AUTOMATIC SEGMENTATION OF CARPAL AREA BONES WITH RANDOM FOREST REGRESSION VOTING FOR ESTIMATING SKELETAL MATURITY IN INFANTS.

Steve A. Adeshina.

Nigerian Turkish Nile University
Abuja, FCT, Nigeria
steve.adeshina@ntnu.edu.ng

Claudia Lindner, Timothy F. Cootes.

Imaging Sciences and Biomedical Engineering
The University of Manchester, UK
t.cootes@manchester.ac.uk

ABSTRACT

We evaluate the utility of a model of the structure of the carpal area bones in the hand for predicting skeletal maturity in infants (0 -7 yrs) using a texture based statistical model of appearance. Skeletal maturity assessment is important for diagnosing and monitoring growth disorders. Statistical models of bone shape and appearance have been shown to be useful for estimating skeletal maturity. In this work we locate the carpal bones with an automatic system which uses a Constrained Local Model with Random Forest Regression Voting. Appearance models were built from the segmented images, and texture parameters extracted were used to estimate skeletal age. By analysing the performance on a data-set of 284 digitized radiographs of normal infants we show that an automatically segmented texture based appearance model of the carpal region produces very satisfactory skeletal age estimation results of a mean absolute error of $(0.43, 0.53)$ years, for male and female respectively.

Index Terms— Skeletal maturity assessment, Random Forest regression, Statistical Models of appearance, Carpal bones' segmentation.

1. INTRODUCTION

Skeletal maturity plays an important role in the diagnosis of growth and endocrine disorders. The two main methods examine the morphology of the bones and the joints of the non-dominant hand in a radiograph. A significant difference between the bone age and the actual age of a child is an indication of growth abnormalities. The predominant methods in clinical practice are those of Greulich and Pyle(GP) [11] and Tanner and Whitehouse(TW2/3) [18].

There have been many attempts to automate the bone age assessment procedure. These range from classical image analysis methods [17], machine learning techniques and model based methods [16]. Thodberg *et al.*[19] showed how Active Appearance Models [5] can be used to determine skeletal maturity. However the estimation was limited to 2 - 17 years and excluded the use of carpal bones.

The key issue here is that the most critical years during which corrective procedure can easily be carried out are excluded in maturity estimation. The reason may be due to the limited availability of images and the radiograph image quality at this early age. The Carpal bones, which are the most useful at this stage, are either in cartilage form or are just appearing as a pin point as shown in Figure 1 . Their order of appearance is not consistent.

Medical research showed that the Carpal bones are most useful in the estimation of skeletal maturity between the ages of 0 - 7 years [18] [11], however because of the difficulty in segmenting these bones and the fact of their lack of correspondence, most algorithms either use one or two bones that are consistent (ie the Capitate and the Hamate). Unfortunately the above statement on the the accuracy of estimating skeletal maturity from carpal bones is predicated on the consideration of all the available bones [18] [11]. This is why we differ from most existing skeletal maturity methodology that utilizes one or two of the bones to estimate skeletal maturity [20]. This lack of correspondence poses particular research problems to Statistical models. This work attempts to find an automatic way of of segmenting the the carpal area bones as well as build a statistical model of the area that does not suffer from lack of correspondence.

In our earlier work [3] we evaluated several structures in estimating skeletal maturity, but the work was limited to ages above 5 years. Further more in [4] we evaluated the use of Carpal bones for determination of skeletal maturity for infants, however all the carpal bones considered were segmented manually. This work builds on [4] to find an automatic way of segmenting the area around the Carpal bones and using the same to predict skeletal maturity of infants.

The closest work is that of Zhang *et al.* [20] where classical image processing techniques and fuzzy classification were used to estimate bone age from the Hamate and Capitate (see figure 1). We differ in our approach from Zhang *et al.* as we use a texture based statistical models of appearance to estimate skeletal maturity from an automatically segmented region around the Carpal bones for children of ages 0 -7 years. Recent work by Harmsen *et al.*[13] used Support vector machine Classification based on Correlation prototypes, however they excluded the use of Carpal bones. Though they recorded some fairly good results, this method could not have been effective for infants. Hark *et al.*[12] used a support vector regression methods smart class mapping, but the work also excluded the use of Carpal bones. Whereas Hsieh *et al.*[14] considered the carpal bones using a fuzzy-based growth model with principle component analysis, they had to use geometrical ratios and values and we recorded better results. Giodarno *et al.*[10] combined the use of both Carpal bone and other Epiphyseal/Metaphyseal region to estimate ages between 0-10 years, but they used extremely complex image processing methodology with very limited number of X-rays which may obviously be of good quality. Our results are comparable.

