

# **Computer Vision Report**

## **Incisor Segmentation**

Orsolya Lukacs-Kisbandi, r0650042  
Stefano Proto, r0650809

June 2, 2017

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Active Shape Model</b>	<b>2</b>
2.1	Mean shape . . . . .	2
2.2	Principle Component Analysis . . . . .	2
<b>3</b>	<b>Image Processing</b>	<b>3</b>
3.1	Pre-processing . . . . .	3
3.2	Detecting edges . . . . .	3
<b>4</b>	<b>Fit model on an image</b>	<b>4</b>
4.1	Hand draw boxes . . . . .	4
4.2	Fine-tune boxes . . . . .	4
4.3	Fitting the active shape models . . . . .	4
<b>5</b>	<b>Evaluation and final considerations</b>	<b>5</b>

# 1 Introduction

The aim of this project was to write an algorithm that is capable of segmenting the upper and lower incisors in panoramic radiographs using a model based approach. The chosen method is called Active Shape Models, and it is described in [1] Cootes et al. The implemented algorithm can be broken down into sub-parts, namely: building the Active Shape Model, described in section 2, pre-processing the input images, explained in section 3, and finally fitting the model on the images, as presented in section 4. In addition, our observations and experimental results will be discussed in the final part of the report.

## 2 Active Shape Model

The reason we are building Active Shape Models for every incisor is to create a general representation of each incisor. By obtaining these models from combining different representations of the same tooth extracted from different peoples' radiographs, we can build a global model which can be fit on the same incisor for new images. This is a good approach for such problems because it creates a general representation of the object in interest, becoming more robust against variability.

Active shape models are based on a prior model of what is expected to be seen on the new image. In our case these are the incisors, represented by a set of characteristic points, the landmarks. Based on these landmarks, a general representation of the expected shape is created, which can be used to find the best match of the model in the new image. This model gives a compact representation of the previous knowledge allowing some variation, but being specific enough to exclude changes which haven't been seen in the training data. The landmarks are specific coordinates on the training images which represent a distinguishable point which appears on every training example.

### 2.1 Mean shape

The training set contains samples from several people. For each person in the training data, every incisor is represented by a set of 40 landmark points describing the shape of the tooth. In order to create the models, the shape of each tooth had to be scaled and translated to the origin, to make the computations easier. Translating the images also included rotating them in a way that their  $x$  and  $y$  axes would be aligned. After the shapes were centered to the origin based on their center of gravity, the scaled mean shape for each tooth was computed. This was necessary because we wanted to build a scale and rotation invariant model which could handle any type of input. The following procedure was done iteratively until the mean shapes converged to a value which did not change after realigning the landmarks to the new mean as proposed in [1]. After obtaining the mean shape, the landmarks were scaled and centered again around the new mean shape. This technique, is called the Generalized Procrustes Analysis. The main idea is that the alignment and re-centering of the shapes is done iteratively until the model converges to a value which does not change after adding another iteration. Our models were considered stable when the squared sum of the distance of each shape to the mean was less than  $10^{-10}$ .

### 2.2 Principle Component Analysis

Principal Component Analysis is a statistical procedure used for identifying a smaller number of uncorrelated variables, called "principal components" from a large set of data to explain the maximum amount of variance with the fewest number of principal components. In other words, the aim of PCA is to reduce the dimensionality of the data. This is an improvement both in terms of memory and computational power.

In our case, we set the number of principal components to 8, which covers 99% of the variance of the data. This way, we reduced the number of landmarks with a significant amount.

