Bone Age Assessment by Content-Based Image Retrieval and Case-Based Reasoning

Benedikt Fischer[*a], Petra Welter[a], Christoph Grouls[b], Rolf W. Günther[b], Thomas M. Deserno[a]
[a]Department of Medical Informatics, [b]Department of Diagnostic Radiology,
RWTH Aachen University, Pauwelsstr. 30, 52057 Aachen, Germany

ABSTRACT

Skeletal maturity is assessed visually by comparing hand radiographs to a standardized reference image atlas. Most common are the methods by Greulich & Pyle and Tanner & Whitehouse. For computer-aided diagnosis (CAD), local image regions of interest (ROI) such as the epiphysis or the carpal areas are extracted and evaluated. Heuristic approaches trying to automatically extract, measure and classify bones and distances between bones suffer from the high variability of biological material and the differences in bone development resulting from age, gender and ethnic origin. Content-based image retrieval (CBIR) provides a robust solution without delineating and measuring bones. In this work, epiphyseal ROIs (eROIS) of a hand radiograph are compared to previous cases with known age, mimicking a human observer. Leaving-one-out experiments are conducted on 1,102 left hand radiographs and 15,428 metacarpal and phalangeal eROIs from the publicly available USC hand atlas. The similarity of the eROIs is assessed by a combination of cross-correlation, image distortion model, and Tamura texture features, yielding a mean error rate of 0.97 years and a variance of below 0.63 years. Furthermore, we introduce a publicly available online-demonstration system, where queries on the USC dataset as well as on uploaded radiographs are performed for instant CAD. In future, we plan to evaluate physician with CBIR-CAD against physician without CBIR-CAD rather than physician vs. CBIR-CAD.

1. INTRODUCTION

To estimate the maturity of patients, bone age assessment (BAA) based on hand radiographs constitutes a frequent yet time-consuming task for radiologists. The two most commonly used methods are based on image comparison. In the method of Greulich & Pyle [1], the radiologist compares all bones of the hand to radiographs in the standard atlas. In the method of Tanner & Whitehouse (TW3) [2], a certain subset of bones is examined.

In order to relieve the radiologist, several approaches have been made to (partially) automate the process [3-7] and recently a commercial application has been introduced [7]. All require a reliable localization and delineation of bone segments. Yet it is well known that automatic segmentation of medical images is error-prone and requires interactive solutions [8] and that previous approaches have only limited abilities to construct the bone borders [7]. Concerning hand radiographs, one can easily identify the following problems:

1. The summation effect projecting 3D structures onto 2D images yields irritations in the segmentation. This affects especially the area between metacarpal and carpal bones, but also regions of bones overlapping in the projection.
2. Varying radiation dose and noise add difficulties for the training of segmentation algorithms.
3. The exact hand pose, aperture, e.g., including radius or ulna, and resolution are not standardized.
4. The differences between reference groups of identical chronological or bone age, gender, and ethnic origin, are high, so that individual methods or models need to be developed (Fig. 1).
5. High variation may frequently occur even within a certain reference group [1,9]. Hence, it becomes difficult to gather enough cases for a statistically sound training.

[*]Corresponding author: Dipl.-Inform. Benedikt Fischer. Department of Medical Informatics, RWTH Aachen University, Pauwelsstr. 30, 52057 Aachen, Germany, email: bfischer@mi.rwth-aachen.de; phone: +49 241 80 85174, fax: +49 241 80 33 5174

Medical Imaging 2011: Computer-Aided Diagnosis, edited by Ronald M. Summers, Bram van Ginneken,
Proc. of SPIE Vol. 7963, 79630P · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.878839

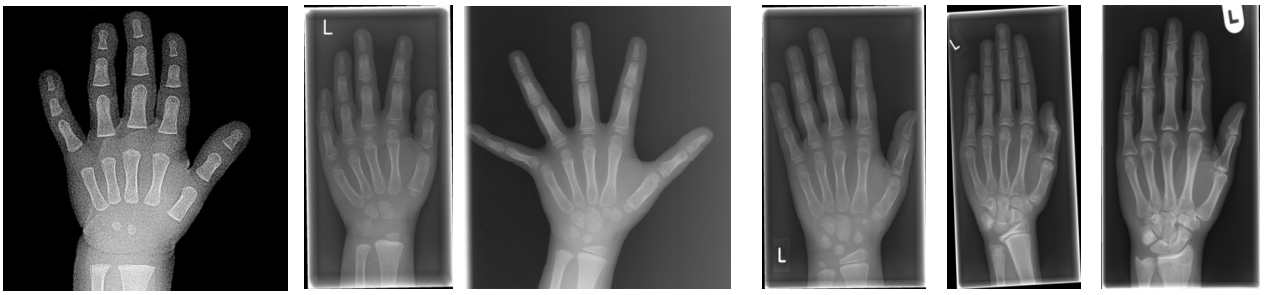

Figure 1: Variances in skeletal development: Shown are successive stages of skeletal development of healthy male subjects. The radiographs have been acquired during radiological routine at the University Hospital Aachen, Germany.

