Automatic recognition of surface landmarks of anatomical structures of back and posture

Jakub Michoński
Wojciech Glinkowski
Marcin Witkowski
Robert Sitnik

SPIE

Automatic recognition of surface landmarks of anatomical structures of back and posture

Jakub Michoński,[a] Wojciech Glinkowski,[b] Marcin Witkowski,[a] and Robert Sitnik[a]
[a]Warsaw University of Technology, Institute of Micromechanics and Photonics, ul. Sw. A. Boboli 8, 02-525 Warsaw, Poland
[b]Medical University of Warsaw, Department of Orthopaedics and Traumatology of Locomotor System, Center of Excellence, "TeleOrto", ul. Lindleya 4, 02-005, and Polish Telemedicine Society, Warsaw, Poland

Abstract. Faulty postures, scoliosis and sagittal plane deformities should be detected as early as possible to apply preventive and treatment measures against major clinical consequences. To support documentation of the severity of deformity and diminish x-ray exposures, several solutions utilizing analysis of back surface topography data were introduced. A novel approach to automatic recognition and localization of anatomical landmarks of the human back is presented that may provide more repeatable results and speed up the whole procedure. The algorithm was designed as a two-step process involving a statistical model built upon expert knowledge and analysis of three-dimensional back surface shape data. Voronoi diagram is used to connect mean geometric relations, which provide a first approximation of the positions, with surface curvature distribution, which further guides the recognition process and gives final locations of landmarks. Positions obtained using the developed algorithms are validated with respect to accuracy of manual landmark indication by experts. Preliminary validation proved that the landmarks were localized correctly, with accuracy depending mostly on the characteristics of a given structure. It was concluded that recognition should mainly take into account the shape of the back surface, putting as little emphasis on the statistical approximation as possible. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.JBO.17.5.056015]

1 Introduction

Spinal deformities affect all age groups, leading to the most frequent postural and pain problems in modern societies.[1–4] Back pain is one of the diseases, which usually lowers the quality of life.[5–8] The adult deformity may relate to pain[9] and/or compression fractures.[10] Faulty postures and associated disorders should be detected as early as possible to apply preventive measures against major consequences in old age.[11,12]

Back examination is performed by physicians based mainly on the observation of the body, sometimes followed by simple linear or other more precise measurements.[13–16] Radiography is carried out in the most suspected cases[16,17]. It provides a widely accepted method of measurement, analysis, documentation and follow-up of clinical cases. However, in most screened cases the aforementioned physical examination is the only type of documentation available and is inevitably subjective. This can lead to errors during comparison of a patient's spine condition in time or even when assessed by another physician. Radiography remains the main method for periodical evaluation of deformity changes during treatment. This implies intensified, repeated x-ray exposure, which should be avoided in general, especially in adolescents.[18–20]

1.1 Deformity Assessment using Back Surface Topography

A few systems for back shape assessment already exist, allowing the accomplishment of the task of documentation. The most of them are marker-dependent or markerless optical systems, based mainly on the moiré technique, laser and structured light illumination (SLI), incorporating assessment methods which analyze mutual position of certain anatomical structures.[21–32]

Back shape indices calculated using mentioned systems were proven to be useful and reliable in clinical settings.[25,33] The conformity of the parameters with the gold standard used in radiography, the Cobb angle measurement, is quite high, e.g., the r-Pearson correlation coefficient was estimated to 0.801 for the Quantec Q-angle parameter[34] and to 0.668 and 0.706 for the posterior trunk symmetry index (POTSI) and deformity in the axial plane index (DAPI).[22]

The approach used in creating these parameters is approximately equivalent to the physical examination, where the physician mainly evaluates the broadly understood symmetry of the back. The vast majority of diagnostic systems are based on manual indication of anatomical back surface landmarks[28,35–37] by operator's palpation. Palpated landmarks are often considered as a principle for quantitative evaluation of deformation, however, landmarks pointed by inspection are accepted and reliable.[28,38]

The aim of the presented algorithm is to provide a more reliable way of anatomical back shape landmarks localization by automatic markerless analysis of the back shape surface. The algorithm was developed as a straightforward generalization of the existing manual deformation assessment methods used in clinical conditions. In the following sections, a detailed description of the algorithm and its preliminary validation are given to show that the accuracy of the proposed solution is comparable with accuracy of the existing methods for given cases (where accuracy is defined as difference with respect to average manual marking and is compared to uncertainty of manual

Address all correspondence to: Wojciech Glinkowski, Medical University of Warsaw, Department of Orthopaedics and Traumatology of Locomotor System, Center of Excellence, "TeleOrto", ul. Lindleya 4, 02-005, and Polish Telemedicine Society, Warsaw, Poland. Tel: +601230577; Fax: +48226190013; E-mail: w.glinkowski@gmail.com

indication by the experts). Further validation for a more diversified group of subjects, including different criteria, will be conducted in the future as indicated in Sec. 5.

1.2 Existing Methods of Automatic Shape Analysis

The methods that deal with automatic shape analysis are usually very much application-dependent, although significant effort has been put into generalizing this problem. A significant factor is the character of data representation used, which in this case is three-dimensional, though it may be approximated as a surface of the form $[x, y, f(x, y)]$ with respect to the frontal plane of the body. This makes it possible to utilize methods intended for two-dimensional data as well, if there is such necessity.

It should be mentioned that several methods exist which deal with "three-dimensional segmentation of anatomical structures,"[39–41] they are nevertheless intended to work with volumetric three-dimensional data, such as data from MRI or CT, and thus are of no use in cases where structures have to be located on an explicitly given surface. It is much easier to generalize the two-dimensional methods into three-dimensional space in most cases.

A separate class of methods deals with broadly understood anatomical shape characterization, but their main goal is finding the best way (i.e., with maximized inter-class variability) to describe a specific object type, usually by indicating landmarks on the surface of the object.[42] Such landmarks are then used for analysis of a given class of objects. This approach cannot be implemented in this case as the set of anatomically meaningful landmarks is given explicitly and has to be recovered in the recognition process. Moreover, some of the methods which could provide a meaningful representation do not support the type of data topology used in our study (e.g., spherical harmonics with point distribution model, SPHARM-PDM, which can only work on data with spherical topology).

The active shape model[43] represents a class of methods which can be used to build a generalized model of the shape based on a learning set of shapes with indicated characteristic points. This model can be further utilized to find instances of new shapes, controlling the extent to what they can deviate from the model using eigen-analysis of the set. The method was originally developed to work with two-dimensional images of arbitrary descent. The point distribution model (PDM), which is used to build the statistical correspondence, can be easily generalized to match the three-dimensional character of input data by adding the third coordinate during analysis. Further generalization, however, is impeded by the fact that the structures have to be found on an explicitly given surface, which provides a strict constraint for the search procedure. It is valid for volumetric data where the model can be fully deformed in all three dimensions.

Moreover, there are doubts whether using statistical correspondence to limit the shape domain would work well, particularly if the final objective is to detect anomalies in the back structure, with equal sensitivity to all kinds of deformations. The author of the ASM technique warns against model "over-training," i.e., providing a training set which is very coherent and does not provide all the variations, because it would prevent the method from finding new shapes which are valid in reality, although do not match the statistically built model. Moreover, since deformed structure of the back constitutes a small percentage of the whole population (even more so considering rare deformations), these modes of varia-

tions could be unintentionally neglected during model building using principal component analysis (PCA).

The same stipulation concerns the active appearance model (AAM) technique.[44] Here, even more statistical constraints exist, as the texture of the image is also taken into account during the statistical model building. This method was also developed for images. It uses a matrix form for the intensity model, particularly because PCA is a linear method. Using this technique would demand generating "texture matrices" from unorganized point clouds for any feature considered as texture – depth, curvature or other – and performing all the operations in the XY projection. Moreover, it is arguable whether global optimization of the model which is utilized by this method when searching for the new model is the right approach in this case.

Summarizing the described approaches, although providing valuable insights into data analysis, they do not solve the given problem well. These general tools may act more as a drawback than as an advantage. The ASM technique, along with the congener active contour model (ACM) technique were used in Ref. 45 to model back and spine variability with moderate success, however, the aspect of landmark recognition was of secondary importance in the mentioned work and did not address the stipulations with respect to some aspects of this method shown here. To overcome some of them, a handcrafted solution was attempted, which uses some of the ideas provided above. However, in future versions the algorithm may incorporate more aspects of the described methods, as presented in Sec. 5.

