# **Original Article**

# Automated Cephalometry: System Performance Reliability Using Landmark-Dependent Criteria

# Chihiro Tanikawa<sup>a</sup>; Masakazu Yagi<sup>b</sup>; Kenji Takada<sup>c</sup>

#### ABSTRACT

**Objective:** The purpose of the present study was to evaluate reliability of a system that performs automatic recognition of anatomic landmarks and adjacent structures on lateral cephalograms using landmark-dependent criteria unique to each landmark.

**Materials and Methods:** To evaluate the reliability of the system, the system was used to examine 65 lateral cephalograms. The area of each system-identified anatomic structure surrounding the landmark and the position of each system-identified landmark were compared with norms using confidence ellipses with  $\alpha = .01$ , which were derived from the scattergrams of 100 estimates obtained according to the method reported by Baumrind and Frantz. When the system-identified area overlapped with the norm area, anatomic structure recognition was considered successful. In addition, when the system-identified point was located within the norm area, landmark identification was considered successful. Based on these judgment criteria, success rates were calculated for all landmarks.

**Results:** The system successfully identified all specified anatomic structures in all the images and determined the positions of the landmarks with a mean success rate of 88% (range, 77%–100%).

**Conclusion:** With the incorporation of the rational assessment criteria provided by confidence ellipses, the proposed system was confirmed to be reliable. (*Angle Orthod.* 2009;79:1037–1046.)

KEY WORDS: Cephalograms; Automatic recognition; Landmarks

#### INTRODUCTION

Lateral cephalograms are essential in contemporary orthodontic diagnosis and treatment planning. Anatomic landmark identification based on the visual assessment of the lateral cephalogram remains a task that requires specially trained clinicians. A fully automated clinical examination of cephalograms would reduce the workload during routine clinical service and would provide orthodontists with more time for optimum treatment planning. Various methods such as a knowledge-based technique with edge tracking,<sup>1-4</sup> model-based approaches,<sup>5-8</sup> pattern-matching techniques,<sup>9-11</sup> and combined algorithms<sup>12-16</sup> have been developed and are available. However, most of these methods have not been adopted in clinical practice.<sup>17</sup>

Recently, a system that recognizes general grayscale images using an automated psychologic brain model,<sup>18</sup> ie, a hardware-friendly algorithm to accomplish real-time recognition by recalling a set of modeled data that is mathematically described using a finite number of traits and previously stored in the system, has been developed. This system employs a new technique called the projected principal edge distribution (PPED) as a means for extracting features from an image, and it has been confirmed that the system demonstrates robust performance in recognizing images, including cephalograms.<sup>19,20</sup> Although experiments have suggested the efficacy of the system in recognizing images, it remains uncertain whether such

<sup>&</sup>lt;sup>a</sup> Assistant Professor, Department of Orthodontics and Dentofacial Orthopedics, Graduate School of Dentistry, Adjunct Assistant Professor, Center for Advanced Medical Engineering and Informatics, Osaka University, Osaka, Japan.

<sup>&</sup>lt;sup>b</sup> Associate Professor, Center for Advanced Medical Engineering and Informatics, Osaka University, Osaka, Japan.

<sup>°</sup> Professor and Chair, Department of Orthodontics and Dentofacial Orthopedics, Graduate School of Dentistry, Adjunct Professor, Center for Advanced Medical Engineering and Informatics, Osaka University, Osaka, Japan.

Corresponding author: Kenji Takada, DDS, PhD, Department of Orthodontics and Dentofacial Orthopedics, Graduate School of Dentistry, Osaka University, Osaka, Japan. 1-8 Yamadaoka, Suita, Osaka, Japan 565-0871

<sup>(</sup>e-mail: opam@dent.osaka-u.ac.jp)

Accepted: October 2008. Submitted: September 2008. © 2009 by The EH Angle Education and Research Foundation, Inc.



Figure 1. Schematic illustration of the system employed in the present study.

a system will detect conventionally used landmarks with high precision. On the other hand, a previous study<sup>21</sup> documented that topographic variations exist in humans' subjective judgments of cephalometric landmarks, and the shapes and size of the variances are unique to each landmark. Mathematical formulation of these landmark-dependent variations in measurement would be help researchers to evaluate objectively the reliability of the automatic cephalogram recognition system.

The purpose of the present study was to examine the reliability of a system that performs automatic recognition of anatomic landmarks and their surrounding anatomic structures in which the landmarks are located on lateral cephalograms using landmark-dependent criteria unique to each respective landmark.

