Foot-Mounted Inertial Measurement Unit for Activity Classification

Mostafa Ghobadi [1] and Ehsan T Esfahani[1]

***Abstract*— This paper proposes a classification technique for daily base activity recognition for human monitoring during physical therapy in home. The proposed method estimates the foot motion using single inertial measurement unit, then segments the motion into steps classify them by template-matching as walking, stairs up or stairs down steps. The results show a high accuracy of activity recognition. Unlike previous works which are limited to activity recognition, the proposed approach is more qualitative by providing similarity index of any activity to its desired template which can be used to assess subjects improvement.**

I. INTRODUCTION

Human activity recognition (HAR) is one of the most attractive fields of study in the areas of healthcare and pervasive computing. Some of HAR applications in pervasive healthcare include monitoring of patients undergoing physical therapies, care services for those having heart disease, diabetes, obesity and dementia and also eldercare. Among different measurement techniques in HAR, wearable sensors are the most desirable solution because of their ubiquitousness and unobtrusiveness nature [1].

Lara and Labrador has conducted a comprehensive literature review on HAR problem from technology, methods and applications angles with concentration on wearable sensors [1]. Zhou and Hu have also made a detailed study on human motion tracking problem with focus on rehabilitation subject, they have considered visual and non-visual sensors and their specific advantages and deficiencies [2]. Hadjidj et al. have reviewed different practical challenges of using wireless multi-sensor HAR [3].

HAR methods can be categorized into feature-based [4]–[6] and template-based [7]–[11] classifications. Feature based HAR mostly use two types of statistical feature sets extracted from a moving time window: 1- Non-directional features such as mean, variance, energy, entropy and power spectral density of the acceleration magnitude [4]. 2- Directional features such as eigenvalues of dominant directions and average velocities in heading and gravity directions [5]. Most of template-based HARs use either manifolds learning [7], [8] or gesture-string mapping that latter group simplifies the classification to a string-matching problem [9]–[11].

Windowing effect is one of the main challenges in feature-based HAR where, changing the size of time window and overlap percentage may cause ambiguity in activity recognition [1]. To overcome this problem, we segment motion trajectories based on correlation analysis of speed profile with a set of predefined templates. This step is done prior to motion classification. Each segmented data represents a single foot step and doesn't overlap with other segments.

Finally, the segmented data is classified into walking, stairs up or stairs down based on template matching of position and velocity profiles. Unlike previous works which are limited to activity recognition, the proposed approach is more qualitative by giving similarity index of any activity to its desired template which can be used to measure the qualitative progress of subjects activity. Although, this approach is used to detect the foot motions, it has the potential to be used as a general framework capable of detecting any predefined body movements.

II. MATERIALS AND METHODS

The experiment consisted of 3 subjects who were instructed to perform 3 different activities (walking straight, climbing stairs up and down). These activities were repeated 3 times and in total about 200 motion segments were recorded for each subject. For data recording an Inertial Measurement Unit (IMU) device consisting of 3 different types of motion sensors was utilized. The motion sensors included a tri-axial gyro with 16-bit resolution and +/-35 rad/sec range , a tri-axial accelerometer with 12-bit resolution and +/-78.5 m/s^2 range and one tri-axial magnetometer with 12-bit resolution and 8.1G range. The size of IMU device was 642 cm^3 and it was attached to a strap band tightened around the shank of subject's dominant leg as shown in Fig.1. For each experiment, 9 set of data (3 axis accelerations, angular velocities and magnetic fields) were recorded on a micro SD card at a sampling rate of 256Hz. All the data were recorded with respect to the body frame (xyz) shown in Fig. 1.

