RESEARCH ARTICLE OPEN ACCESS

Radiographic Model Matching with Markov Graph Shape Model

Steve A. Adeshina[1*] and Timothy F. Cootes [2*]

[1]Department of Electrical & Electronics Engineering,Nile University of Nigeria, Abuja,
E-mail: steve.adeshina@nileuniversity.edu.ng
[2]Center for Imaging Sciences, The University of Manchester, U.K

Abstract—Skeletal maturity assessment is important for diagnosing and monitoring growth disorders in Children and young adults. Statistical models of bone shape and appearance have been shown to be useful for estimating skeletal maturity. One critical requirement in automated skeletal maturity estimation is matching built models to unseen images of the bones of the hand. Oftentimes some form of initialization is required to prevent the model from falling into local minima. In this work we used Markov Graph Shape models (MGSM) to initialize the image of an incoming radiographic image and then fit a global Active Appearance models of the whole hand using the found points from the Markov Graph Shape models as 'weighted' constraints. Having found the approximate positions of the bones of hands, we then fit local models to refine the model fit. By analysing performance on dataset of 70 digitized images of normal children we achieved a model fitting accuracy of an overall point-to-point median error of $0.67mm$.

Keywords—Skeletal maturity assessment, Constrained Active Appearance Models, Markov Graph Shape Models.

1 INTRODUCTION

Skeletal maturity assessment is important in diagnosing growth disorders in Children and young adults. Active Appearance Models (AAMs) are known for their effectiveness in medical image segmentation and model matching. However, they need good initialisation. We use Markov Graph Shape models (MGSM) to find sparse points on the images [14]. These sparse points are used to initialise the global AAMs. With the knowledge of the approximate position of all the bones of the hand, we further used these known points as constraints while finding the exact position of the bones. This was done by fitting local AAMs (models of different bone complexes) using the approximate points from global models as constraints. This work presupposes that Global and Local Statistical models of appearance have been built. This we have done in our earlier works [3] and [1]. In this work we have used a Global-Local model fitting strategy to fit statistical models to Radiographic images. This is with the ultimate objective of extracting model parameters of shape, texture and appearance for estimating skeletal maturity. The efficacy of the extracted parameters for the estimation is well established in literature. [1], [13]. The objective of automatically fitting a model to an unseen image as presented in this work and several others in literature is key to achieving an automated Skeletal maturity assessment system. Without these efforts, skeletal assessment will need a manual intervention with all its attendant inconsistencies.

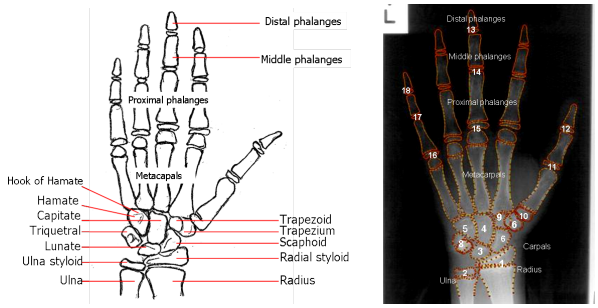
To achieve the above objective we use the global AAMs described in earlier works [3]. Each new image is matched with a global AAM. This gives the approximate position of the different bones of the hand.

With the knowledge of the approximate position of all the bones of the hand, we further use these known points as constraints while finding the exact position of the bones. This was done

Manuscript received September 27, 2017; revised September 31, 2017.

by fitting local AAMs (models of different bone complexes) using the approximate points from global models as constraints.

Figure 1b shows the skeletal maturity growth points based on Tanner and Whitehouse (TW) method of determining skeletal maturity [1] while figure 1a shows the different bones of the hands. Two methods are prevalent in Clinical radiology for the determination of skeletal maturity and they are attributed to Greulich and Pyle [9] and Tanner and Whitehouse [1]. The former considers all the 28 bones of the hand in a subjective manner, while the former considers 13 Radius Ulna and Short (RUS) in an objective manner and it is believed to be more accurate. Our goal is to match built models to the 13 TW RUS bones. These bones are listed in Table 2.



(a) (b)

Fig. 1. (a) Bones of the hand [1].
(b) Skeletal maturity growth points based on TW method. **RUS bones:** Radius(1), Ulna(2), Metacarpal I, III, V, Proximal phalanges(ppha) I, III, V (10,15,16) , Middle phalanges (mpha) III, V (14,17), Distal phalanges (dpha) I, III, V (12,13,18); **Carpal bones:** Capitate(4), Hamate(5), Triquetral(8), Lunate(3), Scaphoid (8), Trapezium(6) and Trapezoid(9).

