A hierarchical approach to model-based skull segmentation in MRI volumes

Marcel Lüthi1, Anita Lerch1, Thomas Albrecht1, Zdzislaw Krol2, Stefan Zimmerer2, and Thomas Vetter1

1 Computer Science Department, University of Basel, Switzerland
2 Hightech Research Center of Cranio-Maxillofacial Surgery, University of Basel

Abstract. We present a model-based approach for segmentation of the skull from T1 weighted MR images of the human head. Segmentation is performed by fitting a statistical shape model of the skull into a crudely pre-segmented version of the image. This yields a segmentation result that is constrained to the normal skull anatomy and thus gives a statistically meaningful approximation of the skull structure, even in places where bone cannot be distinguished from the surrounding tissue. We propose a multi-resolution approach to model fitting and show how a hierarchy of shape models can be used to increase the flexibility of the model. To validate our method, we show fitting results for different MR images and perform a quantitative comparison with a data set where a ground truth segmentation from a CT image is available. Our experiments confirm that even such difficult segmentation tasks as skull segmentation from MR images become feasible given a strong shape prior.

1 Introduction

Models of the human anatomy play an increasingly important role in understanding the human physiology. Building accurate and representative models often requires the fusion of information from different image modalities. While soft-tissue and vessels are relatively easy to segment from Magnet Resonance (MR) images, the bony structure is hardly visible on MR images and requires the use of an additional Computed Tomography (CT) image. This is not only time-consuming, but also exposes a patient to harmful radiation and complicates the model building process by requiring additional registration steps. In this article we present a method for segmentation of the skull directly from T1 weighted MR images, using a statistical shape model, built from a set of CT images, as a shape prior.

Direct segmentation of bones in MR images is a difficult problem, as the bony structure is hard to distinguish from the surrounding tissue and virtually impossible to distinguish from air. The low resolution of typical MR images and the occurrence of fine bony structure in the human skull, makes skull segmentation from MR images a particularly challenging problem. To overcome these difficulties, we exploit prior knowledge about the shape of normal skulls in the
form of a statistical shape model. The statistical model is built from a number of skull surfaces, which are segmented from CT data sets. Hence, our method makes MR segmentation feasible by using knowledge learned from CT data to restrict the solution space to valid skull shapes in the actual MR segmentation.

Our method proceeds in two steps. First, we perform a pre-segmentation of the skull by applying simple thresholding and morphological operations to the image. This yields a rough outline of the skull, which we call the pre-segmentation. In a second step we fit a statistical shape model to this pre-segmented image, which restricts the solution to anatomically valid skull shapes. As the shape model is built from CT data, the fitted model also approximates the skull shape in regions where the bony structure is not visible in the MR image. We propose the use of a hierarchy of models to increase the expressiveness and flexibility of the shape model. For this we divide the shape into several parts, such as the cranium and mandible for the human skull. In a first step, the complete shape model is fitted. Using this fit as an initial solution, the fitting is then performed on the individual parts. This does not only lead to more accurate fitting results, but also allows to segment images in which the mouth is open. The actual fitting is formulated as an energy minimization problem. As the cost function is highly non-linear, we propose to use of a multi-resolution scheme for model fitting, similar to the schemes commonly used in image registration.

The paper is organized as follows: In section 2 we present the mathematical basics of building statistical models and clarify our notation. Section 3 discusses the algorithms we use for pre-segmentation of the MR images. Our approach to model fitting is explained in section 4, where we also introduce the multi-resolution and hierarchical approach. We assess the quality of our method in section 5 by providing fitting result for several MR images. Additionally, we present a quantitative evaluation, by comparing our segmentation results to a ground truth segmentation obtained from a threshold segmented CT image.

