6.1. Introduction

Most methodologies not only differ in the kind of techniques used to pinpoint the location of landmarks but in many other ways like: the size of training and test dataset, quality and variations covered in the dataset based on age, race, gender, etc., different studies use distinctive landmarks and different methodologies to report the results.

In this Chapter, we propose a combination of local and global feature matching techniques to enhance the accuracy of the landmark localization, speed of computation and robustness to image noise. A three stage landmark strategy is used. In the first stage, the approximate landmark locations are identified using a reliable and global shape matching technique using Zernike moments and Procrustes analysis for shape alignment. Four expectation windows for each landmark are derived using these landmark positions. To bring robustness to noise only a subset of Zernike moments is used. It can be seen as a fast way of comparing two patterns in, which the finer details have been smoothed out. The Zernike moments with order five are used to find the 5% most similar images in the training set to the test image in our method. Many landmarks have clear definitions when they are searched in a small window. The RP and CP based methods can efficiently reduce the feature space and has the rotation invariance property. This leads to reduction of execution time significantly.

In the second stage, the system applies rotation invariant template matching using a combination of RPs and CPs on the four expectation windows to get candidate landmark locations very close to the exact location of the landmark. Once candidate landmark points are found, the final refinement of the landmark location is done in stage three of the algorithm with slower, but more robust traditional template matching using SSD and NCC. At this stage multiple, templates are used to achieve robust search of objects in presence of different scales, object deformations, illumination variation, etc. Four templates are used to cover the variations, and the
computational cost is not increased much as this method is used in a very constrained area around the landmark positions.

The main contributions of this work are summarized below:

i. The proposed algorithm uses Zernike features, very good shape descriptors for initial landmark approximation instead of using edge features. Most other methods are based on training using few reference points, found using image enhancement and edge detection. Fig. 6.1 shows the reference points used in [17, 26, and 27].

ii. With increased image enhancement, the subjective perception of X-ray image may improve, but may lead to loss of details. The same is true with edge detectors due to complexity of structure, low contrast and blur in X-rays it is difficult to get good edge detection results for all X-ray images. Moreover, each edge detector has certain parameters, which need to be adjusted.

iii. The rotation of landmark structures is arbitrary in cephalograms, and the templates rotated at fixed intervals as used in the most previous work [26, 27] does not yield good results. This paper uses rotation invariant template matching based on central and RPs.

iv. The proposed algorithm uses a general technique not specific to each landmark. It does not use any handcrafted algorithm as used in [8, 9, 17, 26, 27], nor any structuring element or control and feedback templates fixed for each landmark [22, 24, 30]. Thus, addition of new landmarks is easy.

Figure 6.1: (a) Four reference landmarks used in [26] (b) Five reference landmarks detected using Canny edge detector in [17] and (c) Three reference landmarks detected using Susan edge detector in [27].
6.2. Proposed Algorithm
The algorithm consists of an offline and an online step.

6.2.1. Offline module
Offline module helps reduce the computational costs for the next stages. Zernike moment features are calculated and stored along the hand annotated landmark positions for each training set image. Two orthodontists with 5+ years of clinical experience helped in getting manual landmark locations. No more than ten radiographs were traced in a single session to minimize error owing to examiners fatigue. The CPs and RPs corresponding to each template are pre-computed and stored. Lookup tables for radius and theta values are pre-computed and stored.

6.2.2. Online Module
Online module is initiated by providing a test X-ray as input. We propose a three stage framework.

Stage 1: Finding Expectation Window for each Landmark

Step 1: The system starts by extracting the Zernike moment features for the query image.

Step 2: Compare the Zernike moment magnitude of the test image $|Z_{T_{pq}}|$ with the training set images $|Z_{R_{pq}}|$ using Euclidian distance.

$$d_{s^2} = \sum_{(p,q) \in D} \sum (|Z_{T_{pq}}| - |Z_{R_{pq}}|)^2$$

(6.1)

5% most similar images to the test image are selected.

Step 3: Compute the mean shape for each landmark set by Procrustes shape alignment algorithm [101-103].

The mean shape for each landmark set is obtained by computing the shape that minimizes the sum of squared Procrustes distances to all the landmark sets obtained by considering 5% most similar images from the training data set.

