

Accuracy and Robustness of Kinect Pose Estimation in the Context of Coaching of Elderly Population

Štěpán Obdržálek¹, Gregorij Kurillo¹, Ferda Ofli¹, Ruzena Bajcsy¹,
Edmund Seto², Holly Jimison³ and Michael Pavel³

Abstract—The Microsoft Kinect camera is becoming increasingly popular in many areas aside from entertainment, including human activity monitoring and rehabilitation. Many people, however, fail to consider the reliability and accuracy of the Kinect human pose estimation when they depend on it as a measuring system. In this paper we compare the Kinect pose estimation (skeletonization) with more established techniques for pose estimation from motion capture data, examining the accuracy of joint localization and robustness of pose estimation with respect to the orientation and occlusions. We have evaluated six physical exercises aimed at coaching of elderly population. Experimental results present pose estimation accuracy rates and corresponding error bounds for the Kinect system.

I. INTRODUCTION

Observation of human activity through various sensor technologies is becoming increasingly popular in applications of remote healthcare delivery and disease management. Until recently, the motion capture systems with active or passive markers have been used predominantly in the study of human movement kinematics where one is interested in the temporal advances of joint position or angles. The motion capture systems, however, have been due to their size and cost limited to biomechanics or kinesthetic laboratories rather than physician's offices, therapy outfits, or even homes. The human motion can also be measured through various vision based cameras using marker-based or marker-less methods to extract the kinematics.

Recently released Microsoft Kinect camera [1], primarily intended as a human natural interface for the Microsoft gaming system Xbox 360, has found itself in the mainstream of development of low-cost alternatives for rehabilitation and movement analysis. The Kinect camera captures depth and color images with 30 frames per second (fps), generating a cloud of three-dimensional (3D) points from an infrared pattern projected onto the scene. The Kinect Software Development Kit (SDK) features real-time tracking of human limbs for the gesture-based interaction. The underlying body tracking algorithm [2], which was trained on a large dataset of depth-images from able bodied users for in-game interactions, has several assumptions, such as users are standing,

the view is unobstructed, assumes an able bodied user (e.g. not in a wheelchair), body limbs are away from the trunk and not interacting with objects.

When applying the Kinect system to human pose estimation in a context other than gaming, it is important to examine the accuracy and robustness of the algorithms. In our work we are planning to use the Kinect system for observation and on-line feedback for coaching elderly people. In the paper we compare the Kinect pose estimation with more established techniques for pose estimation from motion capture data and analyze the accuracy and robustness. We report pose estimation accuracy rates and corresponding error bounds for a set of representative exercises, including upper and lower extremities, in standing and sitting position.

II. RELATED WORK

The 3D depth accuracy of the Kinect camera has been evaluated quite extensively [3], [4] showing that after calibration, the camera can provide accuracy of depth reconstruction in the order of 1-4 cm at the range of 1-4 m. On the other hand, the accuracy of the joint positions from the Kinect pose estimation has not been evaluated previously.

Several studies have recognized the advantages of using an inexpensive depth camera, such as the Microsoft Kinect, for rehabilitation and assessment of body function. Stone and Skubic [5] have used several Kinect cameras to measure temporal and spatial gait parameters for in-home assessment. In their study they compared parameters obtained from the Kinect with parameters obtained from a motion capture system; however, they were only interested in determining position of the feet and not whole body. Huang [6] developed Kinerehab system based on the Kinect camera to assist rehabilitation of patients with muscle atrophy and cerebral palsy, however, no evaluation of the accuracy was presented. Lange et al. [7] presented an interactive game-based rehabilitation tool aimed at improving the balance in adults after neurological injury. The system was applied in a pilot study to evaluate the acceptability of Kinect-based therapy, however, no report was given on the accuracy of the results. More extensive evaluation of an alternative open source pose estimation algorithm, Flexible Action and Articulated Skeleton Toolkit (FAAST), was presented by Schonauer et al. [8] who used the Kinect for real-time feedback in treadmill-based training. The authors compared position of hands and feet using motion capture and Kinect, reporting errors in the range of 5-7 cm. Up to date we have not found any

¹Š. Obdržálek, G. Kurillo, F. Ofli and R. Bajcsy are with Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley CA 94720, USA (xobdrzal, gregorij, fofli, bajcsy) @ eecs.berkeley.edu

²E. Seto is with School of Public Health, University of California at Berkeley, Berkeley CA 94720, USA (seto @ berkeley.edu)

³H. Jimison and M. Pavel are with Oregon Health & Science University, Portland OR 97239, USA (jimisonh, pavelm) @ ohsu.edu

evaluations done on the official Microsoft Kinect SDK pose estimation.