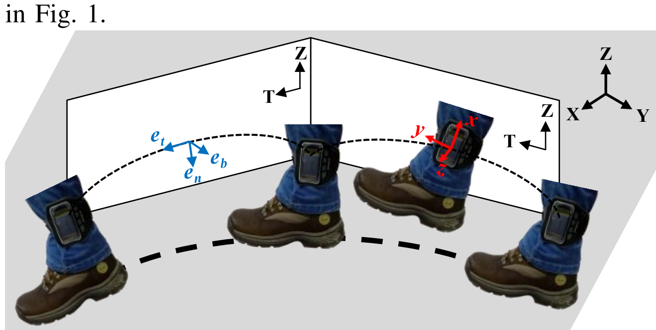

Fig. 1. Experimental setup and different coordinate system

III. DATA ANALYSIS

Block diagram of Fig. 2 shows different steps of the proposed approach to convert raw data into classified activity. In preprocessing stage, the raw data from sensors are

M. Ghobadi and ET Esfahani are with the Mechanical and Aerospace Engineering Department, University at Buffalo SUNY, Buffalo, NY 14221 USA.(e-mail: mostafag@buffalo.edu, ehsanesf@buffalo.edu). Phone: (716)645-2517.

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

converted to motion components such as acceleration ($\vec{a}$), velocity ($\vec{v}$) and position vectors ($\vec{p}$) in world and tangential frames. These motion components are then passed through the recognition stage, where they are first segmented and then classified using a template matching technique.

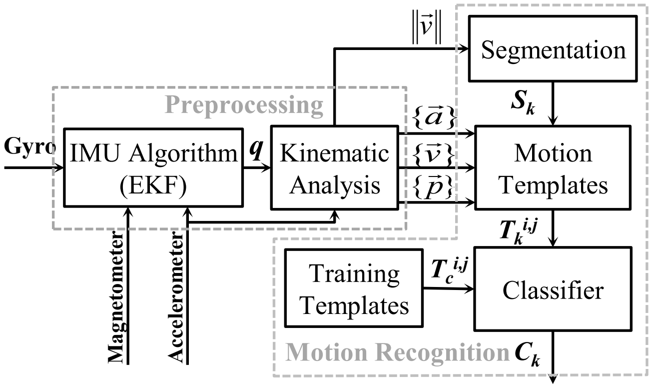

Fig. 2. Block diagram representation of the proposed method

A. IMU Algorithm

An IMU algorithm is developed to obtain spatial orientation of the body frame relative to the world frame. Let $\mathbf{q}$ in (1) be a rotation quaternion that can rotate an arbitrary position vector $\vec{p}$ around a given unit vector $\hat{n}$ with an angle of θ and result in $\vec{p'}$ using (2) [12], [13]. Where, $\bar{\mathbf{q}}$ is quaternion conjugate of $\mathbf{q}$.

$$q = cos(\theta/2) + sin(\theta/2)\hat{n} \tag{1}$$

$$\vec{p'} = \mathbf{q}\,\vec{p}\,\bar{\mathbf{q}} \tag{2}$$

Provided that the bases of the rotating body frame (xyz) is obtained by applying (2) on the bases of the inertial world frame (XYZ), any vector expressed in (xyz) can be transferred to its equivalent representation in (XYZ) using (3).

$$\vec{p}_{XYZ} = \mathbf{q}\,\vec{p}_{xyz}\,\bar{\mathbf{q}} \tag{3}$$

To obtain spatial orientation of the body frame relative to the world frame, the quaternion is calculated in terms of raw data by using an Extended Kalman Filter.

B. Kinematic Analysis

1) Calculation of Inertial Velocity and Position: After obtaining the spatial orientation, body acceleration ($\vec{a}_{xyz}$) is transferred to world acceleration ($\vec{a}_{XYZ}$) using (3). A low pass filter with cut-off frequency of 50 Hz is also applied to remove noises.

In order to calculate world velocity vector, Eq.(4) can be considered as the simplest approach where the standard gravity acceleration ($\vec{\mathbf{g}}$) is subtracted from the filtered acceleration and integrated with respect to time. However, numerical integral error is inevitable in this approach because of the sensor noise, uncertainty in gravity acceleration and computational error that lead to an inaccurate velocity vector.

$$\vec{v}_{XYZ} = \int_0^T \left(\vec{a}^f_{XYZ} - \mathbf{g}\hat{e}_Z\right) dt \tag{4}$$

To avoid this problem, we use the only possible observation that is inferable when the motion is completely stopped and there are very small changes in acceleration. Zero velocity can be observed correctly when the magnitude of acceleration approaches the standard gravity.