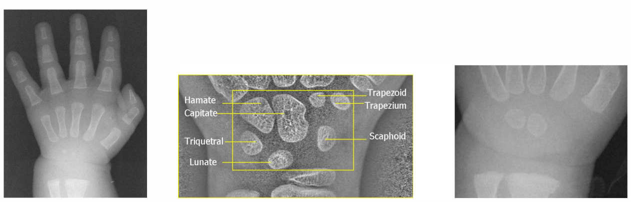
Random Forests (RF) [6] describe an ensemble of decision trees trained independently on a randomized selection of features. They have been shown to be effective in a range of classification and regression problems [6]. Gall and Lempitsky show in Hough Forests [9] that objects can be accurately located using RF regressors to predict the position of a point relative to the sampled region, then running the regressors over the region and accumulating votes for the likely position. Cootes *et al.*[6] show how this method (RF) can be combined with Statistical Shape model to accurately segment a va-

978-1-4799-4106-3/14/$31.00 © 2014 IEEE

riety of complex dataset. Lindner *et al.*[15] using the methods of [6], applied RF regression in a Constrained Local Model(CLM) to accurately segment the femur in a pelvic radiograph. Following the approach of [15] we apply the method of Cootes *et al.*to segment the Carpal areas of hand radiograph by applying RF regression in a Constrained Local Model (CLM) framework to vote for optimal position of each model point. This is done by running feature detectors independently to generate a response image for each point. A shape model is used to find the best combination of points.

In this work we applied the method of Cootes *et al.*[6] to accurately segment the Carpal area region. From the segmented Carpal area we then built Statistical model of appearance and extract texture parameters. These parameters were used in classical linear regression . The innovation of this work is in the application of the works of Cootes *et al.*[6] to segment the Carpal area region. In addition and more importantly we treat the Carpal area as an entity, whose texture features can be characterised to predict skeletal maturity of infants, thereby avoiding the correspondence problem associated with the Carpal bones of young children in predicting the 'common' (race normalized) bone age of the child.

The main problem is that of correspondence which results from inconsistent appearance of the component carpal bones as illustrated in Figure 1. In our approach we extracted the region around the carpal bones and built a texture-based appearance model.



a) Radiograph b) Carpals bones c) Carpals 2

Fig. 1. a) Example of a hand radiograph of a child with no bones, b) Carpal bones labeled and c) carpal region with two bones.

2. METHODS

2.1. Data Set

We have used a publicly available database of radiographs of the non-dominant hand of normally developing children from Ipilab laboratories[1]. We used a subset of 284 images representing ages from 0 -7 years. The dataset are of children from 4 ethnic groups (Caucasians, Asians, African Americans and Hispanics). The data also comes with two independent expert ratings who were blinded to the chronological age and the ethnicity of the children at the time of reading. For the purposes of segmentation evaluation the images were fully annotated.

In studying the structures of the carpal bones, we consider the average of the experts estimates of the bone age as the 'common' bone age.

2.2. Constrained Local Models (CLM)

We segment the region around the carpal bones using Constrained Local Models (CLMs) of [8]. We follow the method of Cootes *et al.*[6]. CLM combines global constraints with local models to segment an object form an image. It does this by considering the pattern of intensities. Based on a number of landmark points outlining the contour of the object in a set of images, we train a statistical shape model by applying PCA to the aligned shapes [7]. This yields a linear model model of shape variation which represent the position of each landmark point using $\mathbf{x}_i = T_o(\bar{\mathbf{x}}_i + \mathbf{P}_i\mathbf{b} + \mathbf{r})$ where $\bar{\mathbf{x}}_i$ gives the mean in the reference frame, $\mathbf{P}_i$ is a set of modes of variation, $\mathbf{b}$ are the shape model parameters, $\mathbf{r}$ allows small deviation from the model, and T_o apply a global transformation with parameters θ. To match a CLM to a new image we seek the shape and pose parameters $\mathbf{p} = \{\mathbf{b}, \theta\}$, which optimize the fit to the model. These parameters seek to optimize $\sum_{i=1}^{n} \mathbf{R}_i(T_\theta(\bar{\mathbf{x}}_i + \mathbf{P}_i\mathbf{b} + \mathbf{r}))$ where at every position i, $\mathbf{R}_i$ is the stored value of the quality of fit at every position representing the similarity between template texture at this landmark leaned from the model and the texture at the same position.

