

be exactly determined by a big-magnitude offset diffusion, and landmarks of tiny structures corresponding to a small offset obtain their approximate positions from the knowledge of point relationship to those distinguished ones. The actual locations will be available after dilations and erosions.

The remainder of this paper is organized as follows. contributions are depicted in Section 2 and Section 3. Performance demonstrated in Section 4 shows that the new method is quite stable to extract feature points in a big scale, and significant for other landmarks detection as well, no matter the landmarks are of a big scale or in a tiny structure. Automatic detected landmarks provide a set of reference points for further template deformation.

II. KNOWLEDGE BASED DIFFUSION FOR LAYERED STRUCTURE REPRESENTATION

Anisotropy diffusion happens in homogenous areas and along edges, but stops when crossing object contour. It produces adaptively sharp boundaries of smooth regions. Adjusting parameters of a diffusion model, we could yield images in different scales for landmark detection.

Perona and Malik firstly suggested diffusion using partial differential equation(PDE) for image f [6]:

$$\partial f(x, y, t) / \partial t = \text{div} (g \parallel \nabla f \parallel \nabla f) \quad (1)$$

where (x,y) denotes a pixel position, $t=0$ corresponds to the initial state, ∇f represents a gradient image of f , and $g(\cdot)$ is a positive monotonous descending function satisfying the following conditions: it is large enough to cause diffusion at a position located in a smooth region, and approaches zero to preserve discontinuity when crossing edges. This diffusion model is computationally expensive because of its iteration. Fortunately B. Fischel et al. proposed a fast realization[7], in which the serial PDE diffusion is replaced with parallel kernel operations performed in one single step, the kernel size is viewed as offsets relatively to discontinuities.

In [7], an offset vector(OV) is estimated by:

$$v(x, y) = m(x, y) \frac{o(x, y)}{|o(x, y)|} \quad (2)$$

where

1) $o(x,y)$ indicates offset orientation normal to image border and usually takes a simple form of gradient as follows:

$$o(x, y) = \nabla f(x, y) \quad (3)$$

2) $d(x,y)$ is offset direction, i.e. the sign for an OV, with $d(x,y)=1$ in the orientation direction and $d(x,y)=-1$ denoting the opposite case. It is expressed as:

$$d(x, y) = -\text{sgn} \left[\int_w (o(x, y) \nabla f(x + x', y + y')) (o(x, y) \text{Dis}(x', y')) dx' dy' \right] \quad (4)$$

Here w is a filtering window for integral, (x',y') is a pixel in w with a relative coordinate to the filter center (x,y) , and $\text{Dis}(x',y')$ is a relative distance as well.

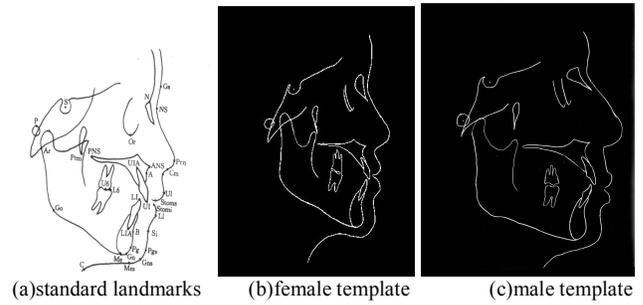


Figure1. Templates.

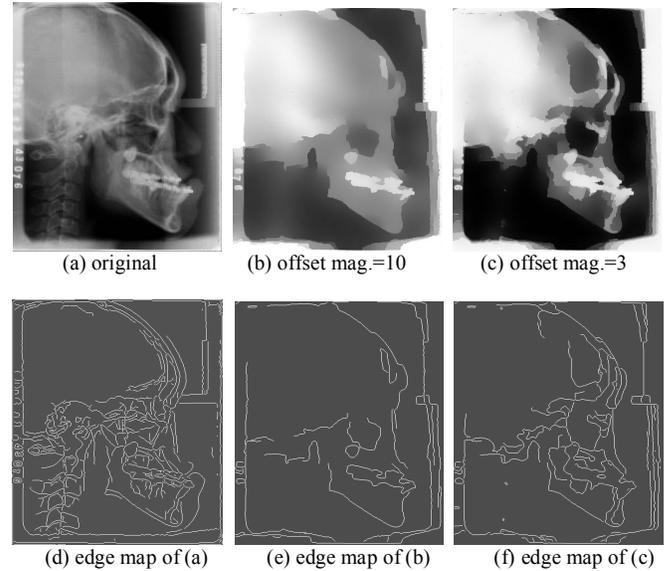


Figure 2. Nonlocal filtering results and their edge maps.

