Transactions of the VŠB – Technical University of Ostrava, Mechanical Series
No. 2, 2013, vol. LIX
article No. 1966

Marcel MORE[*], Ondrej LÍŠKA[**]

RECOGNITION OF GESTURES USING ARTIFICAL NEURAL NETWORK

ROZPOZNÁVANIE GEST POMOCOU NEURÓNOVEJ SIETE

Abstract

Sensors for motion measurements are now becoming more widespread. Thanks to their parameters and affordability they are already used not only in the professional sector, but also in devices intended for daily use or entertainment. One of their applications is in control of devices by gestures. Systems that can determine type of gesture from measured motion have many uses. Some are for example in medical practice, but they are still more often used in devices such as cell phones, where they serve as a non-standard form of input. Today there are already several approaches for solving this problem, but building sufficiently reliable system is still a challenging task. In our project we are developing solution based on artificial neural network. In difference to other solutions, this one doesn't require building model for each measuring system and thus it can be used in combination with various sensors just with minimal changes in his structure.

Abstrakt

Senzory pomocou ktorých je možné merať pohyb sú v súčasnosti stále rozšírenejšie. Vďaka ich parametrom a cenovej dostupnosti už nachádzajú využitie nielen v odbornej sfére, ale tiež v zariadeniach určených na zábavu resp. na bežné používanie. Jednou možnosťou ich využitia je ovládanie zariadení pomocou gest. Systémy, ktoré umožňujú z nameraného pohybu určiť o aké gesto sa jedná majú uplatnenie v medicínskej praxi, ale stále častejšie sa používajú aj v zariadeniach ako sú napr. mobilné telefóny, kde slúžia ako neštandardná forma ovládania. V súčasnosti už existujú viaceré prístupy k riešeniu tohto problému, ale navrhnúť spôsob ktorý by bol dostatočne spoľahlivý je stále náročná úloha. V našom projekte využívame na rozpoznávanie pohybu systém, ktorého základom je neurónová sieť. Tento na rozdiel od iných riešení nevyžaduje model konkrétnych meracích prostriedkov a tak je možné ho s minimálnymi úpravami použiť v kombinácii s rôznymi snímačmi.

Keywords

gesture, recognition, MEMS, gyroscope, neural, network

1 INTRODUCTION

Today there are various ways to track and analyze human motions. There are contactless methods based on cameras, that can recognize key points in image and by change in their position they can determinate motion in two dimensional space, or when using more cameras even in third dimension. Another technique is to use sensors that can measure velocity and direction of motion and therefore determine change in position of an object. Such sensors are for example accelerometers and

[*] Ing., Department of Automation, Control and Human Machine Interaction, Faculty of Mechanical Engineering, Technical University of Košice, Letná 9, 04200 Košice, Slovakia, +4210556022588 marcel.more@tuke.sk
[**] doc. Ing, CSc., ondrej.liska@tuke.sk

gyroscopes which today, mainly thanks to MEMS technology, have miniature dimensions and can be therefore used in devices like wristwatches, wearable sensor networks, cell phones, etc. Data measured by these methods can be evaluated in different ways and for different purposes. One of the first applications of motion analysis was in medicine, both for diagnostic purposes, but also in sports medicine to analyze progress in training. Although from the beginning there were attempts to use computers for this task, in some stage there was always need for human intervention. Of course procedure like this was demanding, time consuming and it was not possible to use it in larger scope.

2 PATTERN RECOGNITION IN MOTION ANALYSIS

With progress in information technologies there was also developed new methods of automatic signal processing. These make whole process of motion analysis easier by transferring routine activities from humans to computers. Today there are systems that can process even complex signals from motion sensors and within them they can identify key components characteristic for different kinds of motions. With complete automation of this process, there appear also new possibilities of its use. With still more affordable sensors for motion measurement and availability of more powerful computers, some parts of motion analysis are being used in devices designed for entertainment or for daily use. Such devices are for example cell phones or remote controllers, where gestures recognized in their movement are used to control them (Fig.1,2). Controlling devices by gestures, mostly by hand, instead of standard input devices have many advantages. Usage of gestures is for many people more natural and requires less fine motor skills than usage of keyboards or touch screens. For some people, this is actually the only way to operate various devices.