2.3. Voting with Random Forest (RF) Regressors

We applied RF similar to the Hough Forest approach of [9],but we did not require voting to be dependent on class labels. We adopted the method used in [6] and [15] where votes are gathered from regions around every point. During training a set of decision trees (a Random Forest) is trained so that each predicts the displacement from a given image patch to the target point. When searching, each tree is scanned over nearby patches in a grid, and produces a vote for where the target point is. All votes for point i are combined in an array, $\mathbf{R}_i$. For further details see [6] and [15].

2.4. Construction of Statistical Appearance Models

Statistical appearance models (SAM) [5] were generated by combining a model of shape variation with a model of texture variation. Each radiograph was manually annotated with points around important structures. Statistical models of shape and texture (intensities in the reference frame) were constructed by applying Principal Component Analysis (PCA) to the resulting annotations, leading to linear models of the form

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s\mathbf{b}_s \qquad \mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g\mathbf{b}_g \tag{1}$$

where $\bar{\mathbf{x}}$ is the mean shape, $\bar{\mathbf{g}}$ is the mean texture, $\mathbf{P}_s$,$\mathbf{P}_g$ are the main modes of shape and texture variation and $\mathbf{b}_s$,$\mathbf{b}_g$ are the shape and texture model parameter vectors. Combining the shape and texture models gives a combined appearance model of the form

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Q}_s\mathbf{c} \qquad \mathbf{g} = \bar{\mathbf{g}} + \mathbf{Q}_g\mathbf{c} \tag{2}$$

where $\mathbf{Q}_s$, $\mathbf{Q}_g$ are matrices describing the modes of variation derived from the training set and $\mathbf{c}$ is a combined vector of appearance parameters controlling both shape and texture.

2.5. Estimation of skeletal maturity

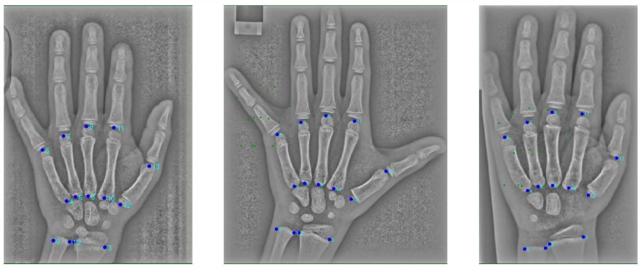
Given the appearance models we can compute shape, texture and appearance parameter vectors for each structure on each image.

We use classical linear regression of the form, $A = \mathbf{w}^T\mathbf{p} + A_0$, where A is the predicted age, $\mathbf{w}$ is a vector of weights, $\mathbf{p}$ is the parameter vector and A_0 is a constant. In the following we describe experiments comparing the performance of different models of the carpal bones [1].

[1] http://www.ipilab.org/BAAweb/

3. EXPERIMENTS

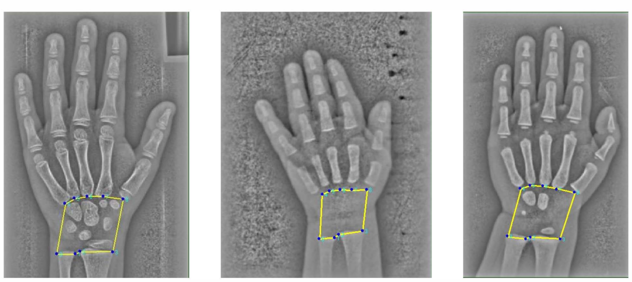
In order to locate the area containing the carpal bones, we built models of 14 points which can be reliably found on almost all hands. These include the 9 points shown on Figure 3 as well as 5 more on the top end of each of the metacarpals shown in Figure 2.



(a) (b) (c)

Fig. 2. Qualitative results of automatically segmented initial 14 points using RF regression voting.

The system combined a Hough Forest-like global search with local refinement (as in [15]). To evaluate the system we trained two models, one on the males, one on the females, and applied each model to the images from the other sex. The images were also manually annotated and the annotation was compared with the automatic points resulting into point to point errors of $\mathbf{1.5mm \pm 0.06}$. It should be noted that this error is higher than the reported error from [15], whose method we adopted. The reported point to curve error of less than 1mm for 98% of their images. The difference may be attributable to the sparse nature in the number of our points. We used 14 points to describe the entire area while they used 65 points due to the nature of their segmentation. However visual observation of the segmented carpal for most of the images are accurate as shown in Figure 3.

