

# Personalized Articulated Atlas with a Dynamic Adaptation Strategy for Bone Segmentation in CT- or CT/MR-Head&Neck Images

Sebastian Steger<sup>a</sup>, Florian Jung<sup>ab</sup>, and Stefan Wesarg<sup>a</sup>

<sup>a</sup>Medical Computing, Fraunhofer IGD, Darmstadt, Germany;

<sup>b</sup>Graphisch-Interaktive Systeme, Technische Universität Darmstadt, Darmstadt, Germany

## ABSTRACT

This paper presents a novel segmentation method for the joint segmentation of individual bones in CT- or CT/MR- head&neck images. It is based on an articulated atlas for CT images that learned the shape and appearance of the individual bones along with the articulation between them from annotated training instances. First, a novel dynamic adaptation strategy for the atlas is presented in order to increase the rate of successful adaptations. Then, if a corresponding CT image is available the atlas can be enriched with personalized information about shape, appearance and size of the individual bones from that image. Using mutual information, this personalized atlas is adapted to an MR image in order to propagate segmentations. For evaluation, a head&neck bone atlas created from 15 manually annotated training images was adapted to 58 clinically acquired head&neck CT datasets. Visual inspection showed that the automatic dynamic adaptation strategy was successful for all bones in 86% of the cases. This is a 22% improvement compared to the traditional gradient descent based approach. In leave-one-out cross validation manner the average surface distance of the correctly adapted items was found to be 0.68 *mm*. In 20 cases corresponding CT/MR image pairs were available and the atlas could be personalized and adapted to the MR image. This was successful in 19 cases.

**Keywords:** Articulated Atlas, Personalization, Dynamic Adaptation Strategy, Bone Segmentation, MRI

## 1. INTRODUCTION

The segmentation of individual bones is required for computer aided diagnosis (CAD) and therapy planning applications. Due to its high contrast, bone can be segmented in CT images with high accuracy. A more complex problem is the segmentation and annotation of individual bones from CT images (e.g. the individual vertebrae). Since neighboring bones are in some cases very close to each other and have very similar appearance, a joint segmentation is required. For MR images, both problems are a lot more challenging due to the lower contrast. However, if both imaging modalities of the same subject are available, a precise segmentation in the CT image can be exploited to obtain a segmentation in the corresponding MR image. An example for that is radiation therapy planning where both imaging modalities may be available, MRI to delineate the tumor and other soft tissue structures and CT to delineate many high contrast organs at risk and do the dose planning. As a by-product the correspondences of the bones in both modalities can be used for the skeleton based multi-rigid registration of both imaging modalities.<sup>1</sup>

For CT images, the joint segmentation of multiple bones can be achieved using an articulated statistical shape model (e.g. the spine<sup>2</sup>) or an articulated atlas (e.g. the entire skeleton of mice<sup>3</sup> or the head and the cervical spine.<sup>4</sup> For MR images, fewer solutions exist (e.g. for the segmentation of the knee<sup>5</sup> and the detection of the spine<sup>6</sup>). Whereas in Ref. 3 and 5 the joints are explicitly modeled, Ref. 2 and 4 learn the articulation from training instances.

The goal of this work is to enrich the articulated atlas presented Ref. 4 with personalized information from a previously segmented CT image of the same subject in order to segment individual bones from a corresponding

---

Further author information: (Send correspondence to Florian Jung)

Florian Jung: E-mail: florian.jung@igd.fraunhofer.de, Telephone: +49 6151 155-520

Sebastian Steger: E-mail: sebastian.steger@igd.fraunhofer.de, Telephone: +49 6151 155-507

Stefan Wesarg: E-mail: stefan.wesarg@igd.fraunhofer.de, Telephone: +49 6151 155-511