Algorithm for Procrustes Analysis

i. Translate each shape so that it is centered on the origin.

ii. Select one shape to be the approximate mean shape.

iii. Align the shapes to the approximate mean shape.

   a. Translate the shape to the origin.
b. Scale the translated shape to unit size.

c. Rotate each shape to align with new approximate mean.

iv. Calculate the new approximate mean from the aligned shapes.

v. If the mean shape from two and three are different then return to step two else consider approximate mean shape as true mean shape and stop.

**Step 4:** Locate four overlapping expectation window of size $2 \times t_s$ ($t_s$ is the template size) for each landmark as shown in Fig. 6.3 (c) by shifting the mean position $(x_m, y_m)$ of the landmarks calculated in step 3 by $(x_m + 0.5 \times t_s, y_m)$, $(x_m - 0.5 \times t_s, y_m)$, $(x_m, y_m + 0.5 \times t_s)$, and $(x_m, y_m - 0.5 \times t_s)$ and finding the window around each of this shifted position. Four expectation windows instead of one large window help in reducing false detection due to the presence of similar structures nearby.

**Stage II: Computing probable landmark positions**

The landmark’s location is extracted using the four expectation windows, CPs and RPs and multiple templates to reject individual variations. Following procedure is used to obtain four different templates for each landmark:

i. Generate a set of templates $T = \{T_1, T_2, ..., T_n\}$ for each landmark, equal to the number of training images in the database.

ii. Compute the Zernike moments for each template.

iii. Save template corresponding to first element in $T$ as $Tm(1)$ and $Temp$.

iv. FOR $i = 2$ to $4$

   a. Compute a set of 20% most similar templates to $Temp$ by comparing their Zernike moment features.

   b. Remove this set from the set $T$.

   c. Select the first element from updated set $T$ and save in $Tm(i)$ and $Temp$.

v. END FOR

**Step 1:** For each landmark pre-compute $CP$ vector $PC^T(\phi_k)$ and $RP$ vector $PR^T(r)$ of the templates.

**Step 2:** Apply template matching for each landmark $\ell$ using $CP$ and $RP$ as given in the following algorithm
i. Calculate CP vectors $PC_i^S(\phi_k)$ and RP vectors $PR_i^S(r)$ for each sub-window obtained by shifting the template window row wise and column wise in each expectation window $w$.

ii. Compute
   a. Euclidian distance $D_1(i)$ between $PC^T(\phi_k)$ and $PC^S(\phi_k)$.
   b. Euclidian distance $D_2(i)$ between $PR^T(r)$ and $PR^S(r)$.

iii. Compute the composite distance measure ($D_c$) based on CPs and RPs as follows
   \[ D_c = \varphi D_1 + (1 - \varphi) D_2. \]

iv. Sort $D_c$ in ascending order.

v. Save three probable landmark locations in $(x_m^i, y_m^i)$ based on first three minimum distances in $D_c$ to avoid false detection due to similar structures present near the actual landmark position.

**Step 3:** Compute the minimum and maximum values of $x$ and $y$ obtained from last step and divide the distributions into groups with class interval of 10. Find the class with highest frequency for $x$ and $y$ respectively. The mid value of these classes is taken as the landmark position $(x_m, y_m)$.

**Stage III: Final Refinement of the Landmark Location**

Once candidate landmark points are found, the final refinement of the landmark location is done by slower pixel $\times$ pixel matching using more accurate metrics, which take into account human perception like SSD and NCC and multiple templates.

**Step 1:** An algorithm for pixel $\times$ pixel template matching to extract the exact landmark location using NCC and SSD is described below. Select a small sub-image around $(x_m^i, y_m^i)$ extracted in stage II.

   i. Consider a sub-window masked by template in the sub-image. Compute the values of NCC and SSD for the selected window around each landmark using the following algorithm

   ii. Save the location of maximum correlation $ncc_{(t,x,y)}$ and minimum distance $ssd_{(t,x,y)}$ for each value of $t$

**Step 2:** The final landmark location is a position at which the two locations $(\max ncc_{(t,x,y)}$ and $\min ssd_{(t,x,y)})$ overlaps else the location at which the value $\left[ 1 - (ncc_{(t,x,y)} - (1 - ssd_{(t,x,y)}) \right]$ is maximum.
Figure showing the stage wise results for three landmarks is shown in Fig. 6.3.

