Computer Vision [H02K5a]
Final Project: Incisor Segmentation

Ramaravind Kommiya Mothilal (r0605648)
Master of Artificial Intelligence (ECS option)
2016-17

TABLE OF CONTENTS

0. Introduction:

The objective of this project is to develop a model-based segmentation method, capable of segmenting the upper and lower incisors in panoramic radiographs. The project can be viewed as consisting of four tasks which are covered from section 1 to 4 in order. Section 1 explains the construction of an Active Shape Model. Section 2 discusses the procedures followed to pre-process the dental radiographs prior to fitting the model. Section 3, first explains how an initial estimate of the model, for the target incisor in the test image, is developed and then explains how the initial estimate is iteratively improved to fit the incisor in the test image. Section 4 discusses the achieved results and how the algorithm is evaluated using leave-one-out analysis. Finally, section 5 briefly discusses some of the possible extensions for this project.

1. Active Shape Model:

Active shape model (ASM), as the name suggests, is a model-based approach which makes use of a prior model of what is to be expected in the image, and typically attempt to find the best match of the model to the data in a new image. More specifically, the model can deform allowing considerable variability that are specific to the class of structures they represent, making it suitable for image structures whose shapes can vary, like our objective in hand where the incisors can differ in time and have variations from one person to another.

It “learns” specific patterns of variability from a training set of image structures represented by a set of points (called landmark points). By examining the statistics of the positions of the labeled points, a Point Distribution Model (PDM) is derived which gives the average positions of the points and the parameters which control their main modes of variation as found in the training set [1]. In this project, an ASM is constructed for each of the incisors separately instead of constructing a model for multiple teeth, that is, a model for four upper incisors or for four lower incisors etc.

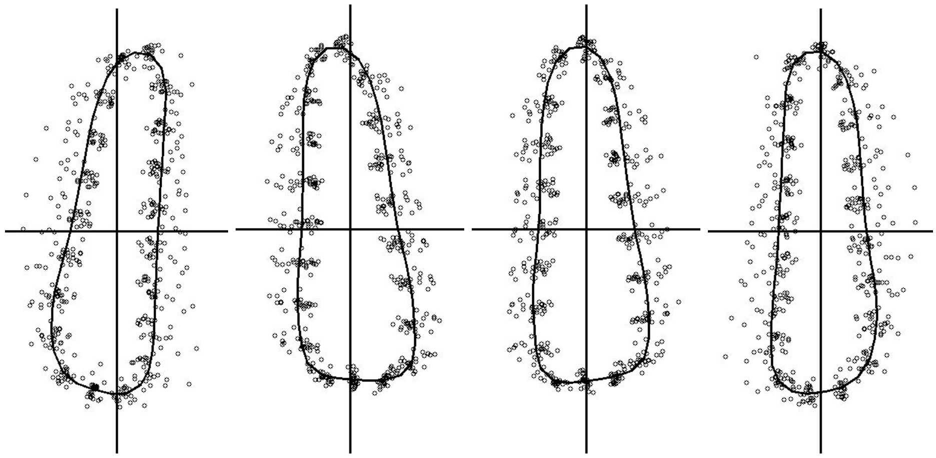
This section is divided into two subsections: Section 1.1 describes the pre-processing of landmark points using Generalised Procrustes Analysis. Section 1.2 then explains how a Principal Component Analysis (PCA) is used to construct an ASM.

1.1 Preprocessing Landmarks:

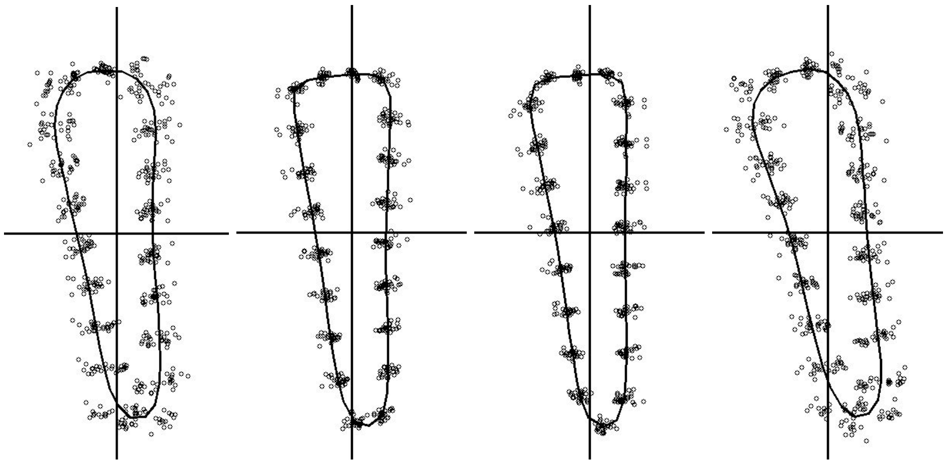
Each incisor in a training image is represented by a set of 40 landmark points. Also, the symmetry in the dental radiographs is exploited to double the number of training examples by considering the mirrored version of the matching incisor.

Prior to performing any statistical analysis on the landmark points, the equivalent points from different shapes must be aligned with respect to a set of axes. This is achieved as suggested in Protocol 4 of [2] using Generalised Procrustes Analyses by translating, scaling and rotating the training shapes such that the sum of squares of distances of each shape to the mean $D = |\Sigma x_i - \overline{x}|^2$ is minimised.

