active appearance models, landmarks,
x-ray images, skull, mandible

Przemysław KOWALSKI, Agnieszka TOMAKA*

USING ACTIVE MODELS FOR FINDING LANDMARKS

The paper presents the results of the experiments in using the active appearance models (AAM) to finding landmarks in the x-ray images of skull (especially of mandible). The landmark pointing is time-consuming in orthodontic analysis. It is useful to initially point the landmarks automatically, to help a physician in his or her work. Because the landmark are pointed not only by local features, but also by geometrical structure of skull, the model have to represent not only the local features (pixels in the image), but also the geometrical constraints. The one of many possible approaches is the active appearance model. The active appearance model was developed from the active shape models, and consists of the model of the shape and the model of appearance (texture). The model is active, that means the model fits to the object in the image. The main principle of the active models is to use two energies – internal energy represents limits of the models (shape, contour, texture), and the external energy represents the fitting the model to the image.

1. THE ACTIVE APPEARANCE MODEL

The active appearance models were developed by Gareth Edwards. The model can be decomposed into two models [2]:

– shape model (**x**);
– texture model (**g**).

The both models are constructed as the average of the shape/texture represented in the training set of objects:

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Q}_s \mathbf{c}$$
$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{Q}_g \mathbf{c}$$

Where **x** represents shape, using the point distribution model:

$$\mathbf{x} = (x_1, y_1, x_2, y_2, ..., x_i, y_i, ...x_n, y_n)^T,$$

Vector **g** represents the texture (normalised values of pixels samples); $\bar{\mathbf{x}}$ and $\bar{\mathbf{g}}$ are average values of (respectively) shape and texture. The **Q** represents the matrix of possible deformations of shape and texture. Vector **c** represents parameter of model. The example of the model is presented on the fig. 2.

* The Institute of Theoretical and Applied Informatics PAS ul. Bałtycka 5, 44-100 Gliwice

1.1. CREATING OF THE ACTIVE APPEARANCE MODEL

For real applications we try to remove correlation of shape and texture parameters:

1) We create covariance matrix:

$$\mathbf{S} = \frac{1}{m}\sum_{i=1}^{m} d\mathbf{x}_i \, d\mathbf{x}_i^T$$

2) Using principal components analysis (PCA) to matrix **S** we find eigenvectors ($\boldsymbol{p}_i$), and eigenvalues λ_i. We use the t maximal eigenvalues, such that the sum of them is bigger than assumed range of variability of the model:

$$\sum_{i=1}^{t} \lambda_i \geq f_v \sum_{i=1}^{2n} \lambda_i$$

3) During the learning phase, both parts of the **Q** matrix are joined, to remove correlation between shape and texture model parameters, and the PCA is used one more time (with repeat of the second step).

1.2. FITTING MODEL TO AN IMAGE

The next problem is how to fit the model to the object in the image. The active appearance model is a generative model, which means, it generates the artificial image, using parameters. The error of fitting is calculated using the difference between image, and texture of model.

$$E(x) = \overline{g}(x) + \sum_{i=1}^{t} \mathbf{c}_{gi} \mathbf{Q}_{gi}(x) - I(\mathbf{W}(x; c_s))$$

The error image is calculated as the difference of the values for all pixel samples in the model (the average texture and the sum of its deformations for parameters) and the appropriate pixels of the image (using warp function $\mathbf{W}(\mathbf{x}; \mathbf{c}_s)$ – warping the position of the pixel x onto image, using parameter $\mathbf{c}_s$).

There are some algorithms used to fitting AAM [5]. The basic algorithm is gradient descend, and its variant to image alignment – Lucas-Kanade algorithm. The Lucas-Kanade algorithm seeks locally "best" alignment by minimizing the mean square error (using *E(x)*).

The active appearance models need previous initialisation. It is not an important problem for the object tracking in the sequence of images (assuming the suitable distance between positions of tracked object in the consecutive images). But in the issue of finding single object in the image, the local minima can be an important problem (see fig. 4). To minimise the problem, there is used a multilayer model fitting – we start from the most general layer (with the least precision) and use the most precise model at the end.

