

Visual and Human-Interpretable Feedback for Assisting Physical Activity

Michel Antunes^(✉), Renato Baptista, Girum Demisse, Djamila Aouada,
and Björn Ottersten

Interdisciplinary Centre for Security, Reliability and Trust (SnT),
University of Luxembourg, Luxembourg, Luxembourg
{michel.antunes,renato.baptista,girum.demisse,
djamila.aouada,bjorn.ottersten}@uni.lu

Abstract. Physical activity is essential for stroke survivors for recovering some autonomy in daily life activities. Post-stroke patients are initially subject to physical therapy under the supervision of a health professional, but due to economical aspects, home based rehabilitation is eventually suggested. In order to support the physical activity of stroke patients at home, this paper presents a system for guiding the user in how to properly perform certain actions and movements. This is achieved by presenting feedback in form of visual information and human-interpretable messages. The core of the proposed approach is the analysis of the motion required for aligning body-parts with respect to a template skeleton pose, and how this information can be presented to the user in form of simple recommendations. Experimental results in three datasets show the potential of the proposed framework.

Keywords: Rehabilitation · Stroke · Feedback · Human interpretable

1 Introduction

Physical activity is vital for the general population for maintaining a healthy lifestyle. It is crucial for elderly people in the prevention of diseases, maintenance of independence and improvement of quality of life [17]. For stroke survivors it is critical and essential for recovering some autonomy in daily life activities [8]. Despite the benefits of physical activity, many stroke survivors do not exercise regularly due to many reasons, such as lack of motivation, confidence, and skill levels [13]. Traditionally, the post-stroke patients are initially subject to physical therapy under the supervision of a health professional aimed at restoring and maintaining activities of daily living in rehabilitation centres [20]. The physiotherapist explains the movement to be performed to the patient, and continuously advises her/him how to improve the motion as well as interrupts the exercise in case of health related risk issues. Unfortunately, and due to the high economical burden [1], the *at site* rehabilitation is usually of a short period of time and prescribed treatments and activities for home based rehabilitation are

usually suggested [9]. Unfortunately, stroke patients, and more frequently older adults, do not appropriately adhere to the recommended treatments, because, among other factors, they do not always understand or remember well enough what and how they are supposed to do the physical treatment.

In order to support the rehabilitation of stroke patients at home, human tracking and gesture therapy systems are being investigated for monitoring and assistance purposes [3, 6, 12, 13, 16, 25]. These home rehabilitation systems are advantageous not only because they are less costly for the patients and for the health care systems, but also because having it at home and regularly available, the users tend to do more exercise. A well accepted sensing technology for these purposes are RGB-D sensors (e.g. Kinect) that are affordable and versatile, allowing to capture in real-time colour and depth information [3, 13].

Existing systems and research either (1) combine exercises with video games as a means to educate and train people, while keeping a high level of motivation [2, 7]; or (2) try to emulate a physical therapy session [13, 16]. These works usually involve the detection, recognition and analysis of specific motions and actions performed. Very recent works tackle the problem of assessing how well the people perform certain actions [13, 14, 18, 23], which can be used in rehabilitation e.g. to evaluate mobility and measure the risk of relapse. The authors of [14] propose a framework for assessing the quality of actions in videos. Spatio-temporal pose features are extracted and a regression model is estimated that predicts scores of actions from annotated data. Tao et al. [18] also describe an approach for quality assessment of the human motion. The idea is to learn a manifold from normal motion, and then evaluate the deviation from it using specific measures. Wang et al. [23] tackle the problem of automated quantitative evaluation of musculo-skeletal disorders using a 3D sensor. They introduce the Representative Skeletal Action Unit framework from which clinical measurements can be extracted. Very recently, Ofli et al. [13] presented an interactive coaching system using the Kinect. The coaching system guides users through a set of exercises, and the quality of execution of these exercises is assessed based on manually defined pose measurements, such as keeping hands close to each other or maintaining the torso in an upright position.

In this work, we want to go one step further and not only evaluate, but also provide feedback in how people can improve the action being performed. There are two main works that tackle this problem. In the computer vision community, the work of Pirsiavash et al. [14] is the most relevant. After assessing the quality of actions using supervised regression, feedback proposals are obtained by differentiating the scoring with respect to the joint locations, and then selecting the joint and the direction it should move to achieve the largest improvement in the score. In the medical community, Ofli et al. [13] provide assistive feedback during the performance of exercises. For each particular movement, they define constraints such as keeping hands close to each other or maintaining the torso in a upright position. These constraints are constantly measured during the exercise for assessing if the movement is performed correctly and in case pre-defined values for metrics on these constraints are violated, then corrective feedback is provided.

While in [14] the corrective feedback is analysed per joint, which involves a complex set of instructions for suggesting a particular body-part motion (e.g. arm moving up), in [13] the motion constraints are action specific and manually defined.

