Contour-Driven Atlas-Based Segmentation

HHS Public Access

Author manuscript

IEEE Trans Med Imaging. Author manuscript; available in PMC 2016 December 01.

Published in final edited form as:

IEEE Trans Med Imaging. 2015 December ; 34(12): 2492–2505. doi:10.1109/TMI.2015.2442753.

Contour-Driven Atlas-Based Segmentation

Christian Wachinger,
Computer Science and Artificial Intelligence Lab (CSAIL) at the Massachusetts Institute of Technology (MIT).

Department of Neurology, Massachusetts General Hospital, Harvard Medical School.

Karl Fritscher,
Department of Radiation Oncology, Massachusetts General Hospital, Harvard Medical School.

Greg Sharp, and
Department of Radiation Oncology, Massachusetts General Hospital, Harvard Medical School.

Polina Golland
Computer Science and Artificial Intelligence Lab (CSAIL) at the Massachusetts Institute of Technology (MIT).

Abstract

We propose new methods for automatic segmentation of images based on an atlas of manually labeled scans and contours in the image. First, we introduce a Bayesian framework for creating initial label maps from manually annotated training images. Within this framework, we model various registration- and patch-based segmentation techniques by changing the deformation field prior. Second, we perform contour-driven regression on the created label maps to refine the segmentation. Image contours and image parcellations give rise to non-stationary kernel functions that model the relationship between image locations. Setting the kernel to the covariance function in a Gaussian process establishes a distribution over label maps supported by image structures. Maximum a posteriori estimation of the distribution over label maps conditioned on the outcome of the atlas-based segmentation yields the refined segmentation. We evaluate the segmentation in two clinical applications: the segmentation of parotid glands in head and neck CT scans and the segmentation of the left atrium in cardiac MR angiography images.

Keywords

Image segmentation; atlas; patch; spectral clustering; Gaussian process; left atrium; parotid glands

I. Introduction

Atlas-based segmentation extracts information from image collections with manual labels to facilitate the automatic segmentation of new images. Methods that use atlas information can be broadly classified into two groups. The first group employs registration to align training

images with the test image and to propagate the training labels. The training images are either summarized in a single probabilistic atlas that is registered to the test image [1]–[3] or the training images are directly registered to the test image [4]–[8]. Multi-atlas approaches tend to outperform single-atlas segmentation strategies when the anatomical variability is too large to be represented by the mean statistics [4]–[6].

The second group of atlas-based techniques infers label maps by comparing local image regions. For each voxel in the test image, the surrounding patch is compared to patches in the training images. With the rationale that similar patches tend to share the segmentation label, weighted voting among the most similar patches promises to produce accurate segmentations [9], [10]. Alternatively, classifiers can be trained on the manually labeled images [11]. Classification of patches in the test image yields the segmentation. Patch- and registration-based approaches are also used in combination to improve the segmentation [12].

High anatomical variability presents a serious challenge for atlas-based segmentation. Registration-based approaches may fail to warp structures that vary significantly in shape due to regularization constraints. Such inaccuracies mainly cause segmentation errors at organ boundaries. Similarly, patch-based methods experience difficulties when labeling regions close to the boundaries. Fig. 1 illustrates this problem for a patch-based segmentation of the left parotid gland. To further investigate the source of errors, we examine patches in the atlas that are the most similar to the one example patch in the image. According to the manual labeling, the selected patch belongs to the left parotid gland. However, all of the closest patches in the repository vote for background, yielding a wrong result. Since all of the nearest neighbor patches have a very similar appearance overall, the problem is not in the retrieval but is inherent in patch-based segmentation; patches that are very similar may not share the same label due to slight variations. This is most problematic close to organ boundaries, where this can cause segmentation errors.

The contributions of this article include a Bayesian framework for creating label maps in atlas-based segmentation and a contour-driven refinement of initial label maps with Gaussian process regression. Bayesian inference with the Laplace approximation enables us to derive location-, region-, and image-wise voting schemes for inferring the final label maps. The combination of these schemes with different distributions for the label and image likelihood, including voxel- and patch-wise image similarities enables modeling commonly used registration- and patch-based segmentation methods. For the refinement of initial label maps, we extract image contours and use them to learn correlations across image locations. The label fusion procedure in atlas-based segmentation generally relies on intensity differences between images and votes independently for each location.[1] We construct kernel functions that model the interaction of locations based on image contours. The kernel acts as a covariance function in a Gaussian process, which defines a distribution over image-specific label maps. Conditioning this distribution on the initial label maps leads to segmentation results that are consistent with image contours while also accommodating the

[1]The label fusion procedure assumes independence. Regularization constraints in registration may introduce dependencies between voxels.

vote of the atlas. We propose two different kernel functions. First, we use intervening contours [13] to define the kernel based on the boundaries in the image. Second, we define a kernel based on the parcellation of the image, yielding a voting scheme on superpixels.

As a motivating example, we present the result of intensity-weighted label fusion [8] for the left atrium of the heart in Fig. 2(a). We observe an undersegmentation of the left atrium. The segmentation does not follow the image contours. Fig. 2(b) shows a refinement of the segmentation by considering image contours, which faces challenges in identifying the correct boundaries especially for smooth intensity transitions. As a result, certain parts are oversegmented while others, e.g., the veins, are cut off. We propose to refine the label map by combining contours extracted from the image and the initial label map. This approach leads to the most accurate segmentation of the left atrium in Fig. 2(c).

A. Clinical Applications

The contour-driven refinement helps to delineate structures of high variability, which are challenging to segment with atlas-based techniques. We present results for two clinical applications, radiation therapy and cardiac ablations, that necessitate segmentation of such structures with strong variations. The first application requires the segmentation of parotid glands of patients undergoing radiation therapy. Radiation therapy planning aims to maximize the dose in the target region while minimizing the radiation dose in the surrounding tissue. Intensity modulated radiation therapy enables more effective administration of the radiation dose to reduce the damage to healthy tissues. During the planning phase, experts delineate most critical structures, also called organs at risk, to ensure low radiation in these regions. The parotid glands are organs at risk for head and neck cancer treatment because they are the most important salivary glands. Irradiation of the parotid glands can lead to xerostomia, resulting in difficulties for mastication, deglutition, and speech of the patients. The low soft-tissue contrast in CT images and the high anatomical variability of the parotid glands make the automatic segmentation challenging.

The second clinical application involves ablation of ectopic foci, which is the common treatment for patients with atrial fibrillation [14]. In atrial fibrillation, the left atrium of the heart no longer pumps blood into the left ventricle efficiently because it quivers in an abnormal rhythm. Atrial fibrillation accounts for 15% of all strokes [15]. Accurate segmentation of the left atrium and its pulmonary venous drainages in contrast-enhanced magnetic resonance angiography (MRA) images is essential for planning and evaluating ablation procedures. The segmentation of the left atrium is challenging because it exhibits strong variations in the shape of the cavity and in the number and location of pulmonary veins [16].

B. Related Work

Our work builds on atlas-based segmentation. We review atlas-based segmentation in greater detail in Section III, when we introduce the Bayesian framework for inference of labels maps. Here, we focus on studies that are relevant to the label refinement. In [17], [18], graph cuts was used to refine the segmentation, where we use Gaussian process regression to infer the refinement. Regression has been previously used to estimate correlations of errors

for atlas-based segmentation [19]. In later work [20], the correlation between atlases is estimated with a linear appearance model. Instead of refining the segmentation through the integration of image information in a post-processing step, we consider image and label information jointly in the graph Laplacian. Moreover, we work on image contours that potentially have advantages over comparing intensities [13]. Label refinement based on the hypothesis of atlas-based under-segmentation was proposed in [21]. Our probabilistic approach uses Gaussian processes, which arise in numerous fields of machine learning [22] and computer vision, e.g., image denoising [23], interpolation [24] and registration [25]–[27]. In [28], Gaussian processes were applied for image segmentation of natural images. In contrast to our work, the identity covariance function was used, samples from the Gaussian process were thresholded, and no atlas information was available.

When focusing on our clinical applications, atlas-based segmentation of parotid glands with deformable registration was demonstrated in [29], [30]. In [31], the atlas images were used for training an active shape model of parotid glands. The refinement of head and neck segmentations based on classification with features was proposed in [32]. For the left atrium, a segmentation method was proposed by extracting the blood pool with intensity thresholding [33]; this method is sensitive to intensity variations. Intensity-weighted label fusion achieved accurate results for the segmentation of the left atrium in [16]. We treat this label fusion technique as a baseline for comparison. A preliminary version of this work was presented at conferences [34], [35].

II. Problem Statement

Given a novel image I, we aim to infer its segmentation S based on an atlas that contains training images $\mathscr{I}=\{I_1,\ldots,I_n\}$ with segmentations $\mathscr{S}=\{S_1,\ldots,S_n\}$. A probabilistic label map $\mathscr{L}=\left\{L^1,\ldots,L^\eta\right\}$ specifies the likelihood for each label $l \in \{1, ..., \eta\}$

$$L^l(x)=p(S(x)=l|I,\mathscr{I},\mathscr{S}). \quad (1)$$

We obtain the estimated segmentation $\hat{S}(x)$ by choosing the label with highest probability for each voxel x on the image grid Ω

$$\hat{S}(x)=arg\max_l L^l(x). \quad (2)$$

A perfect label map assigns probability one to the correct label for each location. In Section III, we discuss the construction of the initial label map L_o based on the atlas. In Section IV, we define a prior over label maps $p(L)$ based on contours in the test image I to improve the initial label map. Maximization of the posterior yields the refined label map

$$\hat{L}^l=arg\max_L p\left(L|L_o^l\right), \quad (3)$$

which is used in Eq. (2) to get a refined segmentation.

