Improving the Robustness in Extracting 3D Point Landmarks Based on Deformable Models

Manfred Alker[1], Sönke Frantz[1], Karl Rohr[2], and H. Siegfried Stiehl[1]

[1] Universität Hamburg, FB Informatik, AB Kognitive Systeme, Vogt-Kölln-Str. 30, D-22527 Hamburg, {alker,frantz,stiehl}@kogs.informatik.uni-hamburg.de
[2] International University in Germany, D-76646 Bruchsal, rohr@i-u.de
http://kogs-www.informatik.uni-hamburg.de/PROJECTS/imagine/Imagine.html

Abstract. Existing approaches to extracting 3D point landmarks based on deformable models require a good model initialization to avoid local suboptima during model fitting. To overcome this drawback, we propose a generally applicable *novel hybrid optimization algorithm* combining the advantages of both conjugate gradient (cg-)optimization (known for its time efficiency) and genetic algorithms (exhibiting robustness against local suboptima). We apply our algorithm to 3D MR and CT images depicting tip-like and saddle-like anatomical structures such as the horns of the lateral ventricles in the human brain or the zygomatic bones as part of the skull. Experimental results demonstrate that the robustness of model fitting is significantly improved using hybrid optimization compared to a purely local cg-method. Moreover, we compare an *edge strength-* to an *edge distance-based fitting measure.*

Keywords: 3D landmark extraction, deformable models, robustness, hybrid optimization

1 Introduction

Extracting 3D point landmarks from 3D tomographic images is a prerequisite for landmark-based approaches to 3D image registration, which is a fundamental problem in computer-assisted neurosurgery. While earlier approaches exploit the local characteristics of the image data using differential operators (e.g., [16],[11]), in [5] an approach based on parametric deformable models has recently been proposed that takes into account more global image information and allows to localize 3D point landmarks more accurately. However, since local optimization is employed for model fitting, a good model initialization is required to avoid local suboptima. To overcome this, we propose a generally applicable new *hybrid optimization algorithm* combining the computational efficiency of the (local) conjugate gradient (cg-)optimization method with the robustness of (global) genetic algorithms. Existing optimization algorithms for fitting parametric deformable models to image data are either purely local (e.g., [13],[18],[1],[5]) or strictly global (e.g., [4],[17]). Moreover, we compare an *edge strength-* (e.g., [18]) with an *edge distance-based fitting measure* (cf., e.g., [1]). We apply our fitting algorithm in order to extract salient surface loci (curvature extrema) of *tip-* and *saddle-like structures* such as the tips of the ventricular horns or the saddle points at the zygomatic bones (see Fig. 1(a),(b)). For representing such structures, we use globally deformed quadric surfaces (Sect. 2). The fitting measures for model

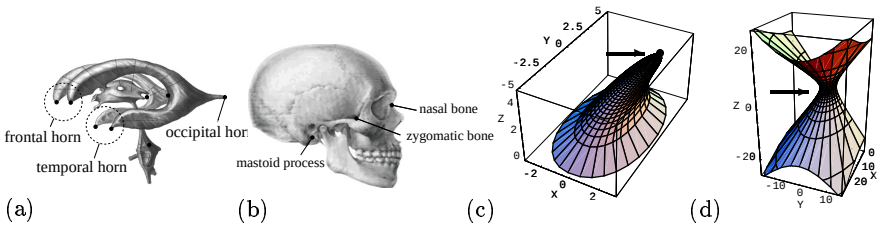


Fig. 1. (a),(b): Ventricular horns of the human brain (from [12]) and the human skull (from [2]). Examples of 3D point landmarks are indicated by dots. **(c),(d)**: Quadric surfaces as geometric models for tips ((c): tapered and bended half-ellipsoid) and for saddle structures ((d): hyperboloid of one sheet). The landmark positions are indicated by a dot.

fitting are then described in Sect. 3, while our hybrid algorithm for optimizing a fitting measure w.r.t. the model parameters is outlined in Sect. 4. Experimental results of studying the robustness of model fitting are presented in Sect. 5. In particular, we analyze the landmark localization accuracy of our new approach and compare it with that of purely local cg-optimization.