1.1 Contributions

As discussed previously, the objective of this paper is not only to assess the quality of an action, but also to provide feedback in how to improve the movement being performed. In contrast to previous works, there are three main contributions:

1. We do not compute feedback for single joints, but for body-parts, defined as configurations of skeleton joints that may or may not move rigidly;
2. Feedback proposals are automatically computed by comparing the movement being performed with a template action, without specifying pose constraints of joint configurations;
3. Feedback instructions are not only presented visually, but also human interpretable feedback is proposed from discretized spatial transformations that can be suggest to the user using, for example, audio messages.

1.2 Organization

The article is organized as follows: Sect. 2 introduces the problem that we want to solve, and briefly discusses the pre-processing that is required for spatially and temporarily aligning skeleton sequences. Section 3 presents the body-part representation, the computation of feedback proposals and how they can be translated to human-interpretable messages. Finally, the experimental results are presented in Sect. 4.

2 Problem Definition and Skeleton Processing

This section discusses the problem that we aim to solve, and describes the processing that is performed for spatially and temporarily aligning two skeleton sequences.

2.1 Problem Definition

Let $S = [\mathbf{j}_1, \dots, \mathbf{j}_n, \dots, \mathbf{j}_N]$ denote a skeleton instance with N joints, where each joint is given by its 3D coordinates $\mathbf{j} = [j_x, j_y, j_z]^T$. Let us define an action or movement as being a skeleton sequence $M = [S_1, \dots, S_f, \dots, S_F]$, where F is the number of frames of the sequence. The objective of this paper is to solve the following problem: given a template skeleton sequence \hat{M} and a subject performing a movement M , we want to provide, at each time instant, feedback proposals such that the movement can be iteratively improved to better match \hat{M} . As a first step, pre-processing on the input skeleton data is required. Existent approaches were previously introduced in the literature (e.g. [21]), and are adapted for our specific problem.

2.2 Data Normalization

The first requirement for comparing two skeletal sequences is that they need to be spatially registered. This is achieved by transforming the joints of each skeleton S such that the world coordinate system is placed at the hip center, and the projection of the vector from the left hip to the right hip onto the x-y plan is parallel to the x-axis. Then, for achieving invariance to absolute locations, the skeletons in M are normalized such that the body part lengths match the corresponding part lengths of the skeletons in \hat{M} . This is performed without modifying the joint angles.

2.3 Temporal Alignment

Different subjects, or the same subject at different times, perform a particular action or movement at different rates. In order to handle rate variations and mitigate the temporal misalignment of time series, Dynamic Time Warping (DTW) is usually employed [15]. In our particular case, we want to align a given sequence M with a template sequence \hat{M} . There are two possibilities, we either align M with respect to \hat{M} , or vice-versa, \hat{M} with respect to M . We assume the subject is trying to replicate the same action as \hat{M} , and given M , we want to provide feedback proposals. Since we want to compute a feedback proposal for each temporal instant of M , it is reasonable to compute the temporal correspondences of \hat{M} with respect to M . Figure 1 shows a temporal alignment example.

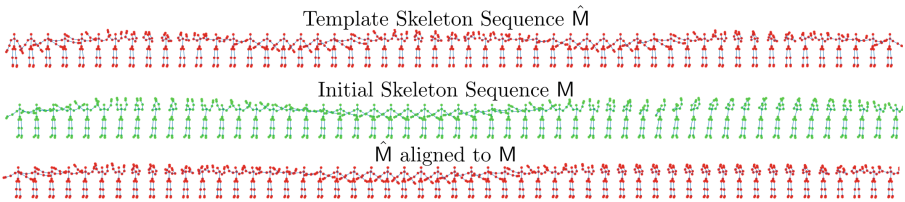


Fig. 1. Temporal alignment of skeleton sequences using DTW. The first row shows the template action \hat{M} (red), the second row shows the skeleton sequence M , and the third row shows \hat{M} aligned with respect to M using DTW. (Color figure online)

3 Human-Interpretable Feedback Proposals

After the spatial and temporal alignment processing described in the previous section, the skeleton instance \hat{S}_f in \hat{M} will be in correspondence with S_f in M . This section explains how to compute the body motion required to align corresponding body-parts of aligned skeletons \hat{S} and S , and proposes a method for extracting human-interpretable feedback from these transformations.

3.1 Body-Part Based Representation

In line with recent research [4, 11, 19, 22], we analyse the human motion using a body-part based representation. A skeleton S can be represented by a set of body-parts $\mathcal{B} = \{b^1, \dots, b^k, \dots, b^N\}$. Each body part b^k is composed by n^k joints $b^k = \{b_1^k, \dots, b_{n^k}^k\}$ and has a local reference system defined by the joint b_r^k . Figure 2 shows the different body-parts defined for the dataset Weight&Balance.

