Dental age estimation from panoramic X-ray images using statistical models

Čular, Luka; Tomaić, Mia; Subašić, Marko; Šarić, Tea; Sajković, Viktorija; Vodanović, Marin

Dental age estimation from panoramic X-ray images using statistical models

Luka Čular, Mia Tomaić, Marko Subašić
Department of Electronic Systems and Information Processing
University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia

Tea Šarić, Viktorija Sajković, Marin Vodanović
Department of Dental Anthropology, School of Dental Medicine, University of Zagreb, University Hospital Centre Zagreb, Croatia

***Abstract*—This paper presents an application of computer vision methods to dental age estimation based on the lower third right molar in panoramic X-ray images. For this purpose, two statistical computer vision models are adjusted and applied: Active Shape Model and Active Appearance Model. Both models use shape and appearance of the object to find the outer contour, with the only difference being in the way appearance is used. Statistical models are used to extract features describing the selected tooth, and neural network is used to provide dental age estimation using the features as input. Our own dataset was created, consisting of panoramic X-ray images with known age. A manual segmentation of the selected tooth has been performed for each image in the training set, and the obtained outer contours were used to train both models. Promising preliminary results are presented.**

Keywords—*Active Shape Model, Active Appearance Model, dental X-ray images, age estimation*

I. INTRODUCTION

The problem of age estimation is common in forensics and law enforcement when age information is not available. Non-invasive methods for age estimation are more applicable, and thus, preferred. The most used and reliable methods for dental age estimation are based on X-ray image analysis. As every part of human body, teeth also have different stages of development, which are visible in X-ray images. A person's age can be estimated by estimating the development stages of a tooth and by linking of estimated stage to the most probable age. One such method is developed by Haavikko K. where sketches of teeth at various development stages and a table linking each stage to an age, are used. Due to inter-sex variations, different age estimates are provided for males and females [1]. This manual estimation of age is tedious and subjective, so an automated age estimation has a potential to increase accuracy and repeatability.

Human experts are making age estimates based on the shape and the appearance of teeth in panoramic X-ray images. These estimates are backed up by their knowledge of teeth and jaw anatomy. Our automatic age estimation is based on the same characteristics visible in panoramic X-ray images, and on statistics of those characteristics. We utilize statistical models of shape and appearance to capture knowledge of common shape and appearance variations from a training data set. The knowledge is used to find the exact shape of teeth in images. To describe and discriminate different teeth shapes and appearances, a descriptor vectors are used. Statistical models also enable convenient dimensionality reduction of descriptor vectors. We estimate age from descriptor vectors using neural networks.

The emergence of the third molars takes the last place in teeth dentition. By the time person is 24 years old, third molar root development is completed, so they are reliable age indicators for age estimate in the range from 10 to 24 years [2]. Hence our age estimates are based on the lower right third molar.

II. IMAGE DATASET

We use panoramic X-ray images as input data. More precisely, we use orthopantomograms (Fig 1.), which are collected by School of Dental Medicine, University of Zagreb. All personal information, except age and gender, are removed from orthopantomograms so there is no violation of privacy present. There are 203 orthopantomograms collected in a database with age range between 10 and 25 years. Images are stored in JPEG format.

The data set doesn't have all ages equally represented, so images are divided into two sets in 3:1 ratio. Every age is represented with ¾ of a total number of orthopantomograms in the training set and with ¼ of orthopantomograms in the test set. We choose this approach due to a limited data set and unbalanced age representation. To obtain segmentation ground truth and to enable model training, manual segmentation was performed on both sets using Gimp 2.8 (GNU *Image Manipulation Program*) [3]. All visible tooth tissues are annotated in one new graphic layer and two characteristic points on the tooth are saved in the additional layer. These points are needed to initialize models in new images, and teeth contours are used to build statistical models of shape and appearance. In the presented research, we focused only on the lower third right molar in panoramic X-ray images.

