Torso Shape Extraction from 3D Body Scanning Data Using Automatic Segmentation Tool

Ola AHMAD*[1,2], Philippe DEBANNÉ[1,2], Stefan PARENT[2], Hubert LABELLE[2], Farida CHERIET[1,2]
[1] Polytechnique Montréal, Montreal QC, Canada;
[2] CHU Sainte-Justine, Montreal QC, Canada

DOI: 10.15221/17.192 http://dx.doi.org/10.15221/17.192

Abstract

The automatic and standardized extraction of the torso shape from 3D body scanning data has an important role in biomedical applications. In scoliosis clinics, the asymmetry analysis of the 3D scoliotic trunk shape relies on a prior cropping of the regions corresponding to the arms and neck. At Sainte-Justine Hospital, a system of four optical digitizers (Capturor II LF, Creaform Inc.) is used to scan the body of scoliosis patients. At present, the cropping of the trunk shapes is a manual process and is therefore operator-dependent, time-consuming and can affect the reliability of subsequent trunk asymmetry analysis. In addition, the inferior portion of the trunk (pelvic region) includes noisy geometric features that are due to the patient's lower body clothing and are irrelevant to the study of scoliotic trunk shape deformations. In this paper, we present a robust and efficient tool to extract the meaningful torso regions based on automatic segmentation. The 3D body scanning system provides a 3D triangulated mesh of the shape accompanied by an RGB color map of the texture. An anatomical landmark placed at the midpoint of the posterior-anterior iliac spines (MPSIS) prior to the acquisition determines the separation level between the pelvic region and the rest of the torso (i.e. the lumbar and thoracic regions). We propose a two-phase segmentation algorithm. In the first phase, a skin-color model is used to separate the pelvic region from the other portions of the torso. The second phase separates the arms and neck regions using relevant geometric features captured by a spectral representation of the shape. We tested our algorithm on a dataset composed of 56 scoliotic body shapes scanned in neutral standing and lateral bending postures by comparing the torsos cropped automatically versus manually by an operator. The results show that our algorithm achieves a 0.95 ($\pm$0.04) degree of overlap, in terms of the average Dice similarity measure, between the extracted torso shapes and their ground truth counterparts. The proposed automatic segmentation method thus constitutes a useful tool to include in the 3D body surface scanning systems used in scoliosis clinics.

Keywords: 3D torso surface model, skin-color model, spectral shape representation, segmentation

1. Introduction

Body shape segmentation is an essential process for shape analysis and understanding in different applications, including computer graphics, fashion design, and medicine. The segmentation of the shape into meaningful parts facilitates the extraction of relevant information associated with each subpart and, subsequently, has a significant impact on the accuracy of processing shape subparts. Our motivation in doing this work is to develop an automatic tool that helps analyzing the 3D deformations of torso shapes in scoliosis clinics. In this application, the 3D body scanning system captures the upper part of the body including the torso, the pelvis, the neck and the upper regions of the arms connected to the shoulders. Consequently, analyzing the deformations of the 3D scoliotic trunk shapes relies on a prior cropping or delineation of the arms, neck and pelvis subparts. At present, this is done manually and based on anatomical landmarks. However, previous studies show that the manual segmentation process affects the accuracy of the scoliosis deformity characterization and torso shape analysis [4, 5]. In this context, a standardized and automatic segmentation method can overcome these limits.

In many applications, 3D mesh models become the conventional representation of body shapes. This requires most automatic methods to focus on mesh segmentation using the geometric properties of shape, as for instance, contour-based methods [7, 8], spectral clustering [6] and diffusion distance [9]. The geometric-based methods are very efficient in segmenting "low-frequency" structural information, such as articulated body parts. However, in our application, we need to segment parts of the body covered by clothing textures; it, therefore, becomes more difficult to rely on the mesh geometry to

* Corresponding author: olasahmad@gmail.com

accurately segmenting those regions. Interestingly, the body shape representation obtained from our acquisition system includes a color map of the texture in addition to the 3D mesh model. The use of this additional information allows us to overcome this problem and to more accurately and efficiently segment the shape's different portions. More particularly, the color map will help in separating the pelvic region from the torso. To this end, we propose in this paper a two-phase segmentation method that distinctly combines geometric and texture features in order to extract the torso part of the body. During the first phase, the method uses the color map of the texture to delineate the pelvic region and separate it from the upper body. In the second phase, we use the spectral embedding of the shape to segment the articulated parts at the arms and neck. At the end of each phase, our method crops off the undesired parts and applies an interpolation on the rest of the shape. Finally, we use a nearest neighbor search to recover the texture information associated with the torso part. We begin this paper by describing the 3D body scanning system and the details of our segmentation method to extract the torso shape. Then, we present validation results of the proposed method on a dataset of scoliotic subjects with different postures to evaluate the segmentation accuracy.

