

# Generalized Multiresolution Hierarchical Shape Models via Automatic Landmark Clusterization

Juan J. Cerrolaza<sup>1</sup>, Arantxa Villanueva<sup>2</sup>, Mauricio Reyes<sup>3</sup>, Rafael Cabeza<sup>2</sup>, Miguel A. González Ballester<sup>4</sup>, and Marius George Linguraru<sup>1,5</sup>

<sup>1</sup> Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Health System, Washington DC, USA, {JCerrola, MLinguraru}@cnmc.org

<sup>2</sup> Public University of Navarra, Pamplona, Spain

<sup>3</sup> Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland

<sup>4</sup> ICREA, Barcelona, Spain – Universitat Pompeu Fabra, Barcelona, Spain

<sup>5</sup> School of Medicine and Health Sciences, George Washington Univ. Washington, DC, USA

**Abstract.** Point Distribution Models (PDM) are some of the most popular shape description techniques in medical imaging. However, to create an accurate shape model it is essential to have a representative sample of the underlying population, which is often challenging. This problem is particularly relevant as the dimensionality of the modeled structures increases, and becomes critical when dealing with complex 3D shapes. In this paper, we introduce a new generalized multiresolution hierarchical PDM (GMRH-PDM) able to efficiently address the high-dimension-low-sample-size challenge when modeling complex structures. Unlike previous approaches, our new and general framework extends hierarchical modeling to any type of structure (multi- and single-object shapes) allowing to describe efficiently the shape variability at different levels of resolution. Importantly, the configuration of the algorithm is automatized thanks to the new agglomerative landmark clustering method presented here. Our new and automatic GMRH-PDM framework performed significantly better than classical approaches, and as well as the state-of-the-art with the best manual configuration. Evaluations have been studied for two different cases, the right kidney, and a multi-object case composed of eight subcortical structures.

**Keywords:** Shape models, multiresolution, hierarchical models, PDM.

## 1 Introduction

Since their inception in the early 1990s, active shape models (ASM) [1] have proven effective for addressing a number of problems where the target structures are consistent in shape but poorly defined by image features, as is often the case in medical images. The success of point distribution models (PDM)-based matching approaches depends on an accurate description of the shape class, the expected shape instances, and their variations. While a limited number of examples may be sufficient when working with relatively simple objects, an adequately large training set is not always available as the dimensionality and complexity of the structures increase, as is usually

the case when working with 3D multi-object structures. This issue is known as the high-dimension-low-sample-size (HDLSS) problem. Trying to overcome this question, some authors have proposed interesting versions of the classical PDM, exploiting the possibilities of incorporating multiresolution-based hierarchical analysis to shape modeling. Davatzikos et al. [2] proposed the hierarchical decomposition of the shape into small pieces of information via the wavelet transform. However, whereas the independent modeling of these bands allows reducing the dimensionality of the problem, and thus the HDLSS effect, it also reduces the robustness of the model as shown in [3]. An interesting attempt to describe the interrelationships between objects at different scales statistically is the multiscale framework proposed by Lu et al. [4], using m-reps as the geometric representation of shapes. In spite of the valuable multiscale properties of m-reps, they are less intuitive than the landmarks-based representation used in PDMs, which is probably one of the simplest and most generic methods used to represent shapes. Yokota et al. [13] proposed an interesting hierarchical statistical model of the femur and pelvis, imposing additional connectivity constraints to control the matching between different subparts. In the recent work of Cerrolaza et al. [3,5], a new multiresolution hierarchical variant of PDM (MRH-PDM) was introduced, able to efficiently characterize the different inter-object relationships, as well as the particular locality of each element separately. Even though the potential of this new method in terms of accuracy and robustness improvement was successfully verified, there are two main drawbacks that limit its practical application. First, the absence of an automatic grouping approach can hinder its use when working with complex data with a large number of objects. On the other hand, the hierarchical decomposition is limited to multi-object structures, since no intra-object analysis is considered within the original framework.