Related work: Automatic segmentation of medical images is an extremely well researched problem and the need for incorporating statistical information to constrain the segmentation process has clearly been recognized [1]. The use of statistical models for identifying structures in (2D) medical images has already been proposed by Cootes et al. in 1993 [2] and since then applied for segmentation of various structures from 3D CT and MR images [3–7]. All these methods have been evaluated on rather simple anatomical structures (such as the femur or pelvic bone). However, no results were shown for more complex structures such as the human skull. The literature on skull segmentation is very sparse in general. Results for the segmentation of the skull from CT images are shown by Kang et al. [8], where, in a multi-step approach, a sequence of standard segmentation techniques is applied to extract the bony structure. Dogdas et al. [9] use techniques from mathematical morphology to segment the skull from MR images. Rifai et al. [10] use a level-set segmentation technique to deform the contour of the scalp to fit the skull structure.
2 Statistical Skull Model

Statistical shape models have become a widely used tool in medical image analysis, computer vision, and computer graphics. The main idea is to span a space of shapes (3D Surfaces) from a set of normal examples by taking linear combinations of these examples [11]. A probability distribution (usually a normal distribution) is defined on this space, which quantifies the probability of observing a particular linear combination. In image processing, the most common way to use this information is to limit the search space of an algorithm to likely instances of the shape. Figure 1 shows an example of the statistical shape model of the human skull used throughout this paper. It shows the mean skull as well as the first two principal modes of variations.

![Fig. 1. (a) The mean skull. (b)-(e) The first two modes of variation of the skull model (σ_i denotes 1 standard deviation in the direction of the i-th principal component).](image)

We now present the details of building a statistical shape model. Let \( \{ \Gamma_i \subset \mathbb{R}^3 \mid i = 1, \ldots, n \} \) be \( n \) surfaces, given in some suitable representation. Define an arbitrary surface, say \( \Gamma_1 \), as a reference surface. We assume that each surface \( \Gamma_i \) was obtained by warping the reference surface \( \Gamma_1 \) with a smooth vector field \( \phi_i : \Gamma_1 \to \mathbb{R}^3 \). That is

\[
\Gamma_i = \{ x + \phi_i(x) \mid x \in \Gamma_1 \}.
\]

Let \( \hat{\Gamma}_i \) be a suitable discretization of \( \Gamma_i \) of \( N \) points (e.g. \( \hat{\Gamma}_i \) is represented as a triangle mesh with \( N \) vertices). Note, that the same discretization is induced by the mapping \( \phi \) for each surface \( \Gamma_i \). We define the shape vector \( s_i \in \mathbb{R}^{3N} \) as

\[
s_i = (v_{i,1}^1, v_{i,1}^2, v_{i,1}^3, \ldots, v_{i,N}^1, v_{i,N}^2, v_{i,N}^3)^T,
\]

where the vector

\[
v^{i,j} = (v_{x}^{i,j}, v_{y}^{i,j}, v_{z}^{i,j}) = (v_{x}^{i,j} + [\phi_i(v^j)]_x, v_{y}^{i,j} + [\phi_i(v^j)]_y, v_{z}^{i,j} + [\phi_i(v^j)]_z)
\]

represents the \( x, y, z \) coordinates of the \( j \)-th vertex of \( \hat{\Gamma}_i \).

Usually we are given the surfaces \( \Gamma_1, \ldots, \Gamma_n \) rather than a reference surface \( \Gamma_1 \) and the corresponding vector fields \( \{ \phi_i \}_{i=2}^n \). Finding a vector field \( \phi \) that maps between a given pair of surfaces is a central problem in medical imaging.
and computer vision and is referred to as the registration or correspondence problem. To find the correspondence, we use the algorithm proposed by Dedner et al. [12].

