

Computer Vision (H02A5A)

Project report

Incisor Segmentation

KRISTOF ARRON - R0467804
ROBIN HAVENEERS - R0450702

PROF. DR. IR. DIRK VANDERMEULEN
2016-2017

KU LEUVEN

Contents

1	Introduction	2
2	Building an Active Shape Model	3
2.1	Generalized Procrustes Analysis	3
2.2	Principal Component Analysis	3
3	Preprocessing the Radiographs	5
4	Fitting the Model to Images	5
4.1	Manual Initialization	6
4.2	Automatic Initialization	7
4.3	Fitting procedure	9
4.3.1	Strongest Edge	9
4.3.2	Grey Level Models	10
5	Evaluation of the Results	10
5.1	Examples of Manual Segmentation	10
5.2	Examples of Automatic Segmentation	11
5.3	Fitting Procedure Evaluation	12
6	Discussion	12
6.1	Preprocessing Alternatives	12
6.2	Automatic Initialization Alternatives	13
6.3	General Discussion	13
	References	14

1 Introduction

This report describes the process of both the manual and automatic segmentation of all eight incisor teeth in different dental panoramic radiographs. We use the Active Shape Model (ASM) method to segment the teeth. First we discuss how to build such a model. We briefly mention the steps that are needed to process the data before a model can be built. Next we analyse the results of ASM via Principal Component Analysis (PCA). This gives us an illustration of the most important modes of variation within the shape of each tooth. A next section describes the preprocessing steps that are taken on the radiographs in order to reduce noise and improve contrast. Thereafter we explain how to fit the model to new images, either manually or automated. We end the report with a short evaluation of the results and a brief discussion on the project.

2 Building an Active Shape Model

2.1 Generalized Procrustes Analysis

In this first part of building the Active Shape Model, it is important that the given landmarks are represented in the same coordinate frame. The shape of the landmarks should be independent of the position, scale and orientation of the tooth. This is where we applied the Procrustes Analysis. As described in [Cootes et al., 2000], we align the set of training shapes into a common frame. The goal is to align each shape such that the sum of distances of each shape to the mean shape is minimised.

Below, on the left hand side in figure 1 you can see a plot of all the landmarks of the first incisor. On the right hand side, in figure 2, you can see the same shapes, centered around the origin, having a mean scale of unity and having a fixed but arbitrary orientation.

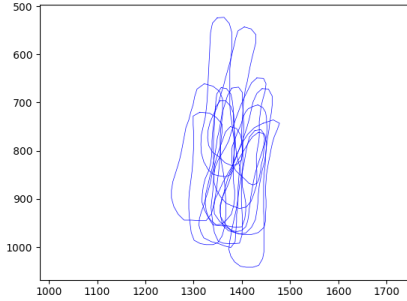


Figure 1: Unaligned shapes
of tooth 1

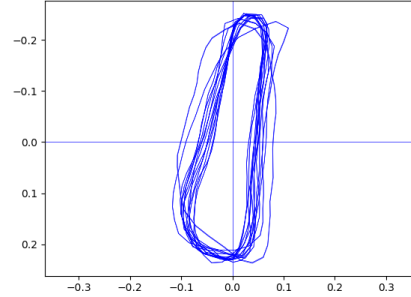


Figure 2: Aligned shapes
of tooth 1

2.2 Principal Component Analysis

Once all shapes are aligned, we can perform Principal Component Analysis on these new shapes. The result of this analysis is a set of principal components or modes of variation. These modes describe the directions with the largest variance within the data. Figure 3 shows the three most important principal components and the shape variations they allow. PCA allows us to reduce the dimensionality of the dataset by keeping only a subset of the principal components, while still explaining a high percentage of the variance. In this application, the amount of modes needed to explain over 99%

of the variance is determined dynamically for each tooth. However, usually about 7 to 9 modes are needed, table 1 shows this for tooth 1 and 8, other teeth are similar.

