Probabilistic Fitting of Active Shape Models

Andreas Morel-Forster(✉), Thomas Gerig, Marcel Lüthi, and Thomas Vetter

University of Basel, 4051 Basel, Switzerland
{andreas.forster,thomas.gerig,marcel.luethi,thomas.vetter}@unibas.ch
http://gravis.dmi.unibas.ch

Abstract. Active Shape Models (ASMs) are a classical and widely used approach for fitting shape models to images. In this paper, we propose a fully probabilistic interpretation of ASM fitting as Bayesian inference. To infer the posterior, we use the Metropolis-Hastings algorithm. We then use the maximum a posteriori sample as the segmentation result. Our approach has several advantages compared to classical ASM fitting: (1) We are left with fewer parameters that we need to choose. (2) It is less prone to get trapped in local minima. (3) It becomes straightforward to extend the approach to include additional information, such as expert annotations. (4) It is even simpler to implement than the classical ASM fitting method.

We apply our algorithm to the SLIVER dataset and show that it achieves a higher segmentation accuracy than the standard ASM approach. We further demonstrate the flexibility and expressivity of the framework by integrating experts annotations along parts of the outline to further increase the accuracy. The code used for fitting is based on open-source software and made available to the community.

Keywords: Active shape model · Statistical shape model
Gaussian process · MCMC · Sampling · Metropolis Hastings
Bayesian · Liver

1 Introduction

To automate medical diagnosis, treatments or the planning of interventions, a segmentation of an organ is a useful preprocessing step. However, analyzing volumetric computed tomography (CT) images is a difficult task, because human organs are highly variable in terms of shape and appearance. Possible shifts from neighboring organs can change the visual appearance of the boundary. An often used approach to organ segmentation is the Active Shape Model (ASM) [2] algorithm.

An ASM consists of two main parts. A point distribution model (PDM) is used to summarize prior knowledge about the shapes of the organs. An instance of an organ's shape is fully described using a set of parameters. In addition, the

M. Reuter et al. (Eds.): ShapeMI 2018, LNCS 11167, pp. 137–146, 2018.
https://doi.org/10.1007/978-3-030-04747-4_13

ASM has a built-in prior for the appearance of the volume around the organ boundary. The appearance is modeled at a sparse set, of so-called profile points. Visual features are extracted in the normal direction and approximated using a Gaussian distribution. When adapting the shape, better locations for the profile points are searched. To increase robustness while fitting, new points are removed when the point or appearance distance is larger than a given threshold. The shape is updated using the remaining points. To keep the segmented shapes plausible the model parameters are restricted to a certain interval.

In this paper, we propose a probabilistic interpretation of this model. The probabilistic formulation has several advantages: (1) It makes many of the implicit assumptions taken in the ASM explicit. (2) It does not rely on seemingly arbitrary parameters. (3) We can apply standard inference procedures developed in the statistics community for fitting the model. (4) It provides a principled way of extending the algorithm with additional information such as expert annotations.

We propose to use a sampling-based strategy for model fitting based on the Metropolis-Hastings (MH) algorithm. This leads to a simple, stochastic algorithm, which is less prone to get stuck in local optima and provides an estimate of the posterior distribution. In this approach, the only parameter to choose is a proposal distribution for the MH algorithm. While in theory, the exact choice of this distribution has no influence on the result, in practice, this can change the efficiency of the algorithm. Our experiments on the SLIVER dataset [5] show that our method is more robust and leads to better segmentation performance compared to the standard, deterministic search-based approach. Finally, we demonstrate how we can incorporate expert annotations, and that such additional information significantly improves on the segmentation performance.

Prior Work: The ASM approach, as introduced by Cootes and Taylor [1,2], is a generic approach to model-based image segmentation. It has many components and parameters, which affect its performance. An overview of different possibilities to tune the algorithm is given by van Ginneken et al. [12]. Consequently, a lot of work has been done. In the following, we concentrate on reviewing some work addressing specific limitations of the standard ASM: Wimmer et al. [13] replaces the Gaussian assumption with a probabilistic likelihood of the boundary profile based on a k-nearest-neighbor estimate using positive and negative boundary profiles. Norajitra et al. replaced the search for better boundary locations in [10] with random forests, which, when compared to the line profiles, take information of a larger volume into account. Kirschner et al. [7] use a non-linear shape prior based on a kernel PCA. They showed the superior performance of the non-linear model for vertebra, which we think holds also for livers. Note, that our proposed method for model adaptation is orthogonal to all formerly mentioned changes and can make use of improved shape or appearance models.