Even when setting any segmentation problems aside, existing approaches to BAA do not always provide plausible results. A radiologist should not have to rely on a black box solution, even when complex algorithms turn out with a good solution but the results are incomprehensible and hindering evidence-based medicine. We therefore not only aim at avoiding the problems of segmentation but we also strive for providing fully comprehensible results.

We achieve this by computing the bone age from similarities to earlier cases. The similarities are obtained by content-based image retrieval (CBIR), and the radiologist is provided visually with what our system regards as the best correspondences in the database. The radiologist then can directly estimate the usefulness of the offered second opinion. This approach is consistent with the case-based reasoning (CBR) paradigm which has been frequently applied in medical decision support [10] and has already come to attention for medical applications in conjunction with CBIR [11, 12].

Based on the image retrieval in medical applications (IRMA) framework (http://irma-project.org), we propose a web-based CAD tool for bone age assessment.

2. METHODS

The methodology for the CBIR-based bone age assessment is based on comparing image content from a new case to earlier cases. All clinically used BAA methods refer to the epiphyseal area between the bones. We therefore extract those areas as epiphyseal regions of interest (eROIS) of each image in a standardized way. Instead of applying a query on the complete image, every eROI is used for an individual query to the database. The BAA process therefore can be performed in four stages (Fig. 2), [13]:

1. The centers of 14 epiphyseal regions are located in the image either automatically or manually.
2. The eROIs are extracted, geometrically oriented into an upright position, and uniformly scaled.
3. Each eROI is used for individual queries to the retrieval engine. For each eROI, a list of best matching correspondences is retrieved from the case database together with their age and similarity score.
4. From these, the respective ages are accumulated at first for each eROI and then for an overall age prediction.

2.1 Center Localization and eROI Extraction

The eROI centers can either be determined automatically [14] or set manually by a customized web interface. If set manually, the user also defines the top and bottom of the hand by clicking above the topmost fingertip and the top of the ulna to ensure proper scaling for the subsequent comparisons. Details on orientation, scaling and extraction of the eROIs are provided in [13].

2.2 Case Comparison

Once the eROIs have been extracted, each one is subjected to an individual query to the case database to retrieve the most similar reference cases. As each eROI is labeled by its position, only corresponding eROIs are compared, e.g. eROIs of position 9 (distal eROI of the middle finger, see Fig. 5c) are only compared to other eROIs with the same position-label. For the similarity computation we use a combination of the cross-correlation function (CCF), the image distortion model (IDM) [15,16] and texture features proposed by Tamura et al. [17].

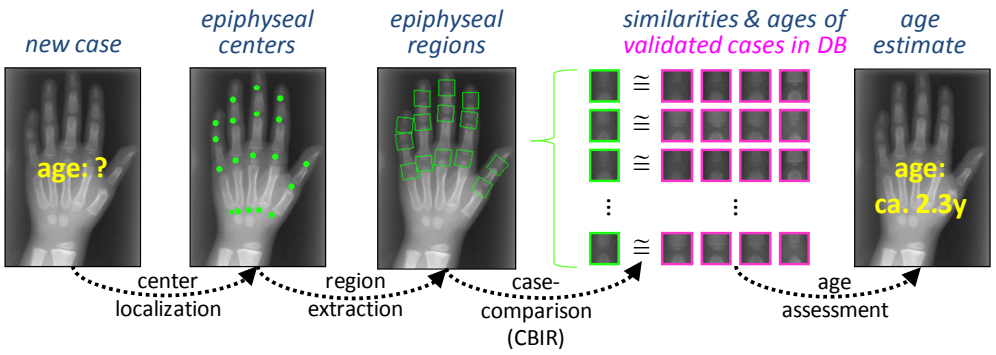

Figure 2: CBIR approach to bone-age assessment: For a hand of unknown age, the epiphyseal centers are localized and the corresponding eROIs are compared to the references in the database. The ages of the *K* most similar ones for each eROI yield the predicted age.

In contrast to other similarity measures, the CCF is easy to compute and robust regarding the radiation dose. In addition, translation within the search window is normalized. As described above, our procedure of eROI extraction ensures invariance to orientation and scaling. The similarity of a query eROI q to an eROI p in the retrieval database is therefore computed by:

$$s_{\mathrm{CCF}}(p,q) = \max_{|m|,|n|\le d} \left\{ \frac{\sum_{x=1}^{X}\sum_{y=1}^{Y}\left(p(x-m,y-n)-\bar{p}\right)\left(q(x,y)-\bar{q}\right)}{\sqrt{\left(\sum_{x=1}^{X}\sum_{y=1}^{Y}\left(p(x-m,y-n)-\bar{p}\right)^2\right)\cdot\left(\sum_{x=1}^{X}\sum_{y=1}^{Y}\left(q(x,y)-\bar{q}\right)^2\right)}} \right\} \tag{1}$$

The variables $\bar{p}$ and $\bar{q}$ denote the respective mean gray values of p and q. In our experiments, we use 16 x 16 scaled versions of the eROIs, i.e., X=Y=16, and a warp range of d=2 to determine the maximum correlation.