Figure 2 shows the process for locating the bones in a new radiograph. Markov Graph Shape models (MGSM) with global AAM are used to find the approximate locations of the bones. The bone complexes of interest are then extracted. Local models of equivalent bone complexes are then fit using points from the global model as constraints. The accuracy of the model fit is evaluated at the global and local levels.

The innovation in this work is the combination of an MRF-based local shape model for guided candidate selection with a PCA-based global shape model for regularization. The method also innovatively uses a two-stage cascade implementation for robustness: in the first step, we determine the global translation, orientation and scale using only the MRF to localize a salient subset of the landmark points; in the second step, we refine the previously localized points and localize a larger set of less salient landmarks in the model by allowing the global model to account for more of the shape variation.

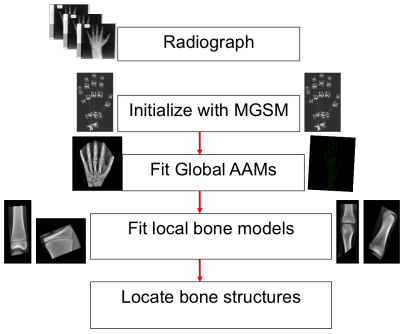


Fig. 2. Process diagram for automatic matching of hand radiographs using Active Appearance Models (AAM)

Additionally the combination of Markov Random Graph Models (MGSM) with Constrained Active Appearance Model with a weighted fit is also an innovation. The original AAM constrained fit assumes equal variances for all points. We however introduced a weighted fit. After the initial AAM fit, we computed the Euclidean points' error difference between the target points and the best model fit while specifying a maximum error threshold. We modified the constraint weights based on how large the errors are.

2 Related Work

A comparable work in literature includes the work of Luis-Garcia *et al* [11] where an adaptive snake was used to find the contours of the hand. An average success rate of 73.9% was reported. Another approach by Pietka *et al* [12] reduced the regions of interest to Epipyseal-Metapyseal Region of Interest (EMROI). A success rate of 96% was reported. It was however

reported that about 14% of their data set was set aside due to poor quality. This action may have had a positive effect on the success rate [11]. The works of Giordano *et al* [8] and Pietka *et al* [12] examined a reduced number of bone complexes while our work and Luis-Garcia *et al*'s considered the entire bones of the hand. Thodberg *et al.* [13], whose work is very closely related, did not quote a result for the localisation of bones. Tresadern *et al.* [14] used the MGSM to find sparse points in an image. This work did not proceed to segmenting and matching local models which are often requred for skeletal maturity. Cootes *et al.* [6] and Lindner *et al.* [10] in their very excellent work presented a robust shape model fitting using Random Forest Regression Voting and they applied their methods to several datasets including hand radiographs with excellent results. However they needed hundreds of images. Adeshina *et al.* [2] presented a similar method with an impressive results using a Part and Geometry models. This work differs from our earlier work by demonstrating that Markov Based Graph models can be used instead of Parts and Geometry models to acheive the same purpose and indeed better results. This method is also more robust as its model building process require fewer number of images. The method will prove very useful when the number of available dataset is limited.

3 METHODS

3.1 Data Set

We have access to a database of radiographs of the non-dominant hand of normally developing children being collated at the University of Manchester [1] for a different bone ageing project. In this work, we used a subset of 142 digitized radiographs of normal children.

3.2 Construction of Markov graph shape model

We performed initial experiments in constructing sparse models capable of being used for initialising global AAM for the localisation of the bones of the hand in hand radiograph. We used a Markov graph (local patch models) shape models.This involves representing an image with a set of patches together with a geometrical model of their relative positions. The geometry is modeled with a global pose and linear shape model. The local displacements from the global model is modeled with a Markov Random Field (MRF). Matching to a new radiograph involves an alternating scheme in which an MRF inference technique selects the best candidate for each point, these are then used to update the parameters of the global pose and shape model. This scheme was made robust by introducing different level of cascade with increasing complexity.

