



Published in final edited form as:

*Proc SPIE Int Soc Opt Eng.* 2016 ; 9784: . doi:10.1117/12.2217337.

## Automatic segmentation and statistical shape modeling of the paranasal sinuses to estimate natural variations

Ayushi Sinha<sup>a</sup>, Simon Leonard<sup>a</sup>, Austin Reiter<sup>a</sup>, Masaru Ishii<sup>b</sup>, Russell H. Taylor<sup>a</sup>, and Gregory D. Hager<sup>a</sup>

<sup>a</sup>Dept. of Computer Science, The Johns Hopkins University, Baltimore, MD 21218, USA

<sup>b</sup>Dept. of Otolaryngology-Head and Neck Surgery, Johns Hopkins Medical Institutions, Baltimore, MD 21287, USA

### Abstract

We present an automatic segmentation and statistical shape modeling system for the paranasal sinuses which allows us to locate structures in and around the sinuses, as well as to observe the natural variations that occur in these structures. This system involves deformably registering a given patient image to a manually segmented template image, and using the resulting deformation field to transfer labels from template to patient. We use 3D snake splines to correct errors in the deformable registration. Once we have several accurately segmented images, we build statistical shape models for each structure in the sinus allowing us to observe the mean shape of the population, as well as the variations observed in the population. These shape models are useful in several ways. First, regular video-CT registration methods are insufficient to accurately register pre-operative computed tomography (CT) images with intra-operative endoscopy video because of deformations that occur in structures containing high amounts of erectile tissue. Our aim is to estimate these deformations using our shape models in order to improve video-CT registration, as well as to distinguish normal variations in anatomy from abnormal variations, and automatically detect and stage pathology. We can also compare the mean shape and variances of different populations, such as different genders or ethnicities, and observe the differences and similarities, as well as of different age groups, and observe the developmental changes that occur in the sinuses.

### Keywords

Segmentation; Statistical shape modeling; Paranasal sinuses

## 1. DESCRIPTION OF PURPOSE

Automatic segmentation of medical images is important for several reasons. Manual segmentation is tedious and not scalable to large datasets. Automatic segmentation allows us to extract shapes from a large number of labeled images, and to study large populations. These shapes or structures play an important role in registration. However, for our

application, video-CT registration during minimally invasive surgery through the paranasal sinuses, simply having a segmented structure can be insufficient. Several structures in the sinus exhibit high variability, for instance structures containing erectile tissue. The three turbinates, superior, middle and inferior, which reside in the nasal passage contain erectile tissue. These turbinates undergo periodical deformation as the complete they nasal cycle, which is the alternating partial congestion and decongestion of the nasal cavities due to the expansion and contraction of the turbinates. These contractions can be further exaggerated due to decongestants administered to patients before surgery to facilitate easy movement of tools in the sinus. These changes in structure make registration of turbinates error prone. In order to tackle this problem, we need models explaining structure variability in addition to high quality segmentations.

Our aim, therefore, is to obtain accurate segmentations with correspondences for a large number of patient images, build shape models, and learn useful information from these models. In order to achieve these goals, we built a template image (Fig. 1) which we manually segmented under the supervision of an otolaryngologist, and deformably registered all given patient images to this template. We then automatically segmented the patient image by deforming the template segmentations to patient space using the deformation fields resulting from the registration. This segmentation is further improved using gradient vector flow (GVF) to drive mesh vertices closer to edges in the patient image. Since we deform one shape to a collection of images, we are guaranteed correspondence between the shapes. Finally, once we obtained accurate segmentations for several images, we built statistical shape models (SSMs) for each segmented structure, and acquired statistics for our patient population. Each of these steps is detailed in the next section.

## 2. METHODS

### 2.1 Template creation

In order to minimize the influence of individual variation in our registration, we built a template image which represents the mean of a collection of images. We use a standard template building technique (Fig. 1), which requires one target image, and  $n$  other images of the same dimensions and resolution as the target image. These  $n$  images are deformably registered to the target image using ANTs registration software,<sup>1</sup> resulting in deformation fields that can transform the target image to each of the  $n$  images. If we average these deformation fields and transform the target image by this mean deformation field, it takes the target image towards the space of the mean of the  $n$  images.<sup>2</sup> This process can be repeated several times using the output from the previous iteration as the new target image. Each iteration further reduces the individual variation of the target image, and moves it closer to the population mean. The resulting image results in a highly symmetrical and ideal looking image of the sinus. This is our template image, which we hand segment under the supervision of an expert, and use to automatically segment any given patient image.

