
Proceedings Paper:
García, H.F., Álvarez, M.A. orcid.org/0000-0002-8980-4472 and Orozco, Á.A. (2017) Bayesian optimization for fitting 3D morphable models of brain structures. In: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2016. 21st Iberoamerican Congress, CIARP 2016, 08/11/2016-11/11/2016, Lima, Peru. Lecture Notes in Computer Science, 10125 . Springer Verlag , pp. 291-299. ISBN 9783319522760

https://doi.org/10.1007/978-3-319-52277-7_36

Bayesian Optimization for Fitting 3D Morphable Models of Brain Structures

Hernán F. García, Mauricio A. Álvarez and Álvaro A. Orozco

Grupo de Investigación en Automática, Universidad Tecnológica de Pereira, Pereira, Colombia,
{hernan.garcia,malvarez,aaog}@utp.edu.co

Abstract. Localize target areas in deep brain stimulation is a difficult task, due to the shape variability that brain structures exhibit between patients. The main problem in this process is that the fitting procedure is carried out by a registration method that lacks of accuracy. In this paper we proposed a novel method for 3D brain structure fitting based on Bayesian optimization. We use a morphable model in order to capture the shape variability in a given set of brain structures. Then from the trained model, we perform a Bayesian optimization task with the aim to find the best shape parameters that deform the trained model, and fits accurately to a given brain structure. The experimental results show that by using an optimization framework based on Bayesian optimization, the model performs an accurate fitting over cortical brain structures (thalamus, amygdala and ventricle) in comparison with common fitting methods, such as iterative closest point.

Keywords: Bayesian optimization, 3D Brain structures, Shape fitting, Morphable model.

1 Introduction

Deep brain stimulation surgery (DBS) in Parkinson's disease, is a surgical procedure used to treat the most common neurological symptoms such as stiffness, slowed movement, tremor and walking problems. In this procedure, the neurosurgeons perform an electrical stimulation of the basal ganglia area (thalamus, ventricles and sub-thalamus) by placing a micro-electrode device over this region [1]. However, a misplacement of the final position of this micro-electrode can induce negative effects in the post-surgery, like abnormal postures, loss of speech and even loss of mobility, among others symptoms [2]. That is why the neurosurgeons need to locating with high accuracy, the cortical brain structures related with the basal ganglia area [3]. The most common planning DBS surgery methods include a registration process in which a given brain atlas (volume of labeled brain structures with respect to a healthy patient), is fitted into the magnetic resonance images (MRI) of the patient to be treated. This step allows to the neurosurgeons localize the target area in the MRI volume [4]. The main problem of these approaches is that the brain atlases used to estimate the target brain structures lack of generalizability (since we need to modeling the brain volumetry in a given population), and the fitting accuracy of the atlases in the brain volumes is still an open research topic [5]. Therefore, methods that can be able to capture the shape variability over a set of MRI

volumes could improve the robustness (generalizability and accuracy) of the target area localization [3].

The model-based approaches such as morphable models (MM), make use of the prior knowledge of the shape variability in a set of images (MRI volumes) and typically finds the best match between the model and the new image (brain structure) [6]. In medical image analysis, the use of prior information (shape contour of the brain structures labeled by a medical specialist), combined with some model that can be able to represent a shape contour, leads to an accurate fitting of a given brain structure [3]. A commonly used method for fitting morphable models, is the Iterative Closest Point (ICP) [6]. This method establishes closest point associations from one shape to another, and finds a rigid transformation that minimizes an Euclidean error between shapes. However, due to the large deformations that the brain volumes present over different patients (i.e. different thickness for ventral brain structures), the performance of the fitting process becomes inaccurate [7]. The reason for this performance, is that the global minimum of the cost function that measures the fitting process (mean square error between the deformed and the target brain structure), often does not corresponds to the optimal fitting (That is, a rigid transform only estimates the scale, rotation and translation parameters, discarding the shape parameters that control the deformations) [8].