While the CCF considers global displacements, the image distortion model (IDM) is able to model local deformations within an image as a distance [15]:

$$d_{\mathrm{IDM}}(p,q) = \sum_{x=1}^{X}\sum_{y=1}^{Y} \min_{|x'|,|y'|\le W_1} \left\{ \sum_{|x''|,|y''|\le W_2} \left\| \left(p(x+x'+x'', y+y'+y'') - q(x+x'', y+y'')\right) \right\|_2 \right\} \tag{2}$$

Each pixel of the database query eROI q is mapped onto some pixel in p, but the opposite does not necessarily have to be the case. W_1 defines the size of the search window. To prevent a totally unordered pixel mapping, the local neighborhood is incorporated as a context when evaluating the correspondence hypothesis [16]. The size of this context within the search window is controlled by W_2.

Tamura et al. proposed a set of texture features to reflect coarseness, contrast, and directionality which is stored in a three-dimensional histogram quantized into M = 6 x 8 x 8 = 384 bins [17]. Two eROIs p and q can then be compared by applying the Jensen-Shannon divergence [18] on their histograms $H(p)$ and $H(q)$:

$$d_{\mathrm{Tamura}}(p,q) = \frac{1}{2}\sum_{m=1}^{M}\left[H_m(q)\log\frac{2H_m(q)}{H_m(q)+H_m(p)} + H_m(p)\log\frac{2H_m(p)}{H_m(q)+H_m(p)} \right] \tag{3}$$

The combination of CCF, IDM, and Tamura features is achieved by their weighted sum after normalizing the values. The normalization for a distance measure $d(q,p)$ such as d_{IDM} and d_{Tamura} is computed over all similarities $d'(p_i,q)$, i=1,…,N between query eROI q and all database eROIs p_i:

$$d'(p_i,q) = \frac{d(p_i,q)}{\sum_{n=1}^{N} d(p_n,q)} \tag{4}$$

Since the CCF is not a distance measure but a similarity, $d(p,q) := 1 - s_{CCF} =: d_{CCF}'(p,q)$ is used in (4) for its normalization. The combined distance measure then results in:

$$d_{\text{combined}}(p,q) = \lambda_{\text{CCF}} \cdot d_{\text{CCF}}'(p,q) + \lambda_{\text{IDM}} \cdot d_{\text{IDM}}'(p,q) + \lambda_{\text{Tamura}} \cdot d_{\text{Tamura}}'(p,q), \text{ where } \lambda_{\text{CCF}} + \lambda_{\text{IDM}} + \lambda_{\text{Tamura}} = 1 \quad (5)$$

In order to be able to preserve our previously published age estimation (7), we transform (5) back into a similarity measure which additionally can stress the importance of more similar eROIs by the exponent α:

$$s_{\text{combined}}(p,q) = \frac{1}{\left(d_{\text{combined}}(p,q)\right)^{\alpha}} \quad (6)$$

2.3 Age Estimation

The database delivers the K most similar correspondences for each eROI r in a sorted list $\partial(r,k)$ for $k = \{1,\ldots,K\}$ and the known BAA as $a^{\text{known}}(k,r)$. The implemented age prediction weighs each eROI r of R regarded positions with identical influence. The ages of K most similar eROIs of known cases are weighed according to their similarity which has been obtained from (6):

$$a^{\text{predict}} = \frac{1}{R}\sum_{r=1}^{R}\left(\frac{1}{\sum_{k=1}^{K}\partial(k,r)}\sum_{k=1}^{K} a^{\text{known}}(k,r)\cdot\partial(k,r)\right) \quad (7)$$

2.4 Validation Experiments

The retrieval database contains a set of 1,102 hand radiographs from the USC database along with the patients' chronological ages and two maturity readings by independent radiologists [6]. For the experiments, manually localized center coordinates of the eROIs are used. As a quality measure, the mean and variance of the absolute difference between the two radiologists' average readings and the bone age as predicted by (7) is computed.

All parameter settings are based on the experiments from [13]: $K=3$ most similar cases are considered, the choice of eROI positions is set to the six eROIs 11, 15, 7, 18, 3, and 10 (Fig. 5c). The CCF- and IDM-computations are performed on 16 x 16 versions of the eROIs, while the texture features are computed on patches of 32 x 32 pixels. The parameters for CCF and IDM are set to d=2, W_1=2, W_2=1.

In the experiments, we analyze the influence of the similarity exponent α in (6) and different weights for CCF, IDM, and Tamura in (5). All experiments are run in leaving-one-out manner with the USC hand atlas as the underlying ground truth.

The prediction quality of our CBIR-based BAA system is measured by the mean absolute prediction error in comparison to the average of the two readings from the USC database and the standard deviation:

$$\mu_{\text{err}} = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{a^{\text{read1}}(n) + a^{\text{read2}}(n)}{2} - a^{\text{predict}}(n)\right|$$
$$\sigma^2_{\text{err}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}\left(\left|\frac{a^{\text{read1}}(n) + a^{\text{read2}}(n)}{2} - a^{\text{predict}}(n)\right| - \mu_{\text{err}}\right)^2} \quad (8)$$

Here, N denotes the number of hand radiographs available for comparison.

