Spine Detection and Labeling Using a Parts-Based Graphical Model

Stefan Schmidt[1,2], Jörg Kappes[2], Martin Bergtholdt[2], Vladimir Pekar[1],
Sebastian Dries[1], Daniel Bystrov[1], and Christoph Schnörr[2]

[1] Philips Research Europe - Hamburg, Germany
[2] CVGPR Group, University of Mannheim, Germany
`Stefan.Schmidt@ti.uni-mannheim.de`

Abstract. The detection and extraction of complex anatomical structures usually involves a trade-off between the complexity of local feature extraction and classification, and the complexity and performance of the subsequent structural inference from the viewpoint of combinatorial optimization. Concerning the latter, computationally efficient methods are of particular interest that return the globally-optimal structure.

We present an efficient method for part-based localization of anatomical structures which embeds contextual shape knowledge in a probabilistic graphical model. It allows for robust detection even when some of the part detections are missing. The application scenario for our statistical evaluation is spine detection and labeling in magnetic resonance images.

1 Introduction

Problem Description. We investigate a new method for automatically locating the human vertebral column (spine) and for simultaneously labeling the intervertebral disks in 3D T1-weighted magnetic resonance (MR) survey images of total spine. The method provides robust input for further automatic processing, e.g. initialization of vertebra models to segment individual vertebrae or extraction of regions of interest for subsequent image acquisition or processing.

The human spine typically consists of 24 vertebrae (7 cervical: C1-C7, 12 thoracic: T1-T12, 5 lumbar: L1-L5) with in-between situated intervertebral disks, aligned in a double-S shaped curve. In T1-weighted fast field-echo (FFE) images, the disks typically appear as bright structures, and the vertebrae themselves give virtually no signal and appear dark. We therefore use the disks as high level features (parts) for localizing the spine column and individual vertebrae. Integration of spatial context is essential, however, for coping with erroneous local detections and missing features, in order to label vertebrae anatomically.

There are several pathologies that may significantly affect the appearance of the spine such as fracture, neoplasm, deformity (e.g. scoliosis) and degeneration. Also the number of vertebrae may differ from 24, e.g. by lumbalization of the cranial sacral segment into an L6. Our objective is the design of a probabilistically sound model for reliably detecting and localizing the spine column, that copes with the complexity of the depicted anatomy and the variability in image quality. In this paper, we concentrate on the standard case of 24 vertebrae.

Approach. We use a probabilistic graphical model for representing both the appearance of local parts and the shape of the anatomy in terms of geometric relations between parts. Features for detecting parts are learned from a set of training data in manually marked image regions.

In the detection stage, a multi-class classifier is applied to detect potential locations for each part in a new image. In a subsequent step, the graphical model provides a contextual decision by fusing these data with the geometrical prior knowledge and inferring the globally optimal configuration of the parts.

2   State of the Art and Contribution

Related Work. At present, not many fully automated methods for locating the spine in MR images have been reported in the literature.

Peng et al. [1] detect the intervertebral disks in MR images in connection with segmenting the whole spine column. The detector used convolves a gray value template of a disk with the image, followed by searching for the disk centers and further local postprocessing. The method is not 3D but processes (sagittal) 2D slices only which may not intersect the whole spine. Furthermore, the approach does not model and employ contextual (non-local) prior knowledge. The study reports good results for 5 subjects, but the performance of the method for low-quality scout data, or when local disk detection fails, is unclear.

Weiss et al. [2] propose a semi-automatic algorithm for localizing the spine and for labeling the intervertebral disks. Operator assistance is used to provide a single seed point in the C2-3 disk. The procedure relies on intensity thresholds for detecting the remaining disks, rendering the method highly dependent on image quality and intensity correction in a preprocessing step. This also applies to the approach of Vrtovec et al. [3] who locate the spinal canal by searching for circular areas of homogeneous intensity in axial slices.

A well-known class of approaches that exploits geometrical prior knowledge for anatomy segmentation tasks is based on the Active Shape Model (ASM) [4]. ASMs are known for their dependency on a good search initialization, and on the reliability of local feature detection. Despite using a non-local shape model, the optimization strategy is merely local. If the search is misled by ambiguous landmarks, the optimization process cannot recover. In principle, the same criticism applies to the Active Appearance Model (AAM) [5,6], that extends the ASM by additionally representing the texture within the shape.