2. THE ISSUE – LANDMARKS IN ORTHODONCY

In the orthodontic diagnostic often are examined shapes of the skull, and the skull deformation [7]. There are many approaches used to deformation analysis. Many of them bases on the set of landmarks. (A dozen of landmarks for the each approach.) The landmarks vary for the different techniques. There are a set of 48 most important cephalometric landmarks (hard tissue) for analysis of the side x-ray view of the skull [1].

There are some computer applications used by physicians. Some of the applications use 3D point position (obtained from the two x-ray images from different projections [7]). There are still no applications, making possible automatic landmark pointing, although the some attempts have been made [3, 4]. We have no ambitions for the all-automatic landmark pointing, and only want to assist physician in pointing. The physician has the "last word" in such problems, although the pointing by physicians also are characterised by errors [6].

3. EXPERIMENTS

In the experiments we use the set of 209 x-ray images, with landmarks pointed by a physician. On the each image are pointed 27 landmarks.

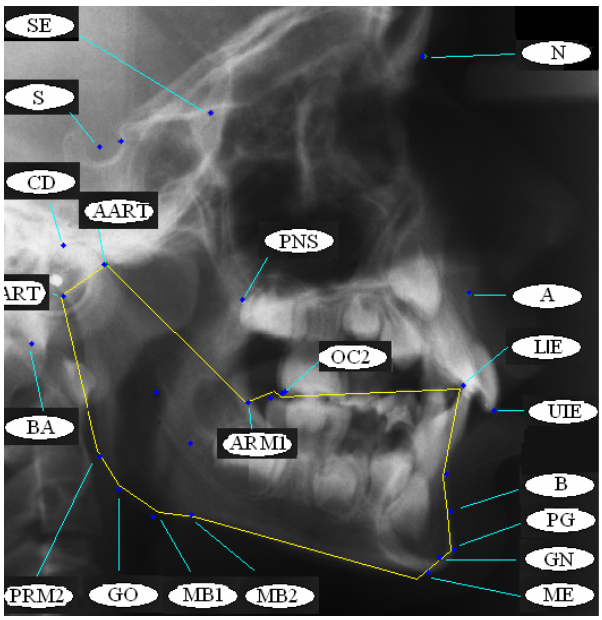

Fig. 1. Landmarks on x-ray image.

The used landmarks are (see fig. 1, [1]):

N – *nasion* – the most anterior point in the front nasal suture.

SE – *spheroidale (spheno-ethomoidal registration point)* – the point of intersection between the greater wings of the sphenoid and the anterior cranial base.

S – *sella* – centre of sella turcica.

CD – *condylion* – the most superior point on the condyle of the mandible.

ART – *articular* – the point of intersection between the posterior border of the mandibular condyle and the lower border of the cranial base.

AART – *anterior articulare* – the point of intersection between the anterior border of the mandibular condyle and lower border of the cranial base.
A – *A-point* – the deepest point on the concave outline of the upper labial alveolar process extending from the anterior nasal spine to prosthion.
PNS – *posterior nasal spine* – the tip of the posterior nasal spine.
OC2 – *the midpoint in the occlusal* space distal to the upper and lower first molars to provide a posterior point for the functional occlusal plane.
ARM1 – *anterior ramus point 1* – the point of intersection, on the anterior ramus vertivally, with an extension posteriorly of the occlusal plane.
BA – *basion* – the lowest point on the anterior border of the foramen magnum.
PRM2 – *posterior ramus point 2* – the most prominent posterio-superior point at the angle of the mandible on the posterior ramus behind gonion.
GO – *gonion* – the midpoint at the angle of the mandible.
MB1 – *mandibular* base point 1 – the most inferior point on the lower border of the mandible behind the antigonial notch and in front of gonion.
MB2 – *mandibular* base point 2 – the highest point on the antigonial notch.
ME – *menton* – the lowest point on the lower border of the mandibular symphysis.
GN – *gnathion* – the most antero-inferior point on the mandibular symphasis.
PG – *pogonion* – the most anterior point on the mandibular symphysis.
B – *B-point* – the deepest point on the bony curvature between the crest of the alveolus (infradentale) and pogonion.
UIE – *upper incisor edge* – the tip of the most prominent upper incisor crown.
LIE – *lower incisor edge* – the tip of the most prominent lower incisor crown.
And also some additional landmarks.