2.5 Web Interfaces

The integration of CAD systems in clinical workflow is facilitated by web-based systems and investigations have shown that physicians prefer web search engines and web-based interfaces to retrieve interesting image data from archives [19,20]. Our web-based system for BAA-CAD features two operating modes:

- In the "demonstration mode", the user browses through available radiographs from the reference database prior to selecting one, which is then analyzed using the remaining ground truth data.

- In the "local mode", the user can upload an appropriate radiograph from a local system to the web server and a BAA estimate is returned as a second opinion. Again, the USC data is used as ground truth.

An overview of the complete web-based scenario is illustrated in Fig. 3. The implementation utilizes the IRMA framework. This provides Smarty templates for web-based GUI design and a large functionality for extended query refinement and relevance feedback including complete query logging of all user/system-interaction [20]. PHP is used to interface a PostgreSQL database with an Apache web server on a Linux-based server PC running a 64bit Debian Linux.

3. RESULTS

3.1 Validation experiments

The results of the validation experiments are summarized in Tab. 1. Further results with intermediate steps on α have been computed but are omitted from the table as they do not contribute to the understanding. The overall best results of a mean absolute error of 0.9655 and 0.9659 are obtained by the parameters (α,λ_{CCF}, λ_{IDM},λ_{Tamura}) = (22, 0.6, 0.4, 0.0) and (32, 0.6, 0.4, 0), respectively. With respect to the parameter α, results improve from α=2 to α=22 and diminish afterwards.

Concerning the individual similarity weightings, i.e., one of the λ-values is set to 1, d_{CCF} delivers the best results for all values of α. The range is approximately 0.97 to 0.99. d_{Tamura} shows values between 1.58 and 1.61, while d_{IDM} shows the poorest performance between 3.15 and 3.18. For a mixed weighting, d_{CCF} is most important for smaller values of α, then d_{IDM} gains more importance while d_{Tamura} is disregarded throughout the course.

3.2 Web interface

The web interface for the presented research prototype is publicly accessible[1]. The graphical user interface (GUI) is partitioned horizontally into distinct sections such as header bar, parameter field, status and navigation bar, and output field, where the icons of available images are displayed. This is consistent with all other IRMA interfaces and ensures easy orientation switching between interfaces [21].