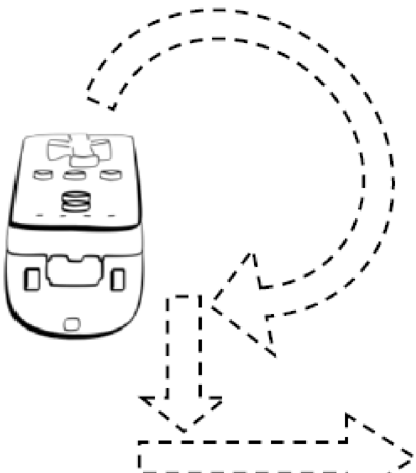


Fig. 1 Concept of gesture controlled TV remote [1]

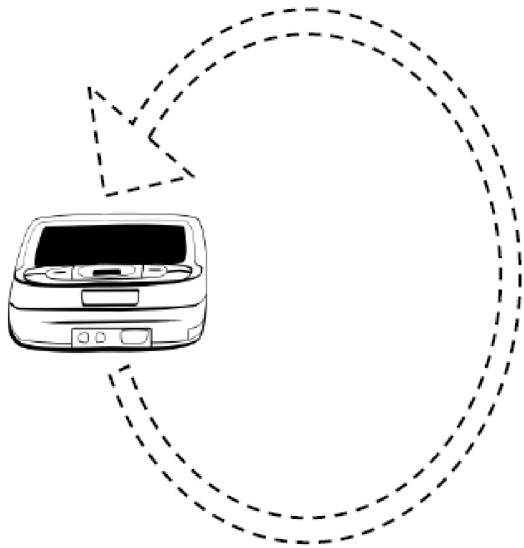


Fig. 2 Concept of cell phone controlled by gestures [1]

Although higher power of today's computers allows us to use in analysis more complex algorithms than before, recognition of match between measured signals and reference model remains still a nontrivial task. It is mainly because of fact, that even if we try to repeat the same movement identically, measured signal will still have larger or smaller deviations compared to previous measurements. Therefore, when designing system for recognition, it is necessary to consider that processed signal can be time-shifted, dilated or it can have some other form of deformation. This effect is even worse when there is more time elapsed between measurements or when the same movement is performed by multiple subjects. This means that the recognition system must be sensitive enough to discern a signal slightly different from the reference, but also robust enough to not return false positive results. Currently, when it comes to pattern recognition, usual choice is to use Hidden Markov Model (HMM) [2], [3]. It is being used in motion recognition as well as in speech and image processing. But this method is complicated because there is need to assemble complex model for each sensor used. This is why we decided to use another approach and that is pattern recognition using neural network.

3 USING NEURAL NETWORK FOR PATTERN RECOGNITION

In our project we use inertial measurement unit (IMU) consisting of 3-axis analog accelerometer MMA7431 from FreeScale and 3-axis digital gyroscope ITG3200 from InvenSense (Fig.3,4). Data from these sensors are acquired by microcontroller unit Atmel ATmega32L and send to computer at speed of 100 kilobauds. With this assembly we can track movement in all six degrees of freedom with sampling frequency much higher than it is necessary to study human motions. Although accelerometers and gyroscopes are principally designed to measure different components of motion, most complex gestures generates waveforms in acceleration and angular velocity domain unique enough to identify them. That is why there is not necessary to use data from both sensors. This paper describes usage of same system to process data from sensors which are usually not interchangeable.