3.1. THE EXPERIMENT FOR ALL IMAGES

In the first experiment[1] we have tested active appearance model for finding landmarks using all the accessible images. The images are divided into two sets: training set (containing 141 images), and out-of-sample testing set (containing 68 images). The images differ in quality, skeleton age, etc.

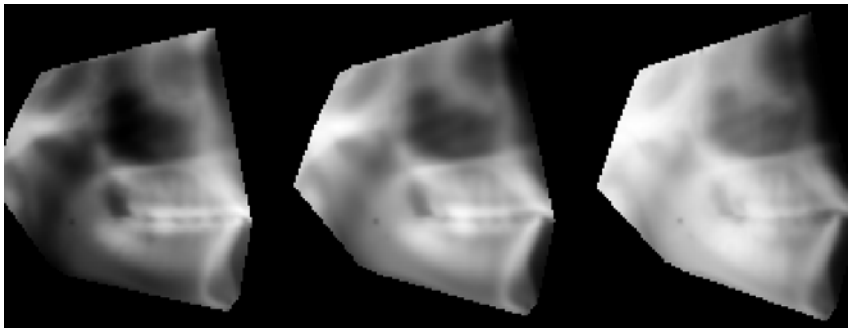

Fig. 2. The active appearance model for x-ray images of skull – in the middle: mean shape and texture.

[1] The used software was developed by Tim Cootes from University of Manchester (http://www.isbe.man.ac.uk/~bim/).