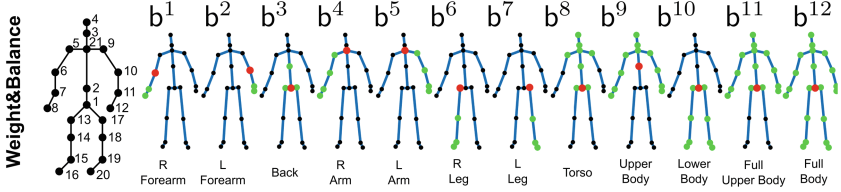


Fig. 2. Proposed body-part representation. The skeleton of the dataset Weight&Balance is composed by 21 joints (left). 12 body-parts were defined. For each body-part, the composing joints were highlighted in green. The red joint corresponds to the local origin b_r^k of a each body-part (R=right, L=left). (Color figure online)

Given the aligned skeletons \hat{S} and S , the objective is to compute the motion that each body-part of S needs to undergo to better match the template skeleton \hat{S} . This analysis is performed for each body-part using the corresponding local coordinate system. As a metric for measuring how similar is the pose of corresponding body-parts, we use the Euclidean distance as the scoring function. Following this, the error between b^k and \hat{b}^k is given by:

$$m^k = \sum_{j=1}^{n^k} \|b_j^k - \hat{b}_j^k\|^2. \quad (1)$$

Remark that $\|b_r^k - \hat{b}_r^k\| = 0$, because the previous computation is performed using the local coordinate systems that are assumed to be in correspondence.

3.2 Feedback Proposals

For providing feedback to the performer of skeleton S on how the movement can be improved to better match \hat{S} , we compute the transformation that each body-part b^k needs to undergo for decreasing the scoring function m^k . We anchor the reference joints b_r^k and \hat{b}_r^k (refer to Fig. 2) of the corresponding body-parts. The aim is then to compute the rotation $R^k \in SO(3)$ that minimizes the following error:

$$e^k(R^k) = \sum_{j=1}^{n^k} \|R^k b_j^k - \hat{b}_j^k\|^2, \quad (2)$$

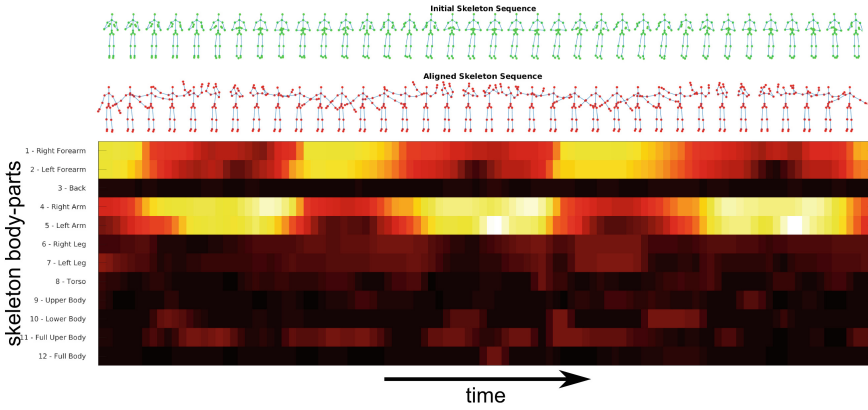


Fig. 3. Intensity of feedback required for each body-part. (Top) Sequence M performing clapping, (middle) target sequence \hat{M} corresponding to the action waving using two hands after spatial and temporal alignment, and (bottom) the cost c_i^k (refer to Method 1) calculated for each temporal instant independently (the vertical axis corresponds to different body-parts, while the horizontal axis is the temporal dimension).

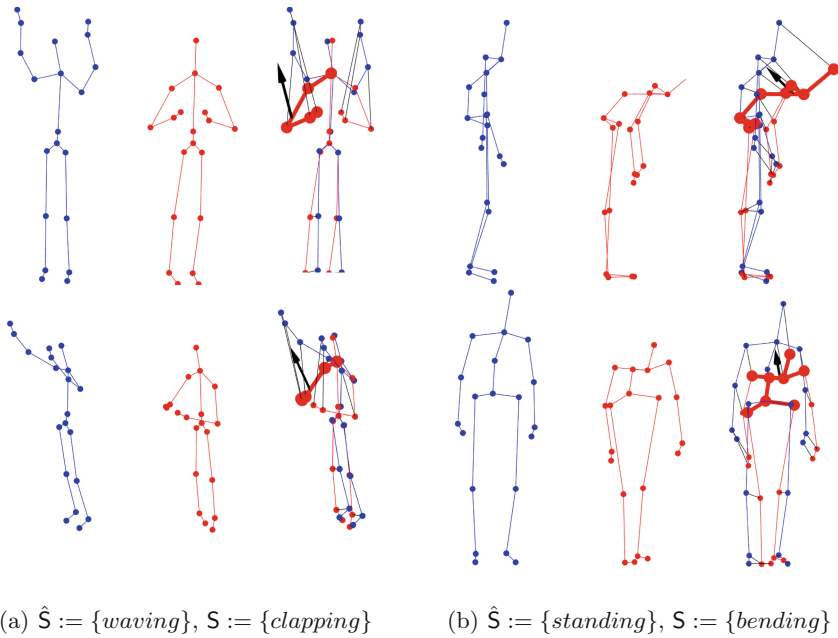


Fig. 4. Two examples for feedback proposals. The target pose \hat{S} is shown in blue and the action being performed is shown in red. For each example, the third column shows superimposed the two skeletons, the matching joints (black lines) and the feedback vectors \mathbf{f}_k (black arrows). Only the feedback proposal for R_1 is shown. The different rows present different viewing angles. (Color figure online)