### 2.2 Automatic segmentation

Given a new patient image, we deformably register it to our segmented template image, again using ANTs registration software.<sup>1</sup> We extract surface meshes<sup>3</sup> from our template

segmentations, and use the deformation fields from deformable registration to move each vertex in our template meshes to patient space. We use GVF to correct errors in registration.<sup>4</sup> Since it is not possible to sequentially order vertices in 3D the way it is done in 2D, we need to find a way to order the vertices that make up our snake splines. Simple blob-like structures have been approximated with parametrized ellipsoids with high accuracy.<sup>5</sup> However, structures in the sinus can often be too complex to be approximated using a simple parametrization. We take a very simple approach to solving this problem, without having to parametrize a complex shape, or having to build and store a large matrix containing per vertex neighborhood information.

### Algorithm 1

Order neighboring vertices

---

```

1: procedure Order_Neighboring_Vertices(Mesh  $m$ ,  $v[][]$ ) ▷  $v[][]$  is a vector of vectors
2:   for each vertex  $i$  in  $m$  do
3:      $f \leftarrow \text{find\_faces}(i)$  ▷ Find faces incident on  $i$ , ordered clockwise
4:     for  $j = 0 \rightarrow f.size()$  do ▷ For each face in  $f$ 
5:        $v[j][i] \leftarrow$  vertex at the end of half-edge6 facing vertex  $i$ 
6:   return  $v[][]$ 

```

---

### Algorithm 2

Sample vertices

---

```

1: procedure Sample_Vertices(Mesh  $m$ ,  $i$ ,  $n$ ,  $s[]$ ,  $v[][]$ ) ▷ Sample  $n$  vertices in a spiral starting at  $i$ 
2:    $visited.resize() \leftarrow m.vertices.size()$ 
3:    $s.resize() \leftarrow n$ 
4:    $s[0] \leftarrow i$ 
5:    $visited[i] \leftarrow 0$  ▷  $visited$  initialized to 0
6:    $k \leftarrow 0$ 
7:    $x \leftarrow 0$ 
8:   for  $j = 1 \rightarrow n$  do
9:     if  $k < v[j].size()$  then
10:      if  $!visited[v[j][k]]$  then ▷ Visit all unvisited vertices in one-ring neighborhood of  $i$ 
11:         $s[j] \leftarrow v[j][k]$ 
12:         $visited[v[j][k]] = 1$ 
13:         $k \leftarrow k + 1$ 
14:      else ▷ Visit all unvisited neighbors of vertices in one-ring neighborhood of  $i$ 
15:         $k \leftarrow 0$  ▷ Once the two-ring neighborhood is exhausted, proceed to the next
16:         $x \leftarrow x + 1$  ▷ Repeat until  $n$  vertices have been sampled
17:         $i \leftarrow s[x]$ 
18:   return  $s$ 

```

---

We use a simple but efficient sampling structure for our 3D snake control points. We sample points in a spiral around each vertex (Algorithms 1, 2), which allows us to store the points on this curve in a sequential order in a vector. These vertices can therefore be moved using simple vector-matrix operations. The user can specify how many points should be sampled at each iteration, which defines the length of the curve. These curves have consistent internal and external directions which make it easy to define internal and external energies<sup>4</sup> on the spline (Fig. 2). The only points of concern with this sampling structure are at the source and end points on the spiral. These two points are attracted to each other despite being far away from each other spatially due to our closed loop formulation. We deal with this by updating all points after propagation, except these two points. These points are updated in other iterations when they are not source or end points (Fig. 3). We use GVF to guide these control points toward edges in the corresponding image.<sup>4</sup> This method gives us highly accurate segmentations (Fig. 4, 5, Table 1), evaluated in the next section. We are working on further improving segmentation by making use of more information in our images and meshes, such as pixel intensity in addition to image gradients, as well as vertex orientation and mean shape information in addition to neighborhood information.