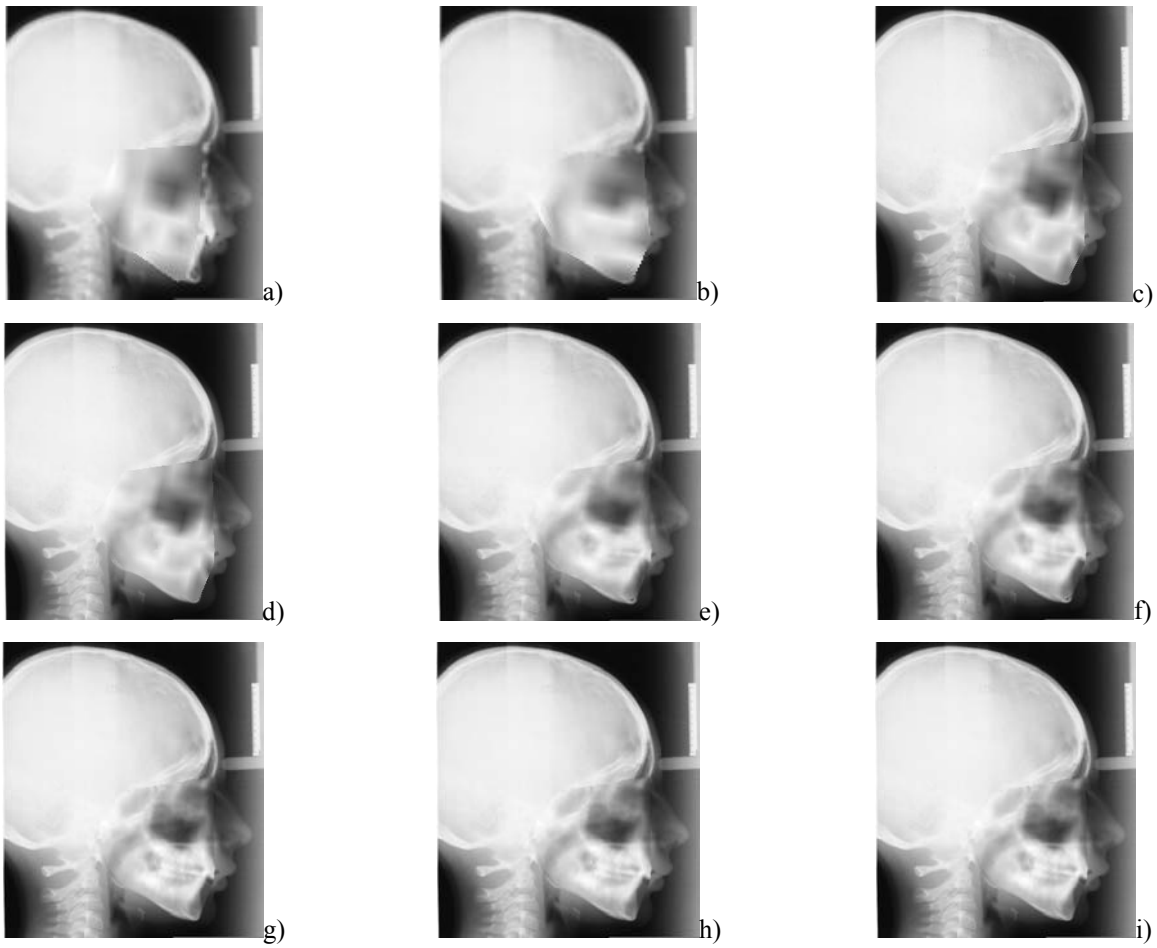

Fig. 3. Some stages of the model fitting – in the stages c) and e), level has changed.

The results (table 1) are slightly worse than in work of Liu, Chan and Chang [4]. Because of a few big mistakes in model fitting (which can be easily removed from the output), we decided to present in the table not only the mean error, but also the median error.

Table 1. Errors for landmarks in the experiment with the all images.

Landmark	Mean error [mm]	Median [mm]
N	3.58	2.62
SE	2.93	2.48
S	2.82	2.58
CD	2.60	2.47
AART	4.00	3.06
ART	2.65	2.20
BA	3.04	2.68
GO	4.08	3.42
PRM2	4.19	3.82
MB1	4.18	3.79
MB2	4.28	3.75
AMR1	4.42	3.45
PG	6.02	5.58
PN	3.66	3.25
ME	2.89	2.46
B	3.33	2.38
LIE	3.36	2.28
OC2	3.29	2.09
PNS	3.48	2.63
A	3.10	2.23
UIE	2.65	1.82

3.2. THE EXPERIMENT FOR SKULLS OF LIMITED BONE AGE

In the next experiments we tested active appearance model for finding landmarks using only the images of skull of one, or two skeleton ages. In the first experiments we use images of skeleton age 2 and 3. The results were even worse (table 2).

Table 2. Comparison of mean and median error for experiments.

Set of images	Mean error [mm]	Median [mm]	Number of samples
All	3.38	2.53	141 for training, and 68 for testing
Skeleton age 2	3.20	2.83	11 for training and 11 for testing
Skeleton age 3	3.56	2.78	19 for training and 19 for testing
Skeleton age 5 & 6	2.83	1.52	61 for training and 30 for testing

The most promising results are obtained for the set of images of mixed two skeleton ages – 5 and 6. It is caused, by the limited variability of the shape (for the group of skulls with limited to two, similar skeleton ages), and relatively high number of training samples (61). The results are for many landmarks better, than in work of Liu, Chan and Chang [4].

Table 3. Errors in the experiment with the images of skulls of skeleton age 5 and 6.