The training set used for explaining in whole of section 1 consists of radiographs numbered from 01 to 14 excluding 02 (test image). Figure 1 shows the variability of some landmark points of the aligned shapes of each incisor with respect to its mean shape. It can be observed that, mostly the points along the crown and the root show little variability while those along the sides form more diffuse "clouds". For the given training set, the top incisors show large variability compared to the lower ones. Also, almost all of the landmark points of incisor 6 and 7 show very less variability.



(a) Incisor 1 (b) Incisor 2 (c) Incisor 3 (d) Incisor 4



(e) Incisor 5 (f) Incisor 6 (g) Incisor 7 (h) Incisor 8

Figure 1: Scatter of some landmark points from the aligned set of each of the incisors, with their respective mean shape overlaid.

1.2 Principal Component Analysis:

After alignment, each example incisor in the training set can be represented by a single point in a $2n$ dimensional space where n is the number of landmark points. Clearly the landmark points are partially correlated as they do not move about independently. Thus the principal axes of a $2n-D$ cloud of data is derived by applying principal component analysis to the data. Each resulting axis gives a "mode of variation", a way in which the landmark points tend to move together as the shape varies. Thus the position of any of the points in a higher dimensional space is approximated using a small number of parameters The procedure described in Protocol 5 of [2] is followed for the implementation of PCA.

The PCA analysis allows us to visualise the variation in the shapes by varying the linearly independent parameters b_k within suitable limits such that the generated shape will be similar to those in training set. The most significant modes of variation in the data is described by the eigenvectors of the covariance matrix corresponding to the largest eigenvalues. It can be observed from Figure 2 that first four principal components explain about 96% of the variation.

Figure 3 shows the plot of b_1 against b_2 for the training set. The lack of structure in the scatter plot suggests that the parameters can be treated as independent. Figure 4 shows the two largest principal components for incisor 3, clearly showing that the incisors vary mostly in width and hardly in length. This also holds in tandem with the results visualised in Figure 1. It can be interpreted that the first PC explains the width of the incisor while the second component explains its skewness. While these interpretations hold for every incisor models, the other PCs explain minor shape variations which are characteristic of each incisors. Thus an Active Shape Model for an incisor, so developed, consists of its mean shape and has number of parameters which control the main modes of variation in ways found in the training set.

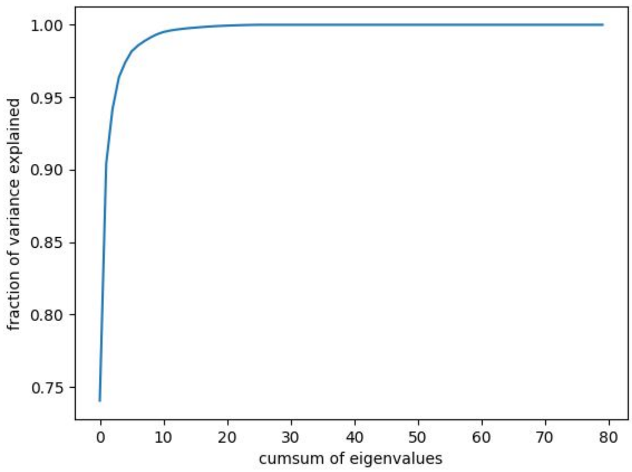


Figure 2: Variance explained by eigenvectors

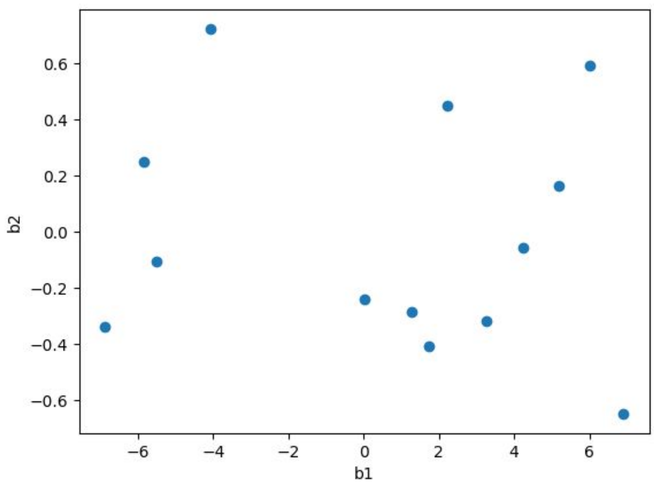
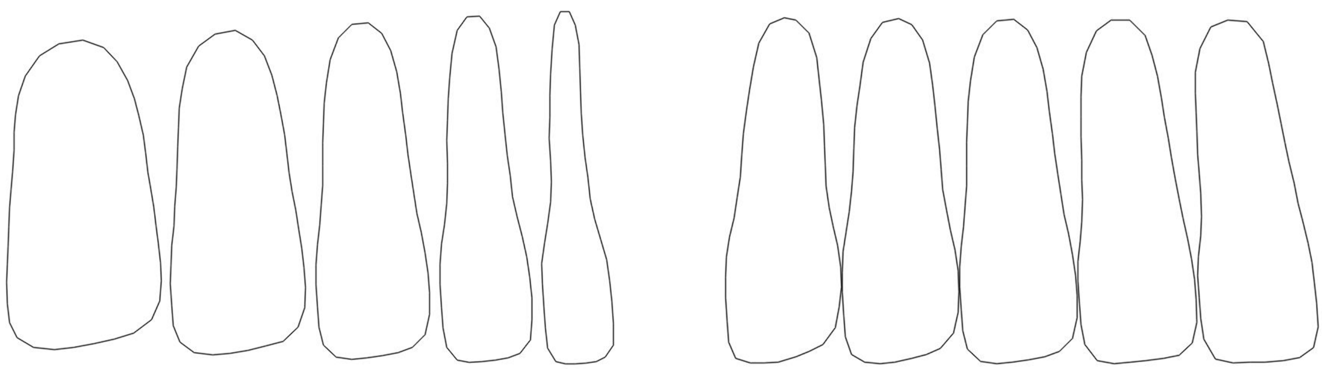