III. Atlas-Based Segmentation

In this section, we propose a Bayesian framework for atlas-based segmentation, which enables a unified review of existing segmentation approaches. We start with multi-atlas approaches [4]–[8], where each training image in the atlas $\mathscr{I}$ is aligned to the test image I, yielding deformation fields $\Phi = \{\varphi_1, ..., \varphi_n\}$. We condition the segmentation probability $p(S|I, \mathscr{I}, \mathscr{S})$ on the deformation fields, leading to $p(S|\Phi, I, \mathscr{I}, \mathscr{S})$. The deformation fields are commonly estimated with registration tools that seek the mode

$$\hat{\Phi} = arg\,\max_{\Phi} p(\Phi|I, \mathscr{I}). \quad (4)$$

The likelihood of the segmentation S is approximated by inserting the mode of the deformation fields $\hat{\Phi}$

$$p(S|I, \mathscr{I}, \mathscr{S}) \approx p\left(S|\hat{\Phi}, I, \mathscr{I}, \mathscr{S}\right). \quad (5)$$

The drawback of working with the mode is that it does not incorporate the uncertainty in registration. For nonlinear registration, we potentially estimate a large number of parameters, yielding a high uncertainty. Bayesian inference circumvents this loss by marginalizing the latent deformation fields

$$p(S|I, \mathscr{I}, \mathscr{S}) = \int_{\Phi} p(S, \Phi|I, \mathscr{I}, \mathscr{S})\, d\Phi \quad (6)$$

$$= \int_{\Phi} p(S|\Phi, I, \mathscr{I}, \mathscr{S}) \cdot p(\Phi|I, \mathscr{I})\, d\Phi. \quad (7)$$

Working with the mode corresponds to approximating the prior on the deformation field with a delta function

$$p(\Phi|I, \mathscr{I}) = \delta\left(\hat{\Phi}, \Phi\right). \quad (8)$$

Since the delta function selects a single deformation field, only a single **location** in the training image affects the voting. This approximation is valid if $p(\Phi|I, \mathscr{I})$ has a sharp peak at $\hat{\Phi}$, which is questionable for deformation fields because there can be more unknowns than observations.

In general, computing the integral in Eq. (7) is intractable. In [36], Markov chain Monte Carlo sampling of deformation fields was presented to improve segmentation results. We propose to apply the *Laplace approximation*, which uses a normal distribution centered at the mode as the prior distribution over deformation fields $p(\Phi|I, \mathscr{I}) = \mathcal{N}\left(\Phi|\hat{\Phi}, \tau^2\mathbf{I}\right)$, yielding

$$p(S|I, \mathscr{I}, \mathscr{S}) = \int_{\Phi} p(S|\Phi, I, \mathscr{I}, \mathscr{S}) \cdot \mathcal{N}\left(\Phi|\hat{\Phi}, \tau^2\mathbf{I}\right) d\Phi, \quad (9)$$

where τ^2 is the variance and **I** is the identity matrix. This approximation weighs segmentations that are proposed by deformation fields Φ according to the distance to the mode $\hat{\Phi}$. Deformations that are closer to the mode have larger influence on the segmentation.

The Laplace approximation reduces segmentation errors that are caused by incorrect registrations. The distribution over similar transformations translates to a distribution over similar locations in the case of deformation fields. Instead of comparing location *x* to $\hat{\phi}_i(x)$, the surrounding **region** $N_{\hat{\phi}_i(x)}$ in the training image is also considered, as illustrated in Fig. 3. $N_{\hat{\phi}_i(x)}$ denotes a local neighborhood centered at the location $\hat{\phi}_i(x)$. Although the normal distribution has in theory infinite support, most of the mass of the distribution lies within a neighborhood dependent on the variance. The variance τ^2 and the size of neighborhood N_x are set to reflect the expected accuracy of the registration. It is possible to estimate them by aligning the training data and using the manual labels to quantify the registration uncertainty. If local estimates of registration accuracy exist, it is also possible to work with a locally varying variance $\tau^2(x)$ and neighborhood size.

Considering the asymptotics of the prior over deformation fields, for $\tau^2 \to 0$ we obtain the delta function and approximation with the mode in Eq. (5). For $\tau^2 \to \infty$ or the approximation with a uniform distribution $p(\Phi|I,\mathscr{I}) \propto 1$, we arrive at non-local methods. This flat prior on deformation fields assigns the same probability to mapping to any location in the training image. The search region is therefore no longer limited to a local neighborhood around the mode, but covers the **entire image** domain of the training image Ω in Fig. 3. Such non-local methods do not require the alignment of images.

A. Label and Image Likelihood

We follow the derivation of locally weighted voting in [8] by splitting the segmentation probability $p(S|I,\mathscr{I},\mathscr{S})$ into the label and image likelihood term. The independence of spatial locations in the test image is part of the derivation. In contrast to [8], our derivation includes the integration over deformation fields

$$p(S|I,\mathscr{I},\mathscr{S}) \propto \prod_{x\in\Omega} p(S(x), I(x)|\mathscr{I},\mathscr{S}) \quad (10)$$

$$\propto \prod_{x\in\Omega}\sum_{i=1}^{n} p(S(x)|S_i)\cdot p(I(x)|I_i) \quad (11)$$

$$\propto \prod_{x\in\Omega}\sum_{i=1}^{n}\int_{\Phi} p(S(x)|S_i,\Phi)\, p(I(x)|I_i,\Phi)\, p(\Phi|I,\mathscr{I})\, d\Phi. \quad (12)$$

We list the corresponding voting methods for different approximations of the deformation prior $p(\Phi|I,\mathscr{I})$ in Tab. I. The integration over discrete image grids yields the summation. For region-wise voting, this summation is limited to a region $N_{\hat{\phi}_i(x)}$, as discussed earlier.

Below, we state possible models for label likelihood $p(S(x)|S_i)$ and image likelihood $p(I(x)|I_i)$.

1) Label Likelihood—The label likelihood $p(S(x) = l|S_i(\varphi_i(x)))$ expresses the conditional probability that location x belongs to organ l, given the segmentation of the i-th training image at the corresponding location $\varphi_i(x)$. The **discrete model** for the label likelihood is

$$p\left(S\left(x\right)=l|S_i\left(\phi_i\left(x\right)\right)\right)=\begin{cases}1 & \text{if } S_i\left(\phi_i\left(x\right)\right)=l\\ 0 & \text{otherwise.}\end{cases} \quad (13)$$

An alternative model uses the logarithm of odds (LogOdds), which is based on the signed distance function [37]. D_i^l denotes the signed distance transform of label l of the segmentation S_i, which is positive inside the organ. The label probability for the **LogOdds model** is

$$p\left(S\left(x\right)=l|S_i\left(\phi_i\left(x\right)\right)\right)=\frac{1}{Z_i\left(x\right)}exp\left(\kappa D_i^l\left(\phi_i\left(x\right)\right)\right), \quad (14)$$

with the slope constant $\kappa > 0$ and the partition function $Z_i\left(x\right)=\sum_{l=1}^{\eta}exp\left(\kappa D_i^l\left(\phi_i\left(x\right)\right)\right)$.

2) Image Likelihood—Fixing the image likelihood across all images to a constant, $p(I(x)|I_i) \propto c$, ignores the image likelihood and yields majority voting. Fixing the image likelihood within a single image to a constant, $p(I(x)|I_i = c_i$, models atlas selection [38], [39]. Instead of globally expressing the similarity between test and training image, a local estimation enables better differentiation across image regions. In a first approach, we consider a **univariate** Gaussian distribution over single intensities

$$p\left(I\left(x\right)|I_i\right)\propto exp\left(-\frac{\left(I\left(x\right)-I_i\left(\phi_i\left(x\right)\right)\right)^2}{2\nu^2}\right) \quad (15)$$

with variance ν^2. Multivariate models promise to be more discriminative than a univariate comparison. We consider a local neighborhood M_x centered at location x and the corresponding patch $P(x) = I(M_x)$. Similarly, a patch in the i-th training image is $P_i(x) = I_i(M_{\varphi i(x)})$. For volumetric images, patches correspond to sub-volumes. We reshape each patch into a vector with linear indexing, yielding the **multivariate** Gaussian distribution over patches

$$p\left(I\left(x\right)|I_i\right)\propto exp\left(-\frac{1}{2}\left[P\left(x\right)-P_i\left(x\right)\right]^T\Psi^{-1}\left[P\left(x\right)-P_i\left(x\right)\right]\right) \quad (16)$$

with covariance matrix Ψ. The covariance matrix accepts patch specific distances. For instance, we can emphasize the center locations of the patch by setting the diagonal entries accordingly. The special case of a spherical Gaussian distribution [9], [10]

$$p\left(I\left(x\right)|I_i\right) \propto exp\left(-\frac{\| P\left(x\right) - P_i\left(x\right) \|_2^2}{2\nu^2}\right) \quad (17)$$

corresponds to the common L_2 distance. Patches are more discriminative than voxels, which is helpful for voting schemes that pool the information from local regions $N_{\varphi_i(x)}$ or the entire image grid Ω.

In a further extension, the patch $P(x)$ in the multivariate distribution can be replaced by a feature vector $F(x)$ that is extracted from the image. Examples of such features include Gabor wavelet features [40], [41], local binary patterns [42], Haar-like features [43], and histograms of oriented gradients [44]. Results of an empirical study that compared several features for non-local means segmentation [45] indicates good results for features that compute image gradients and local entropies [46], [47].

B. Atlas voting methods

In the last sections, we presented three voting strategies (location-, region-, and image-wise), two label likelihood models (discrete and LogOdds), and three image likelihood models (constant, intensity-, and patch-wise). In principle, it is possible to combine all of these variants arbitrarily, although some combinations may not be very sensible. We use this framework to describe several existing voting schemes, grouped by local and non-local approaches.

1) Local—Local methods require the images to be registrated, yielding deformation fields $\hat{\Phi}$. We obtain standard **majority voting** [4], [5] by location-wise voting with a constant image likelihood and discrete label probabilities. Majority voting with LogOdds seems also plausible. The selection of location-wise voting with the univariate image likelihood and the LogOdds label probabilities leads to label fusion with **intensity-weighted voting** [8], also referred to as locally-weighted voting.

An extension of intensity-weighted voting is **patch-weighted voting** [12], [19], [48], which replaces the univariate intensity likelihood with the multivariate patch likelihood and the spherical Gaussian model in Eq. (17). In addition, we consider the extension due to the Laplace approximation, which leads to a region-wise voting. We combine the region-wise voting with two different image likelihoods, the univariate intensity and multivariate patch model. We refer to them as **region intensity voting** and **region patch voting**. A similar voting scheme was described in [49], where the spatial weighting is heuristically integrated in the label likelihood. In contrast, it is treated as an independent factor in our derivation based on the Laplace approximation.

2) Non-local—For non-local approaches, we work with **non-local means** (NLM) segmentation [9], [10] and the application of a **random forest classifier** [11], [50]. Both approaches do not depend on the registration of the images and work with the discrete label likelihood.