which can be computed in closed form. It is important to refer that since the human motion is articulated, depending on the movement being performed, a given body-part \mathbf{b}^k may or may not move rigidly. This is not a critical issue because body-parts that do not moving rigidly have high joint matching error and will be considered not relevant by the method described next. Note that different body-parts \mathbf{b}^k can contain subsets of the same joints, which implies that the transformation \mathbf{R}^k will also have impact on the location of the other body-parts $\mathbf{b}^{l \neq k}$. Taking this into account, we want to compute a sequence of transformations $\mathcal{R} = \{\mathbf{R}_1, \dots, \mathbf{R}_i, \dots, \mathbf{R}_N\}$, one rotation $\mathbf{R}_i = \mathbf{R}^k$ for each body-part \mathbf{b}^k , such that the first rotation \mathbf{R}_1 has the highest decrease in the joint location error until \mathbf{R}_N , which has the lowest impact in the human pose matching. This sorting is performed maximizing the following cost

$$c_i^k = m^k - e^k(\mathbf{R}^k), \quad (3)$$

where in iteration i , the body-parts \mathbf{b}^k selected in the previous $i - 1$ iterations are not taken into account. The pseudo-code of the overall scheme is shown in Method 1. Figure 3 show an example of the intensity pattern c_i^k for actions clapping and waving across time.

Input: S, \hat{S}, \mathbf{B}

Output: Sequence of rotations \mathcal{R} , list of body-part indexes \mathcal{K}

$\mathcal{L} := \mathcal{B}, \mathcal{K} = \{\}, \mathcal{R} = \{\}, i = 1;$

while $\mathcal{L} \neq \{\}$ **do**

foreach $\mathbf{b}^k \in \mathcal{L}$ **do**

 compute \mathbf{R}^k that minimizes $e^k(\mathbf{R}^k)$ (refer to Equation 2);

$c_i^k = m^k - e^k(\mathbf{R}^k)$ (refer to Equation 3);

end

$l := \operatorname{argmax}_k (c_i^k);$

$\mathbf{R}_i = \mathbf{R}^l;$

$\mathcal{K} := \mathcal{K} \cup l, \mathcal{R} := \mathcal{R} \cup \mathbf{R}_i;$

$\mathcal{L} := \mathcal{L} \setminus \mathbf{b}^l;$

$i=i+1;$

end

Method 1. Computation of the sequence of body-part transformations that minimizes the skeleton matching error.

The rotations $\mathbf{R}_i = \mathbf{R}^k$ correspond to the motion required for the best alignment of \mathbf{b}^k and $\hat{\mathbf{b}}^k$. However, it is difficult to present this rigid-body transformation as feedback proposals on, for example, a screen. For overcoming this, we compute feedback vectors for suggesting improvements on the motion. For each body-part, we pre-calculate the spatial centroid \mathbf{c}^k (note that in case of single limbs, this point is located on the body-part itself). Then, the feedback vector anchored to \mathbf{c}^k is defined as

$$\mathbf{f}^k = \mathbf{R}^k \mathbf{c}^k - \mathbf{c}^k. \quad (4)$$

Figure 4 shows feedback vectors for two different pairs of actions being performed.

3.3 Feedback Messages

At this point, we have discussed how to compute the optimal rotation R^k for each body-part b^k , and how this transformation can be presented to a user in form of a feedback vector f^k anchored to the body-part centroid c^k . Nevertheless, not all the persons have the same spatial awareness to realize how to perform the motion suggested by the feedback vector f^k (refer to Fig. 4). This difficulty is even more evident in cognitive impaired individuals [5]. In order to support the patient in improving their movements, we introduce in this section a system for presenting simple human-interpretable feedback messages that can be shown or/and spoken to the patient by the computer system.

Let us analyse the case of the body-part b^k that needs to undergo the largest motion $R_1 = R^k$. Initially, to each b^k was assigned a body-part name BN , e.g.