Recently, a combination of ASM and robust point-matching (RPM) has been proposed [7], that uses soft-assign for matching the model and image features. This work is closer to our approach regarding the ability to revise erroneous local decisions through contextual prior knowledge. The success of the method for finding the globally optimal configuration depends crucially on the annealing schedule that is used in the nonconvex energy minimization whereas the inference algorithm that combines local and global information in our approach guarantees global optimality.

Our model can be regarded as an instance of the object recognition framework suggested in [8] that describes the image as a deformable configuration of local

parts. While this work only considers tree-structured constellations for computational efficiency, we use a fully interconnected model which enables the efficient detection of missing parts. Other instances of this class of approaches include [9], [10] and [11] where the geometry is described by a multivariate normal distribution of part positions relative to a non-occluded reference part. Our model does not rely on any particular reference part and is invariant against translation and rotation. The graphical shape templates proposed by Amit [12] for anatomy detection are defined by potential functions on triples of parts, and require a *decomposable* graph structure enabling the application of a dynamic programming algorithm for second order Markov chains. The specific graph structure causes the approach to be sensitive to missing part detections, a limitation which is overcome in our formulation.

Contribution. The basic features of our approach are:

– The design of the part detectors is completely determined by learning from the data. It does *not* involve any model assumptions (e.g., template).
– Local information is fused with non-local geometrical prior information and results in a globally optimal configuration. Favorable consequences of this globally optimal inference are (i) that problems due to a poor initialization are obsolete, and (ii) that evaluations of the result only judge the model and cannot be misinterpreted as poor local minima.
– Missing parts are explicitly taken into account by the probabilistic graphical model. In case of failure of a local detector, we are still able to predict the most probable position of a missing part.
– Any further available information, e.g., identification of the up-most inter-vertebral space (C1-2) as anchor point, can be easily integrated and fully exploited during inference.

3   Methods

Graphical Model. Our probabilistic model represents the image appearance of single parts as well as the relative geometry of *pairs* of parts, in terms of vertices V and edges E of a graph $G = (V, E)$. Each vertex $s \in V$ indexes a random variable x_s that assigns to the part its unknown image location. Accordingly, $x = \{x_1, x_2, \ldots, x_{|V|}\}$ denotes a *configuration* of parts.

Given an image I, the objective is to localize the object of interest by determining the most probable configuration x that maximizes

$$P(x|I,\theta) = P^A(x|I)P^S(x|\theta) , \tag{1}$$

where P^A represents the appearance of parts and P^S captures shape information using a set of parameters θ. We model $P(x|I, \theta)$ by the Gibbs distribution

$$P(x|I,\theta) = \frac{1}{Z} \exp(-\mathcal{E}(x|I;\theta)) , \quad \mathcal{E}(x|I;\theta) = \sum_{s \in V} \psi_s(I, x_s) + \alpha \sum_{(s,t) \in E} \psi_{st}(x_s, x_t; \theta),$$
$$\tag{2}$$

whose components will be detailed below.

Part Appearance Potentials. The unary potential functions ψ_s in (2) are directly given by the output probability $P(x_s|I)$ provided by the classifier that is used to detect candidates for part s in an image, as described below. Note that we assumed mutual independence of *local* detections in (2).

Relative Geometry of Part Pairs. The potential functions ψ_{st} in (2) represent the relative geometry of pairs s, t of parts.

We evaluate two alternative models: The first version models the distance of the parts by the 1D Gaussian,

$$\psi_{st}(x_s, x_t) = \frac{(\|x_s - x_t\| - \mu_{st})^2}{2\sigma_{st}^2} \ . \tag{3}$$

This representation is invariant against joint translations and rotation of both parts. The second version uses instead of (3) a multivariate normal distribution to represent the part locations relative to each other:

$$\psi_{st}(x_s, x_t) = \frac{1}{2}(x_s - x_t - \mu_{st})^\top \Sigma_{st}^{-1}(x_s - x_t - \mu_{st}) \ . \tag{4}$$

This variant captures the geometry more accurately but is no longer rotational invariant. In applications where the absolute orientation of the structure of interest does not vary too much relative to the image frame (i.e. scanner axis), this second formulation is preferable.