This technique formulates the object matching as a global shape alignment problem combined with MRF-based local modeling. The model can be represented by a set of N points, such that $\mathbf{X} = \{\mathbf{x}_i = (u_i, v_i)\}$. Given a query image, $\mathbf{I}$, the objective is to find the optimal set $\mathbf{X}^*$, that maximises the posterior,

$$p(\mathbf{X} \mid \mathbf{I}) \propto p(\mathbf{I} \mid \mathbf{X})\, p(\mathbf{X}). \tag{1}$$

The number of found patches for each $\mathbf{x}_i$ can be very large and combinatorially intractable. As a result of this, we made assumptions of conditional independence between found features, this enables us approximate the joint prior with a Markov Random Field (MRF). The process reduces the complexity of the problem so that an approximate solution, $\mathbf{Y}$, can be found efficiently. We can then regularise with a global shape prior to push the approximate solution towards the optimum $\mathbf{X}^*$.

The iterative procedure can be summarised as follows:

- Initialise point locations, $\mathbf{X}_0^*$ using fixed locations from the detector output.
- For s = 1 to N
  1) Select the best candidates, $\mathbf{Y}_t = \arg\max_{\mathbf{Y}} p(\mathbf{I} \mid \mathbf{Y})\, p\left(\mathbf{Y} \mid \mathbf{X}_{t-1}^*\right)$ An MRF Solver picks the best combination of candidates by minimising energy function G, the dynamic programming solver was used in the case. $G = \sum \theta(\mathbf{y}_i) + \alpha \sum \psi(\mathbf{y}_i, \mathbf{y}_j)$

1. Special thanks to Prof Judith Adams for providing the dataset

where $\mathbf{y}_i, \mathbf{y}_j$ are patch candidates associated with patches $i, j.\alpha$ is a parameter that weights the influence of the prior and likelihood terms, $\psi(\cdot)$ is an error function that indicates the goodness of fit between a pair of candidates and $\theta(\cdot)$ is an error function that indicates the goodness of fit with the image data.

2) Update points by regularising the candidates by fitting a shape model $\mathbf{X}_t^* = \arg\max_{\mathbf{X}} p(\mathbf{X} \mid \mathbf{Y}_t)$

The algorithm was implemented in a cascade format, with two levels of 12 and 21 points. The candidate points have to be initialised by marking up a number of points on the training images. In this form, the model training remains semi-manual as sparse landmarks are required for each image.

3.3 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 automatically 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

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s\mathbf{b}_s \qquad \mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g\mathbf{b}_g \tag{2}$$

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

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{Q}_s\mathbf{c} \qquad \mathbf{g} = \bar{\mathbf{g}} + \mathbf{Q}_g\mathbf{c} \tag{3}$$

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.

3.4 Active Appearance Model search

The AAM matching algorithm is outlined below. There are two main components: a parameterized model of object appearance, and an estimate of relationship between parameter errors and image residuals [4].

The appearance model parameters, $\mathbf{c}$, and shape transformation parameters, $\mathbf{t}$, define the position of the model points in an image frame, $\mathbf{X}$, which gives the shape of the image patch to be represented by the model. During the matching we sample the pixels in the region of the image, $\mathbf{g}_{im}$ and project them to the texture model frame, $\mathbf{g}_s = T^{-1}(\mathbf{g}_{im})$. The current texture model is given by $\mathbf{g}_m = \bar{\mathbf{g}} + \mathbf{P}_g\mathbf{Q}_g\mathbf{c}$ The model, image difference in the normalized texture frame is

$$\mathbf{r}(\mathbf{p}) = \mathbf{g}_s - \mathbf{g}_m \tag{4}$$

where $\mathbf{p}$ are the parameters of the of the model $\mathbf{p}^T = (\mathbf{c}^T|\mathbf{t}^T|\mathbf{u}^T)$. A scalar measure of difference is the sum of squares of the elements of $\mathbf{r}$, $E(\mathbf{p}) = \mathbf{r}^T\mathbf{r}$. A first order Taylor expansion of equation 4 gives

$$\mathbf{r}(\mathbf{p}+\delta\mathbf{p}) = \mathbf{r}(\mathbf{p}) + \frac{\partial \mathbf{r}}{\partial \mathbf{p}}\delta\mathbf{p} \tag{5}$$

where the ij^{th} element of matrix $\frac{\partial \mathbf{r}}{\partial \mathbf{p}}$ is $\frac{dr_i}{dp_j}$

If our current residual is $\mathbf{r}$, we intend to choose $\delta\mathbf{p}$ so as to minimize $|\mathbf{r}(\mathbf{p}+\delta\mathbf{p})|^2$ by equating equation 5 to zero, we then get a Root Mean Square solution as shown below:

$$\delta\mathbf{p} = -\mathbf{R}\mathbf{r}(\mathbf{p}) \;\; where \;\; \mathbf{R} = \left(\frac{\partial \mathbf{r}^T}{\partial \mathbf{p}}\frac{\partial \mathbf{r}}{\partial \mathbf{p}}\right)^{-1}\frac{\partial \mathbf{r}^T}{\partial \mathbf{p}} \tag{6}$$