We recognize that the accuracy of the body tracking to great extent depends on the type of poses being observed, distance from the sensor and possible occlusions. Due to size limitation of the paper, we focus on six exercises that are more challenging for the pose estimation algorithm since the subject is either seated or positioned next to a chair. The tracking in seating position is especially relevant in a rehabilitation context as users may be bounded to a wheelchair.

III. ACQUISITION SYSTEMS

We have simultaneously captured human poses using a marker-based motion capture system Impulse (PhaseSpace Inc., [9]) and the Microsoft Kinect. Impulse motion capture system is capable of tracking 3D location of active LED markers (with unique IDs) with the frequency of 480 Hz and sub-millimeter accuracy. We used nine infra-red cameras positioned in a circular fashion to cover workspace of about $3\text{ m} \times 3\text{ m} \times 3\text{ m}$. We have built a custom tight-fitting motion capture suit containing 43 markers roughly positioned at standard body landmarks. A skeletonization process then fits the recorded 3D marker data with an articulated kinematic chain corresponding to human body.

We have applied two different skeletonization methods using commercial software. In the first method, using PhaseSpace Recap software, the skeletons are generated at the time of data acquisition. A two-step calibration procedure is required prior to the acquisition: (a) the markers are manually mapped to the corresponding limbs of a predefined skeleton structure and (b) the limb lengths are estimated based on a training procedure. The procedure requires the subject to perform a range of circular movements, one joint at a time (ankles, knees, hips, wrists and elbows).

The second skeletonization was done offline using Autodesk MotionBuilder software [10]. The calibration procedure here was similar to Recap: after the markers were manually associated with different sections of human body, a generic human model had to be adjusted to the captured LED positions by scaling, translating, and rotating the model and its segments. As opposed to Recap, the second step requires manual interaction with the virtual model. One of the drawbacks of the MotionBuilder pose estimation algorithm is in its inability to robustly handle missing or imprecise marker measurements.

Compared to the marker-based system, the Kinect works with much denser but less precise 3D data. A matrix of 320×240 depth measurements with precision in the centimeter range is captured and used for the skeletonization. The pose estimation algorithm is fully automatic and does not require, nor allow, any user interaction, calibration or correction.

In our experiments we have used a single Kinect camera positioned at about 3 m distance from the subject – a minimum distance needed to see whole body of a human. Poses were captured at 30 Hz.

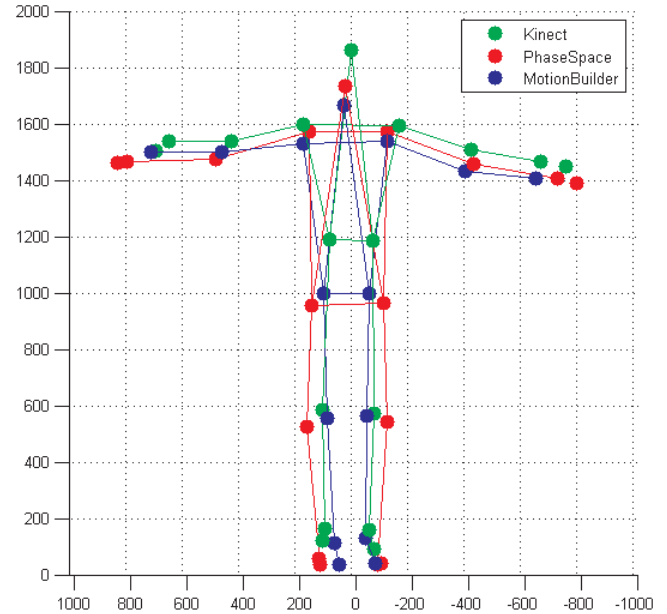


Fig. 1. Comparison of human pose representation by the three systems: Kinect, PhaseSpace Recap and Autodesk MotionBuilder. Noticeable differences are in unnaturally high hip joints by the Kinect and the choice to represent heels instead of ankles by the PhaseSpace system.