2 Methodology

2.1 Input Data Characterization

In real application the 3-D measurement with algorithms of directional merging and conversion (3DMADMAC) system environment was utilized, a complete measurement and calculation environment for three-dimensional objects based on the SLI technique.[46,47] The measurement is carried out by projecting a set of sinusoidal and Gray code images on the surface of the object, capturing the deformed patterns using a digital camera and transforming the obtained values to real coordinates with a calibration performed prior to the measurement.

The data obtained from the measurement is in the form of raw point clouds. A cloud of points represents an unordered set of points, where each point is characterized by its Cartesian coordinates, color (RGB or gray scale, depending on the type of detector), normal vector of the surface and quality factor corresponding to the uncertainty of coordinates' determination. These data form a set of information which fully describes the geometrical properties and color of the measured surface. In this study, only geometry is used in the automatic recognition process, however, color was helpful during the analysis of measurements by experts, as it made the measurement resemble a "3D photograph."

A single point cloud contains information about points on the surface of the measured object which are both illuminated by the projection unit and captured by the acquisition unit. The shape of human back was captured with a single-directional system.

2.2 Selection of Anatomical Structures

There are several anatomical structures used to calculate indices for spine deformity assessment.[28] These structures represent
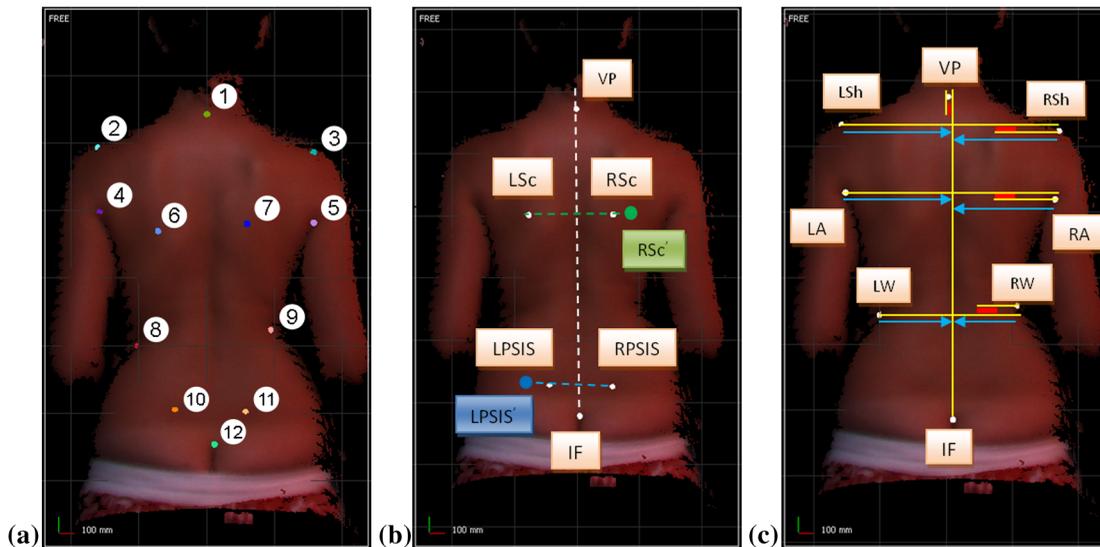
Fig. 1 (a) Anatomical landmarks used in this study: 1—vertebra prominens, 2,3—shoulders, 4,5—axillae, 6,7—scapulae, 8,9—waist line, 10,11—posterior superior iliac spines, 12—natal cleft; (b) POTSI (c) DAPI

skeletal topographic points characteristically visible on the surface (e.g., lower angle of scapula) or unique areas of the body (e.g., axilla). Locations of other landmarks are arbitrarily defined (e.g., point on the shoulder above the axilla). The common feature of those structures is that their position is influenced by the shape of the spine which may facilitate deformity diagnosis using only topographic data. The indices addressed in this study are POTSI[36,48,49] and DAPI,[25] which impose the selection of anatomical structures to discover (Fig. 1). By merging the lists of landmarks required for both indices, the following anatomical structures were selected:

- vertebral prominence of C7,
- top of the intergluteal furrow,
- left and right shoulder (points on the shoulder immediately above the axillae),
- left and right axilla,
- most proximal points on both sides of waist,
- posterior superior iliac spines (left and right),
- inferior angles of scapula (left and right).

Most of the structures are small areas rather than single points. However, the parameters commonly used to assess the deformity should take single positions as an input (POTSI and DAPI). Despite the fact that this approach may be error prone, it is a method medically acclaimed and thus has to be addressed. High quality and resolution of the measurement data suggest that more general parameters would better describe the deformity, but unfortunately such parameters have not been developed yet. What follows is that the output of the described algorithms should also be single points, each point indicating the position of a single anatomical structure.

2.3 Assumptions

The algorithm originally developed for this study does not fully incorporate any existing concepts. Its approach has not been described in the literature yet. It benefits only loosely from certain aspects of other analysis methods. The developed concept of anatomical structures recognition is based on the following empirical assumptions:

- the anatomical structures are associated with the shape of human body, i.e., the topographical properties of the surface;
- it is possible to describe local topography using features calculated for the surface;
- the mutual position of the anatomical landmarks and their relation to the boundary of the body in the frontal plane remains relatively stable across the population.

The anatomical structures are connected with the skeleton and are all placed in the trunk area. The variability of their location can be recognized and utilized by analyzing the population. The population data consists of a set of information about the patient, including a 3-D model of the patient's back, anatomical landmarks indicated on the model by a group of experts and some additional information, such as sex or BMI, which can help to distinguish natural variability between subjects from that caused by spine deformation.

The assumptions together form a mutual relation of equilibrium. Subjects may appear for whom some of the anatomical structures would not manifest sufficiently on the back surface, for example due to adipose tissue or musculature. In such case optimization process based only on the shape and not guided by any other *a priori* information may result in finding points of no anatomical significance, but of some local characterization of shape that by chance matched the given criteria. However, by providing additional *a priori* data in the form of a mean model of structure arrangement, a good starting point and additional constraint for the optimization process are obtained. The mean model is considered only as a very broad approximation of final positions of landmarks.

The strength of the proposed algorithm consists of both combined concepts, namely: the position of the anatomical structure may be approximated using the mean model and further tuned

using shape variability with influence of the statistical model, so that the spatial relations provided by the first approximation are regarded as an accurate guideline. The assumptions are expected to apply to most of the cases and in the preliminary version of the algorithm only general validation is performed, but the potential limitations mentioned above have to be considered in future and the method will be validated with more factors in mind, such as age, weight or posture type.

2.4 Algorithm Formulation

The developed algorithm is composed of two phases. The first phase is based on a learning set of data containing landmarks indicated manually by experts on the surface of point clouds representing patient's back obtained from measurement (Fig. 2). For each measurement the specialists of four physiotherapists and four orthopedists were requested to select all possible landmarks (based on a list supplied within the application used for this purpose, containing all the structures listed in Sec. 2.2) without the knowledge about the position of landmarks selected by the rest of the staff. Altogether, a total of 23 measurements were examined. The average age of the patients was 24.7 (minimum 23, maximum 26), with a sex ratio of 11 females and 12 males. The data were supplemented by information about the region of interest for every measurement, detected automatically using the developed algorithms. The landmarks are aligned using a modified Procrustes method and a mean model is built upon it, including an average region of interest.[50]

It should be mentioned that the mean model was generated only on the basis of healthy (as stated by the experts) nonobese subjects, to present a good reference frame for the recognition process. Precise information about the location of structures in unknown measurements will then be supplied by shape variability.

The second phase uses the mean model from the previous step. First, the measurement is aligned and the region of interest is found within the measurement using the same methods as in phase one. Then, knowing the relation between the mean and current regions of interest, the mean model containing positions of landmarks is adjusted to fit new data and the points are projected onto the measurement surface. Finally, features are calculated for the point cloud, normalized locally and combined with the approximated positions of landmarks with the help of a Voronoi diagram[51] to extract the final location of anatomical structures.

2.4.1 Definition of the region of interest

The region of interest extracted from the measurements has to be defined unambiguously. During this process, information is generated by the segmentation operation. The bounding box is then estimated as follows:

- the left and right surfaces are set by the left and right axillae;
- the top surface is set on the neck, at a point exhibiting a specific width ratio with the shoulders;
- the bottom surface is indicated by the buttocks.

The algorithm takes a preliminarily aligned cloud of points as its input. It may be classified as operating on 2-D data (in fact it is a simple segmentation), but considering information about depth as well. The point cloud is divided into slices perpendicular to the vertical axis (calibrated during measurement) which are then connected into groups. Using this information, characteristic points used to calculate the final bounding box are found. For the set of measurements analyzed in this work the following, experimentally chosen, values were used as input for the algorithm:

- segmentation slice height: 10 mm;
- minimum distance between groups in the X direction: 8 mm;
- minimum distance between groups in the Z direction: 8 mm;
- width drop ratio from arms to neck: 0.5.