Principal Component Analysis Although PCA is a well known method, we will give a rather detailed derivation in order to clarify notation and introduce the concepts used later in section 4. We start by defining the mean free data matrix $X \in \mathbb{R}^{3N \times n}$ as

$$X = [s_1 - \mu \ldots s_n - \mu], \quad (1)$$

with columns $x_i := s_i - \mu$. Let $\tilde{n} = \text{rank}(X)$ be the column rank of $X$. Using the singular value decomposition (see e.g. Demmel [13]), we can write

$$X = UDV^T. \quad (2)$$

where $U = [u_1 \ldots, u_\tilde{n}] \in \mathbb{R}^{3N \times \tilde{n}}$ is a matrix satisfying $U^TU = I$, $D \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ a diagonal matrix with strictly positive diagonal entries $\sigma_i$, $i = 1, \ldots, \tilde{n}$ and $V^T = [v_1 \ldots v_n]^T \in \mathbb{R}^{\tilde{n} \times n}$ is a matrix, satisfying $V^TV = I$. By choosing a $3N$-by-$(3N - \tilde{n})$ matrix $\tilde{U}$ so that $[U \tilde{U}]$ is square and orthogonal we can write

$$\Sigma = \frac{1}{n} \sum_{i=1}^{n} (s_i - \mu)(s_i - \mu)^T = \frac{1}{n} \sum_{i=1}^{n} x_i x_i^T = \frac{1}{n} XX^T$$

$$= \frac{1}{n} UDV^TVDU^T = [U \tilde{U}] \begin{bmatrix} \frac{1}{n}D^2 & 0 \\ 0 & 0 \end{bmatrix} [U \tilde{U}]^T. \quad (3)$$

Equation (3) is an eigendecomposition of $\Sigma$ and hence the columns of $[U \tilde{U}]$ are the eigenvectors of $\Sigma$. We refer to the eigenvector $u_i$ as the $i$-th principal component. It follows that expressed in the basis of principal components, $\Sigma$ becomes a diagonal matrix whose diagonal entries $\sigma_i^2$ correspond to the variance represented by $u_i$. Note that all the variance in the data is captured by the $\tilde{n}$ principal components. Using this fact, we can define a probability distribution over the shapes by expressing a shape $s$ as

$$s = \mu + \sum_{i=1}^{\tilde{n}} \alpha_i u_i$$

and inducing the normal distribution $\alpha \sim \mathcal{N}(0, \frac{1}{n}D^2)$ on the coefficients. The probability density over the shapes $s$ has the simple form

$$p(s) := p(\alpha) = z \exp\left(\frac{1}{n}||D^{2-1}\alpha||^2\right). \quad (4)$$

where $z = \frac{n}{(2\pi)^{(n-\tilde{n})/2}||D||^{1/2}}$ is the normalization factor. The coefficients $\alpha$ are referred to as the PCA coefficients.

We stress at this point that any linear combination of the principal components $u_i$ can be expressed as a linear combination of the example data $x_i$ and
vice-versa. Let \( \alpha = (\alpha_1, \ldots, \alpha_n)^T \), \( \beta = (\beta_1, \ldots, \beta_n) \) be coefficient vectors. Using the relation \( U = XVD^{-1} \) from Equation (2) we can write

\[
\sum_{i=1}^{n} \alpha_i u_i = U\alpha = XVD^{-1}\alpha = X\beta = \sum_{i=1}^{n} \beta_i x_i.
\]

Hence we have the relations

\[
\beta = VD^{-1}\alpha \quad \text{and} \quad \alpha = DV^T\beta.
\]  

This fact allows us to keep the same linear combination of example vectors, while we freely change among different basis representations.

3 Pre-Segmentation

The pre-segmentation step provides an automated procedure for finding a rough approximation of the skull structure, by means of simple image-processing operations. While segmenting bony structure directly from MR images is difficult, the brain and scalp can rather easily be segmented. These regions already constrain the position of the skull relatively well. We mask them in the image and perform a threshold segmentation on the remaining region. Selecting the largest connected component yields the initial segmentation result. Figure 3 illustrates the different steps.

![Fig. 2. The pre-segmentation procedure. The original image (a) is masked with the brain and scalp mask (b). After thresholding, the final pre-segmentation result is obtained (c) and (d). The ellipses in (c) mark places where the segmentation result is wrong. The blue ellipses show areas where air is classified as bone, the yellow ellipses show the same for bone-marrow and the red ellipses for CSF.](image)

**Scalp-segmentation** For scalp segmentation we are using the method proposed by Dodgas et al. [9]. The segmentation is performed by using a combination of thresholding operations, to segment the soft-tissue from the bone, and
mathematical morphology to close the auditory channels and the nasal cavity. Performing a hole filling, leads to the final scalp mask.