IV. EXPERIMENTS

Initially we have recorded a 30-minute coaching session that would be normally applied for daily exercise routine of elderly. From the video recording we have identified representative exercises that focus on the upper and lower extremities and captured four subjects performing these exercises. The captured data include spatial coordinates of 43 LED markers attached to the subjects, measured by PhaseSpace Impulse system, poses derived from these markers by PhaseSpace Recap tool and poses estimated by the Microsoft Kinect. After the recording another set of poses was obtained by the post-processing of marker data using Autodesk MotionBuilder.

The setup and calibration of the LED-based motion capture system took about half an hour for each subject. After putting on the suit, the subject undertook a calibration procedure to calibrate the position of the joints and limb lengths by the motion capture software. For the Kinect camera, a suitable location and elevation was determined to capture the whole body in both standing and sitting positions. No other calibration was needed.

The captured data were post-processed by first aligning all the modalities both in spatial and temporal domain. The temporal synchronization was obtained from a clapping action at the beginning of each sequence and from timestamps assigned to each measurement. For the spatial alignment we took shoulder, elbow and knee locations over all sequences and computed a rigid 3D transformation (rotation and translation) that minimized the square distance of the joints, while ignoring outliers.

The three different methods of pose estimation produced



Fig. 2. The six exercises in the study: ‘Knee Lift’, ‘Cops and Robbers’, ‘Deep Breathing’, ‘Pendulum Legs’, ‘Stand ups’ and ‘Line tapping’.

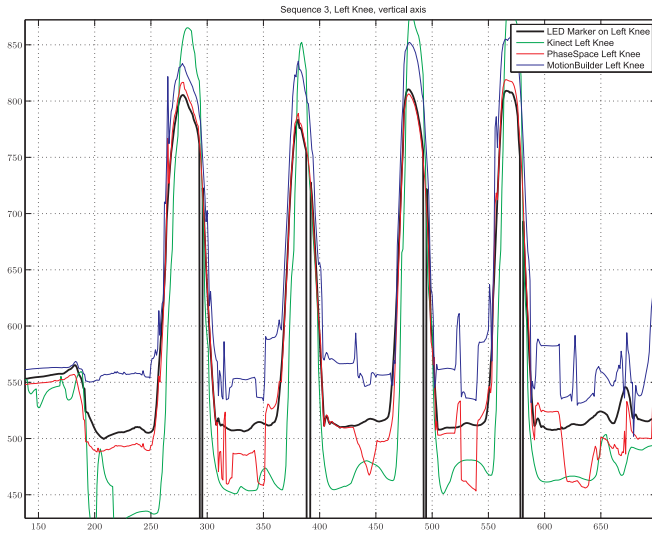


Fig. 3. Vertical position of knee joint (from the ground level) during the Knee Lift exercise. The trajectory of the LED marker attached close to the knee (black line) is close to ground truth trajectory. Of the skeleton fitting algorithms the magnitude of the movement is best captured by PhaseSpace, with Kinect consistently reporting larger and MotionBuilder smaller numbers. Amount of noise is comparable for all three systems.

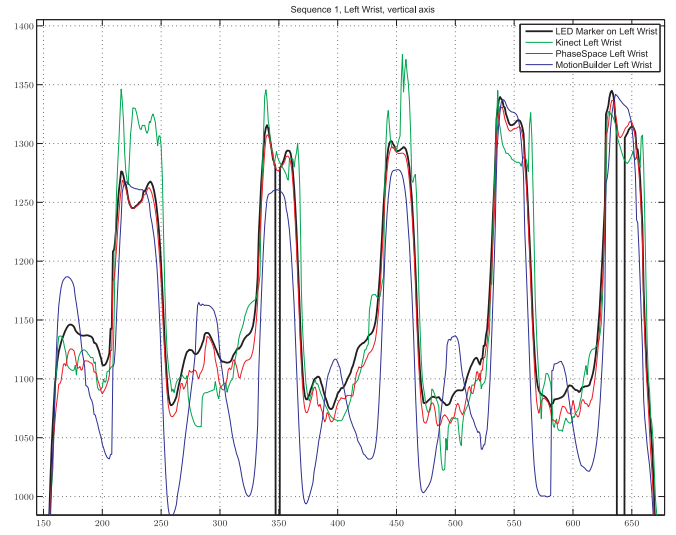


Fig. 4. Vertical position of wrist joint (from the ground level) during ‘Cops and Robbers’ exercise. The joint location is best estimated by the PhaseSpace system.