2. 3D Body Shape Scanning

The human body geometries are acquired using a multi-head optical digitizing system (Capturor II LF, Creaform Inc.) located in the scoliosis clinic of Sainte-Justine Hospital in Montreal, Canada. The system comprises four calibrated scanners (see figure 1) placed around the patient (at the back, at the front, and to the right and left sides of the front angulated about 55°). Each scanner consists of a CCD camera and a structured light projector. They have a large field with 3D view of 1200 mm x 900 mm x 1000 mm, a lateral resolution of 1.2 mm, a depth resolution of 1.0 mm, and a camera resolution of 0.8 megapixels. The patient stands in the middle field of view of the four scanners (in the origin of the coordinate system shown in figure 1), focusing on a point above the front digitizer with arms in slight abduction by the side in order to minimize occlusion of the right and left sides of the torso. Prior to acquisition, anatomical landmarks are affixed manually to the skin of the torso surface. They are located at the left and right anterior-superior iliac spines, the midpoint of the posterior-superior iliac spines and at C7 vertebral prominence. The acquisition process, for each scanner, uses the phase-shifting interferometry to capture four fringe patterns, and then computes a 3D point cloud of the surface depth by triangulation. The entire acquisition time lasts approximately five seconds. Each digitizer reconstructs a portion of the body surface geometry and provides the corresponding RGB color map of the texture (see figure 2(a)). Using dedicated processing software, the 3D partial surfaces are then registered and merged to produce a complete 3D textured surface mesh of the torso, arms, neck and pelvis. The resulting mesh data contains between 50k and 100k vertices and between 100k and 200k triangles, depending on the patient's size. A previous study of this system demonstrated an accuracy of about 1.4 mm [1].

3. Torso Shape Extraction

We propose a sequential algorithm that first segments the regions covered by clothing textures (at the pelvic region) and then the articulated parts of the body.

3.1. First segmentation phase

During this phase, the segmentation relies on the color information of the texture. The color features obtained from acquisition are the default RGB (Red-Green-Blue) components. We use these features to distinguish between skin and non-skin regions of the shape. However, the RGB channels are, by their very nature, highly correlated and include redundant information. In addition, they are sensitive to intensity variations, which affect the segmentation accuracy [10]. It is therefore important to choose a color transformation that allows for better color constancy and discrimination power. It is the primary step in skin-color segmentation. For this goal, several transformations were proposed in the state-of-the art, as surveyed by [11]. Among these, we choose an orthogonal color spaces, namely the YIQ representation (see figure 2). We favor this color space for two reasons: first, as an orthogonal space, it reduces the redundancy existing between the RGB channels. Second, it represents colors by statistically independent luminance (Y) and chrominance (I, Q) features. This allows us to efficiently drop the intensity features and perform skin segmentation in a robust manner with respect to variable illumination conditions. Following this choice, we use the I-component, which includes color features ranging from orange to cyan and covers a significant range of the skin tones [12].