## MATERIALS AND METHODS

#### System Overview

The system employed in the present study, which performs automatic recognition of anatomic landmarks and their surrounding anatomic structures on lateral cephalograms, is shown in Figure 1. A detailed description of the system has been reported elsewhere.<sup>1,19</sup> Briefly, the system incorporates two major tasks: the "knowledge-generation" (system learning) phase and the "recognition" phase. In the knowledgegeneration phase, image data extracted from learning samples are converted into PPED vectors consisting of 64 variables that feature contours of the anatomic structures.<sup>1,18,19</sup> From these vectors, template vectors, ie, the principal information for identifying the landmarks, are generated using a generalized Lloyd algorithm<sup>22</sup> for each landmark, which are stored in the system as the system's knowledge. During the recognition phase, the system is designed to perform pixel-by-pixel film scanning with template-matching operations between PPED vectors that are generated from an input film and template vectors stored on the system. The system recognizes the most matched position as a landmark position.

#### Data

Pretreatment lateral cephalograms were obtained from 465 Japanese orthodontic patients (147 male and 318 female patients; mean age, 22 years 6 months; age range, 10 years 9 months to 60 years) who had visited the university dental hospital between the years 1998 and 2003. Patients were enrolled consecutively. Criteria for selection were permanent dentition, no congenital anomalies, and no missing teeth or metallic restorations. Digital lateral head films of each patient, with a magnification ratio of 1:1.1, were taken with the teeth in habitual maximum intercuspation position and the lips in repose. The sample was divided into two groups at random: 400 cephalograms for system learning and 65 cephalograms for testing the system's performance.

Each record was cross-marked directly by one of the authors (CT) on the top right and bottom left cor-

 Table 1. Definitions of 20 Anatomic Landmarks Employed in the Present Study

Landmark	Definition
S	The center of sella turcica <sup>23</sup>
Ν	The junction of the frontonasal suture at the most posterior point on the curve at the bridge of the nose $^{\rm 24}$
Or	The lowest point on the inferior margin of the orbit <sup>25</sup>
Po	Point on the upper margin of the porus acusticus ex- ternus <sup>25</sup>
Ва	Point where the median sagittal plane of the skull inter- sects the lowest point on the anterior margin of the foramen magnum <sup>25</sup>
ANS	The tip of the anterior nasal spine <sup>23</sup>
PNS	The posterior limit of the palatine bone <sup>23</sup>
Point A	The most posterior point on the curve between ANS and prosthion <sup>24</sup>
Point B	The point most posterior to a line from infradentale (Id) to Pog on the anterior surface of the symphyseal outline of the mandible <sup>24</sup>
Pog	The most anterior point on the contour of the symphy-
0	sis determined by a tangent through Id.
Me	The most inferior point of the symphysis to the line Id- Pog <sup>a</sup>
Gn	Intersection between the perpendicular bisector of Pog- Me and the anterior contour of the symphysis <sup>a</sup>
Go	External angle of the mandible, located on the lateral radiograph by bisecting the angle formed by tan- gents to the posterior border of the ramus and the inferior border of the mandible <sup>25</sup>
Ar	Intersection of the lateral radiographic image of the posterior border of the ramus with the base of the occipital bone <sup>25</sup>
Cd	The top of the head of the condyle <sup>23</sup>
U1	The upper central incisal tip <sup>22</sup>
U1_C	The labiopalatal midpoint of the visible image of the root of the upper central incisor <sup>a</sup>
L1	The lower central incisal tip <sup>23</sup>
L1_C	The labiolingual midpoint of the visible image of the root of the lower central incisor <sup>a</sup>
Ptm	The most anterior inferior confluence of the curvatures, which is a bilateral, upside-down, teardrop-shaped area of radiolucency, the anterior surfaces of which are taken as the posterior surfaces of the maxilla <sup>24</sup>

ners of the film with a pencil for use as reference points. The films were then digitized in a scanner (ES8500, EPSON, Tokyo, Japan) at a resolution of 300 dots per inch to provide an image dataset (2320  $\times$  2960 pixels, 1 pixel = 0.085 mm; hereafter referred to as a raw image dataset). The films were then traced with a pencil on acetate paper overlaid on the films. Twenty anatomic landmarks<sup>23-25</sup> (Table 1) were identified visually, cross-marked by one of the authors, and double-checked by another author. The degree of certainty was recorded for each landmark using the following three subjective judging scores: 1 = absolutely correct; 2 = probably correct; and 3 = difficult to recognize. All procedures were performed on a light box with identical writing tools in an air-conditioned and darkened room.