In this paper, we propose a new Generalized Multiresolution Hierarchical PDM (GMRH-PDM) that addresses these two important issues, automatic grouping and intra-object analysis. The new notation introduced in Section 2 extends the hierarchical modeling of PDM even to single-object structures, which leads to a more versatile and generalizable framework. Finally, the configuration of the algorithm (i.e., the definition of clusters at each resolution) is automatized thanks to the new agglomerative landmark clustering approach described in Section 3. The performance of the new GMRH-PDM method is studied for two different cases, the right kidney, and a multi-object case composed of eight subcortical structures.

## 2 Generalized Multiresolution Hierarchical PDM

In this section we present a new generalization of the original MRH-PDM formulation described by Cerrolaza et al. [3]. In their work the capability to model variability in subparts of a single object was limited, as they considered the single objects as the simplest structure to model at the finest resolution levels. Relaxing this condition, we go one step further in the development of hierarchical PDMs, introducing a more general framework where any possible grouping of landmarks is considered.

Let  $\mathbf{x}$  be the vector form of a 3D shape defined by  $K \in \mathbb{N}$  landmarks. In the general case of a multi-object shape composed of  $M$  ( $M \in \mathbb{N}$ ) single-object structures,  $\mathbf{x}_j$  ( $1 \leq j \leq M$ ),  $\mathbf{x}$  is defined by the concatenation of the 3 coordinates of the  $K_j \in \mathbb{N}$  landmarks ( $K = \sum K_j$ ) that define each object, i.e.  $\mathbf{x} = (\mathbf{x}_1; \dots; \mathbf{x}_M)^T$ . Using the multi-object generalization [5] of the matrix notation initially proposed by Lounsbery et al. [6], the multiresolution analysis of  $\mathbf{x}$  can be formulated as:

$$\mathbf{x}^r = \mathbf{A}^r \mathbf{x}^{r-1} \quad (1)$$

$$\mathbf{z}^r = \mathbf{B}^r \mathbf{x}^{r-1} \quad (2)$$

where  $r \in \mathbb{N}$  indicates the level of resolution (in particular  $r = 0$  defines the finest level of resolution, and thus,  $\mathbf{x}^0 = \mathbf{x}$ ), and  $\mathbf{A}^r$  and  $\mathbf{B}^r$  represent the analysis filters. Equation (1) implements the filtering and downsampling of  $\mathbf{x}^{r-1}$ , providing a lower resolution version of it (i.e.,  $K^{r-1} > K^r$ , where  $K^r \in \mathbb{N}$  represents the number of landmarks at the resolution level  $r$ ), while  $\mathbf{z}^r$  (2) captures the lost detail between  $\mathbf{x}^r$  and  $\mathbf{x}^{r-1}$ . An optimal selection of these analysis filters guarantees that no information is lost during the process, being possible to reverse the analysis process with synthesis equation:  $\mathbf{x}^{r-1} = \mathbf{F}^r \mathbf{x}^r + \mathbf{G}^r \mathbf{z}^r$ . Lounsbery et al. [6] provide a general multiresolution framework to compute the analysis and synthesis filters ( $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{F}$ , and  $\mathbf{G}$ ) for meshes with subdivision connectivity and arbitrary topology. In this work, we define the multiresolution domain using the octahedron as the reference mesh [8], with a 4-to-1 splitting step, and a lifter butterfly scheme for triangular meshes [7].

With this method described above, it is possible to decompose any multi object structure into different levels of resolution. Whereas MRH-PDM established a specific division of the  $M$  objects into  $M^r \in \mathbb{N}$  disjoint subsets at each level or resolution (i.e., only complete objects can be part of a subset), here we propose a more general definition of the disjoint subsets allowing any type of grouping between the total set of landmarks. Thus, at each level of resolution  $r$  we define a particular division of the  $K^r$  landmarks into  $M^r$  separate clusters,  $(\mathcal{S}_1^r, \dots, \mathcal{S}_{M^r}^r)$ , where  $\mathcal{S}_s^r$  ( $s = 1, \dots, M^r$ ) is formed by the indices of the landmarks contained in this subset, and therefore,  $\bigcap_{s=1}^{M^r} \mathcal{S}_s^r = \emptyset$  and  $\bigcup_{s=1}^{M^r} \mathcal{S}_s^r = (1, \dots, M)$ . In addition, we impose the following condition. Suppose  $\mathcal{S}_{s^{r-1}}^{r-1}(i)$  represents the  $i$ -th element of the  $s^{r-1}$ -th subset defined at the  $r-1$ -th resolution level, and  $\widehat{\mathcal{S}}_{s^r}^r$  be the propagation of  $\mathcal{S}_{s^r}^r$  to  $r-1$ , then

