RESEARCH PAPERS
FACULTY OF MATERIALS SCIENCE AND TECHNOLOGY IN TRNAVA
SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA

2014 Volume 22, Number 35

COMPARISON OF MACHINE LEARNING METHODS FOR THE PURPOSE OF HUMAN FALL DETECTION

Maximilián STRÉMY, Andrea PETERKOVÁ

doc. Ing. Maximilián Strémy, PhD., Ing. Andrea Peterková, Slovak University of Technology, Faculty of Materials Science and Technology in Trnava, Institute of Applied Informatics, Automation and Mathematics, Hajdóczyho 1, 917 24 Trnava, Slovak Republic
maximilian.stremy@stuba.sk, andrea.peterkova@stuba.sk

Abstract

According to several studies, the European population is rapidly aging far over last years. It is therefore important to ensure that aging population is able to live independently without the support of working-age population. In accordance with the studies, fall is the most dangerous and frequent accident in the everyday life of aging population. In our paper, we present a system to track the human fall by a visual detection, i.e. using no wearable equipment. For this purpose, we used a Kinect sensor, which provides the human body position in the Cartesian coordinates. It is possible to directly capture a human body because the Kinect sensor has a depth and also an infrared camera. The first step in our research was to detect postures and classify the fall accident. We experimented and compared the selected machine learning methods including Naive Bayes, decision trees and SVM method to compare the performance in recognizing the human postures (standing, sitting and lying). The highest classification accuracy of over 93.3% was achieved by the decision tree method.

Key words

fall detection, human fall, comparison of machine learning methods, Kinect camera

Introduction

The European population is rapidly aging far over last years and, according to some studies, it is anticipated to rise from 17.9 % in 2007 to 53.5 % in 2060 (1). To deal with this fact, it is necessary to introduce the appropriate homecare to ensure independent life of aging inhabitants with the minimum support of the working population, and to give them a sense of security, confidence and independence. According to studies (2, 3, 4), fall is the most dangerous accident in the life of aging population, and it can occur in everyday situations. The elderly fall at least once a year (3) and falls are the leading cause of accidental death in persons older than 65 years on the basis of 90 epidemiological studies (4). Thus, human fall detection has become one of the most popular research problems and it is closely linked with the issue of independent life in modern homecare. In our proposed system, we monitored human users to detect the fall using visual detection. The advantage of visual detection in comparison with the non-visual one is no need of any wearable sensors or tags. The user is monitored without the restriction in the daily activities. In this paper, we describe some of the machine learning methods with the properly selected attributes to maximize the accuracy of fall detection. The aim of this paper is to compare the selected machine learning methods for the purpose of human fall detection. The main human postures to recognize are standing, sitting, process of falling and lying.

The paper is structured as follows: Section 2 gives an overview of related works of the fall detection and activity recognition. Section 3 describes the proposed system, the recordings of users' behaviour and positions used as input data. Section 4 lists the appropriate attributes extracted from the input data that are fed into the machine learning algorithms for fall detection. Section 5 presents the experiments in which various machine learning algorithms are compared to raise the best accuracy. Finally, Section 6 concludes the paper and summarizes the results.

Related systems

There are several systems with different technologies and techniques used to detect the fall. Primary systems of fall detection can be divided into two groups depending on how the information is acquired. Such systems are divided into two groups, non-visual and visual systems (5).

Non-visual systems

These systems include the systems using wearable motion detectors, which include an accelerometer and gyroscope, to record rapid changes in the movement of the user. An accelerometer is a device for detecting the magnitude and direction of the acceleration along a single axis or along multiple axes. Typically, three-axis accelerometers are used. By detecting the acceleration caused by the earth's gravity, the accelerometer's angle with respect to the earth can by also computed. The most common and most simple methodology for fall detection using accelerometer is using a tri-axial accelerometer with threshold algorithms (6, 7). Threshold algorithms simply raise the alarm when the threshold value of acceleration is reached. There are several sensors with built-in fall detection hardware (8, 9), having the accuracy of over 80 %. A gyroscope, which measures orientation, consists of a spinning wheel whose axle is free to take any orientation. It can measure orientation along one axis or multiple axes and it is possible to exactly determine the object's orientation and the changes in orientation, from which the angular velocity can be computed.

A common problem of these detectors tends to the fact that a person who has used the device usually forgets or ignores its significance, and therefore the fall is not detected and the device is inactive.