In NLM, the initial label map is a weighted sum [9]

$$L_o^l(x) = \frac{\sum_{i=1}^{n} \sum_{y \in \Omega} w_{i,y,x} \cdot p(S(x) = l | S_i(y))}{\sum_{i=1}^{n} \sum_{y \in \Omega} w_{i,y,x}} \quad (18)$$

with

$$w_{i,y,x} = exp\left(- \frac{\| P_i(y) - P(x) \|_2^2}{min_{i',y'} \| P_{i'}(y') - P(x) \|_2} \right). \quad (19)$$

In our framework, NLM corresponds to the image-wise voting scheme with a patch image likelihood. While NLM is a non-local method, the search window in [9], [10] is restricted to small regions of 9^3 to 15^3 voxels, which assumes a coarse correspondence between test and training image. This spatial restriction makes the practical implementation of NLM very similar to the region patch voting. Explicit modeling of the spatial distance to the center location, $\mathcal{N}\left(y | \hat{\phi}_i(x), \tau^2 \mathbf{I}\right)$, is the main difference between our approach and NLM. Restricting the search to small regions of interest can be interpreted as implicit location information. Results indicate that increasing the search window size in NLM deteriorates the segmentation accuracy and necessitates explicit integration of location information [45].

In a second non-local approach, a random forest (RF) classifier [11], [50] is trained to predict the initial label map

$$L_o^l(x) = p(S(x) = l | P(x)) = \frac{1}{T} \sum_{t=1}^{T} p_t(S(x) = l | P(x)), \quad (20)$$

where $p_t(S(x)|P(x))$ is the posterior distribution of the t-th tree. Similar to NLM segmentation, the RF classifier considers all patches in the training images, but a training stage is required.

IV. Contour-Driven Label Refinement

In this section, we present a new approach for the contour-driven refinement of the initial label map L_o created in Section III. We assume additive, heteroskedastic Gaussian noise in the initial label maps

$$L_o(x) = \overline{L}(x) + \varepsilon(x), \ \varepsilon(x) \sim \mathcal{N}\left(0, \sigma^2(x)\right). \quad (21)$$

with the underlying true label map $\bar{L}$. Note that label maps are numerical variables that indicate the probability of an organ being present at a specific location. Setting the diagonal entries of the covariance matrix $C_{xx} = \sigma^2(x)$ yields the multivariate distribution $\varepsilon \sim \mathcal{N}(0, C)$. We estimate the variance $\sigma^2(x)$ from L_o by calculating the entropy across label maps, $\sigma^2(x) = -\sum_{l=1}^{\eta} L_o^l(x) \cdot log\, L_o^l(x)$.

A. Contour-Driven Regression with Gaussian Processes

In contrast to most label fusion methods that make decisions at each voxel independently[2] and do not consider contour information, we define a distribution that models the relationship between locations by exploiting image contours. Stochastic processes offer a versatile framework to model interactions between a possibly infinite number of random variables. We treat label maps as realizations of a Gaussian process with mean *m* and covariance *k*, $L \sim \mathcal{GP}(m, k)$. Gaussian processes are entirely characterized by the mean and covariance functions; every finite subset is distributed according to a multivariate Gaussian distribution [22].

To obtain the posterior distribution over label maps, we condition the distribution of label maps $p(L)$ on the initial label map L_o implied by the atlas

$$p(L|L_o) \sim \mathcal{N}(\mu, \Sigma) \quad (22)$$

with mean and covariance

$$\mu = m + K \cdot [K + C]^{-1} \cdot (L_o - m), \quad (23)$$

$$\Sigma = K - K \cdot [K + C]^{-1} \cdot K. \quad (24)$$

The mean vector **m** and the kernel matrix *K* are constructed from the mean function *m* and kernel function *k*, respectively. We use the Cholesky factorization to perform the matrix inversion. The maximum a posteriori label map coincides with the mean label map for Gaussian distributions,

$$\hat{L} = arg\max_L p(L|L_o) = arg\max_L \mathcal{N}(L; \mu, \Sigma) = \mu. \quad (25)$$

Performing this estimation for all label maps yields the segmentation

$$\hat{S}(x) = arg\max_l \mu^l(x). \quad (26)$$

The mean function *m* causes a constant additive shift of all label maps $\boldsymbol{\mu}^l$ and therefore does not influence the segmentation result $\hat{S}$, motivating the choice of $m = 0$. Fig. 4 illustrates the key steps of the segmentation process.

B. Kernel function

The kernel function *k* gives rise to a covariance function that models the relationship between locations in the label map. We propose two different constructions of the kernel function based on image contours and based on image parcellations into superpixels. The computation of image contours and parcellations is described in Section IV-C. Both definitions yield non-stationary covariance functions.

[2]The models presented in Section III combine the information from local neighborhoods in the training image but treat each location in the test image independently, cf. Eq. (10).

First, we employ image contours Γ to compute entries of k for in-plane points x and x', based on the concept of intervening contours [51], by identifying the maximum response along the line $\overline{xx'}$

$$k\left(x,x'\right)=exp\left(-\underset{y\in\overline{xx'}}{max}\left\{\Gamma\left(y\right)\right\}\right). \quad (27)$$

We consider locations within a radius r, where we set r = 10 pixels. A small radius increases the sparsity of the kernel matrix and therefore decreases the computational complexity. A radius of 10 pixels is a compromise between computational efficiency and modeling correlations with a large enough spatial extent. High values are assigned to pairs of points that are not separated by a contour. These points are subsequently encouraged to share the same label. Fig. 4 shows samples drawn from the distribution $p(L)$, where we overlay the manual segmentation for reference. We observe that the distribution promotes label maps that follow image structures. In this example, labels are propagated to the thin ends of the left parotid gland, which improves the segmentation in comparison to the initial labeling.

Second, we employ a parcellation of the image with superpixels Π_j, which are an exact cover of the image domain $\Omega=\cup_j\Pi_j$. We define the kernel for in-plane points x and x' with superpixels Π_j

$$k\left(x,x'\right)=\begin{cases}1/|\Pi_j| & \text{if } \exists j{:}x,x'\in\Pi_j\\ 0 & \text{otherwise.}\end{cases} \quad (28)$$

After permutation, the resulting kernel matrix K has a block diagonal structure, where each entry in the block is the reciprocal of the corresponding superpixel size. This matrix acts as a projection operator with the property that $KK = K$, causing the Moore-Penrose-Pseudoinverse of the kernel matrix to be the kernel matrix itself, $K^{\dagger} = K$. The computation of the mean in Eq. (23) reduces to

$$\mu=K\cdot\left[K+0\right]^{-1}\cdot L_o=K\cdot K^{\dagger}\cdot L_o=K\cdot L_o \quad (29)$$

for a label noise with zero covariance, $C = \mathbf{0}$. Due to the block diagonal structure of the kernel matrix, the multiplication by the label map $K \cdot L_o$ is equivalent to averaging within superpixels

$$\mu\left(x\right)=\frac{1}{|\Pi_j|}\sum_{y\in\Pi_j}L_o\left(y\right),x\in\Pi_j. \quad (30)$$

This identity facilitates the inference of label maps on image parcellations.

C. Contour Extraction and Superpixels

Contours—Essential for the label refinement is the reliable extraction of the contour Γ in the test image I. We use concepts from spectral clustering [52] to perform contour extraction. Image and texture cues are combined across several scales to achieve a reliable contour estimate [51], [53], which outperforms other contour detectors [51]. The global

probability of boundary *gPb* = *mPb* + *sPb* is the sum of a multi-scale probability of boundary *mPb* and a spectral probability of boundary *sPb*, illustrated in Fig. 5. The multi-scale probability of boundary considers intensity and texture information and is described in details in the Appendix. The spectral probability of boundary results from spectral clustering. Given a weight matrix *W*, the generalized eigenvalue decomposition

$$(D - W)\, v = \lambda D v \quad (31)$$

leads to eigenvectors $\{\mathbf{v}_k\}$ that partition the image, with the degree matrix $D_{ii} = \sum_i W_{ij}$. We consider two settings for the weight matrix: $W = W^I$ and $W = W^I + W^{Lo}$, based on the definitions

$$W^I_{xx'} = exp\left(-\max_{y \in \overline{xx'}} \{mPb(y)\}\right) \quad (32)$$

$$W^{L_o}_{xx'} = exp\left(-\max_{y \in \overline{xx'}} \{lPb(y)\}\right), \quad (33)$$

where *lPb* is the label probability of boundary computed on the initial label map L_o (see Fig. 5). The first definition of *W* only considers image contours, the second definition combines the image and the label cues.

We obtain the spectral boundary *sPb* by summing up the gradients of the 16 eigenvectors corresponding to the smallest non-zero eigenvalues λ_k [51]

$$sPb = \sum_{k=1}^{16} \frac{1}{\sqrt{\lambda_k}} \cdot \nabla v_k. \quad (34)$$

While *mPb* contains responses to all edges, *sPb* only captures the most salient structures in the image. Fig. 5 illustrates the probability of the boundary on an image from the left atrium. We employ *mPb* and *gPb* as contours Γ in Eq. (27).

Superpixels—For the kernel function in Eq. (28), we estimate the superpixels of the image. In this work, we compute superpixels based on the image contour Γ, but other superpixel algorithms, e.g., SLIC [54] could be used. We apply the oriented watershed transform on the contour Γ to create the finest partition of the image [51]. To adapt the superpixel size, we use the strength of the contours to build a hierarchical segmentation. The ultrametric contour map (UCM) [51] represents the hierarchy. Fig. 5 shows that superpixels form an oversegmentation of the image. The superpixel size is subject to a trade-off. Large regions provide stability in the face of errors in the initial label map, but they are also more likely to miss the organ boundary. The advantage of the hierarchical segmentation is that thresholding the UCM with ρ enables adaptation of the superpixel size. At the lowest level, we have the finest partition of the image, and the higher levels contain larger regions implied by stronger contours.

V. Experiments

We evaluate the methods by performing segmentation of the parotid glands and of the left atrium. We compare two non-local and five local methods to create initial label maps. For each of these label maps, we infer contour-driven label refinements with image boundaries and superpixels.

A. Parotid Glands

In the first set of experiments, we segment the left and right parotid glands in the head and neck CT images. The standardization with Hounsfield units supports patch-based approaches, since intensity values are comparable among patients. Each of the 16 CT scans contains between 80 and 200 axial slices with a slice thickness of 2.5mm; the in-plane resolution is 0.9mm; the slice size is 512 × 512 pixels. All 16 images have the left parotid gland labeled by a trained anatomist for treatment planning. The right parotid gland was consumed by a tumor in one patient. We quantify the segmentation quality by evaluating the Dice volume overlap score [55] and the modified Hausdorff distance [56] between the automatic and manual segmentations.