III. ACTIVE SHAPE MODEL (ASM)

The model was presented in 1995 by Cootes and Taylor [4]. This is a flexible model, which primary purpose is to learn

978-1-5090-4011-7/17/$31.00 ©

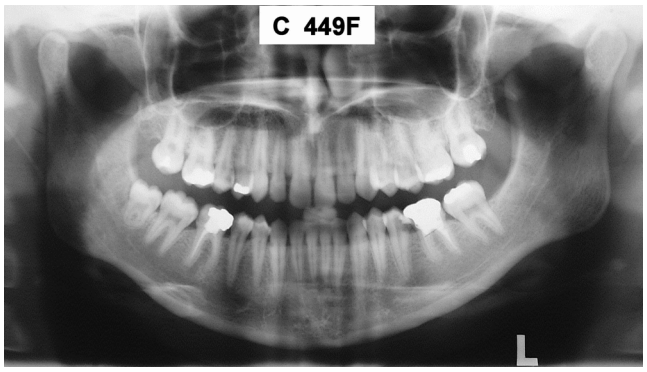

Fig. 1. Orthopantomogram

different shape variations so it can be used for automatic segmentation and object recognition in images. Initial applications were in medicine where ASM was used for accurate locating of the vertebrae on X-ray images of the cervical spine [5]. This model has also shown promising results in face recognition by determining not only gender but, also, person's identity in images, which is the most common use of ASM. The model is learning different shape variations by training on a large number of object contours, which contain mentioned variations. In our particular case, we are interested in variations of the lower third right molar contours. For this purpose, Principal Component Analysis (PCA) [6] is performed. Each component is a combination of contributions to the contour positions which explain additional part of the set variance, and by discarding components with least variance contribution, dimensionality reduction is achieved. PCA provides eigenvalues, eigenvectors and mean shape, and only a subset of eigenvectors, which explains variance to a greater extent, is kept.

Each contour is represented with the same number of equidistant points like demonstrated in Fig. 2, so the mean shape is determined by averaging corresponding points' coordinates from all contours in the training set. The mean shape is used as initial shape when finding contour in a new (test) image. ASM is iteratively trying to find the best fit for each contour point, but contour points are not independent. They all have to move together to new positions in some way, and eigenvectors encode the most probable joint variations of those positions. This allows ASM to take only probable (learned) shape variations in each iteration. New positions are determined by adding a linear combination of eigenvectors, multiplied by appropriate constants, to the mean shape. The constants represent shape parameters of the model, and by variating them, the shape is changed into a probable shape. Variations of shape parameters are allowed only within a range limited by their standard deviations, to constrain shape variations to probable ones. ASM uses pixel intensities to determine the best new position for every contour point. To find the best position, a statistical model of pixel profiles for each contour point is used. The profile lies on a line passing through a contour point in the normal direction.

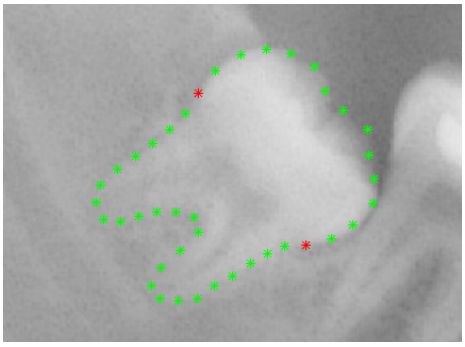

Fig. 2. Right third molar with marked contour points for ASM

We use 10 pixels on both sides of a contour point. Again, PCA is performed in order to determine the mean (initial) profile, eigenvectors, and eigenvalues for every contour point. A new point position is found as the best profile fit between the profile from the current image and statistical model of the profile.

A. Creating ASM

To build ASM, contour points of a tooth have to be provided. For each training image, we use manual segmentation results to obtain an equal number of contour points in a consistent order. 15 equidistant contour points are obtained between two marked points on tooth crown and 25 contour points that describe the root of the tooth in a clockwise direction (Fig 2.). All contour points' coordinates are transformed in the new coordinate system in order to achieve scale, translation and rotation invariance. The first and the fifteenth point of a contour are transformed into coordinates (0,0) and (1,0). Based on this two point pairs, a transformation matrix is determined and all contour points are transformed into the new coordinate system. Transformed contour points' coordinates are fed to PCA and we keep only most significant components that explain 97% of the variance. Models of pixel profiles perpendicular to each of the 40 contour points are also stored.