There is also work on advanced fitting strategies in the recent literature about ASMs. In [3], Esfandiarkhani et al. propose a non-linear fitting scheme. Zhan et al. present a method related to the inverse gradient descent optimization used for active appearance models [14]. In contrast to those, our proposed method is

simple to implement, fully probabilistic and provides a principled mechanism for integrating additional information.

Van Ginneken et al. proposed an extension of ASM to allow for interactive fitting [4]. In contrast to our method, their method requires the use of corresponding landmarks, while we can incorporate information, which is given as line to surface correspondence only. Furthermore, in our approach, the constraint is formulated probabilistically, which allows us to add the user annotation as uncertain observation, in a principled way.

The most closely related work to ours is the one of Schönborn et al. [11]. They propose to use sampling to adapt a 3D face model to 2D photographs using computer graphics.

2 Background : Active Shape Models

An Active Shape Model consists of two main components: (1) A *Point Distribution Model (PDM)*, which represents the normal shape variation of the modeled anatomical structure and (2) an intensity model, which models the intensities in a neighborhood around dedicated points of the PDM. Before discussing the main fitting algorithm, we discuss these two main components in more detail.[1]

Point Distribution Models: The main idea behind a PDM is that given a set of $i = 1, \ldots, N$ typical example surfaces $\{\Gamma_i\}$ of a certain shape, it becomes possible to learn the mean shape and the normal variability of this shape. For this to be possible, it is necessary that the example surfaces are in correspondence. This means that each surface is defined using the same number of boundary points, $\Gamma_i = (x_1^i, \ldots, x_n^i)$, and that the points $\{x_k^i\}$ on each of the surfaces $\{\Gamma_i\}$ are at the same anatomical location. Assuming that all the surfaces $\{\Gamma_i\}$ are rigidly aligned to each other, we can define the mean of a boundary point as $\overline{x}_k = \frac{1}{N}\sum_{i=1}^{N} x_k^i$. The corresponding mean shape is given as $\overline{\Gamma} := (\overline{x}_1, \ldots, \overline{x}_n)$. Furthermore, we can compute from the example surfaces $\{\Gamma_i\}$ a set of $N-1$ principal components, which represent the directions of main variation in the data. In our model, the possible locations of a point x_k is defined as a linear combination of the principal components with coefficient vector $\alpha = (\alpha_1, \ldots, \alpha_{N-1})$ as follows:

$$x_k(\alpha) := \overline{x}_k + \sum_{i=1}^{N-1} \alpha_i u_k^i$$

where $u_k^i \in \mathbb{R}^3$ denotes a displacement direction for the k-th point given by the i-th principal component. The corresponding shape $\Gamma(\alpha)$ with coefficients α is in turn defined as $\Gamma(\alpha) := (x_1(\alpha), \ldots, x_N(\alpha))$.

[1] Note that, while the mathematically concepts are equivalent to the classical ASM papers [1,2], our exposition of PDMs is based on the notation and interpretation of Point Distribution Models as Discrete Gaussian processes, as presented by Lüthi et al. [9].

A probabilistic model of shape variation is obtained by assuming that the coefficients α_i are independent and normally distributed, $\alpha_i \sim N(0,1)$. For this to hold, the eigenvectors forming the basis are scaled by the square root of the corresponding eigenvalue. The model can then be seen as a multivariate normal distribution of shapes centered around $\overline{\Gamma}$, the mean shape.

For explaining an organ's shape in a new image, it is not sufficient to model shape variation, but also the pose needs to be modeled. Thus we define a translation parameter $t \in \mathbb{R}^3$ and a rotation matrix $R(\phi, \psi, \rho)$, which itself is parametrized by the three Euler angles (ϕ, ψ, ρ). For notational convenience, we summarize all parameters in a single vector $\theta = (t, R, \alpha)$. Using this full parametrization, the $k-$th model point becomes

$$x_k(\theta) := R\left(\overline{x}_k + \sum_{i=1}^{N-1} \alpha_i u_k^i\right) + t$$

and the notation $\Gamma(\theta)$ is used again to refer to the full surface induced by θ.