2 Modeling Tip- and Saddle-Like Structures with Quadrics

In the literature, a variety of 3D surface models has been used for different applications, e.g., segmentation, registration, and tracking (e.g., [13],[15],[4],[18], [1]; see [9] for a survey). To extract 3D point landmarks, we here use quadric surfaces as geometric models for tip- and saddle-like structures ([5]) since they well represent the anatomical structures of interest here, but only have few model parameters. It is advantageous here that they may be represented by both a parametric and an implicit defining function. Tapered and bended ellipsoids are utilized for representing 3D tip-like structures such as the ventricular horns, whereas hyperboloids of one sheet are used for 3D saddle-like structures such as the zygomatic bones (see Fig. 1(c),(d)). For tip-like structures, the *parametric form* of our model is obtained by applying linear tapering [4] and quadratic bending [4] as well as a rigid transformation, resp., to the parametric form of an ellipsoid:

$$\boldsymbol{x}_{tip}(\theta,\phi) = \boldsymbol{R}_{\alpha,\beta,\gamma} \begin{pmatrix} a_1 \cos\theta\cos\phi/(\rho_x \sin\theta{+}1){+}\delta\cos\upsilon(a_3\sin\theta)^2 \\ a_2 \cos\theta\sin\phi/(\rho_y \sin\theta{+}1){+}\delta\sin\upsilon(a_3\sin\theta)^2 \\ a_3 \sin\theta \end{pmatrix} + \boldsymbol{t}, \qquad (1)$$

where $0 \le \theta \le \pi/2$ and $-\pi \le \phi < \pi$ are the latitude and longitude angle parameters, resp. Further on, $a_1, a_2, a_3 > 0$ are scaling parameters, $\rho_x, \rho_y \ge 0$ denote the tapering strengths in x- and y-direction, and δ, υ determine the bending strength and direction, resp. For the rigid transformation, α, β, γ denote the Eulerian angles of rotation and $\boldsymbol{t}^T = (X, Y, Z)$ is the translation vector of the origin. Hence, the model is described by the parameter vector $\boldsymbol{p} = (a_1, a_2, a_3, \delta, \upsilon, \rho_x, \rho_y, X, Y, Z, \alpha, \beta, \gamma)$. The landmark position is then given by $\boldsymbol{x}_l = \boldsymbol{x}_{tip}(\pi/2, 0) = \boldsymbol{R}_{\alpha,\beta,\gamma}(\delta\cos\upsilon\, a_3^2, \delta\sin\upsilon\, a_3^2, a_3)^T + \boldsymbol{t}$. The *parametric form* of hyperboloids of one sheet is the same as the one given in [5].

3 Model Fitting with Edge-Based Fitting Measures

In order to fit the geometric models from Sect. 2 to the image data, a fitting measure is optimized w.r.t. the model parameters. Here, we consider an *edge strength-* and an *edge distance-based fitting measure*. For the *edge strength-based measure* M_{ES} (e.g., [18]), [14]), the edge strength e_g is integrated over the model surface M:

$$M_{ES} = -\int_M e_g(\boldsymbol{x})dF = -\int\int_{\theta,\phi} e_g(\boldsymbol{x}(\theta,\phi)) \left\| \frac{\partial \boldsymbol{x}}{\partial \theta} \times \frac{\partial \boldsymbol{x}}{\partial \phi} \right\| d\theta d\phi \rightarrow \text{Min!}, \quad (2)$$

where $e_g(\boldsymbol{x}) = \|\boldsymbol{\nabla} g(\boldsymbol{x})\|$ is the gradient magnitude of the intensity function g and $\boldsymbol{x}$ is a vector at the model surface M which is parameterized by θ, ϕ. To emphasize small surfaces, we additionally apply a *surface weighting factor* to the fitting measure (2) which then takes the form $M_{ES} = -\frac{\int_M e_g(\boldsymbol{x})dF}{\sqrt{\int_M dF}}$.