**Brain-segmentation** For brain segmentation we follow the method proposed by Géraud et al [14, 10]. Similarly to the scalp-segmentation, the method uses a combination of thresholding and mathematical morphology operations. Morphology is used to separate the brain from the surrounding structures. Selecting the largest component yields the final brain segmentation.

**Pre-segmentation results** Figure 3 shows that the overall skull shape is already roughly approximated after this initial step. However, the pre-segmentation fails at several places. For example in the area around the sinuses, thin bones and air are located side by side. The algorithm fails to distinguish them, as air and bones give almost the same signal. Also, the cerebrospinal fluid (CSF) cannot be distinguished from bone by this simple method, as the intensity value of CSF in a T1 weighted image lies close to that of bone. A different problem occurs with the bone marrow, which gives a distinct signal from bone and is therefore not excluded in the thresholding step.

### 4 Fitting

As discussed in Section 2, the main idea behind statistical models is to learn a space of shapes by taking suitable (i.e. probable) linear combinations of the example shapes. The shapes are most conveniently represented as linear combinations of the principal components $u$ (cf. Section 2):

$$s = \mu + \sum_{i=1}^{n} \alpha_i u_i. \quad (7)$$

Recall that the probability of observing $s$ depends solely on the size of $\|D^{2^{-1}} \alpha\|^2$. These properties lead naturally to an approach for improving the result from the pre-segmentation step: We try to find a linear combination of the examples, such that the shape matches the one given from the pre-segmentation step well, but discourages unlikely combinations by penalizing large $\alpha$.

We proceed by stating this idea more formally. Let $I : \Omega \subset \mathbb{R}^3 \rightarrow \{0, 1\}$ be the resulting binary image from the pre-segmentation step and $\tilde{\Gamma} \subset \mathbb{R}^3$ be the surface that induces the partition in $\tilde{I}$. We define the distance image as

$$I(x) = \begin{cases} \text{dist}(x, \tilde{\Gamma}) & x \in \text{outside}(\tilde{\Gamma}) \\ 0 & x \in \tilde{\Gamma} \\ -\text{dist}(x, \tilde{\Gamma}) & x \in \text{inside}(\tilde{\Gamma}). \end{cases} \quad (8)$$

Finding a likely shape that fits the partition boundary of $\tilde{I}$ amounts to finding transformation and model parameters, such that the resulting surface comes to
lie on the zero-level set of $I$. This is satisfied by the solution of the optimization problem

$$
\min_{s,t,R,\alpha} \sum_{i=1}^{N} w_i I(sR[\delta_i(\mu + \sum_{j=1}^{n} \alpha_j u_j)] + t)^2 + \lambda \| D^{-1} \alpha \|^2
$$

where $s \in \mathbb{R}$ is a scaling factor, $t \in \mathbb{R}^3$ a translation, $R \in \mathbb{R}^{3 \times 3}$ a rotation matrix and $\lambda$ a weighting coefficient. The function $\delta_i(x) : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ extracts the coordinates belonging to the $i$-th vertex from the shape vector. Finally, $w_i \in \mathbb{R}$ is a weight for point $i$. Note that the first term is zero, if and only if the shape defined by the optimization parameters lies exactly on the zero-level set that defines the partition boundary $\tilde{\Gamma}$. The second term in (9) serves as a regularizer and penalizes unlikely shapes. As the cost-function is only evaluated at a discrete number of points of the surface, areas that are more densely sampled will have a larger influence on the error. The weight $w_i$ in (9) can be used to compensate for this effect. In our implementation, we set $w_i$ proportional to the area of the size of the neighborhood of vertex $v_i$.

The problem (9) can be minimized by using any gradient based optimization scheme. Unfortunately, the dependence on $I$ makes it highly non-convex, and hence the algorithm is likely to get stuck in a local minimum.