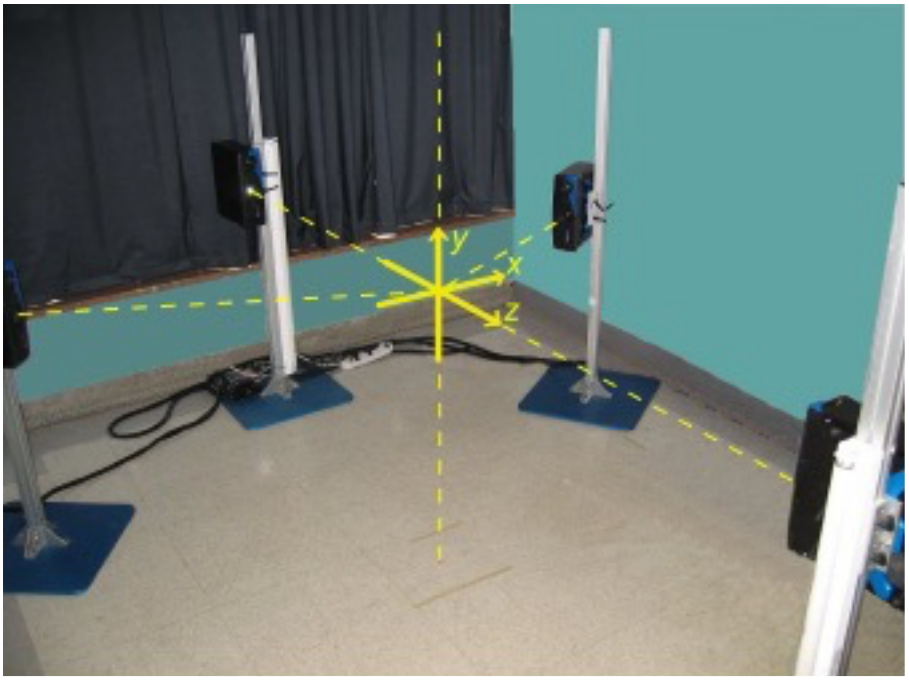

Fig. 1. The Creaform 3D Body scanning system used at Saint-Justine hospital.

Let $\mathrm{x} = (x, y, z)$ be a surface coordinate point of the shape, and $I(\mathrm{x})$ be the I-color component at this point (x). As illustrated in figure 2(c), the histogram function of this component represents a mixture of two models: each one represents the distribution of the chrominance values, in the skin and non-skin regions, respectively, which can be approximated by a Gaussian model. A threshold value can, therefore, be computed from a statistical inference based on the Bayes decision rule. The decision function is formulated in terms of the class-conditional probabilities of the skin and non-skin models, while the likelihood ratio test (LRT) is defined as follows:

$$\frac{P(I(\mathrm{x})\,|\,\Theta_1; H_1)}{P(I(\mathrm{x})\,|\,\Theta_0; H_0)} > \gamma, \quad (\gamma > 0) \tag{1}$$

where Θ_0, Θ_1 are the models parameters under hypotheses H_1 (skin) and H_0 (non-skin), respectively, and γ is the likelihood threshold determined according to a desired level of significance (e.g., $\alpha = 5\%$) that controls the false positive rate. Computing the log-LRT, we get:

$$\log[P(I(\mathrm{x})\,|\,\Theta_1; H_1)] - \log[P(I(\mathrm{x})\,|\,\Theta_0; H_0)] > \delta \tag{2}$$

where $\delta = \log[\gamma]$ and $\log[P(I(\mathrm{x})\,|\,\Theta_k; H_k)] = -\log[|\Sigma_k|]/2 - (\mathrm{x}\text{-}\mu_k)^T \Sigma_k^{-1} (\mathrm{x}\text{-}\mu_k)/2$ is the log of the multivariate Gaussian probability of the kth model, with $\Theta_k = (\mu_k, \Sigma_k)$ the unknown parameter vectors. In order to estimate the model's parameters under both hypotheses, we use the well-known iterative expectation-maximization (EM) algorithm. Finally, the skin and non-skin color regions are separated at 5% significance level according to (2). In the example illustrated in figure 2, the segmentation threshold obtained for $\alpha = 5\%$ equals -0.1. This threshold isolates the regions covered by clothing and the anatomical landmarks fixed on the skin prior to acquisition. To obtain a standardized cropping of the pelvis across all shapes, we choose to segment it below the MPSIS landmark located on the lower back (see figure 2(a)). Figure 2(c) (bottom right of figure) illustrates the segmentation result of the first phase.