The front and back faces of the bounding box are not necessarily required. The bounding rectangle is used only to scale the mean landmark set in the X and Y directions. The rectangle is described using following features (Fig. 3):

- the orientation of the rectangle, in the form of three unit vectors corresponding to the X, Y and Z axes of the coordinate system;
- center of the coordinate system of the area, positioned in the center of the bounding rectangle;
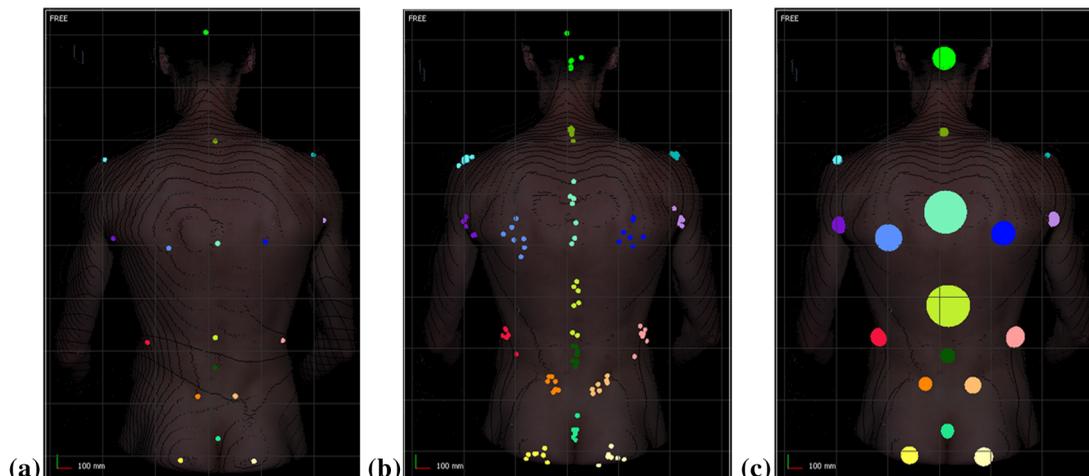- size of the box, in the X and Y directions, reflecting size of the bounding rectangle.

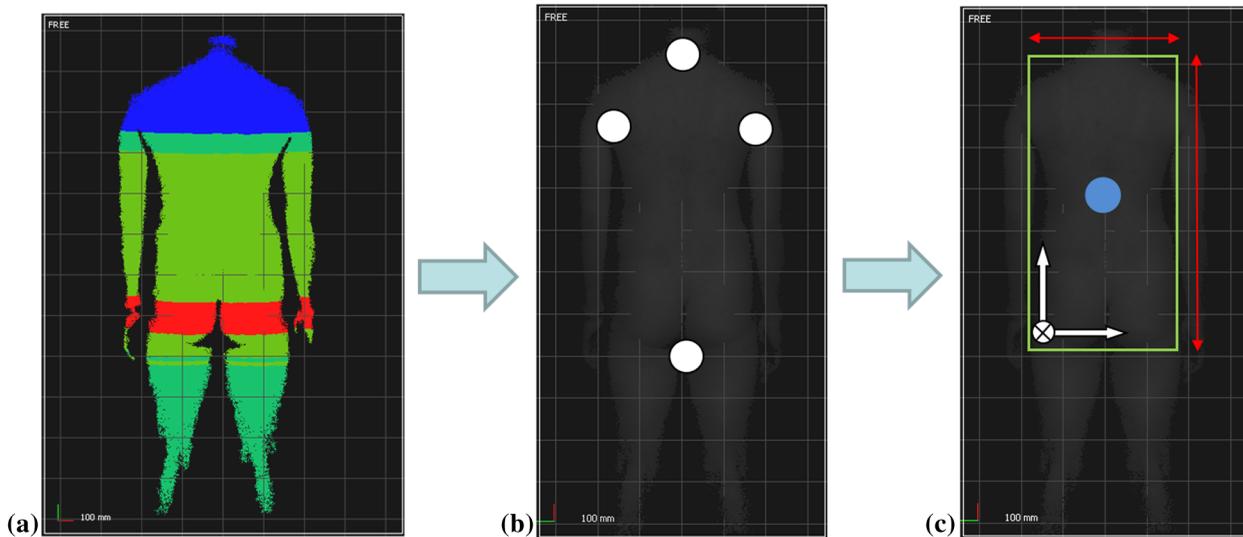**Fig. 2** Landmarks charted by (a) single expert, (b) group of 8 experts, and (c) uncertainty.

Fig. 3 (a) Vertical cross-sections, (b) limiting points, and (c) region of interest.

2.4.2 *Analysis of existing dataset*

Every measurement requires preprocessing, i.e., filtering, alignment and segmentation. For this purpose the same algorithms as in the recognition part are used, so the correspondence between the currently analyzed measurements and their landmarks can be applied in future to reverse the relation and approximate the landmarks based on the features calculated for the new measurement. For each measurement, points indicating characteristic anatomical structures selected by the operators were averaged and projected onto the cloud, creating a mean landmark set. The deviations were determined as well. This landmark set is a local mean landmark set, as opposed to the global mean landmark set, which will be calculated based on all local sets. Because the mean local sets are found in different coordinate systems and come from different patients, the shape cannot be analyzed until the operations mentioned above are used: a full transformation (i.e., translation, rotation and scale) has to be calculated for every set.[52,53]

In order to find the rotation and translation parameters, it suffices to use the Iterative Closest Point (ICP) technique.[54] The method for computing scale was based on,[55] which is calculated before other operations as follows:

- calculate the centers of mass of the two sets:

$$\bar{x}_l = \frac{1}{n} \sum_{i=1}^{n} x_{l,i}, \qquad \bar{x}_r = \frac{1}{n} \sum_{i=1}^{n} x_{r,i},$$

- modify the positions of the landmarks by subtracting the centers of mass

$$x'_{l,i} = x_{l,i} - \bar{x}_l, \qquad x'_{r,i} = x_{r,i} - \bar{x}_r$$

- the optimal scale factor s is then computed as:

$$s = \sqrt{\frac{\sum_{i=1}^{n} \|x'_{r,i}\|}{\sum_{i=1}^{n} \|x'_{l,i}\|}}.$$

Anisotropic scaling perturbs the definition of the model. Further, in the process of recognition, after the model is roughly

segmented, the mean model is adjusted in a more appropriate way using a bounding rectangle based on body boundaries. Simple single-factor scaling would fail to address the variability of the skeletal structure of the body and would give less accurate approximations, because it would only take into account the overall change in size of the model.

All the shapes are first coarsely aligned using the system's vertical, calibrated before the measurements, and the left and right posterior superior iliac spines (PSIS), obtained from the defined landmark set, as the fine alignment using ICP needs a good starting point (i.e., the differences between positions of the stable and aligned model have to be moderate).

The final shapes alignment is carried out with ICP by choosing one arbitrary landmark set as the stable one, to which all the other shapes are aligned (Fig. 4). Then, second iteration of alignment is performed in the same fashion, only this time a mean model can be calculated from all the previously aligned shapes. The scale, pose and central point of the model are normalized as follows:

- the scale was chosen so that the difference between PSIS landmarks is equal to 1;
- pose was chosen so that the line connecting vertebra prominens and top of the natal cleft is the Y axis;
- central point was chosen as the middle point between the PSIS landmarks.

This approach, described in Ref. 43, allows to obtain better models, as opposed to normalizing each individual shape separately.