Visual systems

This group of systems includes the systems that use visual analysis. They mostly require the use of one or more cameras. Their main advantage is that there is no need of device attached to a person. These systems are based on detection of human movement using a computer vision algorithms based on the vision techniques extracting input data from still images or from video sequences. Various computer vision techniques use just vision input information and do not reconstruct the human body or posture explicitly.

In our experiments, we chose the approach of the visual detection with posture reconstruction, which is based on 3D locations of an object, typically human body. The approach also uses video recordings and the visual information is used to reconstruct the 3D location. Additional processing uses the 3D parameters as input data in the fall detection methods. If a sufficient number of parameters is provided, it is possible to reconstruct the coordinates of an object, which in our case means the human posture. To receive data from the real life environment, we used a Kinect camera, which is becoming an increasingly common part of visual systems.

There are several papers devoted to using and processing the obtained information from Kinect sensor. Kinect camera was used to detect the fall in finding unique segments (Nyan, Tay 2012) for the precise determination of the fall with six cameras. Experiments confirmed the correlation between thighs and torso segments (10) and these segments were also chosen as the main segments in determining the fall. Another paper (Olivieri, Conde 2012) presents a system called Motion Vector Flow Instance (11), which captures the relevant information about the rate of extraction of dense optical flow from video sequences. Automatic detection is achieved by projecting each video sequence, consisting of about 100 frames in generalized eigenspace (11) and the subsequent application of machine learning to train and recognize activities from video database. Study (Han, Lee 2013) provides an overview of the different positions of the body and its movement and subsequent assessment of the situation. Kinect at the same time pays attention to the actions assessed as dangerous; the authors focus mainly on handling the ladder (12).

Proposed system

As mentioned above, Kinect camera was used to receive data from the real life environment in the proposed visual system. Kinect includes RGB camera, an infrared sensor and also a depth sensor. The monitoring of a human through the depth sensor protects the identity of individual, as the processed data do not reveal any detailed facial features. Finding the appropriate representation of user's activity was the most necessary part of this research because the activity must be represented with a simple and general activity which will not fail in detection of many people. The system set up with the Kinect, captures the user in front of the equipment at the distance between approximately 1.5 - 2.5 meters; beyond this distance, the data became unreliable.

Dataset and parameters

The goal of this experiment was to classify the user's following activities: standing, sitting and lying. Input data used in our analysis were similar to those in our previous study (13). The Kinect provides data in frames at a rate of 30 frames per second. Each frame is processed by

Kinect to obtain joint positions which are subsequently used for the fall detection algorithm to classify the activity of the user. The information about the user is returned as information about users' 20 joint position. From the Kinect SDK for each joint on the skeleton, three pieces of data are acquired: the x, y and z coordinates of the joint in the Cartesian system. Coordinates are given in meters with the Kinect located at the centre. The algorithm works with raw depth data. The training data, including recording of 450 examples of human activities, were captured by the Kinect sensor. Each recording consisted of multiple activities of five persons. Each recording consisted of the four body joints coordinates. The body joints are the body tags representing head, left shoulder, right shoulder and hip coordinates. These tags were the most representive joint for detecting weather human is lying on the floor.

Data processing

Data collected form Kinect sensor was processed as follow. First of all, joints appropriate for fall detection were selected from the raw data. Data was represented in matrix of 450x5. Each row of the matrix represents one training example. Training example consists of the following attributes: joint coordinates and class which represents human action: standing, sitting and lying. The attributes are expressed in the Kartesian coordinate system. Each training example is represented by a feature vector:

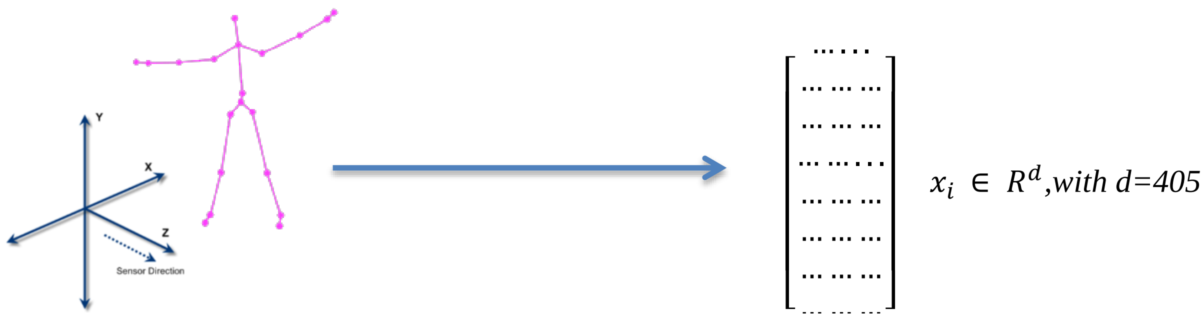

Fig. 1 *Skeletal data representation*

Machine learning algorithms

We have tried various machine learning algorithms for training the classifiers for detecting the activity of the user. We disposed 405 vector attributes. These vectors were used as training data for the following machine learning algorithms: The algorithms were implemented in C#.