α	λ_{CCF}	λ_{IDM}	$\lambda_{Tam.}$	μ	σ^2	α	λ_{CCF}	λ_{IDM}	$\lambda_{Tam.}$	μ	σ^2	α	λ_{CCF}	λ_{IDM}	$\lambda_{Tam.}$	μ	σ^2	α	λ_{CCF}	λ_{IDM}	$\lambda_{Tam.}$	μ	σ^2
2	0.0	0.0	1.0	1.595	1.693	12	0.0	0.0	1.0	1.581	1.751	22	0.0	0.0	1.0	1.597	1.839	32	0.0	0.0	1.0	1.611	1.886
2	0.0	0.2	0.8	1.608	1.672	12	0.0	0.2	0.8	1.580	1.671	22	0.0	0.2	0.8	1.578	1.702	32	0.0	0.2	0.8	1.583	1.748
2	0.0	0.4	0.6	1.602	1.662	12	0.0	0.4	0.6	1.583	1.654	22	0.0	0.4	0.6	1.572	1.659	32	0.0	0.4	0.6	1.569	1.670
2	0.0	0.6	0.4	1.596	1.662	12	0.0	0.6	0.4	1.584	1.655	22	0.0	0.6	0.4	1.575	1.650	32	0.0	0.6	0.4	1.568	1.649
2	0.0	0.8	0.2	1.580	1.597	12	0.0	0.8	0.2	1.576	1.598	22	0.0	0.8	0.2	1.572	1.599	32	0.0	0.8	0.2	1.568	1.600
2	0.0	1.0	0.0	3.184	4.408	12	0.0	1.0	0.0	3.169	4.390	22	0.0	1.0	0.0	3.157	4.368	32	0.0	1.0	0.0	3.145	4.350
2	0.2	0.0	0.8	1.152	0.748	12	0.2	0.0	0.8	1.153	0.756	22	0.2	0.0	0.8	1.171	0.787	32	0.2	0.0	0.8	1.183	0.811
2	0.2	0.2	0.6	1.133	0.712	12	0.2	0.2	0.6	1.119	0.697	22	0.2	0.2	0.6	1.123	0.707	32	0.2	0.2	0.6	1.130	0.725
2	0.2	0.4	0.4	1.094	0.659	12	0.2	0.4	0.4	1.081	0.644	22	0.2	0.4	0.4	1.076	0.641	32	0.2	0.4	0.4	1.077	0.647
2	0.2	0.6	0.2	1.047	0.596	12	0.2	0.6	0.2	1.039	0.591	22	0.2	0.6	0.2	1.982	2.072	32	0.2	0.6	0.2	1.033	0.592
2	0.2	0.8	0.0	0.986	0.628	12	0.2	0.8	0.0	0.982	0.627	22	0.2	0.8	0.0	0.978	0.626	32	0.2	0.8	0.0	0.975	0.625
2	0.3	0.3	0.3	1.041	0.594	12	0.3	0.3	0.3	1.030	0.586	22	0.3	0.3	0.3	1.032	0.591	32	0.3	0.3	0.3	1.039	0.600
2	0.4	0.0	0.6	1.062	0.613	12	0.4	0.0	0.6	1.064	0.630	22	0.4	0.0	0.6	1.080	0.656	32	0.4	0.0	0.6	1.092	0.671
2	0.4	0.2	0.4	1.039	0.597	12	0.4	0.2	0.4	1.031	0.591	22	0.4	0.2	0.4	1.041	0.602	32	0.4	0.2	0.4	1.052	0.614
2	0.4	0.4	0.2	1.017	0.544	12	0.4	0.4	0.2	1.004	0.538	22	0.4	0.4	0.2	1.001	0.544	32	0.4	0.4	0.2	1.004	0.552
2	0.4	0.6	0.0	0.984	0.630	12	0.4	0.6	0.0	0.976	0.627	22	0.4	0.6	0.0	0.969	0.626	32	0.4	0.6	0.0	0.966	0.627
2	0.6	0.0	0.4	1.023	0.559	12	0.6	0.0	0.4	1.027	0.579	22	0.6	0.0	0.4	1.044	0.602	32	0.6	0.0	0.4	1.054	0.614
2	0.6	0.2	0.2	1.003	0.519	12	0.6	0.2	0.2	0.992	0.521	22	0.6	0.2	0.2	0.998	0.540	32	0.6	0.2	0.2	1.006	0.554
2	0.6	0.4	0.0	0.982	0.630	12	0.6	0.4	0.0	0.969	0.626	22	0.6	0.4	0.0	0.966	0.629	32	0.6	0.4	0.0	0.967	0.634
2	0.8	0.0	0.2	0.996	0.515	12	0.8	0.0	0.2	0.994	0.534	22	0.8	0.0	0.2	1.008	0.556	32	0.8	0.0	0.2	1.016	0.572
2	0.8	0.2	0.0	0.982	0.628	12	0.8	0.2	0.0	0.967	0.627	22	0.8	0.2	0.0	0.970	0.639	32	0.8	0.2	0.0	0.975	0.652
2	1.0	0.0	0.0	0.974	0.623	12	1.0	0.0	0.0	0.972	0.640	22	1.0	0.0	0.0	0.983	0.664	32	1.0	0.0	0.0	0.992	0.678

Table 1: Experiment outcome for varying similarity exponent and different feature weights. The best result for each block is indicated by a box, the best global result is highlighted.