The bounding rectangles obtained for every measurement were aligned using transformations found for the shape along with the landmark sets. The global mean bounding rectangle was calculated. Following the formulation of the bounding rectangle, we obtain a set of:

- triples of X, Y, Z unit vectors;
- coordinate system centers;
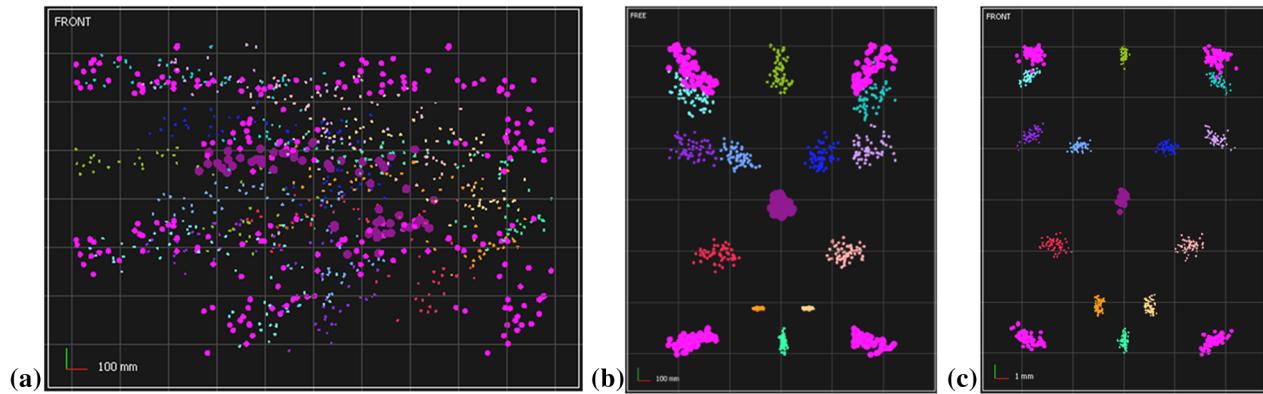- size of the box in X and Y direction of the bounding box coordinate system.

Fig. 4 (a) Unaligned sets of landmarks and bounding rectangles (in magenta), (b) after preliminary alignment, (c) after final alignment.

Next iterations are performed in the same way as the second one until the method converges. The convergence condition is tested by examining the difference between calculated mean shape of two consecutive iterations.

2.4.3 *Finding potential areas for the locations of landmarks*

Having a new measurement, after preprocessing and calculation of the bounding box, the transformation between the current and the mean bounding box can be found easily (in fact this is equivalent to moving between coordinate systems and scaling). This transformation is used to transform the mean shape into the coordinate system of the new measurement (Fig. 5). The role of this approximation is twofold:

- approximation of areas where the anatomical structures will be found; with this knowledge global optimization can be avoided and local optimization performed for every area does suffice;
- the points are potentially good starting positions which may facilitate and accelerate the optimization process.

In order to divide the point cloud into areas corresponding to the calculated landmark set, the landmarks are projected onto the frontal plane. Then, the two-dimensional Voronoi diagram is computed. The main purpose of such a diagram for a given a set of points is to assign to each of these points a segment of space in which a defined distance metric yields smaller values than for any other point. In this case, a simple Euclidean metric in the frontal plane is used. The diagram is then projected back on the cloud with every cell determining a separate group of points connected with each anatomical structure.

An additional advantage of using the Voronoi diagram in this case results from the rigidity of the skeletal structure. The anatomical structures, although their exact positions are unknown, never lie in direct proximity to each other. The diagram assigns weights to points in such a way that points placed halfway between the approximated positions are the least probable to become the final result.

2.4.4 *Surface shape descriptors*

The methods of describing a surface may be divided into based on local curvature and employing statistical features. The former
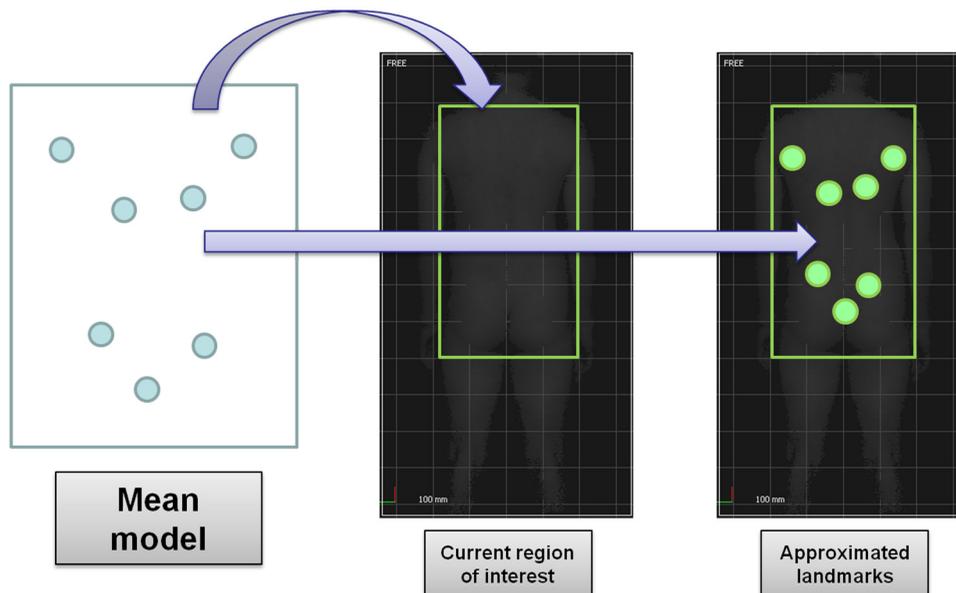

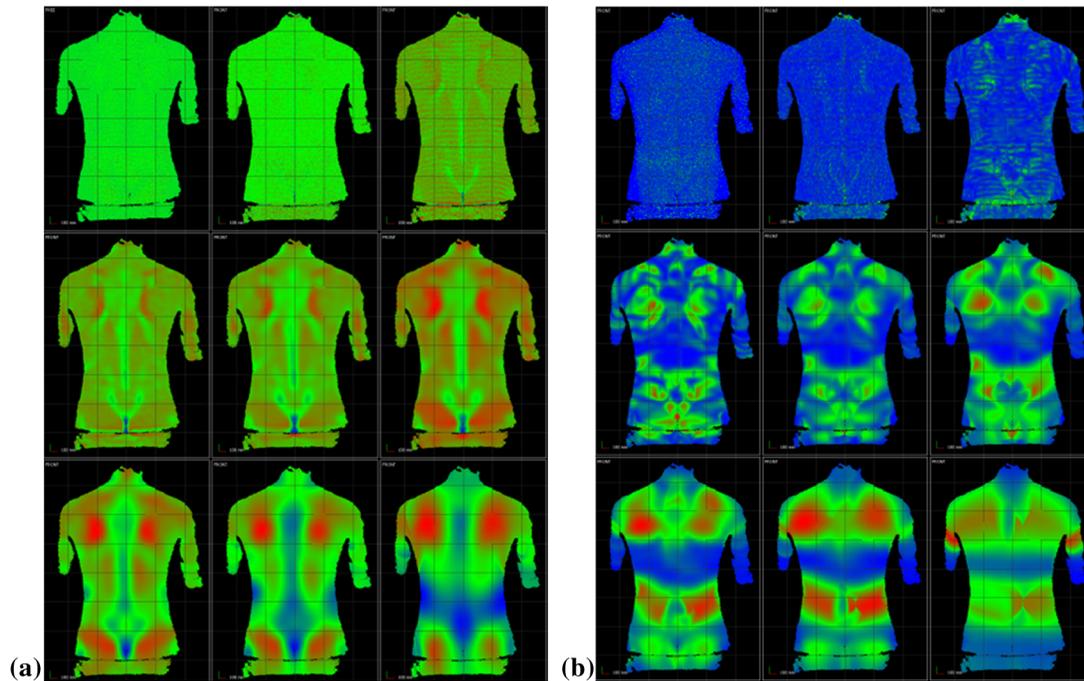**Fig. 5** First approximation by fitting regions of interest.

Fig. 6 Distributions of features for different radii: (a) C1 and (b) C2.

may be further divided into calculated on the basis of analytical definition, mathematically approximated or numerically computed as derived features, semantically similar to classical parameters, though not directly correlated. When computing the mathematically defined curvature, other derived parameters are available, which assist the interpretation of obtained surface description.

In case of a 3-D surface representation in its parametric form, calculation of the curvatures is a straightforward differentiation task. However, when working with real data there is an additional difficulty in approximation of the local shape of the surface before calculating the curvatures.

Methods of curvature approximation are either inaccurate and prone to noise or computationally expensive.[47] Estimators used for triangle meshes work well, but require intensive preprocessing to create the mesh, which may take longer than estimation of the curvature itself for experimental data.

The main advantages of using numerically computed parameters (including simple statistical features) over mean and Gaussian curvatures are full numerical stability for all neighborhood types. There is no need for setting boundary conditions for simultaneous nonlinear differential equations. It is also more resistant to noise, and operates with a less complicated algorithm yielding faster numerical apparatus, because it does not require iterative calculations.

Therefore, five features suited to the type of input data and purposes were selected:

- topographical parameters: C_1 and C_2,
- statistical parameters: RMS, kurtosis and skewness.[56]

Because all these parameters have to be calculated assuming local neighborhood of a specified size, the operator size which could describe the structures to be found has to be determined beforehand. Moreover, usefulness of these features should be evaluated, in order to minimize information necessary to find specific structures. To accomplish this, distributions of all features for nine different operator sizes (neighborhood radius equal to 2, 5, 10, 20, 30, 40, 60, 80 and 120 mm) were tested for a set of measurements (Fig. 6).

