

Computer Vision [H02K5a]

Final Project: Incisor Segmentation

Pieter Robberechts (r0296915)
2Ma CW

June 2016

1 Introduction

This report discusses my solution for the Incisor Segmentation assignment of the Computer Vision course. The goal of this project was to develop a model-based segmentation approach, capable of segmenting the upper and lower incisors in panoramic radiographs. Section 2 explains the construction of an Active Shape Model. Section 3 discusses the different pre-processing steps that were used to enhance the quality of the raw radiographs. Next, Section 4 discusses how the previously generated models are fitted to the enhanced radiographs, completing the segmentation algorithm. This section is split up into two subsections: Section 4.1 first discusses how a rough initial estimate is found for the model position. Section 4.2 then explains how this initial estimate is iteratively improved to better match the incisor, resulting in a final solution. Finally, Section 5 uses leave-one-out analyses to evaluate the complete algorithm and discusses the achieved results.

2 Active Shape Model

The Active Shape Model (ASM) represents a parametric deformable model where a statistical model – called the Point Distribution Model (PDM) – of the global shape variation in the training set of incisors is built. The PDM is then used to fit the mean shape of an incisor to unseen radiographs. The PDM determines which shape variations are plausible.

The symmetry in the dental radiographs was exploited to artificially increase the number of training examples. When building a model, the number of examples of each incisor was doubled by considering the mirrored version of the matching incisor too.

This section is divided into two subsections: Section 2.1 first describes the process that *normalizes* the raw landmark data used to base the Active Shape Model (**ASM**) on. Section 2.2 then explains how a Principal Component Analysis (PCA) is used to derive the PDM for the ASM.

2.1 Pre-processing Landmarks

Each incisor from the training set is represented by a set of 40 landmark points. Before we can perform a statistical analysis on the distribution of these points it is important that all these points are aligned in the same co-ordinate frame. As suggested, this is done by using a Procrustes Analysis. The algorithm is outlined in Appendix A of Cootes et al. [4]. The idea is to translate, rotate and scale each shape so that the sum of distances of each shape to the mean ($D = \sum |x_i - \bar{x}|$) is minimized.

Figure 1 shows the aligned landmarks for incisors 1 (top-left) to 8 (down-right). Each plot shows the mean shape as a thick black line and each set of aligned landmarks as a thin colored line. The mean shape is clearly the mean of all aligned shapes, and the aligned shapes seem to have a matching scale and orientation.