1) Non-local Labels—We first consider non-local, patch-based approaches for the segmentation (Section III-B2). To avoid training and testing on the entire image, we define regions of interest (ROI) around the parotid glands. These ROIs can be obtained either from a coarse registration or by template matching of the mandible bone, which is adjacent to the parotid glands. We train the random forest (RF) classifier on patches within the ROI of 8 randomly selected subjects. Tests are performed on the remaining 8 datasets for the left parotid gland and 7 for the right. We choose 500 trees per random forest with 12 predictors sampled for splitting at each node. We perform experiments with three different patch sizes: 7 × 7 × 3, 7 × 7 × 5, and 9 × 9 × 5. The patch dimensions account for the anisotropic image resolution. Similarly, experiments are performed with non-local means (NLM). Fig. 6 presents the statistical analysis of the results (left bar in each group). The random forest classifier yields better results than the NLM strategy, which is mainly due to many false positives in the NLM labels. Moreover, larger patch sizes tend to slightly improve the segmentations. The left column of Fig. 7 illustrates one slice with the labels created by both methods.

For each of the six initial label maps resulting from three patch-sizes and two methods, we apply the contour-driven regression. We evaluate *mPb* and *gPb* as contour estimates Γ for the kernel in Eq. (27). For *gPb*, we use the image-based weight matrix ($W = W^I$) and the weight matrix that includes label contours ($W = W^I + W^{Lo}$), referred to as *gPb label*. Additionally, we evaluate the kernel based on superpixels in the ultrametric contour map (UCM) in Eq. (28). We consider UCM based on image information only and UCM *label* that integrates label information. We create superpixels from the ultrametric contour maps by thresholding at $\rho = 0.2$ and setting the covariance of the label noise to $C = \sigma^2\mathbf{I}$ with $\sigma = 1.5$. Fig. 6 reports the results for different contour-driven refinement strategies for each of the initial label maps. Fig. 7 illustrates the refinement of the initial label maps with UCM and *mPb*.

The contour refinement leads to a significant improvement over the initial segmentation for almost all techniques. The refinement that lacks statistical significance is UCM on RF labels and NLM labels on the right parotid when evaluated for volume overlap. Generally, the regression with *gPb* and *mPb* leads to a clear improvement in Dice score and Hausdorff distance across all initial segmentations for the left and right parotid glands. The integration of the label information in the contour with *gPb label* does not further improve the results over *gPb*. In contrast, we observe a substantial improvement for UCM *label* over UCM for most cases. In particular, the refinement of NLM label maps with UCM results in an outlier with low Dice score for one subject. This is characteristic for the refinement with superpixels; they can yield large improvements but also produce substantial degradation if the regions are not well defined. The combination of image and label contours in UCM label helps to extract more meaningful superpixels, as indicated by the results. The improvement of UCM *label* over UCM is also illustrated qualitatively in Fig. 2.

We further study the dependence of the segmentation result on the threshold ρ for the creation of superpixels. Fig. 8 reports volume overlap for the refinement of the initial label map with the UCM *label* for different values of threshold ρ. The initial label map was created with the random forest classifier with patch size $7 \times 7 \times 3$. The segmentation results show only slight variations for a range of different threshold parameters. We continue using $\rho = 0.2$ for the remainder of the manuscript.

2) Local Labels—As a second approach, we evaluate local voting techniques (Section III-B1). The deformable registration is performed with Plastimatch using B-splines [57]. The best results were obtained with mutual information (MI) as similarity measure. Following a leave-one-out strategy, one of the images is selected as a test image and all of the remaining 15 images are registered to it. We compare several voting schemes to create the initial labels: majority voting, intensity-weighted voting, patch-based voting, and region-based voting. We work with discrete label probabilities in all cases. For intensity-weighted voting, we set the variance of the image likelihood to $\nu = 45$. For patch-based voting, we work with patches of size $7 \times 7 \times 5$ and $\nu = 20$. For region-based voting, we set the spatial variance $t^2 = 1$ and consider two different region sizes: $7 \times 7 \times 3$ and $17 \times 17 \times 9$ (denoted as “large”). For the large region, we set the step size to 2 to limit computational costs. For the region-based voting, we use single intensities and patches in the image likelihood.

After the initial label maps are created, we apply five different contour-driven refinement techniques. Fig. 9 presents the statistical analysis of the results. Again, we consider the refinement with contours and superpixels. In contrast to the previous experiment, we also evaluate the local noise model for label maps with spatially varying variances $\sigma^2(x)$, cf. Eq. (21), denoted as “local” in the figure. We only report results for UCM *label* in this experiment because of the superior performance in the previous experiment.

We note an improvement in the volume overlap measures for the application of more sophisticated voting schemes. Region patch voting on regions of size $7 \times 7 \times 3$ yield the best results for the left parotid. For the right parotid, the median is highest for region patch voting on large regions. A possible reason for this difference is that the right parotid generally yields lower Dice scores than the left parotid so that searching over a larger region has more

potential to improve results. When examining the Hausdorff distance, the performance is similar across all voting schemes for the left parotid. For the right parotid, region patch voting on the larger region yields the lowest median.

When investigating the Dice scores for the different refinement techniques of the left parotid, we observe a very similar performance of *gPb* and *mPb*. The local noise model and the inclusion of label information lead to similar results. Except for majority and intensity voting, the Dice score for UCM *label* is lower than for the boundary refinement methods. Small regions with intensity- and patch-based voting refined by *gPb* and *mPb* obtain the overall highest Dice score. All improvements for the left parotid gland are statistically significant. Regarding the modified Hausdorff distance for the left parotid, UCM *label* is the only method that yields consistently significant improvements over the initial labels and outperforms other refinement strategies. For the right parotid, voting on large regions with patches in combination with the contour refinement reaches the highest Dice score. Generally, the Dice score of UCM *label* on the right parotid is higher than that for the left parotid. The refinement with *mPb* and *mPb local* is not signifiant for the initial label maps created with majority, intensity, and both region intensity voting schemes. In terms of Hausdorff distance, superpixel voting with UCM *label* yields the significantly best results.

B. Left Atrium

In the second experiment, we automatically segment the left atrium of the heart in a set of 16 electro-cardiogram gated (0.2 mmol/kg) Gadolinium-DTPA contrast-enhanced cardiac MR angiograpy images (CIDA sequence, TR=4.3ms, TE=2.0ms, θ=40°, in-plane resolution varying from 0.51mm to 0.68mm, slice thickness varying from 1.2mm to 1.7mm, 512×512×96, ± 80 kHz bandwidth, atrial diastolic ECG timing to counteract considerable volume changes of the left atrium). The left atrium was manually segmented in each image by an expert.

Contrast agents used in MR angiography cause intensity variations that complicate purely patch-based approaches. With registration-based methods it is possible to compensate for these variations by fitting a transfer function during the registration. Specifically, we work with a variant of the diffeomorphic demons algorithm [58] and a polynomial transfer function [16]. The transfer function is estimated during the registration and corrects for intensity variations caused by the contrast agent. After the intensity correction, we normalize the intensities to be between zero and one. We evaluate various voting strategies together with the contour refinement. We perform leave-one-out experiments by treating one subject as the test image and the remaining 15 subjects as the training set.

We employ majority, intensity, patch, region intensity, and region patch voting. For intensity voting, we set the LogOdds slope constant to $\kappa = 1.5$ and intensity variance to $\nu = 0.5$. For patch voting, we use a patch of size $2 \times 2 \times 1$ with $\nu = 0.1$. For region voting, we set the region size to $7 \times 7 \times 3$ and set the spatial variance to $\tau^2 = 1$. We work with superpixels from the ultrametric contour map by thresholding at $\rho = 0.2$. The statistical analysis of the results in Fig. 10 shows an improvement in the Dice scores for intensity voting with respect to majority voting. Patch voting further improves the result. The best results are achieved for

region voting. The results for intensities or patches within the regions are similar. The evaluation with the Hausdorff distance produces similar results.

We employ the superpixel-based inference for the refinement of the initial segmentations, where the superpixels in UCM are constructed based on image and label contours (UCM *label*). The refinement yields a statistically significant improvement in Dice score and Hausdorff distance for majority, intensity, and patch-based voting. For the region-based voting, the refinement still yields an improvement, but it is no longer significant. The higher initial segmentation accuracy of the region-based voting makes it more difficult to achieve a significant refinement. Region voting with the refinement yields the overall best results.

Fig. 11 illustrates the segmentation results for majority and intensity-weighted voting together with the UCM refinement for several subjects. We see that the refined segmentation better captures the organ boundary. This is supported by the clearly lower Hausdorff distances in Fig. 10. On the images for Subject 1 in Fig. 11, we observe that the refinement achieves a better separation between the veins and atrium. This case is particularly challenging because the gap is small and registration errors of misaligning either the vein or the atrium lead to a closure. Integrating image and label cues together with voting on superpixels yields a more accurate segmentation. Fig. 2 presents another example, where we compare the result of intensity voting to the refinement with UCM and UCM *label*.

VI. Discussion

In general, purely patch-based strategies tend to be well suited for handling structures with significant variations because there exist no regularization constraints as in registration-based techniques. Standardized intensity values, as in CT images, facilitate patch-based approaches because the image intensities are comparable across subjects. However, the independent voting for each voxel increases the risk of false positives and isolated islands. The proposed refinement strategies improve these initial segmentations based on the covariance between voxels based on image contours.

For the segmentation of the parotid glands, the performance of registration-based approaches is better than that of purely patch-based techniques (Random forest, NLM). The integration of patches into the registration-based voting schemes generally improves the results. The region-based voting with patches leads to the best performance, where the optimal region size varies from the left to the right parotid. Region-based voting is a compromise between non-local methods and location-wise voting schemes, promising to benefit from both.

For the refinement techniques, we noted little differences between the *mPb* and *gPb* boundaries as contour estimates. The inclusion of the label information or a local noise model did not change the results much. For the superpixel voting with the ultrametric contour map, we noted a clear improvement by defining the parcellation based on image and label cues. Generally, the contour-driven refinement techniques yield a clear improvement over the initially created label maps.

VII. Conclusion

We presented a Bayesian framework for voting in atlas-based segmentation that spans the range from local to non-local methods. This framework enabled us to review existing methods and to propose novel voting schemes. To improve the initial label maps, we proposed a refinement with Gaussian process regression. The key contribution is a contour-driven distribution over label maps that is supported by structures in the image. Inference of the posterior distribution yields the MAP estimate of label maps and the refined segmentation. Our segmentation experiments for the parotid glands and the left atrium demonstrate improved performance for the new voting schemes and the contour-driven refinement.