In the low range of neighborhood size, the noise and measurement errors were visible, while in the upper range only general description of the shape could be given. Optimal values of the neighborhood were found in the central range, with radius of 30 mm being the most informative. It was also discovered that the C_1 and C_2 parameters present the best insight into shape, with the least ratio of noise. The statistical features become more prone to noise when moments of higher order are computed (RMS giving the lowest amount of noise, with results similar to C_1).

2.4.5 Cost function

The cost function is a weight map, which contains a specific scalar value for every point in the region of interest. This function is composed of:

- weights calculated from the Voronoi diagram; this part of the cost function is influenced by the landmark set approximated using the mean model; the values are in the range [0;1] and are calculated as:

$$V(P) = d(P, E_{i,j})/d(S, E_{i,j}),$$

where $d(P, E_{i,j})$ is the two-dimensional distance of point P from the j-th edge of the i-th cell along the line containing both site S and point P, and $d(S, E_{i,j})$ is the two-dimensional distance of the site S from the edge along the same line; the highest value corresponds to the central point of the cell and the lowest to points on the boundary of the cell;

- weights resulting from features calculated from current measurement data; these weights are computed
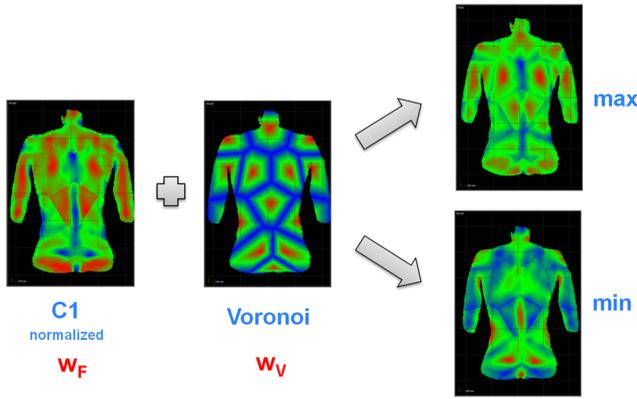
Fig. 7 Generation of cost function using normalized map of C1 parameter and weights resulting from Voronoi diagram.

by normalizing the value of each feature with respect to the maximum and minimum value in each cell of the Voronoi diagram; the final values of every feature are scaled to range [0;1] using maximum and minimum values of the feature in each cell; then, values for all the points belonging to a cell are normalized in the following manner:

$$F(P) = [f(P) - f_{\min,i}]/(f_{\max,i} - f_{\min,i}),$$

where $f(P)$ is the old value in point P, $f_{\min,i}$ and $f_{\max,i}$ are minimum and maximum values for the ith Voronoi cell and $F(P)$ represents the new value assigned to the point P.

The components of the weight function are then summed to yield the final map. Additionally, the weights given by a feature for a specific cell are inverted, depending on whether minimum or maximum is predicted for the specific landmark (Fig. 7). The optimization is also governed by coefficients used during summing of the Voronoi and feature parts:

$$C(P) = w_v V(P) + w_F F(P),$$

where $C(P)$—resulting cost function; w_v, w_F—factors assigned to the Voronoi and feature maps; $V(P)$—Voronoi weight for a given point P, $F(P)$—feature weight for a given point P.

In order to find the optimal factor of influence for the feature and mean model for all landmarks, the recognition process was carried out with step of 0.1, with 0.0 corresponding to a process based only on the local feature and 1.0 corresponding to a process governed only by the mean model.

2.4.6 *Extracting final positions of landmarks*

Final positions of points are obtained by a local optimization method, which simply takes the first approximation as a starting point, and then iteratively searches for the largest value in the local neighborhood of the point of radius equal to three average point-to-point distances in the cloud. The process stops when the last calculated value is the largest and the corresponding geometrical point of the cloud is taken as the final position of the landmark.

3 Validation

Since the information to be found is not exactly specified (i.e., it is not possible to define exact locations in space of anatomical structures, because they are not single points, as described in

Sec. 2.2), it is hard to assess the accuracy of results. What determines the characteristics of the areas searched for is expert knowledge. The most obvious solution for validation would be to compare the estimated position of each structure with its position indicated by a physician, although it has to be assumed that manual indication has high intra-observer and inter-observer reliability.[38]

To become less dependent on the variability of manual indication and to provide more reliable input for comparison, mean positions of structures were calculated over all points supplied by experts. These local average mean landmark sets were found by calculating average points for given landmarks for every measurement and projecting them onto the cloud, so that they represent actual points on the back surface. The uncertainty for each landmark was calculated as the root mean square error between all the manual landmark positions and the obtained average.

The final validation was performed by comparing the Cartesian distance between automatically found positions of landmarks and averaged positions of manually indicated landmarks with the uncertainty of manual indication. Accuracy was tested with different weight factors for the Voronoi diagram and curvature map to determine an optimal value for the current population. The error for every landmark was normalized with respect to the minimum and maximum in all factors to expose the relation trend of the recognition accuracy with respect to weight factor value.

The developed method has to consider anatomical landmarks used in deformation assessment with topographical data, but not all of those landmarks can be related to x-ray, as they are not part of the skeletal system (i.e., axillae or waist). Moreover, trunk radiography can be performed either in the horizontal or vertical position, which may influence on the mutual locations of anatomical structures with respect to standing position, used in three-dimensional measurements. Finally, it is hard to correlate the coordinate systems of x-ray and 3-D models in an unambiguous manner, which further impedes result comparison. Due to these reasons validation of results with radiography data was not attempted in this study but some data already appeared in the literature.[34]

4 Results

Preliminary results representing the accuracy of automatic landmark recognition with respect to accuracy of manual indication by the experts for some of the landmarks is shown in Fig. 8. Accuracy is better in the automatic procedure, although standard
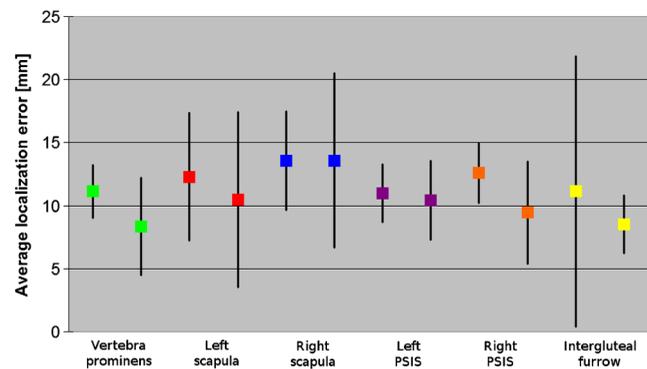

Fig. 8 Average error and standard deviation in manual (left value) and automatic (right value) procedure for selected anatomical structures.

Fig. 9 Representative distributions of the recognition error for (a) left scapula, (b) right PSIS, and (c) vertebra prominens for different values of weight factor for 8 patients. Figures to the left show the relative recognition error, which is the ratio of distance between automatic and manual landmark, figures to the right show the same error normalized to expose the trend (red dashed line).

deviation is usually higher. In all presented cases the error was between 5 and 15 mm.

The trends exposed for landmarks, showing the relation of recognition accuracy against weight factor used in the cost function (Fig. 9), are of various character and can be grouped into:

- trends which exhibit a clear minimum: both scapulae and posterior superior iliac spines; the minimum for scapulae appears between 50% and 60% of influence of the mean model (relative error with respect to manual uncertainty is 58% and 78% for the left and right scapula, respectively); right PSIS is best recognized on average for 60% to 70% of influence of the mean model (relative error with respect to manual uncertainty is 11% to 16%), while left PSIS is best recognized on average for 40% to 50% of influence of the mean model (relative error with respect to manual uncertainty is 34%); trend minimum can also be found for left shoulder for 30% to 40% of influence of the mean model (average relative error with respect to manual uncertainty equal to 20%) and for right axilla—at 60% (average relative error with respect to manual uncertainty equal to 32%);

- trends which attain smaller values with high influence of a feature: right shoulder;

- trends which attain smaller values with high influence of mean shape: vertebra prominens and natal cleft;

- trends which do not show a clear relation: left axilla.