Therefore, an improvement to (4) can be obtained by introducing a new linear system including normal noises ($N(.)$) and a discrete observation (Y) as described in (5) and (6). Here, ϵ and T_s are acceleration and time thresholds used to define zero velocity conditions. In our analysis we used $\epsilon = 1m/s^2$ and $T_s = 0.1s$.

$$\begin{cases} \frac{d}{dt}\vec{v}_{XYZ} = \vec{a}^f_{XYZ} - \mathbf{g}\hat{e}_Z + N\left(\vec{\theta}(t), \sigma_a^2\right) \\ Y = \vec{v}_{XYZ} + N\left(0, \sigma_v^2\right) \end{cases} \tag{5}$$

$$Y = \vec{0} \quad , \quad \|\vec{a}_{XYZ} - \mathbf{g}\hat{e}_Z\| < \mathcal{E} \quad for \quad \Delta t < T_s \tag{6}$$

Observation Y can represent two different states: 1) resting states and 2) active states. Resting states are those with continuously zero-observed speed and the active states are those with continuously nonzero-observed speed values.

Let i^{th} resting state starts at t_i^0 and finishes at t_i, then following period would be the i^{th} active state that starts at t_i and finishes at t_{i+1}^0. Assuming $\vec{\theta}_i$ to be the mean value of acceleration noise within i^{th} active state, it can be estimated via (8). A relatively accurate estimation of velocity vector can then be obtained from piece-wise integration of (9) over active states. The world position vector can also be obtained by integrating the velocity vector over time.

$$\int_{t_i}^{t_{i+1}^0} \left(\vec{a}^f_{XYZ} - \vec{\theta}_i\right) dt = \vec{v}_{XYZ}\left(t_{i+1}^0\right) - \vec{v}_{XYZ}\left(t_i\right) = 0 \tag{7}$$

$$\vec{\theta}_i = \frac{1}{t_{i+1}^0 - t_i} \int_{t_i}^{t_{i+1}^0} \vec{a}^f_{XYZ} dt \tag{8}$$

$$\vec{v}_{XYZ}(t) = \int_{t_i}^{t} \vec{a}^f_{XYZ} dt - \vec{\theta}_i\left(t - t_i\right) \ , \ t_i < t < t_{i+1}^0 \tag{9}$$

2) Transferring Motions to the Tangential Frame: To make the motion profiles independent from both the observer and the orientation of sensor on the body, we transfer the world motion profiles (velocity, acceleration and angular velocity) to the tangential coordinate system. The tangential coordinates as shown in Fig. 1 is defined by $(\hat{\mathbf{e}}_t, \hat{\mathbf{e}}_n, \hat{\mathbf{e}}_b)$ where $\hat{\mathbf{e}}_t$ is a unit vector tangent to the motion trajectory and $\hat{\mathbf{e}}_n$ is directed to center of curvature. All directions are calculated with basic vector transformations.

A 2D projected position vector (P_T, P_Z), in each segment is obtained in TZ. As it is shown in Fig.1, TZ plane is composed of world-frame Z axis and a new axis T that is a horizontal projection of the average of $\hat{\mathbf{e}}_t$ over a segment period.

IV. Motion Recognition

A. Segmentation

Each Segment is defined as the part of motion between two successive resting states. Equation (6) can be used to determine segmentation by detecting the resting states. However, in practice this is not a reliable approach and often results in incorrect segmentation. The problem specially occurs when two or three succeeding steps are observed as a single active state. To avoid this problem, a template matching approach is applied on the velocity magnitude or speed profile to find the segments of the motion. We use the vector functions ($V_{T,j}$) as the matching speed template in segmentation as shown in (10).

$$V_{T,j} = \frac{1}{2} - \frac{1}{2}\cos\left(\frac{t\pi}{N_j}\right); t \in [0 \quad 2N_j], j = 1,2,3 \quad (10)$$

j value in (10) adjusts the acceleration and deceleration magnitudes of the template such that for $j = 1 to 3$ can approximately represent any motion with fast, normal, or slow paste, respectively. In experiments, parameter N_j are set in the range of 150 to 250 for all three fast to slow speed templates.