### 2.3 Statistical Shape Models (SSMs)

Once we have high quality segmentations in images of several healthy individuals, we can study the statistics of this population. We do so by using principal component analysis (PCA).<sup>7</sup> We create shape vectors, stack them into a matrix, and find the eigenvectors and eigenvalues of this matrix. This allows us to observe the mean shape as well as the principal modes of variation in our dataset. This information not only shows us how different structures vary across our sample population, but also reflects natural variations that occur periodically in some structures in the sinus.

## 3. RESULTS

We compare our segmentation results to manual segmentations. Table 1 shows that segmentations after GVF have smaller average error, and almost always have smaller maximum error. Fig. 4 shows edge maps of deformably registered maxillary sinuses (red) and those after snakes splines (green) compared against manual segmentations (blue). The green and blue edge-maps overlap almost perfectly (right) indicating snakes splines minimizes errors in segmentation. Fig. 5 also shows results from deformable registration and GVF on the right maxillary sinus. Errors noticeable after deformable registration (top) are clearly reduced after GVF (bottom).

We build our SSMs using these highly accurate segmentations. However, our point correspondences may contain errors, especially after GVF since we move the vertices in each shape independently to improve segmentations. We use this initial point correspondence to estimate a SSM, which can then be used to re-estimate point correspondences.<sup>8</sup> In this way, we update our point correspondences using SSMs. By repeating this process multiple times, we can iteratively improve our point correspondences. The results from iterative improvement in point correspondences on the middle turbinate can be seen in Fig. 6.

Finally, we evaluate whether or not our shape models can estimate natural variations in the turbinates. We know that the turbinates, superior, middle, and inferior, contain erectile tissue, and facilitate the nasal cycle by expanding and contracting periodically. Since each image in our dataset likely contains turbinates at different points in the nasal cycle, we can hypothesize that the variation we see in our population reflects the natural variation that turbinates undergo. In order to evaluate this hypothesis, we segment two CT images from the same patient, one pre- and the other post-operation, using the automatic segmentation method described in section 2.2. Then, we project the skull from both the pre- and post-operative images onto the skull model, and obtain mode weights for each shape. We can then use the mode weights and the shape model to estimate both the pre- and post-operative skulls. The mode weights for both the pre- and post-operative skulls are similar, and therefore, the two estimated shapes are also very similar. This is what we would expect because since skulls do not exhibit change over a short period of time.

We repeat the same process with the inferior turbinates, and observe a large variation in the mode weights, and hence also in the estimated shapes. This, again, is as we expected, since the turbinates are likely at different points in their nasal cycle in the two CT images. Fig. 7 shows the lack in variation in the pre- and post-operative skulls, whereas Fig. 8 shows the similarity between population variation and natural variation in the inferior turbinates.

## 4. DISCUSSIONS

### Video-CT registration using natural shape variations in the turbinates

Since we can show that population variation is able to reflect natural variation in the turbinates, we hope to attempt to use this information to co-register pre-op patient CT and intra-op endoscopy video. The problem we face during this registration is due to topology changes that occur in the turbinates. During a CT scan, the patient's turbinates are at some stage of the nasal cycle. However, during surgery, the patient is administered de-congestants to facilitate smooth insertion and movement of the endoscope and other surgical tools. This causes a very high amount of contraction in the turbinates. Since the endoscope sees the middle turbinate soon after entering the sinus, we want to be able to use this structure to start our video-CT registration. We aim to use the modes of variation of the middle turbinate to optimize registration (Fig. 9).

### Normal vs abnormal variations

Since our SSMs are built from a set of patient images with “normal” or disease free sinuses, they tell us what types and amounts of variations we should expect in each of the segmented structures. This allows us to define the range of variations that can be described as normal, or observed in a normal population. If, however, we observe an individual who demonstrates variations exceeding this normal range, we can quantify the amount by which this individual exceeds the normal range, and relate this quantification to a grading scale for pathology (Fig. 10). Further, we can evaluate the severity of disease according to our automatic grading scale against the manual grading scale currently in use, as well as against post-surgery patient quality of life.

## Population based statistics

Our current SSM is built from 53 images. Other than the scans, we did not have any additional patient information, such as age, gender, or ethnicity, available to us. Therefore, our SSM describes the statistics of a general population, not specific to age, gender, or ethnicity. We hope to collect more data in the future with additional patient information, and build SSMs for different age groups, genders, and ethnicities. Such specialized SSMs have several advantages. SSMs of different age groups would allow us to observe the developmental changes that structures in the sinus undergo as humans develop. It would also allow surgeons to adapt their tools to fit the size of structures present in different age groups. Similarly, SSMs of different genders and ethnicities also allow us to observe the differences and similarities in the different groups, again allowing surgeons to customize their care.