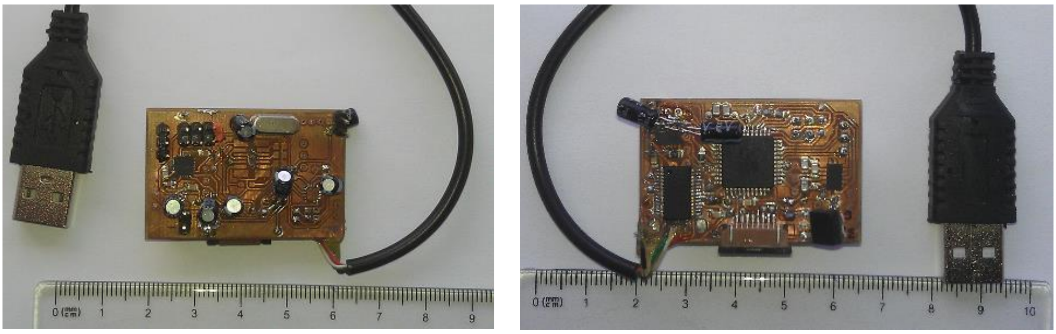
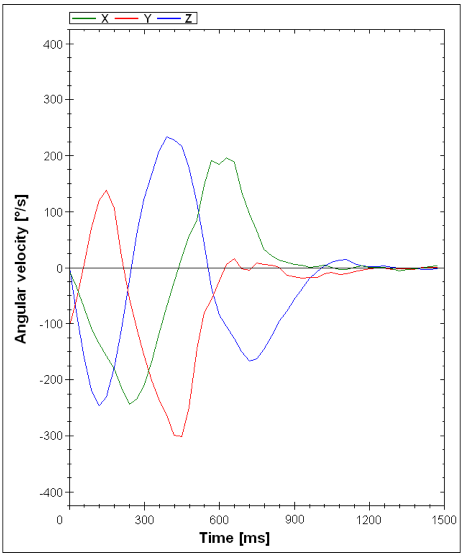


Fig. 3, Fig. 4 Inertial measurement unit used for testing [4]

3.1 Function of recognition system

IMU don’t process data, it just transferred them to computer. Before they enter neural network, they go through very little processing. First, data from both sensors are cleared from offset, which is difference between measured value and zero in state, when the device is not moving. Besides that, data from accelerometer are run through low-pass filter, because signal from it is fairly noisy. This data are continually monitored and program waits for one of the values to exceed threshold values of acceleration or angular velocity, which could mean that there is a gesture in progress. After system detects beginning of motion, measured data are recorded until they fill up given frame. Length of the frame and sampling frequency is adjustable, but for backward compatibility it is convenient to choose and use same values. Determination of right frame length is not critical, but it has significant impact on overall system performance. If the frame is too short, it cannot span over whole gesture and therefore recognition capability is limited. But if it is too long, it contains lot of unnecessary data, which are not important for functionality, but slow down whole algorithm. Therefore right frame length should be determined based on the expected inputs.



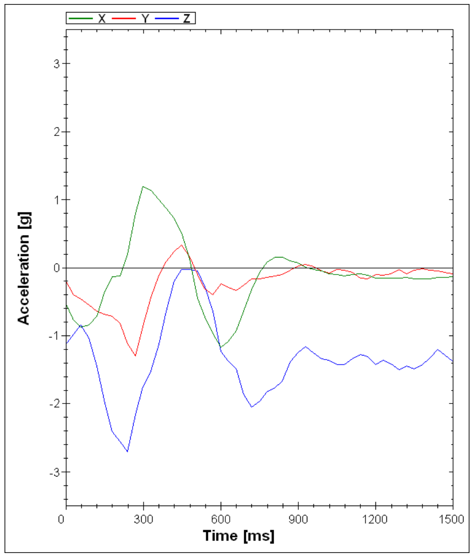


Fig. 5, Fig. 6 Gesture waveform measured by gyroscope (left) and accelerometer (right)

Frame processed like this is input to the neural network, where the number of neurons in the input layer is equal to the number of samples in one frame. Because we use data from all three axes, there is one neural network for each of them and final result is based on mean value of these three partial results. This solution has proved to be more reliable over the use of one neural network for all data. But this way, when we try to use gesture that has not significant waveform in some dimension (like X axis on Fig.7), neural network assign to this axis will be unable to learn properly and will always return uncertain result. But this is sufficiently compensated by averaging this result with other two certain results (Fig.7). Important role in this is also right setup of threshold, which is the value by which the final result is considered positive.