[1] http://irma-project.org/onlinedemos_en.php

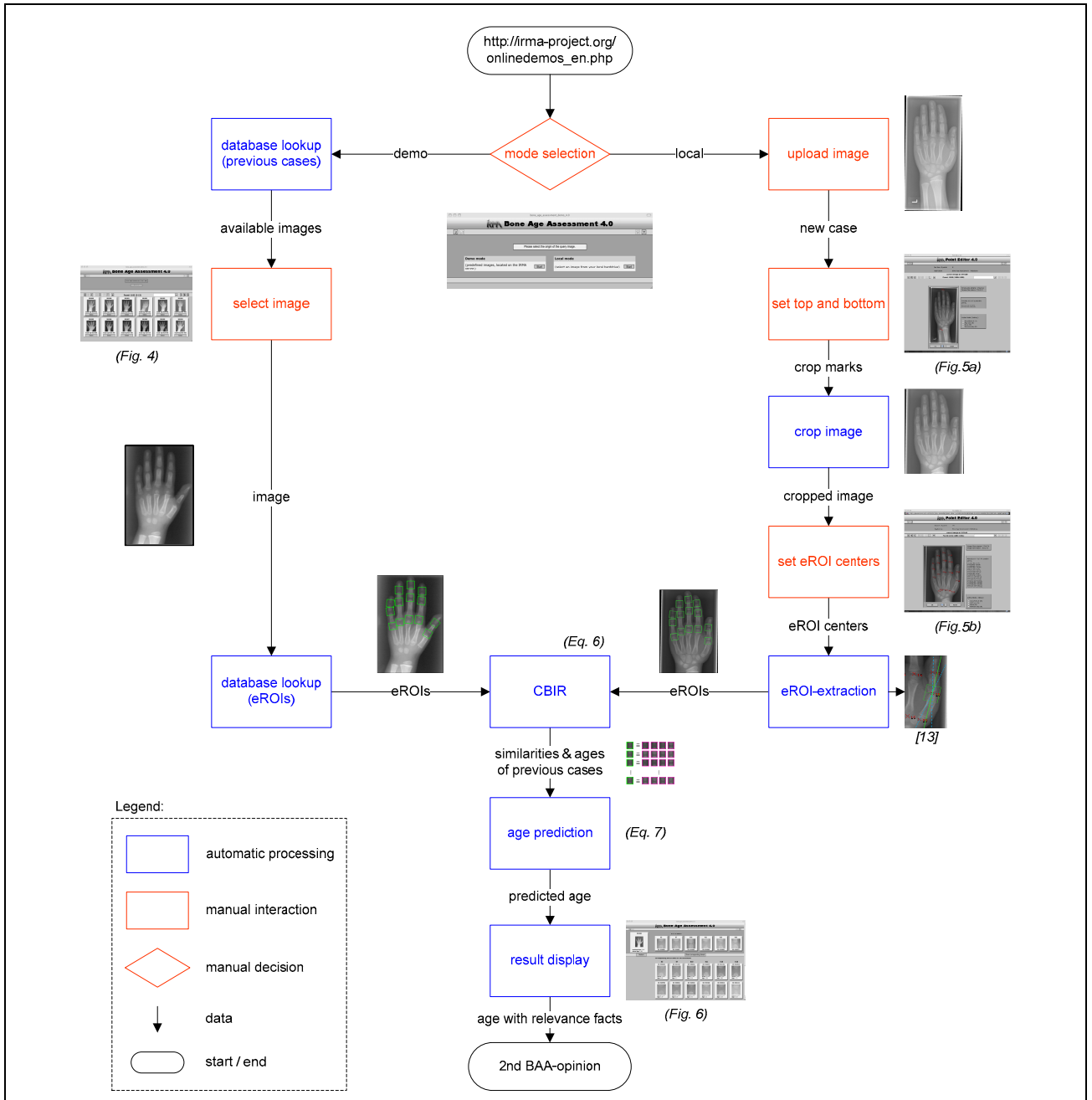

Figure 3: Complete processing pipeline for the web-based BAA: Both modes compute similarities to the cases of the USC hand atlas.

The user either selects a hand radiograph from the database in "Demonstration mode" (Fig. 4) or uploads a new image in the "Local Mode" from the user's local computer. A new image is at first cropped from the distal fingertip to the ulna by two clicks (Fig. 5a). The same interface is used to localize the eROI centers in a new image (Fig. 5b). The eROIs are then extracted automatically, subjected to the CBIR engine and the bone age is estimated.

In both operating modes, the result is displayed with the original image and the extracted eROIs at the top of the browser window. The *K* most similar eROIs for each extracted eROI are shown below in decreasing order of similarity (Fig. 6).

Figure 4: Example screen for browsing the available database before selecting an image for analysis. A detailed view of a hand radiograph is available by clicking on the thumbnails.

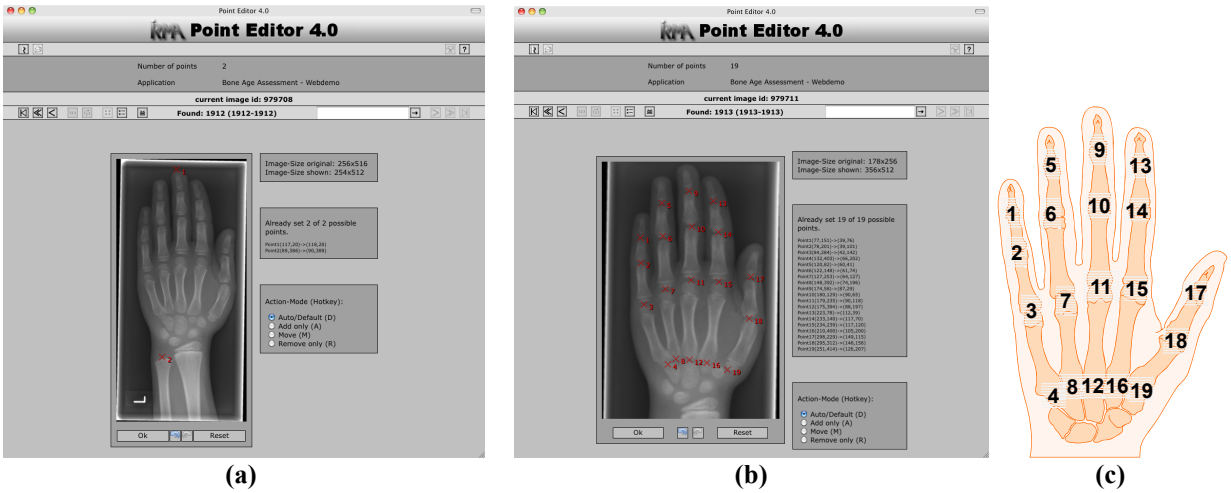

Figure 5: Point Editor interface for setting the top and bottom crop marks (a) and labeling the epiphyseal centers (b), and the labels used throughout the experiments and in the result display (Fig. 6) (c).

Each column corresponds to one eROI position. For each retrieved eROI, the computed similarity to the query eROI as well as the recorded bone age is shown.

