COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 117 (2014) 189–207

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Automated evaluation of physical therapy exercises using multi-template dynamic time warping on wearable sensor signals

Aras Yurtman, Billur Barshan *

Department of Electrical and Electronics Engineering, Bilkent University, Bilkent, TR-06800 Ankara, Turkey

ARTICLE INFO

Article history:
Received 15 October 2013
Received in revised form 9 July 2014
Accepted 14 July 2014

Keywords:
Physical therapy
Motion sensors
Inertial sensors
Dynamic time warping
Pattern recognition
Motion capture

ABSTRACT

We develop an autonomous system to detect and evaluate physical therapy exercises using wearable motion sensors. We propose the *multi-template multi-match dynamic time warping* (MTMM-DTW) algorithm as a natural extension of DTW to detect multiple occurrences of more than one exercise type in the recording of a physical therapy session. While allowing some distortion (warping) in time, the algorithm provides a quantitative measure of similarity between an exercise execution and previously recorded templates, based on DTW distance. It can detect and classify the exercise types, and count and evaluate the exercises as correctly/incorrectly performed, identifying the error type, if any. To evaluate the algorithm's performance, we record a data set consisting of one reference template and 10 test executions of three execution types of eight exercises performed by five subjects. We thus record a total of 120 and 1200 exercise executions in the reference and test sets, respectively. The test sequences also contain idle time intervals. The accuracy of the proposed algorithm is 93.46% for exercise classification only and 88.65% for simultaneous exercise and execution type classification. The algorithm misses 8.58% of the exercise executions and demonstrates a false alarm rate of 4.91%, caused by some idle time intervals being incorrectly recognized as exercise executions. To test the robustness of the system to unknown exercises, we employ leave-one-exercise-out cross validation. This results in a false alarm rate lower than 1%, demonstrating the robustness of the system to unknown movements. The proposed system can be used for assessing the effectiveness of a physical therapy session and for providing feedback to the patient.

1. Introduction and background

Physical therapy is an important type of rehabilitation for patients with various disorders. Cardiopulmonary medicine, neurology, orthopedics, and pediatrics are branches of medicine that may benefit from physical therapy [1,2]. Physical therapy usually requires exercising in a hospital or a rehabilitation center under the supervision of a specialist who assigns one or more exercises to the patient [3]. After learning how to perform the exercises correctly, patients usually need to continue exercising at home, where they receive no feedback [4]. Even at the hospital, specialists cannot follow each patient continuously during their exercise sessions because the

* *Corresponding author*. Tel.: +90 3122902161; fax: +90 3122664192.
E-mail addresses: yurtman@ee.bilkent.edu.tr (A. Yurtman), billur@ee.bilkent.edu.tr (B. Barshan).
http://dx.doi.org/10.1016/j.cmpb.2014.07.003
0169-2607/

former often alternate between several patients or have other tasks to do, a situation that can result in insufficient, inaccurate, and often subjective feedback [5,6].

An important issue in physical therapy is to evaluate the exercises and assess the effectiveness of an exercise session. Previous studies consider energy expenditure [7], the total duration of the exercise session, or the time period during which the patient is physically active [8]. The system in [3] generates a warning whenever relevant parameters exceed predefined thresholds to reduce the risk of over-exercising. These approaches may be misleading or fail if the patient performs the exercises incorrectly or at an unusual pace. Previous studies have not considered detecting sequential exercise executions and providing an objective evaluation of their accuracy to assess the effectiveness of a physical therapy session. Such a process is also difficult and tedious for a specialist, even when he is responsible for only a single patient, and impossible when he is monitoring several. Patients and specialists would highly benefit from an automated system.