It would be necessary to re-calculate $\frac{\partial \mathbf{r}}{\partial \mathbf{p}}$ in standard optimization exercise and this is computationally expensive. It is however considered approximately fixed - the estimation can be done from the training set. Numeric differentiation can be used to estimate $\frac{\partial \mathbf{r}}{\partial \mathbf{p}}$ displacing each parameter systematically from the known optimal value on typical images and computing the average over a set. Residuals at displacements of differing magnitudes measured - say $0.5SD$ for each parameter and this is combined with a Gaussian to smooth them. $\mathbf{R}$ is precalculated and used for subsequent searches [4]. The

images used in calculating the partial residuals can be examples from the training set or the images generated using the appearance model.

3.5 Constrained active appearance models (CAAM)

The AAM as a local search method depends on an update matrix learned near correct solutions. As a result, it depends upon adequate initialisation. Usually this initialisation is provided by prior estimates of some of the shape points, either manually, or through automatic methods as discussed in preceding sections. There may be some prior knowledge of the variances associated with these initialisation points. Cootes *et al.* developed the Constrained AAM [7] to incorporate such constraints.

Essentially the least squares minimisation of the standard AAM is replaced by a maximum a-posteriori (MAP) formulation, which seeks to maximise the probability of the model given the data which (by Bayes theorem) is proportional to:

$$P(data|model)P(model) \tag{7}$$

Assuming a uniform prior on the model parameters. This can be equated to a least squares formulation where Gaussian residuals are not correlated and the variances are equal. A Gaussian prior could be assumed on the model parameters, and Cootes *et al.* showed how the AAM update step can be reformulated to incorporate this prior. In the rest of this section we will be concentrating on incorporating prior knowledge about constrained points. We worked with a simplified version of [7] and assumed a constant model prior [1].

The Jacobian of the residual can be represented as $\mathbf{J}$ (see section 3.4)

$$\mathbf{J} = \frac{\delta \mathbf{r}}{\delta \mathbf{p}} \tag{8}$$

as a result equation 6 in section section 3.4 can be re-written as

$$\mathbf{R} = \left[\mathbf{J}^T\mathbf{J}\right]^{-1}\mathbf{J}^T \tag{9}$$

Suppose we have prior estimates of the positions of some points in the image frame $\mathbf{X}_0$, together with their covariance matrix $\mathbf{S}_X$. Unknown points can be represented by zeroes, together with large upper bounds in $\mathbf{S}_X$, and effectively zeroes in ${\mathbf{S}_X}^{-1}$. Let $\mathbf{d}(\mathbf{p}) = (\mathbf{X} - \mathbf{X}_0)$ be a Vector of the displacements of the current point positions from their prior positions. We assume further that the prior point positions are Gaussian distributed, and also that the texture residuals are independently and identically distributed with variance ${\sigma_r}^2$. $\mathbf{r}$ is the Vector of residuals. Then maximising the logarithm of the MAP is equivalent to minimising (see section 3.4):

$$E_1(\mathbf{p}) = {\sigma_r}^{-2}\mathbf{r}^T\mathbf{r} + \mathbf{d}^T{\mathbf{S}_X}^{-1}\mathbf{d} \tag{10}$$

By using a first order Taylor expansion similar to that used to derive the basic AAM update equation, the parameter update is given by the solution to the equation set:

$$\mathbf{A}\delta\mathbf{p} = -\mathbf{a} \tag{11}$$

where, after defining the Jacobian of $\mathbf{d}$ w.r.t $\mathbf{p}$ as $\mathbf{K}$

$$\begin{aligned} \mathbf{A} &= \left({\sigma_r}^{-2}\mathbf{J}^T\mathbf{J} + \mathbf{K}^T{\mathbf{S}_X}^{-1}\mathbf{K}\right) \\ \mathbf{a} &= \left({\sigma_r}^{-2}\mathbf{J}^T\mathbf{r}(\mathbf{p}) + \mathbf{K}^T{\mathbf{S}_X}^{-1}\mathbf{d}\right) \end{aligned} \tag{12}$$