Figure 3: Plot of b_1 against b_2



$-2\sqrt{\lambda_1}$ <---------------- b_1 ------------------> $+2\sqrt{\lambda_1}$ $-2\sqrt{\lambda_2}$ <---------------- b_2 ------------------> $+2\sqrt{\lambda_2}$

Figure 4: Effect of varying First and Second Principal Component between -2 and 2 standard deviations of third incisor's model

2. Pre-processing of Dental Radiographs:

The dental radiographs are low in contrast and quite noisy. To segment the incisors in a new image, if we ensure that the region of the radiograph containing teeth is high in contrast relative to the rest of the image and is also less noisy, then the task at hand becomes easy. There are several methods found in the literature to do this. A technique for noise suppression and contrast enhancement suitable for X-ray images as suggested in [3] is used for preprocessing the dental radiographs in our case. Figure 5 shows the complete overflow of the technique followed. The noise in the image is first removed before applying any contrast enhancement techniques to ensure that the noises in the image are not amplified avoiding any possible distortion of the image.

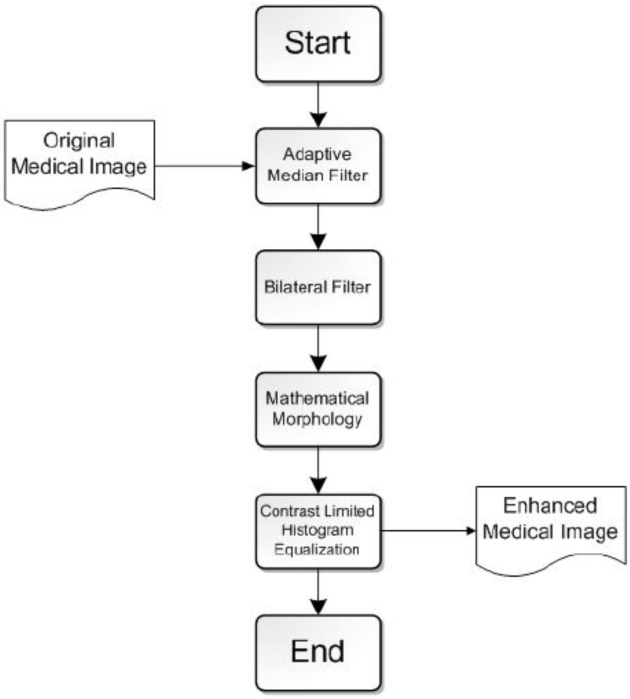


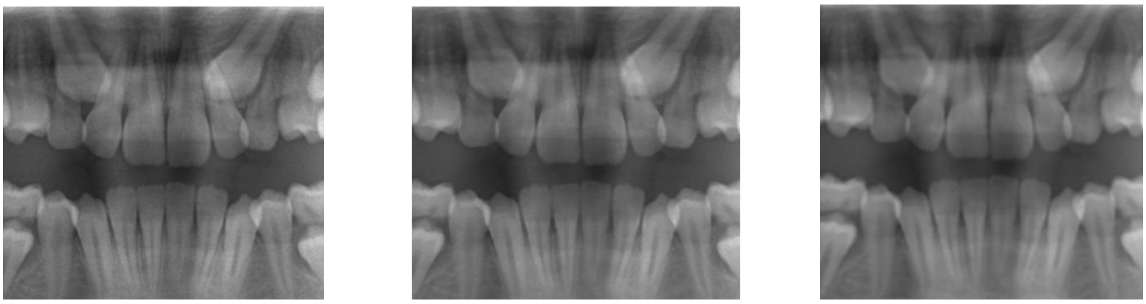
Figure 5: The overflow of the preprocessing technique.

Thus, an adaptive median filter is first applied to the original image, keeping balance between the noise level and image content, to remove any impulsive noise. Further, to handle gaussian noise while preserving the edges in the image (since it is necessary to correctly segment an incisor), a bilateral filter is applied which takes pixel similarity into consideration besides the spatial similarity.

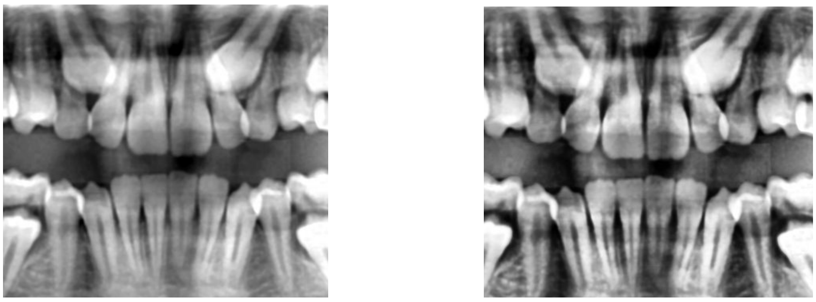
Next, in morphological processing, the top-hat transform and bottom-hat transform is applied separately to extract brighter structures (peaks and ridges) and darker structures (valleys and troughs) respectively. The resultant transformed images are added to the original image to create an enhanced image where the brighter structures are made more brighter and darker structures even darker.