kinematic models (*i.e.* skeletons) which differ in the number and position of joints, arrangements of which are not fully anthropometric. Figure 1 shows the three skeletons obtained for a typical T-pose body configuration. The most significant differences are the unnaturally high placement of hip joints by the Kinect and the choice to represent heels instead of ankles by the PhaseSpace system. Both Recap and MotionBuilder pre-calibrate the limb (bone) lengths which are then kept constant through the entire sequence. The Kinect on the other hand has no such calibration procedure and the limb lengths vary from frame to frame.

We present results measured in six exercises. In ‘Knee Lift’ the sitting subject raises knee as high as (s)he can, alternating between left and right leg. In ‘Cops and Robbers’ exercise, upper arms are kept horizontal (elbows at shoulder

height) and lower arms are periodically raised from horizontal to vertical position and lowered back. Third exercise is ‘Deep Breathing’, with only minimal torso movement during inhalations and exhalations. Fourth is called ‘Pendulum Legs’, when, standing behind a chair for support, left legs are swung to the left and right legs to the right. To avoid occlusion by the chair, we have recorded this exercise from behind. Fifth is a balance exercise ‘Stand ups’ from a chair, with arms crossed on chest. The final sixth exercise is called ‘Line tapping’ in which a standing subject taps a line marked on the ground in front of him/her. Figure 2 illustrates postures typical for the exercises. The exercises were recorded at four different orientations of the subject, ranging from frontal to side view in 30° increment (*i.e.* the angle between camera optical axis and the sagittal plane was 0°, 30°, 60° and 90°).

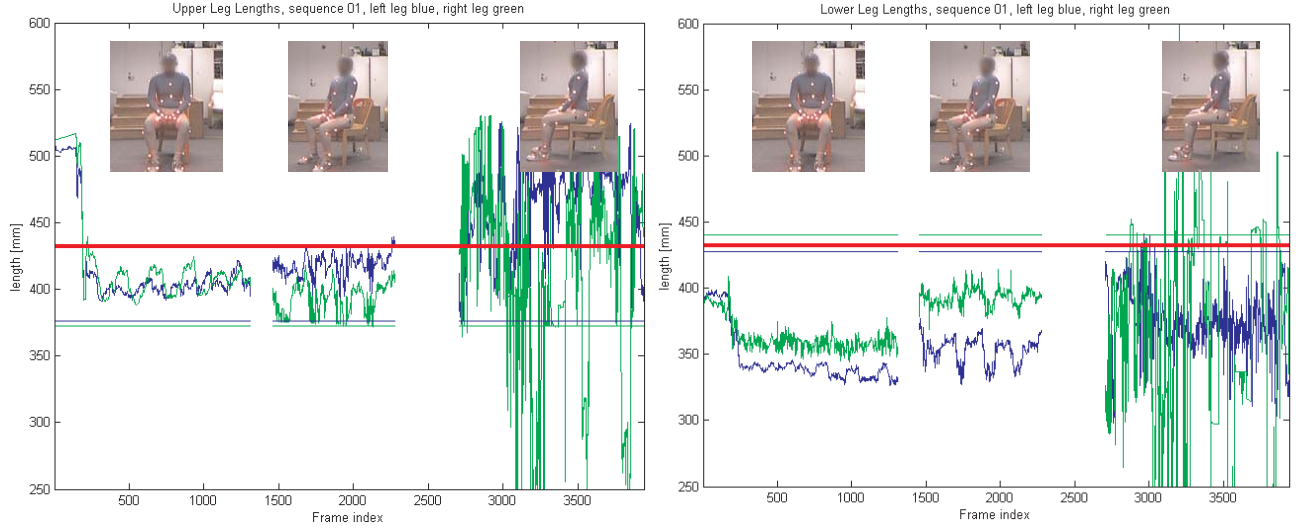


Fig. 5. Upper and lower leg bone lengths determined from the inferred poses as observed in three different orientations of the subject with respect to the camera. The Kinect estimates body geometry in every frame, varying joint to joint distances.