## 5. CONCLUSIONS

In conclusion, we present a system which is able to automatically segment sinus images to sub-millimeter accuracy. We build statistical shape models that allow us to observe the different variations that structures in the sinus undergo, allowing us to estimate the kinds of deformations that happen in the turbinates during the nasal cycle. We are currently working on improving video-CT registration accuracy using these shape models. We also hope to use these models to allow us to identify and stage pathology, amongst other useful applications. In the future, we hope to further improve our segmentations, and build higher quality, population specific SSMs.

## Acknowledgments

This work was funded by NIH R01-EB015530: Enhanced Navigation for Endoscopic Sinus Surgery through Video Analysis.

## References

1. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC. A reproducible evaluation of {ANTs} similarity metric performance in brain image registration. *NeuroImage*. 2011; 54(3):2033–2044. [PubMed: 20851191]
2. Avants BB, Yushkevich P, Pluta J, Minkoff D, Korczykowski M, Detre J, Gee JC. The optimal template effect in hippocampus studies of diseased populations. *NeuroImage*. 2010; 49(3):2457. [PubMed: 19818860]
3. Lorensen, WE., Cline, HE. Marching cubes: A high resolution 3d surface construction algorithm. *Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '87*; New York, NY, USA: ACM; 1987. p. 163-169.
4. Xu C, Prince J. Snakes, shapes, and gradient vector flow. *Image Processing, IEEE Transactions on*. Mar.1998 7:359–369.
5. Delgado-Gonzalo R, Chenouard N, Unser M. Spline-based deforming ellipsoids for interactive 3d bioimage segmentation. *Image Processing, IEEE Transactions on*. Oct.2013 22:3926–3940.
6. Weiler K. Edge-based data structures for solid modeling in curved-surface environments. *Computer Graphics and Applications, IEEE*. Jan.1985 5:21–40.
7. Cootes T, Taylor C, Cooper D, Graham J. Active shape models-their training and application. *Computer Vision and Image Understanding*. 1995; 61(1):38–59.
8. Seshamani, S., Chintalapani, G., Taylor, R. Iterative refinement of point correspondences for 3d statistical shape models. *Proceedings of the 14th International Conference on Medical Image*

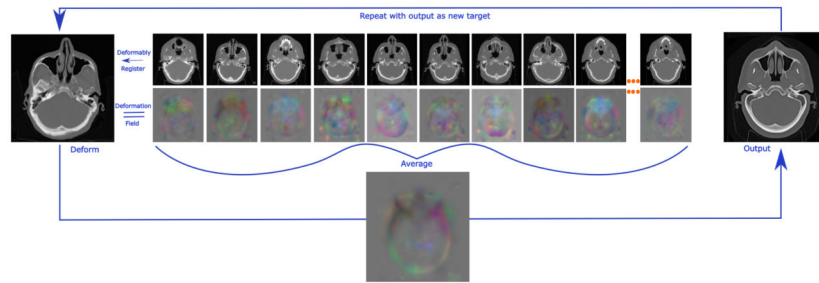
Computing and Computer-assisted Intervention - Volume Part II, MICCAI'11; Berlin, Heidelberg:  
Springer-Verlag; 2011. p. 417-425.