3) $m(x,y)$ refers to offset magnitude, and determines how far the displacement should be forward. $m(x,y)$ is acquired by the following steps:

· Generate initial OV with (5):

$$v_i(x, y) = d(x, y) o(x, y) \quad (5)$$

· Search a zero-crossing point (x',y') in the offset direction to produce offset magnitude with

$$m(x, y) = \min \alpha: v(x + \alpha v_i(x, y), y + \alpha v_i(x, y)), v_i(x, y) \leq 0 \quad (6)$$

The final OV field is formed by letting

$$v(x, y) = m(x, y) \frac{v_i(x, y)}{|v_i(x, y)|} \quad (7)$$

With parallel diffusion, features of various scales are filtered non-locally at the offset positions using kernels of different sizes. The size is a crucial factor in diffusion. A large offset relates to a longer displacement vector. It will reduce the subtle local structure, especially to feature points UIA , UI , LI , LLA , $U6$ and $L6$ etc. as shown in Fig.1(a), whereas it retains the global landmarks on the contours of big structures, such as N , Or , Ans , A , B , Pg , Gn , Me Go and Ar . On the other hand, a detailed region is associated with a small scale and a small magnitude. Thus various diffusion sizes could be used to adapt the structure scale. Since the craniofacial template is knowledge based, the suitable size for different structures are

available by training. Fig. 2 are images accomplished by non-local filtering and edge maps. Fig. 2(a) is the original cephalogram, Fig. 2(b) and Fig. 2(c) shows the results with offset magnitudes 10 and 3. Because of the similarity of human brains, the training sizes are robust to most people. Hence by making use of diffusion results, we could evaluate the appropriate positions for feature points and derive them with a new idea described in the next section.

III. LANDMARK LOCATION BY EDMS OF DILATED SKELETONS

Isotropic pixels inside a region are filtered in PDE processing to produce an integral surface, whereas diffusion decreases rapidly around the edges. This behaviour finally generates segmentation-like result with a visually sharp outline and landmarks located at the junctions of area contour. If we perform dilations on their edge maps, each junction will become a pixel-clique and appear as a thick point in EDM, taking a big value greater than the dilation times. Hence a binary operation could clearly pick up the accurate positions for landmark candidates.

EDM is an elevation model calculating the straight distance between an inside-feature pixel and its nearest boundary, these distance are encoded by gray values and form an EDM image. Assume a set U representing a feature, and ∂U is the boundary of U , an EDM is defined as follows:

$$\{ U_{EDM} : u_i = 0, \forall u_i \text{ not belongs to features};$$

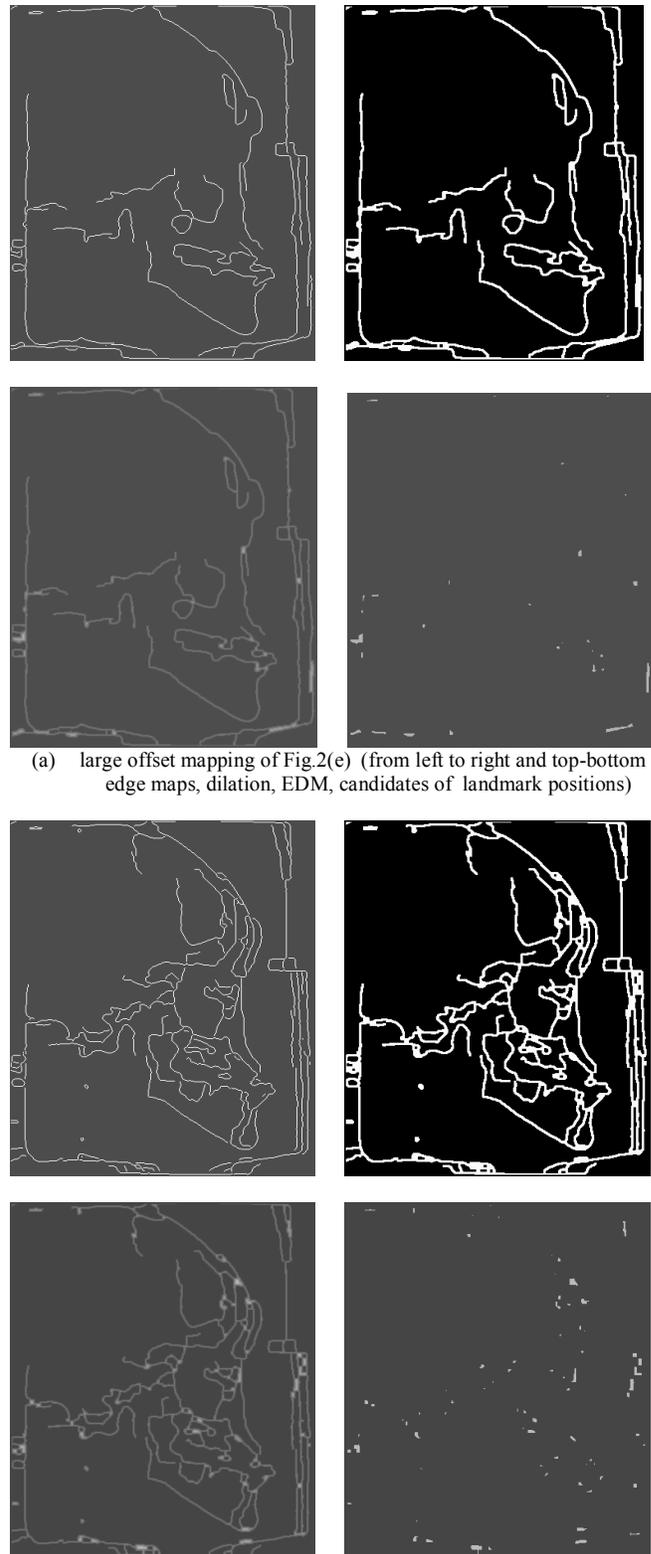
$$AND u_i = d_\epsilon(u_i, \partial U), \forall u_i \text{ belongs to features} \} \quad (8)$$

where $d_\epsilon(u_i, \partial U)$ denotes the smallest distance from the pixel u_i to the feature boundary. Conceptually, the time cost for the construction of an EDM depends on the feature sizes, it may be relative expensive due to the width of an object. Most of the current EDM techniques, such as the ordered propagation through bucket sorting of the pixels are quite complicated and not in very high speed[8].

In this paper, we use our fast and simple algorithm scanning the input image with no more than four passes to create EDM:

- 1) The first scanning computes the minimum distance from a pixel to its feature horizontal boundary. It goes from left to right and top to bottom. For each pixel in an EDM row, if it is a feature start-point, the corresponding map takes value 1. Within the same feature, each pixel after the first one maps its magnitude one greater than the former neighbor. When a feature endpoint is met, pixels from the end of the feature to its half-length are reset from 1 to length/2.
- 2) The second pass accomplishes calculation of the smallest distance in vertical direction. It is the same as the first one except proceeding from top to bottom and left to right.

