Action Recognition using Dynamics Features

Al Mansur, Yasushi Makihara and Yasushi Yagi

Abstract— In this paper, we propose a method of action recognition using dynamics features based on physics model. The dynamics features are composed of torques from knee and hip joints of both legs and implicitly include the gravity, ground reaction forces, and the pose of the remaining body parts. These features are more discriminative than the kinematics features, and they result in a low dimensional representation of a human action which preserves much information of the original high dimensional pose. This low dimensional feature allows us to achieve a good classification performance even with a relatively small training data in a simple classification framework such as HMM. The effectiveness of the proposed method is demonstrated through experiments on the CMU motion capture dataset with various actions.

I. INTRODUCTION

The recognition and interpretation of human actions and activities has gained considerable interest in the robotics and AI communities [1], [2] because there exist a large number of potential applications, e.g., human computer interaction, robot learning and control, and the imitation learning. For example, in the imitation learning task, a teacher may want to show a robot how to serve food. This activity includes several action components: walk to the location of food, hold it, and return to the point by walking. It would be easier for the robot to learn the task if the robot is able to recognize the key action components rather than to imitate the whole task.

Depending on the way of feature extraction, approaches used for human action recognition are roughly classified into two types: sensor-based approach (e.g., motion capture data), and vision-based approach.

In the vision-based approaches, feature commonly used for action recognition include shape [3], [4], optical flow [5], [6], point trajectory [7], [8], and joint angle [9]. Most of these features encode the kinematics of the human motion, and in general they are high dimensional.

Recognizing human actions from a video is a challenging task [14]. Firstly, it is difficult to identify an action independent of viewing direction. Secondly, there is difficulty of extracting stable features due to noise, or model fitting errors.

In the sensor based approaches, sensors are usually placed in the environment or attached to the human body to capture the human motion. Yang et al. [29], proposed an action recognition system using wearable motion sensor networks. Also, there are several works which use acceleration data for

A. Mansur, Y. Makihara, and Y. Yagi are with The Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka, 567-0047 Japan. Emails: `fuji.mansur@gmail.com`; `makihara@am.sanken.osaka-u.ac.jp`; `yagi@am.sanken.osaka-u.ac.jp`

action recognition [30]. In several works [35], [36], action recognition is done using the joint angles or point trajectories obtained from the human motion capture data.

In general, for many actions, kinematics features are high dimensional and have small interclass variation. Moreover, they do not consider the interaction of environment on human motion. On the other hand, recently, physics-based models have been used successfully for 3D people tracking [10], [11]. The physics-based models provide parameterizations for effective modeling of plausible poses and motions. In addition, they are capable to capture the influence of gravity, ground contact, and other physical interactions with the environment on pose and motion. In a related work [12], internal joint torques and external forces were recovered from observed motions in an inverse problem setting.

Inspired by these works, we attempt to use dynamics features obtained from physics model for human action recognition. Use of dynamics allows us to represent an action by the torques and forces governing the human motion. We believe, these kind of feature will result in a more discriminative feature.

Our main contribution in this paper is that we represent actions using dynamics features, namely joint torques. Use of such a feature has several advantages over kinematics features. Firstly, these features are more discriminative than the kinematics features. Secondly, they provide a low dimensional representation of the actions. Low dimensional representation is necessary in many cases to handle with the limited number of training data. We consider only the lower body torques for action representation, as they implicitly includes the configuration of the other limbs of the body and the external forces acting on the body such as gravity and the ground reaction force (GRF). In most actions, human body remains in an upright position while the legs supporting the body. As a result, upper body parts have more influence on the dynamics of the lower body parts. Usually, lower body parts have little influence on the dynamics of the upper body segment. Therefore, we choose to use the lower body torques for action recognition. Due to low dimensional representation, we can use simple learning and classification framework and achieve a good classification performance with a small number of training data.

II. RELATED WORK

Human action and activity recognition is an active area of research in the field of computer vision for over two decades. Large number of literatures exists in this field. A comprehensive review of these works is presented in survey

papers [13], [14]. However, in this section, we limit our discussion some of the papers closely related to our research.