B. Applying ASM

The first step requires placing ASM model on the lower third right molar in orthopantomogram. The transformation matrix is determined based on two manually marked points on tooth crown and two corresponding model points (first and fifteenth contour point). Using this transformation matrix all contour points of the model's initial shape are placed in new (starting) positions. In order to find best contour position, several iterations have to be performed. Each of them includes three steps. Finding contour normals and textures along them, then comparing them with previously stored profile models are part of the first step. The best new position is determined within the profile. The second step includes points transformation back to model coordinate system and determining corresponding shape parameter combination. In the third step, the inverse transformation is performed and contour points are placed back in image coordinate system. Age estimation is performed using model parameters obtained in the last iteration.

IV. ACTIVE APPEARANCE MODEL (AAM)

AAM is a computer vision algorithm for matching a statistical model of object shape and appearance to a new image. It's similar to ASM, but it brings improvements by introducing whole object's texture into the model along with the shape model, which increases the amount of information that the model uses. It was created and presented in 1998 by Cootes et al. [7]. Although the model is primarily used for facial recognition, it is also widely used in interpretation of medical images [8]. Model of the shape is similar to ASM and is built from a fixed number of contour points. Their coordinates must be aligned for each image, and images have to be warped according to the mean shape so that each texture (appearance of the tooth) has the same shape and size. Average texture can then be calculated by averaging corresponding pixel intensities. This transformed texture is addressed as appearance. Pixel intensities are not independent, similar to contour points' coordinates, so texture eigenvectors encode the most probable joint variations of the appearance.

A. Creating AAM

For all the images in the training set with manually segmented third lower right molar, the mean shape is found in a slightly different way than with ASM. Here, the points lying on the contour of the third molar root are ignored because shape change there is rather small. The information contained in that part of the third molar will be included in the appearance model. Also, two points used to position the model are the most right bottom and most left bottom points as shown in Fig 3. Preliminary tests have shown the accuracy of the model is increased if the texture is expanded beyond the tooth contour so that it covers texture just outside the third molar. In this way, the appearance of the surrounding of the tooth will be included into the appearance model. In the particular case, the surrounding area of the tooth does not vary much, so it could provide additional help in guiding the tooth segmentation. To include the surrounding area, the shape was expanded radially from its center of mass by increasing the distance from the center of the mass for each contour point by 30%. Fig. 4. shows expanded model texture with blue line representing the third molar mean shape. The inclusion of texture beyond object's contour is not part of the original AAM setup, and we introduce this modification because it fits our particular problem.

The model, much like ASM, is built using a set of training images with manually annotated contours of the object of interest. The process begins by determining the mean shape of the object. But now the main goal is to also add texture into the model. The texture of each object from the training set is non-rigidly transformed to fit the mean shape so that PCA of texture can be performed. We obtain the mean appearance and principal texture variation components. We retain only most significant principal components that explain 99% of the texture variance. As there is obvious correlation between shape and appearance of the tooth, one can further reduce the number of parameters. One additional PCA is performed on concatenated shape and appearance parameter vectors. Again, most significant principal components that explain 99% of the joint variance are kept.

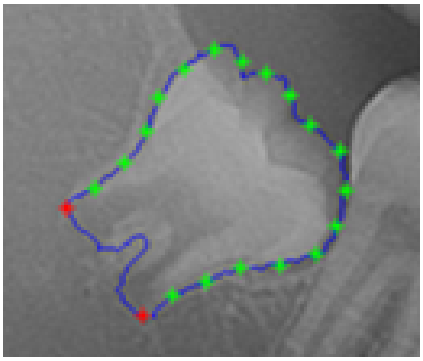

Fig. 3. Right third molar with marked contour points for AAM

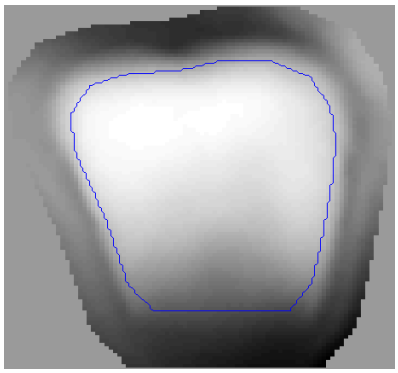

Fig. 4. Appearance model

B. Applying AAM

When mean shape and appearance are obtained, a search of a third molar in the new image can start. In order for the model to know the whereabouts of a third molar, two points are marked manually on the root of the third molar, much like during the training of the model.