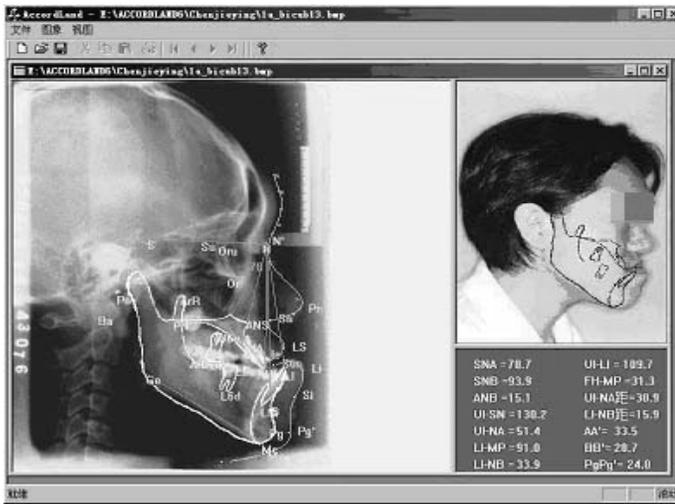
The maps available after the above passes could indicate the 4-neighbor-rule distances of all the pixels within a feature, but it is just a closer approximation in a sense of the 8-neighbor-rule. For pixels diagonally adjacent to the boundaries, the distances will be one pixel biased actually. The next two propagating passes modifies this situation.



(a) large offset mapping of Fig.2(e) (from left to right and top-bottom : edge maps, dilation, EDM, candidates of landmark positions)

(b) tiny structures mapping of Fig.2(f) (from left to right: edge maps, dilation, EDM, candidates of landmark positions)

Figure 3. EDM based edge-map dilation & landmark extraction.



(a)



(b)

Figure 4. Landmark detection results and measurement.

3) The third scanning advances from left to right and top to bottom, it assigns each pixel in a feature a gray value one greater than the minimum of any of its neighbors. i.e., if the 4-neighbor-rule distance maps is U_F , the image after propagation is U_G , N_F is the neighbor field of pixel p in U_F , p' is the pixel in U_G located at the same position as p , the following equation will hold:

$$U_G(p') = \min(N_F(U_F(p))) + 1 \quad (9)$$

The left propagation performs a repeat step, working from bottom to top and right to left.

Since the ridge is constituted by the locally biggest values in EDM, it corresponds closely to the original skeleton, thus the dilation skeleton and the dilation junctions are of different widths. We could immediately gain a straightforward way to get unique landmark positions. Here a repeated dilation and binarization scheme is introduced for the purpose. The location procedures are follows:

Step 1. Set initial dilating times $D_t(0)=1$;

Step 2. Repeat dilations as $D_t(i)$ denoted and perform binary operation with a threshold $D_t(i)-1$, then count the pixel-cliques and record it as $N_c(i)$.

Step 3. if $N_c(i) \geq N_c(i-1)$, $D_t(i+1) = D_t(i)+1$, goto Step 2; else stops.

After dilations in Step2, edges containing obvious terminals and crossing points are obtained with their small intervals filled in. On the other hand, binarization could eliminates the smooth skeleton, and reserve the pixel-cliques which correspond to the key points around landmarks in that scale. The knowledge-based template is used to assure a truly mapping of cliques to landmarks, firstly referred to the large offset result and then to the tiny structure result. Fig.3(a) and Fig.3(b) show the edge-map dilation and landmark extraction based on EDM, acquiring landmark candidates in large structures and in tiny scale. Each clique provides a feature point candidate for template matching The cliques are finally peeled off as shown in Fig. 4. The left side of Fig. 4(a) demonstrates the positions of 12 landmarks $S, N, Or, Po, Ar, ArR, ANS, A, B, Pg, Me$ and Go , and approximate locations of 2 landmarks PNS and LI . The right bottom of Fig. 4(a) presents the evaluation result of feature angles and distances according to the feature point positions. Fig. 4(b) reveals the landmark locations of another patient.

IV. EXPERIMENTS AND CONCLUSIONS

The proposed method was experimented with 30 cases of clinical surgeries to compare the detected landmark positions and those actual parameters measured by orthodontists. We evaluate the match performances with t -testing method, setting its confidence coverage to 0.95, i.e., $\alpha=0.05$, and $100(1-\alpha)\%=0.95$. Table I gives the behavioral meaning: 9 angle parameters are meaningless in statistics with probability intervals ($P>0.05$), this result illuminates that angle measurements in our method are accurate; among 5 distance parameters, only one parameter ($B-B'$) testing has a value $P<0.05$, this also implies that most distance measurements are correct. The wrong position of B reflects intrinsic difficulty of labium location, because mandible is a linkage joint, its landmark positions are affected by many factors.