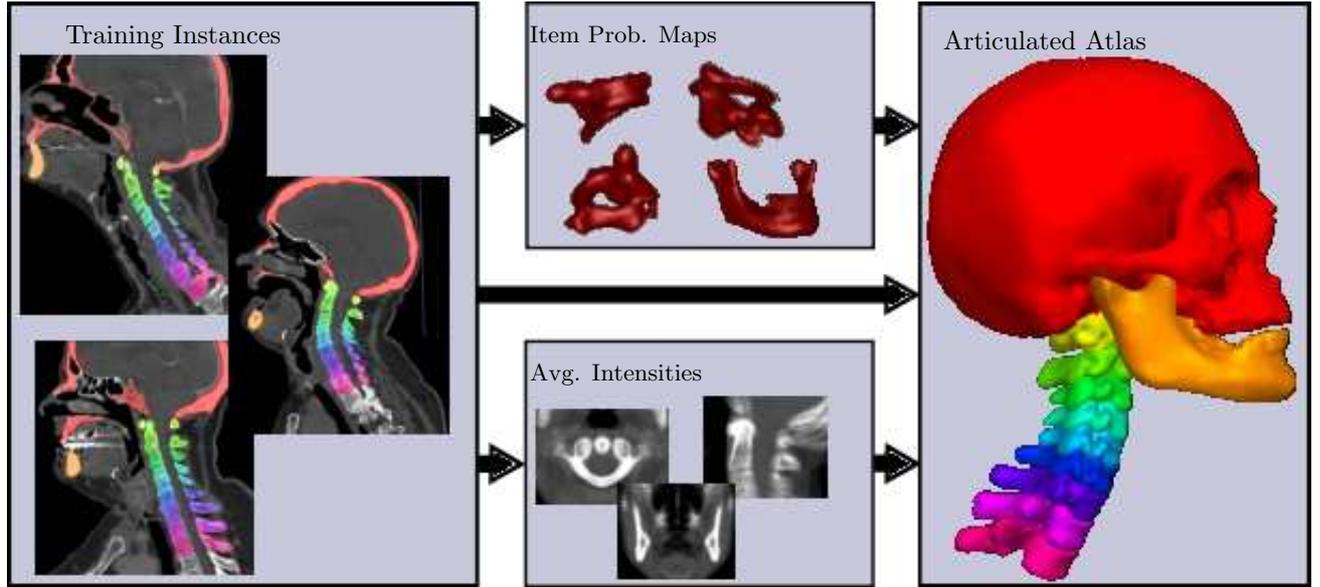


Figure 1. Training of the articulated atlas from annotated images for segmentation of bones in CT images. The articulated atlas consists of shape and appearance of individual bones along with the articulation between them.

MR head&neck image with high accuracy. After a short introduction of the articulated atlas and the presentation of an advanced articulated atlas adaptation strategy, we describe the necessary steps for the atlas personalization in order to be able to successfully adapt it to MR images. Finally, the atlas adaptation strategy and the personalization is evaluated quantitatively on clinical datasets.

## 2. ARTICULATED ATLAS

### 2.1 Articulated Atlas Creation

The articulated atlas consists of  $n$  atlas items, each representing a different bone (e.g. skull, mandible, vertebra C3). The shape and appearance of each atlas item as well as the relative spatial location and orientation of all items (articulation) is learned from  $m$  manually labeled training instances that shall cover the possible shape, appearance, and spatial location space as good as possible.

The shape of an item is represented by a probability map  $s : D \subset \mathbb{R}^3 \mapsto [0, 1]$  indicating the membership to the item computed in a coordinate system normalized by translation, rotation and scaling. This coordinate system is computed by a group wise alignment of the training data masks maximizing the mutual overlap. The probability map is created by averaging the aligned training data masks.

The appearance of an item is given by the average intensities  $g_i : D \mapsto \mathbb{R}$  extracted from the training intensity images aligned to that normalized coordination system.

Each bone item has those 7 degrees of freedom relative to the atlas coordinate system and thus the articulation of the entire atlas can uniquely be described with a  $7n$  dimensional vector  $\mathbf{p}$ . To enable vector space like operations for the transformations all transformations are expressed by linear combinations of unit transformations (unit translation, unit rotation, unit scale) using the matrix exponential representation.<sup>7</sup> The space of all possible articulations can be described by applying principal component analysis (PCA) on the training articulation matrix  $P = (\mathbf{p}_1, \dots, \mathbf{p}_m)$ , where  $\mathbf{p}_j$  is the articulation vector of the specific training instance  $j$ . An arbitrary articulation  $\mathbf{p}$  can then be described with a  $k \ll 7n$  dimensional vector  $\mathbf{b}$ :