$$\mathcal{S}_{s^{r-1}}^{r-1}(i) \in \widehat{\mathcal{S}}_{s^r}^r \implies \forall \mathcal{S}_{s^{r-1}}^{r-1} \subseteq \widehat{\mathcal{S}}_{s^r}^r \quad (3)$$

That is, two sets of landmarks that have been grouped separately at a specific level of resolution, should not be jointly modeled at finer levels; or equivalently, the clusters created in the resolution  $r$ -th derived from the fragmentation of clusters in resolution  $r + 1$ -th (see Fig. 1). Despite the intuitive meaning of (3), there is a challenge yet to be resolved: the propagation of the clusters between two consecutive resolutions. Let  $\mathbf{l}^r$  be a  $(3K^r \times 1)$  vector (i.e., the same size as  $\mathbf{x}^r$ ) containing the labels of the subset to which each landmark of  $\mathbf{x}^r$  belongs; i.e., if  $\mathbf{x}^r(i) \in \mathcal{S}_s^r$  then  $\mathbf{l}^r(i) = s$ . With this notation, we can estimate  $\widehat{\mathbf{l}}^{r-1}$ , the propagation of the subdivision defined by  $\mathbf{l}^r$  to the landmarks of the following resolution,  $r-1$ , by means of the synthesis matrix,  $\mathbf{F}^r$ .

### 3 Automatic Landmark Clustering Using Vector Fields

In this section we introduce a new landmark clustering approach that allows to define automatically the division of the landmarks into separate clusters at each resolution. The clustering process was initially inspired by the work presented by Roy et al. [9], which was originally conducted for vector field segmentation of moving objects in 2D videos, and extended to 3D objects by Reyes et al. [10] to study the anatomical variability of single organs via principal factor analysis. Here we propose a more general approach based on the agglomerative hierarchical clustering method presented by Ward [11], where the criterion for choosing the pair of clusters to merge at each step is based on the minimum value of the tailored objective function:

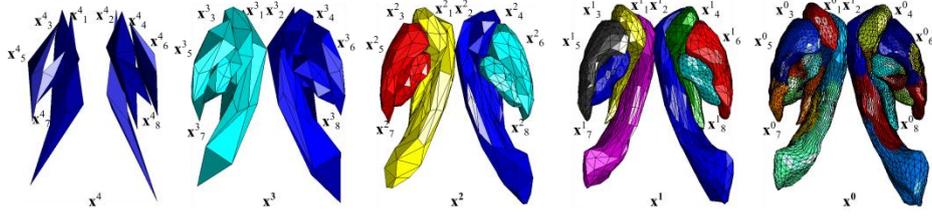
$$J(\Omega) = \alpha_1 \int_{\Omega} \left( \frac{|\mathbf{V}_{\Omega} \times \mathbf{V}_i|}{|\mathbf{V}_i|} \right)^2 \frac{L_{max}}{|\mathbf{V}_i|} di + \alpha_2 \left( 1 - \frac{\int_{\Omega} di}{\int_{\mathcal{S}} di} \right) + \alpha_3 H(\Omega) \quad (4)$$