In general, the recognition error depends on the mutual position of feature's local extremum, manually indicated landmark and mean landmark. The higher the influence of the mean model, the more the error distribution stabilizes to the difference between a manually indicated point and the mean model. When a specific landmark is closer to the average model, the final error will become small. However, this error is not stable and may prove to be equally large or small and depends on how much the specific landmark's position is different from its corresponding position in the mean model. On the other hand, if the manually indicated landmark is close to the local extremum of the feature, errors will decrease with decreasing influence of the mean model, but if the assumption that the true location of the structure is connected with an extremum of shape feature distribution is incorrect, the error will increase. In total, the two models should compensate, yielding more stable results than obtained by the two methods used separately, which can be seen in the case of scapulae and posterior superior iliac spines.

It is also interesting that despite low accuracy of indication of axillae extracted with high influence of the feature, points on the shoulders found using these erroneous results yield higher accuracy with respect to manual landmarks. This may be due to the fact that the points representing axillae are unstable in the Y direction, but are well localized in the X direction, which is used to extract positions of shoulders.

Similarly, better accuracy of localization of spinous process of vertebra prominens and natal cleft with high influence of the mean model results from the fact that these points were used to extract the region of interest (especially natal cleft, which was used as a limit for the bounding rectangle due to incomplete measurement data). The remaining error results from differences of the mean model in the X direction only, as the mean model is directly scaled to fit those landmarks.

Values of the deformation parameters are very unstable for all the influence factors. However, a general conclusion can be drawn in the case of POTSI,[47,57] that balanced values of the factor from the middle of the range provide better correlation with parameters calculated on the basis of manually indicated landmarks.

5 Conclusions and Future Work

All anatomical back surface landmarks have been recognized successfully as there were no coarse errors among processed results. The accuracy of automatic detection of waist, shoulders, or axillae on the boundary of the body in the frontal plane remains lower than the accuracy of manual indication, mainly due to inaccurate formulation of the structures and measurement errors, which result from local surface orientation tangent to the direction of measurement. In future versions of the system, a two-directional measurement system will be used to eliminate some of these errors. Nevertheless, other structures in the central part of the back surface were recognized with comparable or even better accuracy than the average accuracy of operator's detection (ranging from 50% to 90% of average relative error).

The problem of choosing factors for merging feature map and weights calculated on the basis of Voronoi diagram remains significant. The ratio of those factors determines whether we rely more on the local shape properties or the statistical model, which stores the general geometrical relations between landmarks. In this case, the factor was chosen experimentally to be equal to 0.7 for the feature and 0.3 for the Voronoi weights, i.e., an assumption is made that the variability of shape will guide the recognition process better than the rough statistical model. This approach works for the majority of cases, but sometimes ambiguities emerge. Figure 10 shows two cases of the cost function, calculated with influence factor 0.3 and 0.4 for the Voronoi diagram weights. In the former case the scapula is positioned close to the local maximum of the C_1 parameter, while in the latter Voronoi weights dominate and the scapula is placed close to the first approximation of the statistical model. In this case, a serious asymmetry is visible in the region of right scapula, giving rise to the bump close to the spine line. The question remains whether asymmetry should be accounted for during calculation of deformity parameters and signal abnormal shape distribution.

Weights computed on the base of Voronoi diagram not only determine the influence of the approximated locations of landmarks, but also decrease the probability for points near the boundaries of the cell to be selected as anatomical landmarks. It plays a role when the feature takes extreme values for points near the border, which could confuse the optimization process. In specific cases a problem arises, when feature artifacts occur and very high (or low) values of the feature are found in a cell (Fig. 11). The normalization process simply scales values of the feature within each cell with respect to the minimum and maximum and, if an unexpectedly extreme value is used, difference between values near the local extremum for the actual landmark searched for becomes very small. In general, this would not cause big problems to the optimization process. However due to merging of these normalized values with weights from Voronoi diagram, which are computed artificially and are always in the range of [0; 1], very low sensitivity is obtained in the region of the extremum. The extremum is iteratively searched for in the
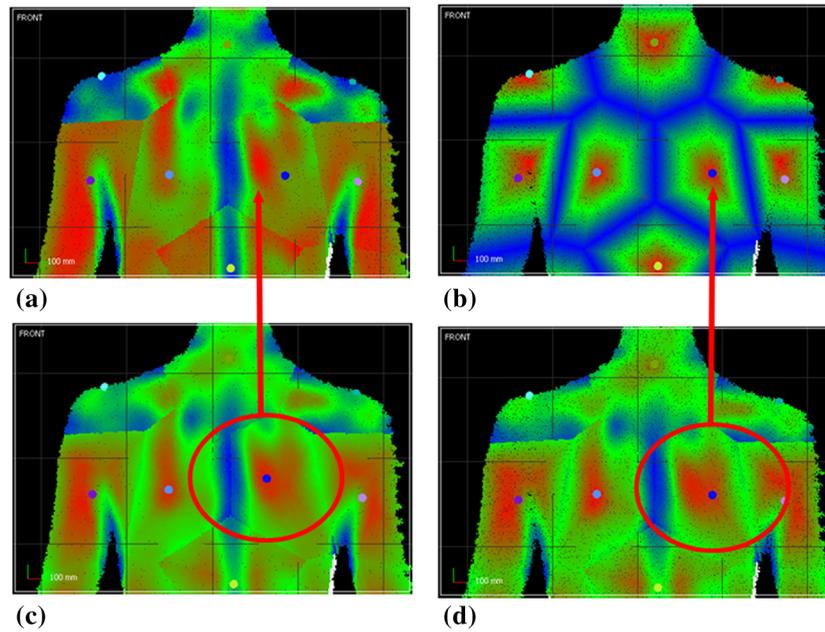
Fig. 10 (a) Map of the C1 parameter normalized in Voronoi cells, (b) Voronoi weights computed on the base of mean shape, (c) both maps merged with factor 0.3, and (d) 0.4 for the Voronoi part, yielding different positions of right scapula.

local neighborhood of a small radius and in extreme cases the optimization process may fail.

Generally, all errors in shape alignment propagate to subsequent stages of analysis, which is rarely addressed in the literature. This corresponds to the mutual relation of measurement system and the data captured during measurement. Analysis of any set of geometrical data should be preceded by accurate alignment with respect to an accurately defined reference coordinate system.

The alignment error in the training set may directly influence the generation of the mean model. It is impossible to give an

absolute estimate of it, because the problem of shape correspondence itself is very vague. Although formulated quite clearly, there is no objective way of solving it. In this study, the most popular and mathematically correct method was used. It is arguable if cumulative distance error of all landmarks used for fine-tuning the transformation of each shape is an objectively better criterion than the one used for coarse alignment process—the pelvis could easily be chosen as an object of reference for the skeletal system with additional orientation correction using the vertical direction, which would give an anatomical reference frame. Moreover, the alignment error affects deformation
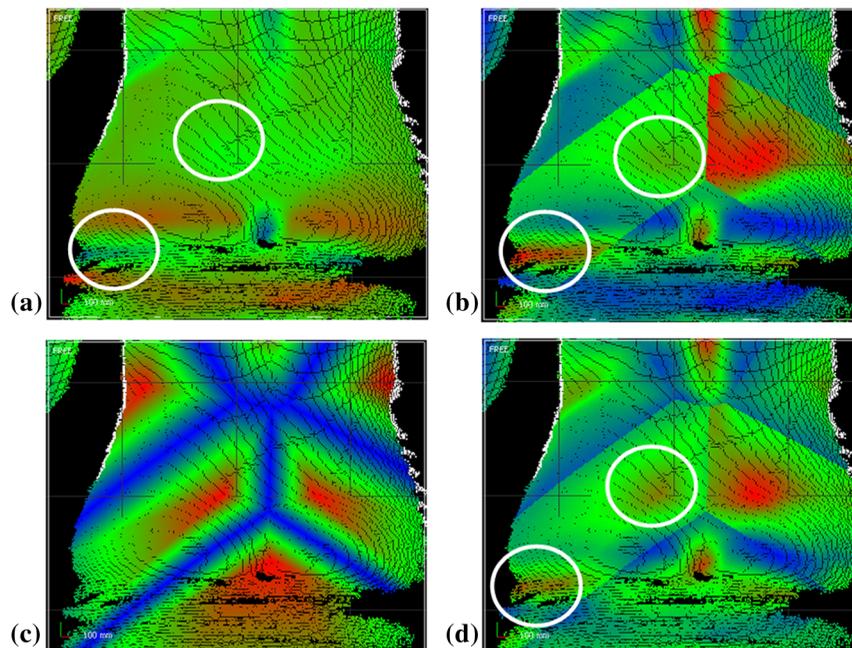