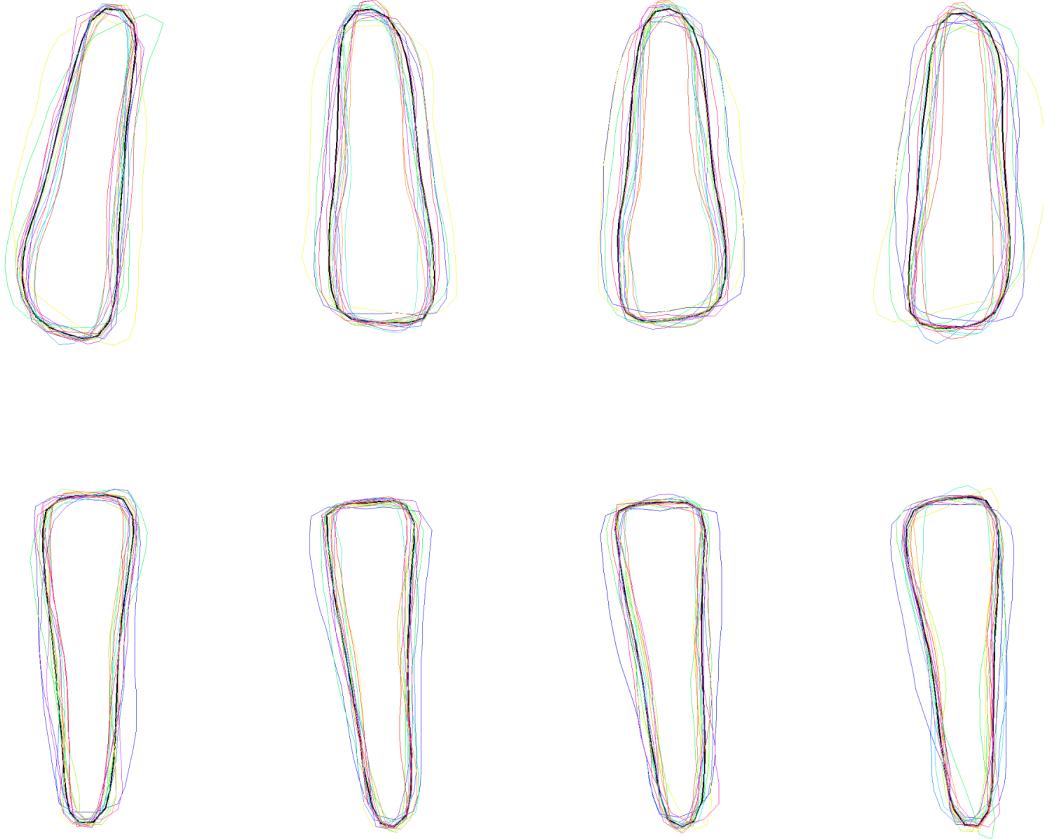


Figure 1: Aligned landmarks for incisors 0 (top-left) to 7 (down-right). Each plot shows the mean shape as a thick black line and each set of aligned landmarks as a thin colored line.

2.2 Principal Component Analysis

The aligned landmarks of each incisor are now used to build an ASM. The landmark points of each incisor form a distribution in a high-dimensional space. By applying Principal Component Analysis (PCA) to the data, the dimensionality is reduced to something more manageable. PCA computes the main axes of variance, allowing one to approximate any of the original shapes with a concise model. Again, for a complete description of the PCA algorithm we refer to Cootes et al. [4]. Figure 2 shows the four largest components for incisor 1. These images clearly show that incisors vary mostly in width and hardly in length. The first component explains the width of the tooth, the second component explains the skewness and the third commonest explains the ratio between the widths of the root and crown. These interpretations hold for every model. The other

components explain minor shape variations, which are different for each incisor. These first four components explain about 96% of the variance. We retained the first eight components which explain 99% of the total variation.

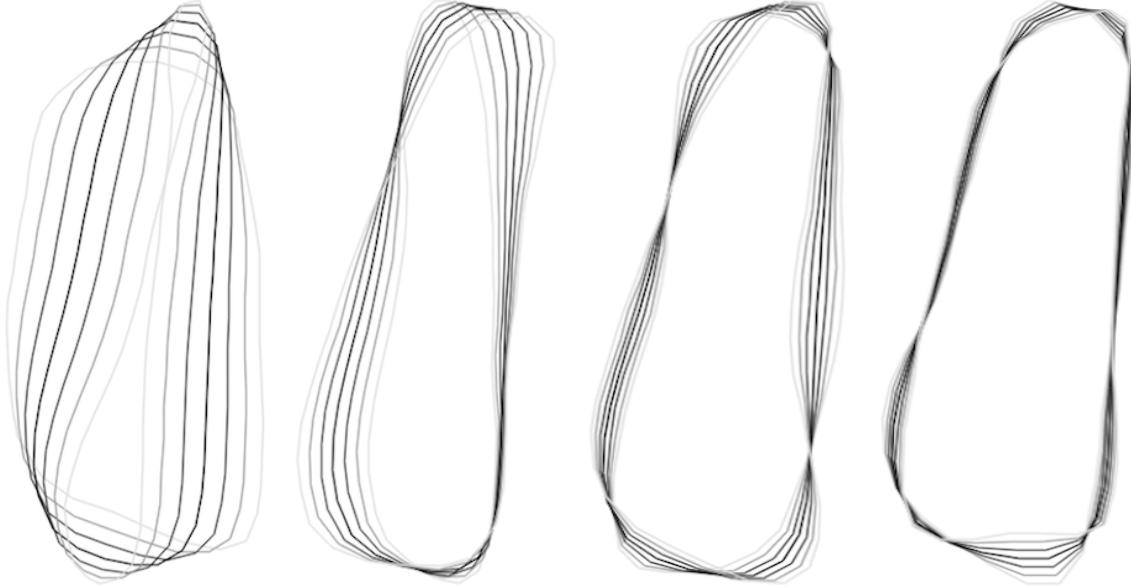


Figure 2: The four strongest principal components in the first incisor’s model. Each plot varies the component between -3 and 3 standard deviations where the darkest - center - shape represents the mean shape.

3 Pre-processing Dental Radiographs

As suggested in the project description, I first looked up some literature regarding the pre-processing of (dental) radiographs. Ahmad et al. [2, 1] compare four techniques for enhancement of dental x-ray images, namely Adaptive Histogram Equalization (**AHE**), Median Adaptive Histogram Equalization (**MAHE**), Contrast Limited Adaptive Histogram Equalization (**CLAHE**) and Sharp Contrast Limited Adaptive Histogram Equalization (**SCLAHE**). Results show that **CLAHE** outperforms the original and the other techniques in enhancing certain pathologies and as such improves the manual diagnostic abilities of the radiograph. Additional and more recent research in [3] shows that **CLAHE** outperforms both **SCLAHE** and **MAHE** for statistical research, which makes it a perfect candidate for this project.