where  $\alpha_1, \alpha_2$  and  $\alpha_3$  are real values such that  $\sum \alpha_i = 1$ .  $\Omega \subseteq \mathcal{S}$  represents a region or subdomain within the set of landmarks  $\mathcal{S}$  we want to divide into an optimal set of clusters. The first component of (4) takes into account the colinearity between deformation vectors within the domain  $\Omega$  and the predominant vector direction  $\mathbf{V}_{\Omega}$  in  $\Omega$ . Here, we define the deformation vector of landmark  $l_i \in \mathcal{S}$ ,  $\mathbf{V}_i$  as the sum of the eigenvectors obtained via PDM over  $\mathcal{S}$ , and weighted by their corresponding eigenvalues. Then  $L_{max} = \max_{\mathcal{S}} \{ \|\mathbf{V}_i\| \}$ , and  $\mathbf{V}_{\Omega}$  is defined as the highest eigenvalue of the matrix  $M(\Omega) = \int_{\Omega} \mathbf{V}_i \mathbf{V}_i^t di$ . The second term in (4) acts as a maximal area constraint. The aim of the third term,  $H(\Omega)$ , defined as the Hausdorff distance between the objects that compose  $\Omega$  normalized by the maximum distance among objects in  $\mathcal{S}$ , is to promote the grouping of objects that are spatially close. When minimizing equation (4), it is desirable that the colinearity between deformation vectors be the dominant term in the generation of clusters, while the second and third term act as additional constraint to guarantee the consistency of the final results, i.e.  $\alpha_1 \gg (\alpha_2, \alpha_3)$ .

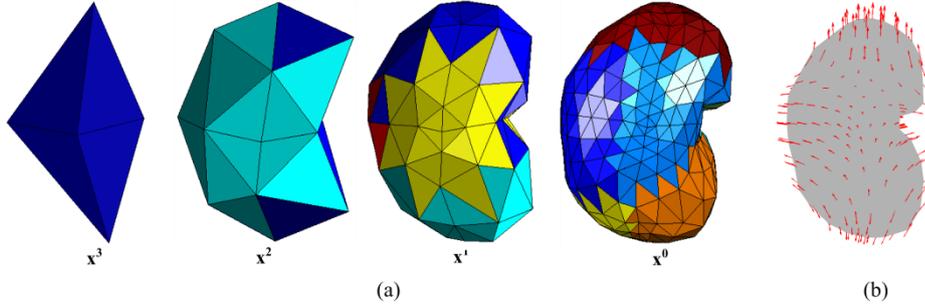
From the family of partitions provided by Ward's [11] algorithm, we define the optimal landmark division based on a tailored version of the Silhouette coefficient defined below. Suppose landmark  $l_i$  is assigned to cluster  $\Omega_i$ . Then, it is possible to define how well  $l_i$  is assigned to its cluster as  $a_i = J(\Omega_i) - J(\Omega_i \setminus l_i)$ , where  $\Omega_i \setminus l_i$  represents the cluster  $\Omega_i$  after removing  $l_i$ . Thus, large  $a_i$  represents a high dissimilarity between  $l_i$  and  $\Omega_i$ . In the same way, we define the dissimilarity of  $l_i$  to any other cluster  $\Omega_j (j \neq i)$  to which  $l_i$  is not member as  $b_i = \min\{b_{i,j}\}$ , where  $b_{i,j} = J(\Omega_j + l_i) - J(\Omega_j)$ , and  $\Omega_j + l_i$  represents the union of  $l_i$  and  $\Omega_j$ . Constraining the value of  $a_i$  and  $b_i$  to the range  $[0,1]$  by means of the logistic function,  $LF(\cdot)$ , we define the Silhouette coefficient for landmark  $l_i$ ,  $s_i$ , as

$$s_i = \frac{LF(b_i) - LF(a_i)}{\max\{LF(b_i), LF(a_i)\}} \quad (5)$$

Since a value of  $s_i$  close to 1 means that  $l_i$  is appropriately clustered in  $\Omega_i$ , the optimal clustering of  $\mathcal{S}$  will be the one that maximizes the average  $s_i$ .



**Fig. 1.** Hierarchical configuration provided by the GMRH-PDM algorithm for the multi-object structure composed by eight subcortical structures. At each level of resolution the set of landmarks depicted with the same color are modeled jointly via PDM. At resolution  $\mathbf{x}^1$  the lateral ventricles are in magenta ( $\mathbf{x}_1^1$ ) and navy ( $\mathbf{x}_2^1$ ), the caudate nuclei in yellow ( $\mathbf{x}_3^1$ ) and green ( $\mathbf{x}_4^1$ ), the putamen in black ( $\mathbf{x}_5^1$ ) and red ( $\mathbf{x}_6^1$ ), and the globus pallidi in blue ( $\mathbf{x}_7^1$ ) and cyan ( $\mathbf{x}_8^1$ ).