Acknowledgements

We thank Michal Depa for algorithmic help, Martin Reuter for discussions, and Ehud Schmidt for providing image data. This work was supported in part by the Humboldt foundation, the National Alliance for Medical Image Computing (NIH NIBIB NAMIC U54-EB005149), the NeuroImaging Analysis Center (NIH NIBIB NAC P41-EB-015902), and the Wistron Corporation.

Appendix

Multi-scale Probability of Boundary

The calculation of the multi-scale probability of boundary *mPb* consists of several steps. First, the oriented gradient signal $G(x, \theta)$ is computed on image I [59]. The oriented gradient signal robustly estimates the image gradient by calculating the χ^2 distance between the histograms of two half-discs at each location x for various orientations θ. Depending on the size of the disc, we obtain gradient estimates on multiple scales. We work with eight equally spaced orientations between 0 and π. A successive Savitzky-Golay filtering enhances local maxima and reduces multiple detection peaks [60]. Texture provides an additional channel to calculate the gradient signal. We convolve the image with 17 Gaussian derivative and center-surround filters [53]. We obtain 64 different textons with a K-means clustering in the 17-dimensional space. The multi-scale probability of boundary for each orientation is calculated by summing the oriented gradient signal of the intensity G^I and texture G^T channel across different scales s

$$mPb\,(x,\theta) = \sum_{s\in\{3,5,10\}} G^I_s\,(x,\theta) + \sum_{s'\in\{5,10,20\}} G^T_{s'}\,(x,\theta)\,. \quad (35)$$

The maximum across orientations leads to the boundary strength for each location

$$mPb\,(x) = \max_{\theta}\, mPb\,(x,\theta)\,. \quad (36)$$

Finally, a non-maximum suppression [51] yields thinned contours, illustrated in Fig. 4.

REFERENCES

1. Ashburner J, Friston K. Unified segmentation. Neuroimage. 2005; 26(3):839–851. [PubMed: 15955494]

2. Pohl K, Fisher J, Grimson W, Kikinis R, Wells W. A Bayesian model for joint segmentation and registration. Neuroimage. 2006; 31(1):228–239. [PubMed: 16466677]
3. D'Agostino E, Maes F, Vandermeulen D, Suetens P. A unified framework for atlas based brain image segmentation and registration. WBIR. 2006
4. Heckemann R, Hajnal J, Aljabar P, Rueckert D, Hammers A. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. NeuroImage. 2006; 33(1):115–126. [PubMed: 16860573]
5. Rohlfing T, Brandt R, Menzel R, Maurer C, et al. Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage. 2004; 21(4):1428–1442. [PubMed: 15050568]
6. Rohlfing T, Brandt R, Menzel R, Russakoff D, Maurer C. Quo vadis, atlas-based segmentation? Handbook of Biomedical Image Analysis. 2005:435–486.
7. Svarer C, Madsen K, Hasselbalch SG, Pinborg LH, Haugbøl S, Frøkjær VG, Holm S, Paulson OB, Knudsen GM. Mr-based automatic delineation of volumes of interest in human brain pet images using probability maps. Neuroimage. 2005; 24(4):969–979. [PubMed: 15670674]
8. Sabuncu M, Yeo B, Van Leemput K, Fischl B, Golland P. A Generative Model for Image Segmentation Based on Label Fusion. TMI. 2010; 29
9. Coupé P, Manjón JV, Fonov V, Pruessner J, Robles M, Collins DL. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. NeuroImage. 2011; 54(2): 940 – 954. [PubMed: 20851199]
10. Rousseau F, Habas PA, Studholme C. A supervised patch-based approach for human brain labeling. IEEE Trans. Med. Imaging. 2011; 30(10):1852–1862. [PubMed: 21606021]
11. Cuingnet R, Prevost R, Lesage D, Cohen L, Mory B, Ardon R. Automatic detection and segmentation of kidneys in 3d ct images using random forests. MICCAI. 2012:66–74. [PubMed: 23286115]
12. Wolz, R.; Chu, C.; Misawa, K.; Mori, K.; Rueckert, D. Medical Image Computing and Computer-Assisted Intervention. Springer; 2012. Multi-organ abdominal ct segmentation using hierarchically weighted subject-specific atlases; p. 10-17.
13. Fowlkes C, Martin D, Malik J. Learning affinity functions for image segmentation: Combining patch-based and gradient-based approaches. Computer Vision and Pattern Recognition. 2003
14. Lesh M, Van Hare G, Epstein L, Fitzpatrick A, Scheinman M, Lee R, Kwasman M, Grogin H, Griffin J. Radiofrequency catheter ablation of atrial arrhythmias. Results and mechanisms. Circulation. 1994; 89(3):1074. [PubMed: 8124793]
15. Investigators AF. Risk factors for stroke and efficacy of antithrombotic therapy in atrial fibrillation: analysis of pooled data from five randomized controlled trials. Archives of Internal Medicine. 1994; 154:1449–1457. [PubMed: 8018000]
16. Depa M, Sabuncu MR, Holmvang G, Nezafat R, Schmidt EJ, Golland P. Robust atlas-based segmentation of highly variable anatomy: left atrium segmentation. Statistical atlases and computational models of the heart. 2010
17. Song, Z.; Tustison, N.; Avants, B.; Gee, J. Integrated graph cuts for brain MRI segmentation. In: ser. LNCS, R. Larsen; Nielsen, M.; Sporring, J., editors. MICCAI. Vol. 4191. Springer; Heidelberg: 2006. 2006. p. 831-838.
18. Wolz R, Aljabar P, Hajnal J, Hammers A, Rueckert D. LEAP: Learning embeddings for atlas propagation. NeuroImage. 2010; 49(2):1316–1325. [PubMed: 19815080]
19. Wang H, Suh JW, Das S, Pluta J, Altinay M, Yushkevich P. Regression-based label fusion for multi-atlas segmentation. CVPR. 2011:1113–1120.
20. Wang H, Suh J, Pluta J, Altinay M, Yushkevich P. Optimal weights for multi-atlas label fusion. Information Processing in Medical Imaging. 2011:73–84. [PubMed: 21761647]
21. Wachinger, C.; Golland, P. Medical Image Computing and Computer-Assisted Intervention–MICCAI. Springer; 2014. 2014. Atlas-based under-segmentation; p. 315-322.
22. Rasmussen, C.; Williams, C. Gaussian processes for machine learning. MIT Press; 2006.
23. Liu, P. Ph.D. dissertation. University of Toronto; 2007. Using gaussian process regression to denoise images and remove artefacts from microarray data.

24. Stytz MR, Parrott RW. Using kriging for 3d medical imaging. Computerized Medical Imaging and Graphics. 1993; 17(6):421 – 442. [PubMed: 8287354]
25. Kim K, Lee D, Essa I. Gaussian process regression flow for analysis of motion trajectories. International Conference on Computer Vision. 2011:1164–1171.
26. Wachinger, C.; Golland, P.; Reuter, M.; Wells, W. Medical Image Computing and Computer-Assisted Intervention–MICCAI. Springer; 2014. 2014. Gaussian process interpolation for uncertainty estimation in image registration; p. 267-274.
27. Gerig, T.; Shahim, K.; Reyes, M.; Vetter, T.; Lüthi, M. MICCAI. Springer; 2014. 2014. Spatially varying registration using gaussian processes," in Medical Image Computing and Computer-Assisted Intervention; p. 413-420.
28. Sudderth EB, Jordan MI. Shared segmentation of natural scenes using dependent Pitman-Yor processes. NIPS. 2008:1585–1592.
29. Han X, Hibbard LS, O'connell NP, Willcut V. Automatic segmentation of parotids in head and neck CT images using multi-atlas fusion. Medical Image Analysis for the Clinic: A Grand Challenge. 2010:297–304.
30. Ramus L, Malandain G. Multi-atlas based segmentation: Application to the head and neck region for radiotherapy planning. Medical Image Analysis for the Clinic: A Grand Challenge. 2010:281–288.
31. Chen A, Noble JH, Niermann KJ, Deeley MA, Dawant BM. Segmentation of parotid glands in head and neck CT images using a constrained active shape model with landmark uncertainty. SPIE. 2012; 8314:83140P.
32. Qazi AA, Pekar V, Kim J, Xie J, Breen SL, Jaffray DA. Auto-segmentation of normal and target structures in head and neck CT images: A feature-driven model-based approach. Medical physics. 2011; 38:6160. [PubMed: 22047381]
33. Karim R, Mohiaddin R, Rueckert D. Left atrium pulmonary veins: segmentation and quantification for planning atrial fibrillation ablation. SPIE. 2009
34. Wachinger, C.; Golland, P. Spectral label fusion. In: ser. LNCS, N. Ayache; Delingette, H.; Golland, P.; Mori, K., editors. MICCAI. Vol. 7512. Springer; Heidelberg: 2012. 2012. p. 410-417.
35. Wachinger, C.; Sharp, G.; Golland, P. Contour-driven regression for label inference in atlas-based segmentation. In: ser. LNCS, editor. MICCAI. Springer; Heidelberg: 2013. 2013.
36. Iglesias JE, Sabuncu MR, Leemput KV. Improved inference in bayesian segmentation using monte carlo sampling: Application to hippocampal subfield volumetry. Medical Image Analysis. 2013; 17(7):766–778. [PubMed: 23773521]
37. Pohl, KM.; Fisher, J.; Shenton, M.; McCarley, RW.; Grimson, WEL.; Kikinis, R.; Wells, WM. Medical Image Computing and Computer-Assisted Intervention–MICCAI. Springer; 2006. Logarithm odds maps for shape representation; p. 955-963.2006
38. van Rikxoort E, Isgum I, Staring M, Klein S, van Ginneken B. Adaptive local multi-atlas segmentation: application to heart segmentation in chest CT scans. Proc. of SPIE Vol. 2008; 6914:691, 407–1.
39. Aljabar P, Heckemann R, Hammers A, Hajnal J, Rueckert D. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. Neuroimage. 2009; 46(3):726–738. [PubMed: 19245840]
40. Jain AK, Farrokhnia F. Unsupervised texture segmentation using gabor filters. Pattern recognition. 1991; 24(12):1167–1186.
41. Liao S, Gao Y, Shi Y, Yousuf A, Karademir I, Oto A, Shen D. Automatic prostate mr image segmentation with sparse label propagation and domain-specific manifold regularization. IPMI. 2013:511–523.
42. Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions. Pattern recognition. 1996; 29(1):51–59.
43. Viola P, Jones M. Robust real-time object detection. International Journal of Computer Vision. 2001; 4
44. Dalal, N.; Triggs, B. Computer Vision and Pattern Recognition. IEEE; 2005. Histograms of oriented gradients for human detection; p. 886-893.