CLAHE is a combination of Adaptive Histogram Equalization (**AHE**) and Contrast Limiting (**CL**). **AHE** divides an image into small 8×8 blocks called “tiles”. The histograms of each of those are then equalized. This allows the algorithm to adapt to local variations in an image like shadows and highlights that would not be treated by global histogram equalization. Using small tiles could lead to noise amplification. To avoid this, **CL** is applied before equalization. **CL** randomly distributes the values of pixels above a certain contrast level among the existing bins. In this project I used the **OpenCV** implementation of **CLAHE**.

The actual pre-processing algorithm used was based on [5]. Before applying **CLAHE**, this algorithm first uses an adaptive median filter and a bilateral filter to suppress the impulsive and Gaussian noise in the image. Afterwards, the image is sharpened by combining the bottom-hat

and top-hat transform of the image. The top-hat transform is able to enhance brighter structures, while the bottom-hat transform enhances the darker structures. The final step is to apply CLAHE. Figure 3 shows the flow chart of this algorithm. As the adaptive median filter is very slow, it was applied once on each radiograph and the results were stored. The other pre-processing actions are performed during execution of the incisor segmentation algorithm.

Figure 4 shows the different pre-processing steps for the upper incisors of Radiograph01. Contrast has clearly been improved, making the boundaries that distinguish teeth more clear.

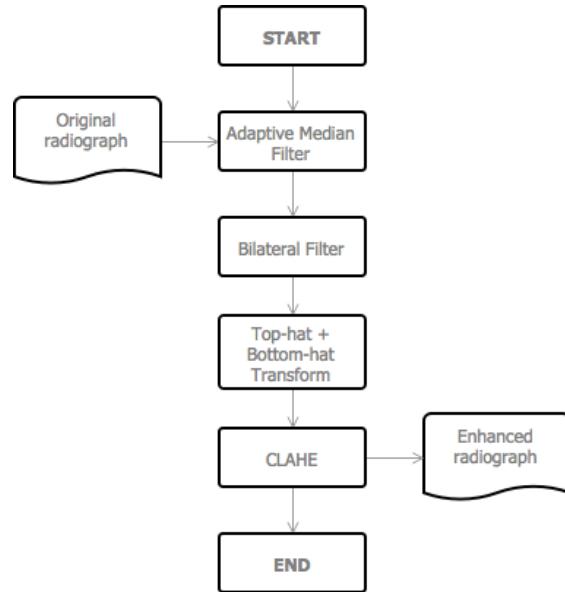
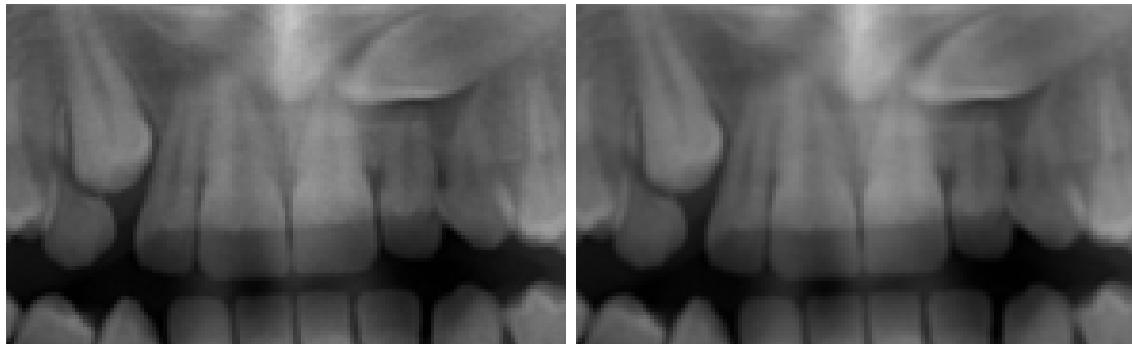


Figure 3: The processing flow of the algorithm in [5], used in this project.