To obtain data from the traced images, the films with the traced papers were digitized using the scanner following the same method as described to obtain the "traced image data." Using a mouse, one of the authors identified the positions of reference points and anatomic landmarks on both the raw image data and the traced image data on a computer monitor (17-inch LCD monitor, 1701FP, Dell, Round Rock, Tex). The images were used at their actual size and magnified to 200% of their original size. The positions of the reference points on the two types of images were mathematically superimposed to provide coordinate values for the landmarks on the raw image data.

The raw image data and the position data thus were used for system's learning. Fifteen template vectors were generated as system knowledge and stored in the system according to a previously reported method.<sup>20</sup> Positions assigned to the degrees of certainty of 2 or 3 were ruled out from the dataset for the system learning/knowledge-generation phase of data collection.

### **Confidence Ellipses**

To evaluate the system's performance reliability, scattergrams that designated errors for manual landmark identification when 10 orthodontists identified a landmark on 10 cephalograms were obtained according to the method reported by Baumrind and Frantz.<sup>21</sup> (For details, see Appendix.) Confidence ellipses with a confidence limit of  $\alpha$  were developed for each landmark from the scattergram, using equation (1).

$$\alpha = CHI2 \left[ \frac{(x/\sigma_x)^2 - 2\rho(x/\sigma_x \cdot y/\sigma_y) + (y/\sigma_y)^2}{1 - \rho^2} \right] \quad (1)$$

where *CHI2* is the function that provides the one-tailed probability  $\alpha$  of the chi-square distribution with 2 degrees of freedom; *x* and *y* are the coordinate values;  $\sigma_x$  and  $\sigma_y$  are standard deviation values for *x* and *y*, respectively; and  $\rho$  is the covariance of correlation between *x* and *y*.<sup>26</sup> The parameter  $\alpha$  was assigned to 1.0 when the input position was at the best estimate position, whereas the value was assigned closer to zero when the input was in a biased position from the best estimate position.

The confidence ellipses with  $\alpha = .01$  for the landmarks are shown in Figure 2A,B. The parameters (the angular measurement between the *x*-axis and the semimajor axis, the lengths of the semimajor and semiminor axes) representing the confidence ellipses with  $\alpha = .01$  are provided in Table 2. The average for the semiminor axis was 2.68 mm (range, 1.35–4.22 mm) and for the semimajor axis the average was 4.66 mm (range, 2.99–10.20 mm). The landmark with the shortest semiminor axis was nasion.

#### System Performance Test

After learning the image characteristics of the 400 cephalograms, the automatic recognition system was asked to identify 65 cephalograms that had been reserved to test the performance of the system. Landmark positions assigned to a 3 for degree of certainty were ruled out of the dataset for this task.

First, for each landmark, the system computed an area that was the most probable location of the anatomic structure surrounding the landmark as a minimum rectangular area that included the first 50 candidate positions of the landmark (hereafter referred to as "search area") in 1/16 downscaled targeted images. If the fiducial zone, designated by a confidence ellipse with  $\alpha = .01$ , was found to overlap the search area, recognition of the anatomic structure was judged successful.

Second, for each landmark, the system computed the most probable position of the landmark by examining targeted images with original resolution in the search area. The success or failure of the assessment by the system was evaluated using confidence ellipses with  $\alpha = .01$ . In short, when a system-identified point was located within a confidence limit of  $\alpha = .01$ , the landmark identification was judged to be successful.

The success rates for the recognition of the landmark and the anatomic structure surrounding that landmark were defined as the proportion of the total samples that could be successfully recognized by the system. The success rates for all the landmarks were calculated. All procedures were carried out on a workstation (Sun Blade 2000, Sun Microsystems, Palo Alto, Calif).

#### RESULTS

The system successfully recognized all anatomic structures surrounding all landmarks (sella turcica, nasofrontal junction, infraorbital area, mandibular symphysis, etc) within a 20.5-  $\times$  20.5-mm search area (range, 11.9-  $\times$  11.9-mm to 27.7-  $\times$  27.7-mm). The positions of the 20 cephalometric landmarks identified by the system are given in Figure 3. The mean suc-



**Figures 2A and 2B.** Confidence ellipses obtained for cephalometric landmarks. Black points indicate coordinate values of landmarks identified by 10 orthodontists on 10 cephalograms. The black lines designate confidence ellipses with  $\alpha = .01$ . Origin indicates the best estimate; *x*-axis, the line that passes through the origin and is parallel to the line S-N; and *y*-axis, the line that is perpendicular to the *x*-axis through the origin.

Figures 2A and 2B. Continued.