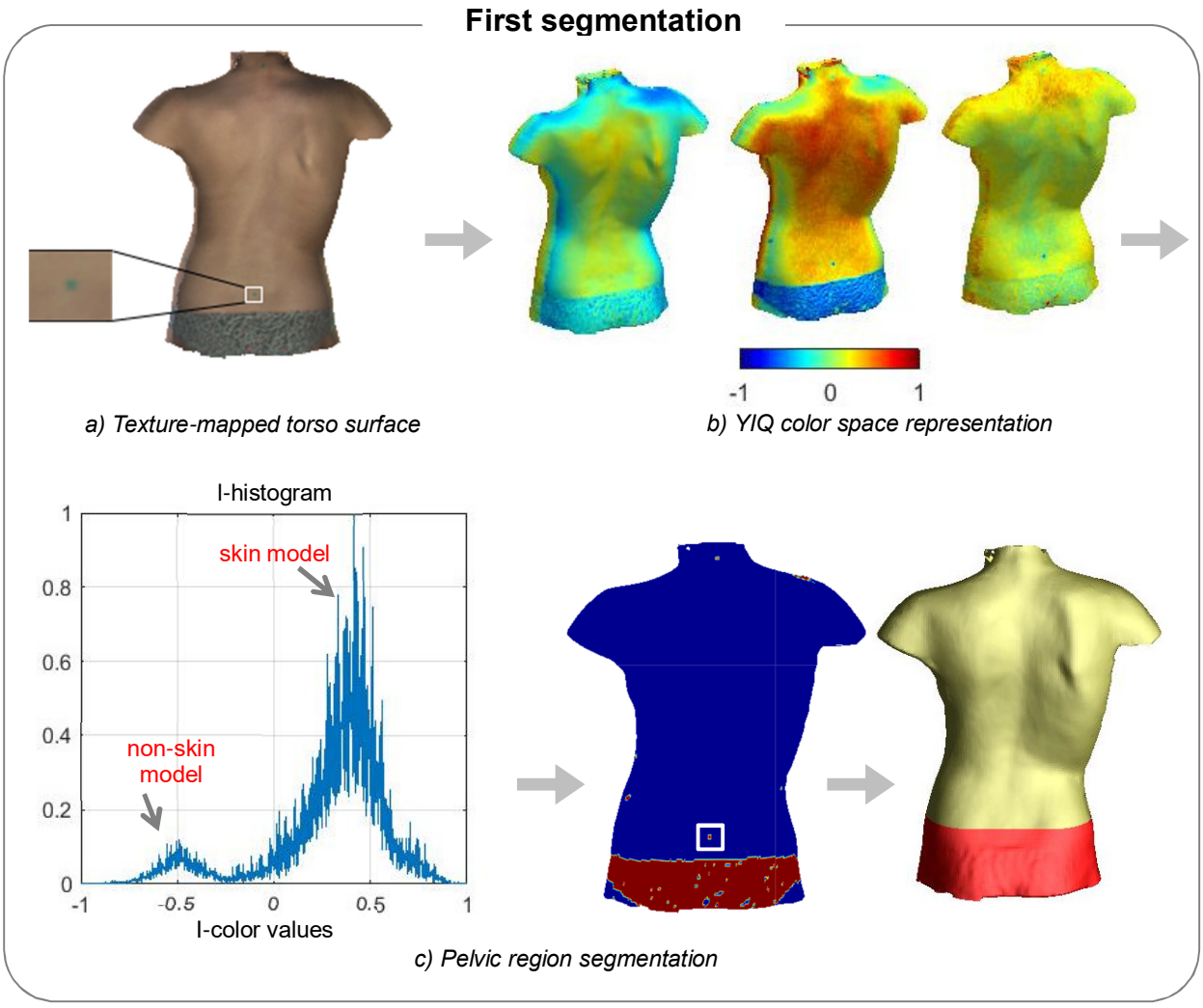

Fig. 2. Illustration of the first segmentation phase to separate the pelvic region from the upper body shape.

3.2. Second segmentation phase

After cropping the pelvis, we apply an interpolation on the rest of the surface mesh using the RBF algorithm [2] to smooth the boundaries and close the holes (see figure 3(a)). This interpolation step is necessary to accurately extract the geometric features needed to segment the articulated parts of the body in the second phase. For this next stage, we use the spectral clustering approach based on the Graph Laplacian and Gaussian mixtures [3]. We describe first the spectral graph decomposition of the shape, and then we show how relevant features can separate the articulated regions using Gaussian mixture modeling.