(a) The original image

(b) The adaptive median and bilateral filtered image

Figure 4: The different steps of the radiograph pre-processing algorithm

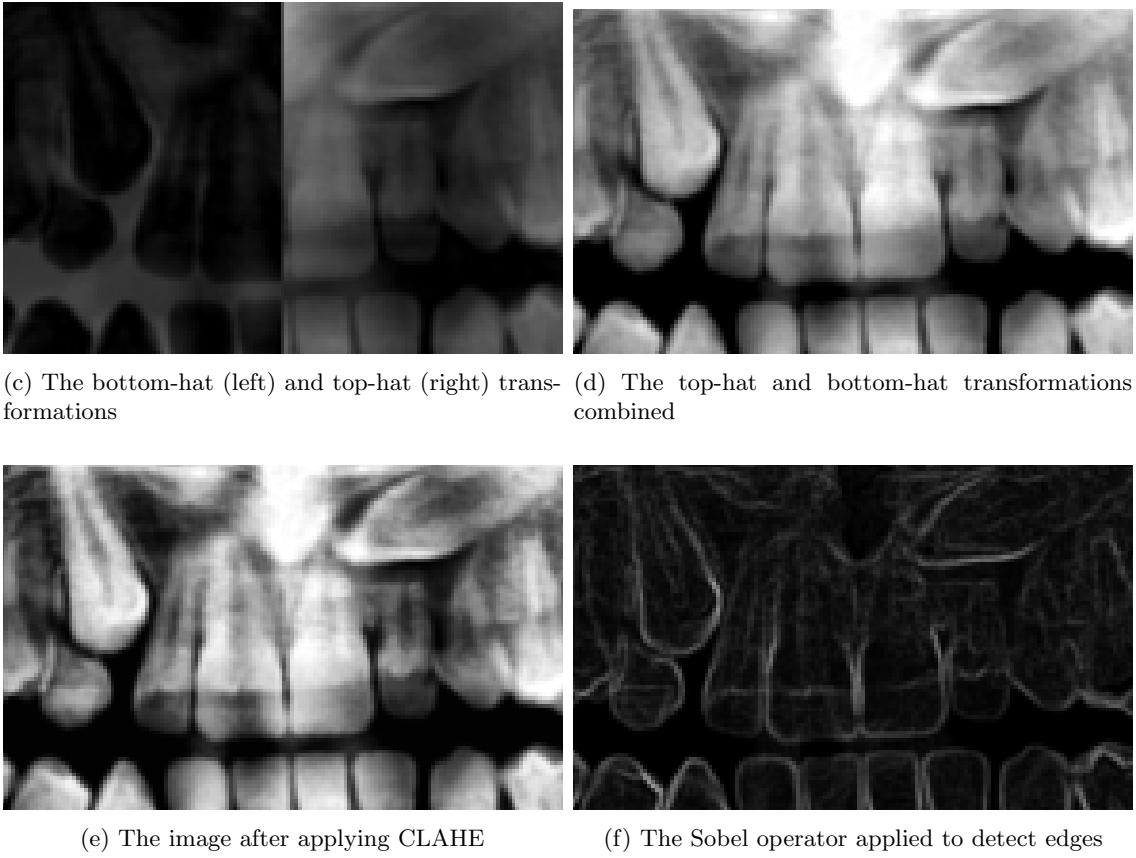


Figure 4: The different steps of the radiograph pre-processing algorithm

4 Model Fitting

Fitting a model to an image happens in two stages. Section 4.1 describes the first stage of looking for an initial model estimate in the test image. This estimate is used to initiate an iterative fitting scheme that improves the model to better match the image, as explained in Section 4.2.

4.1 Initial Estimate

Algorithm 1 outlines the main steps that are undertaken to find an initial estimate. The algorithm is further fleshed out in the following paragraphs.

- First, the split between the upper and lower jaw is detected. This split provides a hint for the vertical position of the teeth. To find the gap, intensity histograms are calculated at intervals over the width of the top-hat transformed radiograph. These histograms are then smoothed using a Fourier transformation. For each interval the four maxima in the histogram are kept. These represent points in the image through which the jaw split might pass. Next, paths are generated between these points. For each point, an edge to a second point is added, such that the intensities along the edge are minimal. Finally, the path with the lowest total intensity corresponds to the jaw split. This approach could detect the jaw split successfully on each training image.

Algorithm 1: Initial Estimate Finding Algorithm

Step 1 Locate the jaw split.

Step 2 Build an appearance model for the four upper incisors and one for the four lower incisors using PCA.

Step 3 Use a moving window at different scales to detect the region of the four upper and the region of the four lower incisors in the enhanced radiograph.

Step 4 Split both found regions in four equally large parts. This results in a bounding box for the initial position of each incisor.

Step 5 Position and scale the mean shape of each incisor in its corresponding bounding box.

- Second, an appearance model for the four combined upper incisors and one for the four combined lower incisors was created from the enhanced (not gradient) radiographs. We extracted the region of the four upper incisors and the region of the four lower incisors from all thirty provided radiographs. By applying a PCA on these regions, a model for the four upper incisors and one for the four lower incisors was built.
- During the initial fitting process, a moving window of different sizes is passed over the image. The search space is limited to the central part of the radiograph and the region right above (to detect an upper incisor) or right below (to detect a lower incisor) the jaw split. Each window is projected onto the first five principal components. The window with the smallest reconstruction error is used as an initial guess for the region of the four incisors.
- Fourth, we assume that each incisor has more or less the same width. Therefore, to create an estimate for each individual tooth, the region is simply divided into four equal parts. This results in a bounding box around the estimated region of the incisor we are looking for.
- Finally, the mean shape of the incisor is translated and scaled to fit within the resulting bounding box.