45. Wachinger C, Brennan M, Sharp G, Golland P. On the importance of location and features for the patch-based segmentation of parotid glands. MICCAI Workshop on Image-Guided Adaptive Radiation Therapy. Midas Journal. 2014
46. Wachinger C, Navab N. Entropy and laplacian images: Structural representations for multi-modal registration. Medical Image Analysis. 2012; 16(1):1 – 17. [PubMed: 21632274]
47. Wachinger C, Golland P, Magnain C, Fischl B, Reuter M. Multi-modal robust inverse-consistent linear registration. Human Brain Mapping. 2015; 36(4):1365–1380. [PubMed: 25470798]
48. Wu, G.; Wang, Q.; Daoqiang; Shen, D. MICCAI Workshop on Sparsity Techniques in Medical Imaging. Springer; 2012. Robust patch-based multi-atlas labeling by joint sparsity regularization.
49. Bai W, Shi W, O'Regan DP, Tong T, Wang H, Jamil-Copley S, Peters NS, Rueckert D. A Probabilistic Patch-Based Label Fusion Model for Multi-Atlas Segmentation With Registration Refinement: Application to Cardiac MR Images. Medical Imaging, IEEE Transactions on. 2013; 32(7):1302–1315.
50. Breiman L. Random forests. Machine learning. 2001; 45(1):5–32.
51. Arbelaez P, Maire M, Fowlkes C, Malik J. Contour detection and hierarchical image segmentation. IEEE Trans. on Pat. Anal. Mach. Intel. 2011; 33(5):898–916.
52. Shi J, Malik J. Normalized cuts and image segmentation. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2000; 22(8):888–905.
53. Malik J, Belongie S, Leung T, Shi J. Contour and texture analysis for image segmentation. International Journal of Computer Vision. 2001; 43(1):7–27.
54. Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Susstrunk S. Slic superpixels compared to state-of-the-art superpixel methods. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 2012; 34(11):2274–2282.
55. Dice L. Measures of the amount of ecologic association between species. Ecology. 1945; 26(3): 297–302.
56. Dubuisson M, Jain A. A modified hausdorff distance for object matching. International Conference on Pattern Recognition. 1994; 1:566–568.
57. Shackleford J, Kandasamy N, Sharp G. On developing b-spline registration algorithms for multi-core processors. Physics in Medicine and Biology. 2010; 55(21):6329. [PubMed: 20938071]
58. Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: Efficient non-parametric image registration. NeuroImage. 2009; 45(1):61–72.
59. Martin D, Fowlkes C, Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues. PAMI. 2004; 26(5):530–549.
60. Savitzky A, Golay MJE. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry. Jul; 1964 36(8):1627–1639.

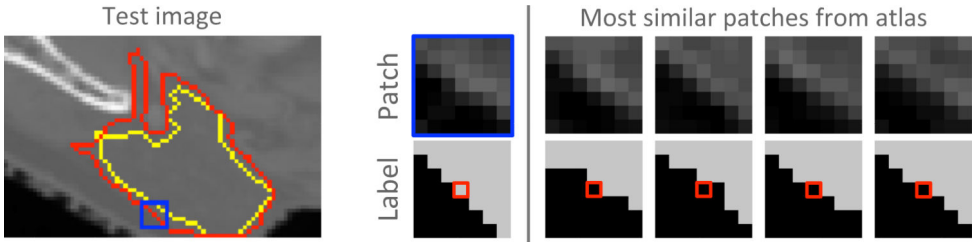
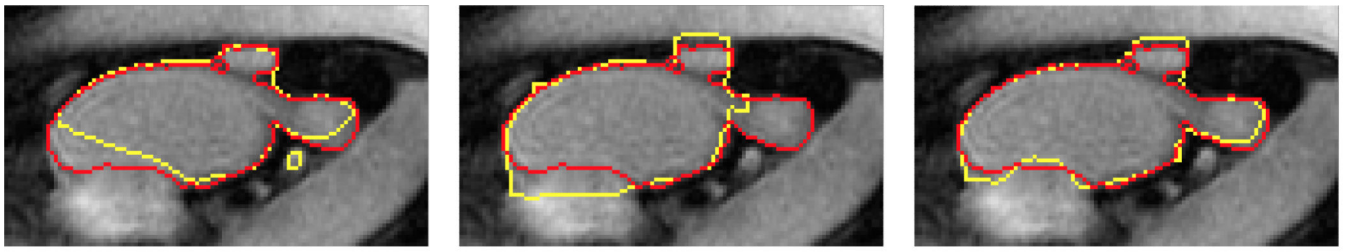


Fig. 1.

Left: CT image with segmentation of left parotid (yellow: patch-based, red: manual). Right: Magnification of the blue patch (top) with manual segmentation (bottom). The four most similar patches in the repository vote for background (black at the center location), although the patch belongs to the left parotid. Image patches are intensity normalized for visualization.



(a) Intensity-weighted label fusion (b) Refinement with image contours (c) Refinement with image and label contours

Fig. 2.

Example segmentations of MR angiography images of the left atrium of the heart (yellow: automatic, red: manual). The initial atlas-based segmentation is established with intensity-weighted label fusion. We compare the refinement that only uses contour information from the image and the refinement that combines contours from the image and those from the label map.

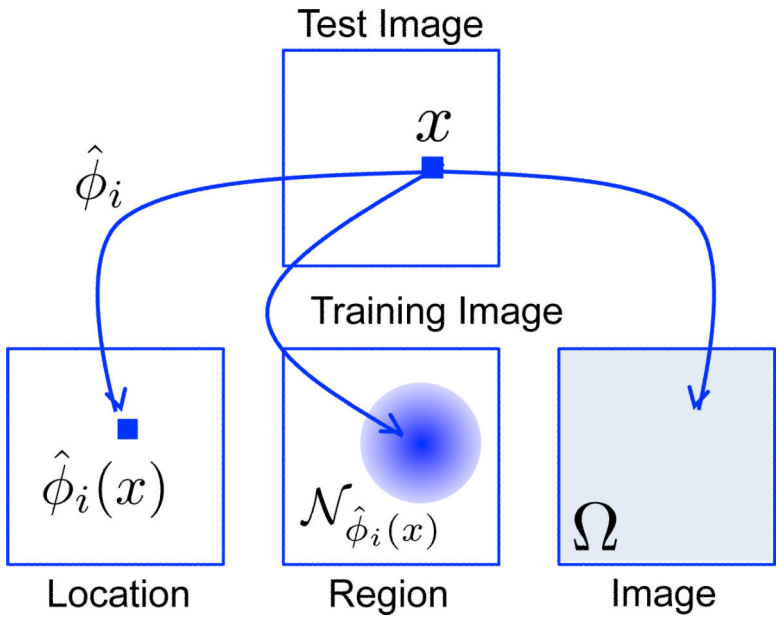


Fig. 3.

Three different voting schemes for labeling the location *x* in the test image (top) using the training image (bottom) and image transformation $\hat{\phi}_i$. Left: Only the information in the training image at the single location $\hat{\phi}_i(x)$ is considered. Middle: Information in a local region $N_{\hat{\phi}_i(x)}$ centered at $\hat{\phi}_i(x)$ is included, with higher weighting towards the center. Right: A non-local approach that integrates information from the entire image grid Ω.

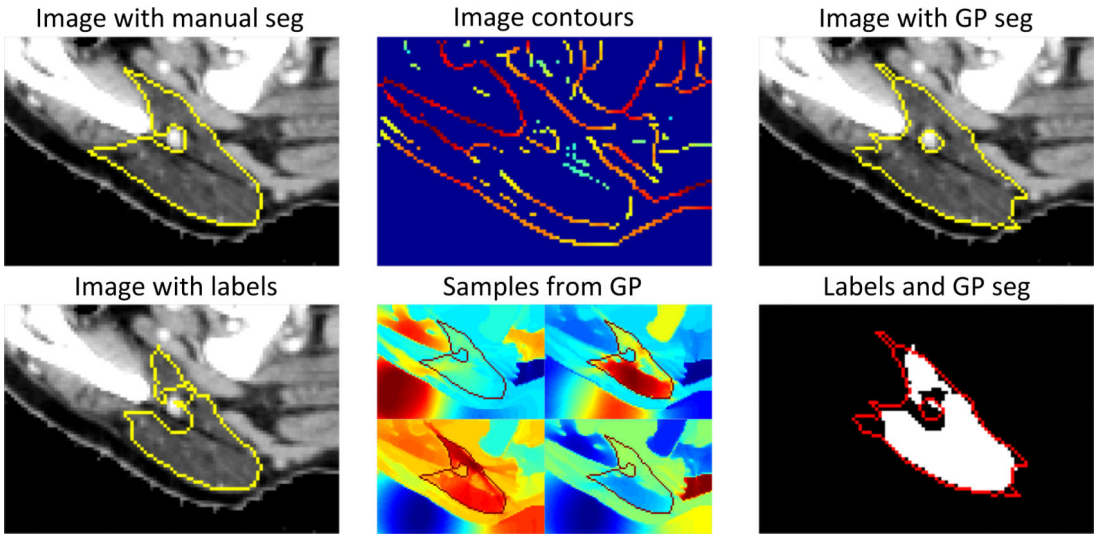


Fig. 4.

Gaussian process segmentation of the left parotid gland. The initial label from the atlas-based segmentation (bottom left) only partially agrees with the manual segmentation (top left). We extract contours from the image (top center) and use them to construct the kernel function *k*, see Eq. (27). The kernel defines a distribution over label maps $L \sim \mathcal{GP}(0, k)$ supported by the image. Samples drawn from the Gaussian process illustrate possible segmentations of the image (bottom center). The manual segmentation is overlaid for reference. The samples exhibit sharp boundaries necessary for segmentation, and the correlation of locations in the parotid gland. The mode of the posterior distribution results in the refined segmentation, overlaid on the intensity image (top right) and on the initial label map (bottom right).