**Fig. 2.** (a) Hierarchical configuration obtained for the kidney model. At each level of resolution the area depicted with the same color is modeled jointly via PDM. (b) Deformation field defined by the first mode of variation when modeling all the landmarks jointly. The clustering configurations proposed by GMRH-PDM (especially in resolutions  $\mathbf{x}^1$  and  $\mathbf{x}^0$ ) are able to identify areas affected by a similar deformation field.

## 4 Shape Modeling via GMRH-PDM

Let  $\mathbf{x}$  be the vector form of any 3D structure (i.e., multi-object or single-object), whose multiresolution decomposition  $\{\mathbf{x} = \mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^R, \mathbf{z}^1, \dots, \mathbf{z}^R\}$ , can be obtained using (1) and (2). Imposing the initial condition that  $M^R = 1$  (i.e., a global statistical shape model of the whole set is built at the coarsest resolution in order to guarantee the coherent disposition of the elements), a new landmark subdivision scheme is calculated at resolution  $r-1$  for each of the  $M^r$  subsets ( $\mathcal{S}_s^r, s = 1, \dots, M^r$ ) obtained at  $r$ . Every new subdivision is obtained automatically using the landmark clustering approach introduced in Section 3. Finally, an efficient statistical model of the shape is created building a different PDM for every  $\mathcal{S}_s^{r-1}$  in which the structure has been divided, allowing to characterize different characteristics of the structure at each scale. Suppose now we want to use the new GMRH-PDM we just created to describe a new case,  $\mathbf{y}$ , i.e., finding the best approximation of  $\mathbf{y}$  in the subspace of allowed shapes described by the statistical model. Starting from the finest resolution,  $\mathbf{y}^0$  is divided into the  $M^0$  subsets previously defined, each of which is corrected by

the corresponding PDM. This process is repeated at each resolution until  $r = R$ . In the transition of each resolution, the high frequency component of the new constrained shape,  $\hat{\mathbf{z}}^1$ , will be used to recover the original resolution at the end of the process using the synthesis equation presented in Section 2.

## 5 Results and Discussion

To evaluate the performance of the new automatic GHMR-PDM approach we use two different datasets. First we use a set of 18 T1-weighted brain MRI volumes obtained from the Internet Brain Segmentation Repository (IBSR) [12] (pixel resolution  $0.94 \times 0.94 \times 1.5$  mm; volumes:  $256 \times 256 \times 128$  voxels). In particular, we work with a multi-object structure composed of eight subcortical structures ( $\mathbf{x}_1, \dots, \mathbf{x}_8$ ), corresponding to the left and right lateral ventricles, left and right caudate nuclei, left and right putamens, and left and right globus pallidi, respectively (Fig. 1). The performance of the new GMRH-PDM is also tested over a single-object database. We use a proprietary dataset of right kidneys from 18 CT abdominal studies (pixel resolution:  $0.58 \times 0.58 \times 1.00$  mm;  $512 \times 512 \times 360$  voxels). Following the general guidelines described in Section 3, the three configuration parameters of GMRH-PDM,  $\alpha_1, \alpha_2$  and  $\alpha_3$ , are set to 0.8, 0.1 and 0.1, respectively. Experimentally, we observed great similarity between the clusters obtained when  $\alpha_1 = [0.7 - 0.9]$  (using  $\alpha_2 = \alpha_3 = (1 - \alpha_1)/2$ ). For  $\alpha_1 < 0.7$ , the landmarks grouped into a single large cluster, being the second and third term of (4) which control the clusterization process. For  $\alpha_1 > 0.9$  landmarks are over-clustered due to the under-penalization of partitions.