4.2 Fitting Estimate to an Image

After finding an initial estimate as described above, we can fit this estimate to a new image. The iterative fitting method used for this project closely follows Protocol 2 of Cootes et al. [4] and is outlined in Algorithm 2. We refer to Cootes et al. [4] for the details of the algorithm. In the next paragraphs we will highlight some small changes we made to the base algorithm.

Removing outliers To find *the best point X_i in the neighboring image region of X'_i* (**Step 2**), a grey-level model of the region around the corresponding model point is used. To create the grey-level model we sample the normalized gradient along a profile k pixels either side of the model point in each training image. Figure 6 shows such a grey-level model. Next, during search we sample the same kind of profile in $m - k$ candidate points either side of X'_i . The candidate which matches best with the grey-level model is used as X_i . Typically, a couple of X_i s are fitted wrong, causing spikes in the surface of the shape formed from the combination of all these X_i s. This complicates **Step 3**. To straighten the shape we store the index of the retained candidate

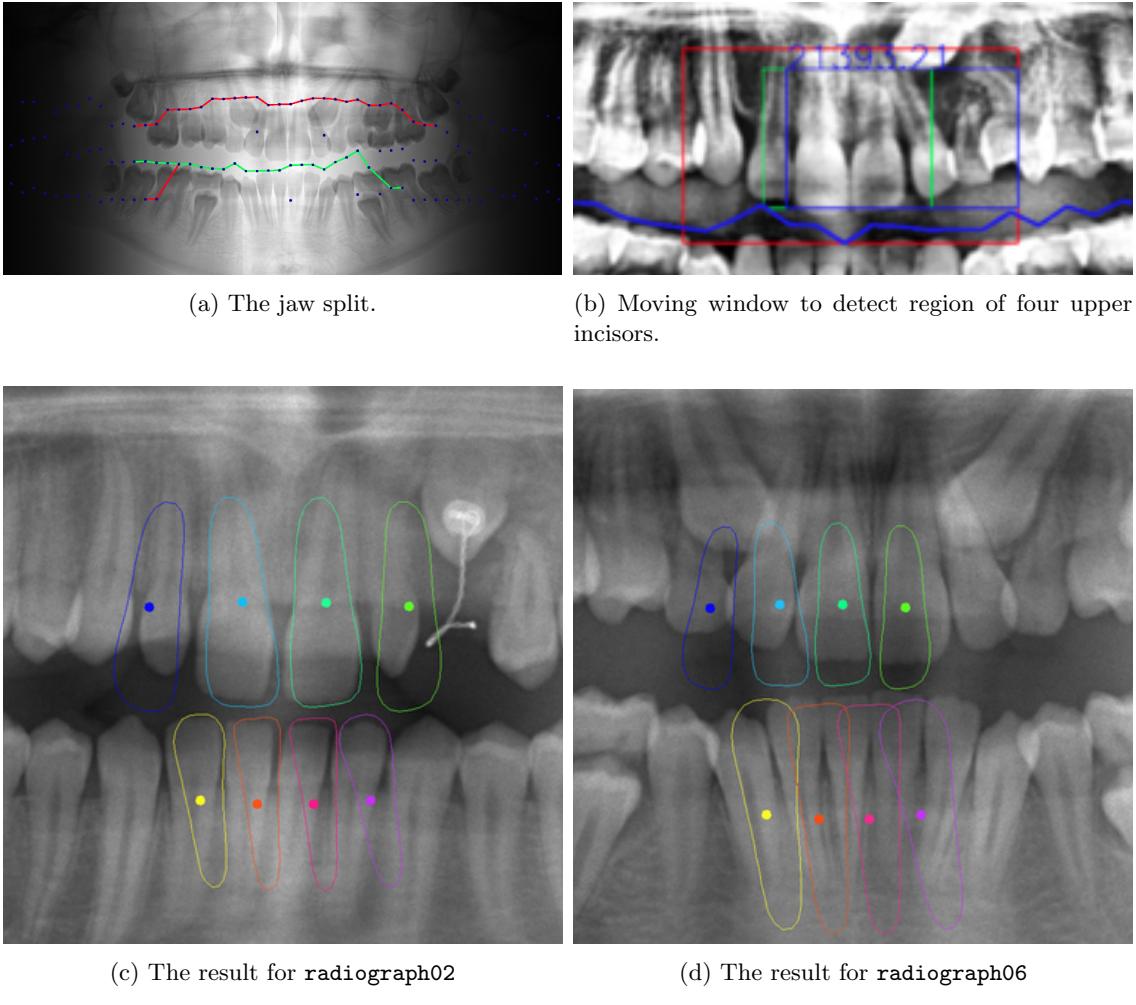


Figure 5: The different steps taken to find an initial model estimate and two results.

points (a number between 0 and $2(m - k)$) in an array and apply a median filter of length 5. This straightens the shape and removes outliers.