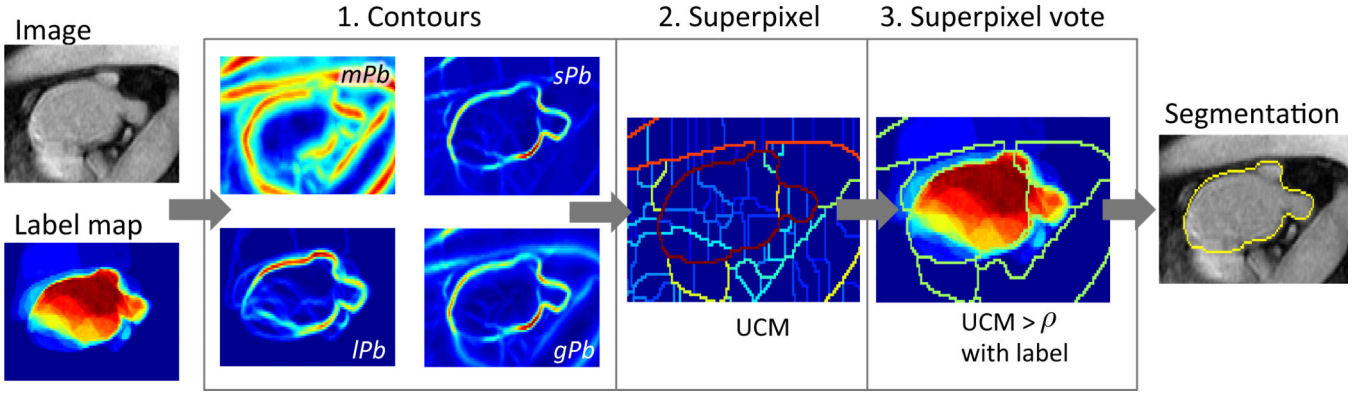


Fig. 5.

Overview of the atlas-based segmentation with region-based voting on superpixels for left atrium. The test image and the initial label map L_o serve as input. First, contours are extracted from both input images, yielding *mPb* and *lPb*. Both are combined for the calculation of the spectral contour *sPb* and global boundary *gPb*. The contour gives rise to a hierarchical parcellation of the image, represented with the ultrametric contour map (UCM). Thresholding the UCM at level ρ yields superpixels at a specific granularity. Gaussian process inference with the kernel in Eq. (28) yields superpixel-wise voting and the final segmentation.

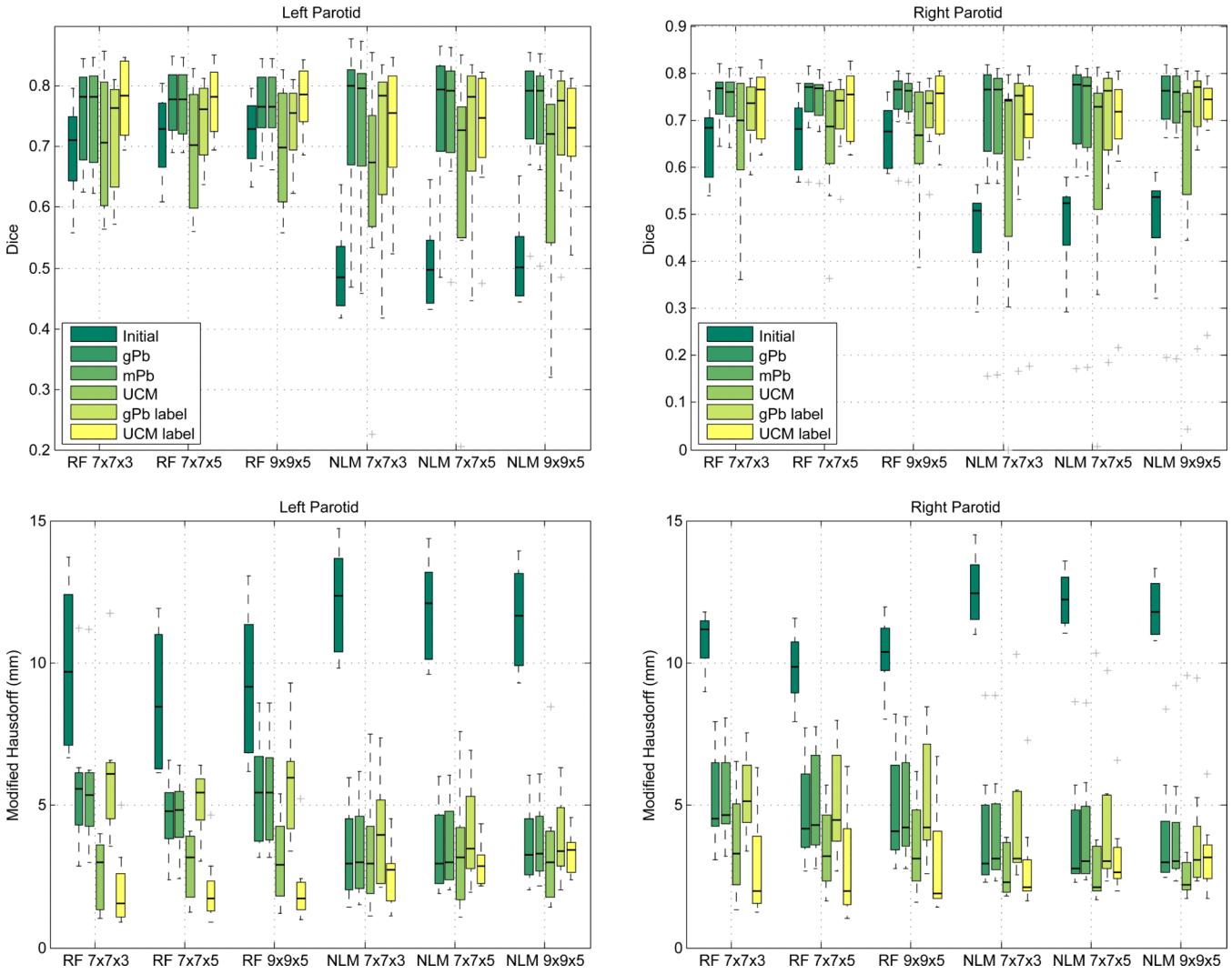


Fig. 6.

Dice volume overlap (top) and modified Hausdorff distance (bottom) for the segmentation of left parotid (left) and right parotid (right) glands. Initial labels are created with a random forest (RF) classifier and non-local means (NLM) for different patch sizes. The refinement of label maps is done with contours (*gPb*, *mPb*) and superpixels (UCM), where we also evaluate the integration of label map contours (label). Center line indicates median, the boxes extend to the 25th and 75th percentiles, and the whiskers reach to the most extreme values not considered outliers (crosses).

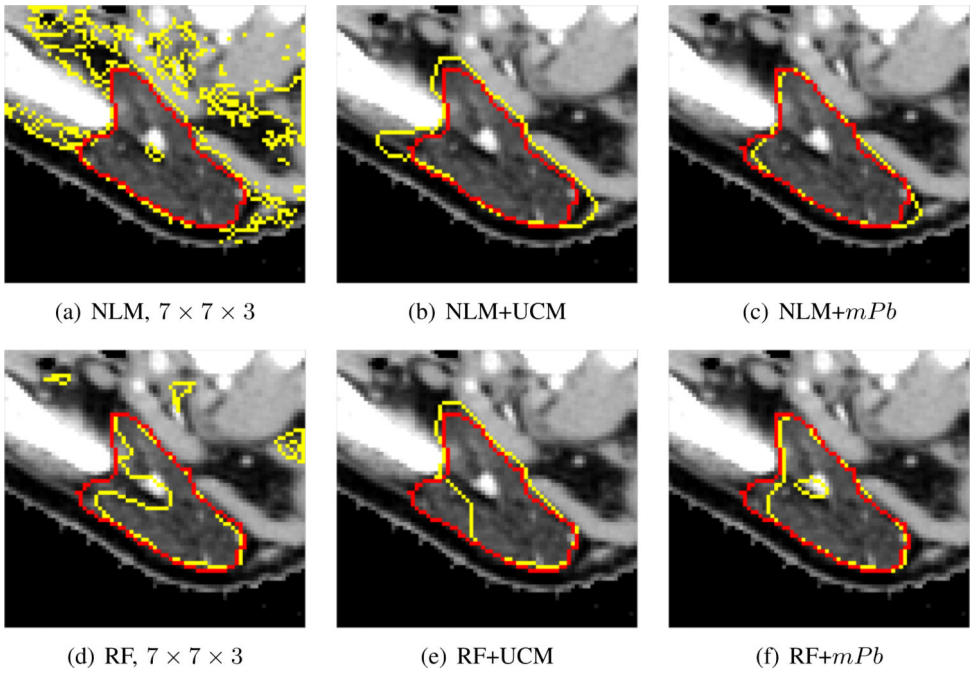


Fig. 7.

Examples slices for the patch-based segmentation and contour-driven refinement of the left parotid gland. Automatic segmentation results are shown in yellow, manual delineations are shown in red. The initial label maps created with non-local means (NLM) and random forests (RF) are shown in the left column. The superpixel refinement with UCM and the contour refinement with *mPb* are shown in the second and third column, respectively.

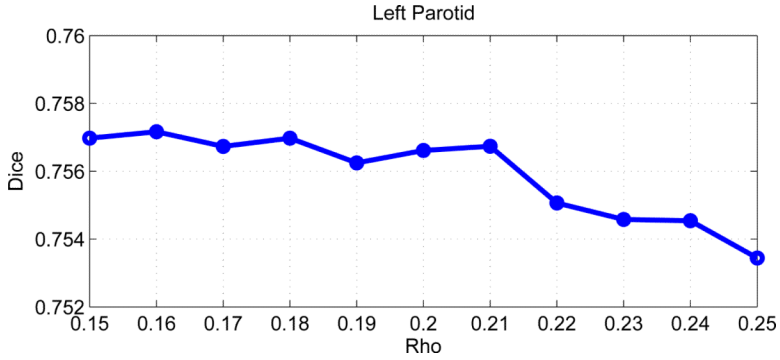


Fig. 8.

Mean Dice score of the refined segmentation with UCM *label* as a function of the superpixel threshold ρ. The initial label map was created with the RF classifier on $7 \times 7 \times 3$ patches for the left parotid gland.