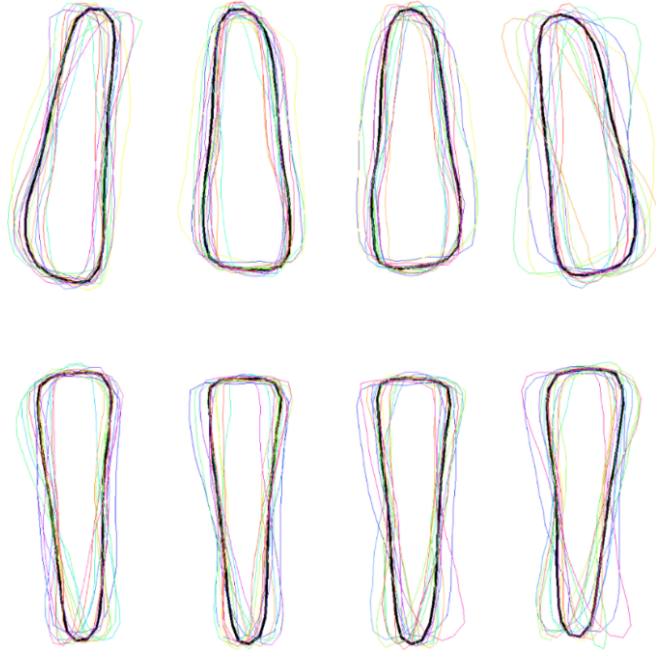


Figure 1: Final results of active shape model(black) form landmarks (other colors)

### 3 Image Processing

#### 3.1 Pre-processing

Radiographs are usually noisy and low in contrast. These characteristics make the image features difficult to extract. In order to overcome this issue, sharpening techniques and contrast enhancement algorithms are applied to the images before they are analyzed.

In order to reduce noise, we experimented with different filters, each having its own advantages and disadvantages. A balance between these filters had to be found, since some of them removed important image features such as edges and high contrast points. For reducing the impulsive and the Gaussian noise without affecting the edges, a median filter and a bilateral filter is applied.

The median filter considers each pixel in the image in turn and looks at the nearby pixels to decide whether or not it is representative of its surroundings. It replaces the pixel value with the median of the neighbouring pixel values. The median is calculated by first sorting all the pixel values from the surrounding neighbourhood into numerical order and then replacing the pixel being considered with the middle pixel value. Instead of the simple median filter, we tried to apply an adaptive median filter but this required too much computational power as well as time, so we decided to use the much simpler version of this.

The Bilateral filter is a non-linear, edge-preserving and noise-reducing smoothing filter. The intensity value at each pixel in an image is replaced by a weighted average of intensity values from nearby pixels, which is based on a Gaussian distribution. Its formulation is simple: each pixel is replaced by an average of its neighbours. It depends only on two parameters that indicate the size and contrast of the features to preserve [5].

After these transformations the contrast of the image is enhanced by means of two gray-level morphology techniques (top-hat and bottom-hat transform) and a histogram equalization technique. The top-hat transform is able to enhance brighter structures, while the bottom-hat transform enhances the darker structures. The obtained picture from top hat transform is added to, from the bottom hat transform is subtracted from the initial radiograph to apply the effect of this transformations.

The last step to enhance the contrast of the image is histogram equalization. For this purpose many approaches have been tried: among all the candidate algorithms, in Ahmad et al. it is shown that the

Contrast Limited Adaptive Histogram Equalization (CLAHE) better enhances the image quality and increases the diagnostic ease, overperforming both the Sharp Contrast Limited Adaptive Histogram Equalization (SCLAHE) and the Median Adaptive Histogram Equalization (SMAHE) [4]. CLAHE is based on the Adaptive Histogram Equalization (AHE). It divides the image in small blocks, *tiles*, and then equalizes the histograms. This allows the algorithm to adapt to local variations in an image such as shadows and highlights which would not be handled by global histogram equalization. AHE could amplify the noise, so the Contrast Limited technique is used: if any histogram bin is above a specified contrast limit, those pixels are clipped and distributed uniformly to other bins before applying histogram equalization.