**Table 2.** Parameters (Angular Measurements Between the *x*-Axis and the Semimajor Axis, the Lengths of the Semimajor and Semiminor Axes, and the Distribution Patterns) That Represented the Confidence Ellipses With  $\alpha = .01$ 

	Angle Between			
	<i>x</i> -Axis and			
	Semimajor	Semiminor	Semimajor	Distribution
Landmark	Axis (deg)	Axis (mm)	Axis (mm)	Pattern
Gn	-19.1	2.8	3.5	Round
Point A	-71.6	3.2	3.4	Round
Po	-30.2	3.5	4.6	Round
Go	56.1	3.7	4.7	Round
Cd	-4.4	4.2	5.2	Round
S	1.9	1.5	3.0	Horizontal
Ν	1.7	1.4	3.3	Horizontal
L1	12.9	1.8	3.2	Horizontal
Ar	-4.2	1.5	3.5	Horizontal
U1	11.0	1.8	3.3	Horizontal
ANS	3.7	1.7	3.5	Horizontal
Or	14.1	1.8	4.2	Horizontal
PNS	2.7	1.6	4.9	Horizontal
Me	9.5	3.0	4.6	Horizontal
Point B	-62.2	3.0	4.4	Right-up
Pog	-63.4	3.1	5.2	Right-up
Ba	-38.7	3.8	5.5	Right-up
L1_C	-49.2	2.9	6.4	Right-up
Ptm	-86.1	3.9	10.2	Vertical
U1_C	68.4	3.4	6.6	Right-down



Figure 3. Positions of the 20 cephalometric landmarks that were identified and exemplified by the system.

**Table 3.** Success Rates in Identifying Landmark Positions with the

 Present System

Landmark	Success Rate (%)		
Ме	100		
Point B	98		
Ptm	98		
Pog	95		
Ва	94		
L1_C	94		
U1_C	94		
Gn	91		
Point A	88		
U1	88		
Po	87		
Cd	86		
ANS	83		
Or	82		
Ar	82		
PNS	80		
L1	80		
S	78		
Go	78		
Ν	77		
Average	88		

cess rate for identifying the landmark positions was 88%, with a range of 77% to 100%. The system demonstrated a 100% success rate in recognizing Me and success rates above 90% in recognizing point B, Ptm, Pog, Ba, L1\_C, U1\_C, and Gn. On the other hand, the success rates for N, Go, S, L1, and PNS were under 80%, with the lowest success rate (77%) for N (Table 3).

Among the cephalograms employed for the system's performance test, 12% had no misidentified points, whereas 75% had fewer than three misidentified points, with a maximum of six.

#### DISCUSSION

For humans, recalling past memory (experiences) in immediate response to a sensory input is assumed to be the very basis of recognition. Based on this postulate, a psychologic brain model employed in the present study was developed in which an image is compressed into a PPED vector so that the system performs a search for the most similar vector stored in the system as template vectors.<sup>1,18,19</sup> Table 4 gives an overview and comparison of earlier systems for automatic landmark recognition on cephalograms with the present system.

To evaluate the accuracy of the landmark identification provided by the systems—in other words, whether the system's definition of a landmark position is clinically acceptable—has been a critical issue in testing the performance reliability of such systems. Three major methods for such an evaluation have been employed so far. In the first method, an individual

Table 4. Overview of neports Documenting Systems for Automatic Cephalometr	Table 4.	Overview of Reports	Documenting S	Systems for	Automatic Ce	phalometry
--	----------	---------------------	---------------	-------------	--------------	------------