In connection with spine detection, it is undesirable to allow disks to overlap. To account for this, we truncate the Gaussians by multiplying them with the indicator functions $I_{C_{st}}(x_s, x_t)$ that take the value 1 inside the sets $C_{st} := \{(x_s, x_t) \mid a_{st} \le \|x_s - x_t\| \le b_{st}\}$, where the parameters a_{st} and b_{st} are determined in relation to the minimum and maximum lengths in the training set. Hence, overlapping parts and pairs too far apart are assigned a pairwise probability of 0 and effectively are no longer considered together.

We point out that, while truncated Gaussian distributions are appropriate for the application studied in this paper, the model adopted from [13] also copes with more general, e.g. multimodal, local distributions.

Part Detection. Randomized classification trees allow for fast evaluation and can be used to detect points of interest [14]. Training such a classifier amounts to creating a set of decision trees and collecting the statistics of the training data under the trees' classifications. The branching tests at tree nodes are chosen at random from a set of very simple tests, each involving only few feature space dimensions.

For our application, we use sub-volumes of 15x15x15 voxels as local feature vectors. The tree tests are plain comparisons of two of their dimensions, translating to simple "brighter than resp. darker than" decisions for pairs of voxels in a neighborhood. The overall performance and robustness against noise results from the aggregation of the statistics over a large number of such tests, that are distributed over the ensemble of decision trees used. As only the ordering of intensity is taken into account, the resulting detector is insensitive to intensity distribution variation as commonly observed in MR images.

Local rotation and scale tolerance is achieved by augmenting the training set by resampled copies of the training images, which have been randomly transformed to vary within the expected range of orientation and scale changes.

Applying the classifier to an image yields a probability volume for each class (Fig. 1). After sorting by descending probability, we pick candidates for each class, starting with the most probable location, and add further candidates as long as a specified number of candidates and a given probability threshold are not exceeded. We impose a minimum spacing constraint between a class's candidate points to suppress non-maxima, and to restrict further the set of candidates.

Modeling Undetected Parts. Missing detections of individual parts are handled by introducing an artificial candidate for each node of the graphical model, the location and appearance of which is declared "hidden". As its position is not known, we would need to marginalize the associated potentials over the image domain, which is computationally infeasible. Instead, we approximate this marginalization by setting the hidden candidate's potentials to their expectation with respect to the training data plus a penalty for the miss:

$$\psi_s(\cdot) = -\log \beta \, \mathbb{E}_{x_s} \big[\exp(-\psi_s(x_s)) \big], \ \ \psi_{st}(\cdot, \cdot) = -\log \gamma \, \mathbb{E}_{x_s, x_t} \big[\exp(-\psi_{st}(x_s, x_t)) \big] \tag{5}$$

Here, $\mathbb{E}_{x_s}$ resp. $\mathbb{E}_{x_s, x_t}$ denote the expectation with respect to the ground truth training data, and the parameters β, γ are empirically determined from the training set such that these "hidden"-probabilities indeed incur a penalty for omitting parts in the majority of training cases.

Inference Algorithm. We use the graphical model as specified above to infer the most likely configuration x (position and overall pose) of the model in any novel image I. This maximum a posteriori (MAP) estimate

$$\hat{x} = \arg \max_{x \in \Omega} P(x|I, \theta) \tag{6}$$

is obtained by searching over all configurations consistent with detected part candidate points. This is done using the A^* algorithm [15] based on a tight upper-bound estimate for efficiently pruning the search space. This estimate is computed by exact inference on a tree-structured subgraph and guarantees not to miss the global optimum [13]. As described below, it turned out empirically that A^*-search terminates after reasonable computation time. Therefore, we preferred this method over alternative approaches to *approximate* graphical inference based on *non*convex optimization [16].