and

$$\mathbf{K} = \frac{\delta \mathbf{d}}{\delta \mathbf{p}} \tag{13}$$

When computing the prior point displacement Jacobian $\mathbf{K}$, it is necessary to take into account the global pose transformation $\mathbf{t}$ as well as the appearance model parameters $\mathbf{c}$. Cootes further developed the special case of isotropic prior point positional variance with zero off-diagonal terms, and when the pose transformation $S_t(\mathbf{x})$ is a similarity transform which scales by s. Let $\mathbf{x}_0$ be the prior point positions mapped into the model frame, so $\mathbf{x}_0 = {S_t}^{-1}(\mathbf{X}_0)$, and let $\mathbf{y} = s(\mathbf{x} - \mathbf{x}_0)$. Then $\mathbf{d}^T{\mathbf{S}_X}^{-1}\mathbf{d} = \mathbf{y}^T{\mathbf{S}_X}^{-1}\mathbf{y}$.

In this case:

$$\begin{aligned} \mathbf{A} &= \left({\sigma_r}^{-2}\mathbf{J}^T\mathbf{J} + {\mathbf{K}_m}^T{\mathbf{S}_X}^{-1}\mathbf{K}_m\right) \\ \mathbf{a} &= \left({\sigma_r}^{-2}\mathbf{J}^T\mathbf{r}(\mathbf{p}) + {\mathbf{K}_m}^T{\mathbf{S}_X}^{-1}\mathbf{y}\right) \end{aligned} \tag{14}$$

The Jacobian $\mathbf{K}_m$ is the concatenation $\left(\frac{\delta \mathbf{y}}{\delta \mathbf{c}} | \frac{\delta \mathbf{y}}{\delta \mathbf{t}}\right)$ and:

$$\begin{aligned} \frac{\delta \mathbf{y}}{\delta \mathbf{c}} &= s\mathbf{Q}_s \\ \frac{\delta \mathbf{y}}{\delta \mathbf{t}} &= -s\frac{\delta\left(S_t{}^{-1}(\mathbf{X}_0)\right)}{\delta \mathbf{t}} - (\mathbf{x}-\mathbf{x}_0)\cdot\left(\frac{s_x}{s}, \frac{s_y}{s}, 0, 0\right) \end{aligned} \quad (15)$$

The update equation can then be solved using standard methods in linear algebra; for example, since the matrix $\mathbf{A}$ is symmetric, Cholesky decomposition can be used for speed to invert $\mathbf{A}$; but if that appears ill-conditioned, then SVD can be used to robustly calculate an inverse (in the least-squares sense).

4 Experiments and Results

The first step in processing a new image is to match a sparse model which is able to locate the main joints. This was achieved using Markov Graph Shape Models.

We perform a constrained fit with AAM on the radiograph, using dense points as constraints. Cootes *et al.* [7] reported lower boundary errors with higher number of constraint points. For this reason we warped 330 points using the points found by our Markov Graph Shape model with a Thin Plate Spline (TPS) deformation field. The original AAM constrained fit assumes equal variances for all points. We however introduced a weighted fit. After the initial AAM fit, we computed the Euclidean points' error difference between the target points and the best model fit while specifying a maximum error threshold. We modified the constraint weights based on how large the errors are. When the errors are greater than the threshold we allocate a weight of 0. This is an infinite variance on such points, equivalent to ignoring the constraints on such points and allowing the formal model constraints to take precedence.

Fitting a global model to these bones could only get approximate positions of the bones. This is because of the global model constraint enforces global fidelity, to ensure shapes do not assume implausible poses. We use a local model to capture the extensive variability which is inherent in the bones of the hand. 13 local models were used, corresponding to each RUS complex, to further refine the location of the bones.