Author Manuscript

Author Manuscript

Author Manuscript

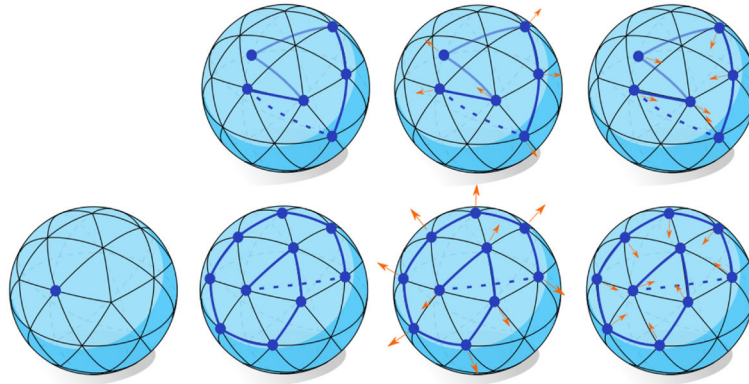
Author Manuscript



**Figure 1.**

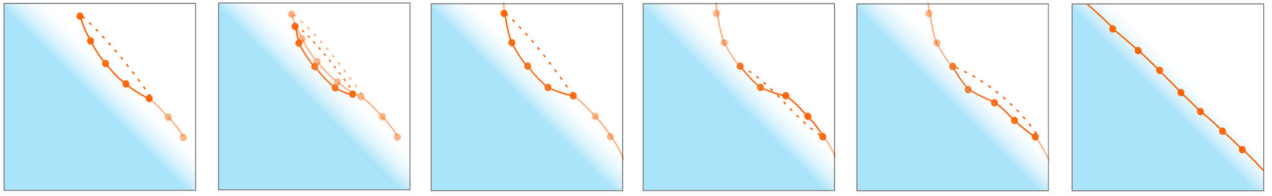
Template creation pipeline: all input images are deformably registered to one target image, which is then deformed by the average of the deformation fields resulting from the registration. The colors in the deformation fields represent the direction of the vectors, whereas the intensity of the colors indicates the magnitude. Deforming the target image by the average deformation field takes the target image towards the mean of our input set of images. This process can be repeated with the output image as the new target image to be registered to. Individual variation from our initial target decreases with every iteration, and the resulting output gets closer to the true average of our input set of images.





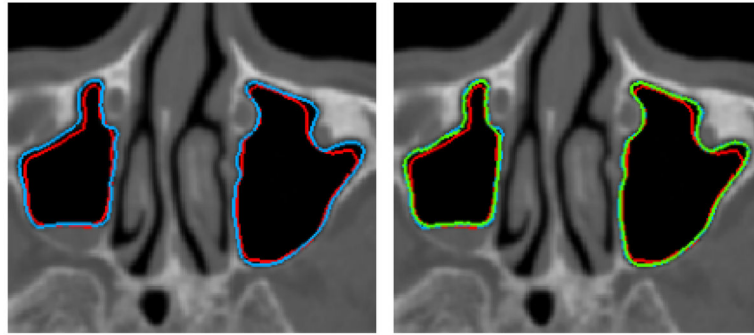
**Figure 2.**

Sampling structure (blue dots represent sampled vertices, blue lines indicate adjacency; dotted lines imply that we do not want the connecting points to be attracted to each other; translucent lines cross through the shape): Random sampling causes inconsistent internal and external directions. The top middle sphere shows a vertex with external energy pointing inwards, and top right shows one with internal energy pointing outwards. Our spiral sampling maintains internal and external direction consistency.

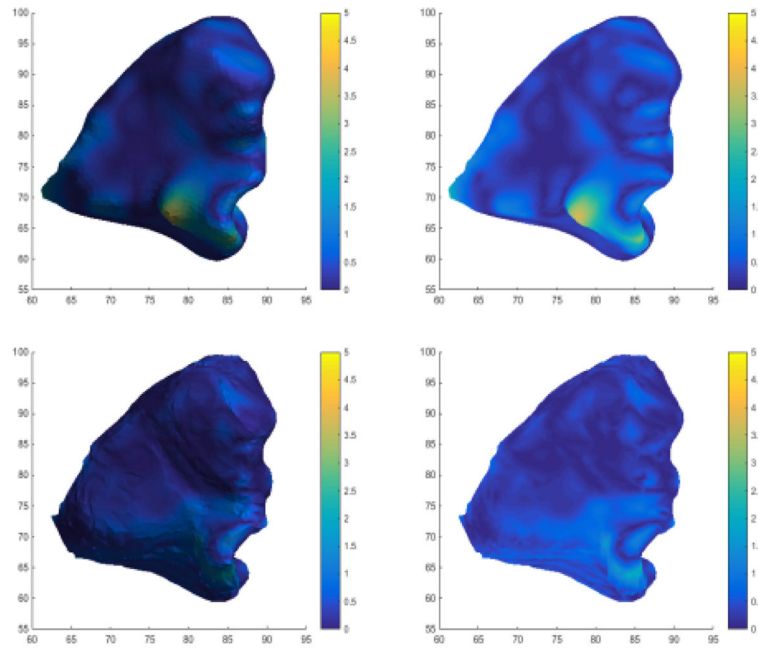


**Figure 3.**

From left to right, we see how the vertices in our snake spline move towards the edge. Since we do not want corner points to get drawn inward toward each other, we do not allow corner points to move. However, they move during following iterations when they are not corner points. Finally, all vertices are on the closest edge.