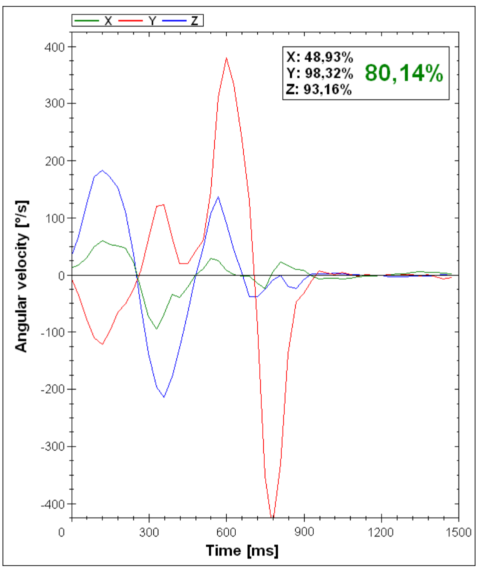


Fig. 7 Final result based on three partial results

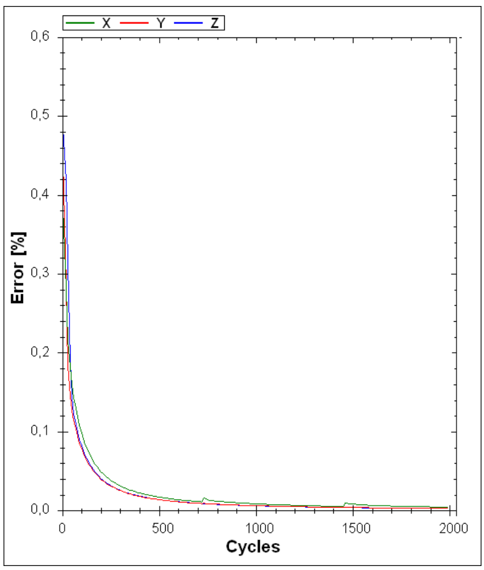


Fig. 8 Change of error during learning cycles

3.2 Training of neural network

In our project we use Backpropagation method to train neural networks. This is a form of supervised training in which the neural network tries to learn the relationship between inputs and outputs based on a given examples. When preparing training sets for the network, procedure of data processing must be same as when we use it. Gesture which we want to be recognized is entered in form of several samples. These samples must have some level of diversity, so that trained network will eventually be sufficiently sensitive. When building training sets, it is preferable to take measurements with more than one subject and also not all of them in one time. This helps to reach required level of variance. To achieve correct results, it is also important to assign negative examples. These are motions and gestures, which we don't want to be recognized. If we know in advance what group of gestures will be used, we can set all other gestures from the group as negative examples and therefore increase reliability of recognition. Selection of appropriate training sets and proper selection of parameters neural network is crucial for right function of whole system.

Training of neural network is pretty time consuming, require lot of computing power and its getting worst with increasing number of training sets, longer frames, etc. But on the other side, use of trained network is pretty quick and saved neural network requires only few kilobytes of memory. Because of this, there isn't problem to build up libraries containing more trained networks and use them as needed.

3.3 Verification of accuracy

As it was mentioned above, correctly working system for recognition must be able to recognize gestures with slight differences, but at the same time must it be solid enough so that it don't return false-positive results. Accuracy of recognition depends on many factors. Besides right setup of the network it is also dependant on types and number of gestures used. To verify attributes of our system, we setup simple experiment. For every variant of input we conducted series of 150 measurements and their results are in Tab 1.

Tab. 1 Accuracy of recognition using different inputs

Sensor	Signal processing	Accuracy	Comment
gyroscope	internal low-pass	**97,3%**	
accelerometer	low-pass	**94,6%**	
accelerometer	none	**78,6%**	
gyroscope	internal low-pass	**0%**	false-positive results test

As it can be seen in this table, best results come from gyroscope. It is mainly because data coming from accelerometer are, regardless filtering, affected by noise. This is noticeably worst when there is no filter used. Last test was aimed to find out if our system generates any false-positive results. As an input in this test were used gestures evidently different from reference and goal was to find out how many times system return positive result. From the 150 measurements, none of the results come back positive, which means that well trained neural network has potential to serve as system for pattern recognition in motion analysis.