The template matching for segmentation is done in two steps. In the first step, the midpoint of each segment i^* is found by maximizing the Pearson correlation $C_{i,j}$ of a moving time window $V_{i,j}$ and each of the speed profile templates $V_{T,j}$ as shown in (11). $V_{i,j}$ is a time window of size $2N_j$ centered at the i^{th} sample of data. i^* is the local maximum values of $C_{i,j}$ that finds the best velocity past (j value) and the time window for segmentation such that the peaks of velocity profile and template coincide on each other.

$$C_{i,j} = \frac{cov\langle V_{i,j}, V_{T,j}\rangle}{\sqrt{var\langle V_{i,j}\rangle . var\langle V_{T,j}\rangle}} \quad (11)$$

$$i^* = \arg\max_{i,j}(C_{i,j}), \quad j = 1,2,3 \quad (12)$$

In second step, the motion speed as well as the starting and ending points of the segments are detected. For this purpose, the sum of point-to-point Euclidean distance of each template $V_{T,j}$ and its corresponding window $V_{i^*,j}$ is calculated. j^* is the index of the template with the smallest distance (V_{T,j^*}) that represent the motion speed. Finally, the length of the best fitted template ($2N_{j^*}$) will be used to determine $S_s^k = i^* - N_{j^*}$ as the starting sample point and $S_e^k = i^* + N_{j^*}$ as the ending sample point of the k^{th} template.

Using S_s^k and S_e^k, we may segment all other motion profiles such as tangential and normal accelerometers, angular velocities, and world velocity and position vectors for further analysis in the classification stage.

B. Classification of Motion Templates

In this section, we first introduce the motion templates and classes, then propose a likelihood metric for template matching.

Any pair consisting two arbitrary jointed components of motion including time is considered as a motion template, e.g. any segment of ($||\vec{v}||$ vs. time) trajectory is a template; another template can be one segment of (a_n vs. a_t) trajectory. However, some intermediate calculations are necessary to constitute the templates. If an array of Motion Components (M) is defined as (13), then vector M_k^i can be defined as k^{th} segment of i^{th} motion component as (14).

$$M = \{time, \|\vec{v}\|, a_t, a_n, v_T, v_Z, p_T, p_Z, \ldots\} \quad (13)$$

$$M_k^i = \left[M^i(S_s^k) \ldots M^i(S_e^k)\right] \quad (14)$$

To remove the effect of the motion speed all templates are normalized by (15). A motion template of (16) is finally composed as a pair of two arbitrary normalized motion components that is resampled to a certain number of sample points.

$$M_k'^i = \frac{M_k^i(p) - M_k^i(k)}{\max(M_k^i) - \min(M_k^i)} \quad (15)$$

$$T_k^{(i,j)} = (M_k'^i, M_k'^j), i \neq j \quad (16)$$

Each class is assigned to a specific type of motion, for example in the current study, we assign three classes to foot motions such as straight walking , stairs up and stairs down steps. Any of the templates belonging to a certain class is expected to demonstrate a unique shape. This shape would not be necessarily distinguishable by its appearance though, it can be detected through distance scoring. An Euclidian distance index of (17) similar to the 1-$ recognizer is used to compare two motion templates [14]. Each time, $T_c^{i,j}$ is one of the classifier templates.

$$d_{k,c}^{(i,j)} = \|T_k^{(i,j)} - T_C^{(i,j)}\| \quad (17)$$

To classify one unclassified template using some known classifier templates, the class of that classifier template which yields the highest likelihood score is assigned to the template. A one-out method is used to classify the motion segments at training in which three classifier templates from three different classes are selected randomly at each training round. At validation stage, templates with highest performance in training are used as predefined classifier templates.

Throughout the classification process, motion segments are categorized into different classes and level of similarity of every segment to each class is obtained.

V. Results

Kinematic output of the preprocessing step is shown in Fig.3. This figure represents position coordinates of eight walking cycles while the subject is moving on a circle.