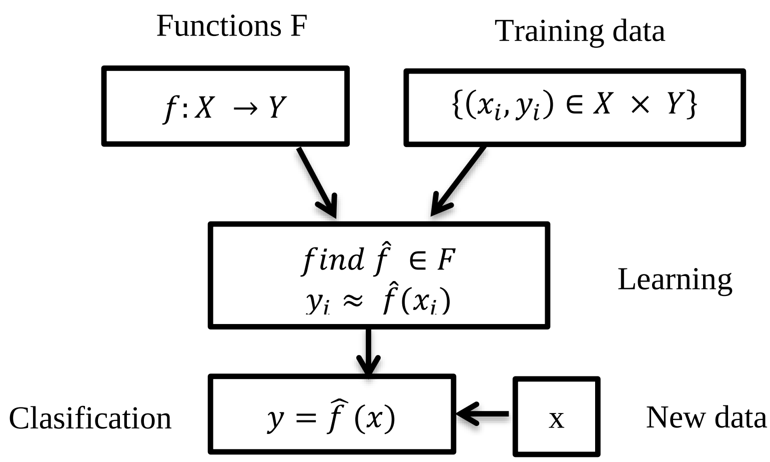

Fig. 2 *Principle of classification algorithms*

SVM

Support vector machine (SVM) is a supervised learning method that can be applied to classification or regression problems. In machine learning, support vector machine is supervised learning model with associated learning algorithms that analyse data and recognize patterns used for classification or regression analysis (14). Given a set of training examples, each is marked as belonging to the category. SVM training algorithm builds a model that assigns new examples into one category.

Naive Bayes

In machine learning, naive Bayes classifiers are family of simple probabilistic classifiers based on applying the Bayes' theorem with strong (naive) independence assumptions between the features (15)

$$P(W|Q) = \frac{P((Q|W)P(W))}{P(Q)} = \frac{P(Q|W)P(W)}{P(Q|W)P(W)+P(Q|M)P(M)} \quad [1]$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad [2]$$

In the simple term, a naive Bayes classifier assumes that the value of a particular feature is unrelated to the presence or absence of any other feature, given the class variable. A naive Bayes classifier considers each of the features to contribute independently to the probability, regardless of the presence or absence of the other features. For some types of probability models, naive Bayes classifiers can be trained very efficiently in a supervised learning setting.

Decision Tree

Decision tree is a simple representation for classifying examples. Decision tree learning is one of the most successful techniques for supervised classification learning. The goal is to create a model that is able to predict the value of a target variable based on several input variables. Attribute selection is the fundamental step to construct a decision tree. The class attribute is referred to as Classifier. Based on joint coordinates (features such as head, left shoulder, right shoulder and hip), we need to decide whether human is sitting, standing or lying, human position class is therefore a classifier to make the decision.

Position algorithm

We also tried the threshold algorithm for better comparison of machine learning methods and the easier way of classifying the postures. We used a position algorithm to detect the person's position based on the joint coordinates.
In Cartesian system, the length of normal vector to the plane ending at a point (x, y, z) can be calculated using the relation (16)

$$d = \frac{Ax + By + Cz + D}{\sqrt{A^2 + B^2 + C^2}}. \qquad [3]$$

Using this relation, the normal distance from the floor to each one joint is obtained. The fall detection algorithm considers the distances for joints which are tracked by the Kinect. If every tracked joint has a normal distance less than the threshold value, the algorithm sets the state to fall, otherwise there is not fall.

Results

In our experiment, we evaluated different algorithms for fall detection using visual systems and machine learning algorithms. The results show that any of these algorithms are efficient for fall detection. Testing data consisted of 120 examples, 40 examples for each position. The highest accuracy was achieved by Decision tree, but the number of positive lying positioned was achieved by position algorithm.