We categorize the approaches for human action recognition on the basis of action representation. Some of the most popular representations include learned geometrical models of human body parts, spatio-temporal templates, appearance or region features, shape information, interest-point-based representations, and motion/optical flow patterns. In the following paragraphs, we provide a brief summary of the related work.

Appearance-based approaches learn the appearance model of the human body or body parts and try to match it to images in a test scene for action or gesture recognition [15], [16]. Actions are learnt by HMM or other variants of it. Appearance-based approach works well for gesture recognition, since the appearance of the human hand do not change much for the change of a person. However, such methods do not work well for whole body actions as they can not handle the problem of clothing or appearance change.

Shape-based representations utilize silhouettes of the human body [3] or the features derived from silhouettes [4]. The basic idea behind shape-based representation is that an action consists of a sequence of poses that can be detected in a single frame. Usually, recognition is based on the single frame; however, to improve the robustness, it may be extended to multiple frames. In [17], silhouette-based representation was extended to characterize the outline of human body in space-time domain. This results in a spatio-temporal 3D volume which is constructed by stacking the silhouettes detected in each frame. Shape-based approaches works well on a number of actions. However, they also suffer from a problem of silhouette changes due to clothing changes or imperfect segmentation.

Another class of works use volumetric analysis of video for action recognition [18], [17], [19]. Yilmaz et al. [18] proposed to use spatio-temporal features in order to simultaneously exploit both shape and motion features, and Blank et al. [17] extended a method developed for analysis of 2D shapes, to deal with volumetric spatio-temporal shapes induced by human actions. The main advantage of the volume-based approach is that it is not necessary to build complex models of body configuration and kinematics. In addition, recognition can be performed directly from the raw video.

Recently, interest-point-based representations has gained a lot of interest. They use spatio-temporal interest points and their trajectories for action and activity analysis [20], [21], [22], [23]. The main strength of this representation is the robustness to occlusion, since there is no need to track or detect the whole human body.

A number of researchers [5], [6], [25], [26], [24] have used features based on motion and optical flow. Bobick et al. [25] proposed motion energy image (MEI) to describe the cumulative spatial distribution of motion energy in a given sequence. Later, the idea of MEI is extended to motion history image (MHI) [27]. In optical flow-based approaches [5], [6], optical flow is used to derive a representation which

is suitable for recognition. It was shown in [27] that MEI and MHI have good discriminative performance for some particular simple action classes. However, it was reported in [24] that MEI and MHI show unsatisfactory performance for complex actions due to over-writing of the motion history.

In the sensor-based approaches, human action is usually described by the output signal obtained from sensors either attached to the human body or installed in the environment. In [28], an effective subject recognition approach is designed using ground reaction force (GRF) measurements of human gait. In [29], Yang et al. proposed an action recognition system using wearable motion sensor networks. Also, there are several works which use acceleration data obtained from the sensors [30] for action recognition. The main disadvantage of the sensor-based methods is that subjects need to wear some specialized devices or sensors or have to perform the action in a specialized environment.

In summary, our proposed framework is different from the previous representations in three ways. Firstly, from the human motion, we compute joint torques using physics model. These features are more discriminative than the kinematics features. Secondly, information about the ground interaction is encoded in the torques as the computation of torques depends on the external forces. Thirdly, we use a low dimensional representation of human actions which allows us to achieve a good classification performance with a relatively few training data in a simple classification framework such as HMM.

III. Feature Extraction

A. Kinematics Model

Our 3D articulated human body model consists of 12 rigid body segments and has a total of 26 degrees of freedom (DOFs), as shown in Fig. 1. We use three kinds of joints to link segments to their parent segments: 1-DOF (hinge), 2-DOF (saddle) or 3-DOF (ball and socket) rotational joints. This ensures that only relevant rotations about specific joint axes are possible. The position and orientation of the root segment is defined in the world coordinate by a 6-DOF global joint. All segments are approximated by sticks of appropriate lengths. We use identical lengths of body segments, mass and inertial parameters [32] for all subjects and actions. We estimate the internal joint torques by applying 3D motion capture data to the model using the method described in the next section. Note that, we are not dealing with the tracking problem in this paper, therefore, we assume that pose tracking data are available to us. This is a realistic assumption as the current markerless tracking algorithms can provide tracking result with good accuracy [33], [34].