To further enhance the image contrast, Contrast Limited Adaptive Histogram Equalization (CLAHE), an improvement of Adaptive Histogram Equalization, is applied. AHE performs histogram equalisation for multiple local windows across the image but it might amplify noise in homogeneous region due to the large slope of mapping function. But CLAHE restricts the slope of the mapping function which is able to reduce the undesired amplification. Since there are many homogeneous regions in medical images, CLAHE is suitable [3]. Also, it has been shown that CLAHE outperforms different AHE based techniques for statistical research making it suitable for this project [4].

The results of the preprocessing stage for radiograph 06 is shown in Figure 6. It can be observed that the contrast of the radiograph has been improved in a useful sense making the boundaries of the teeth clearly distinguishable.



(a) original image (b) after adaptive median filter (c) after bilateral filter



(d) after mathematical morphology (e) final result after applying CLAHE

Figure 6: The different steps of the radiograph preprocessing

3. Model Fitting:

Prior to fitting a model of an incisor to an image, a proper initial estimate for the incisor model in the new image has to be developed which is explained in section 3.1. Then the initial estimate is iteratively improved to better match the actual incisor in the image as explained in section 3.2.

3.1 Initial Estimate:

The task of finding an initial estimate of a model for an incisor in a new image is twofold: (1) an appropriate search region in the new image, which possibly contains our incisor of interest, is constructed; (2) an appearance model is developed to evaluate a possible match found inside the search region. The procedure followed is discussed in detail below.

Finding Initial Estimate

↙ ↘

Appearance Model   Search Region <---- Jaw split

Appearance model:

The bounding boxes containing the four upper incisors from all the training images are extracted and rescaled to their mean shape for which an appearance model is built. Similar procedure is followed to build an appearance model for the lower incisors. The appearance model is built collectively for four upper or lower incisors instead of for each of the incisors. This is mainly because the appearance model of individual incisors are not clearly distinguishable from one another.

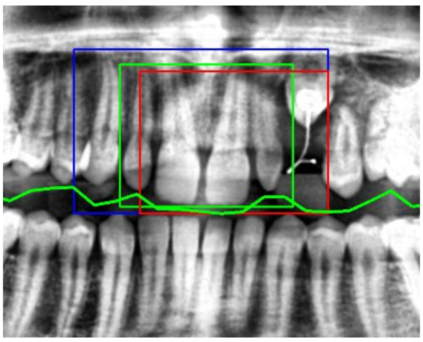
Search region:

The extreme points on the bounding boxes along the width are quite symmetrical about the middle vertical axis of the image. This helps in fixing the width of the search region in the new image. But the extreme points along the height vary significantly from image to image. Thus the jaw split is located in the new image and used as a reference to fix the height of the search region.

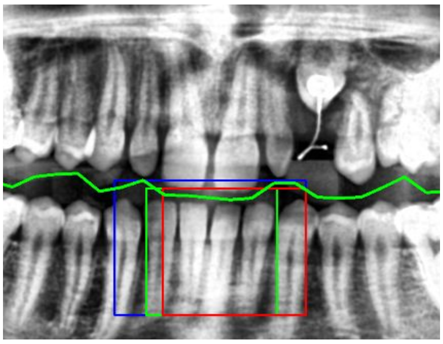
Jaw split:

To locate the jaw split, a gaussian filter is first applied along the width to highlight the middle portion of the image containing all the teeth and underemphasize the rest of the image. Next, intensity histograms are calculated at fixed intervals over the width of the top-hat transformed image. Then for each interval along the width, points representing three local minima in the histogram are stored. The points are then connected to form paths such that the intensities along the edges connecting the points are minimal. The path with lowest total intensity corresponds to the jaw split.

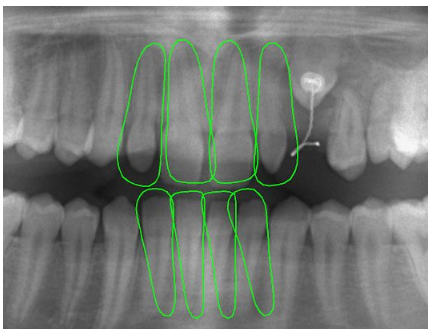
Having constructed the search region and the appearance model for four upper and lower incisors, a moving window of different scale (with respect to the mean shape of the bounding boxes found from the training images) is slided inside the search region. Each window is projected onto the first 5 principal components of the appearance model and the window with lowest reconstruction error is fixed as the initial estimate for the four incisors. Assuming that each incisor has more or less the same width, the window with lowest reconstruction error is then split into four equal parts resulting in a bounding box around each of the incisors. Finally, the mean shape of an incisor is translated and scaled to fit within the estimated bounding box.



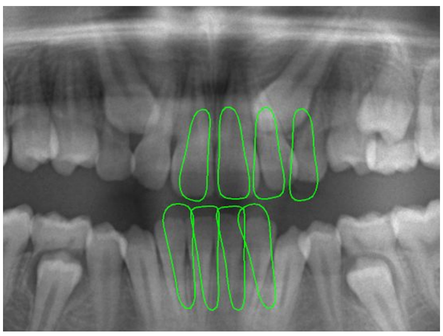
(a) Finding initial estimate for upper incisors



(b) Finding initial estimate for lower incisors



(a) Initial estimate for radiograph 02



(b) Initial estimate for radiograph 06

Figure 7: Finding Initial Estimate. (a),(b) - Process (Blue-search region; Green-Best match so far; Red-Moving window) and (c),(d) - Good and Bad results respectively.