ACCURACY OF FALL DETECTION ALGORITHMS Table 1

RESULTS				
	Number of tests	Precision	Accuracy	Lying position – positive
ALGORITHM				
SVM	120	110/120	91,6 %	34/40
Naive Bayes	120	108/120	90 %	36/40
Decision Tree	120	112/120	93,3 %	36/40
Position algorithm	120	98/120	81,6 %	37/40

Conclusion

In our paper, we compared various machine learning methods. The sensitivity of the fall detection methods is influenced by the accuracy of Kinect equipment. Future research should be focused on the comprehensive solution indepent of the angle of rotation and number of capturing people.

Acknowledgements

This publication is the result of implementation of the project: “UNIVERSITY SCIENTIFIC PARK: CAMPUS MTF STU - CAMBO” (ITMS: 26220220179) supported by the Research & Development Operational Programme funded by the EFRR.

References:

1. Eurostat: Statistic data Eurostat. http://epp.eurostat.ec.europa.eu, 2008-09-09.
2. KIM, K. J., ASHTON-MILLER, J. A. 2003. *Biomechanics of fall arrest using the upper extremity: age differences*. Clin. Biomech. **18**(4), pp. 311–318.
3. DONALD, I., BULPITT, C. 1999. The prognosis of falls in elderly people living at home. *Age Ageing*, **28**(2), 121–5.
4. DEANDREA, S., LUCENTEFORTE, E., BRAVI, F., FOSCHI, R., LA VECCHIA, C., NEGRI, E. 2010. Risk factors for falls in community-dwelling older people: a systematic review and meta-analysis. *Epidemiology*, **21**(5), pp. 658–668.
5. ALWAN M., RAJENDRAN P.J., KELL S., MACK D., DALAL S., WOLFE M., FELDE R. 2006. A smart and passive floor-vibration based fall detector for elderly. In: *Proceedings of ICTTA '06: Information and Communication Technologies*. Damascus, pp. 1003–1007.
6. BOURKE, A. K., SCANAILL, C. N., CULHANE, K. M., O'BRIEN, J. V., and LYONS, G. M. 2006. An optimum accelerometer configuration and simple algorithm for accurately detecting falls. In *Proceedings of the 24th IASTED international Conference on Biomedical Engineering*, pp. 156–160.
7. KANGAS, M., KONTTILA, A., LINDGREN, P., WINBLAD, P., AND JAMSA, T. 2008. Comparison of low-complexity fall detection algorithms for body attached accelerometers. *Gait & Posture*, **28**(2), pp. 285–291.
8. DOUGHTY, K., LEWIS, R., McINTOSH, A. 2000. The design of a practical and reliable fall detector for community and institutional telecare. *Journal of Telemedicine and Telecare*, Vol. 6, pp. 150–154.
9. NOURY, N., BARRALON, P., VIRONE, G., BOISSY, P., HAMEL. M., RUMEAU, P. 2003. A smart sensor based on rules and its evaluation in daily routines. In *Proceedings of the 25th Annual International Conference of the IEEE, Engineering in Medicine Conference of the IEEE, Engineering in Medicine and Biology Society*, Vol. 4, pp. 3286–3289.
10. NYAN, M.N, TAY, F.E.H. 2008. Application of motion analysis system in pre-impact fall detection, 9 Engineering Drive 1, Singapore 117576.
11. OLIVIERI, D., N., CONDE, I.G. 2012. *Eigenspace-based fall detection and activity recognition from motion templates and machine learning* University of Vigo, Spain: Computer Science and Engineering.
12. OLIVIERI, D., N., CONDE, I.G. 2012. Eigenspace-based fall detection and activity recognition from motion templates and machine learning. University of Vigo, Spain, Computer Science and Engineering.
13. HAN S., LEE S. 2013. *A vision-based motion capture and recognition framework for behavior-based safety management.* University of Illinois at Urbana-Champaign and University of Michigan.

14. STRÉMY, M., PETERKOVÁ, A. 2014. Proposed system for human fall detection using Kinect sensor. In: *International Doctoral Seminar 2014*. Zielona Góra.
15. CORTES, C., VAPNIK, V. 1995. Support-vector networks. *Machine Learning*, **20**(3): 273.1 0.1007/BF00994018.
16. RUSSELL Stuart, NORVIG Peter. 2003. (1995). *Artificial Intelligence: A Modern Approach* (2nd ed.). Prentice Hall. ISBN 978-0137903955.
17. WEISSTEIN, E.W. *Point-Plane Distance.* From MathWorld–A Wolfram Web.

Reviewers:

doc. Ing. German Michaľčonok, CSc.
doc. Mgr. Elena Pivarčiová, PhD.