				No. of				
				Cephs for	Cephs for	r		
Author		Techniqueª	No. of Landmarks	Development/ Test	Learning ≠ Test	Picture Size and Resolution	Evaluation Method	Evaluation Result
Levy-Mandel et al1	I	Edge tracking	36	<sup>b</sup> /1	b	$256 \times 256$ pixels (1 pixel = 0.6 mm <sup>c</sup> )	А	64 <sup>d</sup>
Parthasarathy et al <sup>2</sup>	Ι	Edge tracking	9	⊳/5	b	$480 \times 512$ pixels (1 pixel = 0.3 mm <sup>c</sup> )	B <3 pixels	83 <sup>d</sup>
Cardillo et al12	+	Pattern matching	20	<sup>b</sup> /40	$\neq$	$490 \times 512$ pixels (1 pixel = 0.3 mm <sup>c</sup> )	$\rm B<2~mm$	75.8% <sup>d</sup>
Forsyth et al <sup>3</sup>	I	Edge location	19	<sup>b</sup> /10	b	$512 \times 512$ pixels (1 pixel = 0.3 mm)	$B <\!\! 1 \text{ mm}$	<b>73%</b> <sup>d</sup>
Rudolph et al6	II	Spatial spectroscopy	15	13/14	≠	$64 \times 64$ pixels (1 pixel = 2.3 mm <sup>c</sup> )	С	$3.07\pm3.09~\text{mm}$
Ren et al⁴	I	Image layer and edge tracking	b	<sup>b</sup> /10	b	200 dpi (1 pixel = 0.17 mm)	$B <\!\! 1 \text{ mm}$	43 landmarks
Liu et al14	I+II	Edge tracking Self- learning	13	<sup>b</sup> /10	b	b	С	$2.86\pm1.24~\text{mm}$
Hutton et al7	II	Active shape model	16	62/63	¥	$750 \times 950$ pixels (1 pixel = 0.2 mm <sup>c</sup> )	B < 2  mm C	35% <sup>d</sup> ; 4.02 ± 3.74 mm
Grau et al <sup>9</sup>	I+III	Line detection and pat- tern matching	17	20/20	$\neq$	$500 \times 500$ pixels (1 pixel = 0.3 mm)	$\rm B<2~mm$	88.6% <sup>d</sup>
Yang et al13	I+II	Edge detecting Wave- let transform Canny filter	17	Þ/12	b	, , , , , , , , , , , , , , , , , , ,	D	Over 80% <sup>d</sup>
Ciesielski et al11	111	Genetic programming	4	73–74/36	$\neq$	1 pixel = 0.4 mm	B <2 mm	85% <sup>d</sup>
El-Feghi et al10	111	Neuro-fuzzy system	20	200/b	b	1 pixel = 0.33 mm	B <6 pixels	90% <sup>d</sup>
Romaniuk et al8	II	Statistical patter recog- nition	1	40/ <sup>b</sup>	b	b	С	1.2 mm
Giordano et al15	I+III	Cellular neural net- works	8	<sup>b</sup> /26	$\neq$	b	B < 2 mm	85% <sup>d</sup>
Saad et al⁵	II	Active appearance model	18	20/7	$\neq$	100 dpi (1 pixel = 0.25 mm)	С	3.12 mm
Yue et al16	+  +	Active shape model	12	200/79	$\neq$	b , , , , , , , , , , , , , , , , , , ,	B <2 mm	<b>71%</b> <sup>d</sup>
Tanikawa et al (present study)	III	PPED vector	20	400/65	¥	$\begin{array}{l} \text{2320}\times\text{2960 pixels} \\ \text{(1 pixel}=0.085 \text{ mm)} \end{array}$	E	88% <sup>d</sup> (77%–100%)

<sup>a</sup> I indicates knowledge-based algorithm with edge tracking; II, model-based algorithm; III, pattern-matching algorithm.

A indicates orthodontist's visual judgment; B, when the distance between the position located by the system and that by 2 orthodontists were less than a certain amount, it was assigned to success; C, the distance between the position located by the system and that by an orthodontist/orthodontists; D, if the error for linear measurements is less than 2.0 mm, or if the error for angular measurements is less than 2.0 deg, it was assigned to success; E, when a system-identified point was located within a confidence limit of  $\alpha = .01$ , it was assigned to success.

<sup>b</sup> No description; <sup>c</sup> Estimated value given that the size of cephalogram is  $150 \times 150 \text{ mm}^{5.9}$ ; <sup>d</sup> Success rate.

orthodontist makes a visual judgment as to whether or not the system's recommendation is acceptable.<sup>1</sup> This method, which is also used for evaluating whether an orthodontic resident's reading of a cephalogram is acceptable, is limited to the extent that it inevitably incorporates intrajudge and interjudge variations. The second approach involves describing mean recognition errors, ie, the mean distance between the point provided by an orthodontist(s) and the point determined by the system.<sup>6–8,14</sup> The third method is to examine whether the system-identified landmark is located in a circle with a 2-mm radius.<sup>2,3,7,9–12,15,16</sup> This approach is meaningful in the sense that it provides an objective judgment as to whether or not the system's recommendation is correct, but it leaves room for argument as to whether it is reasonable to apply a circle with a 2-mm radius to all cephalometric landmarks, given that such landmarks are located in varying anatomic structures. In the present study, confidence ellipses were developed from the scattergrams that represent topographic variations in experts' subjective judgments of cephalometric landmarks, and these were employed to assess the system performance as the landmark-dependent criteria.