3.2 Iterative improvement of initial estimate:

Given an initial estimate of the model for an incisor, the next task is to iteratively improve the model parameters to best describe the target incisor in the image. To solve this task, a suitable 'Fit function' has to be developed such that the best set of model parameters to interpret the incisor in the image is then the set which optimises this function. The iterative fit method, as explained in Protocol 2 of [2], is used for this project as discussed in detail below.

Grey Level Model:

First, the region of the image around the current model points is examined to find the best nearby points for each of them. To do this, a method described in [2] is followed. For every current model point in the image, the gradient of k pixels on either side in each training image is sampled and normalised to build a grey-level model. During search, the same kind of profile of m pixels ($m > k$) is sampled on either side of each model point of the test image and the quality of fit (using Mahalanobis distance) is tested against the model profile (Figure 8). The point giving the least cost is chosen as the next best pixel. For this project, $k = 10$ and $m = 15$ are used. Also, a median filter of length 5 is applied to the set of newly found points to smooth the boundary of the shape so as to make the subsequent step easier.

Update of model parameters:

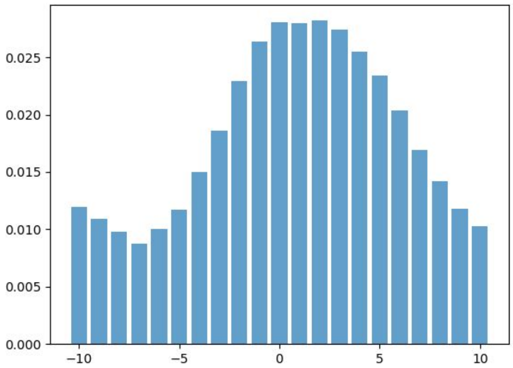
After getting a new position in the image for every model point, the current pose and shape parameters of the model is updated to best match the model to the new points. The protocol 1 of [2] is used for matching current model points to the target points. This uses Procrustes Analysis to find the best translation, scale and rotation to best fit the newly found points. Also, to ensure allowable shape variations, the parameters b_i are constrained such that they vary no more than three standard deviations. Similarly, the scaling parameter is limited to a maximum of 20% change compared to the mean shape and the rotation parameter is limited to a maximum change of 45 degree.

Multi-Resolution Active Shape Models:

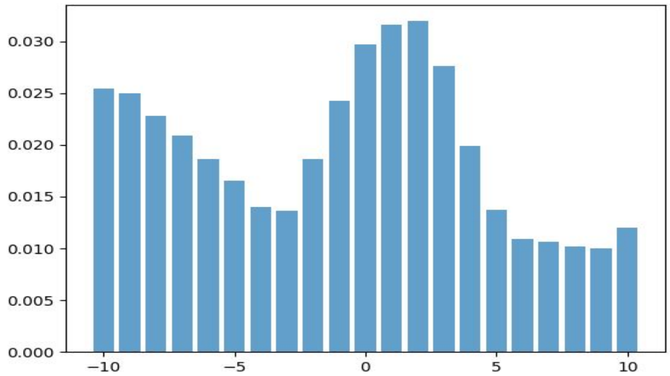
To improve the efficiency and robustness of the algorithm, it is implemented in a multi-resolution framework (Figure 8). For each training and test image, a gaussian image pyramid of two level is built. The base image (level 0) is the original image and the next image (level 1) is formed by smoothing the original then subsampling to obtain an image with half the number of pixels in each dimension [2]. Increasing the level further makes the incisor boundaries diffuse into each other making the fitting process difficult. Thus, this process allows the algorithm to first find a rough estimate of the model points in a coarse image and then refining them in a high resolution image.

Convergence Decision:

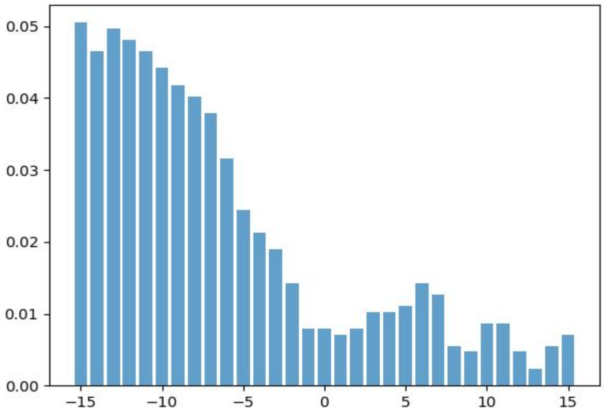
The algorithm is checked for convergence by recording the number of times that the best found pixels along a search profile is within the central 25% of the profile. In general, the crown part of the incisors is easily distinguishable compared to the root part. Hence the best found pixels of the crown part have much higher probability to lie within the central 25% of the profile than the root part. Therefore, if at least 50% of the points are found within the central 25% of the profile, the algorithm is declared to have converged at that resolution.