Missing Part Postprocessing. We can estimate the approximate location of parts detected as missing in a postprocessing step. This is done by greedily searching for these locations while keeping the already inferred parts fixed. For each missing part h, treated individually now, we predict its position by aggregating the evidence from all neighbors N_h of h that have been successfully identified. Specifically, using the learned distributions $x_t - x_s \sim \mathcal{N}(\mu_{st}, \Sigma_{st})$ and the identity $\mathcal{N}(a, A) \cdot \mathcal{N}(b, B) \propto \mathcal{N}(c, C)$ where $C = (A^{-1} + B^{-1})^{-1}$, $c = CA^{-1}a + CB^{-1}b$
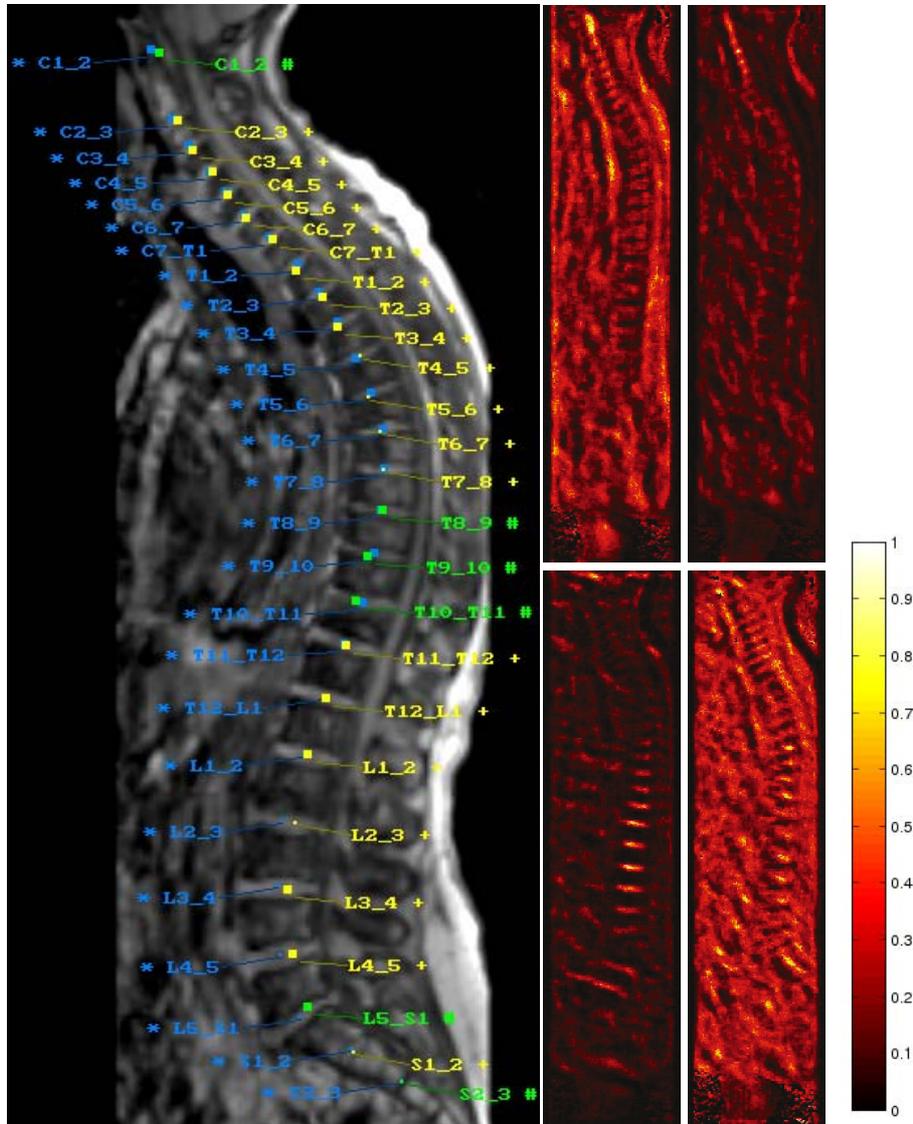
Fig. 1. Left: Labeling result. Yellow labels denote the graphical model's MAP estimate(+), green labels represent parts found by postprocessing(#), and blue labels show the ground truth annotation(*). Larger dots indicate positions located in the viewing plane and smaller dots positions next to it. **Right: Classifier probability maps** for four classes in a new image, produced by the randomized tree classifier that was trained on independent datasets. Clockwise, the corresponding classes are C1-2, C2-3, L1-2, and S1-2. The depicted slices from the probability volumes were normalized for visualization such that the value 1 represents the most likely pixels with respect to these classes. Note that particularly in case of the L1-2 class, the neighboring disks are highly similar in appearance. Thus, local classification cannot discriminate between them, but has to be complemented by "geometrical context" as studied in this paper.