Several different sensor modalities are used in physical rehabilitation, including inertial [9–12], visual [13–15], strain [4,16], medical [6], physiological, kinetic, and environmental sensors [17], or some combination of these [18,19]. Automatic monitoring of the people performing physical therapy exercises should be done without restricting their independence, intruding on their privacy, or degrading their quality of life. A commonly used approach is to fix cameras in the environment, which intrudes on privacy and usually has a relatively high installation cost. The main advantage of this approach is that the person does not have to wear or carry any sensors or devices on their body. This approach may also eliminate problems related to misplacing sensors on the body, although some vision-based systems require wearing/pasting special tags or markers. This method may be acceptable when the person always performs exercises at the same place, but when the exercises are performed in different places, e.g., indoors and outdoors, this approach becomes unsuitable. We believe that wearable sensors are superior to camera systems in these respects.

In [20], participants 3–9 months post-stroke with mild to moderate motor impairment of one arm wore an accelerometer on each arm outside the laboratory for three days before and after treatment or an equivalent no-treatment period. The use of the more impaired arm in daily life was assessed using low-pass filtered accelerometer recordings. Other studies that focus on post-stroke rehabilitation are [21–23].

In numerous studies, a 3D real-time human body model is constructed to observe movements [5,24,25]. In [24,26,27], patients perform the given exercises to complete tasks in video-game-like virtual environments, making exercise sessions more enjoyable.

Another approach to observing movements is to use *biofeedback* devices that transform sensor measurements from the body into sound, a blinking LED, or an observable shape on a screen [28–30]. Biofeedback provides detailed information regarding the lengthening, shortening, and physical exertion of a muscle. It also allows comparing the data of a healthy muscle to a non-healthy one when performing the same exercise. Biofeedback devices are sometimes combined with electromyography. Although these devices have become portable recently, older devices are immobile and costly, and are mostly used in hospitals or rehabilitation centers [31]. In addition, most devices either do not evaluate a patient's performance or they evaluate the results using simple thresholding. Hence, they require a specialist or the patient to evaluate the feedback [31], both of whom may be highly subjective. Therefore, biofeedback devices cannot replace an attendant specialist most of the time.

Below, we provide a summary of studies aiming to assess the accuracy of physical therapy exercises or classify them as correct/incorrect:

Fergus et al. [5] propose a tele-rehabilitation system that collects and stores the patient's motion data, utilizing body area and sensor networks, including inertial sensors. The system virtually simulates the patient's body motions on a 3D human body model in real time or using the stored data. The physician or physical therapist monitors the motions remotely to evaluate the patient's progress. This proposed solution is impractical and does not significantly improve inspection time because the system provides no information regarding the patient's movement capability, movement accuracy, or progress.

Using five body-worn tri-axial accelerometers, Taylor et al. [11] build a classifier that labels incorrectly performed exercises for knee osteoarthritis, a degenerative disease of the knee joint. Three exercises are performed by nine healthy subjects. The exercises are performed in the correct way as well as with particular errors. Several features extracted from the accelerometer data are used in the AdaBoost classifier to classify the exercises as correct or as having a particular error type. However, multiple errors are not allowed by the methodology and the classification accuracy is about 70% in most cases, which is not very high.

In [12], an Android application estimating the accuracy (i.e., score) of balance board exercises is developed, using a smartphone's internal accelerometer and magnetometer. A complex rule-based algorithm is proposed to obtain a score value closest to the score given by an expert; the difference between human and automatic scores is found to be fewer than 10 points in more than 75% of the exercises on a 0 to 100 scale. However, the proposed methodology does not yield an optimal solution, and different rule-based scoring algorithms are used for different exercise types, which makes it difficult to add a new exercise to the system.

In the MyHeart system [16], the accuracy of arm movements for post-stroke rehabilitation is determined using strain sensors. Healthy subjects wearing tight-fitting garments with printed strain sensors imitate how post-stroke patients might perform each of seven exercise types correctly and incorrectly under the supervision of a physician or therapist. An exercise is considered correctly performed if the similarity between the recorded signal and a pre-recorded template, calculated by the open-end DTW (see Appendix A.2), exceeds a threshold [12]. The system provides real-time feedback to the patient with an average classification accuracy of 85%. The main disadvantage of the system is the difficulty of putting on the garment for a post-stroke patient, even with help.

In [4], strain sensors worn on the arm are used to provide real-time feedback to patients undergoing motor rehabilitation. Seven exercises are executed correctly and incorrectly