Bayesian optimization (BO) provides an elegant approach for the global optimization problem, in which a given cost function is minimized in a probabilistic way [9]. For continuous functions, Bayesian optimization assumes that the unknown function is sampled from a Gaussian process, and maintains a posterior distribution for this function as observations [10]. In the fitting process of the morphable model, these observations are the measure of generalization performance (matching accuracy) under different settings of the hyperparameters that we wish to optimize (shape parameters that control the deformations) [11]. In this paper we propose a Bayesian optimization framework for fitting 3D morphable models of brain structures. We use as morphable model, a point distribution model (PDM) in order to capture the shape variability in a set of brain volumes (brain structures related with Parkinson's disease). From the trained model, we perform a fitting process based on Bayesian optimization in order to find the optimal match between the morphable model and a given brain structure. The main contribution of our work, is based on the Bayesian optimization process that finds the shape parameters that control the deformations of a morphable model in a probabilistic way.

2 Materials and Methods

2.1 Database

In this work we use a MRI database from the Universidad Tecnológica de Pereira (DB-UTP). This database contains recordings of MRI volumes from ten patients with Parkinson's disease. The database was labeled by neurosurgeons from the Institute of Parkinson and Epilepsy of the Eje Cafetero, located in Pereira-Colombia. The database contains $T1$, $T2$ and CT sequences with $1\ mm$ x $1\ mm$ x 1 mm voxel size and slices of 512x512 pixels. Also a set of detailed label maps related to the brain structures (i.e. basal ganglia area) are provided. Moreover, three-dimensional models of the labeled anatomical structures for each patient are available in the dataset.

2.2 3D Brain Model

We use a 3D brain model based on a PDM to represent the shape of the brain structures. The shape information is captured by the vertexes information (point-cloud data in the $\mathbb{R}^3$) from the mesh data that represents each brain structure. Besides, the model uses statistical information of the shape variation across the training set in order to model the brain volumetry of a given brain structure [6]. The PDM is a parametrized model, $\mathbf{S} = \gamma(\mathbf{b})$, where $\mathbf{S} = [\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N]$, with $\mathbf{w}_i = [x_i, y_i, z_i]^\top, \quad \mathbf{w}_i \in \mathbb{R}^{3\times 1}$ representing each landmark (point data in $3D$ space). The vector $\mathbf{b}$ holds the parameters which can be used to vary the shape across the surface and γ defines the function over the parameters. We use N landmarks representing the points of the surface related to a given brain structure, from a training set of L brain meshes, where each shape is $\mathbf{S}^k = \left[\mathbf{w}_1^k, \mathbf{w}_2^k, \ldots, \mathbf{w}_N^k\right], \quad \mathbf{S}^k \in \mathbb{R}^{3\times N}$.

In order to eliminate the global transformations for the training shapes, we use Procrustes analysis [6]. The alignment process is carried out by minimizing the square-distance of each shape $\mathbf{S}_k$ with respect to their mean $\bar{\mathbf{S}} = \frac{1}{N}\sum_{k=1}^{L}\mathbf{S}_k$ ($\bar{\mathbf{S}}$ is scaled such that $|\mathbf{S}| = 1$). We use Principal Component Analysis (PCA) to model the shape variations of the brain structures in the training set. The model estimate these variations by computing the eigenvalues and eigenvectors of the covariance matrix of the training set defined by, $\mathbf{C} = \frac{1}{L-1}\sum_{k=1}^{L}(\mathbf{s}_k - \bar{\mathbf{s}})(\mathbf{s}_k - \bar{\mathbf{s}})^\top$, where $\mathbf{s}_k$ and $\bar{\mathbf{s}}$ are the vectorized forms of the shape $\mathbf{S}_k$ and the mean shape $\bar{\mathbf{S}}$ respectively. Les us define $\boldsymbol{\phi}_i$ and λ_i as the ith eigenvector and eigenvalue of the covariance matrix $\mathbf{C}$. If $\boldsymbol{\Phi}$ holds the t eigenvectors corresponding to the largest eigenvalues, a given plausible shape (similar to the brain structures in the training set) can be estimated by

$$\hat{\mathbf{s}} \approx \bar{\mathbf{s}} + \boldsymbol{\Phi}\mathbf{b}, \tag{1}$$

where $\boldsymbol{\Phi} = (\boldsymbol{\phi}_1|\boldsymbol{\phi}_2|\ldots|\boldsymbol{\phi}_t)$ and $\mathbf{b}$ is a t dimensional vector representing the shape parameters (those who controls the shape variability). The value of t is chosen such that the model represents the 98% of the shape variance [6].[1]

