Posture Analysis and Range of Movement Estimation using Depth Maps

Miguel Reyes[1,2], Albert Clapés[1,2], Luis Felipe Mejía[1,2], José Ramírez[3], Juan R. Revilla[3], Sergio Escalera[1,2]

[1]*Dept. Matemàtica Aplicada i Anàlisi, UB, Gran Via de les Corts Catalanes 585, 08007, Barcelona*
[2]*Computer Vision Center, Campus UAB, Edifici O, 08193, Bellaterra, Barcelona*
[3]*Instituto de Fisioterapia Global Mezières, Guillem Tell 27, 08006, Barcelona*

Abstract. World Health Organization estimates that 80% of the world population is affected of back pain during his life. Current practices to analyze back problems are expensive, subjective, and invasive. In this work, we propose a novel tool for posture and range of movement estimation based on the analysis of 3D information from depth maps. Given a set of keypoints defined by the user, RGB and depth data are aligned, depth surface is reconstructed, keypoints are matching using a novel point-to-point fitting procedure, and accurate measurements about posture, spinal curvature, and range of movement are computed. The system shows high precision and reliable measurements, being useful for posture reeducation purposes to prevent musculoskeletal disorders, such as back pain, as well as tracking the posture evolution of patients in rehabilitation treatments.

Keywords: Depth maps, Physiotherapy, Posture Analysis, Range of Movement Estimation, Rehabilitation, Statistical Pattern Recognition

1 Introduction

World Health Organization has categorized disorders of the musculoskeletal system as the main cause for absence from occupational work and one of the most important causes of disability in elders in the form of rheumatoid arthritis or osteoporosis. It is estimated that 80% of world population will suffer from musculoskeletal disorders during their life.

The body posture evaluation of a subject manifests, in different degrees, his level of physic-anatomical health given the behavior of bone structures, and especially of the dorsal spine. For instance, common musculoskeletal dysfunctions or disorders (MSDs) such as scoliosis, kyphosis, lordosis, arthropathy, or spinal pain show some of their symptoms through body posture. This requires the use of reliable, noninvasive, automatic, and easy to use tools for supporting diagnostic. However, given the articulated nature of the human body, the development of this kind of systems is still an open issue.

The solution more frequently applied to measure body posture consists of the synchronization of multiple cameras, applying stereo vision techniques [3, 5]. This kind of systems use to be expensive and invasive. Moreover, it uses to require specific and restricted illumination conditions. The main alternative is accelerometers. These systems also use to be expensive, invasive, and inaccurate because of the spatial measurements of multi-axial articulations. Most of these systems only treat specific areas of the body with little configurability, which implies that therapists cannot use their own methods of analysis. A recent alternative is the use of the depth maps provided by the Microsoft Kinect device [1]. The Kinect camera uses a structured light technique to generate real-time depth maps containing discrete range measurements of the physical scene [2].

In this work, we present a novel semi-automatic system that uses RGB-Depth information to elaborate a clinical postural analysis through the examination of anthropometric values. Given a set of keypoints defined by the user, our proposed method performs the following steps: a) RGB and depth data are aligned, b) noise is removed and depth surface is reconstructed, c) user keypoints and predefined protocols are matched using a novel point-to-point fitting procedure, d) static measurements about posture and spinal curvature are accurately computed, and d) dynamic range of movement is robustly estimated. Compared to standard alternatives and supported by clinical specialists, the system shows high precision and reliable measurements to be include in the clinical routine.

The paper is organized as follows: Section 2 present the system for posture analysis and range of movement estimation. Section 3 presents the validation of the proposal, and finally, Section 4 concludes the paper.

2 Posture analysis system

We designed a full functional system devoted to help in the posture reeducation task with the aim of preventing and correcting musculoskeletal disorders. The system is composed by three main functionalities: a) static posture analysis (SPA), b) spine curvature analysis (SCA), and c) range of movement analysis (RMA). The architecture of the system is shown in Figure 1. First, a pre-processing step to remove noise and reconstruct surfaces is performed. Next, we describe each of these stages.

