

Measurement of Angles in Time-of-Flight Data for the Automatic Supervision of Training Exercises

S. Soutschek, J. Kornhuber

Psychiatrische und Psychotherapeutische Klinik, Universitätsklinikum Erlangen, Germany

A. Maier, S. Bauer, P. Kugler, J. Hornegger

Lehrstuhl für Mustererkennung, Universität Erlangen-Nürnberg, Germany

M. Bebenek, S. Steckmann, S. von Stengel, W. Kemmler

Osteoporoseforschungszentrum, Institut für Medizinische Physik, Universität Erlangen-Nürnberg, Germany

stefan.soutschek@informatik.uni-erlangen.de

Abstract—In this paper we present a novel automatic system to supervise and support rehabilitation and fitness exercises. Goal of the system is to enable autonomous training of rehabilitation patients to support their recovery from movement restrictions as well as the prevention of age-related diseases such as osteoporosis. The system uses time-of-flight technology in order to measure movement angles in real-time. This novel setup is compared to a standard tracking system using visual markers. The measurements of both systems correlate with r values between 0.99 and 0.97 depending on the task domain. The mean error of the angle measurements is between 13 and 17 degrees.

Index Terms—rehabilitation, ambient assisted living (AAL), automatic supervision of exercises, camera, time-of-flight.

I. INTRODUCTION

Sarcopenia, defined as increased loss of muscle mass and strength due to advanced age, is associated with a functional decline, affecting independent living [1], [2]. It is well established that exercise is effective in counteracting the negative consequences of aging on muscle mass and function [3], [4] throughout life (and thus positively impacts functional capacity and health).

However, due to orthopedic and cardiac limitations many elder people are unable to perform vigorous exercise programs [5], which are known to be most effective to improve strength and power. Furthermore, despite the high relevance for this group, in the elderly the acceptance of high-intensity programs to reverse the loss of muscle mass and function is rather low. Thus, for older people a gentle though effective program would be desirable. In this context, the application of high-frequency mechanical stimuli could be a possible solution.

Several have shown that whole body vibration training has a positive effect on neuromuscular parameters such as maximum strength [6], [7], [8], muscle performance [6], [7] and flexibility [9], [10]. Our own studies [11], [12], [13] confirm this effect and furthermore prove a significant positive influence on pain in the lower back and the quality of life [13]. The acceptance is, however, an essential success guarantor besides the effectiveness of an intervention measure.

Whole body vibration as an independent and autonomous training program in public facilities or homes for the elderly has the potential of a widespread implementation because of low personnel expenditures and temporal flexibility. In the ELVIS II-trial (Erlangen Longitudinal Vibration Study), a randomized, controlled study with 108 community dwelling postmenopausal women, we determined the effect of a whole body

vibration on osteoporotic risk factors. The video-animated training program consisted of dynamic exercises which were performed on the vibration plates for 20 min three times per week. The first three sessions were supervised by a qualified trainer. In the following, the participants performed the training on their own. All six weeks, the training was controlled by a trainer. The study results demonstrate a good feasibility of the video-guided training program which is shown amongst other things by above average compliance rate [13].

However, in settings with hardly attainable or completely missing staff a video-animated “do it yourself”-training with older, technical not experienced persons with little movement experience might result in several problems. The development of a vibration device which can be handled easily even by the elderly and which accurately instructs the performance of the exercises on the plate would be the solution of many of these problems. Based on acoustic and visual instruction, the operation of the device should be so simple that even technically unexperienced people can handle it without problems. Further, to avoid a wrong execution of the exercises and therewith to reduce the risk of injury, a feedback mechanism which controls the movements would be of high importance when developing a new vibration device.

Figure 1 displays the idea of the new training system. The ViVidomat is equipped with a vibration plate, a 3-D range camera, and a 2-D camera for the tracking of the exercises. At present the 2-D camera is only used to acquire lateral reference data to evaluate the quality of the tracking with the 3-D camera. In the final system, only the 3-D camera will be used. Furthermore, the 3-D camera will be inserted into the tower at the front of the system. In the current study, this is not yet the case.

II. METHODS

The novel tracking algorithm is based on time-of-flight (TOF) camera data. With TOF cameras it is possible to collect a dense depth image. Using the depth data, we track and compute different angles during the exercise to provide supervision of the training.