Fig. 11 Map of the C1 parameter (a) normalized in Voronoi cells (b), Voronoi weights (c) and the final cost function computed by merging the two previous maps (d). In the cell corresponding to left posterior superior iliac spine, extreme values of the feature are found on the left gluteus, caused by underwear. After normalization, values in the area of the PSIS were flattened.
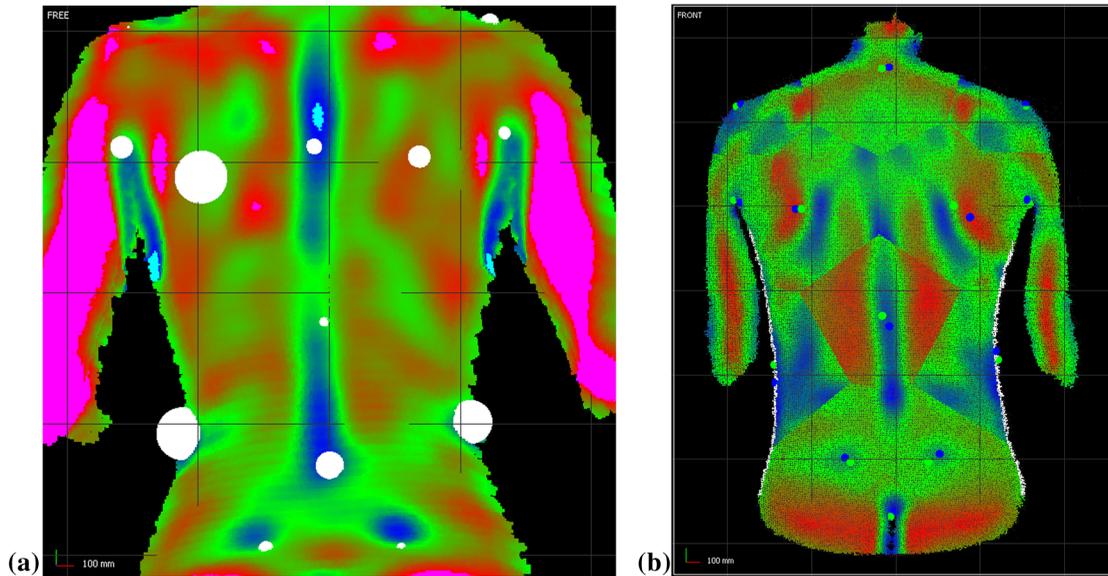
Fig. 12 (a) A distribution of the C1 parameter with average positions of landmarks indicated manually (in white, the radius corresponds to uncertainty of indication). (b) Automatically located landmarks (in blue), extracted using influence factor of 0.1, found at the local extremum, and manually indicated landmarks (in green); it is arguable which positions of the PSIS are more correct.

indices, calculated based on positions of localized landmarks. Currently, values of the indices are inaccurate not only because of high intra-observer variability, but also due to misalignment of the back surface. Some systems use depth-alignment,[22] in cases where the prominence of landmarks is relevant, but most often a correct position is assumed. Finally, it may be related to the problem of manual indication of landmarks by experts. Those cannot be validated and have to be considered correct, which holds true in most cases.

However, if the anatomical structures are not well visible on the surface of the back (the landmarks were indicated on the point cloud without any additional help), the indication is sometimes incorrect and could provide better results when correlated with the features at the time of indication. Figure 12(a) presents the distribution of the C_1 parameter with average positions of landmarks indicated manually (in white, the radius corresponds to uncertainty of indication). Clearly the posterior superior iliac spines should be placed in the minima of the distribution (blue areas above), while, for some reason, all the experts positioned them below the correct region (small radius refers to low uncertainty). Moreover, in this measurement the scapulae are not well visible, which impedes any topography-based recognition, be that manual or automatic.

Automatic analysis should provide more repeatable results, when the structures actually do manifest themselves on the surface of the back. It is questionable in case of scapulae and vertebra prominens, where surface shape varies with changes in the positions of arms and posture in general. If this condition is not fulfilled, the result will be very coarsely approximated and will raise even bigger errors. However, the authors state that the recognition process should be guided mainly by the shape of the back surface, putting as little emphasis on the first approximation as possible. This may be regarded as a more radical approach to the analysis process, because recognition will fail if anomalies are present. On the other hand, if more influence will be given to the mean model, which represents an average patient (i.e., healthy in a normal population), cases with an

actual deformation present will be regarded as closer to the average and the magnitude of the deformation will be diminished. It is therefore a choice between a false positive and a false negative, which should be considered by specialists utilizing the system [Fig. 12(b)]. In this study a balanced compromise was presented, which may provide guidelines for future uses of the developed algorithms. It has to be noted that the algorithm was only tested on a small dataset and should be validated with a larger amount of measurement data, with diverse distributions of patients.

One of the considered directions of development may be the incorporation of Eigen-analysis of the data (e.g., PCA), firstly, in the analysis of landmarks indicated by the experts and secondly, in the analysis of feature distribution for the corresponding measurement data. Former path probably will not enhance much the quality of recognition. However, it may turn out to be a useful tool for further research into deformation type assessment based on the anatomical landmarks, especially since it would cover all three dimensions. This is a direct consequence of the analysis method, since the resulting Eigen modes are dimensions containing the most variability of the shape. Experts have to further examine the modes of variability for correspondence with specific types of deformation or correlation techniques should be used to filter out other, irrelevant variability, resulting from sex, age or BMI. Similar approaches have already been tried, but they were either not further developed[55] or treated as raw input for further stages of analysis without any interpretation attempts.[58] Additionally, the chosen shape alignment method would have a great impact on this analysis as mentioned above.

In the latter case, dimensionality of the description could be reduced using Eigen-analysis, but in this study manual localization of the structures in the training set is determined with moderately low accuracy, which could provide confusing results. In case of a holistic approach the back shape would be analyzed as a whole, which would introduce irrelevant information to the description (as in the case of AAM). On the other hand if the neighborhood of each landmark was to be analyzed separately, it

would be difficult to choose the boundary of each structure (both shape and size) containing information significant for that particular structure.

An interesting aspect of the Active Appearance Model technique is that it takes into account not only the distribution of object's shape, but also the statistical model of its appearance in space. If the appearance would be interpreted differently, i.e., as the distribution of some feature on the surface, statistical analysis of such distributions across the population can provide a useful tool for building correspondence between the landmarks and its actual position in the model. This however has to be analyzed locally, not globally, which would be considered a hybrid approach, connecting some of the approaches of ASM and AAM.