$$\mathbf{p} = \bar{\mathbf{p}} + A \cdot \mathbf{b} + \mathbf{r} , \quad (1)$$

where  $\bar{\mathbf{p}}$  is the average articulation of all training articulations  $\mathbf{p}_j$ ,  $A$  is a matrix consisting of the eigenvectors of the covariance matrix  $C = \frac{1}{m-1} \bar{P} \bar{P}^T$ ,  $\bar{P} = (\mathbf{p}_1 - \bar{\mathbf{p}}, \dots, \mathbf{p}_m - \bar{\mathbf{p}})$ , and  $\mathbf{r}$  is the residual vector. The more likely an articulation  $\mathbf{p}$  is, the smaller the magnitude of  $\mathbf{b}$  and  $\mathbf{r}$  will be.

## 2.2 Articulated Atlas Adaptation

The atlas is adapted to a CT image by minimizing an energy functional depending on the 6 dimensional vector  $\mathbf{x}$  describing a global rigid transformation and the articulation parameter vector  $\mathbf{p}$ :

$$\begin{aligned} E(\mathbf{p}, \mathbf{x}) &= E_{\text{external}}(\mathbf{p}, \mathbf{x}) + \alpha E_{\text{internal}}(\mathbf{p}) \\ &= \sum_{i=1}^n E_{\text{external}}^i(\mathbf{p}, \mathbf{x}) + \alpha E_{\text{internal}}(\mathbf{p}) . \end{aligned} \quad (2)$$

The empirically determined parameter  $\alpha$  balances the external and internal energy.

The external energy  $E_{\text{external}}$  ensures gray value similarity between the test image  $I : \mathbb{R}^3 \mapsto \mathbb{R}$  and the transformed trained intensities of the atlas items  $g_i$ . For a single atlas item it is defined as the weighted mean squared error:

$$E_{\text{external}}^i(\mathbf{p}, \mathbf{x}) = \sum_{\forall \mathbf{y} \in D} (G_\sigma \star s_i)(\mathbf{y}) [g_i(\mathbf{y}) - I((T_{\mathbf{x}} \circ T_{\mathbf{p}_i})(\mathbf{y}))]^2 , \quad (3)$$

where  $G$  is the Gaussian kernel,  $s_i$  the item's probability map describing the shape,  $T_{\mathbf{x}}$  is the atlas transformation described by external parameters  $\mathbf{x}$ , and  $T_{\mathbf{p}_i}$  is the atlas item transformation described by the item's internal parameters  $\mathbf{p}_i$ . The Gaussian smoothing of the item's probability image ensures that not only gray values within the model item but also in the neighborhood have an impact on the external energy.

The internal energy  $E_{\text{internal}}$  ensures that the atlas is within or at least close to the trained articulation space. To achieve this, we model the probability density of the articulations by an approximate Gaussian distribution using the approach of Moghaddam and Pentland.<sup>8</sup> The internal energy is defined as

$$E_{\text{internal}}(\mathbf{p}) = \sum_{j=1}^k \frac{b_j^2(\mathbf{p})}{\lambda_j} + \frac{m-k-1}{\sum_{j=k+1}^{m-1} \lambda_j} \|\mathbf{r}(\mathbf{p})\|^2 , \quad (4)$$

where the  $\lambda_j$ ,  $\lambda_1 \geq \dots \geq \lambda_{m-1} > 0$  are the non-zero eigenvalues of  $C$ . The first term is the Mahalanobis distance in the linear subspace  $\mathcal{U}$  spanned by the first eigenvectors  $A$ , and the second term approximates the Mahalanobis distance in the subspace complementary to  $\mathcal{U}$ . The latter term ensures that articulations are not completely restricted to the linear subspace  $\mathcal{U}$ , which increases the flexibility of the model.

A gradient descent optimizer first finds the global transformation parameters  $\mathbf{x}$  and in a second step the articulation and external parameters  $\mathbf{p}, \mathbf{x}$  jointly. Once the articulated model converged, the segmentations for each bone item are extracted based on the probability maps and the CT intensities. Please refer to Ref. 4 for details.