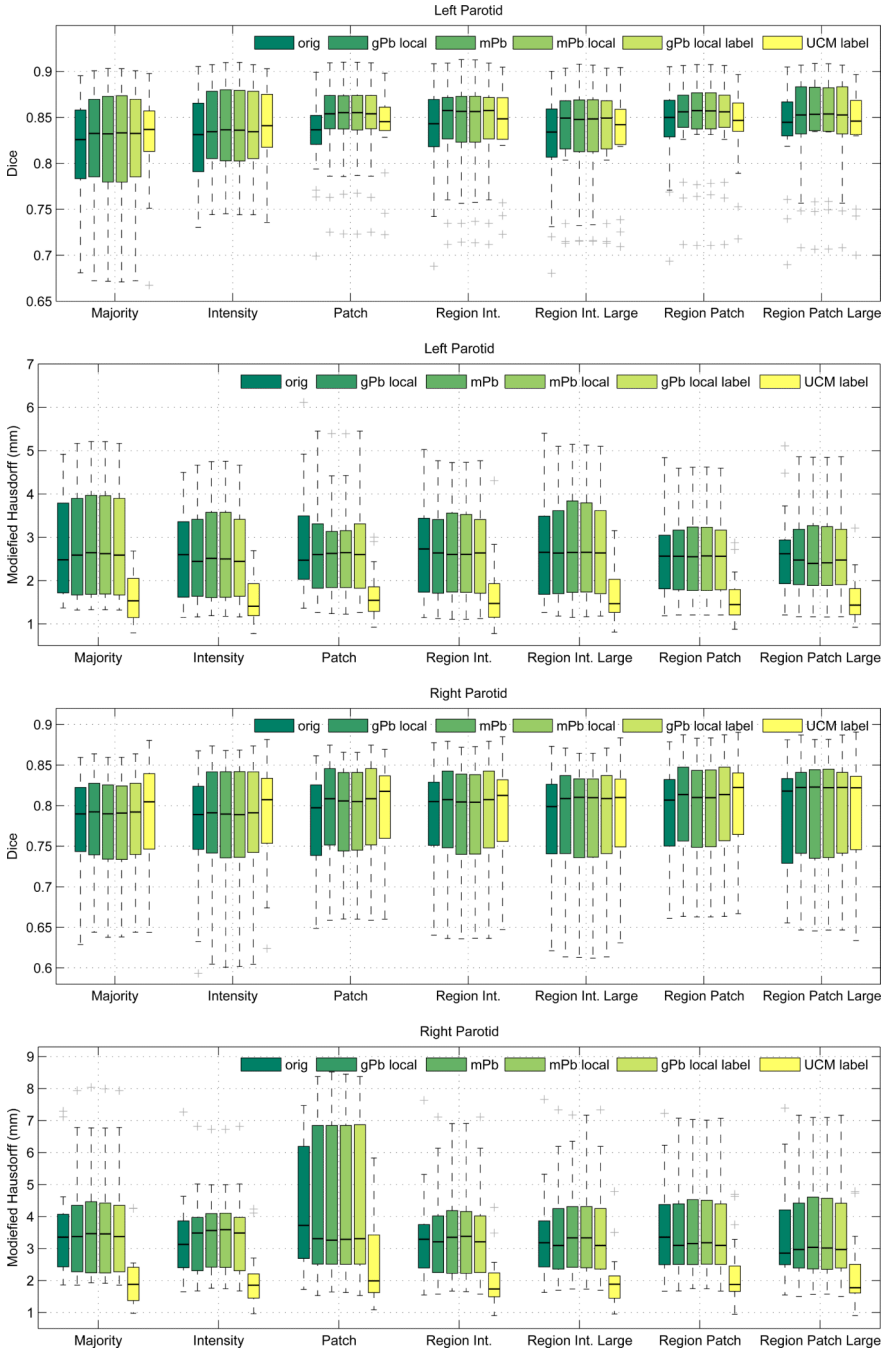


Fig. 9.

Dice volume overlap and modified Hausdorff distance for the segmentation of left (top two graphs) and right (bottom two graphs) parotid glands. Initial label maps are created with atlas-based voting schemes (majority, intensity, patch, region intensity, and region patch voting). For the contour-driven refinement, we set *gPb* and *mPb* as contour estimates. We further compare to using a local noise model (local) and incorporating label contours (label). For the superpixel refinement, we use UCM with label contours. Center line indicates the

median, the boxes extend to the 25^{th} and 75^{th} percentiles, and the whiskers reach to the most extreme values not considered outliers (crosses).

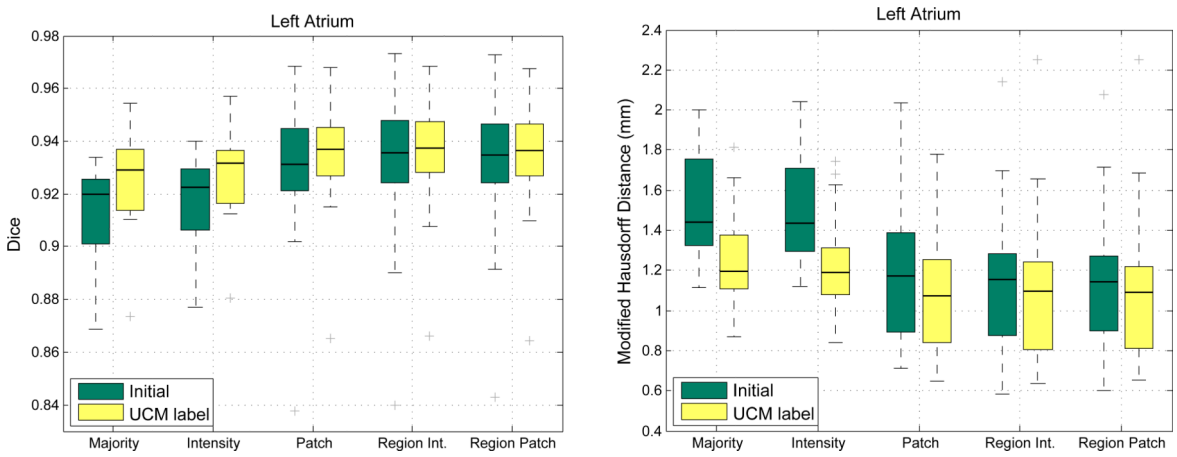


Fig. 10.

Dice volume overlap (left) and modified Hausdorff distance (right) for the segmentation of the left atrium. Initial label maps are created with atlas-based voting schemes (majority, intensity, patch, region intensity, and region patch voting). The label maps are refined with the superpixels in UCM with the label map contours information (UCM label). Center line indicates median, the boxes extend to the 25^{th} and 75^{th} percentiles, and the whiskers reach to the most extreme values not considered outliers (crosses).

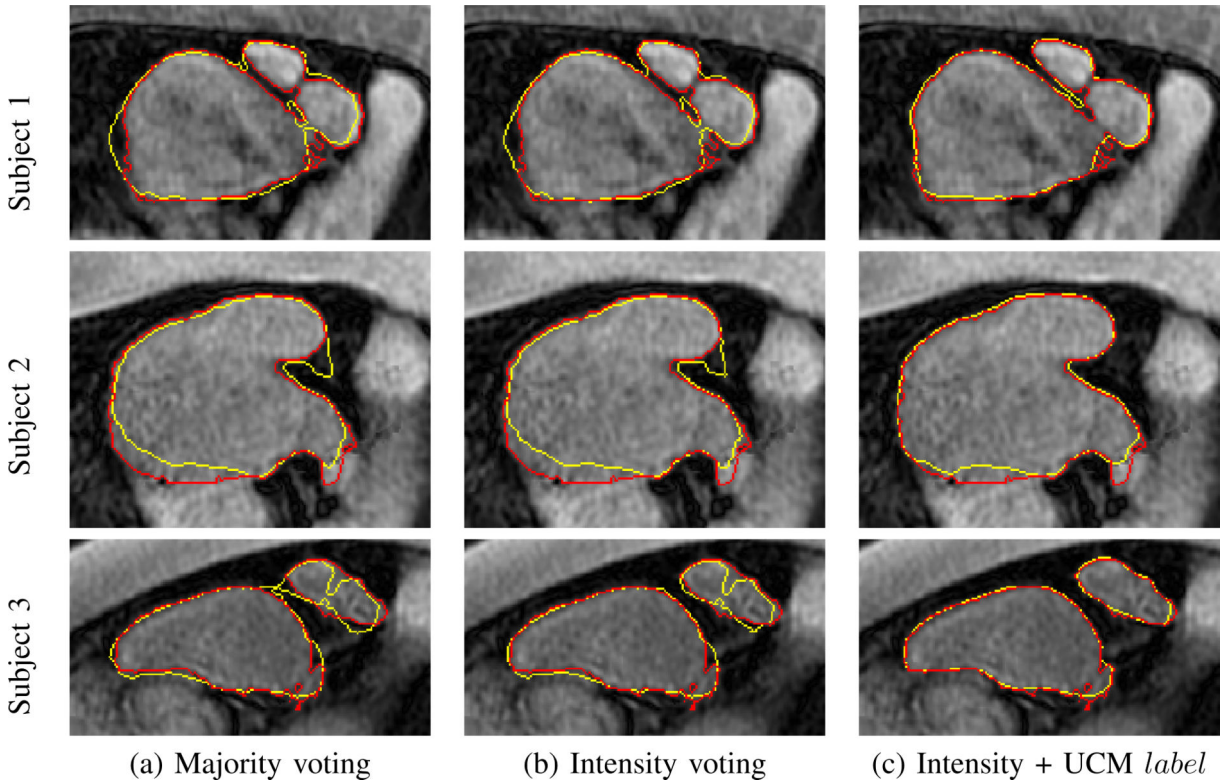


Fig. 11.

Examples slices for the atlas-based segmentation and contour-driven refinement of the left atrium. Automatic segmentation results are shown in yellow, manual delineations are shown in red. The initial label maps were created with majority and intensity-weighted voting. The label maps for intensity-weighted voting are refined with the superpixels in UCM also considering label map contours (UCM *label*).

TABLE I

Overview of different voting schemes for atlas-based segmentation obtained by varying approximations of the deformation field prior $p(\Phi|I,\mathscr{I})$.

Voting scope	Approximation of $p(\Phi \mid I,$ I)	Type	Posterior probability $p(S(x) \mid I,$ I, S)
Location	Delta function	Local	$\sum_{i=1}^{n} p\left(S(x) \mid S_i(\hat{\phi}_i(x))\right) \cdot p\left(I(x) \mid I_i(\hat{\phi}_i(x))\right)$
Region	Normal distribution	Local	$\sum_{i=1}^{n} \sum_{y \in N_{\hat{\phi}_i(x)}} p\left(S(x) \mid S_i(y)\right) \cdot p\left(I(x) \mid I_i(y)\right) \cdot \mathrm{N}\left(y \mid \hat{\phi}_i(x),\ \tau^2\mathbf{I}\right)$
Image	Uniform distribution	Non-Local	$\sum_{i=1}^{n} \sum_{y \in \Omega} p\left(S(x) \mid S_i(y)\right) \cdot p\left(I(x) \mid I_i(y)\right)$