3.2.1. Spectral graph decomposition of shapes

Let $G=(V,E)$ be a connected undirected graph where V represents a set of N nodes corresponding to the mesh elements (i.e., triangles), and E is the set of edges connecting pairs of neighboring vertices. We present a graph Laplacian on the surface mesh points $\mathbf{x}$ with its $N\times N$ weighted adjacency matrix $\mathbf{W}$, where the edge weights are given by $\mathrm{w}_{ij}=\exp\left(-\left\|\mathrm{x}_i-\mathrm{x}_j\right\|^2/2\sigma^2\right)$ if $(i,j)\in E$ and 0 otherwise. The Laplacian matrix $(\mathbf{L})$ of the graph is semi-definite positive and symmetric defined by $\mathbf{L}=\mathbf{D}^{-1}\left(\mathbf{D}-\mathbf{W}\right)$, where $\mathbf{D}$ is the diagonal degree matrix, $D_{ii}=\sum_{j=1}^{N}\mathrm{w}_{ij}$. The spectral decomposition of the Laplacian matrix, $\mathbf{L}=\Phi^T\Lambda\Phi$, defines the set of the eigenvalues

$\Lambda = \left(\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N\right)$, where λ_1 is the smallest non-zero eigenvalue, and the associated eigenvectors $\Phi = (\phi_1, \cdots, \phi_N)$, which are normalized in the range $[-1,+1]$. Without loss of generality, we focus here on the first 6 eigenvectors since they capture the "low-frequency" geometric properties of the shape corresponding to the articulated subparts. More particularly, we select the second and forth eigenvectors to extract segments at the arms and the neck. Figure 3(b) shows how these selected eigenvectors behave on the body shape; positive and negative values of the 2nd eigenvector near ± 1 are located at left and right arms (shoulders), while high positive values of the 4th eigenvector are located in the neck region. As can be noticed, in spite of the local noise at the boundaries of the shape and posture changes, these features are stable and describe the same geometric structures of the shape. This is an important benefit of the spectral shape representation: the first embedding components capture intrinsic and global geometric structures that are invariant to the position changes of the surface points in the Euclidean space [13].

3.2.2. Spectral Clustering using Gaussian Mixtures

To segment the articulated shape subparts, we propose a multi-threshold approach based on the distributions of the selected eigenvectors. In an analogous manner to the segmentation in the first phase, we use the Gaussian mixtures to estimate the distribution models. Note that the Gaussian mixture model (GMM) was proposed in spectral mesh segmentation with the Bayesian information criterion to cluster local regions along the extremities of the eigenvectors [3]. In our case, we will use the second and fourth spectral features to crop the arms and the neck articulated to the torso shape. In contrast to the segmentation in section 3.1, the GMM model will be composed of multiple components and a multiple hypotheses must be considered.