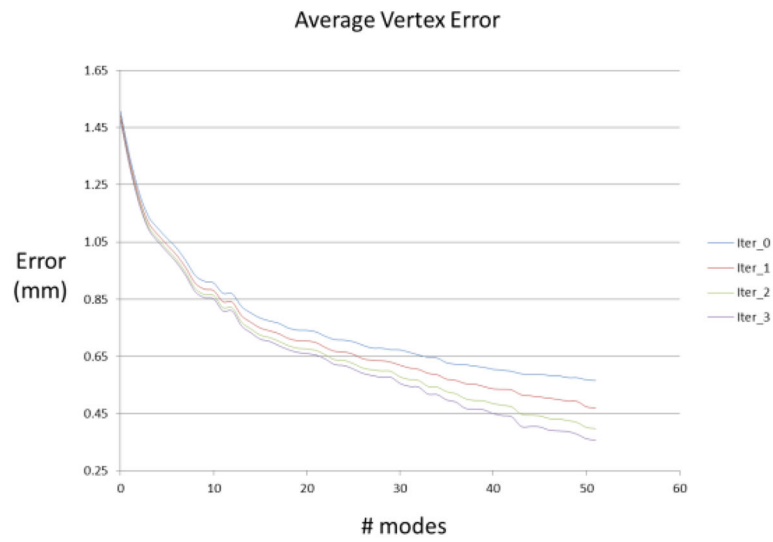


**Figure 4.** Left: Edge maps of hand-segmented left and right maxillary sinuses (blue) and deformedly registered maxillary sinuses (red); Right: Edge map of maxillary sinuses after snakes splines (green) overlaps almost perfectly with the hand-segmented edge map.

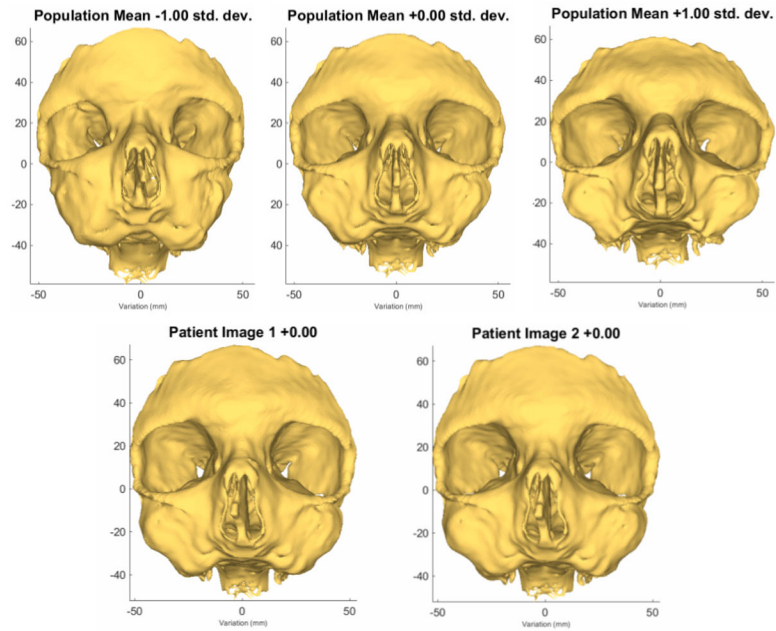


**Figure 5.**

Top: Errors from deformable registration visualized on the mesh with lighting (left), to show structure, and without lighting (right), to show errors without distractions from secularity or shadows; Bottom: Errors after GVF, visualized similarly as above.