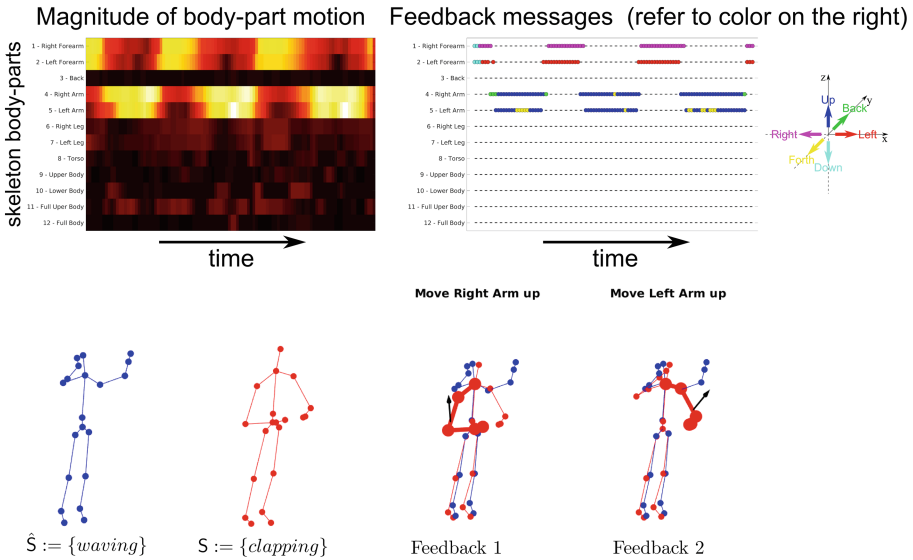


Fig. 5. Feedback message proposals. The target action is waving using two hands and the movement being performed corresponds to clapping. (Top, left) The intensity c_i^k for each body-part b^k ; (top, right) the feedback message proposals for the body-parts corresponding to R^1 and R^2 . Each point corresponds to a particular message at a given time instant using the body-part name identified on the left and the color coding on the right, e.g. a blue point on the fourth dotted line corresponds to the message *Move Right Arm Up*. (Bottom) a particular instance of the template skeleton \hat{S} (blue), an instance of the skeleton S (red), the feedback vectors for the body-parts corresponding to R^1 and R^2 (black arrows), and the corresponding feedback messages at the top. (Color figure online)

\mathbf{b}^1 is the *Right Forearm* and \mathbf{b}^8 is the *Torso* (refer to Fig. 2). These labels are used directly for informing the user which body-parts should be moved. Then, the feedback vector $\mathbf{f}^k = [f_x^k, f_y^k, f_z^k]^\top$ is discretized by selecting the dimension d with highest magnitude $|f_d^k|$. The messages regarding the direction of the motion BD are then defined as:

- if $d = x$
 - if $f_x^k < 0$, then $BD = Right$
 - if $f_x^k > 0$, then $BD = Left$
- if $d = y$
 - if $f_y^k < 0$, then $BD = Forth$
 - if $f_y^k > 0$, then $BD = Back$
- if $d = z$
 - if $f_z^k < 0$, then $BD = Down$
 - if $f_z^k > 0$, then $BD = Up$

The feedback proposal messages are represented as the concatenation of strings:

$$\text{Feedback message} := \text{“Move”} + \text{BN} + \text{BD}. \quad (5)$$

Refer to Fig. 5 for an example of feedback messages, where a color coding is used for identifying the directions BD .

4 Experiments

In this section, we experimentally evaluate the proposed system using three different sets of data. The first is called **ModifyAction**, and we use pairs of

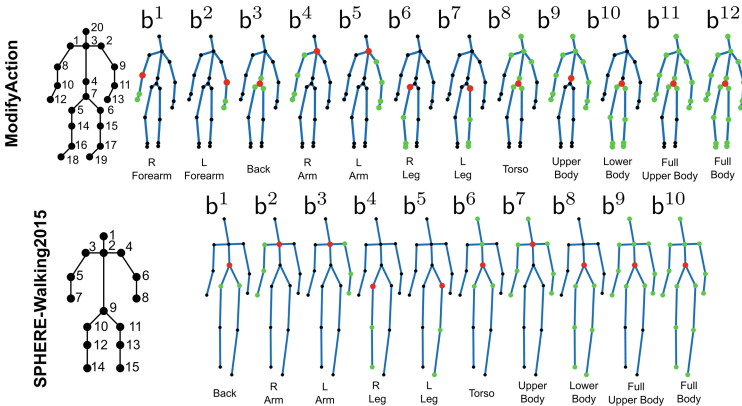


Fig. 6. Proposed body-part representations. Each row shows the skeleton (black) and body-part configurations used for two different datasets. For each body-part, the composing joints were highlighted in green. The red joint corresponds to the local origin \mathbf{b}_r of a each body-part. (Color figure online)

actions instances from the datasets *UTKinect* [24] and *MSR-Action3D* [10]. The objective is: given a person performing a particular action M , provide feedback proposals such that the person is able to perform a different action \hat{M} . The skeleton and body-parts used for this dataset are shown in Fig. 6.

The second dataset is **SPHERE-Walking2015** that was introduced in [18]. The skeleton and body-parts used for this dataset are shown in Fig. 6. It contains people walking on a flat surface, and it includes instances of normal walking and subjects simulating the walking of stroke survivors under the guidance of a physiotherapist. The objective in this regard is to analyse the difference in the walking pattern of normal subjects when compared to people with stroke.