2.1 Noise removal and surface reconstruction

After aligning RGB and depth data [7, 11], and even though the used depth information is compelling it is still inherently noisy. Depth measurements often fluctuate and depth maps contain numerous holes where no readings are obtained. In order to obtain a valid and accurate depth map, we perform a depth preprocessing step to eliminate erroneous information caused by noise and to reconstruct surfaces not well defined. We perform the following methodology:

Noise removal: For each point we compute the mean distance from it to all its neighbors. By assuming that the resulted distribution is Gaussian with a mean and a standard deviation, all points whose mean distances are outside an interval defined by the global distances mean and standard deviation are considered as outliers.

Surface reconstruction: We use a resampling algorithm [10], which attempts to recreate the missing parts of the surface by higher order polynomial interpolation between the surrounding data points. By performing resampling, these small errors can be corrected. Figure 2 shows an example of this process [1].

[1] We experimentally found that our approach for noise removal and background reconstruction obtained better results than standard approaches based on accumulating temporal images (e.g. 30 frames of a stationary subject) for noise reduction and hole filling.

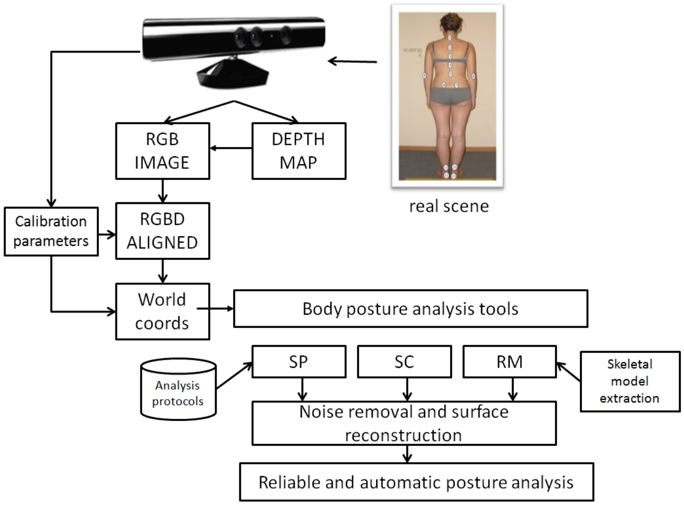
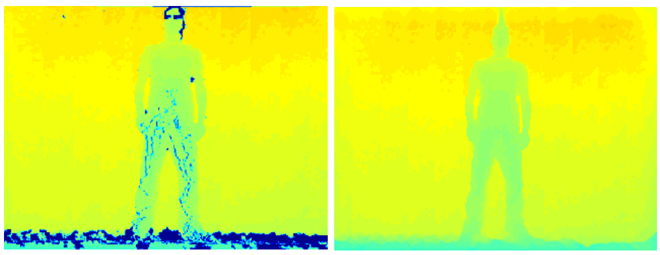


Fig. 1. Posture analysis system.



(a) (b)

Fig. 2. (a) Original depth map. (b) Filtered and resampled.

Once the system is calibrated, data is aligned, and depth maps are filtered, the user can access to the three posture facilities.

2.2 Static posture analysis (SPA)

This module computes and associates a set of three-dimensional angles and distances to keypoints defined by the user. These keypoints correspond to manual interactions of the user with the RGB data displayed in the screen (which internally is aligned with the corresponding depth data). The module also allows the therapist the possibility of designing a protocol of analysis. That is, a predefined set of angular-distance measurements among a set of body keypoints, all of them defined and saved by the user for posterior automatic matching. Figure 3 shows an example of a predefined protocol (the set of manual annotated keypoints together with the list of distance and angle relations to be computed).