Determining alignment parameters **Step 3** updates the model parameters to best fit the newly found image points X (now called Y). Here we use the approach as suggested in Protocol 1 of [4]. We only consider the crown part of the tooth to determine the translation ($[X_t, Y_t]$), scale (s) and rotation (θ) parameters. The crown of the incisors is mostly easy to distinguish on the radiographs, while the root is not. There exists some correlation between the length of a tooth and the width of the crown, such that the shape of the root can often be approximated by only fitting the crown. The shape parameters b are computed for both the crown and root.

Ensuring plausible shapes To ensure plausible shapes we allow $|b_i|$ to vary no more than three standard deviations, we limit scaling to 20% increase or decrease compared to the mean shape and limit the rotation to 45° .

Algorithm 2: Model Fitting Algorithm

- Step 1** Use the previously obtained initial estimate as the first current fit X'_i .
- Step 2** For each point X'_i of the current fit, find the best point X_i in the neighbouring image region.
- Step 3** Update the model's parameters (X_t, Y_t, s, θ, b) such that the new fit X is best approximated.
- Step 4** Apply constraints to shape parameters b in the object-centred coordinate frame. Ensure plausible shapes by allowing $|b_i|$ to vary no more than three standard deviations.
- Step 5** If not converged, then re-iterate from **Step 2**.
-

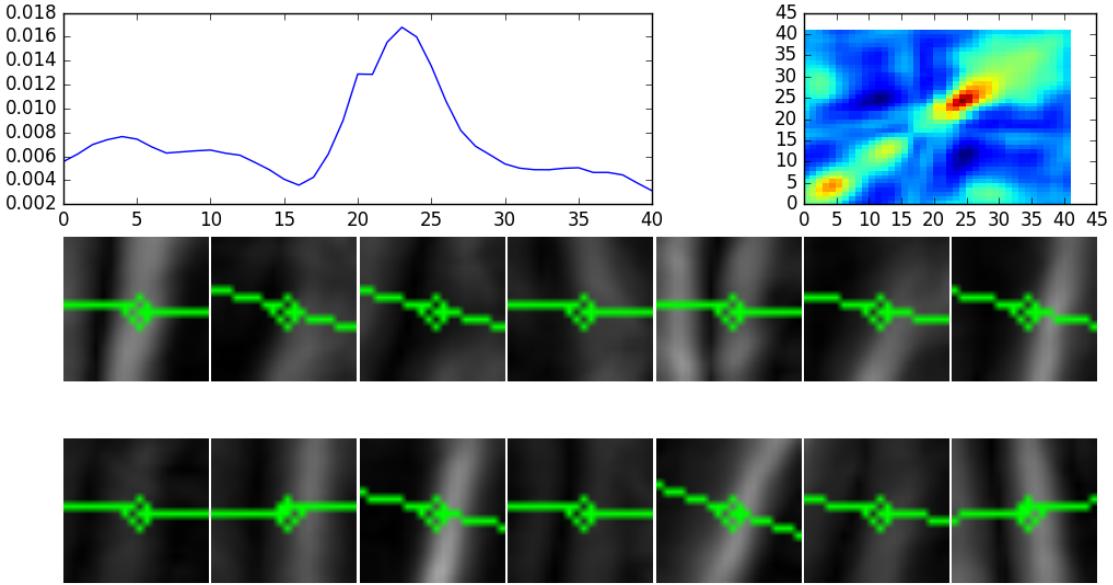


Figure 6: The grey-level model for the tenth model point of incisor 2. Top-left shows the average profile, top-right the correlation matrix and the small images on the bottom show the profiles extracted from all training images from which the model is build.

Multi-resolution framework To improve the efficiency and robustness of the algorithm, it is implemented in a multi-resolution framework. A Gaussian image pyramid of two levels is constructed. This allows the algorithm to first find a rough estimate of the location in a coarse image, then refining the location in a high resolution image. Below level two, incisor boundaries already start flowing over in each other which makes fitting impossible. Therefore, we only use a two-level pyramid.

Stopping criterion To determine when to change to a finer resolution or stop the search we record the number of times that the best found pixel along a search profile is within the central 25% of the profile. Here again, we only consider the points outlining the crown of the tooth. When

90% of the points are so found, the algorithm is declared to have converged at that resolution. If after 50 iterations the algorithm is still not converged, it probably diverged and we backtrack to the best fit found during these iterations.