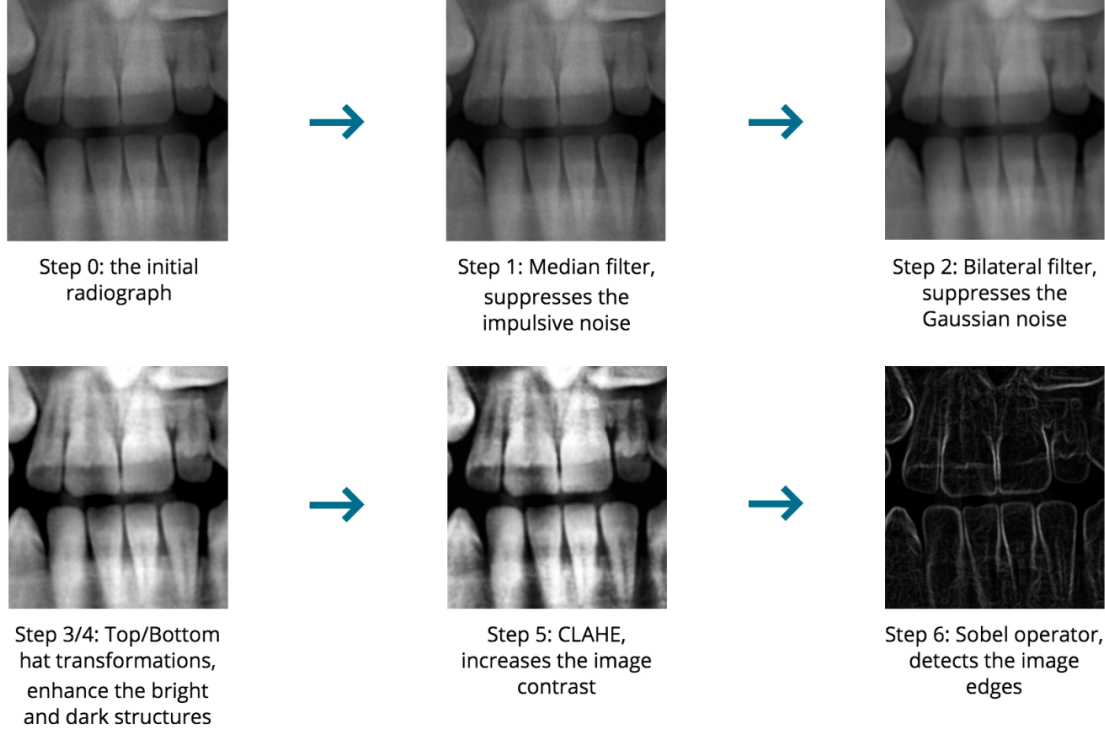


Figure 2: Preprocessing stages

### 3.2 Detecting edges

In order to find the edges over the images, the Sobel operator is used. The Sobel operator performs a 2-D spatial gradient measurement on an image and emphasizes regions of high spatial frequency that correspond to edges. Typically it is used to find the approximate absolute gradient magnitude at each point in an input grayscale image.  $G_x$  is representing the horizontal changes  $G_y$  the vertical ones on the image.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * I \quad G_y = \begin{bmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix} * I$$

At each point of the image we calculate an approximation of the gradient in that point by combining both results above with the following formula.

$$G = \sqrt{G_x^2 + G_y^2} \quad (1)$$

## 4 Fit model on an image

Fitting the models on the image consists of multiple stages. At first hand drawn boxes have to be created, than they are used in *best\_box* function to fine tune them. After returning these boxes the active shape models are fitted on the image.

### 4.1 Hand draw boxes

In the training part, the user is asked to draw boxes around the upper and lower incisors for every sample image in the training set.

The obtained boxes are stored in a file which are going to be used as an initial value for the automatic estimate. For testing, we select the coordinates which are the most extreme, the combination of these representing the largest box, in which all upper or lower incisors will fit in. These two boxes (one for the upper and one for the lower incisors) are used as initial estimates on the location of incisors when we are applying our model on an unseen radiograph.