4.1 Markov graph shape model experiments

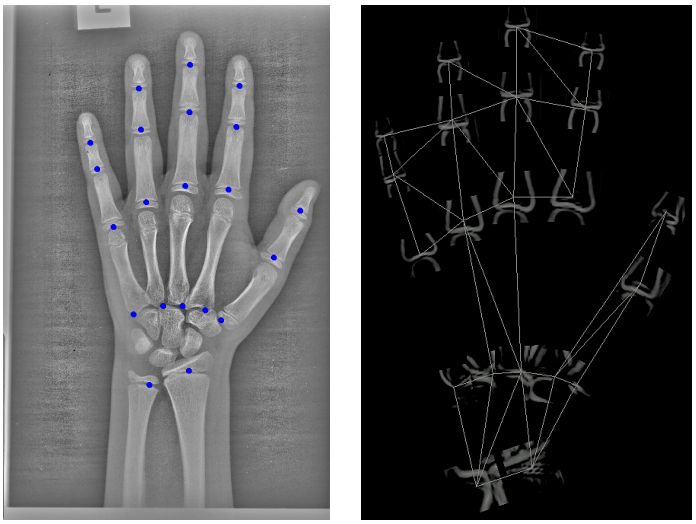
We used a subset of 142 radiographs for experimentation. We built a Markov graph shape model of salient structure of the bones of the hand using 72 images, while we tested the models on previously unseen 70 images. 21 landmarks were manually annotated on each of the 72 images. These 21 landmarks are a subset of a 330 landmarks earlier put on the 72 training images. Typical landmarks on a single image is shown in Figure 3a. A model of two level of cascade was built with 12 and 21 points respectively. The model as shown in Figure 3b was tested on the remaining 70 images. Preliminary experiments were performed to evaluate the influence of the number of cascade and the structure of MRF prior. We evaluated the performance of a given image by computing $d\left(\mathbf{x}_i^*, \mathbf{z}_i\right)$ which is the mean distance between the found point, $\mathbf{x}_i^*$, and its manually labelled counterpart $\mathbf{z}_i$. To make the euclidean error distance invariant to scale, it is normalised with respect ti a reference distance, d_{ref}. This reference distance was defined as the length of the fifth metacarpal. The normalised scale invariant median error m_e is given as;

$$m_e = \frac{1}{N}\sum_i \frac{d\left(\mathbf{x}_i^*, \mathbf{z}_i\right)}{d_{ref}} \quad (16)$$

The sparse model and initial landmark points are shown in figure 3a and 3b respectively. The performance evaluation is shown in Figure 4. We have applied the method of Tresadern *et al.* [14] to acheive this results.

4.2 Constrained Active Appearance global Model fitting experiments

We have applied the CAAM to Markov Graph Shape model built by using the 21 found points from the part based model to warp all 330 points (marked on a reference image) to the incoming radiograph and then fit an AAM with modified version of Constrained AAM fit (refine with AAM) [7]. In this way we have



a b

Fig. 3. (a) Hand radiograph images with landmark points, (b) A trained model.

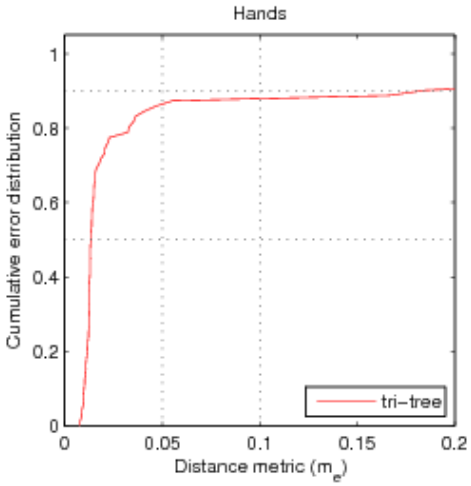


Fig. 4. Performance evaluation on 70 images.

two sets of dense annotations which identified the position of the bones. The first is that from using the 19 found points to propagate 330 points to all images in the set (with TPS). For clarity we call this 'Initial points from TPS warp'. The second is the result of fitting Appearance Models using the propagated points as constraints. Again for clarity we call this 'After CAAM refinement'. This was done using the global model of the entire hand. In theory the result of the second should be better than the first.

The table below shows the results of points to point errors for a global model fit. These errors were found by computing the point to point errors between a set of 330 automatic points and their manually annotated equivalent as shown in table 1. This is also illustrated in Figure 5.

	MGSM Global Model Accuracy	
	Mean (mm)	Median (mm)
Initial points from TPS warp	2.07±0.1	1.75
After CAAM refinement	1.19 ±0.1	0.91

TABLE 1
Search error statistics. Point to point errors (mm) of the global model fit using MGSM based constraints with CAAM.