In order to obtain an intelligent and automatic estimation of posture measurements, we define a correspondence procedure among manually placed virtual markers and protocol markers. We formulate markers matching as an optimization problem. Sup-

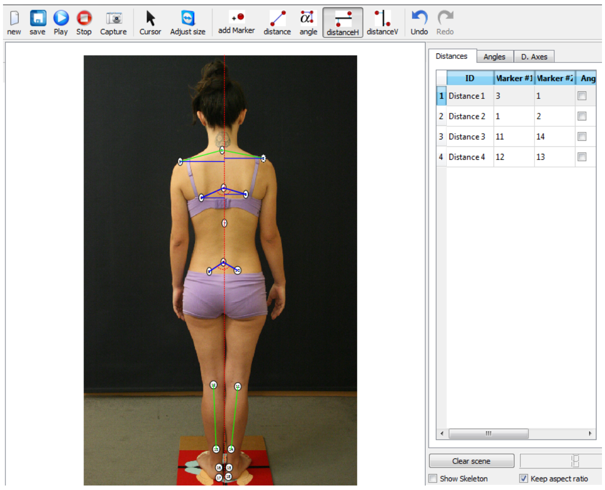


Fig. 3. Static posture analysis example.

pose a protocol analysis (template) T composed by N markers, $T = \{T_1, T_2, ..., T_N\}$, $T_i = (x_i, y_i, z_i)$, and the current analysis C composed by the same number of markers, $C = \{C_1, C_2, ..., C_N\}$ (predefined template and current set of keypoints defined by the user, respectively). Our goal is to make a one-to-one correspondence so that we minimize the sum of least square distances among assignments as follows:

$$\mathrm{argmin}_{C'} \sum_{i=1}^{N} \|C'_i - T_i\|^2 , \tag{1}$$

where C' is evaluated as each of the possible permutations of the elements of C. For this task, first, we perform a soft pre-alignment between C and T using Iterative Closest Point (ICP) [8], and then, we propose a sub-optimal approximation to the least-squares minimization problem. ICP is based on the application of rigid transformations (translation and rotation) in order to align both sequences C and T. This attempts to minimize the error of alignment $E(.)$ between the two marker sequences as follows:

$$E(\mathcal{R}, \mathcal{T}) = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} \|T_i - \mathcal{R}(C_j) - \mathcal{T}\|^2, \tag{2}$$

being $\mathcal{R}$ and $\mathcal{T}$ the rotation matrix and translation vector, respectively. $w_{i,j}$ is assigned 1 if the i-th point of T described the same point in space as the j-th point of C. Otherwise $w_{i,j} = 0$. Two things have to be calculated: First, the corresponding points, and second, the transformation $(\mathcal{R}, \mathcal{T})$ that minimizes $E(\mathcal{R}, \mathcal{T})$ on the base of the corresponding points. For this task, we apply Singular Value decomposition (SVD). At the end of the optimization, the new projection of the elements of C is considered for final correspondence. Then, Eq. 1 is approximated as follows: Given the symmetric matrix of distances M of size $N \times N$ which codifies the set of $N \cdot (N-1)/2$ possible distances among all assignments between the elements of C and T, we set a distance threshold

θ_M to define the adjacency matrix A:

$$A(i,j) = \begin{cases} 1 & \text{if} \quad M(i,j) < \theta_M \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Then, instead of looking for the set of $N!$ possible assignments of elements of C and T that minimizes Eq. 1, only the possible assignments (C_i, T_j) that satisfies $A(i,j) = 1$ are considered, dramatically reducing the complexity of the correspondence procedure[2].