The *edge distance-based fitting measure* M_{ED} used here is written as (cf., e.g., [13],[1],[17])

$$M_{ED}(\boldsymbol{\Xi}, \boldsymbol{p}) = \sum_{i=1}^{N} e_g(\boldsymbol{\xi}_i)\ \rho\left(\frac{1-\hat{\mathrm{F}}(\boldsymbol{\xi}_i, \boldsymbol{p})}{\|\boldsymbol{\nabla}\hat{\mathrm{F}}(\boldsymbol{\xi}_i, \boldsymbol{p})\|}\right) \rightarrow \text{Min!} \quad (3)$$

The sum is taken over all N image voxels $\boldsymbol{\Xi} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N\}$ which – in order to eliminate the influence of neighbouring structures – lie within a *region-of-interest (ROI)*. and whose edge strength $e_g(\boldsymbol{\xi}_i)$ exceeds a certain threshold value. The vector of model parameters is denoted by $\boldsymbol{p}$. Further on, we use $\rho(x) = |x|^{1.2}$ for all $x \in \mathbb{R}$ as a distance weighting function to reduce the effect of outliers ([19]). The argument of ρ is a first order distance approximation between the image voxel with coordinates $\boldsymbol{\xi}_i$ and the model surface ([17]), where $\hat{\mathrm{F}}$ denotes the *inside-outside function* of the tapered and bended quadric surface after applying a rigid transform (cf., e.g., [13],[4],[1]). The *inside-outside function* of both, undeformed superellipsoids (F_e; upper plus sign in eqn. (4)) and undeformed superhyperboloids of one sheet (F_h; lower minus sign in eqn. (4)), can be written as

$$\mathrm{F}_{e/h}(\boldsymbol{\xi}) := \mathrm{f}_{\epsilon_1/2}\left(\left(|\xi_1/a_1|^{2/\epsilon_2} + |\xi_2/a_2|^{2/\epsilon_2}\right)^{\epsilon_2/\epsilon_1} \pm |\xi_3/a_3|^{2/\epsilon_1}\right), \quad (4)$$

where $\mathrm{f}_\epsilon(x) := \mathrm{sign}(x)|x|^\epsilon$ for $x \in \mathbb{R}$ is the signed power function for $\epsilon \geq 0$ and $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)^t$. Since superellipsoids have a closed surface, there is a simple interpretation of the *inside-outside function* that explains its name:

$$\begin{cases} \text{If} & \mathrm{F}(\boldsymbol{\xi}) = 1, & \boldsymbol{\xi} \text{ is on the surface,} \\ \text{if} & \mathrm{F}(\boldsymbol{\xi}) > 1, & \boldsymbol{\xi} \text{ is outside, and} \\ \text{if} & \mathrm{F}(\boldsymbol{\xi}) < 1, & \boldsymbol{\xi} \text{ is inside.} \end{cases} \quad (5)$$

Also, a volume factor is used in conjunction with (3) to emphasize small volumes. This factor has been chosen as $1 + \frac{a_1 a_2 a_3}{a_{1,est} a_{2,est} a_{3,est}}$, where the weighting factor $a_{1,est} a_{2,est} a_{3,est}$ is coarsely estimated to a value of $10^3\, vox^3$ (vox denotes the spatial unit of an image voxel). For volume weighting factors (or size factors), see also, e.g., [13].