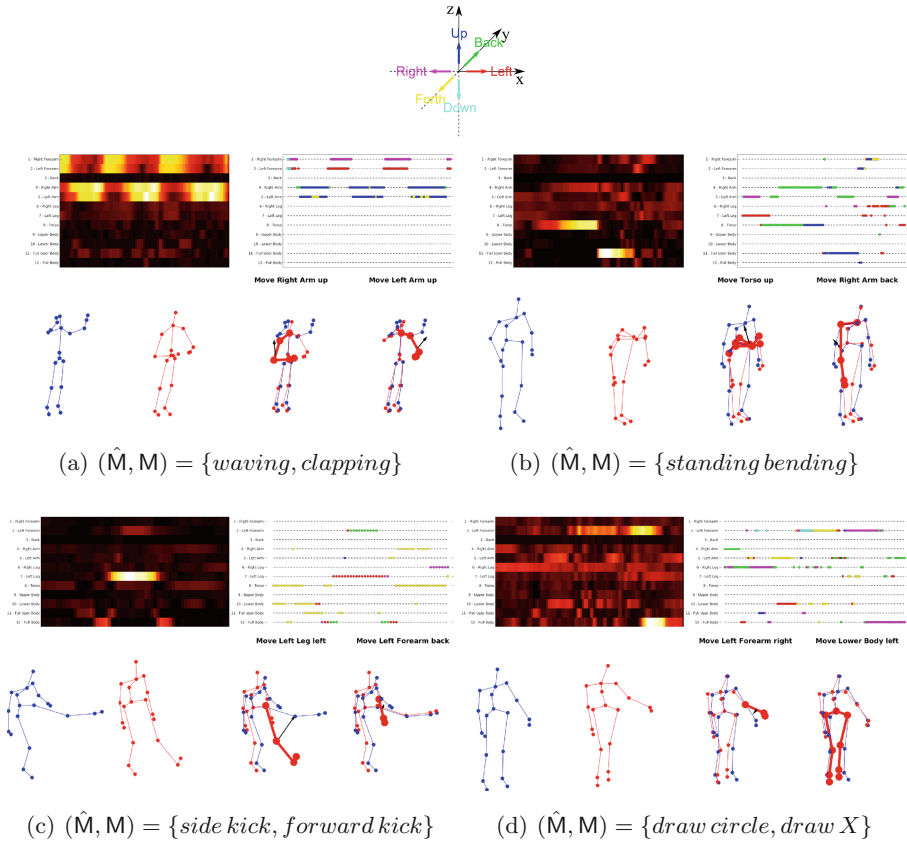


Fig. 7. Four experimental results for the **ModifyAction** dataset are shown. For each example, we show the magnitude of the motion for each body-part (top, left); the feedback messages corresponding to R^1 and R^2 (top, right), refer to the color coding at the top; and the feedback vectors and messages for a particular temporal instant (bottom). (Color figure online)

Finally, the third dataset is new and is called **Weight&Balance**. This data was captured using the Kinect version 2. Refer to Fig. 2 for a detailed description of the body-parts used. The idea is to simulate a person who suffered a stroke (refer to Fig. 10): the *bad arm* issue due to the paralysis of an upper limb is simulated by lifting a kettle-bell using one of the arms, and the *balance* problem is replicated using a balance ball.

Figure 7 shows experimental results of the proposed coaching system for the **ModifyAction** dataset.

4.1 Experiments in SPHERE-Walking2015

In the experiment of Fig. 8, we compared the walking pattern of all the subjects with respect to the walking of healthy people (template action). It shows the intensity profile defined as the sum c_i^k across time for each subject. It is evident that stroke patients have a balance problem, because the body-part corresponding to the torso has high skeleton matching error, while also the stronger paralysis

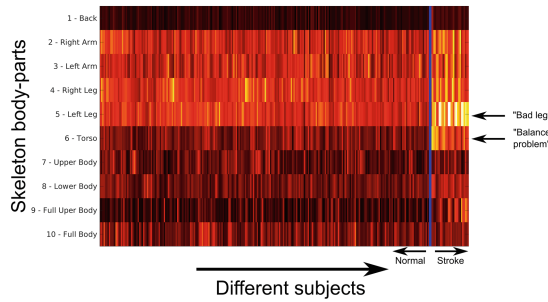


Fig. 8. Motion intensity of different body-parts for different subjects. The subjects on the left of the blue line are healthy people, while the subjects on the right are the (simulated) stroke survivors. (Color figure online)

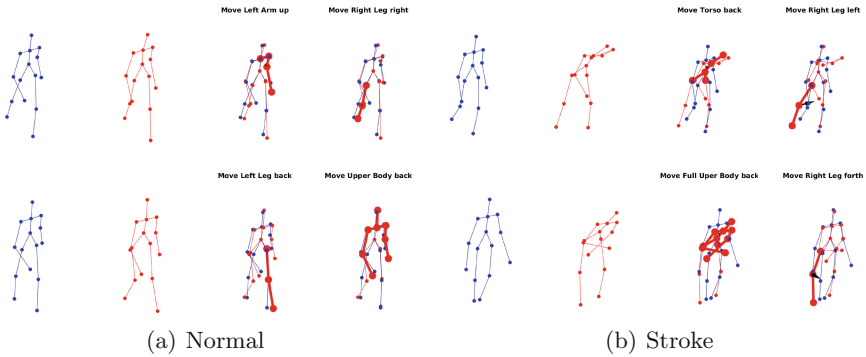


Fig. 9. Feedback proposals. The two subjects on the left are normal people, while the two subjects on the right are stroke survivors

of one of the lower limbs can be identified. Figure 9 shows feedback proposals for normal people and stroke patients.