Intensity Models: On top of the PDM, an Active Shape Model describes an intensity model. The intensity model summarizes the intensity distribution around a subset of the N points that define the model. We refer to these points as profile points. For every such profile point x_k, an ASM models the variation of some intensity feature $\rho_k \in \mathbb{R}^d$. Usually, it is assumed that ρ_k follows a normal distribution $\rho_k \sim N(\mu_k, \Sigma_k)$, with a mean μ_k and covariance matrix Σ_k. The parameters are estimated during training time from a set of example images. For a given point x and image I, we can extract the corresponding feature vector $\rho(x, I)$ and use the model to evaluate how likely this point is to correspond to the boundary point x_k:

$$p(\rho(x, I)) = \frac{1}{Z} \exp\left((\rho(x, I) - \mu_k)^T \Sigma_k^{-1} (\rho(x, I) - \mu_k)\right) \tag{1}$$

Hence we can use this intensity model to select the most likely boundary point during model fitting.

Active Shape Model Fitting: With these concepts defined, we now formulate the ASM algorithm in Algorithm 1. Note that once we have found the best matching points, the optimal rotation and shape parameters on line 10, can be computed in closed form (see e.g. [1] for details). The seemingly arbitrary constraint on line 12 involving the threshold κ_α is usually motivated by assuming that α_i follow a standard normal distribution, which implies that values of α_i which are more than three standard deviations away from the mean are very unlikely. New point locations are dropped if they are unlikely under the shape model using κ_S, or if they are unlikely under the appearance model using κ_T.

3 Method

In this section, we introduce our fully probabilistic interpretation of the ASM fitting. We use the ASM in a Bayesian setting, where our goal is to compute the

Algorithm 1. Active Shape Model fitting

$\theta_0 \leftarrow$ intialization
for $i = 1$ **to** *max-iterations* **or** *converged* **do**
  for $x_k \in \{x_1, \ldots, x_K\}$ **do**
    generate candidate locations $x_c, c = 1, \ldots, C$
    find best new candidate x'_k as $\arg\max_{x_c} p(\rho_k(x_c, I))$
    if $||\arg\min_{\alpha} Rx_k(\alpha) + t - x'_k|| > \kappa_S$ **or** $\sqrt{(\rho_k(x'_k, I) - \mu_k)^T \Sigma_k^{-1} (\rho_k(x'_k, I) - \mu_k)} > \kappa_T$ **then**
      drop the correspondence pair (x_k, x'_k)
    end if
  end for
  calculate rigid alignment R, t of remaining sets $\{x_k\}$ and $\{x'_k\}$
  find best α given R, t and remaining sets $\{x_k\}$ and $\{x'_k\}$
  $\alpha_i \leftarrow \max(\min(\alpha_i, \kappa_\alpha), -\kappa_\alpha)$
end for

posterior distribution over the parameters θ given the observed image I, which we would like to segment. The posterior is defined as:

$$p(\theta|I) = \frac{p(\theta)p(I|\theta)}{\int p(\theta)p(I|\theta)d\theta} \propto p(\theta)p(I|\theta). \tag{2}$$

As we will see later, our inference procedure makes it possible to work with the unnormalized posterior, hence we only have to specify the prior $p(\theta)$ and likelihood function $p(I|\theta)$.

Prior Distributions: The prior distribution defines our assumptions about the value of the parameters, before we have seen any data. We assume independence between components of the prior, i.e.

$$p(\theta) = p(t, R, \alpha) = p(t)p(R)p(\alpha). \tag{3}$$

For the pose parameters t and R we assume that every possible value is equally likely, so $p(\phi) = p(\psi) = p(\rho) = p(t_1) = p(t_2) = p(t_3) = \mathcal{U}(-\infty, \infty)$. For the shape prior, we use the PDM distribution, which states that the coefficients α_i are independent and follow a standard normal distribution $\alpha_i \sim N(0, 1)$.

Likelihood Functions: Let $x_k \in \Gamma_R$ denote the k-th point of the model, which has an associated intensity distribution $N(\mu_k, \Sigma_k)$ (Cf. Eq. 1). The likelihood for a fixed point x_k of the model is then given by

$$p(I|\theta, x_k) = \frac{1}{Z} \exp(-(\rho(x_k(\theta), I)) - \mu_k)^T \Sigma_k (\rho(x_k(\theta), I)) - \mu_k)$$

where Z is a normalization constant. Assuming independence between the observed values, we can define the shape likelihood as

$$p(I|\theta) = \prod_{k=1}^{N} p(I|\theta, x_k) \tag{4}$$

We can integrate additional constraints in the posterior formulating them as an additional likelihood. We demonstrate this in terms of a few strokes drawn by an expert along the perceived boundary of the liver. We integrate the lines as a discrete set of points $\{x_l\}, l = 1, \ldots, L$ using the likelihood

$$p(\{x_l\}|\theta) = \prod_{l=1}^{L} \mathcal{N}(x_l|CP(x_l, \Gamma(\theta)), \sigma_l). \tag{5}$$

Here, the function CP returns the point from the surface $\Gamma(\theta)$ which is closest to the point x_l.