4 A Novel Hybrid Optimization Algorithm

Most optimization algorithms considered in the literature for model fitting are *local* algorithms such as the *conjugate gradient (cg-)method* (e.g., [13],[18],[1],[5]) and thus are prone to run into local suboptima. The cg-method combines problem specific search directions of the *method of steepest descent* with optimality properties of the method of *conjugate directions* (e.g., [6]). However, since it is a purely local method, it is prone to run into local suboptima. On the other hand, global optimization methods such as genetic algorithms (GAs); e.g., [7]) have been proposed to avoid local suboptima (e.g., [4],[17]), but are plagued with slow convergence rates. We here propose a *hybrid algorithm* which combines the advantages of both methods. Similar to GAs, we consider a whole *population* of parameter vectors, but we differ in the *mutation strategy* since we do not use *bit-flips* and *crossovers* to imitate natural mutation ([7]). By contrast, we use several most promising local optima resulting from a *line search* after each cg-step as candidate solutions. In order to obtain a generally applicable strategy, we adapt the population size to the complexity of the problem by increasing the maximal population size each time a candidate solution converges to a local optimum, i.e. when its objective function value does not improve for a given number of cg-iterations. Consequently, several parameters can be adapted to the specific optimization problem at hand:

- the maximum population size that must not be exceeded (here: 20),
- the number of cg-iterations after which the least successful population members (measured by their value of the fitting measure) are discarded (here: 5),
- the minimum number of population members that are retained after each such 'survival of the fittest' step (here: 5), and
- the number of cg-iterations with no significant improvement of the value of the fitting measure after which a population member is marked convergent and is not subject to further cg-iterations (here: 80).
- a difference threshold for two parameter vectors of the deformable model below which they are considered as being equal.

The mentioned parameters have been used in all our experiments. Except for the need of adjusting these parameters, the optimization strategy presented here is a general-purpose method for poorly initialized nonlinear optimization problems. and its applicability is not confined to model fitting problems in medical image analysis. Only one example of *hybrid optimization* in image analysis is known to us: In [8], discontinuity preserving visual reconstruction problems, e.g. sparse data surface reconstruction and image restoration problems, are described as coupled (binary-real) optimization problems. An informed GA is applied to the binary variables (describing the discontinuities), while a cg-method is applied to the real variables for a given configuration of the binary variables visited by the GA. By contrast, in our approach the local and the global part are treated uniformly.

5 Experimental Results for 3D Tomographic Images

Scope of experiments In our experiments, the deformable models were fitted to tip-like and saddle-like anatomical structures and our hybrid optimization algorithm has been compared to purely local cg-optimization w.r.t. poorly initialized model parameters using

- different *types of image data*: two 3D T1-weighted MR images and one 3D CT image of the human head,
- different *types of landmarks*: frontal/occipital horn of the left/right lateral ventricle, left/right zygomatic bone as part of the skull,
- different *fitting measures*: edge distance-based and edge strength-based, and
- different sizes of the *region of interest (ROI)*:
  ROI radius of 10 *vox* and 15 *vox* (*vox*: spatial unit of an image voxel).

Experimental strategy For each 3D MR and 3D CT image, an initial good fit is determined by repeated model fittings and visual inspection of the fitting results. For obtaining poor initial estimates for model fitting, the parameter values of the initial good fit are varied by adding Gaussian distributed random numbers with zero mean and large variances. In order to determine the landmark localization error e, the landmark positions calculated from the fitted deformable models are compared to ground truth positions that were manually specified in the 3D images in agreement with up to four persons. In addition, we consider the root-mean-squared distance between the edge points of the image and the model surface, e_{RMS}, using a Euclidean distance map ([10]) from the image data after applying a 3D edge detection algorithm based on [3]. This procedure is iterated sufficiently often (here: 100 times) with different, randomized model initializations. The mean values and RMS estimates of the resulting values of e and e_{RMS} are tabulated then. For evaluating the fitting measures (2),(3), the derivatives of the intensity function are computed numerically using cubic B-spline interpolation and Gaussian smoothing (see [5]).

General results Common to all experiments is that the final value of the fitting measure is better by about 10-50% for hybrid optimization than for purely local cg-optimization. In most cases, the landmark localization and the model fitting accuracy also improve significantly. Thus, hybrid optimization turns out to be superior to purely local cg-optimization at the expense of an increase in computational costs of a factor of 5-10 (30s–90s for local cg-optimization and 150s–900s for hybrid optimization on a SUN SPARC Ultra 2 with 300MHz CPU).