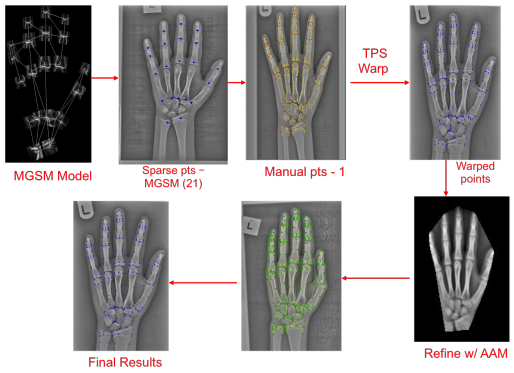
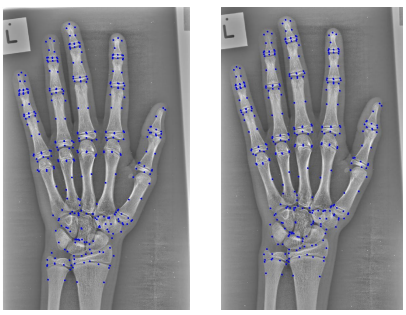


Fig. 5. Process flow for CAAM fit. From 21 automatically found points on 70 images, 330 manual points on a reference image, a TPS warp of manual points to other images, a further refine fit and final points' result.

Figure 6a shows the result of initial point from TPS warp and Figure 6b shows the result after AAM refinement.



a b

Fig. 6. (a) Qualitative result of initial point from TPS warp on an image (b) Resulting points after CAAM refinement.

4.3 Local model fitting experiments

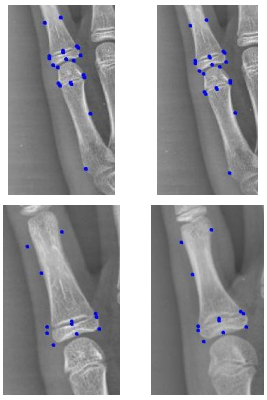
We extracted the 13 bone complexes which have been proven sufficient for skeletal maturity assessment as shown in Adeshina *et al.* [3]. Earlier clinical studies by Tanner *et al.* [1] also confirm the findings. It must be noted that we located all the bones of the hand, and only report on the complexes that are relevant to our work based on our earlier experiments.

The extracted bone complexes are from our global CAAM fit, having recorded better results when compared to a direct thin plate spline propagation as shown in Table 1. We then fitted local models to each of the complexes using the extracted points as constraints. Table 2 shows the result of points extracted without a local AAM refinement and that with local AAM refinement when compared with manual ground truth annotation.

Having fit the AAMs to the bone complexes, the shape, texture and appearance parameters of the models are extracted for further analysis. These parameters have been found to have a linear correlation with skeletal maturity . Qualitative results of automatic points' and refined points' errors are shown in Figure 7. For clarity we define 'automatic warp points' as dense bone border points obtained from a TPS warp of the 21 MGSM points. While 'refined with CAAM' are points after a CAAM have been applied.

	Automatic Warp points		Refine AAM points	
	Mean	Median	Mean	Median
Ulna	1.14±0.03	1.15	1.02±0.05	0.80
Radius	1.04 ±0.03	1.02	1.03±0.02	0.86
ppha5	0.84 ±0.02	0.79	1.77±0.61	0.64
ppha3	0.83 ±0.03	0.79	1.60 ±0.34	0.90
ppha1	0.98 ±0.12	0.71	0.92±0.11	0.64
mpha5	0.84 ±0.04	0.75	0.65±0.05	0.46
mpha3	0.86 ±0.04	0.80	0.56 ±0.02	0.44
mcarpal5	0.90 ±0.02	0.86	0.92±0.06	0.73
mcarpal3	0.94 ±0.03	0.86	1.03 ±0.06	0.84
mcarpal1	1.06 ±0.02	1.02	0.90±0.04	0.85
dpha5	0.85 ±0.03	0.73	0.62 ±0.05	0.49
dpha3	0.79 ±0.02	0.78	0.41±0.02	0.56
dpha1	1.11 ±0.26	0.87	1.16±0.29	0.50
Average	0.93±0.05	0.85	0.97±0.13	0.67

TABLE 2
Search error statistics. Point to point errors (mm) of local models without and with refinement with CAAM



a) b)

Fig. 7. Qualitative results of final points location for mcarpal5, ppha5. Column (a) with automatic warp and (b) with CAAM refinement

5 Discussion and Conclusions

This work presents a method to automatically fit a model to a radiograph with some human intervention. We dealt with the issues of AAM initialisation by using Markov Graph Shape models (MGSM). We formulated the necessary framework to ensure that an accurate localisation of the bones is achieved.

We had used the initial points, found by the Markov Graph Shape model to propagate a detailed annotation on one image (330 points) and compared the results of this with a CAAM global model fit as shown in Table 1. The improved result shows that the use of a CAAM was justified. We acheive an overall median of about $0.67mm$ which is a maginal increase over our earlier work which used a Part and Geometry model [2].