As reference data, we observe the same scene from the side using a normal 2-D color camera. Figure 2 displays the differences between the two modalities: While the 2-D color camera provides a conventional 2-D color image, the TOF

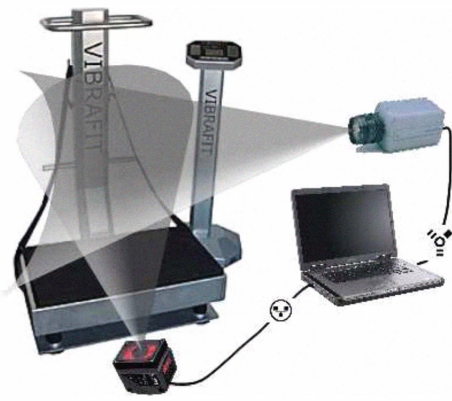


Fig. 1. The current setup of the ViVidomat System: A 3-D camera is placed in front of the subject in order to track his or her movements. For evaluation purposes this is currently also tracked from a lateral position using a 2-D camera. The 2-D camera will be removed in the final system and the 3-D range camera will be integrated into the VIBRAFIT system.

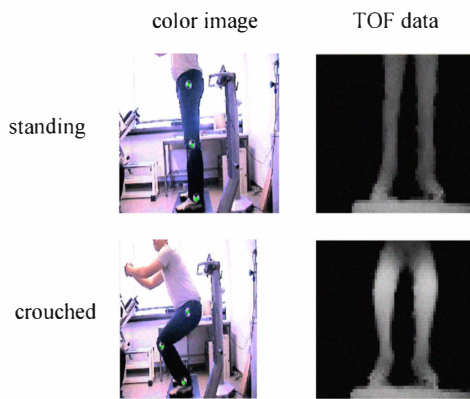


Fig. 2. Comparison of the two image modalities: The 2-D color camera is used to compute reference angle values. In the final system, only the 3-D camera will be used. It provides a dense surface image which is also used to compute the target angles.

camera is able to measure a dense depth image of the scene. Although the scene is seen from the front, the angle at the knee can be computed from the 3-D surface data.

In the following sections we will describe the 2-D and the 3-D camera systems and the algorithms which were used to extract the angle information.

A. 2-D Tracking

To calculate the reference knee angles, three black-and-white markers were attached to the pelvis, the knee and the foot ankle of the subjects. The scene was recorded from a lateral position using a standard color camera (cf. Figure 2 left). The rough positions of the markers were selected manually in the first image of the video stream. The positions were then automatically refined using a sub-pixel accurate corner detector available in OpenCV [14]. After the initialization, the marker-positions were tracked during the video stream using the Lucas-Kanade optical flow method [15], which is a very common tracking method in computer vision. The optical flow between two adjacent image frames assigns each pixel a image velocity, which describes where the pixel has moved

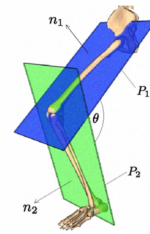


Fig. 3. Illustration of the dihedral angle θ between the intersecting planes of the point clouds P_1 , P_2 of the upper and lower limb, respectively.

from the first image to the second. We computed the optical flow using an iterative implementation available in OpenCV [16]. Using the flow vector a new position of the marker can be calculated from the old position and the marker is tracked throughout the video stream. Whenever tracking is lost, the marker-position is again manually initialized. The markers were tracked throughout the video stream and the 2-D positions were saved for each frame. The knee angles were then calculated from the 2-D positions using simple trigonometry.

III. 3-D TRACKING

Time-of-flight imaging is an emerging technology that provides a direct way to acquire 3-D surface information with a single sensor. Active light sources attached to the camera emit an incoherent cosine-modulated optical signal in the non-visible spectrum of the infrared range (850 nm). The light is reflected by the scene and enters the monocular camera where each TOF sensor element performs a correlation of the local optical signal with the electrical reference of the emitted signal [17]. Based on this correlation, the phase shift ϕ representing the propagation delay between both signals is measured. The distance d can then be computed straightforward,

$$d = \frac{c}{2 \cdot f_{mod}} \cdot \frac{\phi}{2\pi} \quad (1)$$

where f_{mod} denotes the modulation frequency and c the speed of light. For reasons of periodicity of the cosine-shaped modulation signal, the validity of this equation is limited to distances smaller than $c/(2 \cdot f_{mod})$. At a typical modulation frequency of 20 MHz, the non-ambiguity range is about 7.5 m. This is convenient for the application we want to address.