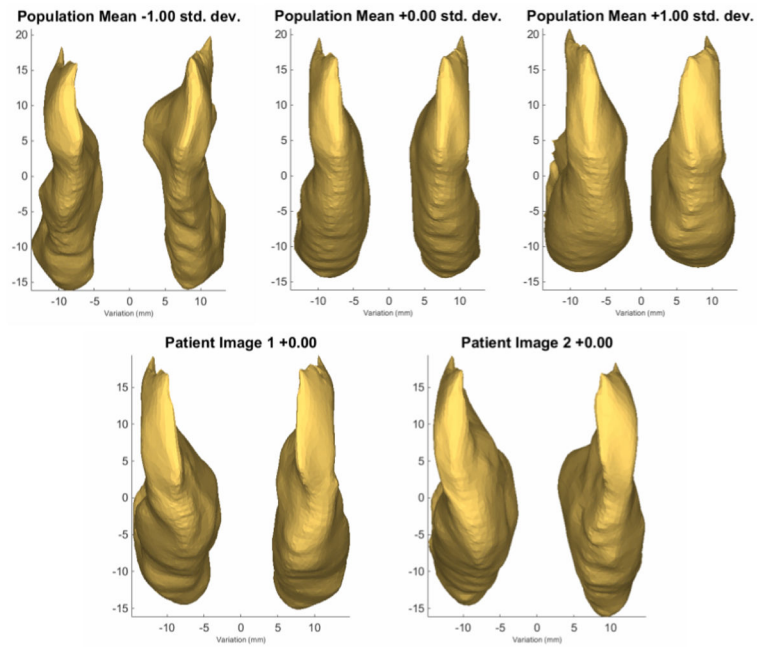


**Figure 6.** Improving vertex correspondence with each iteration for the middle turbinate.



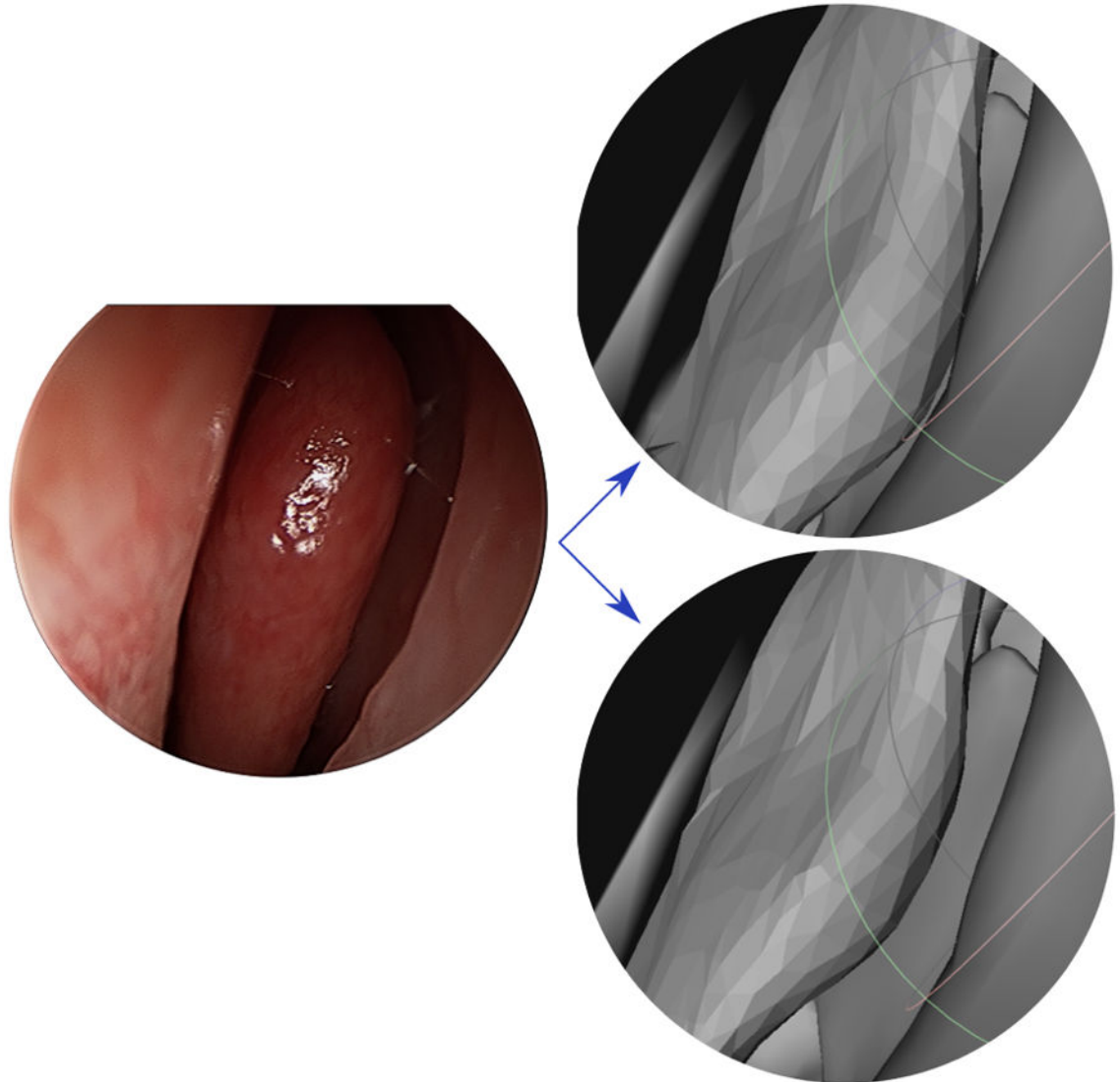
**Figure 7.**

Top: Population variation in the skull model. The middle shape is the mean shape, the left shape shows mean shape with  $-1\sigma$ , where  $\sigma$  is the standard deviation, and the right shape is the mean shape with  $+1\sigma$ . Bottom: The left image shows the pre-op patient skull, and the right image shows the post-op patient skull. The two images show no, or negligible, difference, where minute difference can sometimes be observed due to errors in registration. However, we can see that the population variation is not reflected in the two patient images.



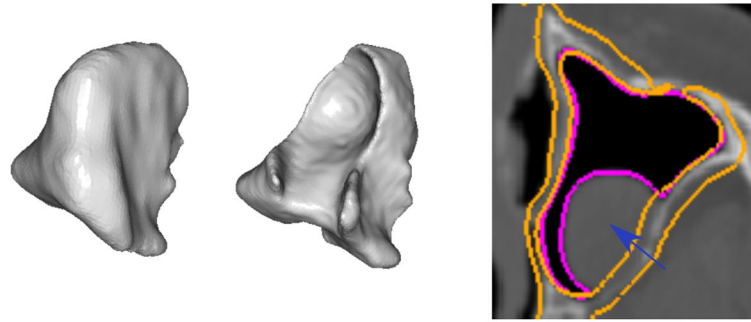
**Figure 8.**

Top: Population variation in the inferior turbinate (IT) model. The middle shape is the mean shape, the left shape shows mean shape with  $-1\sigma$ , and the right shape is the mean shape with  $+1\sigma$ . Bottom: The left image shows the pre-op patient IT, and the right image shows the postop patient IT. The two images show significant differences, allowing us to conclude that the population variation is reflected in the patient images.