4 CONCLUSIONS

Application of MEMS sensor for motion measurement today is still wider. Thanks to their affordability and useful characteristics, they find their usage in many devices. One of the ways they can be used in modern gadgets is for gesture control. Today there are systems that can recognize patterns in waveforms generated by motion sensors and base on them they can identify different gestures. But it is still difficult task to build such an algorithm and it involves several complications. In our project we decided to take less conventional approach and we use artificial neural network instead of standard algorithm. We've put together system that, thanks to the specific characteristics of neural networks can work with different kinds of sensors without need for significant changes in his structure. To verify his properties, we conducted simple experiment, which showed that well trained neural network is capable to reliably recognize simple gestures. A further development of this system should focus on recognition of complex gestures and also sequential movements. In the future we would also like to explore the possibilities of its use in portable devices, in which recognition of gestures has most significant use.

ACKNOWLEDGMENT

The research work is supported by the Project of the Structural Funds of the EU, Operational Programme Research and Development, Measure 2.2 Transfer of knowledge and technology from research and development into practice. Title of the project: „Research and development of intelligent nonconventional actuators based on artificial muscles”, ITMS code: 26220220103.

REFERENCES

[1] BORZA, P. Motion-based Gesture Recognition with an Accelerometer : Bachelor’s Thesis [online] Babes-Bolyai University of Cluj-Napoca, Romania, 2008 [cit. 2013-04-23]. Available from www: <URL:accelges.googlecode.com/files/ThesisPaper.pdf>

[2] BENBASAT, A., PARADISO, J. An Inertial Measurement Framework for Gesture Recognition and Applications [online].[cit. 2013-04-23]. Available from www: <URL:http://pdf.aminer.org/000/238/224/an_inertial_measurement_framework_for_gesture_recognition_and_applications.pdf>

[3] HIESEY, J., MELLINA, C., ZAVAIN, D. Incorporating Weighted Clustering in 3D Gesture Recognition [online] Standford, 2011,. [cit. 2013-04-23]. Available from www: <URL:http://cs229.stanford.edu/proj2011/HieseyMellinaDar-IncorporatingWeightedClusteringIn3DGestureRecognition.pdf>

[4] ŠIMŠÍK, D. et al. Design of inertial module for rehabilitation device . In *SAMI 2013*. Herľany SR : Budapest IEEE, 2013, ISBN 978-1-4673-5926-9.

[5] ŽIDEK, K., ŽUPA, T., FERIANČIK, M. Využitie MEMS snímačov pre meranie nebezpečného náklonu mobilných zariadení. In *Automatizácia a riadenie v teórii a praxi ARTEP 2010*. Stará Lesná SR : Košice TU, 2010, ISBN 978-80-553-0347-5.

[6] SHI, G. et al. Towards HMM based Human Motion Recognition using MEMS Inertial Sensors . In *Proceedings of the 2008 IEEE International Conference on Robotics and Biomimetics* [online] Bangkok Thailand, 2008,[cit. 2013-04-23]. Available from www: <URL:http://edge.rit.edu/content/P10010/public/PDF/getPDF2.pdf>

[7] ŠIMŠÍK, D. et al. Trendy vo využívaní MEMS snímačov. In *Automatizácia a riadenie v teórii a praxi ARTEP 2013*. Stará Lesná SR : Košice TU, 2013, ISBN 978-80-553-1330-6.

[8] ŽIDEK, K., HOŠOVSKÝ, A., MAXIM, V. Real-time safety circuit based on combined MEMS sensor data for rehabilitation device. In ICCC 2012. Podbanské SR : Piscataway IEEE, 2012, ISBN ISBN 978-1-4577-1866-3.