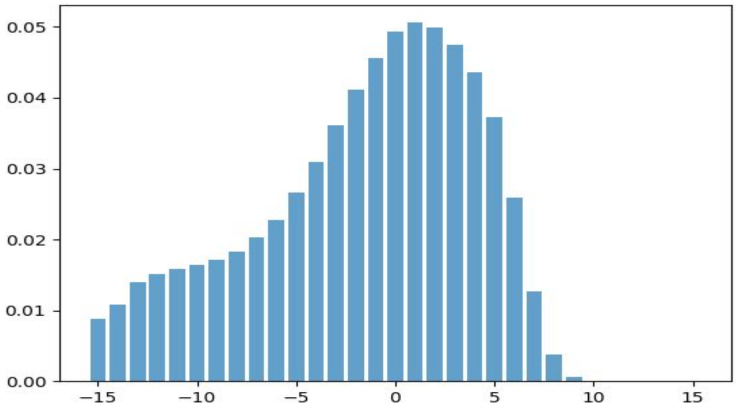
(a) Model profile - Level 0



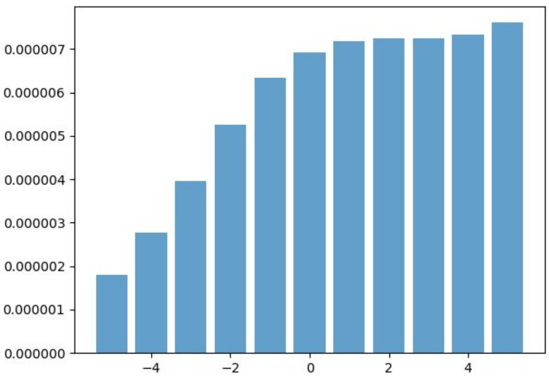
(b) Model profile - Level 1



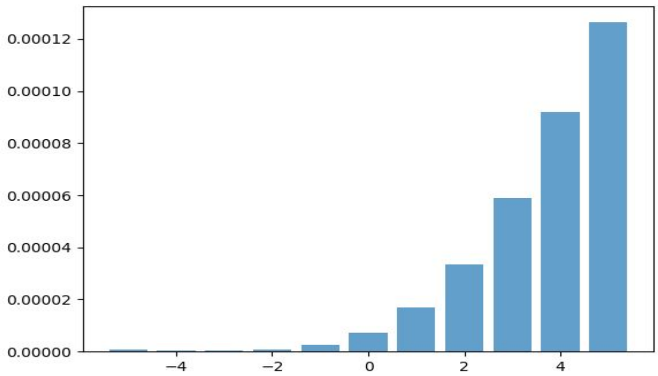
(c) Sample profile - Level 0



(d) Sample profile - Level 1



(e) Cost of fit - Level 0



(f) Cost of fit - Level 1

Figure 8: Grey Level Models of 10th model point of incisor 2

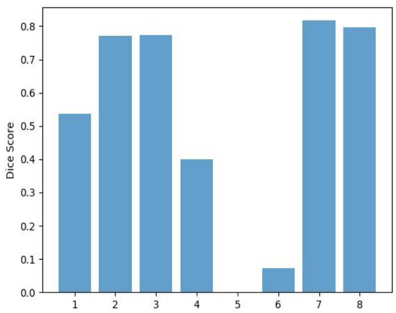
4. Results Evaluation:

After the model is fitted to the target incisor in the test image, the landmark points given by the model are checked for similarity with the ground truth landmarks of the test image. Dice coefficient, which measures the similarity between two objects, is used to quantify the performance of our algorithm. Let $'A'$ be the binary image with ground truth where the pixels of the incisor defined by the true landmark points of the test image is white and rest are black and, $'B'$ be the similar binary image defined using the final landmark points given by the model. Then, the Dice coefficient, which ranges from 0 (no similarity) to 1 (identical), is given by ,

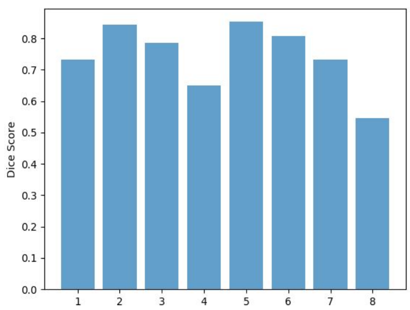
$$D = \frac{2|A \cup B|}{|A| + |B|} = \frac{2 * TP}{2 * TP + FP + FN}$$

Where TP, FP, FN are True positives, False positives and False Negatives respectively. The Dice score is not only a measure of the number of true positives found but also penalizes for false positives and false negatives. Thus it is more similar to precision than accuracy which is the appropriate evaluation metric for this project.

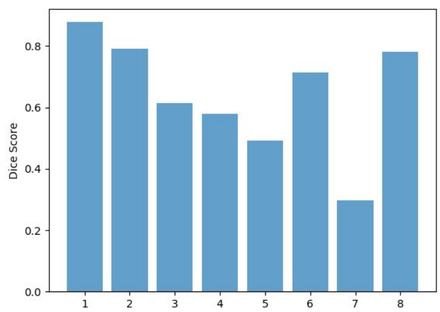
Given 14 radiograph images and their landmark points, the active shape model developed for each of the incisors is evaluated using Leave-One-Out cross validation method. Each one of the 14 provided images is used once as the testing image, while the others are used as training set.
Figure 9 shows the Dice score for the estimated fit for each radiograph image.