**Figure 9.** The image from video (left) looks less like the middle turbinate segmented from CT (top right), and more like the middle turbinate estimated from PCA (bottom right).





**Figure 10.**

This image shows larger than expected difference between the mean (left) and patient (middle), implying disease. The volume between the orange and pink segmentations (right) can be related to the amount of disease.

Mean, minimum and maximum vertex errors (in mm) for deformable registration and after GVF. RMS and LMS indicate right and left maxillary sinuses, and boldface indicates smaller errors.

**Table 1**

		Deformable Registration				GVF			
		Mean	Min	Max		Mean	Min	Max	
Patient 1	RMS	0.4808	<b>0.0000</b>	4.0460		<b>0.3579</b>	0.0003	<b>0.2020</b>	
	LMS	0.3726	<b>0.0001</b>	3.2085		<b>0.2913</b>	<b>0.0001</b>	<b>1.3471</b>	
Patient 2	RMS	0.3419	<b>0.0001</b>	3.8089		<b>0.3401</b>	0.0002	<b>2.3474</b>	
	LMS	0.4161	0.0002	4.5978		<b>0.3115</b>	<b>0.0001</b>	<b>3.7671</b>	
Patient 3	RMS	0.3055	0.0002	<b>1.4191</b>		<b>0.2583</b>	<b>0.0000</b>	1.5183	
	LMS	0.3286	0.0005	2.9758		<b>0.3089</b>	<b>0.0002</b>	<b>2.0215</b>	
Mean	RMS	0.3761	<b>0.0001</b>	3.00914		<b>0.3188</b>	0.0002	<b>0.0226</b>	
	LMS	0.3724	0.0003	3.5940		<b>0.3039</b>	<b>0.0001</b>	<b>2.3819</b>	