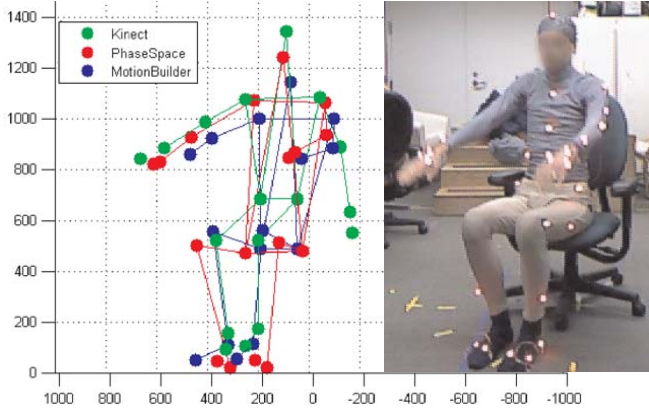


Fig. 6. Example of a failed skeletonization by the Kinect (green skeleton). The chair rest was mistaken for the left arm.

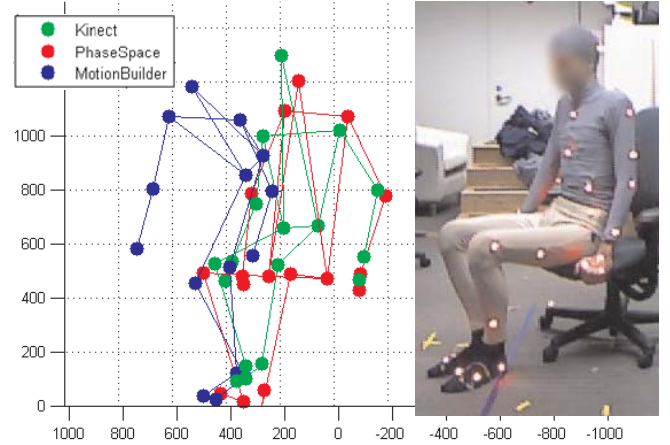


Fig. 7. Example of a failed skeletonization by the MotionBuilder (blue skeleton). An incorrectly interpolated trajectory of an occluded marker displaced the overall location of the skeleton.

A. Pose Estimation Accuracy

In this section we analyze the accuracy of the estimated skeletons. With exact location of physical joints unavailable, we are using the position of selected LED markers as ground truth. We have manually identified parts of the recorded sequences where knee and wrist markers were well vertically aligned with the physical joints, and used the position on the vertical axis (ground distance) as a ground truth.

Figures 3 and 4 show ground distance of left knee and left wrist for the first two exercises. Ground truth trajectories of the LED markers are drawn with thicker black line, the joints computed by Kinect are green, joints by PhaseSpace's Recap are red and joints by MotionBuilder are blue. From the three, the best accuracy is achieved by Recap where the typical error between the LED marker and estimated joint is about 1 cm most of the time. Typical errors of the Kinect and MotionBuilder skeletons are about 5 cm.

B. Robustness of Pose Estimation

The markerless skeleton tracking of Kinect depends solely on dense depth information and thus frequently fails due to occlusions (*e.g.* self-occlusion by other body parts, especially if only a single Kinect device is used), non-distinguishing depths (limbs close to the body) or clutter (other objects in the scene, *e.g.* a chair). Figure 6 shows an example where an armrest of the chair was mistaken for the left arm.

The marker-based systems infer the skeleton from only a small number of sparse measurements and as such are also susceptible to occlusions of individual markers. For example, sitting in a chair with the arms folded on the lap hides large number of markers, consequently providing insufficient information for proper skeleton fitting. The example in Figure 7 shows a situation where the MotionBuilder did not obtain a reasonable pose due to missing readings of markers

on hips. On the other hand, proximity of other body parts, or other markers, does not influence the localization of markers, as they are uniquely identified. Nor do other objects in the scene intervene with the capture unless they occlude the markers.