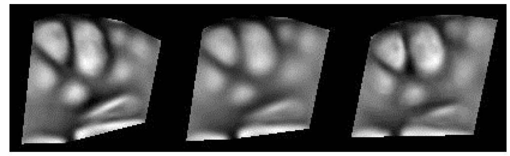


a) Good b) Acceptable c) Good

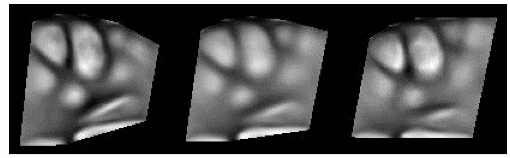
Fig. 3. Qualitative results of automatically segmented 9 points (extracted from the 14 points) using RF regression voting.

In our earlier work [4] we evaluated the utility of models of several structures of Carpal bones to estimate skeletal maturity, however the models were built from manual annotations. In this work we built models from fully automatic annotations as shown in Figure 3 and compare with the performance of equivalent model built from manual annotations. The resulting models are shown Figure 4a for automatically built models and Figure 4b for manually built models. below

Images of males and females were pooled to create both the automatic and manual models. For each model we computed the texture parameters for every image. In our earlier work [4] we had com-



(a)



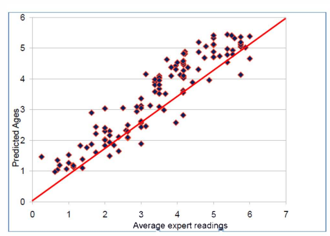
(b)

Fig. 4. Texture based appearance models (a) The first mode of variation of texture models from the 9 automatic points. (b)The first mode of texture models variation from the 9 manual points.

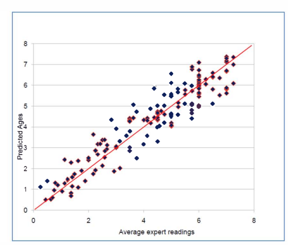
puted for each model of different Carpal structures, the shape, texture and appearance parameters for every image and using a method described below estimated bone age based on these three features.

	Texture model predictions	
	Tex	No of images
Female texture - manual	0.41±0.02	123
Female texture - automatic	0.43 ±0.02	123
Male texture - manual	0.52 ±0.03	161
Male texture - automatic	0.53 ±0.03	161

Table 1. Average performance various carpal based models - Mean absolute prediction error for male and female (years)from texture based models



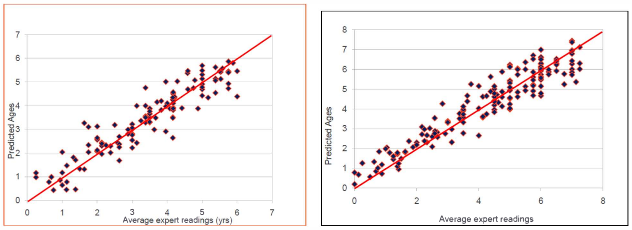
(a)



(b)

Fig. 5. Scatter plot for texture model from automatic point. (a) The plot of predicted ages against average bone age readings for female (b)The plot of predicted ages against average bone age readings for male

In this work, we evaluated the utility of linear age prediction models using a Leave-One-Out (LOO) paradigm. We trained linear regressors to predict age on all but one image, then tested the prediction on the left-out image. Since male and female children are known to develop at different rates, different regressor models were used for the male and the female sets. We evaluated performance using the mean absolute error between prediction and the average of the expert readings which we refer to as 'common' bone age. Figure 5 shows



(a) (b)

Fig. 6. Scatter plot for texture model from manual point. (a) The plot of predicted ages against average bone age readings for female (b)The plot of predicted ages against average bone age readings for male

the scatter plot of the performance of the extracted texture parameters from the Statistical appearance models. Table 1 also show the summary of the same results

4. DISCUSSION AND CONCLUSION

The results in Table 1 and Figures 5 and 6 show that there is a linear correlation between the predicted ages and the experts' ground truth ages. These results are also better than the results reported in [3] where we manually annotated different combinations of carpals bones to predict skeletal maturity. The fully automatically built texture model result of mean absolute errors of 0.43, 0.53 years in Table 1 compares favourably with other figures from the literature [20],[19], [4],[13], [14] and [10]. This shows that the simple automatic texture based Statistical Model of appearance method can used be to accurately estimate skeletal maturity of infants. This approach, also opens up possibilities for a variety of other datasets where correspondence is an issue. It should be possible to improve the accuracy of the point segmentation beyond the $1.5mm$ that we currently obtain, especially having earlier reported a point placement error of $0.7mm$, [2] with a similar data-set. However, this positional error does not seem to have negatively affected the skeletal age prediction results. We hope to apply this method to other growth related medical datasets where correspondence is a significant problem.