2.3 Bayesian optimization with Gaussian process priors

Since we want to compute the model parameters in a probabilistic way, the goal is to find the minimum of a cost function $f(\mathbf{x})$ (i.e. Euclidean distance between the ground truth landmarks and the landmarks deformed by the 3D-MM model) on some bounded set $\mathcal{X}$ that controls the shape variations. To this end, Bayesian optimization constructs a probabilistic framework for $f(\mathbf{x})$ with the aim to exploit this model to make predictions of the shape parameters $\mathcal{X}$ evaluated in the cost function [9].[2]

The main components of the Bayesian optimization framework are the prior over the function being optimized and the acquisition function that will allow us to determine the next point to evaluate the cost function [10]. In this work we use Gaussian process prior,

[1] The variance of the ith parameter, b_i, across the training set is given by λ_i.

[2] The main reason of the Bayesian optimization framework is to use all of the available information from previous evaluations of $f(\mathbf{x})$.

due to its flexibility and tractability. A Gaussian Process (GP) is an infinite collection of scalar random variables indexed by an input space such that for any finite set of inputs $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$, the random variables $\mathbf{f} \triangleq [f(\mathbf{x}_1), f(\mathbf{x}_2), \cdots, f(\mathbf{x}_n)]$ are distributed according to a multivariate Gaussian distribution $\mathbf{f}(\mathbf{X}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}'))$. A GP is completely specified by a mean function $m(\mathbf{x}) = \mathbb{E}[f(\mathbf{X})]$ (usually defined as the zero function) and a positive definite covariance function given by $k(\mathbf{x}, \mathbf{x}') = \mathbb{E}\left[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))^\top\right]$ (see [12] for further details).

Let us assume that $f(\mathbf{x})$ is drawn from a Gaussian process prior and that our observations are of the form $\{\mathbf{x}_n, y_n\}_{n=1}^N$, where $y_n = \mathcal{N}(f(\mathbf{x}_n), v))$ and v is the variance of noise introduced into the function observations. The acquisition function is denoted by $u : \mathcal{X} \in \mathbb{R}^+$ and establish the point in $\mathcal{X}$ that is evaluated in the optimization process as $x_{\text{next}} = \arg\max_{\mathbf{x}} u(\mathbf{x})$. Since the acquisition function depends on the GP hyperparameters, $\boldsymbol{\theta}$, and the predictive mean function $\mu(\mathbf{x}, \{\mathbf{x_n}, \mathbf{y_n}\}, \boldsymbol{\theta})$ (as well as the predictive variance function), the best current value is then $x_{\text{best}} = \arg\min_{\mathbf{x}_n} f(\mathbf{x}_n)$.

2.4 Morphable Model Fitting Using Bayesian Optimization

Our approach is based on the Bayesian optimization process for estimating the shape parameters that fit accurately a given brain structure in a probabilistic way. In this work, we used the 3D models of the labeled anatomical structures for the ten patients of the dataset. We trained three PDM models by capturing the shape variability for the thalamus, amygdala and ventricles (due to their importance in the Parkinson's disease). Since the brain structures are non-rigid shapes and their size vary along the training set, we need to decimate all the 3D shapes (with respect to the smallest brain structure) to perform the eigenvalue decomposition. We use *leave-one-out* validation (we train the 3D-MM with $L-1$ brain structures) to measure the fitting accuracy. Besides, we initialize the fitting process by performing rigid iterative closest point (ICP) in order to find the scale, rotation and translation parameters between the model and the given brain structure. The main reason for this prior initialization is to let the BO process explore the shape parameters that deform the PDM model. For the BO process we use as cost function, the Euclidean distance between the landmarks of the model and the ground truth landmarks of the brain structure. We use the $GPyOpt$[3] toolbox for python. In this work, we report results for the expected improvement (EI), and the probability of improvement (PI) acquisition functions [9]. Figure 1 shows the block diagram of the proposed model used in this work.