C. Comparing Kinect To PhaseSpace's Recap

From the observations presented above we concluded that the Recap system by PhaseSpace provides the most reliable and accurate pose estimation of the available methods. In this section we compare the skeletons computed by Kinect to the ones by PhaseSpace. For each pose t and joint $j \in \{\text{shoulder, elbow, wrist, hip, knee, ankle}\}$ we compute the euclidean distance $D_{t,j} = \|\mathbf{x}_{t,j,\text{Kinect}} - \mathbf{x}_{t,j,\text{PhaseSpace}}\|_2$ between 3D positions \mathbf{x} of Kinect's and PhaseSpace's joints. The distance is not expected to be zero due to different underlying human models, but rather of a constant value μ_j . The differences in distances due to inaccurate pose estimation are modeled as Gaussian with standard deviation σ_j . In case of failed pose detection, *e.g.* due to occlusion, the distances between Kinect and PhaseSpace joints are assumed to have a uniform distribution on interval $(0, 600)$ mm.

Combined, we consider the distances $D_{t,j}, t \in (1 \dots T)$ to be realizations of random variable

$$\mathbf{D}_j \sim \rho_j \cdot \mathcal{N}(\mu_j, \sigma_j^2) + (1 - \rho_j) \cdot \mathcal{U}(0, 600\text{mm}),$$

a mixture of Gaussian and Uniform distributions. The mixture weight ρ_j corresponds to the ratio of inliers (cases where the pose estimation did not fail).

Joint	mean dist μ_j	std. dev σ_j	mix ρ_j	outliers $\frac{ O^P }{T}$	indicated I $\frac{ I^P \cap I^K }{ I^P }$	indicated O $\frac{ O^P \cap O^K }{ O^P }$
Left Hip	213 mm	27 mm	1.00	0%	100%	—
Right Hip	234 mm	23 mm	1.00	0%	100%	—
Left Knee	79 mm	16 mm	0.78	19%	100%	0%
Right Knee	77 mm	12 mm	0.74	24%	100%	1%
Left Ankle	146 mm	39 mm	0.87	9%	97%	8%
Right Ankle	193 mm	38 mm	0.86	11%	99%	6%
Left Shoulder	49 mm	18 mm	0.85	12%	100%	0%
Right Shoulder	44 mm	17 mm	0.85	12%	100%	0%
Left Elbow	57 mm	25 mm	0.82	14%	95%	8%
Right Elbow	76 mm	31 mm	0.83	12%	96%	0%
Left Wrist	67 mm	30 mm	0.81	14%	97%	86%
Right Wrist	76 mm	42 mm	0.81	13%	93%	61%

TABLE I

PARAMETERS OF DISTRIBUTION \mathbf{D} OF L_2 DISTANCES BETWEEN JOINTS OF PHASESPACE AND KINECT SKELETONS, AND THE SUCCESS RATE IN IDENTIFYING INLIERS AND OUTLIERS BY THE KINECT. ALL EXERCISES, ALL SUBJECTS, FRONTAL VIEW.

Maximum likelihood estimates of parameters ρ_j, μ_j and σ_j of the distribution, accumulated over the six exercises and all subjects, and split by the camera-to-user orientation angle, are shown in Tables I, II, III, and IV. Largest distances are seen for hips and ankles, corresponding to the largest differences between the respective human models. The outlier ratio, as well as variations in distances, are higher for joints down the kinematic chain (wrist and ankles) than for joints at torso.

Joint	mean dist μ_j	std. dev σ_j	mix ρ_j	outliers $\frac{ O^P }{T}$	indicated I $\frac{ I^P \cap I^K }{ I^P }$	indicated O $\frac{ O^P \cap O^K }{ O^P }$
Left Hip	216 mm	30 mm	1.00	0%	100%	—
Right Hip	229 mm	22 mm	1.00	0%	100%	—
Left Knee	74 mm	20 mm	0.80	17%	99%	2%
Right Knee	88 mm	21 mm	0.81	15%	99%	2%
Left Ankle	148 mm	36 mm	0.86	10%	92%	10%
Right Ankle	158 mm	52 mm	0.87	8%	92%	2%
Left Shoulder	46 mm	18 mm	0.83	14%	100%	1%
Right Shoulder	55 mm	21 mm	0.84	13%	100%	1%
Left Elbow	53 mm	25 mm	0.73	22%	96%	4%
Right Elbow	72 mm	29 mm	0.82	13%	98%	42%
Left Wrist	69 mm	29 mm	0.74	21%	98%	38%
Right Wrist	69 mm	36 mm	0.76	18%	90%	60%

TABLE II

PARAMETERS OF DISTRIBUTION \mathbf{D} OF L_2 DISTANCES BETWEEN JOINTS OF PHASESPACE AND KINECT SKELETONS, 30° VIEW.