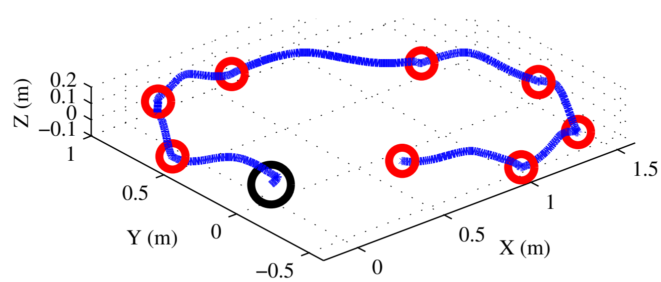

Fig. 3. XYZ coordinates subject foot while walking on a circle.

The results for the correlation matching between a sample speed profile and speed template 2 are illustrated in Fig. 4. The local maximum points are most probable place of peaks.

Four types of templates are used for classification purposes. The accuracy of the classification for each template-based classifier is shown in Table 1. The evaluation outputs of the classifier based on (p_T, p_Z) and (v_T, v_Z) templates are listed in Table 2, 3. These templates show strong capability in classifying motion segments to their true classes. However,

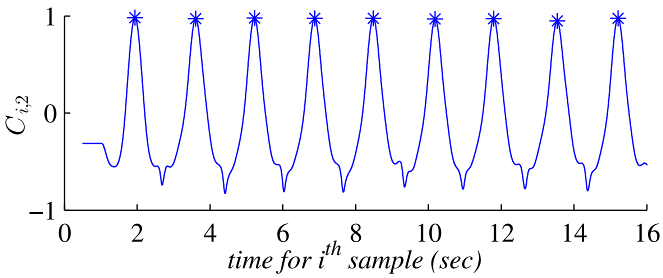

Fig. 4. Correlation coefficient between a speed profile and template 2. Points with maximum correlation are marked by star as segments midpoint.

the first classifier have 94.30% accuracy and sometimes cannot detect straight walking steps that happens for low scoring value, the second classifier gives 99.70% accuracy in detecting the same class. Templates for (p_T, p_Z) are shown in Fig. 5. For testing purpose, some predefined trajectories are used as classifying templates which are highlighted.

TABLE I
CLASSIFICATION RESULTS FOR CONSIDERED TEMPLATES

	(p_T, p_Z)	(v_T, v_Z)	(t, a_t)	(t, a_n)	(a_t, a_n)
Accuracy(%)	96.89	99.80	82.43	73.99	82.06
Deviation(%)	5.09	1.56	10.87	10.87	10.87

TABLE II
CLASSIFICATION RESULTS FOR (p_T, p_Z) TEMPLATE

	Walking	**Stairs Up**	**Stairs Down**	**Undetected**
Walking	94.30%	0.08%	0.00%	5.62%
Stairs Up	0.00%	100.00%	0.00%	0.00%
Stairs Down	0.00%	0.00%	100.00%	0.00%

TABLE III
CLASSIFICATION RESULTS FOR (v_T, v_Z) TEMPLATE

	Walking	**Stairs Up**	**Stairs Down**	**Undetected**
Walking	99.70%	0.22%	0.08%	0.00%
Stairs Up	0.05%	99.95%	0.00%	0.00%
Stairs Down	0.00%	0.00%	100.00%	0.00%

VI. CONCLUSION

The proposed method can recognize the predefined foot motions with a high accuracy close to 100% for position template. The results comparing previous works show a serious progress to approach to perfect classification by using template matching. Templates including accelerations components lead to lower accuracy which is probably because of high sensitivity to input noise which vanishes after integration in velocity and position components. Although, the preprocessing stage is not the main goal of study, its performance has a major effect on the recognition stage. The performance of preprocessing steps including IMU algorithm and kinematic analysis can be evaluated in the quality of results in the classification stage.