### 4.2 Fine-tune boxes

For fine-tuning the boxes the function *estimate* was used. This function takes as parameter a preprocessed radiograph we want to fit the box on, a boolean variable weather we are looking for an upper or lower incisor, the preprocessed radiograph, the "biggest box" coordinates, which represent the biggest image segment selected when the boxes are hand-drawn. The next parameter is all the coordinates from the hand drawn boxes for the particular region and finally *b\_model*, whether we want to build the model or just load the pre trained model. After getting the largest box, the radiographs are cut, resized and flattened (create a  $N \times 1$  array from the image segment) based on these coordinates. We apply PCA on the cut images and take the top six components, which capture most of the variance in the data. After these steps, we call the function *best\_seg* which is taking as parameters the mean returned from PCA, the eigenvectors, the bounding box coordinates, the preprocessed radiograph, and the width and height. In this function, sliding windows of different scales are shifted over the region of interest on the radiograph and the error of reconstructing the segment using the eigenvectors and mean acquired from the cut shapes is calculated. The segment with the smallest error is returned. This box is returned to the main.

### 4.3 Fitting the active shape models

To fit the models, first the upper and lower incisor boxes which now are only containing the four incisors are divided in four equal parts, assuming that the teeth in the same jaw substantially have the same width. Following this, each tooth model is shaped in the corresponding box: the models are centred in the respective tooth box, scaled to fit in it, and then centered again.

After this initial positioning and scaling of the models, they are parsed in order to look for a better estimation of the tooth in their neighbourhood. For every point in the model, the normal line is computed. In case another point is a better fit than the current point, we update the model. The criteria for a point to be better than another is to have a higher value, since edges are represented by light regions. For this reason brightness along the line is compared with the actual value of the model in point crossing that line.

In order to have a greater resistance to noise not only the points on the line are considered, but for each point all its neighbours are evaluated. An average of the region is computed, and then compared with the current brightest point for the line: if the value of the region is greater than a certain threshold (directly depending on the current brightest point), then the centre of the evaluated region will become part of the model.

The parameters that allow flexibility in the model are the length of the norm we calculate, this can decide on the region we are looking for points which are more likely to be edges, the step we traverse the shape of the incisor, larger step is more resistant to noise, while smaller steps are following the shape more precisely.

Due to this fitting, the shape of the model is affected, and it has to be smoothed. After the final model is found, the *smooth\_model* function is invoked. This function smooths the model, computing an average of the shape over three consecutive points, reducing the effect of outliers in the model and creating a better representation of the incisor.

## 5 Evaluation and final considerations

In order to evaluate the obtained result, the Sum of Squared Errors is computed. This value represents how much the obtained model differs from the original landmarks, and it is computed for every sample of which we have the corresponding landmarks.

Based on these, we can observe that the accuracy of the hand drawn boxes in the training has a significant affect on the outcome of the estimation. The estimated position of the incisors in the radiographs is also affected by this. In addition, unusual teeth positions might affect the model as well, because the built model was not trained with such examples.

## References

- [1] Cootes, Tim, E. Baldock, and J. Graham. "An introduction to active shape models." *Image processing and analysis* (2000): 223-248.
- [2] Cootes, Timothy F., et al. "Active shape models-their training and application." *Computer vision and image understanding* 61.1 (1995): 38-59.
- [3] Blanz, Volker, et al. "A statistical method for robust 3D surface reconstruction from sparse data." *3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium on*. IEEE, 2004.
- [4] Ahmad, Siti Arpah Bt, et al. "Utilizing contrast enhancement algorithms (CEAs) in identification of dental abnormalities." *Technological Advances in Electrical, Electronics and Computer Engineering (TAECE), 2013 International Conference on*. IEEE, 2013.
- [5] Paris, S., Kornprobst, P., Tumblin, J. and Durand, F., 2007, August. *A gentle introduction to bilateral filtering and its applications.*, In ACM SIGGRAPH 2007 courses (p. 1). ACM.
- [6] Model reconstruction