The edge distance-based fitting measure in (3) turns out to be more suitable for 3D MR images of the ventricular horns with high signal-to-noise ratio since it incorporates *distance* approximations between the image data and the model surface (*long* range forces, cf. [15]). However, in comparison to (2), it is relatively sensitive to noise. Moreover, it is not suitable for hyperboloids of one sheet due to inaccuracies of the first order distance approximation associated with it.

Results for the ventricular horns The tips of the frontal and occipital horns of the lateral ventricles are considered here. Typical examples of successful model

fitting, which demonstrate the robustness of model fitting, are given in Fig. 2. Here, contours of the model initialization are drawn in black and the results of model fitting using purely local cg-optimization are drawn in grey, while the results of model fitting using our new hybrid optimization algorithm are drawn in white. The ground truth landmark positions are indicated by a $\oplus$-symbol. As can be seen from the averaged quantitative results in Table 1, hybrid optimization is superior to purely local cg-optimization and yields not only better values of the fitting measures, but in most cases also better model fitting (e_{RMS}) and landmark localization (e) results (cf. also Figs. 2(a),(b)). Note that rather coarsely initialized model parameters have been used ($\overline{e}_{initial} \approx 7 \ldots 9\, vox$), and thus some unsuccessful fitting results – particularly in the case of the less pronounced occipital horns – deteriorate the average accuracy of model fitting as shown in Table 1. The average accuracy of landmark localization for the frontal horn of the right lateral ventricle in Table 1 is less than that for the left ventricle in the same 3D MR image due to its comparably thin anatomical structure.

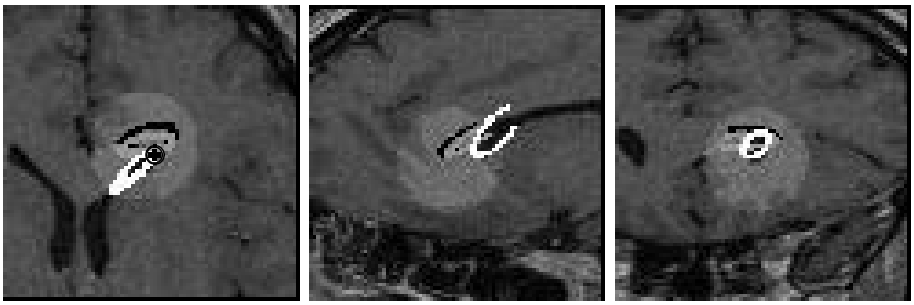
		Model initialization	Edge dist.-b. fitt. meas.		Edge strength-b. fitt. meas.	
			local cg-opt.	hybrid opt.	local cg-opt.	hybrid opt.
Frontal horn (left)	$\overline{e}$	7.71 ± 3.16	3.28 ± 2.99	1.40 ± 1.18	3.54 ± 2.18	2.49 ± 2.21
	$\overline{e}_{RMS}$	2.22 ± 1.10	1.00 ± 0.63	0.65 ± 0.22	1.04 ± 0.31	0.87 ± 0.35
Frontal horn (right)	$\overline{e}$	6.57 ± 3.18	3.87 ± 2.16	3.15 ± 2.18	6.55 ± 3.53	5.19 ± 3.70
	$\overline{e}_{RMS}$	2.12 ± 1.11	1.05 ± 0.60	0.78 ± 0.25	1.56 ± 1.26	1.28 ± 0.79
Occipital horn (right)	$\overline{e}$	9.08 ± 4.42	6.90 ± 3.89	6.68 ± 3.93	4.74 ± 4.33	4.61 ± 4.31
	$\overline{e}_{RMS}$	3.00 ± 1.40	2.06 ± 0.93	2.04 ± 0.87	1.34 ± 0.87	1.29 ± 0.78

Table 1. Fitting results averaged over 100 model fittings with randomized poor model initializations for 3D MR images of the *frontal/occipital* ventricular horns using *13 model parameters* ($\overline{e}$: mean landmark localization error (in *vox*), $\overline{e}_{RMS}$: RMS distance between deformable model and image data (in *vox*), voxel size $= 0.86 \times 0.86 \times 1.2\text{mm}^3$).