B. Recovery of Torques

Given the mass, and the measurement of the acceleration, the net force acting on a particle can be determined by Newton's second law of motion. This net force has two components: force due to gravity and force due to contact. This idea can be extended to analyze the dynamics (to consider the forces and torques responsible for a motion)

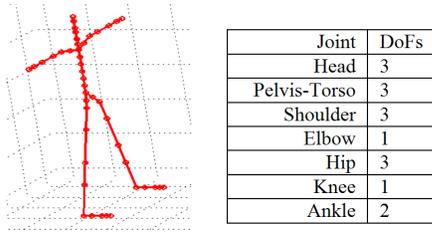| Joint	DoFs
Head	3
Pelvis-Torso	3
Shoulder	3
Elbow	1
Hip	3
Knee	1
Ankle	2

Fig. 1. Kinematic model and the DOF of each part

of human body. However, unlike a particle, human body is articulated with a large degree of freedom. To compute the contact forces and joint torques, we have to take account of all the body segments which are constrained by rotational joints. In addition, we need to consider the multiple contact points with the environment.

To recover the ground reaction forces and the joint torques, we follow the approach presented in [12]. For the sake of completeness of this paper, we briefly describe this approach. Let us consider an articulated human body with P parts and N degrees of freedom (DoF). The root of the body has three translational and three rotational degrees of freedom in world coordinate. All other joint angles are defined relative to their parent segments. This results in a $(N-6)$ joint angles. Using a Lagrangian formulation, we expresses the configuration of the body in terms of its generalized coordinates, $\mathbf{q} \in \mathbb{R}^N$, and N second order differential equations which govern the motion of the body:

$$M(\mathbf{q})\ddot{\mathbf{q}} = F(\mathbf{q}, \dot{\mathbf{q}}) + A(\mathbf{q}, \dot{\mathbf{q}}) \qquad (1)$$

In the equation above, $\dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ denote the first and second time derivatives of $\mathbf{q}$, M is the generalized mass matrix, F is the vector of generalized forces acting on the N DoFs (which includes contact, gravity, and joint torques), and A consists of remaining terms including those necessary to enforce joint constraints. These equations can be derived using the TMT method described in [31]. Our objective is to recover the $(N-6)$ internal torques τ_{int} from the observed N generalized accelerations. We first express F in terms of the internal torques τ_{int}, and the external forces acting on the body:

$$F(\mathbf{q}, \dot{\mathbf{q}}) = A_{int}\tau_{int} + \tau_{ext}(\mathbf{q}, \dot{\mathbf{q}}) \qquad (2)$$

Here the matrix A_{int} is responsible for mapping the joint torques into the vector of N generalized forces. As the estimates of internal torques depend strongly on the external forces, we need to consider them to recover the joint torques. We consider gravity and the GRFs as the external forces and for GRF, we consider only two contact points in each foot. We use an identical ground contact model as proposed in [12] and we only used the vertical component of the ground reaction force.

C. Feature Selection

We use HMM with output densities represented by mixture of Gaussians, and this sometimes make the model incomputable due to small training dataset and we cannot learn a

good model from high dimensional data. Therefore, we use only four joint torques from lower body parts: two knee joint torques and two hip joint torques. These torques represent an action with good consistency, and implicitly include the pose of the upper body parts as well as the external forces.

