ANNALS of Faculty Engineering Hunedoara – International Journal of Engineering
Tome XIV [2016] – Fascicule 2 [May]
ISSN: 1584-2665 [print; online]
ISSN: 1584-2673 [CD-Rom; online]
a free-accessmultidisciplinarypublication
of theFaculty of Engineering Hunedoara

[1.]Ivan ĆIRIĆ, [2.]Žarko ĆOJBAŠIĆ, [3.]Vlastimir NIKOLIĆ, [4.]Milica ĆIRIĆ, [5.]Mladen TOMIĆ, [6.]Emina PETROVIĆ, [7.]Miloš SIMONOVIĆ

NEURAL NETWORK PREDICTION OF PERSON POSITION FOR HUMAN FOLLOWING MOBILE ROBOT PLATFORM

[1-3,6-7.]University of Niš, Faculty of Mechanical Engineering, Niš, SERBIA
[4.]University of Niš, Faculty of Civil Engineering and Architecture, Niš, SERBIA
[5.] The School of Higher Technical Professional Education Niš, Niš, SERBIA

ABSTACT: This paper addresses an important and challenging problem of real-time vision-based human tracking to enable mobile robots to follow a human and help him as a co-worker in hazardous environment. A novel approach that combines intelligent stereo vision-based human detection with neural network estimator that helps tracking is presented. Stereo vision-based detection combines features extracted from 2D stereo images with intelligent classification algorithms to detect humans in a robot's environment. Prediction of the 3D coordinates of a human in the robot's camera coordinate system based on recurrent neural network reduces image region of interest and ensures more reliable human tracking. Within a working scenario of a mobile robot intended to follow a human co-worker in indoor applications, collected video data will be collected and compared with simulation results obtained from the recurrent neural network.
Keywords: robot vision, mobile robotics, service robotics, neural networks; human tracking

1. INTRODUCTION

In the next few years, personal service robots are expected to become part of our everyday life, playing an important role as our appliances, servants and assistants. The future of smart homes points clearly towards the ambient intelligence paradigm. We expect to build an intelligent environment that discovers and adapts itself automatically to the user's needs. In this environment, service robots are completely integrated in the home and it is easy to imagine scenarios in which robots and smart home systems cooperate. Unfortunately, the road toward ambient intelligence is full of obstacles [1,2].

As research robots move closer towards real applications in populated environments, these systems will require robust algorithms for tracking people to ensure safe human–robot cohabitation. These robots will need to be able to track human motion to avoid colliding with them, to acquire a sufficient understanding of the environment, to be aware of different situations, to detect and track people with minimum instruction and with high quality and precision and also to be able to make decisions based on their perception of human state in order to become functional collaborators [1,2,3].

Most vision-based people recognition systems concern nonmobile applications e.g., surveillance or identity verification systems, where detection of persons can be solved easily by background subtraction methods that cannot be applied to mobile systems. Existing vision-based people recognition systems on mobile robots typically use skin color and face detection algorithms. However these approaches assume the frontal pose of a person and require that the person is not too far from the robot.

In this work one application of neural networks for human tracking is presented. First, we described the NAR neural network and it's advantages in time series prediction and then we have compared the results of human tracking obtained from the NN with results previously obtained from the Kalman filter.

2. NONLINEAR AUTOREGRESIVE NEURAL NETWORK (NAR)

According to capability to learn about its environment through an iterative process of adjustments applied to its synaptic weights and thresholds [4], nonlinear autoregressive neural network - NAR is the most straightforward network in time series prediction. Nonlinear autoregressive - NAR method is derived from the linearly autoregressive - AR method to solve many real life applications, in which many nonlinearities usually appear. The input-output connection of a linearly AR model is given by the following equation:

$$r(k) = \sum_{i=1}^{p} a(i) r(k-i) + e(k) \qquad (1)$$

where r(k) is the input signal and e(k) denotes the independent, identically distributed zero that means an error signal often termed the "driving noise". The a(i) represents model parameters and p denotes the order of the model. According to the Yule-Walker equations [5], which are used to estimate the parameters of a(i), the optimal value of a(i) is obtained by minimizing the independent identically distributed error signal. Observed from the (1) the optimal values of the parameters a(i) are considered constant for all samples of signal r(k). However, this can be used only for stationary signals. Non-stationary signals change their parameters over the time in which many nonlinearities usually appear. To cross this barrier, the linear input-output relation can be expanded to a nonlinear one, but the difficulty in this case is the estimation of the model parameters. From that point the NAR model takes the role. As in the linear model, the nonlinear autoregressive model of order p will be provided by the following input-output equation:

$$r(k) = g(r(k-1),\ldots,r(k-p)) + e(n) \qquad (2)$$

where e(n) corresponds to the independent identically distributed error, and g(x) is an unknown non-linear function. This function describes the non-linear relations of the current signal and the previous signal samples p [6].