The object search in the new image is done by placing the mean model of shape and appearance on the position where the object is located using two manually selected points. Texture located in the new image, under the positioned model, is compared to the current model's appearance. Based on the difference of the two, new joint parameters are calculated that should reduce that difference by defining a new shape and new appearance. Variations of the joint parameters are limited by their variance to constrain model variations to probable ones. This procedure is repeated a fixed number of times, after which we obtain the model parameters that are most appropriate for the given image.

V. ARTIFICIAL NEURAL NETWORK

Artificial neural network is a computer model made from a large number of interconnected elements called artificial neurons [9]. The parameters of each link affect the exchange of information among the neurons and are most often determined through the learning process. The neural network

type used in the proposed method is Radial Basis Network [10] which is a common choice for regression problems. It has one nonlinear hidden layer where all neurons are based on radial distance functions. The number of hidden neurons is equal to the number of samples in the training set (157 in our experiments), so the distances for a sample in the input layer are calculated from each sample from the training set. The output layer is linear and consists of one classical neurons, where the input values are multiplied by the corresponding weights and summed up. The weights are determined trough the training process by solving a system of equations whose solution is given in the training set. Our initial experiments include Multilayer Perceptron neural network [10] which provided similar results, but Radial Basis Network provided slightly better results.

VI. RESULTS

Three tests were performed in order to determine models' accuracy. Two tests compare shapes given by ASM and AAM models with shapes which are annotated manually to obtain an estimate of segmentation accuracy. The same model parameters that provide segmentation, are used as input for age estimation, so in the third test, we compare age estimation given by artificial neural network and the real age. Tests were performed on the prepared test data set. Number of iterations during AAM search is set to 30, while, for ASM, the number of iterations is set to 40. The number of iterations for AAM is lower due to it's greater complexity. Joint shape and appearance parameters of AAM were capped at ±1 standard deviation. For ASM, shape parameters were capped at ±3 standard deviation. The limits were determined experimentally, and the allowed range for AAM parameters is smaller to avoid too large shape variations.

A. Area segmentation error

We compare areas of the model's shape and manual segmentation results' shape. Pixels that are part of shape given as result of the model and are not part of manual segmentation, make false positive area. Other pixels that are part of manual segmentation and are not part of shape given by the model, make false negative area. Corresponding percentages are calculated according to equation (1) and (2). Pfp and Pfn are total numbers of pixels that are part of two areas and N is the total number of pixels that are part of the manually segmented area. FP stands for False Positive percentage and FN stands for False Negative percentage.

$$FP = \frac{Pfp}{N} * 100 \quad (1)$$

$$FN = \frac{Pfn}{N} * 100 \quad (2)$$

Table I. shows averaged results of ASM and AAM models on test images. A result is considered to be the best when the sum of percentages is the smallest. One can see that ASM is more likely to include surrounding of the tooth in the model area, while AAM is more likely to stay within the tooth and never reach the tooth contour.

Table I. Average false positive & false negative segmentation error

	ASM	**AAM**
False positive	24%	13%
False negative	13%	35%

B. Contour segmentation error

To determine the contour segmentation error (CSE), Euclidean distance between each contour point (obtained with the model) and the manually annotated contour is calculated according to equation (3).

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (3)$$

Here (x_1,y_1) are coordinates of a contour point of the model, and (x_2,y_2) are coordinates of the closest point on the annotated contour. Average ASM and AAM contour errors are presented in Table II. The average is calculated over all contour points over all images in the test data set. Results show that AAM sets model shape closer to the real shape than ASM does.

Table II. CSE - Average contour error in pixels

	ASM	**AAM**
Average Euclidean distance	11.58	4.76

C. Age estimation error

Age estimates are obtained by a Radial Basis Network. Though an optimal subset of the training set can be used to form radial functions, we used the complete training set. Thus, the hidden layer has 157 radial neurons. The transfer function of each hidden neuron is

$$e^{-dist^2} \quad (4)$$

where *dist* is the distance to one training sample.