IV. CLASSIFICATION FRAMEWORK

HMM is a good probabilistic framework for modeling the dynamics of human action [15]. In our HMM configuration, we have one hidden state variable S and each state can emit a vector valued observation O (torque). This model is based on two dependency assumptions: (1) hidden variable S_t at time t depends only on the hidden variable at previous time step S_{t-1} (2) observation variable O_t at time t depends only on S_t. During the training step, the model learns the parameter set λ consisting of prior probabilities $P(S_0)$, transition probabilities $P(S_t|S_{t-1})$, and the observation probabilities $P(O_t|S_t)$ using available observed data. Learning the parameters of these distributions corresponds to maximizing the joint probability $P(O, S)$. For each action class, we learn a separate HMM model λ_i. During recognition, given a test observation sequence $O_{1:T}$, we select the class label as:

$$c = \arg\max_i P(\lambda_i|O_{1:T}) \qquad (3)$$

Due to the use low dimensional torque and joint angle trajectories, we did not face difficulty learning good models from small training set using a simple HMM framework.

V. EXPERIMENTS

A. Dataset

Experiments were performed on the data set containing 3-dimensional motion capture sequences obtained from CMU mocap data [37]. Fig. 2 shows some typical sequences of seven actions from this data set. In total, we use 171 sequences of 7 action classes, namely, walk, march, run, sit, jump forward, jump in place, and hop with 27, 21, 25, 21, 25, 32, and 20 instances, respectively. All seven classes have significant intra-class variations in terms of speed and style. In addition, for some actions, inter-class variation is very low. For example, joint angle trajectories of 'walk and march' and 'jump in place and hop' are quite similar for some sequences.

Walk, march, and run classes have variations in terms of speed, stride length, bounce, and arm swing. In the sit class, the subjects move their legs randomly after sitting on a stool. This results in uncorrelated knee joint angles. In case of jump forward and jump in place, knee and hip joint angles have good similarity, and sometimes it is difficult to differentiate if translational motion of the body is not considered. In summary, all the action classes contain significant intra-class variations, and therefore, this is a very challenging data set.

B. Baseline Method

We compare the results of dynamics-feature based approach with the baseline method which uses the joint angle features and the same HMM classifier. In the baseline method, we use two different feature sets. First set of features consists of the 6 principal components obtained by reducing
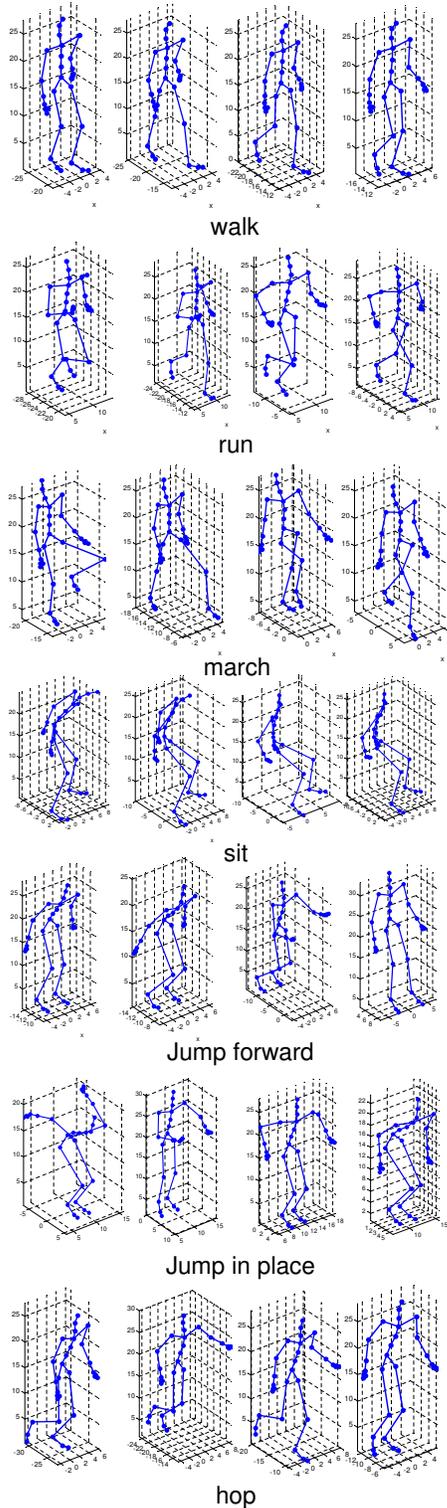
Fig. 2. Sample sequences of seven actions from CMU motion capture dataset.

the original 26 dimensional joint angles by PCA. More than 98% of the variance is accounted for by these 6 principal components. In the second set, we manually selected two knee joint angles and two hip joint angles. These particular joint angles have good consistency for a particular action class whereas the other joint angles are noisy. Dimension reduction was necessary to avoid difficulty in learning a good model from a relatively small training dataset.