2.3 Spine curvature analysis (SCA)

The objective of this task is to evaluate sagittal spine curvatures (curves of the spine projected on the sagittal plane) by noninvasive graphic estimations in kyphotic and lordotic patients. Kyphosis and lordosis are, respectively, conditions of over-curvature of the thoracic spine (upper back) and the lumbar spine (lower back). The methodology proposed by Leroux et al [4] offers a three-dimensional analysis valid for clinical examinations of those conditions. In order to perform this analysis we proceed as follows. First, the therapist places the markers on the spine. Then, a few markers are selected and the 3D curve that represents the spine is reconstructed by linear interpolation (Figure 4(c)). Finally, the anthropometric kyphosis K_a and lordosis L_a are obtained.

The geometric model to compute K_a is represented in the Figure 4(a). F divides the curve representing the thoracic spine in two asymmetric arcs with different radius. Note that the F component begins at the farthest marker (apex, corresponding to T5) and it ends at the intersection with the T2-to-T12 line. *h1* and *h2* are the distances from T2 to the intersection and the distance from the intersection to T12, respectively. Then, the summation of two angles, φ_1 and φ_2, represents the kyphosis curve value, where:

$$\begin{aligned} \varphi_1 &= 180 - 2 \cdot \arctan\left(\tfrac{h_1}{F}\right), \\ \varphi_2 &= 180 - 2 \cdot \arctan\left(\tfrac{h_2}{F}\right). \end{aligned} \tag{4}$$

L_a is calculated in a similar way, though the therapist should note the markers in the lumbar spine region. The capacity analysis of the spine is reinforced by a three-dimensional environment for a thorough examination by the therapist (Figure 4(d)). An example of spine interaction and computation are shown in Figure 4(a) and (b), respectively.

2.4 Range of movement analysis (RMA)

In order to complement the posture analysis procedure, we compute the range of movement of different body articulations. For this purpose, we perform user detection using the Random Forest approach with depth features of Shotton et al [9] and compute the skeletal model. This process is performed computing random offsets of depth features as follows:

$$f_\theta(D, \mathbf{x}) = \mathbf{D}_{\left(\mathbf{x} + \frac{\mathbf{u}}{\mathbf{D_x}}\right)} - \mathbf{D}_{\left(\mathbf{x} + \frac{\mathbf{v}}{\mathbf{D_x}}\right)}, \tag{5}$$

where $\theta = (\mathbf{u}, \mathbf{v})$, and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{\mathbf{2}}$ is a pair of offsets, depth invariant. Thus, each θ determines two new pixels relative to $\mathbf{x}$, the depth difference of which accounts for the

[2] We experimentally found that high values of θ_M obtain optimal results and reduces the computational cost in comparison to other approaches, such as Shape Context [6].

value of $f_\theta(D, \mathbf{x})$. Using this set of random depth features, Random Forest is trained for a set of trees, where each tree consists of split and leaf nodes (the root is also a split node). Finally, we obtain a final pixel probability of body part membership l_i as follows:

$$P\left(l_i|D, \mathbf{x}\right) = \frac{1}{\tau}\sum_{j=1}^{\tau} P_j\left(l_i|D, \mathbf{x}\right), \tag{6}$$

where $P\left(l_i|D, \mathbf{x}\right)$ is the PDF stored at the leaf, reached by the pixel for classification $(D, \mathbf{x})$ and traced through the tree j, $j \in \tau$. Computing the intersection borders among mean shift clusters estimated after Random Forest procedure, we obtain a three-dimensional skeletal model composed by nineteen joints. The physician then selects joint articulations and automatically obtains their maximum opening and minimum closing values measured in degrees for a certain period of time (Figure 4(e)).

3 Results

3.1 Software details

The video data uses a 8 bits VGA resolution at 30Hz, and we capture frames at 640×480 pixels, like the infrared camera. Regarding the implementation we used the Kinect SDK framework. We also used the PCL-Library to treat cloud points, and to support a free and three-dimensional visualization we used the VTK library. The user interface has been developed in multi-platform Nokia Qt technology.