The sizes of the confidence ellipses for sella (the center of a pseudo-ellipse) and nasion (a junction between two bones) were relatively smaller because of their simple definitions and the good contrast found around the imaged landmarks. The large vertical variation in the distributions for Ptm is well-known empir-

	Liu et al14	Hutton et al <sup>7</sup>	Saad et al⁵	Tanikawa et al
Landmark	n = 10	n = 62	tn = 7	n = 65
S	$0.94 \pm 0.54$	$5.5\pm6.8$	$3.24 \pm 2.85$	2.10 ± 2.52
N	$2.32 \pm 1.14$	$5.6\pm3.9$	$2.97 \pm 1.85$	1.70 ± 1.18
Or	5.28 ± 4.10	$5.5\pm3.4$	$3.42 \pm 2.43$	$2.24 \pm 1.35$
Po	$2.43 \pm 2.10$	$7.3\pm6.5$	$3.48 \pm 2.46$	$3.63 \pm 2.53$
ANS	2.9 ± 1.12	3.8 ± 2.2	$2.70 \pm 1.05$	2.32 ± 1.34
Point A	4.29 ± 1.56	$3.3 \pm 2.4$	$2.55 \pm 0.97$	$2.13 \pm 1.36$
Point B	$3.69 \pm 1.55$	$2.6 \pm 2.7$	2.23 ± 1.24	$3.12 \pm 9.46$
Pog	2.53 ± 1.12	$2.7 \pm 3.4$	$3.66 \pm 1.74$	1.91 ± 1.42
Me	$1.9 \pm 0.57$	$2.7\pm3.6$	4.41 ± 2.03	$1.59 \pm 1.07$
Gn	$1.74 \pm 0.86$	$2.7 \pm 3.4$	4.22 ± 1.78	$1.45 \pm 0.84$
Go	4.53 ± 3.13	$5.8\pm6.0$	$3.64 \pm 1.76$	5.06 ± 3.41
U1	2.36 ± 2.01	$2.9\pm3.8$	$3.65 \pm 1.59$	1.78 ± 2.29
L1	2.29 ± 1.00	3.1 ± 2.3	$3.15 \pm 2.30$	1.81 ± 1.66
Total mean error	2.86	4.1	3.33	2.37

 Table 5.
 Mean Recognition Errors (mm) and Standard Deviations for Each Landmark

ically but was tolerable because in cephalometric analysis only the horizontal position of Ptm is used to determine the position of the posterior limit of the maxilla. By the same token, the greater horizontal variances in the distributions for ANS and PNS were tolerable because in cephalometric analysis only the vertical positions of ANS and PNS are used to determine the orientation of the palatal plane. In light of the fact that these factors are likely to cause errors and variations, we employed confidence ellipses for assessing whether the determinations of anatomic landmarks on cephalograms were correctly made by the automatic recognition system. The allowable determination domain has been expressed as clinically "correct" by specialists in terms of statistical probability, and the fiducial zones established were therefore considered to be suitable for assessing the performance of the automatic recognition system.

Identification of sella, nasion, gonion, L1, and PNS achieved success rates in the range of 77% to 80%. The relatively lower success rates for sella and nasion were associated with narrow confidence ellipses with distances of 1.5 mm and 1.4 mm, respectively, in the semiminor axis direction. The smaller zoning made the judgment more stringent, resulting in relatively lower rates of successful recommendations by the system compared with other anatomic landmarks. The landmark PNS, misidentified by the system, was located within the confidence ellipse horizontally but fell out of the fiducial zone vertically in most cases. The zone considered "correct" for PNS was as short as 1.6 mm in terms of the length of the semiminor axis, whose direction is almost vertical; this appeared to lower the rate of successful identifications by the system.

In a case involving excessive negative overjet, L1, the lower incisor tip, was incorrectly recognized as positioned posterior to the upper incisor tip. In a case involving reduced overbite, L1 was placed inferior to the fiducial zone. The relatively lower success rates for L1 were caused by great variation in the spatial relationship between the upper and lower incisors. Similarly, there was great variation in where the system located gonion because of its position in an obscured image area with complicated overlapping of the pharyngeal region, the cervical vertebrae, and the mandibular angle area. The lower rates of successful recognition of these anatomic landmarks may be explained by the presumption that the cephalograms used for testing the system's performance did not employ records of patients whose dentoskeletal/soft tissue relationships were similar to those seen in the record set employed for the system's learning. In a pilot study with sella, nasion, and orbitale, landmarks that have fewer dentoskeletal/soft tissue relationship patterns, 400 cephalograms were considered to be the optimum number for teaching the system. With regard to gonion, L1, and PNS, which have great variation in their locations, however, recognition performance would be improved by increasing the amount of image data available for building the system's knowledge base.