C. Parameter Selection

To fine tune the parameters of HMM, we use 10 fold cross validation. We tune the number of states N_s and the number of mixture components K for the best performance. For the experiments with joint torques we used $K = 3$ and $N_s = 3$. In case of joint angles, we used $K = 4$ and $N_s = 2$ for best classification result.

D. Results

The classification results obtained by our approach and those from kinematics features are shown in Table I. To produce the results, we compute the average over 10 runs with random permutation of training set. Using the dynamics feature, we achieved mean accuracy of 95% on the entire data set. Walk, run and march action classes were classified with 100% accuracy. There was some confusion in case of classifying jump in place, which was confused with jump forward (15%). Remaining confusions are negligible ($\leq 2\%$). This is a reasonable performance considering the similarity between these actions.

Using the baseline method, we obtained mean classification accuracy of 78.4% in the first case (PCA reduced, 6 components) and 85.4% in the second case (4 joint angles). Dynamics feature outperforms both of these feature sets. These results also demonstrate that PCA selected features were not robust for discrimination. Confusion matrices for kinematics features are shown at the top (PCA) and the middle of the Table I.

In case of kinematics feature, there are significant confusions during the classification of run, sit, jump forward, and jump in place classes. However, confusion was significantly small in case of dynamics-based features.

To evaluate the discriminative power of the dynamics and kinematics based features, we compare some of these features for two action pairs: (1) jump forward and jump in place (Fig. 3) and (2) run and march (Fig. 4). It can be noticed that inter-class distance is small in case of joint angles, whereas the distance is large in case of joint torques. We notice that PCA reduced kinematics features also show poor discrimination.

We have done another set of experiments to test the robustness of the method in case of noisy motion capture data. To simulate the noisy measurements, we added Gaussian noises to all joint angles. However, for vertical distance of the foot from the ground, we used clean motion capture data because the accurate determination of the ground contact event is required for computing the torques accurately. In our method, we need to differentiate the joint angles to obtain the angular velocities and accelerations. However, due to the abrupt change in the noisy joint angles, we obtain unrealistic values for velocities and accelerations. To overcome this problem, we sampled the noisy joint angles at regular intervals and

True labels

	walk	run	march	sit	jump_f	jump_ip	hop
walk	0.94	0.02	0	0.02	0	0	0.03
run	0.01	0.81	0.02	0.06	0.02	0.03	0.08
march	0.03	0.15	0.9	0.02	0.08	0.14	0
sit	0	0	0.02	0.52	0.03	0.01	0.03
jump_f	0.02	0	0.02	0.26	0.8	0.03	0
jump_ip	0	0.02	0.04	0.1	0.03	0.66	0
hop	0	0	0	0.02	0.03	0.13	0.88

(Predicted labels)

True labels

	walk	run	march	sit	jump_f	jump_ip	hop
walk	0.98	0	0	0	0	0	0
run	0	0.93	0.04	0	0	0	0
march	0.02	0.07	0.96	0	0	0	0
sit	0	0	0	0.76	0.01	0.03	0
jump_f	0	0	0	0.1	0.83	0.28	0
jump_ip	0	0	0	0.14	0.16	0.69	0
hop	0	0	0	0	0	0	1

(Predicted labels)

True labels

	walk	run	march	sit	jump_f	jump_ip	hop
walk	1	0	0	0	0	0	0
run	0	1	0	0	0	0	0
march	0	0	1	0	0	0	0
sit	0	0	0	0.98	0	0.01	0
jump_f	0	0	0	0.02	0.98	0.15	0
jump_ip	0	0	0	0	0.02	0.84	0.02
hop	0	0	0	0	0	0	0.98