## 3. DYNAMIC ADAPTATION STRATEGY

Due to similar appearance of neighboring items (e.g. vertebra C4 looks very similar to vertebra C5), using a gradient descent based atlas adaptation strategy<sup>4</sup> may move multiple atlas items into the wrong structures in the CT image. The internal energy may not be able to recover this and thus, one atlas item covers two vertebrae in the test image or vice versa. To overcome this issue, we propose a dynamic adaptation strategy. The basic idea is to start the image driven adaptation with a reliable structure (here: the skull) and then exploit relations between neighboring atlas items. Figure 2 shows the order in which the external energy of atlas items are activated.

This idea is implemented by modifying the external energy given by equation (2):

$$E_{\text{external}}(\mathbf{p}, \mathbf{x}, t) = \sum_{i=1}^n w_i(t) E_{\text{external}}^i(\mathbf{p}, \mathbf{x}) , \quad (5)$$

where  $t$  is a parameter describing the progress of the adaptation process and  $w_i(t) \in [0, 1]$  are element weights depending on the adaptation progress. At the beginning ( $t = 0$ ), all weights are set to zero except for the most reliable item (here: the skull) whose weight is set to one. As the gradient descent based adaptation proceeds ( $t$  increases), the partial derivative magnitudes of the energy functional for the parameters of the skull eventually become smaller. As soon as they fall below a certain threshold, the neighboring atlas items (here: the mandible and vertebra C1) are *activated* by setting their weight to one as well. This process continues until all items

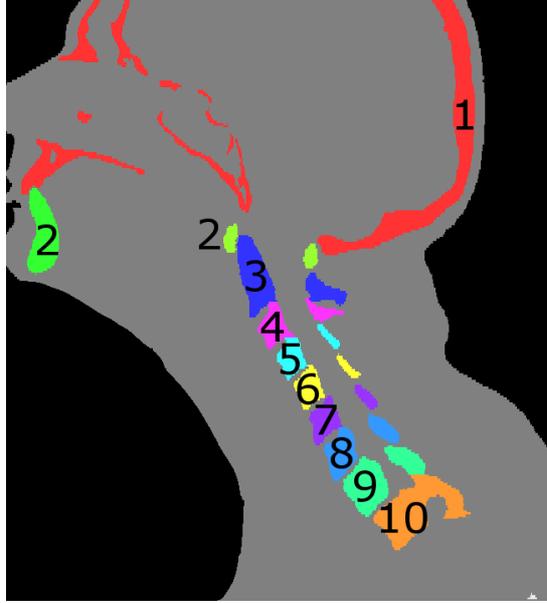


Figure 2. Order of item activation. Item gets activated as soon as the predecessor has converged. Inactive items are dragged along with active items based on learned articulation.

are activated and converged. Items that are not activated at a certain time during the process are driven by the internal energy only and consequently are dragged along with the activated items according to the trained articulation space. At the end, equation (5) becomes identical to the external energy described in equation 2 and thus, not the energy functional itself, but only the adaptation strategy has changed.

#### 4. PERSONALIZED ARTICULATED ATLAS

Unfortunately, the articulated atlas cannot directly be used for the segmentation of bones from MR images because the appearance model (the average intensities) depends on absolute image intensities and high image contrast of the structures. Based on the CT image of the same subject, the atlas adaptation parameters  $\mathbf{x}^*$ ,  $\mathbf{p}^*$  to that image and the resulting segmentations of the bones, the articulated atlas is personalized to enable an adaptation to MR images nonetheless:

- **Shape.** The probability map representing the shape is replaced by the binary mask given by the CT segmentation.
- **Appearance.** The average intensity image representing the appearance is replaced by the intensities from the CT image.
- **External Energy.** Instead of computing image intensity similarity, the mutual information (MI) between the transformed item appearance within a slightly dilated version ( $\approx 5mm$ ) of the item's segmentation and the MR image is computed for the external energy. Note that the MI must be computed for all items jointly based on a single joint histogram in order to increase statistical relevance compared to independent computation for each item.
- **Adaptation.** All components of  $\mathbf{p}$  representing item scale are extracted from  $\mathbf{p}^*$  and not considered during the adaptation process. This ensures that the atlas items are effectively rigidly transformed from the CT to MR image.