Approximating the Posterior Distribution: The posterior distribution from Eq. (2) cannot be expressed in closed form. However, it is possible to draw samples from it, which we can use to find shapes that are likely under the posterior distribution. For this, our method uses the Metropolis-Hastings algorithm. The main idea is, that instead of sampling from the posterior $p(\theta|I)$ directly, we sample from a proposal distribution $Q(\theta'|\theta)$, which proposes a new sample θ' given the current sample θ. This generated proposal is then accepted or rejected using an acceptance criterion based on the unnormalized posterior probability $p(\theta|I)$.

The MH algorithm we propose to use for ASM fitting is shown in Algorithm 2. The individual terms in the ratio a, are given by the shape prior and the appearance model of the ASM. Note that we do not need to choose any parameter. The only part we have to provide is the proposal distribution Q. Also note that the denominator in (2), the normalization constant which is usually intractable, is equal for all θ and hence cancels in the ratio on line 4.

Algorithm 2. Metropolis Hastings sampling

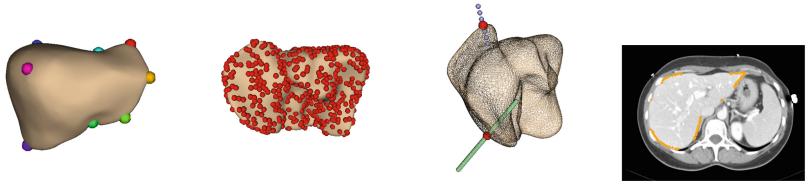
1: $\theta_0 \leftarrow$ arbitrary initialization
2: **for** $i = 1$ to S **do**
3: $\quad\theta' \leftarrow$ sample from $Q(\theta'|\theta)$
4: $\quad a \leftarrow \frac{q(\theta|\theta')p(\theta'|I)}{q(\theta'|\theta)p(\theta|I)} = \frac{q(\theta|\theta')p(\theta')p(I|\theta)}{q(\theta'|\theta)p(\theta)p(I|\theta)}$. {acceptance threshold}
5: $\quad r \leftarrow$ sample from $\mathcal{N}(0, 1)$
6: $\quad$**if** $a > r$ **then**
7: $\quad\quad\theta_i \leftarrow \theta'$
8: $\quad$**else**
9: $\quad\quad\theta_i \leftarrow \theta_{i-1}$
10: $\quad$**end if**
11: **end for**

The most commonly used proposal distribution is a random walk proposal, defined as $Q(\theta'|\theta) \sim N(\theta, \sigma^2 I)$. We change this to a block-wise proposal distribution, meaning that we update only either, α, ϕ, ψ, ρ or t at a time. Further, we use for each block a mixture of Gaussians with three different σ to account for the initial phase of convergence before we get the samples from the true posterior. When we generate a proposal, a part of the mixture is selected with a predefined probability ω.

4 Experimental Setup

In this section, we experimentally compare our proposed fitting method to the standard ASM method. For both methods, we use the same basic ASM. The methods are compared on the SLIVER dataset. In order to retain reproducibility, we use only the 20 livers from the training set with the provided ground-truth. We report the errors for the results of the model using all examples, also including the test item, as well as for all the leave-one-out setups.

We rigidly align the provided data using eight manually-clicked landmarks. The landmarks are shown in Fig. 1a on the surface of the handcrafted reference. We register the livers using the model-based registration presented by Lüthi et al. [9].



(a) landmarks (b) sampled points (c) profile points (d) annotations

Fig. 1. We show the reference liver with the annoated 8 landmarks in (a), the 1k sampled profile anchor points in (b), the locations for the appearance feature extraction (blue) and the sampled search points (green) for different profile anchor points (red) in (c) and the expert annotations in (d). (Color figure online)

We model the appearance at 1000 evenly sampled points (see Fig. 1b). The values of the image gradient are taken at seven points with a spacing of 8 mm to from the appearance feature (see Fig. 1c). The full appearance model consists of all 1000 individual local Gaussian appearance models[2].