A comparable work to this aspect of localisation of the outline of the bones in literature is the work of Luis-Garcia *et al* [11] where an adaptive snake was used to find the contours of the hand. An average success rate of 73.9% was reported. Giordano *et al* [8] reported 86% success rate for EMROI extraction. The works of Giordano and Pietka *et al* examine a reduced number of bone complexes while our work and Luis-Garcia *et al*'s considered the entire bones of the hand. It is difficult to compare results because of uncertainty about the metric used (i.e.N% success in the cited papers, a numerical error in this work).

We achieved a median error of 0.85mm with automatically warped points without CAAM refinement and a median error of 0.67mm with CAAM refinement. We achieved a bones' matching accuracy of $0.67mm$ and this is comparable to what is obtained in the literature [11].

Cootes *et al.* [6] and Lindner *et al.* [10] in their work with Random Forests and Constrained Local Models reported a mean point error of 'less than $1mm$ 90% of the time. The requirement of several hundreds of images is a drawback (they used 200 for training and 200 for testing). Though their results are excellent as a generic framework for several types of datasets (and we hope to further explore its application to skeletal maturity), our results are comparable for this dataset.

The closest work to our is that of Adeshina *et al.* [2] where a Part and Geometry model was used to initialize a global of Radiographic images. They achieved a median error of $0.87mm$ with automatically warped points without CAAM refinement and a median error of $0.71mm$ with CAAM refinement. A bones' matching accuracy of $0.71mm$. We achieved a median error of $0.85mm$ with automatically warped points without CAAM refinement and a median error of $0.67mm$ with CAAM refinement. We achieved a bones' matching accuracy of $0.67mm$ This is just a marginally better result. Whereas Adeshina *et al.* [2] model building process require very sparse annotation of just a single image to build the Parts and Geometry model, in this method we are required to manually sparsely annotate all the images require for building the MGSM models. However the advantage of this work is that, as a result of 'a more supervised learning' the method require a comparatively lower number of images for building a robust MGSM models.

ACKNOWLEDGMENTS

The authors would like to thank Prof Judith Adams for providing the Radiograph images. Dr Phil Tresadrn for codes and experiments on Makov Graph Shape Models.

Steve A. Adeshina Dr Steve A Adeshina is with the Nile University of Nigeria, Abuja Nigera.

Timothy F. Cootes Prof Tim Cootes is with Centre for Imaging Sciences, The University of Manchester, Manchester, UK.

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. Automatic model matching using part based model constrained active appearance models for skeletal maturity. In *Proc. 2015 IEEE International Conference on Electronics Computers and Computation*, pages 733–740, 2015.

[3] S. A. Adeshina, T. F. Cootes, and J. Adams. Evaluating different structures for predicting skeletal maturity using statistical appearance models. In J. Dehmeshki, A. Hoppe, and D. Greenhill, editors, *Proceedings of the 13th Medical Image Understanding and Analysis Conference*, pages 62–66, London, England, 2009. BMVA Press.

[4] T. Cootes and C. Taylor. Statistical models of appearance for medical image analysis and computer vision. In *SPIE Medical Imaging*, Feb 2001.

[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, M. C. Ionita, C. Lindner, 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 and C. J. Taylor. Constrained active appearance models. In 8^{th} *International Conference on Computer Vision*, volume 1, pages 748–754. IEEE Computer Society Press, July 2001.

[8] D. Giordano, R. Leonardi, F. Maiorana, G. Scarciofalo, and C. Spampinato. Epiphysis and metaphysis extraction and classification by adaptive thresholding and d.o.g filtering for automated skeletal bone age analysis. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - EMBS 2007*, pages 6551 – 6556, 22-26 Aug. 2007.

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

[10] C. Lindner, P. Bromiley, M. Ionita, and T. Cootes. Robust and accurate shape model matching using random forest regression-voting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 01(99):1862–1874, 2015.

[11] R. D. Luis-Garcia, M. Martin-Fernandez, J. Arribas, and C. Alberola-Lopez. A fully automatic algorithm for contour detection of bones in hand radiographs using active contours. In *Proc. International Conference on Image Processing*, pages 421–424, Sept. 2003.

[12] 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.

[13] 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.

[14] P. Tresadern, H. Bhaskar, S. Adeshina, C. Taylor, and T. Cootes. Combining local and global shape models for deformable object matching. In A. C. Daniel Alexander and S. Prince, editors, *Proc. British Machine Vision Conference*, pages 1–1. BMVA, Sept. 2009.