For the website, the number of retrieved most similar eROIs has been set to K=10 to provide a better overview than using only K=3. The number of eROI positions has been pre-selected to six (eROI no. 11, 15, 7, 18, 3, and 10 in Fig. 5c), which corresponds to the above experiments and provides a good fit into the typical display resolutions with 1,280 pixels width. All shown thumbnails are clickable to open a new window with the original resolution for detailed inspection.

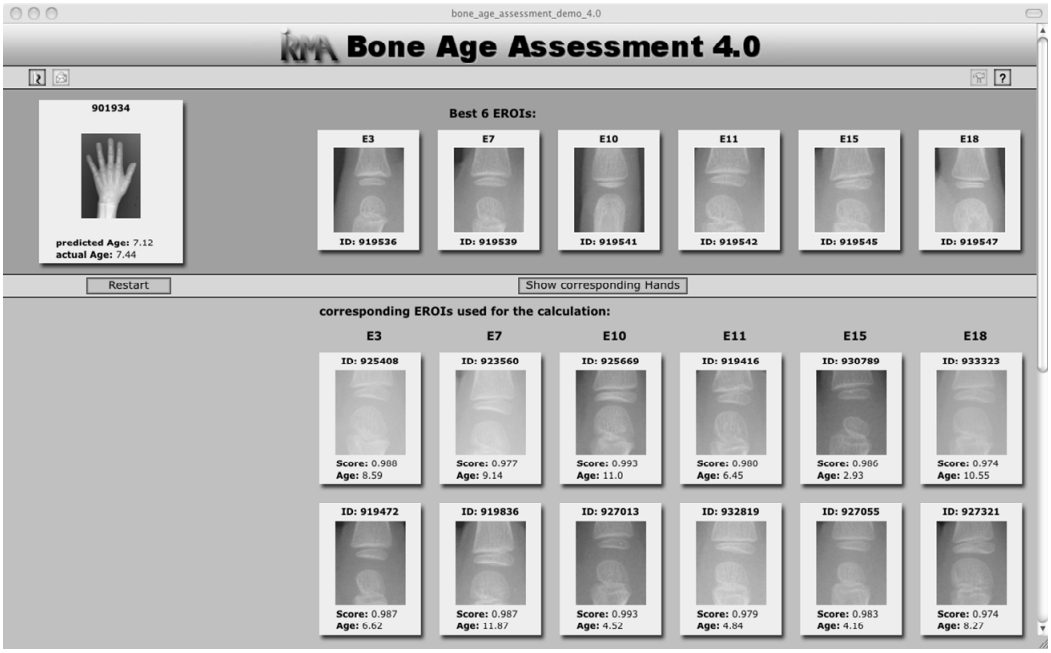

Fig. 6: Result display in demonstration mode for six eROIs. Query image and extracted eROIs are shown at the top. Their most similar counterparts retrieved from the database are shown below (scrollable) in decreasing similarity and with the previously diagnosed bone age. The estimated and the known bone age are shown right below the query image. In the local mode with a newly uploaded image, the known bone age is not available.

4. DISCUSSION

The extension of the similarity computation by the combination of CCF, IDM and Tamura features and the similarity exponent α have improved the quality measure in comparison to the previously published results [13].

The CCF alone almost reaches the performance of the optimal set combination. If only two decimal places are used, the results are even identical at an absolute mean prediction error of 0.97. Interestingly, the IDM, which individually does not perform very well, leads to a better error rate in conjunction with the CCF. On the contrary, the Tamura features do not add to the prediction quality of a set even if they have only half of the error rate of IDM if taken individually. The results for the combination of IDM and CCF with increasing α maybe a hint, that the linear normalization of CCF and IDM as in (4) is not just. With increasing α, IDM is gaining importance in the combination and improves individually, while CCF results are going in the opposite direction. Yet higher values of α have not been computed and are not considered to be useful because of the limitations in computation accuracy. Nevertheless, the introduction of α can be seen as successful as it improves the estimation results and the best overall set has been reached with α=22.

In contrast to other approaches, our web-based system visually provides relevance facts, i.e. the previous cases that are considered most similar for each region. This makes the prediction comprehensible and understandable to the user. The research prototype is available to the scientific community on the internet, and we are hoping to gain important feedback from independent users. The system's performance can be tested within the USC database or on new radiographs uploaded by the user. Being a prototype, the web-based interface still suffers from a number of drawbacks: so far, only uploads of PNG images of at least 256 pixels height are supported and the number of required mouse-clicks still needs optimization to increase the usability of the system.

Our results of a mean error of below one year may not yet have reached the accuracy of the commercial system BoneXpert [7]. But keeping in mind, that the maximum inter-observer difference for the USC readings is 2.5 years and variations of the appearance of two years are considered natural [22], and considering our superior relevance feedback

mechanism, our results suggest readiness for further evaluation in clinical settings. For this, 800 cases taken from daily routine at the University Hospital Aachen are currently being prepared in terms of ground truth and eROI centers. We will then evaluate the performance of a physician with CBIR-CAD against a physician without CBIR-CAD – rather than a physician vs. CBIR-CAD. Additional similarity computations on age prototypes as well as the extension of the web frontend by feedback mechanisms will enhance the age prediction further.