Once the personalized atlas is adapted to the MR image, the segmentations from the CT image are propagated. Figure 3 illustrates this approach.

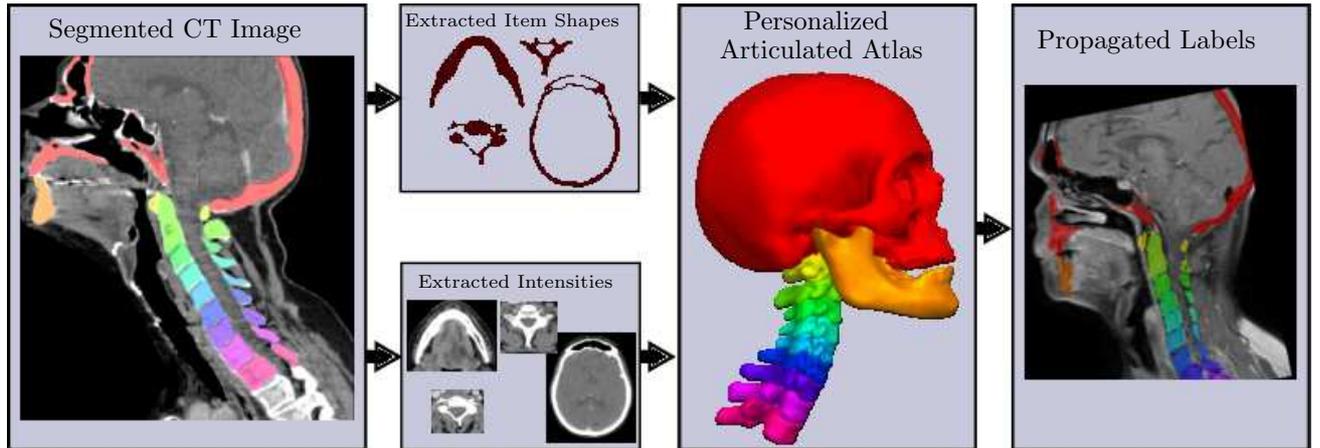


Figure 3. Personalization of the articulated atlas using a CT image for segmentation of bones from MR images. Shape and Appearance is extracted from corresponding CT image. Changes in item size are prohibited. Mutual information is used to measure similarity between item appearance and target MR image.

## 5. EXPERIMENTS & RESULTS

For evaluation, an articulated atlas has been created from 15 manually annotated head&neck CT images. It includes the skull, the mandible, the cervical vertebrae and the upper two thoracic vertebrae. This atlas was then adapted to 58 further head&neck CT images with i) the traditional gradient descent approach<sup>4</sup> and ii) with the dynamic adaptation strategy. Some of the images had strong imaging artifacts and/or were truncated.

As ground truth segmentation was not available, only qualitative results can be reported. An observer rated the location and rotation of the adapted atlas items and assigned them into three categories: i) correct alignment of all items, ii) at least one shifted vertebra, iii) complete misalignment. Table 1 shows the results. A success rate improvement of approximately 22% confirms the design goal of the dynamic adaptation. The main reasons for the still unsatisfactory cases are unusual anatomy (e.g. a missing inter-vertebral disk) or an articulation insufficiently covered by the learned articulation space.

The segmentation accuracy was assessed using leave-one-out cross validation. 14 of the training datasets were used to create the articulated atlas which was then applied to the remaining dataset. For the correctly adapted items (86% of all items) the segmentation was compared to ground truth segmentation. The average dice similarity coefficient and the average surface distance was 0.84 and 0.68 *mm* respectively.

For 20 of the 58 CT images, corresponding T1-weighted MR images with sagittal slices (3 – 5 *mm* slice distance) were available. The personalized articulated atlas was applied to those MR images. Visual inspection of the adapted atlas showed that in 19 of the 20 cases (95%), all atlas items are located correctly. Thus, even though adapting the atlas to a different imaging modality based on mutual information is a more difficult problem than the single modal case for CT images, the personalization of the atlas was able to increase the success rate of the atlas adaptation. Since the anatomy was assumed to be identical in both images (only the pose of the skeleton was different in general), the segmentation accuracy on the MRT images was identical to the previously reported CT case in case of perfect adaptation of the personalized articulated atlas.