for the multiplication of Gaussians, we compute the maximum at the mean of the product

$$\hat{x}_h = \max_{x_h} \prod_{s \in N_h} \mathcal{N}(x_s + \mu_{sh}, \Sigma_{sh}) \quad = \big(\sum_{s \in N_h} \Sigma_{sh}^{-1} \big)^{-1} \big(\sum_{s \in N_h} \Sigma_{sh}^{-1}(x_s + \mu_{sh}) \big). \quad (7)$$

Optionally, we incorporate the appearance term by multiplying the product of Gaussians by the classifier's probability map of the missing part, and then locate the maximum.

Scale Estimation and Invariance. The learned geometry representation exhibits a high variance if the training examples are not normalized in scale. Hence, tighter geometry constraints are obtained if both training and test data are on a common scale. For training images, this can be achieved by Procrustes analysis, for instance. But for the test images, the scale information is not readily available, because the correspondences have not been established prior to detection.

In order to compensate for global changes in scale in a new image, we extend our algorithm by a second run of the graphical model inference. After the first run, which is performed with the original geometry model without scale normalization, we obtain an estimate for the global scale correction parameter as follows: For each pair of detected parts, the ratio of their distance to the model's mean distance is entered into a weighted histogram $H\big(\frac{(\|x_s - x_t\|)}{\mu_{st}}; \frac{1}{\sigma_{st}}\big)$ with the weight derived from the model's standard deviation, so that less reliable pairs of parts have lesser influence on the estimate. After smoothing the histogram H using kernel density estimation with a Gaussian kernel, we use its mode as the global scale compensation parameter for this image.

A second inference run is then performed using the scale-compensated interpart geometry together with a geometry model learned from scale-normalized training examples. The scale estimates for the training data needed to build this latter model are extracted using the same histogram-based procedure, using ground truth pairs of parts instead of part detections.

4   Experimental Results

We evaluate the performance of our graphical model and its components for spine detection and localization, based on annotated ground truth data.

Data Sets. T1-weighted MR images were acquired on a Philips 1.5T Achieva scanner with a multi-station 3D FFE sequence. These 30 datasets each consist of two station scans ($224 \times 224 \times 180$ voxels of extent 1.96mm $\times$ 1.96mm $\times$ 1.5mm), which were combined into a single image and resampled to obtain an isotropic voxel size of 2mm edge length. The image quality of these MR acquisitions varied greatly. Artefacts and intensity distribution fluctuations due to field inhomogeneities as well as pathologies such as vertebra fractures were present in these data.

Training and Parameter Values. The tree classifiers were trained on 5000 positive examples for each of the 26 classes (intervertebral disks C1-2 to S2-3)
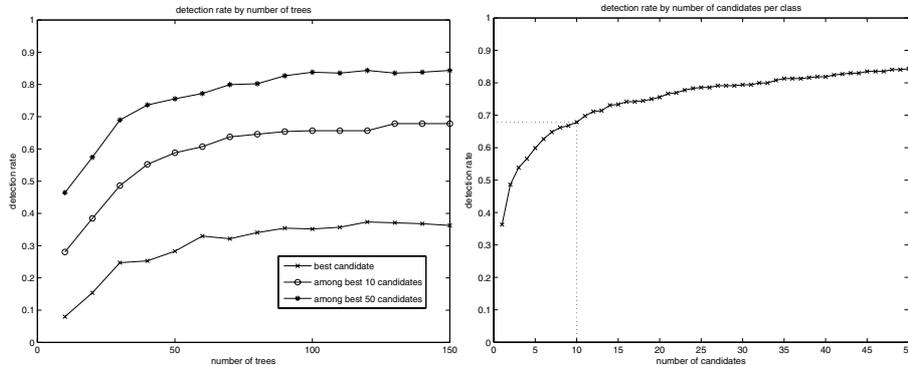
Fig. 2. Local detection performance. A part is detected if any of the best k candidates is within 14mm of the ground truth location. **Left:** Detection performance increases as more classification trees are used. Searching among the best 10 candidates for each class gives about 70% detection rate with 150 trees. **Right:** Producing more candidates increases the probability of detection, shown here for 150 trees. The line indicates our choice of 10 candidates used in subsequent experiments. Note that using a single local decision (the best candidate only) yields a poor detection rate of 36%.

and 50000 background examples (randomly picked patches from other image regions), all generated by resampling from a training set of 16 images. During resampling, the images were anisotropically scaled ($\pm 20\%$) and rotated ($\pm \pi/4$) about a randomly defined axis.