5. CONCLUSIONS

We have presented a novel method for bone age assessment that is based on image content comparison. Avoiding the error-prone segmentation that is required with all other methods of automated BAA, we provide the radiologist a second opinion rather than a machine-based "diagnosis". The system, however, was evaluated on the computer-aided diagnosis paradigm and has proven sufficient reliability for routine testing. We have introduced a web-based interface to foster integration into the radiological workflow.

ACKNOWLEDGEMENT

This research was supported (in part) by the German Research Foundation (DFG), grant no. Le 1108/9.

REFERENCES

[1] Greulich WW, Pyle SI. Radiographic atlas of skeletal development of hand wrist. Stanford University Press, Stanford CA. 1971.
[2] Tanner JM, Healy MRJ, Goldstein H, Cameron N. Assessment of skeletal maturity and prediction of adult height (TW3). WB Saunders, London. 2001
[3] Martin-Fernandez MA, Martin-Fernandez M, Alberola-Lopez C. Automatic bone age assessment: A registration approach. Proc SPIE. 2003; 5032: 1765–76
[4] Pietka E, Gertych A, Pospiech S, et al. Computer-assisted bone age assessment: Image preprocessing and epiphyseal/metaphyseal ROI extraction. IEEE Trans Med Imaging 2001; 20(8): 715-29
[5] Park KH, Lee JM, Kim WY. Robust epiphyseal extraction method based on horizontal profile analysis of finger images. Proc ISITC 2007; 278–82
[6] Gertych A, Zhang A, Sayre J, et al. Bone age assessment of children using a digital hand atlas. Computerized Medical Imaging and Graphics. 2007; 31(4-5):322–31
[7] Thodberg HH, Kreiborg S, Juul A, Pedersen KD. The BoneXpert method for automated determination of skeletal maturity. IEEE Trans Med Imaging 2009;28(1):52-66
[8] Olabarriaga SD, Smeulders AWM (1997) Setting the mind for intelligent interactive segmentation: overview, requirements, and framework. Proc IPMI: 417–422
[9] Gilsanz V, Ratib O. Hand bone age (2005) A digital atlas of skeletal maturity. Springer Verlag Berlin Heidelberg
[10] Berner ES, McGowan JJ (2010) Use of diagnostic decision support systems in medical education. Methods Inf Med; 49(4):412-7
[11] Muller H, Michoux N, Bandon D, Geissbuhler A (2004) A review of content-based image retrieval systems in medical applications. Clinical benefits and future directions. Int J Med Inform;73(1):1-23
[12] Eakins JP, Graham ME (1999) Content-based image retrieval -- A report to the JISC technology applications programme. Technical Report, Institute for Image Data Research, University of Northumbria at Newcastle, http://www.jisc.ac.uk/uploaded_documents/jtap-039.doc Accessed 28 January 2010
[13] Fischer B, Brosig A, Welter P, Grouls C, Guenther RW, Deserno TM. Content-based image retrieval applied to bone age assessment. Proc SPIE 2010; 7624. DOI: 10.1117/12.844392
[14] Fischer B, Brosig A, Deserno TM, Ott B, Günther RW. Structural scene analysis and content-based image retrieval applied to bone age assessment. Proc SPIE 2009; 7260: 041-11
[15] Keysers D, Dahmen J, Theiner T, Ney H. Experiments with an Extended Tangent Distance. Proc ICPR 2000, 2:38-42
[16] Güld MO, Thies C, Fischer B, Lehmann TM. Combining global features for content-based retrieval of medical images. In: Peters C, Quochi V (eds): Results of the CLEF 2005 Cross-Language System Evaluation Campaign - Electronic Working Notes for the CLEF 2005 Workshop, 21-23 September, Vienna, Austria, 2005
[17] Tamura H, Mori S, Yamawaki T. Textural features corresponding to visual perception. IEEE Transactions on Systems, Man, and Cybernetics; SMC-8(6), 460-472, 1978
[18] Puzicha J, Rubner Y, Tomasi C, Buhmann J. Empirical evaluation of dissimilarity measures for color and texture. Proceedings International Conference on Computer Vision, 2, 1165-1173, 1999
[19] Hersh W, Muller H, Gorman P, Jensen J (2005) Task analysis for evaluating image retrieval systems in the ImageCLEF biomedical Image Retrieval Task. Proc SOL, Portland, OR

[20] Muller H, Despont-Gros C, Hersh W et al. (2006) Health care professionals' image use and search behaviour. Proc MIE; 24-32
[21] Deserno TM, Güld MO, Plodowski B, Spitzer K, Wein BB, Schubert H, Ney H, Seidl T (2008) Extended query refinement for medical image retrieval. J Digit Imaging 2008; 21(3): 280-9
[22] Zhang A. A computer-aided-diagnosis (CAD) method combining phalangeal and carpal bone features for bone age assessment of children. Dissertation at University of Southern California; 2007