Joint	mean dist μ_j	std. dev σ_j	mix ρ_j	outliers $\frac{ O^P }{T}$	indicated I $\frac{ I^P \cap I^K }{ I^P }$	indicated O $\frac{ O^P \cap O^K }{ O^P }$
Left Hip	216 mm	32 mm	1.00	0%	81%	—
Right Hip	225 mm	24 mm	0.99	0%	82%	96%
Left Knee	78 mm	22 mm	0.80	16%	99%	8%
Right Knee	92 mm	21 mm	0.79	17%	98%	15%
Left Ankle	138 mm	35 mm	0.86	10%	93%	22%
Right Ankle	159 mm	55 mm	0.86	8%	85%	20%
Left Shoulder	45 mm	17 mm	0.80	17%	80%	21%
Right Shoulder	64 mm	26 mm	0.75	20%	87%	47%
Left Elbow	52 mm	24 mm	0.72	23%	97%	6%
Right Elbow	73 mm	32 mm	0.79	16%	91%	47%
Left Wrist	73 mm	31 mm	0.71	22%	98%	32%
Right Wrist	67 mm	34 mm	0.70	23%	85%	50%

TABLE III

PARAMETERS OF DISTRIBUTION \mathbf{D} OF L_2 DISTANCES BETWEEN JOINTS OF PHASESPACE AND KINECT SKELETONS, 60° VIEW.

As expected, the percentage of failed joint estimates increases, more for the right limbs, as the subject turns away from the Kinect camera to the right. The Kinect works best when faced frontally, but the decrease in performance with the view angle is only gradual.

For each joint the Kinect provides an indication whether the joint was directly observed or not, in which case it may have been inferred from past positions and assumptions about body geometry. Let us denote I_j^P the set of Kinect measurements $\mathbf{x}_{t,j,\text{Kinect}}$ that we classify as inliers based on the PhaseSpace's skeleton, *i.e.* joints at distances $D_{t,j}$ for which $\rho_j \cdot \mathcal{N}(\mu_j, \sigma_j^2) > (1 - \rho_j) \cdot \mathcal{U}(0, 600\text{mm})$. Similarly, let I_j^K be the set of Kinect's joints indicated as directly observed by the Kinect. For algorithms subsequently processing the human pose it would be ideal if the Kinect reliably indicated which measurements are imprecise, *i.e.* if the two sets were identical. Unfortunately they are not. The last two columns in the tables show the percentage of inliers that were marked as directly observed joints by the Kinect, and the percentage of outliers that were marked as unobserved (the closer to 100% the better). In the tables we have denoted the sets of outliers, complementary to the inlier sets, as O^P and O^K .

To reiterate, the results are presented under the assumptions that the skeletons by Recap are correct and accurate,

Joint	mean dist μ_j	std. dev σ_j	mix ρ_j	outliers $\frac{ O^P }{T}$	indicated I $\frac{ I^P \cap I^K }{ I^P }$	indicated O $\frac{ O^P \cap O^K }{ O^P }$
Left Hip	212 mm	36 mm	1.00	0%	57%	—
Right Hip	228 mm	31 mm	0.96	3%	67%	100%
Left Knee	77 mm	21 mm	0.82	15%	97%	19%
Right Knee	92 mm	21 mm	0.69	26%	95%	40%
Left Ankle	132 mm	35 mm	0.88	8%	93%	27%
Right Ankle	153 mm	56 mm	0.82	10%	78%	36%
Left Shoulder	47 mm	19 mm	0.78	19%	62%	48%
Right Shoulder	60 mm	22 mm	0.51	42%	86%	76%
Left Elbow	53 mm	22 mm	0.70	26%	97%	20%
Right Elbow	74 mm	32 mm	0.64	28%	84%	73%
Left Wrist	75 mm	32 mm	0.72	21%	98%	37%
Right Wrist	63 mm	30 mm	0.54	38%	84%	49%

TABLE IV

PARAMETERS OF DISTRIBUTION D OF L_2 DISTANCES BETWEEN JOINTS OF PHASESPACE AND KINECT SKELETONS, 90° VIEW.

and that the differences between Recap and Kinect are reasonably modelled by the chosen Gaussian-Uniform mixture.