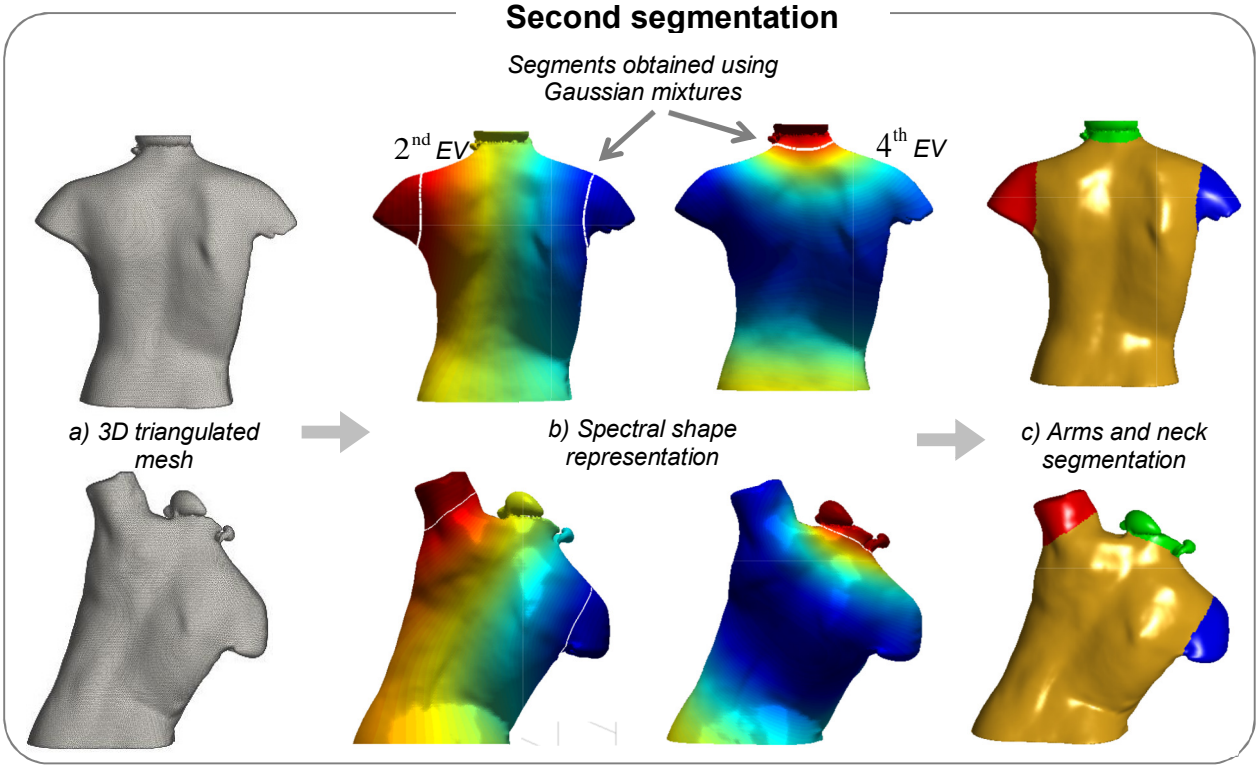

Fig. 3. Illustration of the second segmentation phase using the spectral features (2^{nd} and 4^{th} eigenvectors) of the surface mesh and Gaussian mixtures.

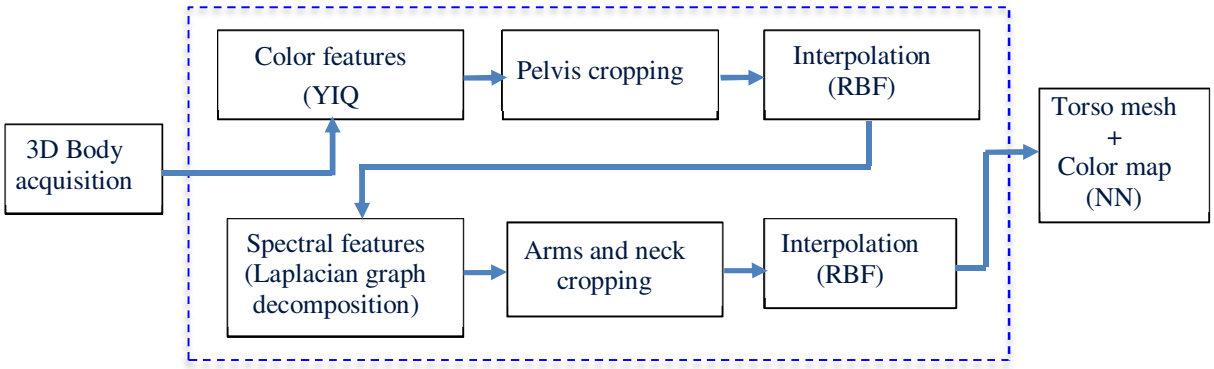

Fig. 4. Flowchart of the torso extraction pipeline.

Let $\phi_j(\mathrm{x})$, with $j \in \{2,4\}$, be the selected Laplacian eigenvector defined at the surface point coordinate $\mathrm{x} = (x, y, z)$. Accordingly, we formulate the likelihood ratio test (LRT) of multiple hypotheses for the jth eigenvector as follows:

$$\frac{P(\phi_j(\mathrm{x}) \mid \Theta_k; H_k)}{\sum_{c=1}^{C} P(\phi_j(\mathrm{x}) \mid \Theta_c; H_c)} > \gamma_k, \quad (\gamma_k > 0) \tag{3}$$

where γ_k is the threshold corresponding to the hypothesis (model) H_k, Θ_c is the parameter vector of the model H_c $(c = 1, \cdots, C)$, and C is the number of models or mixture components. We set this number according to the structure of $\phi_j(\mathrm{x})$ such that $C = 4$ for $j = 2$ and $C = 3$ for $j = 4$. This can be observed in figure 3(b); the positive and negative values of the second eigenvector correspond to the right and left sides of the trunk shape, respectively. The total number of the GMM components is, therefore, twice the number of components on one side, which is two in our case. The fourth eigenvector concentrates its extreme positive and negative values at the head and around the waist, respectively. The segmentation threshold is determined by finding the maximum LRT estimates for each of the model's parameters Θ_k, Θ_c using the log-LRT with EM optimization. Consequently, for the second eigenvector, we obtain two thresholds that delineate the arms (shoulders), and one threshold from the fourth eigenvector that separates the neck (or head) from the body (see figure 3(b)). Figure 3(c) illustrates the segmented shape regions in different colors.