5 Result Evaluation

This section evaluates the algorithm by means of *Leave-One-Out Analyses* and discusses the results. Each one of the 14 provided images is used once as the testing image, while the others are used as a training set. Section 5.1 discusses the used evaluation metric, while Section 5.2 evaluates the fitting algorithm.

5.1 Evaluation metric

We evaluated segmentation results by assessing its consistency with the ground truth segmentation. Therefore, we used the F- measure. Denote by P and R the precision and recall values of a particular segmentation then the F-measure is defined as

$$F = \frac{2 * R * P}{P + R}$$

Or in terms of True Positives (TP), False Positives (FP) and False Negatives (FN) as

$$F = \frac{2 * TP}{2 * TP + FP + FN}$$

The F-measure reaches its best value at 1 and worst at 0. It is illustrated visually in Figure 7.

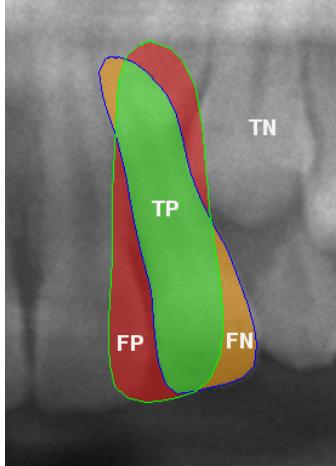


Figure 7: Graphical illustration of the evaluation metric. The ground truth is indicated with a blue line, the estimated fit with a green line. Red is mismatched (False Positive), orange should have been matched (False Negative) and green is correctly matched (True Positive). The F-measure is the ratio between the green zone and the green+orange+red zone.

5.2 Fitting Evaluation

To evaluate the implemented algorithm we performed a leave-one-out cross validation on the training examples. Figure 8 shows the F-measure for the estimated fit of each example. An

initial position of the model was estimated with the method discussed in Section 4.1 and next iteratively improved with the algorithm from Section 4.2. Used parameters are $k = 10$, $m = 15$, $max_{iter} = 50$.

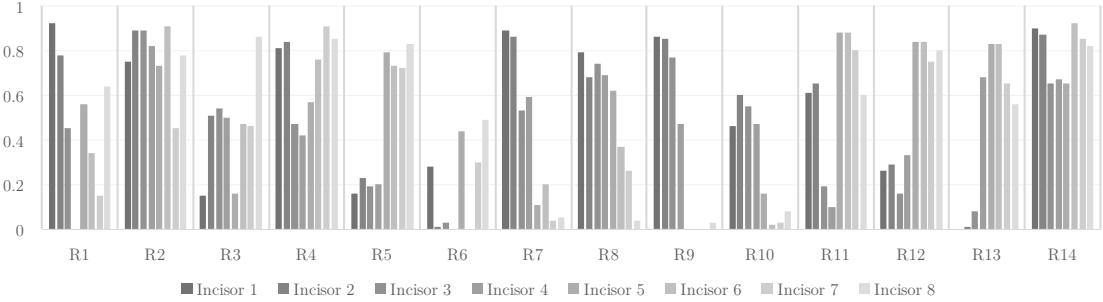


Figure 8: Precision of each fit.

Some incisors are fitted really well. 1/3 examples have an F-measure of 75 or above. These look like a good fit visually. Others have a rather poor fit. The most important cause is a wrong initial estimate of the position. This estimate has to be quite accurate, otherwise the iterative procedure will diverge to one of the neighboring teeth. A second limitation is the rather small training set. Some teeth have a quite unique shape. If such a teeth is left out of the training shape, the model will not recognize it as a regular shape and by consequence will not fit well. Figure 9 shows the radiographs on which the algorithm respectively performed best and worst. Figure 10 shows the result for the same radiographs when the initial fit is provided manually. This manual initialization is done through a GUI that allows to drag the mean shape to the right position on a radiograph. The GUI does not allow to change the rotation or the scale of the shape, resulting in still a sub-optimal initialization. Nevertheless, as the figure shows, the results are a lot better.

