1993, Newport Beach, California/USA, Proc. SPIE 1898, M.H. Loew (Ed.), 549-558
4. S. Frantz, K. Rohr, and H.S. Stiehl, "Localization Of 3D Anatomical Point Landmarks In 3D Tomographic Images Using Deformable Models", *Proc. MICCAI'2000*, Pittsburgh, Pennsylvania/USA, Oct. 11-14, 2000, *Lecture Notes in Computer Science* 1935, S.L. Delp, A.M. DiGioia, and B. Jaramaz (Eds.), Springer-Verlag Berlin Heidelberg, 2000, 492-501
5. D.T. Gering, A. Nabavi, R. Kikinis, W.E.L. Grimson, N. Hata, P.Everett, F. Jolesz, and W.M. Wells, "An Integrated Visualization System for Surgical Planning and Guidance using Image Fusion and Interventional Imaging", *Proc. MICCAI'99*, Cambridge England, Sep. 19-22, 1999, *Lecture Notes in Computer Science* 1679, C. Taylor and A. Colchester (Eds.), Springer-Verlag Berlin Heidelberg, 1999, 808-819
6. T. Hartkens, K.Rohr, and H.S. Stiehl, "Evaluierung von Differentialoperatoren zur Detektion charakteristischer Punkte in tomographischen Bildern", *Proc. 18. DAGM-Symposium Mustererkennung (DAGM'96)*, 11.-13. Sept. 1996, Heidelberg/Germany, *Informatik aktuell*, B. Jähne, P. Geißler, H. Haußecker, and F.Hering (Eds.), Springer-Verlag Berlin Heidelberg, 1996, 637-644
7. T. Hartkens, K.Rohr, and H.S. Stiehl, "Evaluation of 3D Operators for the Detection of Anatomical Point Landmarks in MR and CT Images", *Computer Vision and Image Understanding* 85, 2002, 1-19
8. R.M. Kessler, J.R. Ellis, Jr., and M. Eden, "Analysis of Emission Tomographic Scan Data: Limitations Imposed by Resolution and Background", *Journal of Computer Assisted Tomography* 8:3, 1984, 514-522
9. K. Rohr, "Recognizing Corners by Fitting Parametric Models", *International J. of Computer Vision* 9:3, 1992, 213-230
10. K. Rohr, "On 3D differential operators for detecting point landmarks", *Image and Vision Computing* 15:3, 1997, 219-233
11. J. Sobotta, *Atlas der Anatomie des Menschen. Band 1: Kopf, Hals, obere Extremität, Haut*, Urban & Schwarzenberg, München, 19th edition, 1988
12. J.-P. Thirion, "New Feature Points based on Geometric Invariants for 3D Image Registration", *Int. J. of Computer Vision* 18:2, 1996, 121-137