(Predicted labels)

True labels

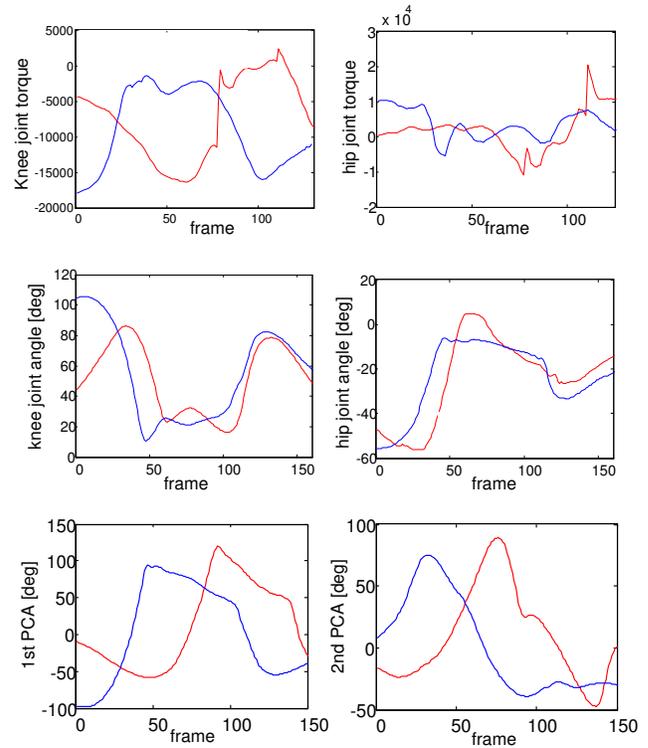|	walk	run	march	sit	jump_f	jump_ip	hop
walk	1	0	0	0	0	0	0
run	0	0.94	0	0	0	0	0
march	0	0.06	1	0	0	0	0
sit	0	0	0	0.78	0.01	0.05	0
jump_f	0	0	0	0.1	0.9	0.15	0
jump_ip	0	0	0	0.12	0.09	0.8	0
hop	0	0	0	0	0	0	1

(Predicted labels)

Fig. 3. Comparison of joint angle and torque feature for two action classes (jump forward and jump in place). Top row: knee and hip joint torques, middle row: knee and hip joint angles, bottom row: 1st and 2nd principal component of PCA reduced joint angles. Red: jump forward, blue: jump in place. The computed torques were not calibrated and therefore we do not use any absolute unit for them.

reconstructed the smoothed versions by using polynomial curve fitting. With these noisy joint angles, we computed the joint torques and the mean classification accuracy was 91.9%. Confusion matrix for this experiment is shown in Table I. This result gives us an idea about the performance of the method on the output of a tracker.

In our current implementation in Matlab, the system takes few seconds to compute the torques from an action of about 100 frames. However, the processing could be made faster by using efficient code and implementing in C/C++.

VI. CONCLUSION AND FUTURE WORK

A. Conclusion

In this paper we introduced dynamics features which reflects the physics model for human action recognition. These dynamics features are computed from the available
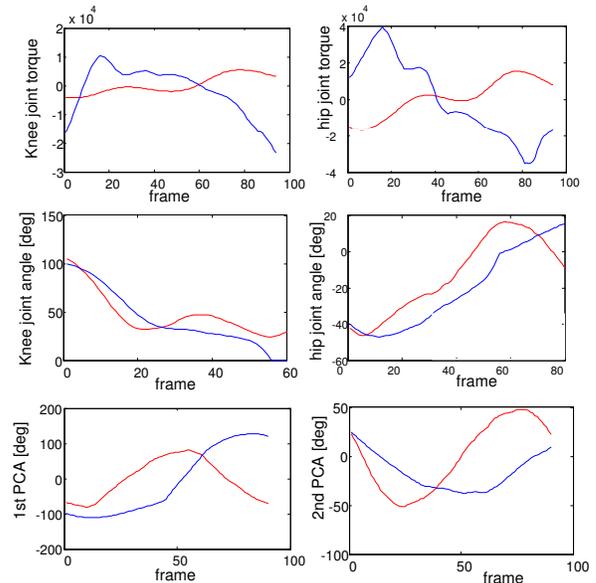