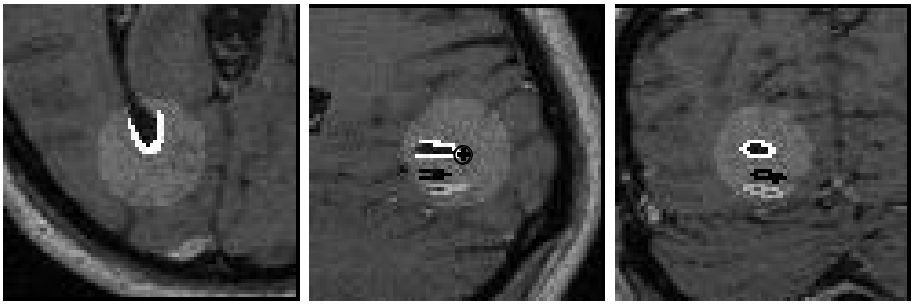
Results for the zygomatic bones All results for the zygomatic bones were obtained with the edge strength-based fitting measure (2). Model fitting for the saddle points at the zygomatic bones (e.g., Fig. 2(c)) in general is not as successful as it is for the tips of the ventricular horns since our geometric primitive does not describe the anatomical structure at hand with comparable accuracy. However, the mean landmark localization error $\overline{e}$ can be reduced from initially $\overline{e}_{initial} = 6.4 \ldots 6.9\, vox$ to $\overline{e} = 2.5 \ldots 3.2\, vox$ and the accuracy of model fitting is $\overline{e}_{RMS} = 1.5 \ldots 1.8\, vox$ (voxel size $= 1.0\text{mm}^3$).

6 Conclusion

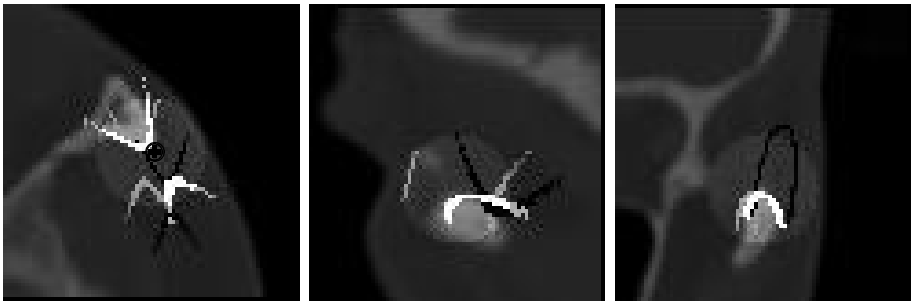
In this paper, landmark extraction based on deformable models has been investigated in order to improve the stability of model fitting as well as of landmark localization in the case of poorly initialized model parameters. To this end, a generally applicable novel hybrid optimization algorithm has been introduced and edge strength- and edge distance-based fitting measures have been compared. Experimental results have demonstrated the applicability of our hybrid



(a) 3D MR image of the *frontal* horn of the *left* lateral ventricle, edge *distance*-based fitting measure, ROI size 15.0 *vox*



(b) 3D MR image of the *occipital* horn of the *right* lateral ventricle, edge *strength*-based fitting measure, ROI size 15.0 *vox*



(c) 3D CT image of the *left* zygomatic bone, edge *strength*-based fitting measure, ROI size 15.0 *vox*