The torso mesh surface is thus obtained by cropping off the segmented regions at the arms and neck. We then interpolate the mesh surface to smooth the boundaries and close the holes left by the removed subparts. Note that interpolating the mesh leads to slight changes of the positions of the nodes and their connections and, therefore, leads to losing the color information. In order to recover the color features for the extracted torso shape, we find the closest positions between the nodes of the original and processed meshes using a nearest-neighbor search. Finally, we associate the torso mesh nodes with the color values of their corresponding nodes from the original mesh. Figure 4. Illustrates the pipeline scheme of the proposed approach starting from the raw data and leading to the final result.

5. Validation Results

We applied the proposed method to a dataset of 56 scanned body shapes in three different postures: left and right lateral bending and neutral standing postures. These samples belong to scoliotic patients who were examined and operated at Sainte-Justine Hospital in Montreal. To validate our automatic segmentation approach, an experienced operator provided us with a ground truth by processing the same set of trunk shapes using a manual segmentation tool. This manual process consists in digitizing

the 3D coordinates of a set of anatomical landmarks: the skin markers at the VP, left and right ASIS and MPSIS to delimit the upper and lower torso limits, and the armpits to delimit the shoulders. Cutting planes are then automatically computed to remove the neck, pelvis and shoulders. In the lateral bending case, additional landmarks must be digitized on top of the shoulders and the clavicle to orient the arm and neck cutting planes correctly.
To visualize the comparison results, we projected the torso meshes extracted from both segmentations onto the original shape and measured their degree of overlap. Figure 5 Illustrates three examples of the segmentation results representing the tree different postures. The average Dice overlap using the Dice similarity measure, computed as $(2|A\cap B|/(A+B))$, is 0.95 ± 0.04 over all the shapes. This result is considered satisfactory for subsequent scoliotic shape analysis.

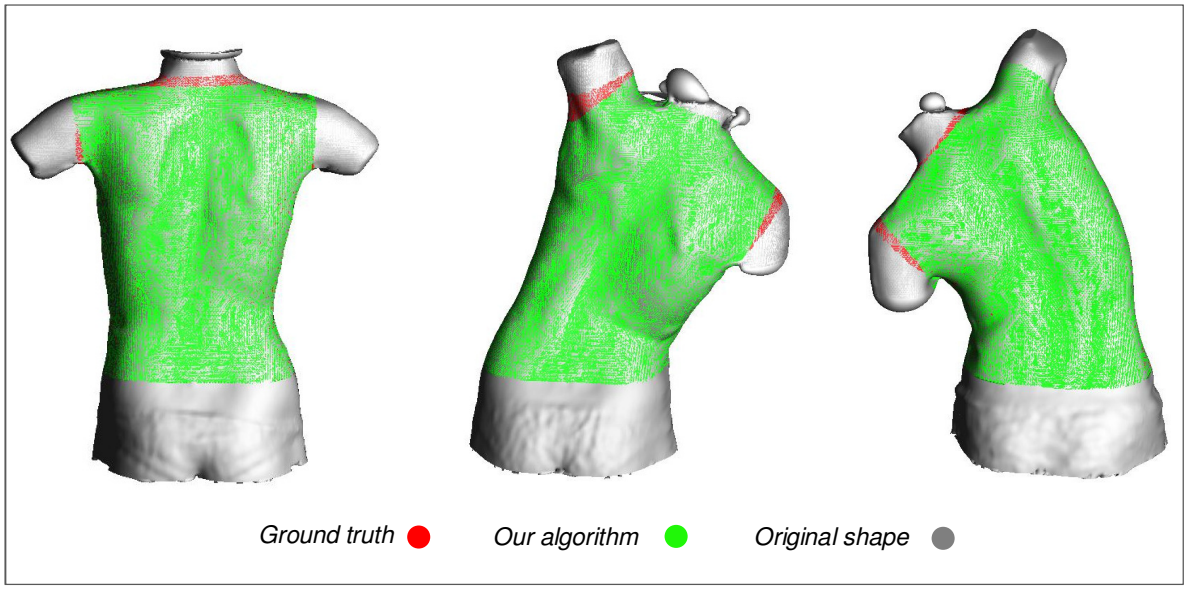

Fig. 5. Examples of torso shape extraction from upper body mesh surfaces in three postures.