The TOF imaging technology benefits from several advantages over other 3-D surface acquisition techniques. The device is compact, portable and easy to integrate. It delivers precise metric information in the sensor coordinate system in real-time, no calibration steps are necessary. In respect to potential applications in the security, automotive and consumer electronics industry [18], a decrease of manufacturing costs can be expected with mass production being an all-solid-state off-the-shelf technology. Recently, first medical applications using TOF sensors have been proposed [19], [20]. In this paper, we investigate the possibility of extracting joint angles of the human body from a frontal point of view in real-time.

The proposed algorithm for 3-D limb tracking and joint angle calculation consists of two main parts. First, the point cloud D provided by the camera with a resolution of $w \times h$,

$$D = [d_{i,j}], \quad i \in \{0, 1, \dots, w-1\}, j \in \{0, 1, \dots, h-1\} \quad (2)$$

is partitioned into two sets, one for the m three-dimensional points of the upper limb and one for the n points of the lower limb.

$$P_1 \in \mathbb{R}^{m \times 3}, P_2 \in \mathbb{R}^{n \times 3}, \quad P_1, P_2 \subseteq D \quad (3)$$

Second, we fit a plane into each point cloud and calculate the dihedral angle θ between the intersecting planes, cf. Figure. 3. This estimation is fast and delivered reliable results in the experiments.

In terms of preprocessing, a bilateral filter [21] is applied on the raw data. From the setup of the ViVidomat system, the coarse volume of interest (VOI) is known a priori. Hence, non-relevant points that are located outside of this VOI are neglected from scratch to reduce the amount of data and to identify 3-D points that belong to the lower body. The subdivision into upper and lower leg is initiated with thresholds motivated from anthropometry as the upper third of the TOF image contains the upper part of the leg and the second third shows the lower part of the leg.

Based on the two data sets P_1, P_2 , the best fitting plane in terms of least squares is determined using the singular value decomposition (SVD). Initially, the data is translated to the respective centroid coordinates. Computing the SVD of P_i ,

$$P_i = U_i \Sigma_i V_i^T \quad (4)$$

then yields the principal axes of a confidence ellipsoid. The least squares 3-D plane contains the centroid of the respective point cloud c_i and its normal vector n_i is the third and last column of V [22]. In case of inconsistent data ($\sigma_3 = 0$) the calculation is aborted. With a PMD O3D camera delivering 3072 points (64×48 pixels), each plane P_i contains approximately 250 points allowing for a robust plane estimation. Eventually, the dihedral angle θ between the planes P_1, P_2 is given via the scalar product of its normals:

$$\cos(\theta) = n_1 \cdot n_2 \quad (5)$$

IV. MATERIAL

In this first preliminary study, we collected data of a healthy person using the VIBRAFIT System. The data were collected at the Osteoporosis Research Center of the University Erlangen-Nuremberg. As first exercise to be automatically tracked by the system we selected squats. This exercise has several advantages:

- The exercise is also suitable for elderly persons. This is important as this group is one of the focus groups of the system.
- Ground truth is obtained easily as the angle at the knee of the test person from the side using 2-D tracking.
- There are no occlusions in this exercise. This problem is omitted for the time being. It will be topic of future work.

TABLE I

OVERVIEW ON THE RESULTS: THE CORRELATION BETWEEN THE ANGLES MEASURED FROM THE SIDE AND FROM THE FRONT IS BETWEEN 0.97 AND 0.99 DEPENDING ON THE TASK. THE ABSOLUTE MEAN ERROR IS BETWEEN 13 AND 17 DEGREES.

sequence	vibration	# of frames	r	ϵ [degrees]
T1	off	242	0.990	16.9
T2	on	149	0.976	13.6

In the following, we will describe the recording setup, its geometry and the data which were obtained.

A. Recording Setup

The recording setup consists of four devices, a Sony DFW-VL500 camera to record the 2-D reference data, a PMD O3D video range camera to acquire the 3-D frontal depth map, the VIBRAFIT System to simulate the training session and a standard PC to store, synchronize and process the collected data. The complete setup is arranged as shown in Figure 1.

The range camera is placed at a position of 1.12 m in front of the VIBRAFIT System and at a height of 0.52 m. This position was chosen, as it is already a possible place for the final system. The 2-D camera is placed at 1.80 m lateral to the VIBRAFIT System at a height of 0.80 m. This position allows for acquiring the complete profile of the test subject. This is necessary to capture the markers, which are attached to the test subject as seen in Figure 2.

B. Data

In order to perform a first evaluation, we recorded two evaluation sequences. The 3-D and the 2-D camera images were taken synchronously. There is always a matching 2-D image to the respective 3-D data. The frame rate of all sequences was five frames per second.