**Multi-Resolution** A common technique in image processing to escape local minima is to use a multi-resolution scheme. We will show how this strategy can be adapted for statistical model fitting.

For the distance image $I$ we build an image pyramid to obtain a set of images $I^{(0)}, \ldots, I^{(L)}$, with decreasing level of details. $I^{(0)}$ corresponds to the original image and $I^{(l)}$ is obtained from $I^{(l-1)}$ by Gaussian smoothing and subsampling. Unfortunately, the zero-level set in the subsampled images $I^{(l)}, i > 0$, does not necessarily represent the original surface anymore, as fine structures are smoothed away. Therefore, we need to apply the same procedure to the reference surface $\Gamma_0$. This can be achieved by forming the distance image of $\Gamma_0$ and compute the image pyramid in exactly the same manner. Using a contour extraction algorithm we extract the zero level set to obtain the reference surface $\Gamma_0^{(l)}$ for level $l$. Repeating the procedure described in section 2 with $I_0^{(l)}$ as the reference surface, we build a separate statistical model for each level.

As the statistical model is built separately for each resolution level, the PCA components may be different and the coefficients $\alpha$ do not represent the same shape anymore. We need to perform a change of basis to the new principal components. Let $\alpha^{(l)}$ be the final PCA coefficients vector obtained from optimizing (9) for level $l$. The relation between the PCA coefficients $\alpha$ and the example-coefficients $\beta$ given in Equation (6), allows us to compute the coefficients for the next level from the previous solution. We compute from the PCA coefficients $\alpha$ the corresponding coefficients $\beta$, which determine the linear combination of the examples $x_1, \ldots, x_n$

$$
\beta^{(l)} = V^{(l)} D^{-1} \alpha^{(l)}.
$$

(10)
Algorithm 1: The fitting algorithm

Input: Number-of-levels \( L \)
Reference-Mesh \( \Gamma_0 \)
Model \( (\mu^{(l)}, U^{(l)}, V^{(l)}, D^{(l)}), l \in \{0, \ldots, L\} \)
Distance-Image \( I^{(l)}, l \in \{0, \ldots, L\} \)

Output: \((s, R, t, \alpha) = (s^{(L+1)}, R^{(L+1)}, t^{(L+1)}, \alpha^{(L+1)}) := (1, I, 0^3, 0^n)\)

1 \( D^{(l+1)} = V^{(l+1)} \)
2 \( \text{foreach } l \in [L, \ldots, 0] \) do
3 \( \alpha := D^{(l+1)} V^T (l+1) V^{(l)} D^{(l)} - 1 \alpha^{(l+1)} \)
4 \( \text{SetInitialParameters} (s^{(l+1)}, R^{(l+1)}, t^{(l+1)}, \alpha^{(l+1)}) \)
5 \( \text{min}_{\sum_i w_{i}} I^{(l)} (s R [\delta^{(l)} (\mu^{(l)} + \sum_j \alpha_j u^{(l)}_j)] + t)^2 + \lambda \| D^{(l)} \|_2 \alpha \) \)
6 \( \text{return} (s^{(0)}, R^{(0)}, t^{(0)}, \alpha^{(0)}) \)

The same linear combination \( \beta^{(l)} \) of the examples is used as the initial solution for the next level. In terms of the PCA coefficient \( \alpha^{(l-1)} \) for the next level, this becomes

\[
\alpha^{(l-1)} = D^{(l-1)} V^T \beta^{(l)} = D^{(l-1)} V^T \beta^{(l)} \]

We proceed iteratively until we reach the finest resolution. Algorithm 1 shows the detailed steps of the multi-resolution fitting.

Multi-parts fitting Representing such complex shapes as the human skull with a single statistical model requires a large number of normal examples for model building. As high quality CT images are often scarce, we try to increase the flexibility of the model without resorting to additional data. This can be achieved by dividing the shape into several, simpler structures, which are easier to represent. Each such part is then fitted independently. A natural segmentation for the skull would, for instance, be to separate the mandible and the cranium.

