Pattern Recognition 43 (2010) 3605–3620

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Comparative study on classifying human activities with miniature inertial and magnetic sensors

Kerem Altun, Billur Barshan *, Orkun Tunçel

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

ARTICLE INFO

Article history:
Received 6 October 2009
Received in revised form
30 March 2010
Accepted 22 April 2010

Keywords:
Inertial sensors
Gyroscope
Accelerometer
Magnetometer
Activity recognition and classification
Feature extraction
Feature reduction
Bayesian decision making
Rule-based algorithm
Decision tree
Least-squares method
k-Nearest neighbor
Dynamic time warping
Support vector machines
Artificial neural networks

ABSTRACT

This paper provides a comparative study on the different techniques of classifying human activities that are performed using body-worn miniature inertial and magnetic sensors. The classification techniques implemented and compared in this study are: Bayesian decision making (BDM), a rule-based algorithm (RBA) or decision tree, the least-squares method (LSM), the *k*-nearest neighbor algorithm (*k*-NN), dynamic time warping (DTW), support vector machines (SVM), and artificial neural networks (ANN). Human activities are classified using five sensor units worn on the chest, the arms, and the legs. Each sensor unit comprises a tri-axial gyroscope, a tri-axial accelerometer, and a tri-axial magnetometer. A feature set extracted from the raw sensor data using principal component analysis (PCA) is used in the classification process. A performance comparison of the classification techniques is provided in terms of their correct differentiation rates, confusion matrices, and computational cost, as well as their pre-processing, training, and storage requirements. Three different cross-validation techniques are employed to validate the classifiers. The results indicate that in general, BDM results in the highest correct classification rate with relatively small computational cost.

1. Introduction

Inertial sensors are self-contained, nonradiating, nonjammable, dead-reckoning devices that provide dynamic motion information through direct measurements. Gyroscopes provide angular *rate* information around an axis of sensitivity, whereas accelerometers provide linear or angular velocity rate information.

For several decades, inertial sensors have been used for navigation of aircraft [1,2], ships, land vehicles, and robots [3–5], for state estimation and dynamic modeling of legged robots [6,7], for shock and vibration analysis in the automotive industry, and in telesurgery [8,9]. Recently, the size, weight, and cost of commercially available inertial sensors have decreased considerably with the rapid development of micro electro-mechanical systems (MEMS) [10]. Some of these devices are sensitive around a single axis; others are multi-axial (usually two- or three-axial). The availability of such MEMS sensors has opened up new possibilities for the use of inertial sensors, one of them being human activity monitoring, recognition, and classification through body-worn sensors [11–15]. This in turn has a broad range of potential applications in biomechanics [15,16], ergonomics [17], remote monitoring of the physically or mentally disabled, the elderly, and children [18], detecting and classifying falls [19–21], medical diagnosis and treatment [22], home-based rehabilitation and physical therapy [23], sports science [24], ballet and other forms of dance [25], animation and film making, computer games [26,27], professional simulators, virtual reality, and stabilization of equipment through motion compensation.

Early studies in activity recognition employed vision-based systems with single or multiple video cameras, and this remains the most common approach to date [28–31]. For example, although the gesture recognition problem has been well studied in computer vision [32], much less research has been done in this area with body-worn inertial sensors [33,34]. The use of camera systems may be acceptable and practical when activities are confined to a limited area such as certain parts of a house or office environment and when the environment is well lit. However, when the activity involves going from place to place, camera

* Corresponding author.
E-mail address: billur@ee.bilkent.edu.tr (B. Barshan).

0031-3203/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2010.04.019

systems are much less convenient. Furthermore, camera systems interfere considerably with privacy, may supply additional, unneeded information, and cause the subjects to act unnaturally.

Miniature inertial sensors can be flexibly used inside or behind objects without occlusion effects. This is a major advantage over visual motion-capture systems that require a free line of sight. When a single camera is used, the 3-D scene is projected onto a 2-D one, with significant information loss. Points of interest are frequently pre-identified by placing special, visible markers such as light-emitting diodes (LEDs) on the human body. Occlusion or shadowing of points of interest (by human body parts or objects in the surroundings) is circumvented by positioning multiple camera systems in the environment and using several 2-D projections to reconstruct the 3-D scene. This requires each camera to be separately calibrated. Another major disadvantage of using camera systems is that the cost of processing and storing images and video recordings is much higher than those of 1-D signals. 1-D signals acquired from multiple axes of inertial sensors can directly provide the required information in 3-D. Unlike high-end commercial inertial sensors that are calibrated by the manufacturer, in low-cost applications that utilize these devices, calibration is still a necessary procedure. Accelerometer-based systems are more commonly adopted than gyros because accelerometers are easily calibrated by gravity, whereas gyro calibration requires an accurate variable-speed turntable and is more complicated.