5. REFERENCES

[1] S. A. Adeshina. *Automatic Determination of Skeletal Maturity using Statistical Models of appearance*. PhD thesis, University of Manchester, 2010.

[2] S. A. Adeshina and T. F. Cootes. Constructing part-based models for groupwise registration. In *Proc. IEEE International Syposium on Biomedical Imaging*, 2010.

[3] S. A. Adeshina, T. F. Cootes, and J. Adams. Evaluating different structures for predicting skeletal maturity using statistical appearance models. In *Proc. MIUA 2009*, 2009.

[4] S. A. Adeshina, T. F. Cootes, and J. Adams. Evaluating the use of carpal bones for determination of skeletal maturity for infants. In *Proc. MIUA 2010*, 2010.

[5] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):681–685, 2001.

[6] T. F. Cootes, C. L. Mircea C Ionita, and P. Sauer. Robust and accurate shape model fitting using random forest regression voting. In *ECCV 2012*, pages 278–291. Springer-Verlag, 2012.

[7] T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham. Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59, Jan. 1995.

[8] D.Cristinacce and T.F.Cootes. Automatic feature localisation with constrained local models. *Pattern Recognition*, 41(10):3054–3067, 2008.

[9] V. Gall. J., Lempitsky. Class specific hough forests for object detection. In *IEEE Proc Computer Vision and Pattern Recognition*, pages 415–422, 2009.

[10] D. Giordano, C. Spampinato, G. Scarciofalo, and R. Leonardi. An automatic system for skeletal bone age measurement by robust processing of carpal and epiphysial/metaphysial bones. *Instrumentation and Measurement, IEEE Transactions on*, 59(10):2539–2553, October 2010.

[11] W. W. Greulich and S. I. Pyle. *Radiographic Atlas of Skeletal Development of Hand Wrist*. Palo Alto, CA: Stanford Univ. Press, 1971.

[12] D. Haak, J. Yu, H. Schramm, T. Seidl, and T. M. Deserno. Bone age assessment using support vector regression with smart class mapping. In *Proc. SPIE 2013*, 2013.

[13] M. Harmsen, B. Fischer, H. Schramm, T. Seidl, and T. M. Deserno. Support vector machine classification based on correlation prototypes applied to bone age assessment. *IEEE J. Biomedical and Health Informatics*, 17(1):190–197, 2013.

[14] C. Hsieh, T. Liu, and T. Jong. A fuzzy-based growth model with principle component analysis selection for carpal bone-age assessment. *Med Biol Eng Comput.*, 48(6):579–588, 2010.

[15] C. Lindner, S. Thiagarajah, J. M. Wilkinson, G. A. Wallis, and T. F. Cootes. Accurate fully automatic femur segmentation in pelvic radiographs using regression voting. In *MICCAI*, pages 353–360, 2012.

[16] M. Niemeijer, B. van Ginneken, C. Maas, F. Beek, and M. Viergever. Assessing the skeletal age from a hand radiograph: automating the tanner-whitehouse method. In *Proceedings of SPIE – Volume 5032*, pages 1197–1205, May 2003.

[17] E. Pietka, A. Gertych, S. Pospiech, F. Cao, H. K. Huang, and V. Gilsanz. Computer-assisted bone age assessment: image preprocessing and epiphyseal/metaphyseal roi extraction. *IEEE Transactions on Medical Imaging*, 20(8):715–729, 2001.

[18] J. M. Tanner, R. H. Whitehouse, W. A. Marshall, M. R. Healy, and H. Goldstein. *Skeletal Maturity and Prediction of Adult Height (TW2 Method)*. New York, NY: Academic, 1975.

[19] H. Thodberg, S. Kreiborg, A. Juul, and K. Pedersen. The bonexpert method for automated determination of skeletal maturity. *Medical Imaging, IEEE Transactions on*, 1(1309):52–66, 2009.

[20] A. Zhang, A. Gertych, and B. Liu. Automatic bone age assessment for young children from newborn to 7-year-old using carpal bones. *Computerized Medical Imaging and Graphics*, 31:299–310, 2007.