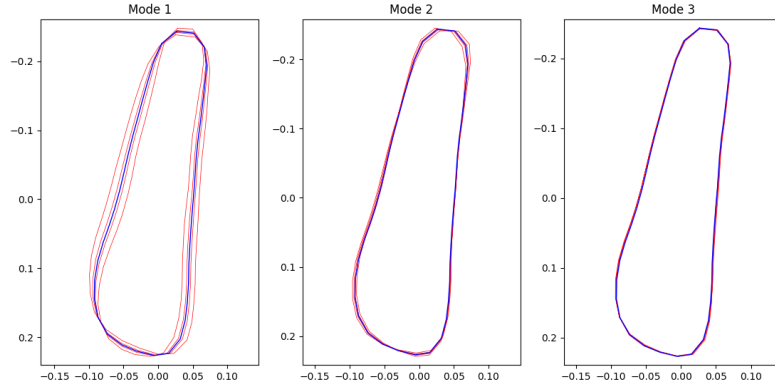


Figure 3: Variation of the 3 most important principal components. The blue shape is the mean shape, the red shapes indicate possible variations.

	Tooth 1	Tooth 8
1	0.646400530498	0.42062844717
2	0.898917160953	0.752692890297
3	0.949902073345	0.868618769304
4	0.971813974816	0.927944903926
5	0.980536831643	0.957612924558
6	0.98736569597	0.972696635066
7	0.991154402009	0.981115590281
8		0.987446059787
9		0.992000005508

Table 1: Examples of the amount of principal components needed to reach 99% variance.

3 Preprocessing the Radiographs

Radiographs are prone to noise. In order to use the images for detection, we need to reduce the noise. In our application, we apply different filters to the images according to tactics described in [X. Wanga, 2005] and [Alavijeh and Mahdavi-Nasab, 2015]. All images are first filtered with a Gaussian filter of size 3. This reduces noise but also blurs the image, making edges less distinct. The second filter is a median filter of size 7. This filter aims to reduce noise while keeping edges clear. The next filter is a hat-filter. This is composed of two separate filters. A white hat and a black hat filter. The black hat filter enhances dark features, while the white hat filter enhances bright features. The last filter that is applied to the radiographs is an adaptive histogram equalisation filter. This filter increases contrast.

The result of applying these filters is an image that is more suitable for edge detection. This is advantageous since we primarily use the strongest edge method to fit the tooth model to the images. The last step in the process is to detect the edges with the Sobel operators. Examples of the results of each step of the filtering process can be seen in figures 4, 5 and 6. The median and Gaussian filter steps are not included.



Figure 4: Result of
hatfilter

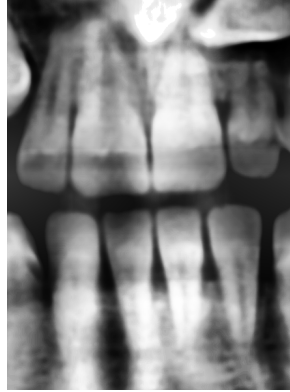


Figure 5: Result of
AHE filter

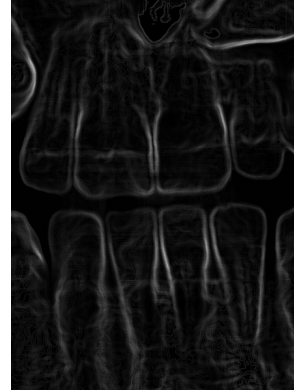


Figure 6: Result of
Sobel detector

4 Fitting the Model to Images

Fitting the found model to the images starts by finding an initial estimate for the model in the image. We developed a manual and an automatic initial-

ization method, as described below. Next, the model is fitted to the image in an iterative manner. Our approach is primarily based on the methods described in [Cootes et al., 2000] and [Cootes et al., 1995].

4.1 Manual Initialization

In the manual initialization, the user is presented with the radiograph of the person for whom he is trying to fit the model, as can be seen in figure 7. He can use his mouse to drag the shape to the correct incisor and when a key is pressed, the current selection is used. The selection defines 40 initial landmark coordinates.