3. EXPERIMENTAL SETUP AND NETWORK TRAINING

The research presented in this paper is a part of a greater research conducted in Niš at Faculty of Mechanical Engineering and in Bremen at Institute for Automation as a part of Serbian-German bilateral scientiffic project entitled "Robust Action Recognition for Human-Robot Synergy" and joint research since 2009. An illustration of the visually controlled robotic system for human tracking used in this research is shown in Figure 1.

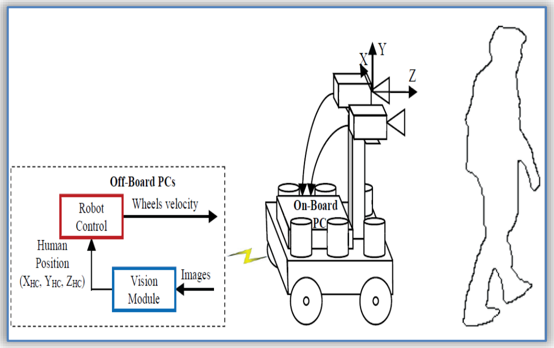

Figure 1. Principal layout of the robotic system for following the human co-worker

The used robotic system [7,8] has a Point Grey Bumblebee XB3 stereo camera system on board, as well as a Nexcom NISE 3500 P2 low-power industrial PC (Intel i7-620M @ 2.66 GHz, 4GB of RAM, 64 GB SSD, ~65 Watt). As the platform is intended to run on batteries in order to move independently of external power supplies, low power consumption of the on-board components is critical. Therefore, the on-board PC functionality is limited to capturing images from the stereo camera and sending them compressed over the wireless link to the off-board vision module as well as sending direct commands obtained from the robot control to the wheels controllers. In order to ensure that the control commands are sent at regular time intervals, the vision module and the robot control module run on different PCs as illustrated in Figure 1 with differently colored blocks. The computational expensive vision algorithms run on a dedicated high-end PC (Intel Xeon E5520 @ 2.26 GHz, 6GB of RAM, NVidia Tesla C1060 GPU) with the goal of tracking the human co-worker in front of the robot and determining its 3D position with respect to the coordinate system of the left stereo camera, as depicted in Figure 1.

After obtaining the 3D position, it is sent to the robot control module located on a separate desktop PC (Intel E4700 @ 2.6 GHz, 2GB of RAM), which computes the required velocities for each wheel such that the robot keeps following the human. The velocities are sent at regular time intervals over the wireless link to the on-board PC, which sends direct commands to the wheels controllers. If the robot control module would run on the same PC as the vision module, it might happen that it would not be able to send the new velocities out in time due to the vision module blocking the CPU.

The communication between the three computers is done via Robot Operative System (ROS) [7,8], which is a widely used communication framework that, among other features, allows easily configuring multiple computers for cooperation in order to improve the time effectiveness of the system. This time effectiveness is achieved by splitting computationally expensive tasks into modules, which run on different computers.

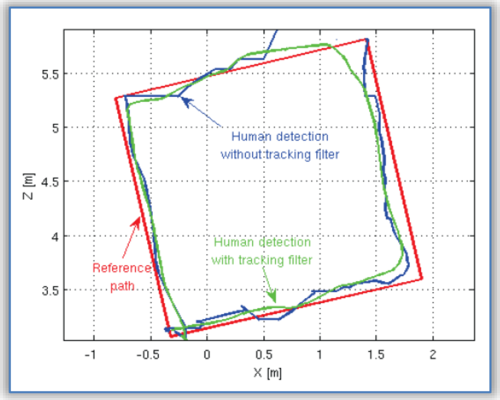

Figure 2.Comparison of the reference human's path and the human's path reconstructed with the stereo-vision based tracker with and without tracking filter

In order to acquiretraining data and later to validate recurrent network, indoor human motion was captured by presented mobile robotic system, and comparison between neural network estimator and Kalman filter is done. However results and measurements obtained during previous research [7,8] shown in Figure 2 were used and there was no online testing of the neural network estimator. The frame rate of the used camera is 15 frames per second (fps) and each pair of stereo frames was processed so to extract information for stereo vision based reconstruction of human walking with respect to camera coordinate system. The state of the person is presented in Cartesian coordinates X, Y, Z and three dimensional velocity.

The purpose of trainingnetwork is to forecast the human trajectory and its X,Y and Z coordinates in real time. The developed estimator is based on feed forward topology of the NAR NN with Levenberg-Marquardt method of training. The NN had ten hidden neurons and two delays as presented in Figure3.