3 Results and Discussions

In the following sections we show the results for our BO framework for fitting 3D morphable models of brain Structures. On section 3.1, we show the results of the 3D shape modeling for the brain structures using a PDM as morphable model. Section 3.2 shows the results for the fitting process of the morphable model using BO.

[3] *Gpyopt* is a BO framework in python available at `http://github.com/SheffieldML/GPyOpt`

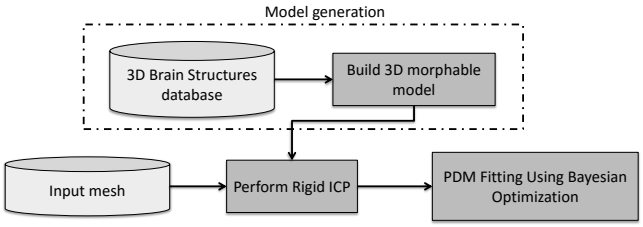
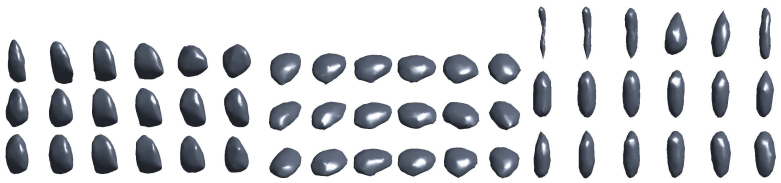


Fig. 1. Block diagram of the proposed model fitting approach based on BO.

3.1 Training results for 3D-MM

Figure 2 shows the training process results for the morphable models. From the figure, it can be noticed that for the three brain structures (thalamus, amygdala and ventricle), the morphable models capture the shape variability along the training set. Moreover, the results also show that by modeling the covariance matrix of the training dataset using PCA, the morphable model capture the relevant information in the latent space by analyzing the shape parameters that controls the deformation of a given brain structure. Besides, figure 2 shows that by changing the shape parameters (eigenvalues of the PDM) the models deform a given brain structure from thin shapes (upper left corner of the figure 2(c)) to curvy shapes (lower right corner of the figure 2(c)). This shape variability can be related with the range of ages of the subjects in the database(between $35 - 65$ years), due to the fact that the brain volume decreases their mass over time in patients with Parkinson's disease (thin shapes modeled by the first eigenvalues of the morphable models) [2].



(a) PDM for the thalamus (b) PDM for the amygdala (c) PDM for the ventricle

Fig. 2. Effects of varying the first 3 shape parameters of the PDM for the analyzed brain structures. Figures 2(a), 2(b) and 2(c) show the model variation for the most relevant eigenvalues: (top) $b_1 = \left\{-3\sqrt{\lambda_1}, \ldots, 3\sqrt{\lambda_1}\right\}$; (middle) $b_2 = \left\{-3\sqrt{\lambda_2}, \ldots, 3\sqrt{\lambda_2}\right\}$; (bottom) $b_3 = \left\{-3\sqrt{\lambda_3}, \ldots, 3\sqrt{\lambda_3}\right\}$. Each shape parameter (structures for each row in the subfigures) ranges from $-3\sqrt{\lambda_i}$ starting with the left column till $3\sqrt{\lambda_i}$ for the right column.