Table 5 gives mean recognition errors in the present study, ie, the distances between the coordinate values determined by our system and the corresponding fiducial coordinates, and a comparison with three previous studies that reported the mean error values for 13 major landmarks (sella, nasion, orbitale, porion, ANS, point A, point B, pogonion, menton, gnathion, gonion, U1, and L1).<sup>5,7,14</sup> The total mean error of landmark identification in the present study was the smallest, and it was similar to that reported by Liu et al.<sup>14</sup> As for sella and porion, the errors in the present study were larger than those reported by Liu et al. Our recognition mean errors for orbitale and point A, however, were significantly smaller than those reported by Liu et al. In addition, our system showed the smallest

mean errors for nasion, orbitale, ANS, point A, pogonion, menton, gnathion, U1, and L1 compared with the three previous reports.<sup>5,7,14</sup> Since our system showed a relatively lower success rate of 78% for sella, it is supposed that the use of the edge-based technique, which was employed by Liu et al, might improve the performance of our system. As for gonion, which also showed relatively lower success rate of 78%, the error reported by Saad et al<sup>5</sup> was smaller than that in the present study. This result implies that the performance of our system for gonion could be improved by using the model-based technique.

Finally, it should be noted that the system correctly recognized the positions of all 20 anatomic landmarks in 12% of the cephalograms used in the test. The system incorrectly identified the positions of three or fewer landmarks in 75% of the cephalograms tested. The maximum number of misidentified landmark positions per cephalogram was six, which occurred in just 4% of the cephalograms used for the test. In addition, in the present study, all anatomic features were correctly recognized. This means accuracy of performance is guaranteed if the current system is given the task of automatically extracting images near to the landmark of an input cephalogram, magnifying them, and displaying them. Thus, the development of an interface that automatically searches anatomic features using our system could help reduce the workload in clinical practice as well as an increase educational efficiency for orthodontic residents. In summary, the results of the performance test obtained in the present study suggest that the proposed system is effective and has potential for possible clinical application.

#### CONCLUSIONS

- The fiducial zones established by the panel of experienced orthodontists are considered valid for evaluation of the ability of the automatic recognition system to recognize anatomic features.
- With the incorporation of the rational assessment criteria provided by confidence ellipses, the proposed system was confirmed to be reliable. The system successfully recognized anatomic features surrounding all the landmarks. The mean success rate for identifying the landmark positions was 88% with a range of 77% to 100%.

#### ACKNOWLEDGMENTS

This study was partially supported by the Ministry of Education, Science, Sports, and Culture under Grant-in-Aid for Scientific Research (No. 14370695, No. 14370696, No. 16791284), by Japan Science and Technology Agency (JST) in the program of Core Research for Evolution of Science and Technology (CREST) and by the 21st-century Centers of Excellence (COE) program "Creation of Frontier Bio-dentistry." The authors are also deeply indebted to Dr Nellie Kremenak for meaningful and productive advice in manuscript preparation.

#### REFERENCES

- Levy-Mandel AD, Venetsanopoulos AN, Tsotsos JK. Knowledge-based landmarking of cephalograms. *Comput Biomed Res.* 1986;282–309.
- Parthasarathy S, Nugent ST, Gregson PG, Fay DF. Automatic landmarking of cephalograms. *Comput Biomed Res.* 1989;22:248–269.
- Forsyth DB, Shaw WC, Richmond S, Roberts CT. Digital imaging of cephalometric radiographs, part 2: image quality. *Angle Orthod.* 1996;66:43–50.
- Ren J, Liu D, Feng D, Shao J, Zhao R, Liao Y, Lin Z. A knowledge-based automatic cephalometric analysis method. In: Chang HK, Zhang YT, eds. *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Piscataway, NJ: IEEE Service Center,* 1998;2:723–727.
- Saad AA, El-Bialy A, Kandil AH, Sayed AA. Automatic cephalometric analysis using active appearance model and simulated annealing. *ICGST Int J on Graphics, Vision and Image Processing, Special Issue on Image Retrieval and Representation.* 2006;6:51–67.
- Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop.* 1998;113:173– 179.
- Hutton TJ, Cunningham S, Hammond P. An evaluation of active shape models for the automatic identification of cephalometric landmarks. *Eur J Orthod.* 2000;22:499–508.
- Romaniuk B, Desvignes M, Revenu M, Deshayes MJ. Shape variability and spatial relationships modeling in statistical patter recognition. *Pattern Recogn Lett.* 2004;25: 239–247.
- Grau V, Alcaniz M, Juan MC, Monserra C, Knoll C. Automatic localization of cephalometric landmarks. *J Biomed Inform.* 2001;34:146–156.
- El-Feghi I, Sid-Ahmed MA, Ahmadi M. Automatic localization of craniofacial landmarks for assisted cephalometry. *Proc Int Symp Circuits and Systems*. 2003;3:630–633.
- Ciesielski V, Innes A, Mamutil J. Genetic programming for landmark detection in cephalometric radiology images. Int J Knowl Base Intell Eng Syst. 2003;7:164–171.
- Cardillo J, Sid-Ahmed MA. An image processing system for locating craniofacial landmarks. *IEEE Trans Med Imaging*. 1994;13:275–289.
- Yang J, Ling X, Lu Y, Wei M, Ding G. Cephalometric image analysis and measurement for orthognathic surgery. *Med Bio Eng Comput.* 2001;39:279–284.
- Liu JK, Chen YT, Cheng KS. Accuracy of computerized automatic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop.* 2000;118:535–540.
- Giordano D, Leonardi R, Maiorana F, Cristaldi G, Distefano ML. Automatic landmarking of cephalograms by cellular neural networks. *Lect Notes Artif Int.* 2005;3581:342–352.
- Yue W, Yin D, Wang G, Xu T. Automated 2-D cephalometric analysis on X-ray images by a model-based approach. *IEEE Trans Biomedical Eng.* 2006;53:1615–1623.
- Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis. *Angle Orthod.* 2008;78: 145–151.
- 18. Shibata T. Intelligent VLSI systems based on a psychological brain model. *Proc IEEE Int Sym on* Intelligent *Signal*