Landmark	Mean error [mm]	Median [mm]
N	3.68	1.74
SE	3.10	1.47
S	2.92	1.43
CD	3.21	1.90
AART	2.39	1.27
ART	2.50	1.31
BA	3.32	2.19
GO	3.29	2.50
PRM2	3.44	2.49
MB1	2.84	1.97
MB2	2.98	1.92
AMR1	2.10	1.47
PG	3.02	1.30
PN	2.95	1.24
ME	3.06	1.27
B	2.84	1.20
LIE	2.64	1.04
OC2	2.65	1.70
PNS	1.94	1.23
A	2.75	1.23
UIE	2.82	1.30

4. CONCLUSIONS

The active appearance models can be used for finding landmarks, but the results still does not confirm the expectations. There are some possibilities to improve quality of the active models fitting:

1) Do additional pre-processing for the x-ray images. (The quality of the images is varied.)
2) Creating model for one skeleton age using the high number of training samples.
3) Use more complex hierarchical fitting for the model – finding in the consecutive stages the skull, the mandible, and the regions of the landmarks.
4) Use not only the area of mandible, to find the landmarks. (Active appearance model uses only the area of convex hull of the mean shape model. Sometimes it results in ambiguity – see fig. 4, for an example of using the model of too small area.)
5) Mix the active appearance and active shape models, using different approach, to the different parts of the skull.

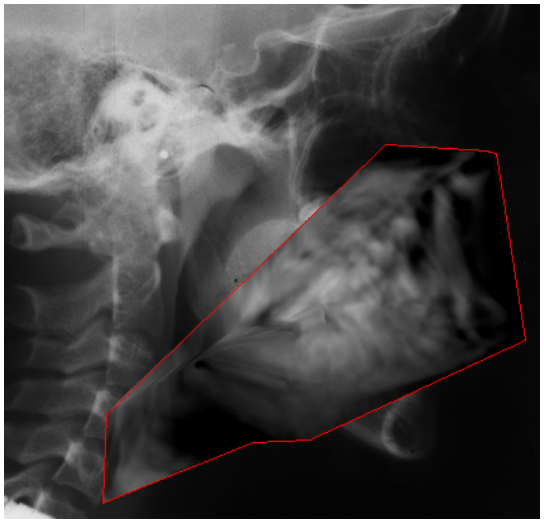

Fig 4. The ambiguity of fitting only the area of mandible results in wrong position of found "mandible". The found "mandible" is almost orthogonal, to the real one.

The more complex solution uses the 3D model of skull. In some images, the mandible deformation in 3D, gives the double landmarks. The only way to manage it, is to use 3D model. The 3D solution cannot use only the one image – the information has to be completed by the images of different projections, and (optionally) the other kinds of medical imagining sources – NMR, CT, etc.

BIBLIOGRAPHY

[1] BHATIA S.N., LEIGHTON B.C., A Manual of Facial Growth, Oxford University Press, 1993
[2] COOTES T.F., EDWARDS G.J., TAYLOR C.J., Active Appearance Models. Proc. European Conference on Computer Vision 1998, vol. 2, pp. 484-498, Springer, 1998.
[3] EL-FEGHI I., SID_AHMED M.A., AHMADI M., Automatic localization of craniofacial landmarks for assisted cephalometry, Pattern Recognition 37 (2004, pp. 609-621
[4] LIU J-K., CHEN Y-T., CHENG K-S., Accuracy of computerized automatic identyfication of cephalometric landmarks, American Journal of Orthodontics and Dentofacial Orthopedics, November 2000, pp. 535-540
[5] MATTHEWS I., BAKER S., Active Appearance Models Revisited. Tech. report CMU-RI-TR-03-02, Robotics Institute, Carnegie Mellon University, April, 2003
[6] MCCLURE S.R., SADOWSKY P.L., FERREIRA A., JACOBSON A., Reliability of Digital Versus Conventional Cephalometric Radiology: A Comparative Evaluation of Landmark Identyfication Error, Seminars in Orthodontics, pp. 98-110, Elsevier, 2005
[7] TOMAKA A., Detekcja i opis deformacji struktur na podstawie sekwencji obrazów cyfrowych. PhD Thesis, Gliwice, 2001