The resulting automatic configurations for the multi-object and single-object case are shown in Figs. 1 and 2 respectively. The behavior of the new modeling approach is compared with two alternative methods for the multi-object case: the classical PDM [1], and the previous multiresolution hierarchical approach, MRH-PDM, proposed in [3, 5]. In particular, we chose the configuration that exhibited best results from all the hierarchical configurations manually defined in [5]. Due to the inability of MRH-PDM to deal with single-object structures, only PDM is considered in the comparison for the second/kidney data under study. The accuracy of the different methods to model new instances of the underlying population is evaluated in terms of the average landmark-to-landmark distance (L2L), and the Dice coefficient (DC), using leave-one-out cross-validation. Table 1 shows the results obtained for the multi-object case. Compared with the classical PDM (avg. L2L error:  $1.20 \pm 0.49$  vox.; avg. DC:  $0.87 \pm 0.06$ ), both multiresolution hierarchical approaches provide substantial improvements in accuracy for all the subcortical structures. With the exception of globus pallidi, all improvements over PDM are statistically significant according to the Wilcoxon signed rank test (p-value  $< 0.05$  for all). Although the new GMRH-PDM performed similarly to the previous hierarchical version, MRH-PDM, in terms of accuracy (avg. L2L error:  $0.94 \pm 0.19$  vs.  $0.95 \pm 0.39$  and avg. DC:  $0.90 \pm 0.04$  vs.  $0.90 \pm 0.05$ , respectively), it provides a significant advantage over the latter. The GMRH-PDM framework introduced in this paper is fully automatic, while the original MRH-PDM requires the hierarchical configuration to be manually defined by the

user. As the number of possible configurations can be considerably high when working with large number of objects, it is a nontrivial challenge to find an optimal one by simple manual supervised selection. Thanks to the landmark clustering approach presented in Section 3, GMRH-PDM is able to automatically provide an optimal hierarchical decomposition of the structure, while performing as well as the best manual configuration of MRH-PDM. But GMRH-PDM has an additional major advantage over MRH-PDM as it allows single-object and intra-object analysis.

**Table 1.** Accuracy Evaluation. Landmark-to-landmark (L2L) distance and Dice’s coefficient (DC) (average / standard deviation) for the three studied methods (PDM, MRH-PDM, and GMRH-PDM) over eight subcortical structures ( $\mathbf{x}_1, \dots, \mathbf{x}_8$ ) (see Fig. 1). (•) marks significant improvements over classic PDM.

L2L (vox.)	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$	$\mathbf{x}_8$	Avg.
PDM	1.63/0.41	1.60/0.56	1.30/0.69	1.21/0.39	1.13/0.30	0.97/0.19	0.86/0.24	0.93/0.25	1.20/0.49
MRH-PDM	• 1.28/0.50	• 1.24/0.49	• 0.98/0.48	• 0.86/0.37	• 0.88/0.18	• 0.77/0.20	0.79/0.21	0.82/0.19	• 0.95/0.39
GMRH-PDM	• 1.26/0.45	• 1.23/0.50	• 0.97/0.47	• 0.85/0.36	• 0.88/0.17	• 0.77/0.19	0.78/0.20	0.82/0.20	• 0.94/0.19

DC	$\mathbf{x}_1$	$\mathbf{x}_2$	$\mathbf{x}_3$	$\mathbf{x}_4$	$\mathbf{x}_5$	$\mathbf{x}_6$	$\mathbf{x}_7$	$\mathbf{x}_8$	Avg.
PDM	0.80/0.05	0.80/0.07	0.86/0.04	0.86/0.04	0.89/0.02	0.91/0.02	0.91/0.02	0.90/0.03	0.87/0.06
MRH-PDM	• 0.84/0.05	• 0.85/0.06	• 0.88/0.04	• 0.90/0.04	• 0.92/0.02	• 0.93/0.02	0.91/0.02	0.91/0.02	• 0.90/0.05
GMRH-PDM	• 0.85/0.04	• 0.87/0.06	• 0.89/0.03	• 0.90/0.03	• 0.93/0.03	• 0.94/0.02	0.91/0.02	0.92/0.03	• 0.90/0.04

The superiority of GMRH-PDM over PDM to model subparts in single-object structures was also proven in the kidney database. In this case, the average L2L errors were  $0.35 \pm 0.19$  vs.  $0.47 \pm 0.2$ , and the average DCs were  $0.99 \pm 0.05$  vs.  $0.97 \pm 0.05$  for GMRH-PDM and PDM, respectively (p-value = 0.03 and 0.02 respectively).