4.2 Experiments in Weight&Balance

The objective in this section is to simulate a simple physiotherapy session at home, and test if the feedback proposals are able to guide the user. We assume

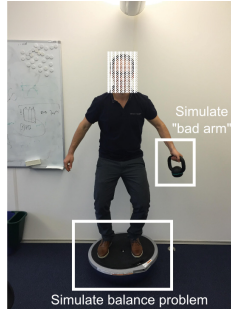


Fig. 10. Weight&Balance dataset. We simulate the motion behaviour of a person who suffered a stroke: the *bad arm* issue due to the paralysis of an upper limb is simulated by lifting a kettle-bell using one of the arms, and the *balance* problem is replicated using a balance ball.

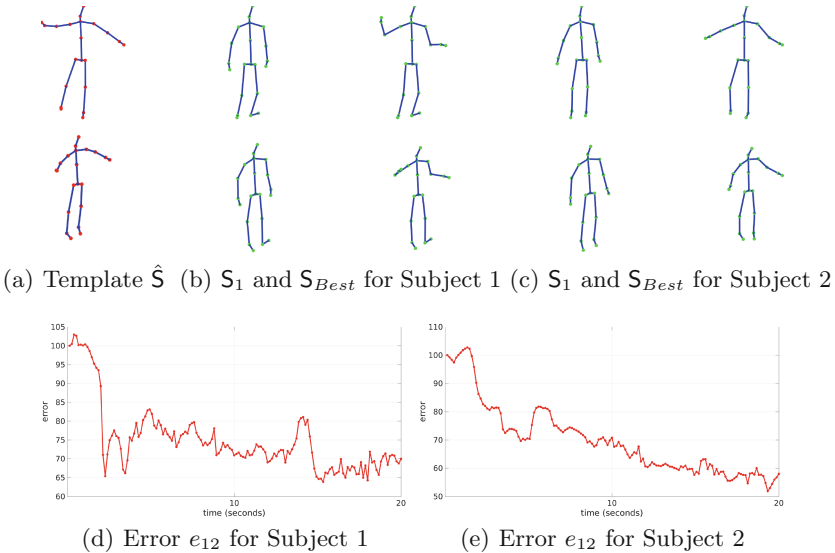


Fig. 11. Example 1 of Weight&Balance. (Top) two views of the template pose \hat{S} , and first pose S_1 and best pose S_{Best} for two subjects are shown. The best pose S_{Best} is the one that minimizes the error m^{12} . (Bottom) the relative error (difference between initial and current error divided by the initial error) in % for b^{12} is shown.

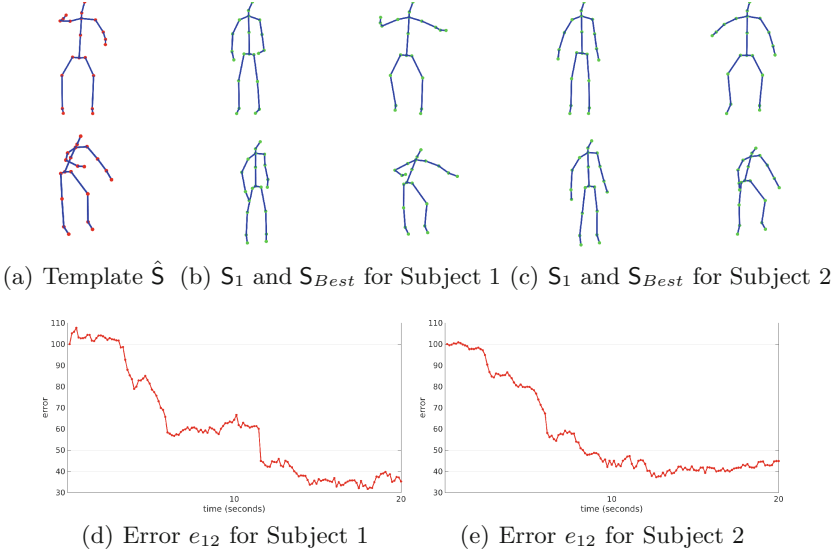


Fig. 12. Example 2 of **Weight&Balance**. (Top) two views of the template pose \hat{S} , and first pose S_1 and best pose S_{Best} for two subjects are shown. The best pose S_{Best} is the one that minimizes the error m^{12} . (Bottom) the relative error (difference between initial and current error divided by the initial error) in % for \mathbf{b}^{12} is shown.

that a person needs to perform a template human pose \hat{S} . The subject puts himself above the balance ball and lifts the kettle-bell. Giving only the guidance of the feedback vectors, body-part motion intensity and feedback messages, the objective is to converge to the template pose without actually seeing it. The exercise lasts for 20s and feedback proposals are shown at each time instant. The experimental results are shown in Figs. 11 and 12.

5 Conclusions

In this paper, we have introduced a system for guiding a user in correctly performing an action or movement by presenting feedback proposals in form of visual information and human-interpretable feedback. Preliminary experiments show that the provided feedbacks are effective in guiding users towards given human poses. As future work, we intend to incorporate physiotherapy practices in the computation of feedback proposals, and validate the proposed framework using real data.

Acknowledgements. This work has been partially funded by the European Union’s Horizon 2020 research and innovation project STARR under grant agreement No.689947. This work was also supported by the National Research Fund (FNR), Luxembourg, under the CORE project C15/IS/10415355/3D-ACT/Björn Ottersten. The authors would like to thank Adeline Paiement and the SPHERE project for sharing the SPHERE-Walking2015 dataset.