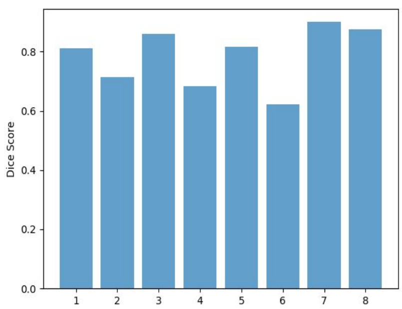
(a) Radiograph 01



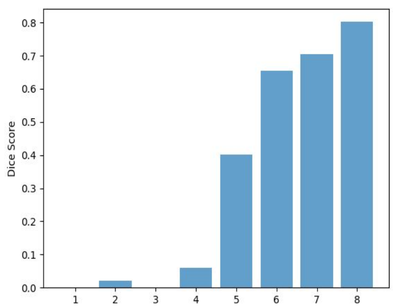
(b) Radiograph 02



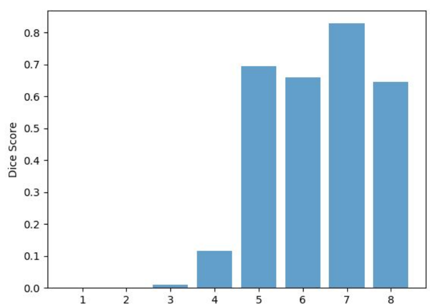
(c) Radiograph 03



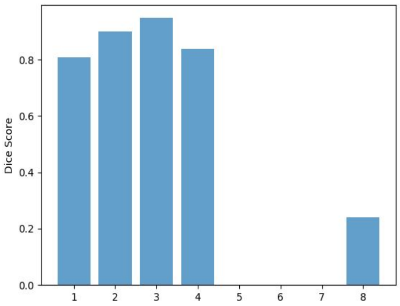
(d) Radiograph 04



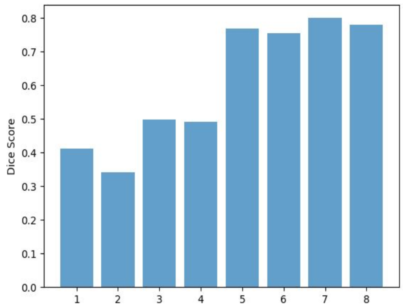
(e) Radiograph 05



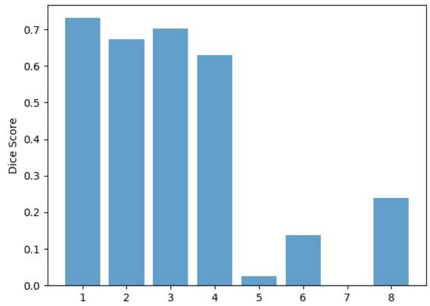
(f) Radiograph 06



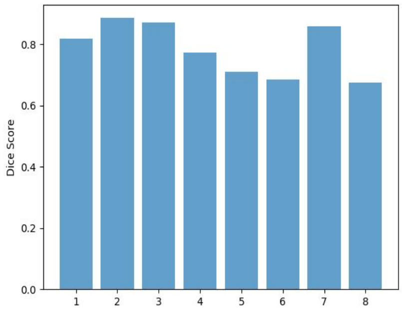
(g) Radiograph 07



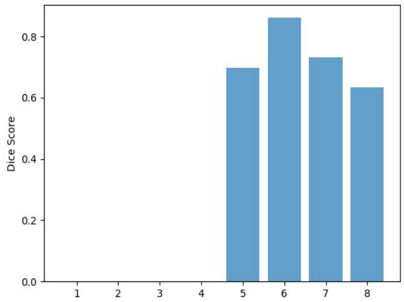
(h) Radiograph 08



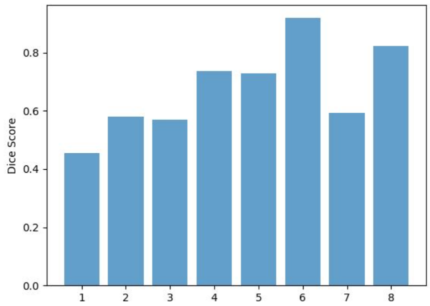
(i) Radiograph 09



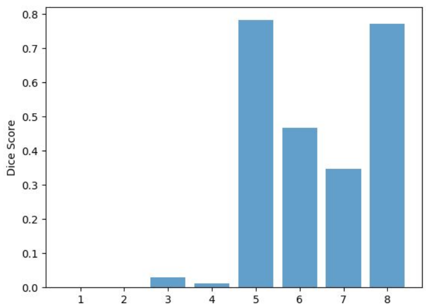
(j) Radiograph 10



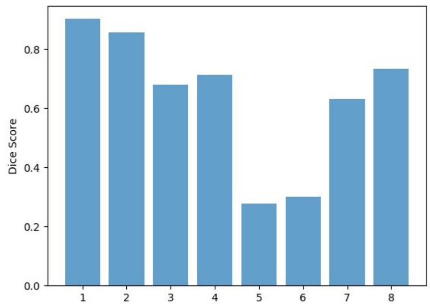
(k) Radiograph 11



(l) Radiograph 12



(m) Radiograph 13



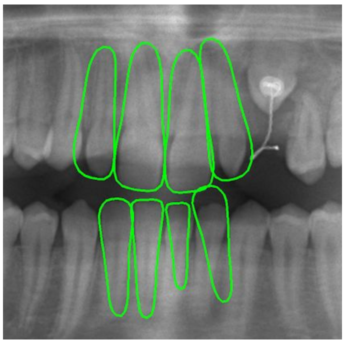
(n) Radiograph 14

Figure 9: Dice Coefficient for the estimated fit for each incisor of radiograph 01 to 14

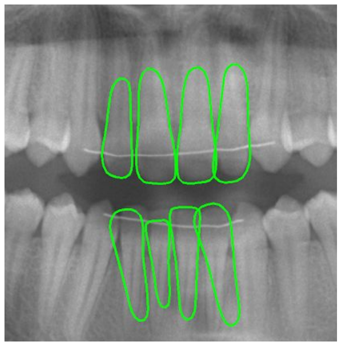
Active Shape Model is a model-based method that makes use of a prior model of what is to be expected in the image and then attempts to find the best match of the model to the data in a new image. But each of the given training radiographs are completely different as it can be easily observed that these medical images don't follow any fixed center or axes. In addition tuning the parameters for automatic initialisation is a time consuming task. Hence without an appropriate initial estimate of the model, finding the optimal model parameters seems like a difficult general optimisation problem. This is the main cause for poor fit in some of the examples (Figure 10). But still, some incisors are fitted really well in almost all the images and almost one-third of the examples have a Dice score of above 0.70.