Fig. 4. Comparison of joint angle and torque feature for two action classes (run and march). Top row: knee and hip joint torques, middle row: knee and hip joint angles, bottom row: 1st and 2nd principal component of PCA reduced joint angles. Red: run, blue: march. The computed torques were not calibrated and therefore we do not use any absolute unit for them.

kinematics and the known mass and inertia properties of

human body. Using these features, we expressed seven action classes in terms of two hip and two knee joint torques, and used them for action recognition. This low dimensional representation allowed us to achieve a good classification accuracy with few training data. We carried out experiments on motion capture dataset containing different human actions and demonstrated the superiority of these dynamics feature over the kinematics features.

B. Future Work

We analyzed GRFs (Fig. 5) for several actions to evaluate their potential application for action recognition. We noticed that walk action produce similar GRFs from each foot and the reactions from each foot has small overlap. On the other hand, run action has picked and isolated reactions. In jump forward and hop, GRFs from both feet varies simultaneously. However, in case of hop, one foot does not produce any GRF as it is hanging above the ground. There are a few actions (such as jump forward and jump in place) for which we got similar GRFs. Therefore, it seems possible to use GRF for action recognition as used in [28] for subject recognition. However, we need a better estimation of GRF and have to consider an improved ground contact model.
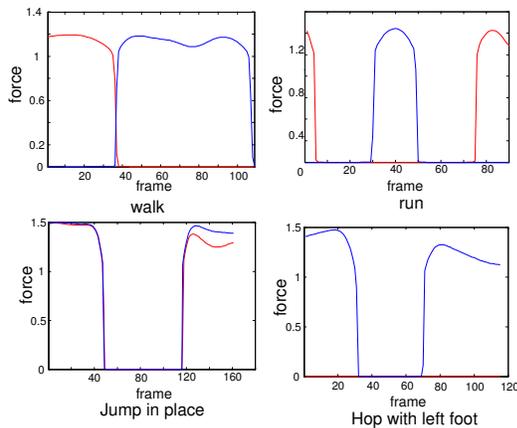

Fig. 5. Computed ground reaction forces for different actions. Red: right foot, blue: left foot. The computed forces were not calibrated and therefore we do not use any absolute unit for them.

REFERENCES

[1] Q. Zhou, S. Yu, X. Wu, Q. Gao, C. Li, and Y. Xu, *HMMs-based human action recognition for an intelligent household surveillance robot*, in proc. International conference on Robotics and Biometrics, 2009.

[2] O. Jenkins, G. Gonzlez, and M. Loper, *Tracking human motion and actions for interactive robots*, In proc. Human-Robot Interaction, 2007.

[3] L. Wang and D. Sutar, *Recognizing human activities from silhouettes: Motion Subspace and factorial discriminative graphical model*, in proc. CVPR 2007.

[4] D. Weinland, E. Boyer, *Action recognition using exemplar-based embedding*, in proc. CVPR 2008.

[5] D. Tran and A. Sorokin, *Human activity recognition with metric learning*, in proc. ECCV 2008.

[6] S. Ali and M. Shah, *Human Action Recognition in Videos Using Kinematic Features and Multiple Instance Learning*, IEEE TPAMI, February 2010.

[7] R. Messing, C. Pal, and H. Kautz, *Activity recognition using the velocity histories of tracked keypoints*, in proc. ICCV 2009.

[8] J. Sun, X. Wu, S. Yan, L.F. Cheng, T.S. Chua, and J. Li, *Hierarchical spatio-temporal context modeling for action recognition*, in proc. CVPR 2009.

[9] A. Bissacco, A. Chiuso, S. Soatto, *Classification and Recognition of Dynamical Models: The Role of Phase, Independent Components, Kernels and Optimal Transport*, IEEE TPAMI, June 2007.