3.2 Data and validation

In order to measure the precision of the proposed methodology in the different modules of the system, a battery of 500 simple tests has been labeled by three different observers, with an inter observer correlation superior to 99% for all planes (X, Y, Z). Each test contains a set of angles and distances in order to simulate an analysis protocol for the study of posture, placing twelve infrared led markers on the body of the subject. A total of 20 subjects participated in the validation of the method. In order to perform automatic validation of the tests, infrared markers are detected by means of thresholding a HSV infrared-filtered image.

Results for different distance of the device to the scene are shown in Table1. AAV and '$\circ$' correspond to the average absolute value and degree, respectively. This analysis validates the accuracy of the SPA and RMA in millimeters and degrees, respectively. Note the high precision in both tests. In addition, in order to validate the curvature analysis of the spine (SCA), we used a group of 10 patients and performed the Leroux protocol [4], placing nine markers over the spine. The relationship between lateral radiographic and anthropometric measures was assessed with the mean difference. It has used Cobb technique on the lateral radiograph in order to obtain the coefficients of kyphosis and lordosis. The results of the SPA validation are shown in Table 3.2. Moreover, after discussing with specialists in physiotherapy they agreed that the accuracy of the results is more than sufficient for diagnostic purposes.

4 Conclusion

We presented a system for semi-automatic posture analysis and range of movement estimation using depth maps. The aim of the system is to assist in the posture reeducation task to prevent and treat musculoskeletal disorders. Given a set of keypoints

Distance subject-device (m)	1,3	1,9	2,2
AAV (○ movement)	2,2	3,8	5,2
AAV (mm)	0,98	1,42	2,1
AAV (○ angles)	0,51	1,04	1,24
AAV (%)	0,46	0,77	1,3
Standard Error (%)	1,01	1,18	1,71

Table 1. Pose and range of movement precision.

	Khyposis range	Lordosis range
AAV (○)	5	6

Table 2. Validation of spinal analysis.

defined by the user, RGB and depth data are aligned, depth surface is reconstructed, keypoints are matching using a novel point-to-point fitting procedure, and accurate measurements about posture, spinal curvature, and range of movement are obtained. The system showed high precision in terms of distance, degree, and range of movement estimation. Supported by clinical specialists, the system shows high precision and reliable measurements to be include in the clinical routine.

Acknowledgements

This work is partly supported by projects TIN2009-14404-C02, IMSERSO-Ministerio de Sanidad 2011 Ref. MEDIMINDER and RECERCAIXA 2011 Ref. REMEDI.

References

1. M. Corporation. Kinect for windows sdk beta programming guide beta 1 draft version 1.1. 2012.
2. I. Freedman Barak, U. Shpunt Cambridge Ma, I. Machline Meir Ashdod, and I. Arieli Yoel Jerusalem. 2008 - depth mapping using projected patterns.pdf. page 12, 2010.
3. A. M. Koontz, Y.-S. Lin, P. Kankipati, M. L. Boninger, and R. A. Cooper. Development of custom measurement system for biomechanical evaluation of independent wheelchair transfers. volume 48, pages 101528, 2011.
4. M. A. Leroux and K. Zabjek. A noninvasive anthropometric technique for measuring kyphosis and lordosis: application for scoliosis. volume 25, pages 168994, 2000. 5. I. Mikic, M. Trivedi, E. Hunter, and P. Cosman. Articulated Body Posture Estimation from Multi-Camera Voxel Data. In CVPR, pages 455460, 2001.
5. I. Mikic, M. Trivedi, E. Hunter, and P. Cosman, Articulated Body Posture Estimation from Multi-Camera Voxel Data, CVPR, pp. 455-460, 2001.
6. G. Mori, S. Belongie, and J. Malik. Efficient shape matching using shape contexts. TPAMI, 27:18321837, 2005.
7. M. Potmesil and I. Chakravarty. A lens and aperture camera model for synthetic image generation. volume 15, pages 297305, New York, NY, USA, Aug. 1981. ACM.
8. R. San Jose Estepar, A. Brun, and C.-F. Westin. Robust generalized total least squares iterative closest point registration. MICCAI04, pages 234241, Rennes - Saint Malo, France, 2004.
9. J. Shotton, A. W. Fitzgibbon, M. Cook, and T. Sharp. Real-time human pose recognition in parts from single depth images. In CVPR, pages 12971304, 2011.
10. J. Yao, M. R. Ruggeri, and P. Taddei. Automatic scan registration using 3d linear and planar features. volume 1, pages 22:122:18, Secaucus, NJ, USA. Springer-Verlag.
11. Z. Zhang. A flexible new technique for camera calibration. TPAMI, 22(11):13301334, 2000.