The computational complexity of the new landmark clusterization is  $O(n^2)$ , taking  $\sim 100$  min. to process the most complex multi-object case with 8208 landmarks (code written in Matlab®). However, this is not a determining factor for the practical application of the method, since the clusterization can be performed off-line.

## 6 Conclusions

In this paper, we present a new Generalized Multiresolution Hierarchical PDM (GMRH-PDM) to address the high-dimension-low-sample-size challenge of great relevance when modeling complex structures with the classical PDM. The general framework introduced here creates different statistical models that allow to describe efficiently the variability of the shape at different levels of resolution. The new GMRH-PDM also tackles the two main drawbacks observed in previous hierarchical approaches: the difficulty of manually defining the hierarchical configuration that provides optimal performance, and the impossibility of dealing with single-object structures by considering entire objects as the minimum modeling unit. The general notation used in GMRH-PDM extends the hierarchical modeling of PDM to any set of landmarks, leading to a more versatile framework able to deal with all types of struc-

tures, even single-object shapes. Finally, the hierarchical configuration of the algorithm is automatically defined by means of a new agglomerative landmark clustering approach, whose optimization is controlled by a tailored definition of the Silhouette coefficient. The algorithm is compared with two different alternatives, PDM and the MRH-PDM. The results show how the new automatic GMRH-PDM significantly outperform the classical PDM in terms of accuracy, while providing similar results to the best manual configuration of MRH-PDM. GMRH-PDM allows the automatic hierarchical modeling of structures, from the multi-object level to the inter- and intra-object resolution, which can be of great interest in the context of full body computational anatomy. In the near future, we plan to continue exploring this capability to study population variability and the temporal anatomical variability of organs.

**Acknowledgment.** This project was supported by a philanthropic gift from the Government of Abu Dhabi to Children’s National Health System, and by the European Union FP7 project HEAR-EU (grant agreement 304857).

## References

1. Cootes, T.F. et al.: Active Shape Models Their Training and Application. *Comput. Vis. Image Underst.* 61(1), 38-59 (1995)
2. C. Davatzikos, et al.: Hierarchical active shape models, using the wavelet transform. *IEEE Trans. on Med. Imag.* 22 (3), 414–423 (2003)
3. Cerrolaza, J.J., et al.: Hierarchical Statistical Shape Models of Multiobject Anatomical Structures: Application to Brain MRI. *IEEE Trans. Med. Imaging* 31(3), 71–724 (2012)
4. C. Lu, et al.: Statistical multi-object shape models. *Int. J. Comput. Vis.* 75, 387–404 (2007)
5. Cerrolaza, J.J. et al.: Multiresolution Hierarchical Shape Models in 3D Subcortical Brain Structures. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) *MICCAI 2013, Part II*, LNCS 8150, pp. 641–648 (2013)
6. Lounsbery, M., et al.: Multiresolution Analysis for Surfaces of Arbitrary Topological Type. *ACM Trans. Graph.* 16(1), 34–73 (1997)
7. Dyn, N. et al.: A Butterfly Subdivision Scheme for Surface Interpolation with Tension Control. *ACM Trans. Graph.* 9(2), 160–169, (1990)
8. Praun, E., and Hoppe, H.: Spherical Parametrization and Remeshing. *ACM Trans. Graph. (SIGGRAPH)* 22(3), 340–349 (2003)
9. Roy, T. et al.: Segmentation of a vector field: dominant parameter and shape optimization. *Journal of Mathematical Imaging and Vision.* 24(2) 259-76 (2006)
10. Reyes, M. et al.: Anatomical variability of organs via principal factor analysis from the construction of an abdominal probabilistic atlas. *IEEE Symp. Bio. Imag.* 682-685(2009)
11. Ward, J. H. : Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association.* 58, 236-244 (1963)
12. IBSR. The Internet Brain Segmentation Repository (IBSR). Available from <http://www.cma.mgh.harvard.edu/ibsr/>
13. Yokota, F. et al.: Automated segmentation of the femur and pelvis from 3D CT data of diseased hip using hierarchical statistical shape model of joint structure. *MICCAI’09.* 811-819 (2009).