As we aligned the SLIVER dataset initially before we built the models, we do not have to align the model to the data at test time. For the standard ASM fitting, we sample 61 search points over a distance of 60 mm around the current profile point location (see Fig. 1c). We choose $\kappa_T = 6$, $\kappa_S = 3$ and $\kappa_\alpha = 3$ to prevent unlikely updates.

For the sampling-based approach, we use the PDM prior from Eq. 3 and the appearance likelihood from Eq. 4. In the experiment including the expert annotated lines, we additionally include the line likelihood from Eq. 5 in the posterior. The expert annotations depicted in Fig. 1d mark parts of the organ boundary on three axis-aligned slices. We use the introduced multiscale, block-wise Gaussian distribution for generating the proposals (See footnote 2). As

[2] The code for the model adaptation is available online at github.com/unibas-gravis/probabilistic-fitting-ASM.

this proposal distribution is symmetric the correction term of the transition probability ratio cancels on line 4 in the Algorithm 2.

For the standard ASM fitting, we take the last state after a maximum of 1000 iterations or convergence. When sampling, we draw 10k samples and use the one with the highest posterior value as result. We use a higher number of samples compared to the standard fitting steps. This is motivated by the fact that the standard approach looks at the appearance of 61 locations per iteration while for one sample we evaluate only one. We report the dice coefficient, the bi-directional average surface distance, and the Hausdorff distance to compare the results.

5 Results

Simplicity: What is striking is the simplicity of the sampling-based approach. We can use the exact same model in both approaches. For the standard approach, we need to define a search strategy, choose the search distance, the point distance threshold, the feature distance threshold and the model coefficient threshold. In contrast, for the sampling, the posterior is completely specified by the model itself. We need only to define the proposal distribution. Further, the standard approach has a fixed search distance. In contrast, when we generate samples from the proposal distribution, the model deformations depends on the local variance of the PDM and hence is locally adaptive.

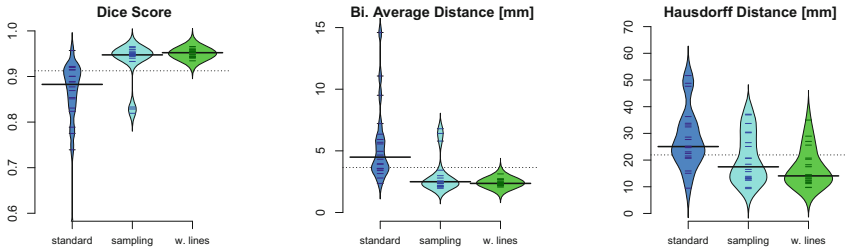


Fig. 2. Segmentation accuracy using the model including the test item. This figure shows that sampling finds a better segmentation compared to the standard ASM fitting. The sampling has the higher dice score as well as the lower bi-directional average distance value and Hausdorff distance. Including expert annotations further improves the result.

Full Model: In the Fig. 2, we show the evaluation of the results for the first experiment, where the target shape is contained in the model. For the Dice coefficient, the bi-directional average surface distance and the Hausdorff-distance one can observe that the sampling outperforms the standard fitting. Including also the expert annotations further improves the result.

Leave-One-Out: The leave-one-out experiments in Fig. 3 show the same trends. Comparing the values to the last experiment we can observe that the Hausdorff distance drops much more than the other measures. This was to be expected, as in this experiment the test item is excluded from the model, but often has a very specific local shape compared to the training items. The expert annotations have a stronger impact on the leave-one-out experiment. Note, also for this experiment, the used model for a specific test case is the same for all methods.

We conclude from the experiments that sampling is a better strategy to adapt an ASM to data. In addition, integrating additional constraints in a straightforward manner, we showed the example of expert annotations along the boundary, helps to improve the accuracy further.

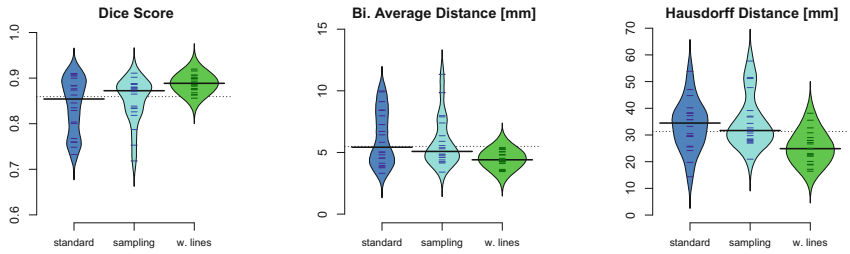


Fig. 3. Segmentation accuracy of the leave-one-out experiment. The sampling outperforms the standard ASM fitting approach. Again, the result further improves when expert annotations are provided.