Our approach is as follows: In a first step, we fit a statistical model of the whole skull. This yields a good approximation of the overall structure and localizes the different parts. Using this result as an initial solution, we fit a separate statistical model for each part into the image. From a technical point of view, this hierarchical fitting does not pose additional challenges. We can employ the same method for changing among the different bases as described for the multi-resolution approach, provided that we keep the same order of the examples \( x_i \) for all the parts. Algorithm 2 illustrates this procedure in more detail.

5 Results

Experimental setup. We implemented our method using the Insight Registration and Segmentation Toolkit (ITK) [15]. Before the skulls are fitted they are roughly aligned with the reference surface using landmark based rigid registration [16]. Our statistical model has been built from 23 segmented skulls from
Algorithm 2: Multi-part fitting algorithm

Input: Number of parts $P$
Reference-Mesh $\Gamma_0$
Model $(\mu(p), U(p), V(p), D(p)), p \in \{0, \ldots, P\}$
Distance-Image $I$

Output: $(s(p), R(p), t(p), \alpha(p)), p \in \{0, \ldots, P\}$

1. $(s(0), R(0), t(0), \alpha(0)) = \text{Fit}(\Gamma_0)$
2. \textbf{foreach} part $\Gamma_{0}(p) \subset \Gamma_0$ \textbf{do}
3. $\pi := D(p) V V(p) V(0) D(0)^{-1} \alpha(0)$
4. SetInitialParameters $(s(0), R(0), t(0), \pi)$
5. $(s(p), R(p), t(p), \alpha(p)) = \text{Fit}(\Gamma_{0}(p))$
6. \textbf{return} $(s(p), R(p), t(p), \alpha(p)), p \in \{0, \ldots, P\}$

Fig. 3. Skull model fitting: The places where the pre-segmentation (a) fails, are segmented correctly by the fitting procedure (Figures (b) and (c)). Multi-part fitting (Figure (c) and (d)) gives clearly better results than fitting of the complete skull (b).

Experiments and Results. In the first experiment we performed a fit of the skull model into the automatically pre-segmented MR image. It can be seen in Figure 5 that the skull shape is generally well approximated and also gives meaningful results at places where the intensity information does not allow to distinguish the bone from the surrounding tissue. However, for the mandible the fitting error is quite large. This can be explained by the limited flexibility of the model, which does not allow for a perfect fit of the complete structure. Using the result of the complete skull fit as an initial solution, we performed a fitting of the separate mandible and cranium model. While for the cranium the improvement is small (the difference is mainly in the area of the sinuses), the result for the mandible is clearly much better. The flexibility of the multi-part fitting scheme is also nicely illustrated in the example shown in Figure 4, in which a skull with open mouth

---

diverse sources, out of which 7 stem from the Bosma collection [17]. For testing the multi-part fitting, we built separate models of the mandible and the cranium. The fitting is performed on 4 levels with 2000 iterations each. In all the experiments the regularization parameter $\lambda$ is set to 0.1.
has to be fitted. As the mandible is decoupled from the cranium, the method can fit both parts individually and match them accurately.

![Fig. 4. Multi-part fitting: (a) A mask for a skull with open mouth. The fitting of the complete model shown in (b) cannot match the lower jaw. In the multi-part fitting (c), the model has the flexibility to open the jaw.](image)

To assess the results quantitatively, we compared the segmentation results on an MR image to the one obtained on a CT image of the same person. As the ground truth we use a segmentation of the CT image, which was obtained using threshold segmentation. In the images used for this comparison the mouth is open. Therefore the results from automatic pre-segmentation are worse than this is usually the case. We therefore manually refined the pre-segmentation to remove the most severe artifacts. We fitted the model to the segmented CT image, the manually refined pre-segmentation, and the fully automatic pre-segmentation of the MR image. For each of our experiments, we computed the root mean squared error (RMSE) and the Hausdorff Distance (HD) between the surfaces. To assess the bias of the model, we also fitted a model where the ground truth is part of the model. The results are summarized in Table 1. The improvements of the multi-parts fitting can be clearly seen. Comparing the RMSE and the Hausdorff-Distance, we see that the overall structure is fitted well, but there are places where the fitting fails. This can partly be explained by the fact that the ground truth itself contains artifacts (such as missing teeth), whereas the model is complete. However, the results also make clear that the expressiveness of the model needs to be increased further if all the fine structures found in the human skull need to be captured accurately.