References

1. Andlin-Sobocki, P., Jönsson, B., Wittchen, H.U., Olesen, J.: Cost of disorders of the brain in europe. *Eur. J. Neurol.* (2005)
2. Burke, J.W., McNeill, M., Charles, D., Morrow, P.J., Crosbie, J., McDonough, S.: Serious games for upper limb rehabilitation following stroke. In: *Conference on Games and Virtual Worlds for Serious Applications, VS-GAMES 2009*. IEEE (2009)
3. Chaaaroui, A.A., Climent-Pérez, P., Flórez-Revuelta, F.: A review on vision techniques applied to human behaviour analysis for ambient-assisted living. *Expert Syst. Appl.* **39**(12), 10873–10888 (2012)
4. Chaudhry, R., Offi, F., Kurillo, G., Bajcsy, R., Vidal, R.: Bio-inspired dynamic 3d discriminative skeletal features for human action recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2013)
5. Cicerone, K.D., Langenbahn, D.M., Braden, C., Malec, J.F., Kalmar, K., Fraas, M., Felicetti, T., Laatsch, L., Harley, J.P., Bergquist, T., et al.: Evidence-based cognitive rehabilitation: updated review of the literature from 2003 through 2008. *Archives of physical medicine and rehabilitation* (2011)
6. Hondori, H.M., Khademi, M., Dodakian, L., Cramer, S.C., Lopes, C.V.: A spatial augmented reality rehab system for post-stroke hand rehabilitation. In: *MMVR* (2013)
7. Kato, P.M.: Video games in health care: Closing the gap. *Rev. Gen. Psychol.* (2010)
8. Kwakkel, G., Kollen, B.J., Krebs, H.I.: Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review. *Neurorehabilitation Neural Repair* (2007)
9. Langhorne, P., Taylor, G., Murray, G., Dennis, M., Anderson, C., Bautz-Holter, E., Dey, P., Indredavik, B., Mayo, N., Power, M., et al.: Early supported discharge services for stroke patients: a meta-analysis of individual patients' data. *The Lancet* (2005)
10. Li, W., Zhang, Z., Liu, Z.: Action recognition based on a bag of 3d points. In: *Workshop on Human Activity Understanding from 3D Data* (2010)
11. Lillo, I., Soto, A., Niebles, J.: Discriminative hierarchical modeling of spatio-temporally composable human activities. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014)
12. Mousavi Hondori, H., Khademi, M.: A review on technical and clinical impact of microsoft kinect on physical therapy and rehabilitation. *J. Med. Eng.* **2014**, 16 (2014)
13. Offi, F., Kurillo, G., Obdrzálek, S., Bajcsy, R., Jimison, H.B., Pavel, M.: Design and evaluation of an interactive exercise coaching system for older adults: lessons learned. *IEEE J. Biomed. Health Inf.* **20**(1), 201–212 (2016)
14. Pirsiavash, H., Vondrick, C., Torralba, A.: Assessing the quality of actions. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) *ECCV 2014, Part VI*. LNCS, vol. 8694, pp. 556–571. Springer, Heidelberg (2014)
15. Rabiner, L., Juang, B.H.: *Fundamentals of speech recognition*. Prentice hall (1993)
16. Sucar, L.E., Luis, R., Leder, R., Hernandez, J., Sanchez, I.: Gesture therapy: a vision-based system for upper extremity stroke rehabilitation. In: *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (2010)

17. Sun, F., Norman, I.J., While, A.E.: Physical activity in older people: a systematic review. *BMC Public Health* (2013)
18. Tao, L., Paiement, A., Aldamen, D., Mirmehdi, M., Hannuna, S., Camplani, M., Burghardt, T., Craddock, I.: A comparative study of pose representation and dynamics modelling for online motion quality assessment. *Comput. Vis. Image Underst.* **11** (2016)
19. Tao, L., Vidal, R.: Moving poselets: A discriminative and interpretable skeletal motion representation for action recognition. In: *ChaLearn Looking at People Workshop*. 2015 (2015)
20. Veerbeek, J.M., van Wegen, E., van Peppen, R., van der Wees, P.J., Hendriks, E., Rietberg, M., Kwakkel, G.: What is the evidence for physical therapy poststroke? a systematic review and meta-analysis. *PloS one* (2014)
21. Vemulapalli, R., Arrate, F., Chellappa, R.: Human action recognition by representing 3d skeletons as points in a lie group. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014)
22. Wang, C., Wang, Y., Yuille, A.: An approach to pose-based action recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2013)
23. Wang, R., Medioni, G., Winstein, C., Blanco, C.: Home monitoring musculo-skeletal disorders with a single 3d sensor. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2013)
24. Xia, L., Chen, C.C., Aggarwal, J.K.: View invariant human action recognition using histograms of 3d joints. In: *Workshop on Human Activity Understanding from 3D Data* (2012)
25. Zhou, H., Hu, H.: Human motion tracking for rehabilitationa survey. *Biomed. Signal Process. Control* **3**(1), 1–18 (2008)