6.3. Materials and Methods used in this Study

The 18 landmarks considered in this study as shown in Fig. 6.2 and defined in Table 6.1.
The landmarks differ depending on their type (soft tissue landmarks or hard tissue landmarks lying on the bony structure). Another criterion to classify landmarks is their location (landmarks lying on contour turning points both convex and concave, on centre points of regions and between soft and hard tissue profile). The selected landmarks are of different types and locations and should provide a reasonable test set for assessing the landmarks [1, 2].

Total 85 training images were collected randomly from the data set without judgment of their quality, sex, and age. The system was tested using a drop-one-out algorithm. Each time 84 images were used for training, and one image was excluded for testing. To verify the performance two assessment methods were used. The first method is comparing the results obtained by the proposed algorithm with the manually traced landmark positions. This is the evaluation method most commonly used in all research papers on automatic cephalometric analysis. Two orthodontist with 5+ years, of clinical experience helped in getting manual landmark locations. No more than ten radiographs were traced in a single session to minimize error owing to
examiners fatigue. In the second assessment method to demonstrate the behavior of other promising techniques on our database Wei et al. method [17] was implemented and its results compared with the proposed algorithm.

If the difference between the two landmark positions compared is less than ±2 mm it is considered successful. Three performance metrics used are, success rate (landmark detected within a range of ±2 mm), consistency and precision. Consistency of each landmark is evaluated by considering its distance in pixels from the ground truth position and averaging over all the test images to obtain mean error for each landmark. Lower the mean error higher is the consistency. The standard deviation of these distances gives the precision of landmark identification. Lower SD (standard deviation) corresponds to higher precision.