6 Conclusion

We presented a fully probabilistic interpretation of ASM based segmentation as Bayesian inference. Using a Metropolis-Hastings sampling approach, we determine the maximum a posteriori segmentation. Our method is simple to implement and leads to better results compared to the standard ASM algorithm. As all the terms in our posterior formulation are motivated by the model, no arbitrary thresholds are needed. Furthermore, the probabilistic formulation provides a principled way of integrating additional information, such as expert annotations. For future work, additional constraints, such as regions in the image, which the fitting result should not enter could be integrated. The crucial component for the performance of our method is the proposal distribution. The better it reflects the (unknown) target distribution, the more efficient the sampling is while convergence is always guaranteed asymptotically. Smarter choices of the proposal distribution than the used random walk proposals, which take the image intensities into account, could improve the convergence rate. Such proposals could be based on random forest regression steps [8], on a global estimated parameter distribution [6] or even include deep learning.

Acknowledgment. This work was supported by the Innosuisse project 25622.1 PFLS-LS.

References

1. Cootes, T., Baldock, E., Graham, J.: An introduction to active shape models. In: Image Processing and Analysis, pp. 223–248 (2000)
2. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models-their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995)
3. Esfandiarkhani, M., Foruzan, A.H.: A generalized active shape model for segmentation of liver in low-contrast CT volumes. Comput. Biol. Med. **82**, 59–70 (2017)
4. van Ginneken, B., de Bruijne, M., Loog, M., Viergever, M.A.: Interactive shape models. In: Medical Imaging 2003: Image Processing, vol. 5032, pp. 1206–1217. International Society for Optics and Photonics (2003)
5. Heimann, T., van Ginneken, B., et al.: Comparison and evaluation of methods for liver segmentation from CT datasets. IEEE Trans. Med. Imaging **28**(8), 1251–1265 (2009). https://doi.org/10.1109/TMI.2009.2013851
6. Jampani, V., Nowozin, S., Loper, M., Gehler, P.V.: The informed sampler: a discriminative approach to Bayesian inference in generative computer vision models. Comput. Vis. Image Underst. **136**, 32–44 (2015)
7. Kirschner, M., Becker, M., Wesarg, S.: 3D active shape model segmentation with nonlinear shape priors. In: Fichtinger, G., Martel, A., Peters, T. (eds.) MICCAI 2011. LNCS, vol. 6892, pp. 492–499. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-23629-7_60
8. Lindner, C., Thiagarajah, S., Wilkinson, J., Consortium, T., Wallis, G., Cootes, T.: Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE Trans. Med. Imaging **32**(8), 1462–1472 (2013)
9. Lüthi, M., Gerig, T., Jud, C., Vetter, T.: Gaussian process morphable models. IEEE Trans. Pattern Anal. Mach. Intell. **40**, 1860–1873 (2017)
10. Norajitra, T., Maier-Hein, K.H.: 3D statistical shape models incorporating landmark-wise random regression forests for omni-directional landmark detection. IEEE Trans. Med. Imaging **36**(1), 155–168 (2017)
11. Schönborn, S., Egger, B., Morel-Forster, A., Vetter, T.: Markov chain Monte Carlo for automated face image analysis. Int. J. Comput. Vis. **123**(2), 160–183 (2017). https://doi.org/10.1007/s11263-016-0967-5
12. Van Ginneken, B., Frangi, A.F., Staal, J.J., ter Haar Romeny, B.M., Viergever, M.A.: Active shape model segmentation with optimal features. IEEE Trans. Med. Imaging **21**(8), 924–933 (2002)
13. Wimmer, A., Soza, G., Hornegger, J.: A generic probabilistic active shape model for organ segmentation. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 26–33. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04271-3_4
14. Zhang, Q., Bhalerao, A., Helm, E., Hutchinson, C.: Active shape model unleashed with multi-scale local appearance. In: 2015 IEEE International Conference on Image Processing (ICIP), pp. 4664–4668. IEEE (2015)