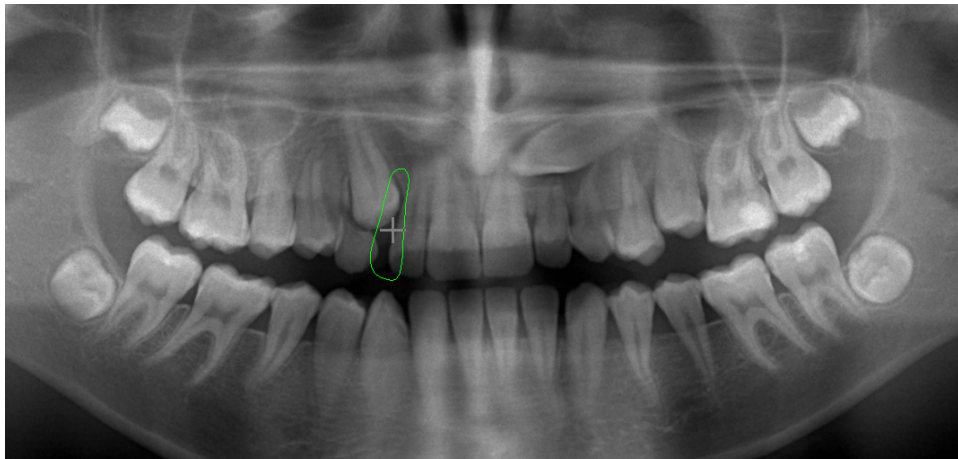


Figure 7: Manual initialization

4.2 Automatic Initialization

Our automatic initialization is a bit more complex. It consists of different steps, described (broadly) in algorithm 1 below.

Algorithm 1: Automatic initialization

- 1 Find mean position of tooth over all test images.
 - 2 For each test image, create a template of the incisors as a whole.
 - 3 Crop the input image to the mean position, found in step 1.
 - 4 For each template, find the score of the template matching the cropped image from the previous step.
 - 5 For each template, find the best matching location in the cropped image. Given the score,
 - (a) If the score is $\geq 99\%$, break and used this score and this location.
 - (b) Else, register this score and location.
 - 6 For each score and location add score*location to the location we are going to return, such that we have a weighted average. If step (a) has occurred, the location of only that template is returned and no weighted average is calculated.
 - 7 Take the landmarks associated with the templates and average them out using the scores obtained from the previous step. If step (a) was followed, it will use one set of landmarks.
 - 8 Return the necessary landmark.
-

We will now go deeper into the steps described in algorithm 1. We want to find an initial position to place the model we want to fit on the radiograph. To achieve the optimal position, we perform a number of steps. Firstly, all training landmarks are loaded and a minimal fitting bounding box is calculated, such that we have a bounding box over all test images in which all teeth are located.

Next, we load all the test images and we crop the image to the bounding box created from the landmarks associated with this image. The resulting image thus contains all the 8 incisors for the image.

Now we load the image of the person we are trying to fit which we crop to the global bounding box, calculated in the first step, assuming that in this new image the teeth are distributed in a similar fashion as in the training images. We now have a number of templates and the input image all cut out (as adequately as possible) to only have the incisors in the image.

In this next step we use OpenCV's `matchTemplate`¹. This method will take the templates created and slides them over the image that you want to match. In our case the teeth templates are matched against the input image. It returns, amongst others, the matching score (as a percentage) and the 'best matching location' as an offset compared to the image over which you are sliding.

Next, we evaluate the scores and location that were obtained before. Firstly, if the method finds a score that is greater than or equal to 99%, this location and score will be used solely, without considering the rest. Otherwise, all scores and locations are registered. We calculate an average location by calculating the weighted average of all the locations found when matching the template using the associated scores as the weights. We now obtain an average location of where the teeth in this image are located.

Lastly, we take all the landmarks associated with the templates we used before and translate them such that they match the location/offset calculated in the previous step. Again, if there was a template with a match greater than 99%, we only use this landmark. If not, we average the landmarks, again, using the scores from before, such that we obtain a set of average landmarks of all the teeth for this image. We now return the landmark of the tooth we have to initialize.

Below you can see the automatic initialization for two example images, which are not in the training data.