For scale detection, we always used model (4). It identifies sufficiently many parts such that the estimated scale correction factor differs from an estimate using ground truth by at most 4%. Further parameters used were $\beta = 0.01$ and $\gamma = 0.1$ in Eq. 5. In (2), we set $\alpha = \frac{2}{|V|}$. The parameters a_{st}, b_{st} for truncating the Gaussians (3) and (4) were determined by multiplying by 0.8 (1.15) the minimum (maximum) distance between parts observed in the training data.

Part Detector Performance. The part detectors are parametrized by the number of trees and the tree depth. We limited the maximum tree depth to 30 and pruned at branches with less than 10 training samples, so that most branches are actually much shorter. In order to set the number of trees to a reasonable value, we conducted a series of experiments for assessing the detection accuracy relative to annotated ground truth positions.

A detection is considered as "correct" if it occurs within a radius of 14mm from the ground truth annotation, resulting in good overall localization.[1] Fig. 2, left, shows the results.

Another important parameter is the number of candidate points we produce for each part class. This number determines the average running time of the inference algorithm based on the graphical model. Motivated by the experimental findings (cf. Figs. 2 right, and Fig. 3), we set the number of candidates per class

---

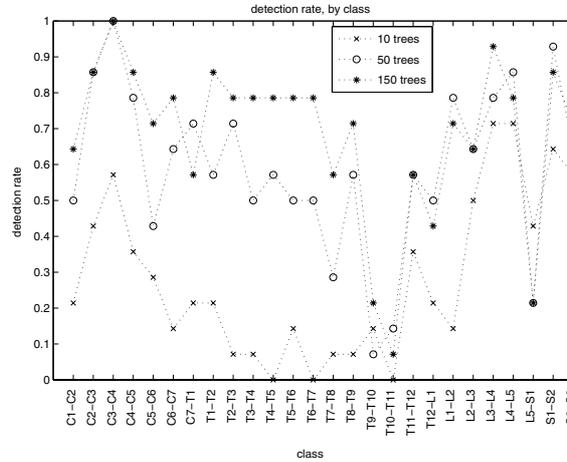[1] This value equals the minimum intervertebral disk distance in our data.

Fig. 3. Local detection performance for each part. While the detection probability increases with the numbers of trees, it varies significantly for different parts. The subsequent non-local inference using the graphical model is therefore essential in order to cope with parts that are more difficult to detect.
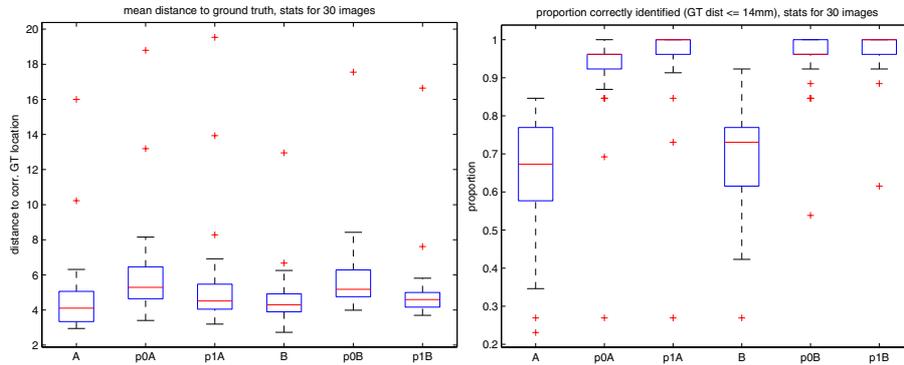

Fig. 4. Localization accuracy of the overall model variants. The box plots show the median (center line), the lower and upper quartiles (box), and the extent of the data (whiskers extending by at most 1.5 times the interquartile distance) excluding outliers, which are separately plotted as crosses. Symbol A indicates the use of orientation information in the geometry model (4); symbol B indicates the rotational invariant geometry model (3). Symbol p1 indicates missing-part postprocessing with appearance information, and p0 without. **Left:** Distances to ground truth locations, averaged over all classes, for each model variant A, p0A, p1A, B, p0B, p1B. **Right:** Proportion of correctly identified parts for each model variant.

to 10 in all subsequent experiments. An obvious modification is to let this number vary for different part classes. We did not investigate this option, however.