Acknowledgments

References

1. A. M. Briggs et al., "Thoracic spine pain in the general population: prevalence, incidence and associated factors in children, adolescents and adults. A systematic review," *BMC Musculoskel. Disord.* **10**(77) (2009).
2. G. S. Dieck et al., "An epidemiologic study of the relationship between postural asymmetry in the teen years and subsequent back and neck pain," *Spine* **10**(10), 872–877 (1985).
3. J. Nowotny et al., "Body posture and syndromes of back pain," *Orthop. Traumatol. Rehabil.* **13**(1), 59–71 (2011).
4. F. Schwab et al., "SRS-Schwab adult spinal deformity classification: a validation study," *Spine* (2011).
5. J. A. Anderson, E. O. Otun, and B. J. Sweetman, "Occupational hazards and low back pain," *Rev. Environ. Health* **7**(1–2), 121–160 (1987).
6. A. C. Johansson et al., "Clinic-based training in comparison to home-based training after first-time lumbar disc surgery: a randomised controlled trial," *Eur. Spine J.* **18**(3), 398–409 (2009).
7. B. K. Kwon et al., "Systematic review: occupational physical activity and low back pain," *Occup. Med. (Lond).* (2011).
8. P. Janwantanakul et al., "Development of a risk score for low back pain in office workers—a cross-sectional study," *BMC Musculoskel. Disord.* **12**(23) (2011).
9. C. R. Good et al., "Adult spine deformity," *Curr. Rev. Musculoskel. Med.* **4**(4), 159–167 (2011).
10. W. B. Katzman et al., "Age-related hyperkyphosis: its causes, consequences, and management," *J. Orthop. Sports Phys. Ther.* **40**(6), 352–360 (2010).
11. Y. A. Shelton, "Scoliosis and kyphosis in adolescents: diagnosis and management," *Adolesc. Med. State Art Rev.* **18**(1), 121–139 (2007).
12. M. Sharma and P. K. Majumdar, "Occupational lifestyle diseases: An emerging issue," *Indian J. Occup. Environ. Med.* **13**(3), 109–112 (2009).
13. P. Cote et al., "A study of the diagnostic accuracy and reliability of the Scoliometer and Adam's forward bend test," *Spine (Philadelphia, PA 1976)* **23**(7), 796–802. Disc. 803 (1998).
14. L. Balzini et al., "Clinical characteristics of flexed posture in elderly women," *J. Am. Geriatr. Soc.* **51**(10), 1419–1426 (2003).
15. D. M. Kado et al., "Comparing a supine radiologic versus standing clinical measurement of kyphosis in older women: the Fracture Intervention Trial," *Spine (Philadelphia, PA 1976)* **31**(4), 463–467 (2006).
16. A. Gilmore and G. H. Thompson, "Radiographic evaluation of children and adolescents with a spinal deformity," *Semin. Musculoskel. Radiol.* **4**(3), 349–359 (2000).
17. J. Cheung et al., "The reliability of quantitative analysis on digital images of the scoliotic spine," *Eur. Spine J.* **11**(6), 535–542 (2002).
18. M. M. Doody et al., "Breast cancer mortality after diagnostic radiography: findings from the Scoliosis U. S. Cohort Study," *Spine (Philadelphia, PA 1976)* **25**(16), 2052–2063 (2000).
19. C. M. Ronckers et al., "Cancer mortality among women frequently exposed to radiographic examinations for spinal disorders," *Radiat. Res.* **174**(1), 83–90 (2010).
20. C. M. Bone and G. H. Hsieh, "The risk of carcinogenesis from radiographs to pediatric orthopaedic patients," *J. Pediatr. Orthop.* **20**(2), 251–254 (2000).
21. K. Ashizawa et al., "Approximate surface development of the left side half-trunk by a free-formed model," *Am. J. Phys. Anthropol.* **76**(2), 165–173 (1988).
22. B. Drerup, "The measurement of the kyphotic angle by contact free registration of back-shape (author's transl)," *Z. Orthop. Ihre. Grenzgeb.* **120**(1), 64–70 (1982).
23. X. C. Liu et al., "Effects of trunk position on back surface-contour measured by raster stereophotography," *Am. J. Orthop. (Belle Mead, NJ)* **31**(7), 402–406 (2002).
24. A. I. Tsirikos "Development and treatment of spinal deformity in patients with cerebral palsy," *Indian J. Orthop.* **44**(2), 148–158 (2010).
25. M. F. Minguez et al., "Quantifier variables of the back surface deformity obtained with a noninvasive structured light method: evaluation of their usefulness in idiopathic scoliosis diagnosis," *Eur. Spine J.* **16**(1), 73–82 (2007).
26. A. Zubovic et al., "New method of Scoliosis Deformity Assessment: ISIS2 System," *Stud. Health Technol. Inform.* **140**, 157–160 (2008).
27. H. L. Mitchell and K. S. Ang, "Non-rigid surface shape registration to monitor change in back surface topography," *Stud. Health Technol. Inform.* **158** 29–33 (2010).
28. P. Patias et al., "A review of the trunk surface metrics used as Scoliosis and other deformities evaluation indices," *Scoliosis* **5**(12) (2010).
29. I. Weisz et al., "ISIS scanning: a useful assessment technique in the management of scoliosis," *Spine (Philadelphia, PA 1976)* **13**(4), 405–408 (1988).
30. T. N. Theologis et al., "Early detection of progression in adolescent idiopathic scoliosis by measurement of changes in back shape with the Integrated Shape Imaging System scanner," *Spine (Philadelphia, PA 1976)* **22**(11), 1223–1227. Disc. 1228 (1997).
31. J. G. Thometz et al., "Relationship between Quantec measurement and Cobb angle in patients with idiopathic scoliosis," *J. Pediatr. Orthop.* **20**(4), 512–516 (2000).
32. A. Bujakiewicz, L. M. Krzesniak, and D. Zawieska, "The moire technic in epidemiological studies of spinal curvatures," *Pol. Tyg. Lek.* **40**(35), 1002–1005 (1985).
33. N. J. Oxborrow, "Assessing the child with scoliosis: the role of surface topography," *Arch. Dis. Child.* **83**(5), 453–455 (2000).
34. C. J. Goldberg et al., "Surface topography, Cobb angles, and cosmetic change in scoliosis," *Spine* **26**(4), E55–E63 (2001).
35. T. Kotwicki, "Evaluation of scoliosis today: examination, x-rays and beyond," *Disabil. Rehabil.* **30**(10), 742–751 (2008).
36. T. Kotwicki et al., "Discrepancy in clinical versus radiological parameters describing deformity due to brace treatment for moderate idiopathic scoliosis," *Scoliosis* **2**(18) (2007).
37. T. Kotwicki et al., "Methodology of evaluation of morphology of the spine and the trunk in idiopathic scoliosis and other spinal deformities—6th SOSORT consensus paper," *Scoliosis* **4**(26) (2009).
38. W. Glinkowski et al., "Reproducibility of structural light assessment 3D posture telediagnostics: preliminary reliability study," *Global Telemedicine and eHealth Updates: Knowledge Resources* **4**, 532–535 (2011).
39. L. K. Arata et al., "Three-dimensional anatomical model-based segmentation of MR brain images through Principal Axes Registration," *IEEE Trans. Biomed. Eng.* **42**(11), 1069–1078 (1995).
40. L. Soler et al., "Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery," *Comput. Aided Surg.* **6**(3), 131–142 (2001).
41. H. E. Cline et al., "Three-dimensional segmentation of MR images of the head using probability and connectivity," *J. Comput. Assist. Tomogr.* **14**(6), 1037–1045 (1990).
42. M. A. Styner et al., "Evaluation of 3D correspondence methods for model building," *Inf. Process Med. Imaging.* **18**, 63–75 (2003).
43. T. F. Cootes et al., "Active shape models-their training and application," *Comput. Vis. Image Understand.* **61**(1), 38–59 (1995).
44. H. Sun et al., "Automatic cardiac MRI segmentation using a biventricular deformable medial model," *Med. Image Comput. Comput. Assist. Interv.* **13** (Pt 1). 468–475 (2010).

45. T. Huysmans et al., "A 3D active shape model for the evaluation of the alignment of the spine during sleeping," *Gait Posture* **24**(1), 54–61 (2006).

46. R. Sitnik and M. Witkowski, "Locating and tracing of anatomical landmarks based on full-field four-dimensional measurement of human body surface," *J. Biomed. Opt.* **13**(4), 044039 (2008).

47. W. Glinkowski et al., "Method of pectus excavatum measurement based on structured light technique," *J. Biomed. Opt.* **14**(4), 044041 (2009).

48. E. A. Ferreira et al., "Quantitative assessment of postural alignment in young adults based on photographs of anterior, posterior, and lateral views," *J. Manipulative Physiol. Therapeut.* **34**(6), 371–380 (2011).

49. M. Asher et al., "Spine deformity correlates better than trunk deformity with idiopathic scoliosis patients' quality of life questionnaire responses," *Stud. Health Tech. Informat.* **91**, 462–464 (2002).

50. J. Badawi-Fayad and E. A. Cabanis, "Three-dimensional Procrustes analysis of modern human craniofacial form," *Anat. Rec. (Hoboken)* **290**(3), 268–276 (2007).

51. F. Giavazzi et al., "Optical generation of Voronoi diagram," *Opt. Express* **16**(7), 4819–4823 (2008).

52. X. Qian et al., "Optimal embedding for shape indexing in medical image databases," *Med. Image Anal.* **14**(3), 243–254 (2010).

53. A. J. Naftel and M. J. Trenouth, "Stereo-assisted landmark detection for the analysis of changes in 3-D facial shape," *Med. Inform. Internet Med.* **29**(2), 137–155 (2004).

54. B. Kamgar-Parsi and B. Kamgar-Parsi, "Algorithms for matching 3D line sets," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(5), 582–593 (2004).

55. T. Huysmans et al., "Three-dimensional mathematical reconstruction of the spinal shape, based on active contours," *J. Biomech.* **37**(11), 1793–1798 (2004).

56. X. Q. Jiang, L. Blunt, and K. J. Stout, "Three-dimensional surface characterization for orthopaedic joint prostheses," *Proc. Inst. Mech. Eng. H* **213**(1), 49–68 (1999).

57. M. Asher et al., "The influence of spine and trunk deformity on preoperative idiopathic scoliosis patients' health-related quality of life questionnaire responses," *Spine (Philadelphia, PA 1976)* **29**(8), 861–868 (2004).

58. L. Ramirez et al., "A support vector machines classifier to assess the severity of idiopathic scoliosis from surface topography," *IEEE Trans. Inf. Technol. Biomed.* **10**(1), 84–91 (2006).