6. Conclusion

In this paper, we have proposed an automatic segmentation tool to extract the torso region from the 3D upper body scanning data in two sequential phases. The advantage of our method is that it combines texture color and geometric information to facilitate the discrimination of the articulated shape subparts as well as the regions covered by clothing. Furthermore, each segmentation phase relies on a feature representation space that captures relevant information to accurately separate shape subparts. The YIQ color space allows us to reduce the sensitivity of our approach to the differences in illumination conditions, while the graph Laplacian embedding captures the intrinsic geometry of the shape and is robust with respect to posture changes. Our proposed segmentation method constitutes a very useful tool to include in 3D body surface scanning systems used in scoliosis clinics and other applications that involve torso shape analysis.

References

[1] V. Pazos et al., “Accuracy Assessment of Human Trunk Surface 3D Reconstructions From An Optical Digitizing System,” in *Journal of Medical & Biological Engineering & Computing*, Vol. 43, No. 1, 2005, pp. 11-15, https://doi.org/10.1007/BF02345117.

[2] J. C. Carr et al., “Reconstruction and Representation of 3D Objects with Radial Basis Functions,” in *Proc. of the 28th Ann. Conf. on Computer Graphics and Interactive Techniques*, New York, USA, 2001, pp. 67-76, http://doi.acm.org/10.1145/383259.383266.

[3] A. Sharma et al. “Mesh Segmentation using Laplacian Eigenvectors and Gaussian Mixtures,” in *AAAI Fall Symp. On Manifold Learning and its Applications*, Arlington, VA, USA, 2009, pp. 50-56.

[4] L. Seoud et al. “Multilevel Analysis of Trunk Surface Measurements for Noninvasive Assessment of Scoliosis Deformities,” *Spine*, Vol. 37, No. 17, 2012, pp. E1045-1053, http://dx.doi.org/10.1097/BRS.0b013e3182575938.

[5] O. Ahmad et al. “Longitudinal Scoliotic Trunk Analysis via Spectral Representation and Statistical Analysis,” in *Proc. Spectral and Shape Analysis in Medical Imaging (LNCS Springer)*, Athens, Greece, 2016, pp. 79-91, https://doi.org/10.1007/978-3-319-51237-2_7.

[6] R. Liu et al. “Segmentation of 3D Meshes Through Spectral Clustering,” in *Proc. of the 12th Pacific Conf. on Computer Graphics and Applications, Seoul, South Korea,* 2004, pp. 298-305, https://doi.org/10.1109/PCCGA.2004.1348360.

[7] A. P. Mangan and R. T. Whitaker. “Partitioning of 3D Surface meshes using Watershed Segmentation,” in *IEEE Trans. Vis. Comput. Graphics,* Vol. 5, No. 4, 1999, pp. 308-321, https://doi.org/10.1109/2945.817348.

[8] R. Liu and H. Zhang. “Mesh Segmentation via Spectral Embedding and Contour Analysis,” in *Comput. Graph. Forum (Eurographics),* Vol. 26, No. 3, 2007, pp. 385-394, http://dx.doi.org/10.1111/j.1467-8659.2007.01061.x

[9] F. De Goes et al. “Hierarchical Segmentation of Articulated Bodies,” in *Proc. of the Symp. on Geometry Processing,* Copenhagen, Denmark, 2008, pp. 1349-1356.

[10] J. Brand and J. S. Mason. “A Comparative Assessment of Three Approaches to Pixel-level Human Skin Detection,” in *Proc. of the International Conf. on Pattern Recognition,* Washington, DC, USA, 2000, pp. 50-56.

[11] P. Kakumanu et al. “A Survey of Skin-color Modeling and Detection Methods,” in *Pattern Recognition,* Vol. 40, No. 3, 2006, pp. 1106-1122, https://doi.org/10.1016/j.patcog.2006.06.010.

[12] D. Ying and N. Yasuaki. “Face Texture Model Based on SGLD and its Application in a Color Scene,” in *Pattern Recognition,* Vol. 29, No. 6, 1996, pp. 1007-1017, https://doi.org/10.1016/0031-3203(95)00139-5.

[13] H. Lombaert et al. “FOCUSR: Feature Oriented Correspondence Using Spectra Regularization--A Method for Precise Surface Matching,” in *IEEE Trans. Pattern Anal. Mach. Inell.,* Vol. 35, No. 9, 2013, pp. 2143-2160, http://dx.doi.org/10.1109/TPAMI.2012.276.