¹http://docs.opencv.org/2.4/modules/imgproc/doc/object_detection.html

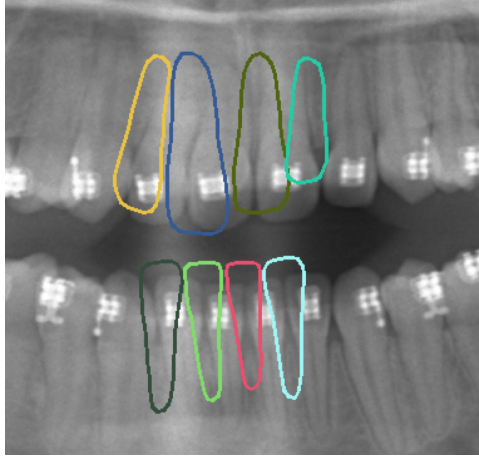


Figure 8: Automatic initialization for person 27

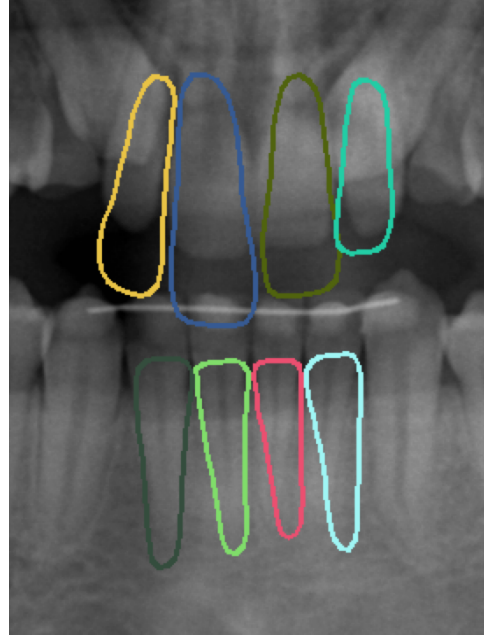


Figure 9: Automatic initialization for person 17

For the teeth in figure 8, this method works adequately, however this is not the case in figure 9. This is due to the fact that this radiograph is slightly different from the ‘average radiograph’: the teeth are positioned a bit higher. In hindsight, we realise that this approach is not the most optimal one. In general: if we present a radiograph which has different proportions or where the teeth are aligned in a very different manner, this initialization behaves poorly.

4.3 Fitting procedure

4.3.1 Strongest Edge

As proposed in [Cootes et al., 2000], we could fit the model to a new image by moving the initial points to the strongest edge in their surrounding. To find this edge, we consider a region of k pixels along the normal on the curve defined by the landmarks. The position with the strongest edge is then selected as the next position for the landmark point.

This method has its flaws, since the real edge of a tooth is not necessarily

the strongest edge present. Neighbouring teeth may have stronger edges, causing the model to be malformed. Another problem occurs when fitting the root of a tooth. The roots are less clearly delineated and therefore harder to fit to.

4.3.2 Grey Level Models

Another method proposed in [Cootes et al., 2000] is to include local grey level information into the fitting procedure. For each landmark, a profile of the grey levels of a region along the normal is created. When fitting, a similar profile is extracted and matched with the ground truth profiles. The point with the best match is then selected as the next point to move towards.

5 Evaluation of the Results

5.1 Examples of Manual Segmentation

Figure 10 shows the result (in blue) of the segmentation when using the strongest edge method and with manual initialisation. Figure 11 shows the result (in blue) of the segmentation when using the grey level models and manual initialisation. Both figures are from radiograph 03. The green shapes show the ground truth for each tooth.



Figure 10: Segmentation with manual initialisation and strongest edge fitting



Figure 11: Segmentation with manual initialisation and grey level model fitting

5.2 Examples of Automatic Segmentation

Figure 12 shows the result (in blue) of the segmentation when using the strongest edge method and with automatic initialisation. Figure 13 shows the result (in blue) of the segmentation when using the grey level models and automatic initialisation. Both figures are from radiograph 02. The green shapes show the ground truth for each tooth.