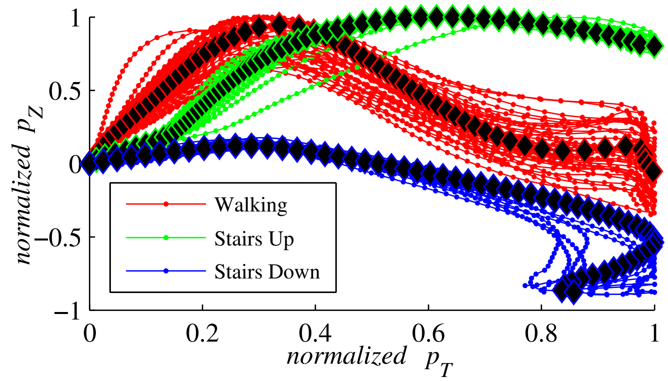

Fig. 5. (p_T, p_Z) testing templates. Training templates are highlighted.

REFERENCES

[1] O. D. Lara and M. A. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, pp. 1192–1209, Jan. 2013.

[2] H. Zhou and H. Hu, "Human motion tracking for rehabilitationA survey," *Biomedical Signal Processing and Control*, vol. 3, pp. 1–18, Jan. 2008.

[3] A. Hadjidj, M. Souil, A. Bouabdallah, Y. Challal, and H. Owen, "Wireless sensor networks for rehabilitation applications: Challenges and opportunities," *Journal of Network and Computer Applications*, vol. 36, pp. 1–15, Jan. 2013.

[4] T. Huynh and B. Schiele, "Analyzing features for activity recognition," in *Proceedings of the 2005 joint conference on Smart objects and ambient intelligence innovative context-aware services: usages and technologies*, (New York, NY, USA), pp. 159–163, ACM Press, 2005.

[5] M. Zhang and A. A. Sawchuk, "A feature selection-based framework for human activity recognition using wearable multimodal sensors," in *Proc. Int. Conf. Body Area Networks*, pp. 92–98, ICST, Nov. 2011.

[6] N. Alshurafa, W. Xu, J. Liu, M.-C. Huang, B. Mortazavi, C. Roberts, and M. Sarrafzadeh, "Designing a Robust Activity Recognition Framework for Health and Exergaming using Wearable Sensors.," *IEEE Journal of Biomedical and Health Informatics*, Oct. 2013.

[7] M. Zhang and A. A. Sawchuk, "Manifold Learning and Recognition of Human Activity Using Body-Area Sensors," in *10th International Conference on Machine Learning and Applications and Workshops*, (Honolulu, Hawaii, USA), pp. 7–13, IEEE, Dec. 2011.

[8] L. Schwarz, D. Mateus, V. Castaneda, and N. Navab, "Manifold Learning for ToF-based Human Body Tracking and Activity Recognition," in *Procedings of the British Machine Vision Conference 2010*, (Aberystwyth, Wales, UK), pp. 80.1–80.11, 2010.

[9] T. Stiefmeier, D. Roggen, and G. Tröster, "Gestures are strings: efficient online gesture spotting and classification using string matching," in *Proceedings of the ICST 2Nd International Conference on Body Area Networks*, (Florence, Italy), pp. 16:1–8, June 2007.

[10] L.-V. Nguyen-Dinh, D. Roggen, A. Calatroni, and G. Troster, "Improving online gesture recognition with template matching methods in accelerometer data," in *12th International Conference on Intelligent Systems Design and Applications*, (Kochi, India), pp. 831–836, IEEE, Nov. 2012.

[11] H. Ghasemzadeh, V. Loseu, and R. Jafari, "Structural action recognition in body sensor networks: distributed classification based on string matching.," *IEEE transactions on information technology in biomedicine*, vol. 14, pp. 425–35, Mar. 2010.

[12] J. Diebel, "Representing Attitude : Euler Angles , Unit Quaternions , and Rotation Vectors," *Matrix*, vol. 58, pp. 1–35, 2006.

[13] E. Pennestrì and P. Valentini, "Dual Quaternions as a Tool for Rigid Body Motion Analysis: A Tutorial with an Application to Biomechanics," *Archive of Mechanical Engineering*, pp. LVII(2):187–205, 2010.

[14] J. O. Wobbrock, A. D. Wilson, and Y. Li, "Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes," in *Proceedings of the 20th annual ACM symposium on User interface software and technology*, UIST '07, (Rhode Island, NY, USA), pp. 159–168, ACM Press, 2007.