3.2 Fitting Results Using Bayesian Optimization

From the trained morphable models, we deform each PDM in order to match a target brain structure. The BO process estimate the best shape parameters that fit each

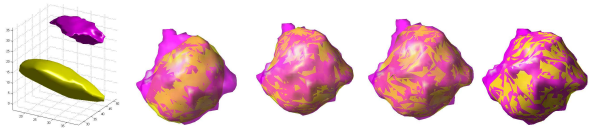
of the brain structures (amygdala, thalamus and ventricle). Figure 3 shows the results for the fitting process using BO. The results show that by using a fully Bayesian treatment of the optimization process, the model can estimate the shape parameters that performs accurately the fitting process between shapes. Figure 3 shows that by adding a prior registration process through rigid ICP, the BO method can explore an exploit the probabilistic search of the shape parameters that minimizes the cost function (Euclidean distance between the deformed model landmarks and the target shape). Besides, the results in figure 3(a) prove that the optimization step estimates the plausible shape parameters that deforms a given brain structure, even if the target shape has large curvatures (significant changes over the shape surface, see figures 3(c) and 3(e)). Beside, the figure 4 shows the convergence of the BO process for both the EI and the PI acquisition functions. The results show that EI has better global convergence, due to the fact that by using this acquisition function the model fits more accurately the target shape and take less iteration to converge (200 iterations for the EI and more than 300 iterations for the PI). The main reason is that the EI estimates best shape parameters in the exploration step than the PI acquisition function. However, both acquisition functions exploit the probabilistic search of the shape parameters with lower values in the evaluated cost function (mean square error of the Euclidean distance, 52 for EI and 60 for the PI). Finally, table 1 shows the accuracy of the BO process compared against the common fitting method such as Rigid ICP. The results show that by using a given acquisition function the optimization process improves the selection of those values that control the deformation with high variance (shape parameters in regions not well explored) and values with high mean value (shape parameters worth exploiting that increase the fitting accuracy, MSE error of 10.295 for the amygdala) in comparison with the Rigid ICP which only removes the translation, rotation and scale between two shapes (MSE error of 23.971 for the amygdala).

Table 1. Accuracy of the BO process for fitting the three morphable models using EI and PI acquisition functions. The table shows the mean square error of the Euclidean distance between the deformed model and the target brain structure.

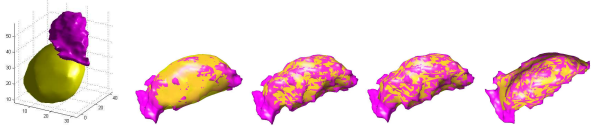
	Fitting method		
Brain Model	**Rigid ICP**	**BO with EI**	**BO with PI**
Amygdala	23.971 ± 4.163	10.295 ± 2.575	13.927 ± 4.665
Thalamus	32.486 ± 4.169	19.907 ± 1.340	22.351 ± 4.170
Ventricle	26.623 ± 4.287	15.924 ± 2.079	16.686 ± 4.289

4 Conclusions and Future Works

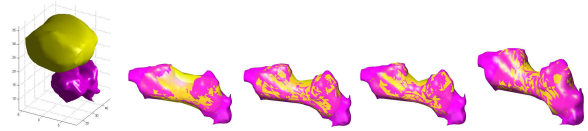
In this paper we propose a Bayesian optimization framework for fitting 3D morphable models of brain structures. Our method deforms a trained point distribution model in order to match a given brain structure. Besides, the shape parameters that control the model variations (relevant eigenvalues derived from the training set), are optimized in a



(a) Initial shapes (b) Fitting process for the amygdala

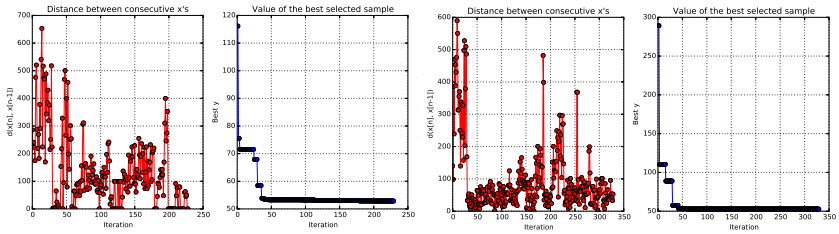


(c) Initial shapes (d) Fitting process for the thalamus



(e) Initial shapes (f) Fitting process for the ventricle

Fig. 3. Fitting process for the PDM models of the amygdala, thalamus and ventricle brain structures (shapes with yellow color are related to the model and those with color magenta depict the target shapes). The figures show the model to be deformed and the target brain structure. Also the figure shows the deformed models at 10, 50, 100 and 200 iterations.