In both sequences, the same test person performs squats. In order to investigate the influence of the vibration on the angle estimation, the vibration plate was turned off for the first sequence T1. In the sequence T2 the vibration plate was turned on.

V. RESULTS

For the evaluation of the accuracy of the system we employed two different measurements. In order to show that the measured angles relate well to each other we use the correlation r of the measured values between each other after Pearson [23].

The other evaluation measure is the average mean error ϵ . It is computed as follows:

$$\epsilon = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2} \quad (6)$$

Table I reports the results which were obtained for both sequences. The correlations are high and significant with 0.990 without vibration and 0.976 with vibration. Also the absolute angular measurements of the 3-D data agree well with the 2-D data. The mean angular error is 16.9 degrees without vibration and 13.6 with the vibration plate turned on. Figure 4 displays

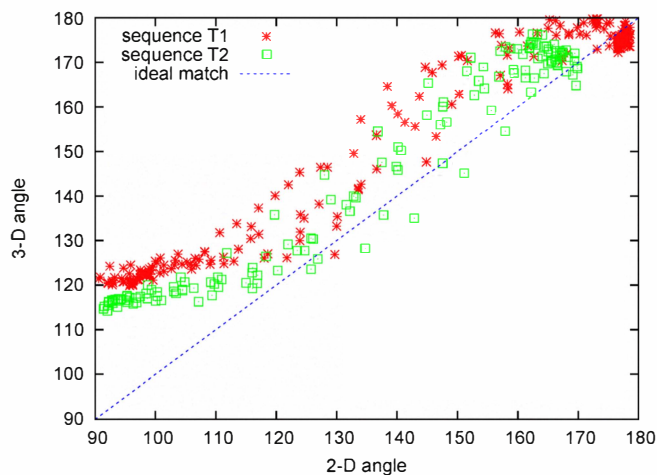


Fig. 4. 2-D and 3-D measurements in comparison: The measured 3-D angle fits well to the measured 3-D angle in both sequences. However, if the angle is close to 90 degrees, the 3-D camera returns too high values for the angle.

a plot of 2-D measured angles against 3-D measured angles of the T1 and the T2 sequence. Again, the high agreement is demonstrated. Note, that the measurements with the 3-D camera degrade slightly when the 2-D angle is close to 90 degrees. This is caused by the segmentation procedure: The hip of the subject enters the field-of-view and as we do not compensate for this, it affects the upper regression plane. Hence, the regression plane is moved upwards and the angular computation between both regression planes returns higher values.

VI. SUMMARY

In this paper, we presented two algorithms for angular measurements of the knee joint from lateral 2-D and frontal 3-D surface data. The 2-D algorithm relied on manual identification of the markers and was used as reference. The 3-D algorithm detected candidate points from the TOF range data for the upper and the lower part of the leg. Both point clouds were used to estimate regression planes. The angle was then determined as the angle between the two surface normals of the regression planes.

The 3-D measurements showed stable results. The obtained correlation coefficients between the reference angle and the 3-D measured angle ranged from 0.97 to 0.99. The average error in the angle was between 13 and 17 degrees. In general, the correlations showed to be slightly worse if the vibration plate of the ViVidomat was active. However, the decrease in accuracy was not severe. The training exercise could be tracked successfully in both states.

In the future, we want to expand the tracking algorithms by more sophisticated models of the lower body. Such models will be able to improve the accuracy of the system, as they will be able to detect body parts which enter the field-of-view, e.g., the hip. This information can then be used to adjust angular computations accordingly and will allow the expansion of the system to more complex training exercises which also may include occlusions.