Processing and Systems (ISPACS), 2000 Nov 5–8, Honolulu, Hi, USA. 2000:323–332.

- Yagi M, Shibata T. An image representation algorithm compatible to neural-associative-processor-based hardware recognition systems. *IEEE Trans Neural Netw.* 2003;14(5): 1144–1161.
- Takada K, Yagi M, Tanikawa C. Dependable systems for decisions in clinical dentistry- how electronics implements what dentists know. Proc Int Sym on Advanced Electronics for Future Generations, "Secure-Life Electronics" for Quality Life and Society. 2005 Oct 11–12, Tokyo, Japan. 2005:15– 21.
- Baumrind S, Frantz RC. The reliability of head film measurements. 1. Landmark identification. *Am J Orthod.* 1971; 60:111–127.
- 22. Fujita K, Takada K, QianRong G, Shibata T. Patterning of human dental arch wire blanks using a vector quantization algorithm. *Angle Orthod.* 2002;72:285–294.
- 23. Raymond T. *Atlas of Orthodontic Principles.* St Louis, MO: Mosby; 1970:279–289.
- 24. Enlow D. *The Human Face: An Account of the Postnatal Growth and Development of the Craniofacial Skeleton.* New York, NY: Harper & Row; 1968:277–283.
- Broadbent BH. Bolton Standards of Dentofacial Development Growth. St. Louis, MO: Mosby; 1975:133–135.
- Walpole RE, Myres RH. Probability and Statistics for Engineers and Scientists. 2nd ed. New York, NY: Macmillan; 1978.

#### **APPENDIX**

Ten lateral cephalograms were selected at random from the learning samples. A panel of 10 dentists (7 men and 4 women; age range, 30 years 1 month to 41 years 2 months), with clinical careers in orthodontics of more than 5 years, was selected. Each film was overlaid with acetate paper and anatomic contours traced in pencil by each panel member, who visually identified and marked 20 anatomic landmarks. All procedures were performed on a light box with identical writing tools in an air-conditioned and darkened room.

Given that the coordinate values determined by a judge *e* for a cephalometric landmark *m* on a cephalogram *c* are designated as V(m, c, e) [c = 1, 2, ..., 10, e = 1, 2, ..., 10] where m = 1(S), 2(N), ..., 20(Ptm), the following computations were made to generate a scattergram SG(*m*).

Step 1. The mean coordinate value V(m, c, \*) was calculated by equation (2) and defined as the 'best estimate' position data of landmark *m* on the cephalogram *c*.

$$V(m, c, *) = \frac{1}{n} \sum_{e=1}^{n} V(m, c, e)$$
(2)

Step 2. A new coordinate system V' was developed, where V(m, c, \*) was the origin, and a line through the origin and parallel to the line S-N connecting the mean coordinate values V(1, c, \*) and V(2, c, \*) for S and N, respectively, was chosen as the *x*-axis. The *y*-axis was a line perpendicular to the *x*-axis through the origin.

Step 3. V(m, c, e), a set of the coordinate values, was projected onto the coordinate system V' for each landmark *m* to generate a scattergram SG(*m*) consisting of 100 points.