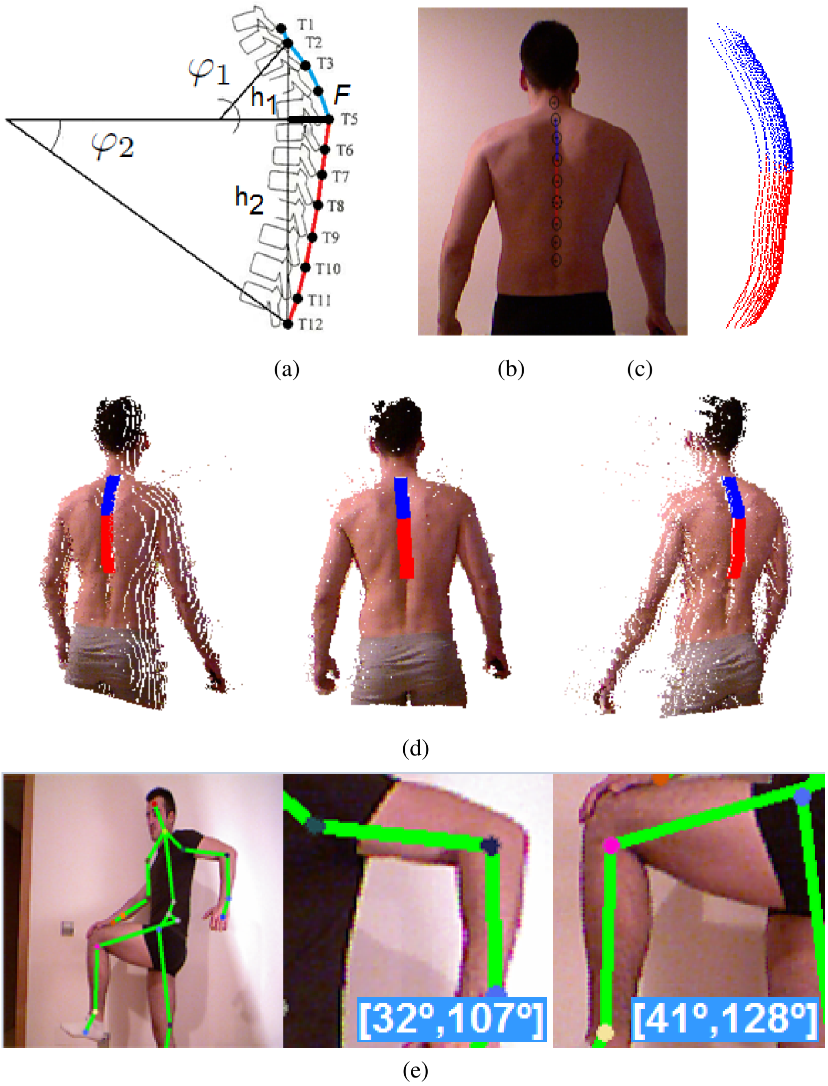


Fig. 4. (a) Geometric model to obtain anthropometric kyphosis, and lordosis value. (b) Sample of analysis. (c) Automatically reconstructed 3D spinal cloud. (d) Three-dimensional examination environment. (e) Skeletal model and example of selected articulations with computed dynamic range of movement.