	Gradient descent	Dynamic adaptation
Correct alignment	41(71%)	50(86%)
Shifted vertebra	14(24%)	7(12%)
Misalignment	3(5%)	1(2%)

Table 1. Atlas adaptation quality for different strategies. The adaptation of atlas items to neighboring structures is significantly reduced by the dynamic adaptation strategy and overall results in a 22% increase of the success rate.

## 6. CONCLUSION

We used a (personalized) articulated atlas for the joint segmentation of individual bones from CT or CT/MR images. The contributions are

- a **dynamic adaptation strategy** for the articulated atlas. Using this strategy, a 86% success rate for the correct adaptation of all atlas items can be achieved. This is a 22% improvement compared to the traditional gradient descent based approach.<sup>4</sup>
- the **personalization** of the articulated atlas which enables adapting it to MR images if a corresponding CT image is available. Even though bones have a much lower contrast in MR images, the atlas personalization may achieve sub-voxel segmentation accuracy, because shape information from the CT image with a typically finer inter-slice resolution is incorporated. In the case of perfect atlas adaptation, the same segmentation accuracy (0.68 mm average surface distance) as in the CT case is achieved by design.

The subsequent adaptation of the articulated atlas first to a CT image and then the personalized articulated atlas to a corresponding MR image enables an interesting application. It enhances the deformable registration of both images with the multi-rigid approach we presented in Ref. 1. From the adaptations to both images, a rigid transformation can be computed for each atlas item and based on that, a dense deformation field can be computed. The particularly interesting aspect of using the articulated atlas is that the articulation is at all times restricted to the learned articulation space and hence, the resulting registration scheme is subject to global regularization. This hardly seen property of a deformable registration method ensures globally plausible deformation fields.

Future work includes rebuilding the articulated atlas with more training instances, include more (possibly deformable) structures and extend it to other parts of the body.

## ACKNOWLEDGMENTS

This work is partially funded within the OraMod project (FP7-ICT-2013-10-611425) by the European Commission.

## REFERENCES

- [1] Steger, S. and Wesarg, S., “Automated skeleton based multi-modal deformable registration of head&neck datasets,” in [*Medical Image Computing and Computer-Assisted Intervention MICCAI 2012*], LNCS 7511, 66–73, Springer Berlin / Heidelberg (2012).
- [2] Klinder, T., Wolz, R., Lorenz, C., Franz, A., and Ostermann, J., “Spine segmentation using articulated shape models,” in [*Medical Image Computing and Computer-Assisted Intervention MICCAI 2008*], LNCS 5241, 227–234, Springer Berlin / Heidelberg (2008).
- [3] Baiker, M., Milles, J., Dijkstra, J., Henning, T. D., Weber, A. W., Que, I., Kaijzel, E. L., Lwik, C. W., Reiber, J. H., and Lelieveldt, B. P., “Atlas-based whole-body segmentation of mice from low-contrast micro-ct data,” *Medical Image Analysis* 14(6), 723 – 737 (2010).
- [4] Steger, S., Kirschner, M., and Wesarg, S., “Articulated atlas for segmentation of the skeleton from head&neck ct datasets,” in [*Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*], 1256 –1259 (2012).
- [5] Bindernagel, M., Kainmueller, D., Seim, H., Lamecker, H., Zachow, S., and Hege, H.-C., “An articulated statistical shape model of the human knee,” *Bildverarbeitung für die Medizin 2011*, 59 – 63 (2011).
- [6] Zhan, Y., Maneesh, D., Harder, M., and Zhou, X., “Robust mr spine detection using hierarchical learning and local articulated model,” in [*Medical Image Computing and Computer-Assisted Intervention MICCAI 2012*], LNCS 7510, 141–148, Springer Berlin/Heidelberg (2012).
- [7] Alexa, M., “Linear combination of transformations,” in [*Proceedings of the 29th annual conference on Computer graphics and interactive techniques, SIGGRAPH '02*], 380–387, ACM, New York, NY, USA (2002).
- [8] Moghaddam, B. and Pentland, A., “Probabilistic visual learning for object representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19, 696–710 (1997).