(a) Experiment using EI (b) Experiment using PI

Fig. 4. Convergence of the BO process for the EI (up) and PI (down) acquisition functions. The figure shows the distance between consecutive selected $\mathbf{x}$ values (left column), the mean of the current model in the selected sample (middle) and the variance of the model in the selected sample (right column).

probabilistic way. The results show that placing a Gaussian process prior over the function being optimized, the proposed model can estimate the best parameters that perform the fitting process. Moreover, by using the *Expected improvement* as acquisition function, the optimization process improves the matching accuracy by exploiting the model parameters bounded by the eigenvalues of the training set.

As future works, we want to analyze the Bayesian optimization framework in high dimensional problems that includes the whole optimization process of a raw point-cloud data. Finally, we want to analyze a morphable model that adds both shape an appearance information in order to model tissue properties related to a given brain structure.

Acknowledgments

This research is developed under the project with code $FP44842-584-2015$, financed by Colciencias and SINCLECLICK SOLUTIONS S.A.S. H.F. García is funded by Colciencias under the program: *Formación de alto nivel para la ciencia, la tecnología y la innovación - Convocatoria 617 de 2013*.

References

1. Hill, D.: Neuroimaging to assess safety and efficacy of ad therapies. Expert Opinion on Investigational Drugs **19**(1) (2010) 23–26 PMID: 19947893.
2. Ibarretxe-Bilbao, N., Tolosa, E., Junque, C., Marti, M.J.: Mri and cognitive impairment in parkinson's disease. Movement Disorders **24**(S2) (2009) S748–S753
3. Cabezas, M., Oliver, A., Lladó, X., Freixenet, J., Cuadra, M.B.: A review of atlas-based segmentation for magnetic resonance brain images. Computer Methods and Programs in Biomedicine **104**(3) (2011) e158 – e177
4. Cosa, A., Canals, S., Valles-Lluch, A., Moratal, D.: Unsupervised segmentation of brain regions with similar microstructural properties: Application to alcoholism. In: Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE. (July 2013) 1053–1056
5. Sotiras, A., Davatzikos, C., Paragios, N.: Deformable Medical Image Registration: A Survey. Medical Imaging, IEEE Transactions on **32**(7) (July 2013) 1153–1190
6. Nair, P., Cavallaro, A.: 3-d face detection, landmark localization, and registration using a point distribution model. IEEE Trans. Multimedia **11**(4) (2009) 611–623
7. Sjöberg, C., Ahnesjö, A.: Multi-atlas based segmentation using probabilistic label fusion with adaptive weighting of image similarity measures. Computer Methods and Programs in Biomedicine **110**(3) (2013) 308 – 319
8. Sjöberg, C., Johansson, S., Ahnesjö, A.: How much will linked deformable registrations decrease the quality of multi-atlas segmentation fusions? Radiation Oncology **9**(1) (2014) 1–8
9. Snoek, J., Larochelle, H., Adams, R.P.: Practical bayesian optimization of machine learning algorithms. In: Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems 2012. Proceedings of a meeting held December 3-6, 2012, Lake Tahoe, Nevada, United States. (2012) 2960–2968
10. Srinivas, N., Krause, A., Seeger, M., Kakade, S.M.: Gaussian process optimization in the bandit setting: No regret and experimental design. In FÃ$\frac{1}{4}$rnkranz, J., Joachims, T., eds.: Proceedings of the 27th International Conference on Machine Learning (ICML-10), Omnipress (2010) 1015–1022
11. Wang, Z., Zoghi, M., Hutter, F., Matheson, D., de Freitas, N.: Bayesian optimization in high dimensions via random embeddings. In: International Joint Conferences on Artificial Intelligence (IJCAI) - Distinguished Paper Award. (2013)
12. Carl Edwards Rasmussen, C.W.: Gaussian Processes for Machine Learning. The MIT Press (2006)