Input for the neural network are parameters obtained by applying models on test images sets. Network inputs for ASM approach are eight shape parameters, while for AAM they could be shape parameters (seven of them), appearance parameters (104 of them) or joint shape and appearance parameters (111 of them). The best results are obtained when just shape parameters of the AAM were used, and that result is presented in Table III. Network outputs a real number which represents an age estimate in years. Comparing the neural network output with actual age determines neural network error. In Table III we provide average errors in years obtained after applying two models on test images. MAE stands for Mean Absolute Error in years. Results for both ASM and AAM are almost the same, with AAM being slightly better. The spread parameter of Radial Basis Network which regulates the area of influence of each hidden neuron is set to 1000 for ASM and for AAM.

Table III. Age estimation

	ASM	**AAM**
MAE	2.481 (std = 2.148)	2.283 (std = 2.168)

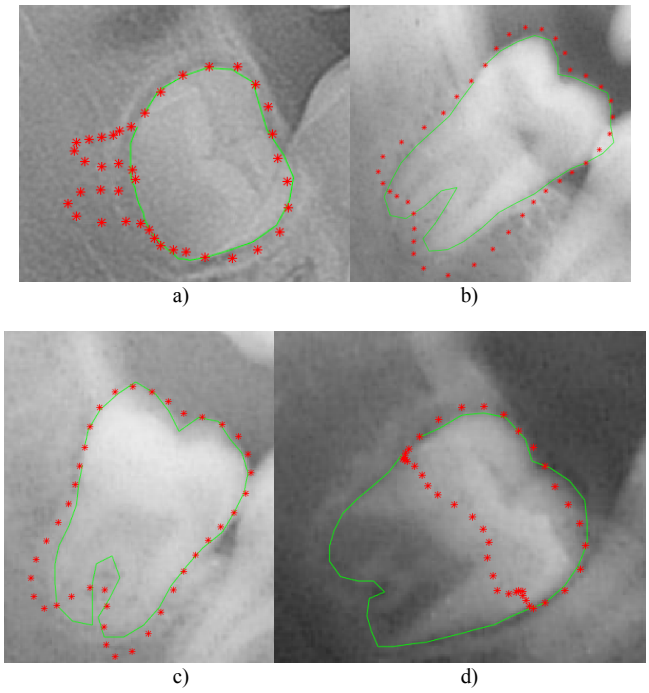

Fig. 5. Examples of results obtained by ASM approach

Fig. 5 provides several examples of ASM outputs. Red dots present contour points obtained in last ASM iteration, while green contour represents manual annotation. The described errors are given in Table IV.

Table IV. Sample results obtained by ASM

Fig. 5	FP	FN	CSE	Age error
a)	21%	3%	12.65 px	2.06 yr.
b)	27%	3%	7.09 px	4.61 yr.
c)	15%	4%	4.27 px	5.48 yr.
d)	2%	56%	14.47 px	0.22 yr.

Fig. 6 shows some segmentation results using AAM. Red dots represent model shape for the last iteration, while green contours represent manual annotation. Corresponding error estimates are provided in Table V.

Table V. Sample results obtained by AAM

Fig. 6	FP	FN	CSE	Age error
a)	14%	10%	1.36 px	0.77 yr
b)	3%	18%	3.38 px	0.81 yr
c)	5%	20%	1.40 px	0.64 yr
d)	287%	1%	24.14 px	3.97 yr