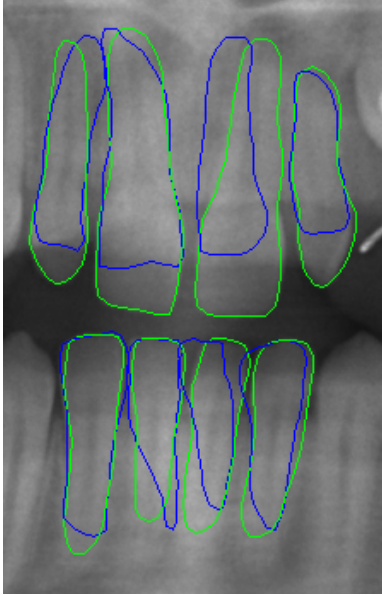


Figure 12: Segmentation with automatic initialisation and strongest edge fitting

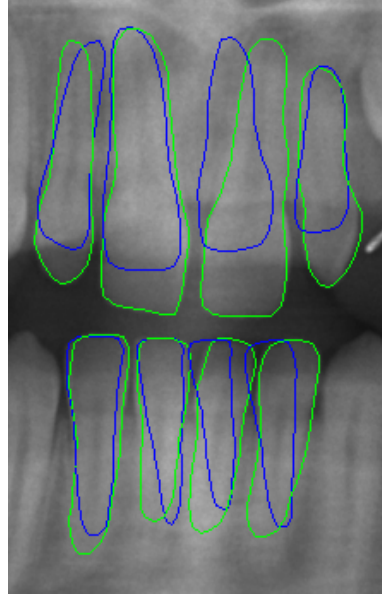


Figure 13: Segmentation with automatic initialisation and grey level model fitting

5.3 Fitting Procedure Evaluation

As mentioned in section 4.3.1, fitting a model to a new image using the strongest edge has some drawbacks. To combat this issue, we tried to implement the grey level model technique as described in section 4.3.2. However, this method did not produce great results. In most cases the shape would shrink or expand out of control. For this reason, the strongest edge method - while being a poor fitting method - still performed better than the grey level model method. Possible explanations for this poor performance might be a bad implementation of the method, or a badly chosen stop condition (the stop condition used to indicate whether or not convergence has occurred.)

6 Discussion

6.1 Preprocessing Alternatives

A number of different filtering techniques could be applied to the radiographs in order to reduce noise and enhance the images. For example, the Radiographic Impulsive Noise removal filter proposed by Iuri Frosio

([I. Frosio, 2009]) could be applied. Other filters such as a bilateral filter or different histogram equalization methods could also be applied, depending on the desired result. For our application, the currently used filters sufficed.

6.2 Automatic Initialization Alternatives

Our approach for automatic initialization behaves poorly when the radiograph presented is different from the radiographs in the training set. In our initialization we try to match the template for the 8 incisors as a whole. In this regard we should have opted to match each teeth separately or, perhaps even better, separately for the upper and lower jaw. An alternative to using the template matching method provided by OpenCV, we could have tried to use a more general approach using PCA.

Another possibility could have been to create a detection algorithm, such as the face detection algorithm, and use that for detecting teeth in a bounded area of a picture. This would only be possible if we had a larger set of teeth to our disposal. In the same manner we could have opted to create a (deep) learning algorithm to detect teeth in an image.

6.3 General Discussion

Overall, we feel like the solution we came up with performs well. Of course, using manual initialization this makes sense: you practically ‘tell’ the program where to search for the teeth. In hindsight, our automatic initialization might have been constructed using an approach that is too brute force or ‘hacky’, if you will. To conclude: we are positive about our approach and along the way we have learned a lot of new techniques and approaches regarding detection algorithms and computer vision.

References

- [Alavijeh and Mahdavi-Nasab, 2015] Alavijeh, F. S. and Mahdavi-Nasab, H. (2015). Multi-scale morphological image enhancement of chest radiographs by a hybrid scheme. *Journal of Medical Signals and Sensors*, 5:59—68.
- [Cootes et al., 2000] Cootes, T., Baldock, E., and Graham, J. (2000). An introduction to active shape models. *Image processing and analysis*, pages 223–248.
- [Cootes et al., 1995] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J. (1995). Active shape models - their training and application. *Comput. Vis. Image Underst.*, 61(1):38–59.
- [I. Frosio, 2009] I. Frosio, N. A. B. (2009). Statistical based impulsive noise removal in digital radiography. *IEEE Transactions on Medical Imaging*, 28(1):3–16.
- [X. Wanga, 2005] X. Wanga, B. S. W. (2005). Image enhancement for radiography inspection. *Proceedings of SPIE - The International Society for Optical Engineering*, 5852.