6 Discussion

In this paper we presented an approach for the segmentation of the human skull by means of fitting a statistical shape model to a pre-segmented MR image of the head. To efficiently solve the resulting optimization problem, we proposed
Table 1. Quantitative comparison of the different fitting results: For each fitting and pre-segmentation method, the distance (in mm) of the fitted surface to the ground truth is computed.

<table>
<thead>
<tr>
<th></th>
<th>CT RMSE</th>
<th>HD</th>
<th>MRI manual RMSE</th>
<th>HD</th>
<th>MRI automatic RMSE</th>
<th>HD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete skull</td>
<td>2.7</td>
<td>21.4</td>
<td>3.5</td>
<td>22.4</td>
<td>3.9</td>
<td>20.8</td>
</tr>
<tr>
<td>Complete skull (in model)</td>
<td>1.3</td>
<td>13.7</td>
<td>2.1</td>
<td>13.1</td>
<td>3.0</td>
<td>13.3</td>
</tr>
<tr>
<td>Multi-parts</td>
<td>1.9</td>
<td>11.9</td>
<td>2.9</td>
<td>16.3</td>
<td>3.4</td>
<td>12.5</td>
</tr>
<tr>
<td>Cranium</td>
<td>1.8</td>
<td>11.8</td>
<td>3.0</td>
<td>17.6</td>
<td>3.5</td>
<td>12.5</td>
</tr>
<tr>
<td>Mandible</td>
<td>2.4</td>
<td>11.6</td>
<td>2.39</td>
<td>12.3</td>
<td>3.2</td>
<td>11.2</td>
</tr>
</tbody>
</table>

a multi-resolution approach to model fitting. Our method limits the possible solutions to those shapes lying in the span of the statistical skull-models, i.e. linear combinations of the examples used to build the model. This strong restriction enables the method to yield meaningful segmentation results, even in places where the bone cannot be distinguished from the surrounding tissue in the image. At the same time, it is this strong bias that limits the flexibility of the model. It would require a large number of example skulls in order to span the space of all normal skull shapes. We therefore proposed the use of a hierarchy of statistical models to increase the flexibility of the model, without requiring more example data. Our experiments showed that this greatly improves the accuracy of the segmentation.

The statistical shape model has been built from example skulls that were semi-automatically segmented from CT data. Hence our approach can be seen as a means of transferring shape information from CT data to MR images. At some point this might reduce the need for acquiring separate CT image of the patient and make it possible to obtain all information that is needed for the medical application from a single MR image.

The full potential of the hierarchical model has not been exploited in this paper. One could consider more complex hierarchies and use a separate model for every part. For example, a separate statistical model for each tooth could be built. Combined with an intelligent algorithm to automatically detect missing parts in data (using a scheme like the one described by Toews et al. [18]) it would be possible to fit images where no one-to-one correspondence to the reference-structure exists. For more complex hierarchies, however, a problem that has to be addressed is how continuity among the segments can be guaranteed.

The fitting procedure described here is by no means limited to skull segmentation, but can be applied to any segmentation task, where a good enough pre-segmentation can be performed. A direct application is the segmentation of bone from CT images, where the pre-segmentation can easily be performed by a simple thresholding operation. The development of more complex hierarchical models will be subject of future work.
Acknowledgements We thank Dr. Philippe Cattin, Medical Image Analysis Center, University of Basel, for providing us with the MR and CT data set for the validation. This work was funded by the Swiss National Science Foundation in the scope of the NCCR CO-ME project 5005-66380 (http://co-me.ch).

References