REFERENCES

- [1] M. Abate, A. Di Iorio, D. Di Renzo, R. Paganelli, R. Saggini, and G. Abate, "Frailty in the elderly: the physical dimension," *Eura Medico-physis*, vol. 43, no. 3, pp. 407–415, 2007.
- [2] W. J. Evans, "Reversing sarcopenia: how weight training can build strength and vitality," *Geriatrics*, vol. 51, no. 5, pp. 46–47, 1996.
- [3] T. J. Doherty, "Invited review: Aging and sarcopenia," *J Appl Physiol*, vol. 95, no. 12, pp. 1717–1727, 2003.
- [4] A. Macaluso and G. De Vito, "Muscle strength, power and adaptations to resistance training in older people," *Eur J Appl Physiol*, vol. 91, pp. 450–472, 2004.
- [5] G. R. Hunter and M. S. Treuth, "Relative training intensity and increases in strength in older women," *J Strength and Cond Res*, vol. 9, pp. 188–191, 1995.
- [6] C. Delecluse, M. Roelants, and S. Verschueren, "Strength increase after whole-body vibration compared with resistance training," *Med Sci Sports Exerc*, vol. 35, no. 6, pp. 1033–41, 2003.
- [7] M. Roelants, C. Delecluse, M. Goris, and S. Verschueren, "Effects of 24 weeks of whole body vibration training on body composition and muscle strength in untrained females," *Int J Sports Med*, vol. 25, no. 1, pp. 1–5, 2004.
- [8] M. Sigrist, C. Lammel, and D. Jeschke, "Krafttraining an konventionellen bzw. oszillierenden Geräten und Wirbelsäulengymnastik in der Prävention der Osteoporose bei postmenopausalen Frauen," *Dtsch Z Sportmed*, vol. 57, no. 7, pp. 182–188, 2006.
- [9] V. B. Issurin, D. G. Liebermann, and G. Tenenbaum, "Effect of vibratory stimulation training on maximal force and flexibility," *J Sports Sci*, vol. 12, no. 6, pp. 561–566, 1994.
- [10] W. A. Sands, J. R. McNeal, M. H. Stone, E. M. Russel, and M. Jemni, "Flexibility enhancement with vibration: Acute and long-term," *Med Sci Sport Exerc*, vol. 38, no. 4, pp. 720–725, 2006.
- [11] S. von Stengel, W. Kemmler, K. Engelke, M. Bebenek, J. L. Mayhew, and W. A. Kalender, "Effects of whole body vibration training on the fracture risk of postmenopausal women. Preliminary results of the Erlangen Longitudinal Vibration Study (ELVIS)," *Med Sci Sports Exerc*, vol. 40, p. 313, 2008.
- [12] S. von Stengel, W. Kemmler, K. Engelke, and W. A. Kalender, "Einfluss eines 18-monatigen Ganzkörpervibrationstrainings auf Frakturrisikofaktoren postmenopausaler Frauen - Vorläufige Ergebnisse der Erlanger Längsschnitt Vibrations-Studie (ELVIS)," *Osteologie*, vol. 17, no. 1, p. 7, 2008.
- [13] S. von Stengel, W. Kemmler, W. A. Kalender, and K. Engelke, "Einfluss von unterschiedlichen Ganzkörper-Vibrationstrainings-Geräten auf Knochenmarker und Parameter der neuromuskulären Leistungsfähigkeit bei postmenopausalen Frauen," *Osteologie*, vol. 18, no. Suppl. 1, p. 14, 2009.
- [14] Intel, "OpenCV open source computer vision library," 2009, <http://www.intel.com/technology/computing/opencv/index.htm>.
- [15] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *From Proceedings of Imaging Understanding Workshop*, 1981, pp. 121–130.
- [16] J. Bouguet, "Pyramidal implementation of the lucas-kanade feature tracker: description of the algorithm," 2000, technical report, OpenCV Document, Intel Microprocessor Research Labs.
- [17] Z. Xu, R. Schwarte, H. Heinol, B. Buxbaum, and T. Ringbeck, "Smart pixel - photometric mixer device (pmd) / new system concept of a 3d-imaging-on-a-chip," in *5th International Conference on Mechatronics and Machine Vision in Practice*, 1998, pp. 259–264.
- [18] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight sensors in computer graphics," in *Eurographics 2009 - State of the Art Reports*, M. Pauly and G. Greiner, Eds., Eurographics Association. Eurographics, March 2009, pp. 119–134.
- [19] C. Schaller, J. Penne, and J. Hornegger, "Time-of-Flight Sensor for Respiratory Motion Gating," *Medical Physics*, vol. 35, no. 7, pp. 3090–3093, 2008.
- [20] C. Schaller, C. Rohkohl, J. Penne, M. Stürmer, and J. Hornegger, "Inverse C-arm Positioning for Interventional Procedures Using Real-Time Body Part Detection," in *MICCAI 2009, Part I, LNCS 5761*, G.-Z. Y. et al., Ed., Berlin - Heidelberg, 2009.
- [21] S. Paris and F. Durand, "A fast approximation of the bilateral filter using a signal processing approach," *Int. J. Comput. Vision*, vol. 81, no. 1, pp. 24–52, 2009.
- [22] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 3rd ed. The John Hopkins University Press, 1996.
- [23] K. Pearson, "Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity and Panmixia," *Philosophical Transactions of the Royal Society of London*, vol. 187, pp. 253–318, 1896.