Owing to edge-preserving diffusion and multi-layer separation, feature points, which are usually missed in the conventional methods, could be perfectly extracted in this paper. Furthermore, since the dilation key points take obviously different EDM values to those of smooth skeleton pixels, their positions could be simply obtained by a binary operation. As we could observe from the discussions and experimental results, the proposed approach is feasible and superior to most classic methods. It is robust and efficient for easy and accurate landmark location.

ACKNOWLEDGEMENT

The authors acknowledge the kindly supports of China NNSF(Grant No: 60472063.60325310) and GDNSF/CNLF (04020074-CN200402).

TABLE I
LANDMARK LOCATIONS AND MEASURED PARAMETERS COMPARISON

| Parameters | \bar{x} | s | t | P |
|----------------|-----------|---------|---------|-------|
| $\angle SNA$ | 0.068 | 0.5034 | - 0.592 | 0.561 |
| $\angle SNB$ | 0.032 | 0.2673 | 0.599 | 0.555 |
| $\angle ANB$ | - 0.015 | 0.1515 | - 0.518 | 0.609 |
| $\angle U1-SN$ | 0.110 | 0.2845 | 0.729 | 0.100 |
| $\angle U1-NA$ | 0.030 | 0.3922 | 0.342 | 0.736 |
| $\angle L1-MP$ | - 0.104 | 0.2746 | -1.894 | 0.070 |
| $\angle L1-NB$ | 0.052 | 0.3537 | 0.735 | 0.469 |
| $\angle U1-LI$ | - 0.127 | 0.3353 | -1.930 | 0.065 |
| $\angle FH-MP$ | - 0.059 | 0.2666 | -1.184 | 0.246 |
| U1-NA | 0.195 | 0.8835 | 0.987 | 0.366 |
| L1-NB | -0.020 | 0.1225 | -0.816 | 0.422 |
| A-A' | -0.040 | 0.1930 | -0.927 | 0.366 |
| B-B' | 0.364 | 0.73612 | 0.320 | 0.030 |
| Pg-Pg' | -0.040 | 0.2406 | -0.083 | 0.934 |

REFERENCES

- [1] I. El-Feghi, M.A. Sid-Ahmed and M. Ahmadi, "Automatic Localization of Craniofacial Landmarks for Assisted Cephalometry," *Patt. Recog.*, vol. 37, no.2, pp. 609-621, 2004
- [2] Y.T. Chen, K.S. Cheng, J.J. Huang, "An Intelligent View Box System for Cephalograms," *IEEE Trans. on Instr and Meas.*, vol..54, no. 1, pp..98-104.,2005
- [3] J.L.Contereras-Vidal, J.Garza, "Knowledge-Based System For Image Processing and Interpolation of Cephalograms," *Proc. of the Canadian Conf. on Elect. and Comput. Eng.*, vol. 75, pp.11-14. 1990
- [4] T.J. Hutton, S. Cunningham and P. Hammond, "An Evaluation of Active Shape Models for the Automatic Identification of Cephalometric Landmarks," *Eur. J Orthod*, vol..22, pp.499-508, 2000
- [5] I. El-Feghi, M.A. Sid-Ahmed and M.Ahmedi, "Location of Craniofacial Landmarks on X-Ray Image by Employing Fuzzy Neural Network," *IEEE Conf.-.*, pp. III348-III351, 2002
- [6] P. Perona and J. Malik, "Scale-Space and Edge Detection Using Anisotropic Diffusion," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.12 , no. 7, pp..629-639, 1990
- [7] B. Fischl and E.L. Schwartz, "Adaptive Nonlocal Filtering: A Fast Alternative to Anisotropic Diffusion for Image Enhance.," *IEEE Trans.on Patt. Anal. Mach. Intell.*, vol.21, no. 1, pp.42-48,1999
- [8] O.Cuisenaire,B.Macq,"Fast Euclidean Morphological Operator Using Local Distance Transform. by Propagation & Application," *Conf. of Image Proces. and Its Application*, pp.856-960, 1999