6.4. Results and Discussions

The performance of the proposed algorithm’s success rate, consistency and precision is calculated for each derived landmark location with respect to the ground truth values obtained as the mean coordinates positions located by all orthodontist for each landmark.
Figure 6.3: Results for three landmarks (Sor, Cd and Prn) after applying different steps of the proposed algorithm on a test image (a) Stage I (Step 2) Landmarks positions corresponding to best matching five images (b) Stage I (Step 3) Landmark position after alignment (c) The four expectation windows around the mean (d) Blue star gives initial approximation and red star gives the position found using projection based template matching and (e) Blue star initial landmark position for stage three and yellow star gives the final position found using the combination of NCC and SSD.
Table 6.2: Comparison of success rate of landmark detection with previous methods (when reported).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pr</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>95</td>
</tr>
<tr>
<td>Sella</td>
<td>53</td>
<td>77</td>
<td>65</td>
<td>76</td>
<td>90</td>
</tr>
<tr>
<td>Na</td>
<td>83</td>
<td>91</td>
<td>85</td>
<td>86</td>
<td>100</td>
</tr>
<tr>
<td>Or</td>
<td>40</td>
<td>74</td>
<td>65</td>
<td>-</td>
<td>85</td>
</tr>
<tr>
<td>Pns</td>
<td>86</td>
<td>100</td>
<td>80</td>
<td>83</td>
<td>88</td>
</tr>
<tr>
<td>Ans</td>
<td>68</td>
<td>92</td>
<td>90</td>
<td>79</td>
<td>86</td>
</tr>
<tr>
<td>Go</td>
<td>61</td>
<td>87</td>
<td>85</td>
<td>86</td>
<td>100</td>
</tr>
<tr>
<td>Uie</td>
<td>79</td>
<td>81</td>
<td>-</td>
<td>90</td>
<td>95</td>
</tr>
<tr>
<td>Me</td>
<td>78</td>
<td>84</td>
<td>100</td>
<td>98</td>
<td>84</td>
</tr>
<tr>
<td>Sor</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>83</td>
<td>80</td>
</tr>
<tr>
<td>Cd</td>
<td>89</td>
<td>91</td>
<td>-</td>
<td>69</td>
<td>72</td>
</tr>
<tr>
<td>Prn</td>
<td>94</td>
<td>100</td>
<td>-</td>
<td>94</td>
<td>85</td>
</tr>
<tr>
<td>Rgn</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>86</td>
<td>85</td>
</tr>
<tr>
<td>Pog</td>
<td>97</td>
<td>100</td>
<td>95</td>
<td>-</td>
<td>77</td>
</tr>
<tr>
<td>A</td>
<td>77</td>
<td>94</td>
<td>95</td>
<td>-</td>
<td>85</td>
</tr>
<tr>
<td>B</td>
<td>71</td>
<td>85</td>
<td>-</td>
<td>-</td>
<td>100</td>
</tr>
<tr>
<td>Ls</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>95</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.2 shows the success rates that are reported in different studies on automatic cephalometric landmark detection and the proposed algorithm with respect to manually located landmarks. Table 6.3 gives the results in terms of consistency and precision respectively. Table 6.2 gives a first appraisal of the relative effectiveness of our approach. 89% of the localization of 18 selected landmarks is within a window of \( \pm 2 \) mm. The average ME is 1.86 mm, and average SD of mean error is 1.24 mm. The results are better or comparable to the existing techniques. However, this comparison is not adequate enough as the result may vary owing to quality of cephalograms, randomness and size of database used in each of the studies. Furthermore, distinct studies use different landmarks for analysis and distinctive reporting method.
Table 6.3: Comparison of recognition results in terms of mean location error (ME) and standard deviation (SD) in millimeter (mm) when reported.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Landmark</td>
<td>ME</td>
<td>ME± SD</td>
<td>ME± SD</td>
<td>ME</td>
<td>ME± SD</td>
<td>ME± SD</td>
</tr>
<tr>
<td>Pr</td>
<td>5.13</td>
<td>5.67±3.37</td>
<td>7.3±6.5</td>
<td>3.48±2.46</td>
<td>1.57±0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sella</td>
<td>1.41</td>
<td>5.06±3.37</td>
<td>5.5±6.8</td>
<td>1.92</td>
<td>3.24±2.85</td>
<td>1.26±1.11</td>
<td></td>
</tr>
<tr>
<td>Na</td>
<td>1.83</td>
<td>2.57±2.18</td>
<td>5.6±2.4</td>
<td>1.40</td>
<td>2.97±1.85</td>
<td>2.30±0.90</td>
<td></td>
</tr>
<tr>
<td>Or</td>
<td>1.28</td>
<td>2.46±3.77</td>
<td>5.5±3.4</td>
<td>1.92</td>
<td>3.42±2.43</td>
<td>1.64±1.13</td>
<td></td>
</tr>
<tr>
<td>Pns</td>
<td>2.16</td>
<td>5.0±4.1</td>
<td>1.32</td>
<td>3.04±1.38</td>
<td>2.20±1.77</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ans</td>
<td>2.36</td>
<td>2.64±3.06</td>
<td>3.8±2.2</td>
<td>0.75</td>
<td>2.70±1.05</td>
<td>1.93±1.12</td>
<td></td>
</tr>
<tr>
<td>Go</td>
<td>1.00</td>
<td>2.72</td>
<td>5.8±6.0</td>
<td>1.11</td>
<td>3.64±1.76</td>
<td>1.77±1.02</td>
<td></td>
</tr>
<tr>
<td>Uie</td>
<td>2.09</td>
<td>2.02±1.99</td>
<td>2.7±3.4</td>
<td>0.48</td>
<td>4.40±2.03</td>
<td>1.94±0.91</td>
<td></td>
</tr>
<tr>
<td>Me</td>
<td>3.09±3.46</td>
<td></td>
<td>2.7±3.4</td>
<td>0.48</td>
<td>4.40±2.03</td>
<td>1.94±0.91</td>
<td></td>
</tr>
<tr>
<td>Sor</td>
<td>1.85±2.26</td>
<td>2.7±3.4</td>
<td></td>
<td></td>
<td>2.17±1.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cd</td>
<td></td>
<td>2.55±0.97</td>
<td>2.95±2.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prn</td>
<td></td>
<td>2.23±1.24</td>
<td>1.58±1.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rgn</td>
<td></td>
<td>1.97±1.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pog</td>
<td>2.91</td>
<td></td>
<td></td>
<td></td>
<td>1.34±0.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.78</td>
<td>2.33±2.63</td>
<td>3.3±2.4</td>
<td>0.90</td>
<td>1.59±1.24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>3.29</td>
<td>1.85±2.09</td>
<td>2.6±2.7</td>
<td>0.95</td>
<td>2.18±1.36</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ls</td>
<td>0.54</td>
<td></td>
<td></td>
<td></td>
<td>1.11±2.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Li</td>
<td>0.34</td>
<td></td>
<td></td>
<td></td>
<td>1.83±1.07</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To better compare our method with other techniques the method proposed by Yue et al. [17] was implemented. The results were verified by considering 65 images from our database for training and 20 images for testing. The comparison of results is presented in Table 6.4. Yue reported a success rate of 71% on his database of 200 training images and 86 testing images. However, when the method is applied on our database, the success rate comes down to 50%. Following could be some reasons for the decrease in success rate:

Yue’s method employs ASM to localize landmarks. It combines global shape model with a local search approach.

i. The smaller size of database is not able to capture the variability of the structure in ASM model effectively.

ii. The method to estimate the initial shape for the model is not robust. It uses five landmarks extracted using edge based handcrafted method to compute features for finding the most similar training set images. The landmarks on these training set images are used to find the initial shape. This technique is highly dependent on the quality of cephalograms and database. Many images
in our database suffer from double edges and nonuniform illumination effects. Furthermore, the database size is much smaller.

iii. Two parameters needed for an optimal local search, the size of the gray level model that is used to define the search at each point and how far from each template point the search should extend are mentioned in the paper. The proposed algorithm achieves higher success rate by considering.

i. Zernike moment based global matching for finding 5% the most similar training set images for finding initial landmark locations.

ii. Use of multiple smaller search windows instead of one large window to avoid false detections and more than one template for each landmark to capture the variability of structures using many measures to find the similarity.

The structural complexity of different landmarks is different and affects their accurate extraction. The detection of landmark Cd is difficult owing to superimposition of multiple structures and very low contrast. Landmark Go is on clear curve but sometimes the presence of double edges and overlap with other structures affects its accuracy. Points A, Point B, Pog and Rgn lie on vertical curves and Me lie on the horizontal curve with generally well defined curves but the landmarks on chin sometimes suffer from nonuniform illumination and broad curves. For Pog and Rgn the error is higher in y plane and Me in the x plane. In Pns, the reason for inconsistence is due to lack of clarity of structure. Ans is difficult to identify because it lies on thin and pointy structure, which has a low contrast from the soft tissue. The source of error in case of Or is very broad and cluttered edges. Sor also suffers from cluttered edges. Soft tissue landmarks Li and Ls lie very close and have a similar structure and one is classified as the other that leads to decrease in consistence and precision. The performance can be enhanced by adding a simple constraint Prn is detected easily in the y plane, but as it lies on a curved surface, it’s difficult to pinpoint its location in the x plane. The variability in Uie is the highest and thus it’s difficult to capture it through four templates. Na lies at the intersection of two clearly identifiable structures. But incorrect position of cephalostat can affect its extraction. Sella is easily identified but sometimes suffer from false detections due to similar structures nearby. Landmark Pr is detected easily. The results corresponding to SSD and NCC in terms of ME and SD of mean error as obtained using stage three of the proposed algorithm are given in Fig. 6.4 and Fig. 6.5 respectively.
Figure 6.4: Mean location error (ME) in millimeters (mm) for template matching using SSD and NCC.

Figure 6.5: Standard deviation (SD) of location error in case of SSD and NCC.
For some landmarks, SSD gives better results and for others NCC. Thus, the overall detection improves by using their combination.

6.5. Conclusions

In this study, we investigate the automatic cephalometric algorithm proposed by Yue et al. [17] and propose a new idea to detect cephalometric landmarks with improved performance. Unlike most existing methods that used features based on few reference landmarks (extracted using edge based handcrafted methods) to select the most similar training set images for initial approximation of landmarks.

The proposed algorithm use Zernike moment based features for comparing the query image to the training set images. This leads to improved initial approximation of landmark's location, which further help improve the results in the next steps. A success rate of 89% within a precision of ±2 mm is achieved with average ME of 1.86 mm and SD of mean location error 1.24 mm.

The results may be further improved by

i. Removing the cephalograms with incorrect head position from the data set.
ii. By correcting nonuniform illumination effects in the cephalograms.
iii. Using larger data set that covers the variability of anatomical structure better.
iv. By the addition of simple easily available information regarding the patient like gender and age. This will help in limiting the search to a subset of the data set that results in closer matches.
v. The accuracy is improved by using higher-resolution images. However, the computational complexity increases.
vi. By exploiting the relationship between the landmarks accuracy improves. However, addition of new landmarks will become difficult.

The problem is very challenging. Each landmark is unique with its own set of features and level of complexity. Optimizing the algorithm for one landmark may affect the accuracy of some other landmarks. For making it clinically viable it will always need an option where the orthodontist can adjust the located landmarks if he desires to.