The use of camera systems and inertial sensors are two inherently different approaches that are by no means exclusive and can be used in a complementary fashion in many situations. In a number of studies, video cameras are used only as a reference for comparison with inertial sensor data [35–40]. In other studies, data from these two sensing modalities are integrated or fused [41,42]. The fusion of visual and inertial data has attracted considerable attention recently because of its robust performance and potentially wide applications [43,44]. Fusing the data of inertial sensors and magnetometers is also reported in the literature [38,46,47].

Previous work on activity recognition based on body-worn inertial sensors is fragmented, of limited scope, and mostly unsystematic in nature. Due to the lack of a common ground among different researchers, results published so far are difficult to compare, synthesize, and build upon in a manner that allows broad conclusions to be reached. A unified and systematic treatment of the subject is desirable; theoretical models need to be developed that will enable studies designed such that the obtained results can be synthesized into a larger whole.

Most previous studies distinguish between sitting, lying, and standing [18,35–37,39,45,48–50], as these postures are relatively easy to detect using the static component of acceleration. Distinguishing between walking, and ascending and descending stairs has also been accomplished [45,48,50], although not as successfully as detecting postures. The signal processing and motion detection techniques employed, and the configuration, number, and type of sensors differ widely among the studies, from using a single accelerometer [18,51,52] to as many as 12 [53] on different parts of the body. Although gyroscopes can provide valuable rotational information in 3-D, in most studies, accelerometers are preferred to gyroscopes due to their ease of calibration. To the best of our knowledge, guidance on finding a suitable configuration, number, and type of sensors does not exist [45]. Usually, some configuration and some modality of sensors is chosen without strong justification, and empirical results are presented. Processing the acquired signals is also often done ad hoc and with relatively unsophisticated techniques.

In this work, we use miniature inertial sensors and magnetometers positioned on different parts of the body to classify human activities. The motivation behind investigating activity classification is its potential applications in the many different areas mentioned above. The main contribution of this paper is that unlike previous studies, we use many redundant sensors to begin with and extract a variety of features from the sensor signals. Then, we use an unsupervised feature transformation technique that allows considerable feature reduction through automatic selection of the most informative features. We provide an extensive and systematic comparison between various classification techniques used for human activity recognition based on the same data set. We compare the successful differentiation rates, confusion matrices, and computational requirements of the techniques.

The paper is organized as follows: In Section 2, we introduce the activities classified in this study and outline the experimental methodology. Describing the feature vectors and the feature reduction process is the topic of Section 3. In Section 4, we briefly review the classification methods used in this study. In Section 5, we present the experimental results and compare the methods' computational requirements. We also provide a brief discussion on selecting classification techniques and their advantages and disadvantages. Section 6 addresses the potential application areas of miniature inertial sensors in activity recognition. In Section 7, we draw conclusions and provide possible directions for future work.

2. Classified activities and experimental methodology

The 19 activities that are classified using body-worn miniature inertial sensor units are: sitting (A1), standing (A2), lying on back and on right side (A3 and A4), ascending and descending stairs (A5 and A6), standing in an elevator still (A7) and moving around (A8), walking in a parking lot (A9), walking on a treadmill with a speed of 4 km/h (in flat and 15° inclined positions) (A10 and A11), running on a treadmill with a speed of 8 km/h (A12), exercising on a stepper (A13), exercising on a cross trainer (A14), cycling on an exercise bike in horizontal and vertical positions (A15 and A16), rowing (A17), jumping (A18), and playing basketball (A19).

Five MTx 3-DOF orientation trackers (Fig. 1) are used, manufactured by Xsens Technologies [54]. Each MTx unit has a tri-axial accelerometer, a tri-axial gyroscope, and a tri-axial magnetometer, so the sensor units acquire 3-D acceleration, rate of turn, and the strength of Earth's magnetic field. Each motion tracker is programmed via an interface program called MT Manager to capture the raw or calibrated data with a sampling frequency of up to 512 Hz.

Accelerometers of two of the MTx trackers can sense up to $\pm 5g$ and the other three can sense in the range of $\pm 18g$, where $g = 9.80665\ \mathrm{m/s^2}$ is the gravitational constant. All gyroscopes in the MTx unit can sense in the range of $\pm 1200^{\circ}/\mathrm{s}$ angular velocities; magnetometers can sense magnetic fields in the range of $\pm 75\ \mu\mathrm{T}$. We use all three types of sensor data in all three dimensions.

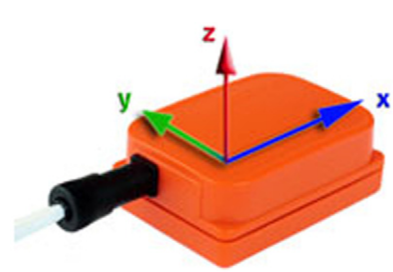

Fig. 1. MTx 3-DOF orientation tracker (reprinted from http://www.xsens.com/en/general/mtx).