Overall Performance and Cross Validation. We determined the localization accuracy and robustness of the complete model by leave-one-out cross validation with 30 annotated images. In each of the 30 runs, 29 datasets were used for training the randomized tree classifier as well as either of the two geometry models (3) and (4), and the remaining dataset was used for testing the resulting model, utilizing the classifier together with graphical model inference, scale estimation, and missing part postprocessing.

Localization accuracy results are summarized in Fig. 4. Due to the restriction to interest points on the voxel grid as candidate locations, our model is not expected to yield sub-voxel accuracy. Nevertheless, Fig. 4, left, shows that most detected part locations are within a few millimeters of the ground truth position.

Using the geometry model with orientation (4) yielded an average ground truth distance of 6.2mm (lower quartile 4.6mm, upper quartile 6.5mm) and a part detection rate of 0.91. The rotational invariant model (3) resulted in an average ground truth distance of 5.8mm (lower quartile 4.8mm, upper quartile 6.3mm) and a part detection rate of 0.95. Including appearance in the search for missing parts improved this to 5.5mm and 5.1mm (part detection rates 0.94 and 0.97), respectively. Fig. 5 shows the localization errors for each class separately.
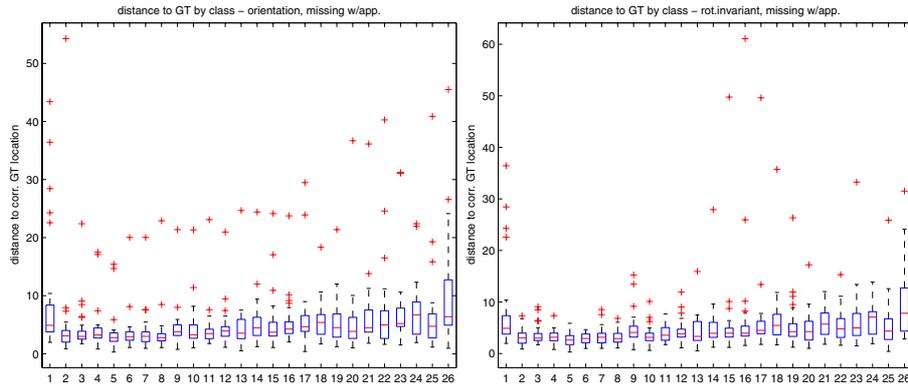

Fig. 5. Localization accuracy for each part class. The class numbers correspond to intervertebral disks C1-2 (class 1) to S2-3 (class 26). **Left:** Model using a geometry model with orientation, and missing part postprocessing with appearance information. **Right:** Rotational invariant geometry model, and missing part postprocessing with appearance information.

Using the rotational invariant geometry model results in more accurate localizations with less outliers. An example result of the overall model fit, i.e. the most probable configuration x, is visualized in Fig. 1.

Generalization to Novel Images. We tested the algorithm also on a set of 37 new images, that were not used for training nor for optimizing the parameters. Using the rotational invariant geometry model and appearance information in

the search for missing parts, yielded an average ground truth distance of 7.8mm (lower quartile 5.3mm, upper quartile 9.5mm). This indicates that our models generalizes well to novel data.

Semi-automated Incorporation of Expert Knowledge. We investigate if, and to what extent the incorporation of expert knowledge (as used in semi-automatic algorithms by "clicking" on relevant anatomical structures) improves the localization accuracy. We repeated the same series of experiments as described above, but replaced the candidates for C1-2 with the expert's annotation (ground truth) and assigned to them an appearance probability of 1, while disallowing to declare it as missing. The results improved slightly: The model including orientation yielded an average ground truth distance of 5.8mm (lower quartile 3.9mm, upper quartile 5.5mm) and a part detection rate of 0.94. The rotational invariant model resulted in an average ground truth distance of 5.7mm (lower quartile 4.2mm, upper quartile 5.3mm) and also a part detection rate of 0.94.