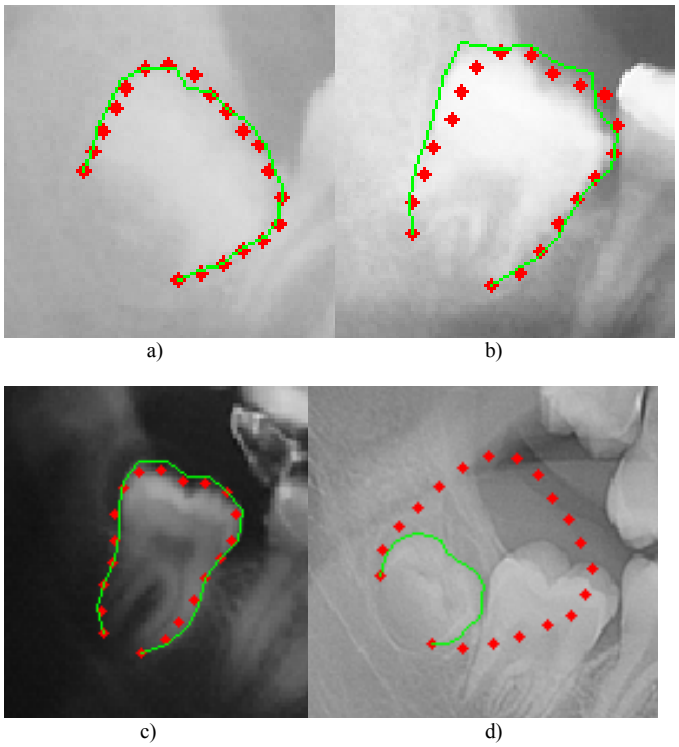

Fig. 6. Examples of results obtained by AAM approach

One can see that segmentation is newer perfect, and some time models miss greatly. It is reasonable to expect that better age estimates can be obtained when segmentation results are good, but the neural network age estimator should be insensitive to limited segmentation imperfections. It is interesting to notice that even when a segmentation error is great (Fig. 6. d), the neural network can sometimes still provide pretty close age estimate. Segmentation accuracy is influenced by image quality, but even in good quality images, tooth edges might be poorly visible. Statistical shape model can bridge moderate gaps in the contour, but such situations can mislead the model. Such poor conditions could even influence human experts and their manual segmentation which also influences the estimation of segmentation accuracy.

VII. CONCLUSION

The main advantage of the automated dental age evaluation is the increase in the speed of estimation, as well as high repeatability. The user input required for the proposed method is minimal and it doesn't require an experienced dentist. The average age estimation error that the proposed system achieves is less than three years. The obtained results for dental age estimation look promising and indicate that such system could be used for law enforcement and forensics, especially if estimation error could be reduced even further. It is interesting to notice that even though the segmentation errors are not insignificant, imperfect shape descriptors were able to produce good and consistent age estimates.

Observing the results, we notice that AAM performs only slightly better compared to ASM which is something that was expected, but it come at a cost of greater complexity. One surprising thing is that appearance parameters of AAM did not provide any advantage when compared to the shape parameters. In fact, when appearance parameters are used together with shape parameters, age estimation deteriorates. We expect that different model evolution strategies, like gradient descent, might change that and increase accuracy for both models. Also, future work will include an extension of the image data set, and comparison to manual age estimation.

REFERENCES

[1] K. Haavikko, "The formation and the alveolar and clinical eruption of the permanent teeth. An orthopantomographic study," *Suomen Hammaslaakariseuran toimituksia= Finska tandlakarsallskapets forhandlingar*, 1970; 66(3): 103.

[2] Brkic, H., Vodanovic, M., Dumancic, J., Lovric, Z., Cukovic-Bagic, I., Petrovecki, M.,2011. The chronology of third molar eruption in the Croatian population. Coll.Antropol. 35, 353–357.

[3] Gimp- GNU Image Manipulation Program [Internet]. 2017. Available at: *https://www.gimp.org/*

[4] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, "Active shape models-their training and application. Computer vision and image understanding", 1995;61(1): 38-59.

[5] A. Lanitis, C.J. Taylor, T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," IEEE Transactions on Pattern Analysis and machine intelligence, 1997;19(7):743-756.

[6] S. M. Holland, "Principal components analysis (PCA)," University of Georgia, 2008.

[7] G.J. Edwards, C.J. Taylor, T.F. Cootes. "Interpreting face images using active appearance models. Proceedings," Third IEEE International Conference on Automatic Face and Gesture Recognition 1998. 300 p.

[8] S.C. Mitchell, J.G. Bosch, B.P.F Lelieveldt, R.J. van der Geest, J.H.C. Reiber, M. Sonka, "3-d active appearance models: Segmentation of cardiac MR and ultrasound images," IEEE Trans. Med. Imaging 2001;21(9):1167–1178.

[9] K. Gurney " An Introduction to Neural Networks," London: Routledge

[10] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice Hall PTR, Upper Saddle River, NJ, 1998