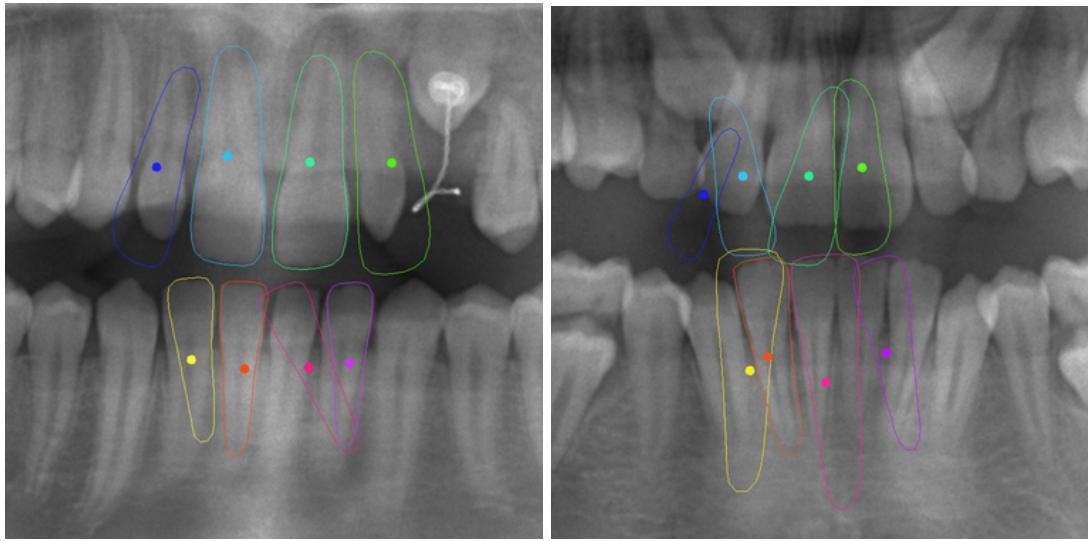


Figure 9: The best (radiograph02) and worst (radiograph06) fitting result.

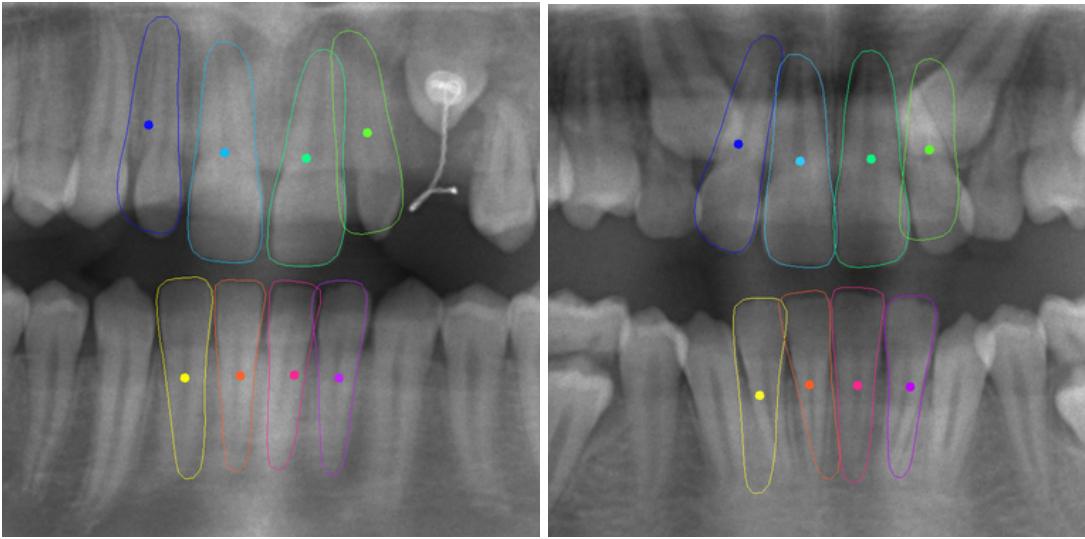


Figure 10: The fitting results for `radiograph02` and `radiograph06` after manual initialization.

References

- [1] S. A. Ahmad, M. N. Taib, N. E. A. Khalid, and H. Taib. An Analysis of Image Enhancement Techniques for Dental X-ray Image Interpretation. *International Journal of Machine Learning and Computing*, pages 292–297, Jan. 2012.
- [2] S. A. Ahmad, M. N. Taib, N. E. A. Khalid, H. Taib, and N. M. Ramli. The performance of contrast enhancement based on sharp filter for digital intra-oral dental radiograph images. In *Proceedings of the 15th WSEAS International Conference on Computers*, pages 344–349, Stevens Point, Wisconsin, USA, 2011. World Scientific and Engineering Academy and Society (WSEAS).
- [3] S. A. Bt Ahmad, M. N. Taib, N. E. A. Khalid, and H. Taib. Utilizing contrast enhancement algorithms (CEAs) in identification of dental abnormalities. In *2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAAECE)*, pages 218–223. IEEE, May 2013.
- [4] T. Cootes. An introduction to Active Shape Models. In *Image Processing and Analysis*, pages 223–248. 2000.
- [5] R.-Y. Huang, L.-R. Dung, C.-F. Chu, and Y.-Y. Wu. Noise removal and contrast enhancement for x-ray images. *Journal of Biomedical Engineering and Medical Imaging*, 3(1):56, 2016.