[10] M. Vondrak, L. Sigal and O. C. Jenkins, *Physical Simulation for Probabilistic Motion Tracking*, in proc. CVPR 2008.

[11] M. Brubaker and D. J. Fleet, *The kneed walker for human pose tracking*, in proc. CVPR 2008.

[12] M. Brubaker, L. Sigal and D. Fleet, *Estimating Contact Dynamics*, in proc. ICCV 2009.

[13] P. Turaga, R. Chellappa, V.S. Subrahmanian, and O. Udrea, *Machine Recognition of Human Activities: A survey*, IEEE Transactions on Circuits and Systems for Video Technology, Nov. 2008.

[14] R. Poppe, A, *survey on vision-based human action recognition*, Image and Vision Computing 28 (2010).

[15] J. Yamato, J. Ohya, and K. Ishii, *Recognizing Human Action in Time Sequential Images Using Hidden Markov Model*, in proc. IEEE CVPR 1992.

[16] T. Starner and A. Pentland, *Visual Recognition of American Sign Language using Hidden Markov Model*, Internation Workshop on Automatic Face and Gensture Recognition, 1995.

[17] M. Blank et al., *Actions as Space-Time Shapes*, in proc. IEEE ICCV 2005.

[18] A. Yilmaz and M. Shah, *Actions Sketch: A Novel Action Representation*, in proc. IEEE CVPR 2005.

[19] E. Shechtman and M. Irani, *Space-Time Behavior Based Correlation*, in proc. IEEE CVPR 2005.

[20] A. Oikonomopoulousm, I. Patras and M. Pantic, *Spatiotemporal Saliency for Human Action Recognition*, in proc. IEEE ICME 2005.

[21] J. C. Niebles, H. Wang, and L. Fei-Fei, *Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words*, in proc. BMVC 2006.

[22] P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie, *Behavior Recognition via Sparse Spatio-Temporal Features*, in proc. IEEE International Workshop on VS-PETS, 2005.

[23] I. Laptev and T. Lindeberg, *Space Time Interest Points*, in proc. ICCV 2003.

[24] A. F. Bobick, *Movement, activity, and action: The role of knowledge in the perception of motion*, Philosophical Transactions of the Royal Society of London B, vol. 352, 1997.

[25] A. F. Bobick and J. Davis, *An Appearance-Based Representation of Action*, in proc. IEEE CVPR 1996.

[26] P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie, *Behavior recognition via sparse spatio-temporal features*, in proc. IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005.

[27] A. F. Bobick and J. W. Davis, *The recognition of human movement using temporal templates*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 3, 2001.

[28] S.P. Moustakidis, J.B. Theocharis, and G. Giakas, *Subject Recognition Based on Ground Reaction Force Measurements of Gait Signals*, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Dec. 2008.

[29] A. Y. Yang, R. Jafari, S. S. Sastry, and R. Bajcsy, *Distributed Recognition of Human Actions Using Wearable Motion Sensor Networks*, Journal of Ambient Intelligence and Smart Environments 1 (2009) 1-5 1.

[30] N. Ravi, N. Dandekar, P. Mysore, and M. Littman, *Activity recognition from accelerometer data. In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence (IAAI)*, 2005.

[31] M. Wisse, A. L. Schwab, and R. Q. V. Linde, *A 3D passive dynamic biped with yaw and roll compensation*, Robotica 19(3):275284, 2001.

[32] P. de Leva, *Adjustments to Zatsiorsky-Seluyanov's segment inertia parameters*, J. Biomech. 29(9):1223-1230, 1996.

[33] B. Rosenhahn, and T. Brox, *Scaled motion dynamics for markerless motion capture*, in proc. CVPR 2007.

[34] A. Balan et al., *Detailed human shape and pose from images*, in proc. CVPR 2007.

[35] Saad Ali, Arslan Basharat and Mubarak Shah, *Chaotic Invariants for Human Action Recognition*, in proc. ICCV 2007.

[36] Y. Shen, and H. Foroosh, *View-invariant Action Recognition from Point Triplets*, IEEE TPAMI, October 2009.

[37] http://mocap.cs.cmu.edu/