Fig. 2. Examples of successfully fitting tapered and bended half-ellipsoids to 3D MR images of the frontal and occipital horns of the lateral ventricles (Fig. 2(a-b)) as well as of fitting a half-hyperboloid with no further deformations to a 3D CT image of the zygomatic bone (Fig. 2(c)). Contours of the model surfaces in axial, sagittal, and coronal planes are depicted here (from left to right). *Black*: model initialization, *grey*: fitting result for local cg-optimization, and *white*: fitting result for our hybrid optimization algorithm. The ground truth landmark positions are indicated by a $\oplus$-sign here.

algorithm as well as its increased robustness as compared to a purely local cg-method. However, the experimental results do not clearly favour one fitting measure. However, for the frontal ventricular horns, our edge distance-based fitting measure is more successful, while for the less pronounced occipital horns and for the zygomatic bones, the edge strength-based fitting measure is more suitable.

References

1. E. Bardinet, L.D. Cohen, and N. Ayache. Superquadrics and Free-Form Deformations: A Global Model to Fit and Track 3D Medical Data. *Proc. CVRMed'95*, LNCS 905, pp. 319–326. Springer, 1995.
2. R. Bertolini and G. Leutert. *Atlas der Anatomie des Menschen. Band 3: Kopf, Hals, Gehirn, Rückenmark und Sinnesorgane.* Springer, 1982.
3. J.F. Canny. A Computational Approach to Edge Detection. *PAMI*, 8(6):679–698, 1986.
4. K. Delibasis and P.E. Undrill. Anatomical object recognition using deformable geometric models. *Image and Vision Computing*, 12(7):423–433, 1994.
5. S. Frantz, K. Rohr, and H.S. Stiehl. Localization of 3D Anatomical Point Landmarks in 3D Tomographic Images Using Deformable Models. *Proc. Int. Conf. on Medical Image Computing and Computer-Assisted Intervention (MICCAI2000)*, LNCS 1935, pp. 492–501, Springer, 2000.
6. G.H. Golub and C.F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, 1996.
7. D. Goldberg. Genetic Algorithms in Search, Optimization and Machine Learning. Addison Wesley, 1989.
8. S.H. Lai and B.C. Vemuri. Efficient hybrid-search for visual reconstruction problems. *Image and Vision Computing*, 17(1):37–49, 1999.
9. T. McInerney and D. Terzopoulos. Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis*, 1(2):91–108, 1996.
10. D.W. Paglieroni. A Unified Distance Transform Algorithm and Architecture. *Machine Vision and Applications*, 5(1):47–55, 1992.
11. K. Rohr. On 3D differential operators for detecting point landmarks. *Image and Vision Computing*, 15(3):219–233, 1997.
12. J. Sobotta. *Atlas der Anatomie des Menschen. Band 1: Kopf, Hals, obere Extremität, Haut.* Urban & Schwarzenberg, 19th edition, 1988.
13. F. Solina and R. Bajcsy. Recovery of Parametric Models from Range Images: the Case for Superquadrics with Global Deformations. *PAMI*, 12(2):131–147, 1990.
14. L.H. Staib and J.S. Duncan. Model-Based Deformable Surface Finding for Medical Images. *IEEE Trans. Med. Imag.*, 15(5):720–731, 1996.
15. D. Terzopoulos and D. Metaxas. Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics. *PAMI*, 13(7):703–714, 1991.
16. J.-P. Thirion. New Feature Points based on Geometric Invariants for 3D Image Registration. *Internat. Journal of Computer Vision*, 18(2):121–137, 1996.
17. V. Vaerman, G. Menegaz, and J.-P. Thiran. A Parametric Hybrid Model used for Multidimensional Object Representation. *Proc. ICIP'99*, vol. 1, pp. 163–167, 1999.
18. B.C. Vemuri and A. Radisavljevic. Multiresolution Stochastic Hybrid Shape Models with Fractal Priors. *ACM Trans. on Graphics*, 13(2):177–207, 1994.
19. Z. Zhang. Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting. *INRIA Rapport de recherche*, No. 2676, 1995.