5   Conclusion and Further Work

We presented a probabilistic graphical approach to the localization of spinal structures, by fusing local part detection with non-local geometrical context. The latter enables to revise local detections in case of ambiguous image data or missing features. The inference algorithm returns always the globally optimal configuration of parts conditioned on the observed image data, and does therefore not suffer from initialization problems.

While our approach enables robust localization and appears to generalize well to novel image data, it fails in cases of severe fractures that lead to geometrical configurations not covered by the training data. Enlarging the latter will most likely fix this problem.

Surprisingly, regarding computational complexity, the limiting component of our current implementation are the local part detectors. For a full volume, they take several minutes computation time. The subsequent combinatorial search for the optimal configuration, on the other hand, runs below 1 second on the average, despite the size of 10^{26} possible configurations. This proves empirically the tightness of our upper-bound estimate used within the A^*-algorithm.

Our approach can be extended in several ways. For example, subvoxel localization can be achieved by fitting more elaborate part models in a post-processing step. Reduction of the processing time mainly depends on a more sophisticated implementation of the part detectors. We also would like to point out, that our model is sufficiently flexible to be adapted to the localization of other anatomical structures.

An important topic of our future work concerns to handle also cases where the total number of vertebrae differs from the standard of 24. The latter can be achieved by extending the graphical model to include a latent unknown variable that represents the major cases of anatomy deviation and has to be inferred as part of the overall process.

References

1. Peng, Z., Zhong, J., Wee, W., Lee, J.h.: Automated vertebra detection and segmentation from the whole spine MR images. In: Proc. Engineering in Medicine and Biology, IEEE-EMBS 2005, pp. 2527–2530 (2005)
2. Weiss, K.L., Storrs, J.M., Banto, R.B.: Automated spine survey iterative scan technique. Radiology 239(1), 255–262 (2006)
3. Vrtovec, T., Ourselin, S., Gomes, L., Likar, B., Pernus, F.: Generation of curved planar reformations from magnetic resonance images of the spine. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 135–143. Springer, Heidelberg (2006)
4. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models – their training and application. Comp. Vision Image Underst. 61(1), 38–59 (1995)
5. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Burkhardt, H., Neumann, B. (eds.) ECCV 1998. LNCS, vol. 1407, pp. 484–498. Springer, Heidelberg (1998)
6. Matthews, I., Baker, S.: Active appearance models revisited. International Journal of Computer Vision 60, 135–164 (2004)
7. Abi-Nahed, J., Jolly, M.P., Yang, G.Z.: Robust active shape models: A robust, generic and simple automatic segmentation tool. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4191, pp. 1–8. Springer, Heidelberg (2006)
8. Felzenszwalb, P.F., Huttenlocher, D.P.: Pictorial structures for object recognition. International Journal of Computer Vision 61, 55–79 (2005)
9. Pham, T.V., Smeulders, A.W.M.: Object recognition with uncertain geometry and uncertain part detection. Computer Vision and Image Understanding 99(2), 241–258 (2005)
10. Weber, M., Welling, M., Perona, P.: Unsupervised learning of models for recognition. In: ECCV (1) 2000, 18–32 (2000)
11. Fergus, R., Perona, P., Zisserman, A.: Object class recognition by unsupervised scale-invariant learning. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR). vol. 2, p. 264 (2003)
12. Amit, Y.: Graphical shape templates for automatic anatomy detection with applications to MRI brain scans. IEEE Trans. Med. Imaging 16(1), 28–40 (1997)
13. Bergtholdt, M., Kappes, J.H., Schnörr, C.: Learning of graphical models and efficient inference for object class recognition. In: Franke, K., Müller, K.-R., Nickolay, B., Schäfer, R. (eds.) Pattern Recognition. LNCS, vol. 4174, pp. 273–283. Springer, Heidelberg (2006)
14. Lepetit, V., Fua, P.: Keypoint recognition using randomized trees. IEEE Trans. Pattern Anal. Mach. Intell. 28(9), 1465–1479 (2006)
15. Minoux, M.: Mathematical programming: theory and algorithms. Wiley, Chichester (1986)
16. Pakzad, P., Anantharam, V.: Estimation and marginalization using Kikuchi approximation methods. Neural Computation 17(8), 1836–1873 (2005)