D. Stability of Kinect Body Geometry

We have investigated how stable is the frame-to-frame estimation of body geometry (bone lengths) by Kinect. Figure 5 shows the measured lengths of the upper and lower leg during the ‘Deep Breathing’ exercise performed in a sitting position. The subject was captured in three different orientations, relative to the camera, approximately 45° apart. The red line indicates bone lengths manually measured on the subject. Constant lines (blue for left leg, green for right) are the lengths from Recap, reflecting the accuracy of the pre-calibration.

During the three repetitions of the exercise the legs were not moving significantly, yet the variation in Kinect’s bone lengths was about 2 cm for the frontal view and about 5 cm for the 45° view. In the final 90° view the Kinect often failed to recover a meaningful pose since half of the body was occluded.

To demonstrate changes in geometry between significantly different postures, the T-pose, and a transition from standing to sitting position, is included at the beginning of the test sequence (leftmost part of Figure 5). The leg length variability is much higher there, about 10 cm, indicating that without markers it is difficult to determine the exact knee and hip location when the legs are straightened. Similar behavior was observed for elbows when the arms are straight.

V. CONCLUSIONS

In this paper we compared the Kinect pose estimation with more established techniques relying on motion capture data. We believe system such as Kinect has significant potential as a low-cost alternative for real-time motion capture and body tracking in health applications. In the context of physical exercise of elderly population we observed that the Kinect skeleton tracking struggles with occluding body parts or objects in the scene (e.g. a chair). One of the main drawbacks of the Kinect skeleton for the purpose of healthcare is in a

very non-anthropometric kinematic model with variable limb lengths.

In a more controlled body posture (e.g. standing and exercising arms), the accuracy of the joint estimation is comparable to motion capture. However, in general postures, the variability of the current implementation of the pose estimation is about 10 cm. The measurements could be used to assess general trends in the movement, but for quantitative estimation an improved skeletonization with an anthropometric model is needed. Such algorithms should also address occlusions and self-occlusions, unconventional body postures or use of wheelchairs or walkers.

ACKNOWLEDGMENT

The authors would like to thank Po Yan for assistance with the data acquisition. This research was supported in part by National Science Foundation (NSF) grant 1111965 and by Grant Number HHS 90TR0003/01. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the HHS.

REFERENCES

- [1] Microsoft, “Kinect,” 2010. url: <http://www.xbox.com/en-us/kinect> (accessed March 7, 2012).
- [2] J. Shotton, A. W. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake, “Real-time human pose recognition in parts from single depth images,” in *CVPR*, pp. 1297–1304, 2011.
- [3] K. Khoshelham and S. O. Elberink, “Accuracy and resolution of kinect depth data for indoor mapping applications,” *Sensors*, vol. 12, no. 2, pp. 1437–1454, 2012.
- [4] J. Smisek, M. Jancosek, and T. Pajdla, “3d with kinect,” in *ICCV Workshops*, pp. 1154–1160, 2011.
- [5] E. E. Stone and M. Skubic, “Evaluation of an inexpensive depth camera for passive in-home fall risk assessment,” in *PervasiveHealth*, pp. 71–77, 2011.
- [6] J.-D. Huang, “Kinerehab: A kinect-based system for physical rehabilitation a pilot study for young adults with motor disabilities,” in *In Proceedings of ASSETS11*, (Dundee, Scotland, UK.), pp. 319–320, October 2426 2011.
- [7] B. Lange, C.-y. Chang, E. Suma, B. Newman, A. S. Rizzo, and M. Bolas, “Development and evaluation of low cost game-based balance rehabilitation tool using the microsoft kinect sensor,” *Conference Proceedings of the International Conference of IEEE Engineering in Medicine and Biology Society*, vol. 2011, pp. 1831–1834.
- [8] C. Schonauer, T. Pintaric, H. Kaufmann, S. Jansen Kosterink, and M. Vollenbroek-Hutten, “Chronic pain rehabilitation with a serious game using multimodal input,” in *Virtual Rehabilitation (ICVR), 2011 International Conference on*, pp. 1–8, june 2011.
- [9] PhaseSpace, “Impulse motion capture.” url: <http://www.phasespace.com> (accessed March 15, 2012).
- [10] Autodesk, “Motionbuilder.” url: <http://usa.autodesk.com> (accessed March 15, 2012).