Also, very less proportions of the incisors in the images are closer to their respective mean incisor model. This could be due to the small training set. Some incisors have quite unique shape. If such a teeth is left out of the training then the model will not perform well on the new image. It can be observed from Figure 10 that, in examples where the model performs poorly, the iterative procedure has made the model diverge to one of the neighbouring teeth. This is probably due to the bad initial estimate or due to the small training set as explained before.

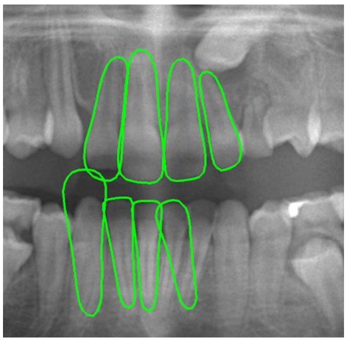
To justify the statements made above, when the initial fit is provided manually by dragging the mean shape of the incisor to the right position on the image, the results improved very significantly as shown in Figure 11.



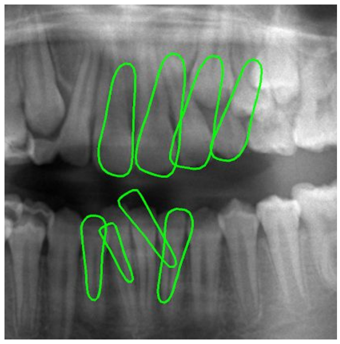
(a) Radiograph 02



(b) Radiograph 10

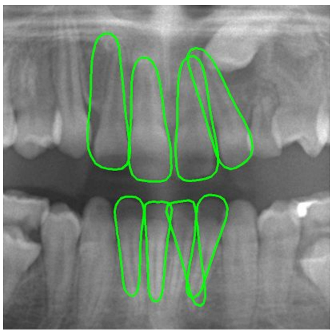


(c) Radiograph 07

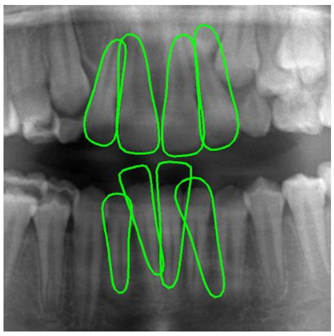


(d) Radiograph 13

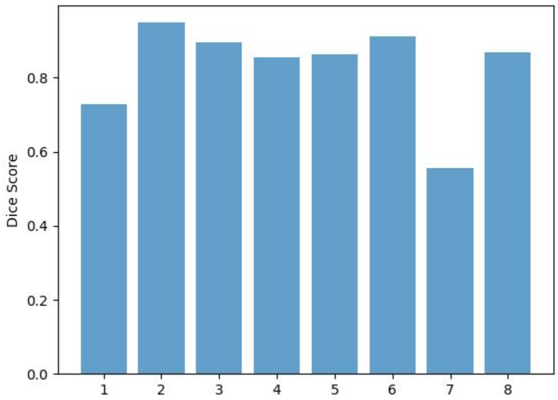
Figure 10: {(a),(b)} - some good results. {(c),(d)} - some bad results. Models diverged to neighbouring lower incisors in (c) and to neighbouring upper incisors in (d)



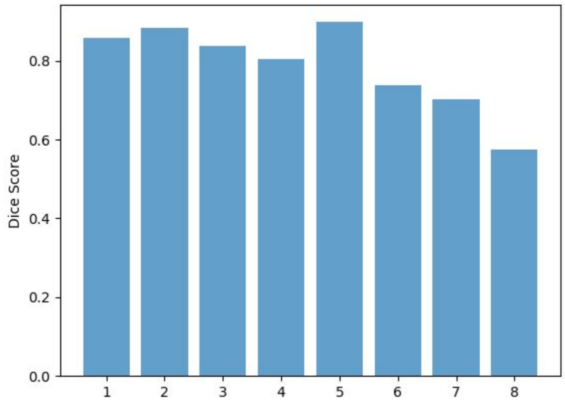
(a) Radiograph 07



(b) Radiograph 13



(c) Dice Coefficients - radiograph 07



(d) Dice Coefficients - radiograph 13

Figure 11: Fitting results after manual initialisation

5. Possible extensions:

The problem complexity could possibly be reduced by estimation of the initial incisor model using Convolutional Neural Networks.

The parameters of the grey level model, namely, k and m were kept fixed throughout the project by trial. But the optimal choice of these parameters could have been selected by cross-validation techniques.

Further, some of the training sets could have some non-linear dependencies between the shape parameter $b_i's$ in which case non-linear PCA using Artificial Neural Network or Support Vector Machines could be performed.

Also, some of the incisors are very similar resulting in divergence of the model of one incisor to the neighbouring incisor. Thus, one possible extension could be to first try fitting the ASM for all four upper (or lower) incisors and then iteratively half the number of incisors to fit and so on.

References:

[1] T. F. Cootes, C. J. Taylor, D.H. Cooper, J. Graham. Active Shape Models - Their Training and Application. In Computer Vision and Image Understanding, 61(1):98-59, Jan. 1995

[2] T. Cootes. An introduction to Active Shape Models. In Image Processing and Analysis, pages 223–248. 2000.

[3] 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):56-67, Feb. 2016.

[4] 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, 2(3):292–297, June 2012.