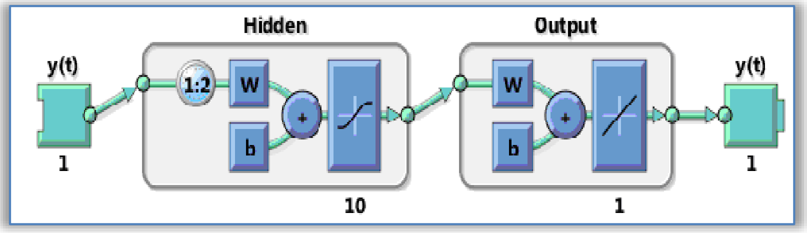

Figure 3. Used NAR NN topology with weights and biases at each node

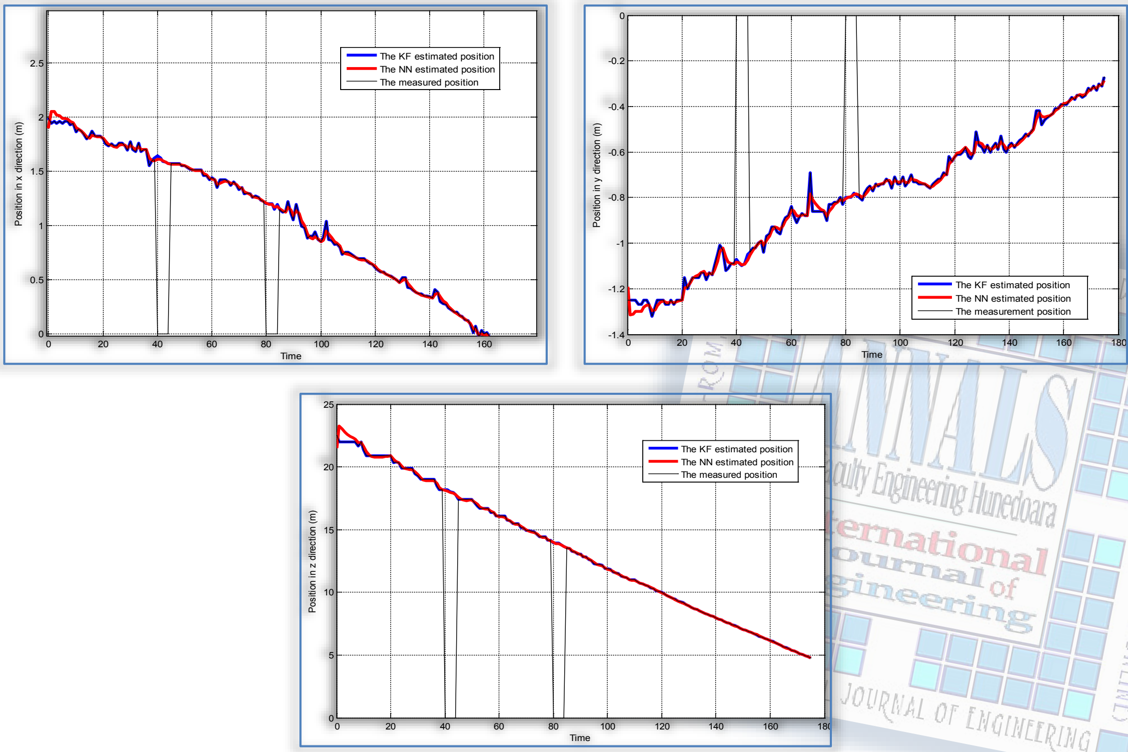

Figure 4. Measured and estimated X,Y and Z coordinate of the tracked human

Figure 4 presents estimated and measured X, Y, and Z coordinates of the tracked human with Kalman filter and NAR NN (output vector y(t) in network topology) respectively.

After reviewing these results from the figures below we can see that the both estimators made prediction accurately and both approaches were able to reduce noise in measurements as well as to predict state when the robot's vision system did not give measurements in some intervals of time. But the main disadvantage of Kalman filter is that we must provide its dynamic model before tracking.

4. CONCLUSION

In this work we have demonstrated the performance of tracking system using the results obtained from the NAR Neural network and based on the video sequences and compared with the results obtained from the Kalman filter. Experimental results showd us that NAR Neural networks give good results, NN made predictions accurately and reliably as well as the Kalman filter. On the other hand NN can work directly with raw inputs and data. Using that fact and characteristic of the NN that they are able to learn the dynamic model in real time, NN can be easily implemented in human tracking algorithm in order to have more robust and reliable tracking result.

Note: This paper is based on the paper presented at The 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology – DEMI 2015, organized by the University of Banja Luka, Faculty of Mechanical Engineering and Faculty of Electrical Engineering, in Banja Luka, BOSNIA & HERZEGOVINA (29th – 30th of May, 2015), referred here as [10].

REFERENCES

[1] Alvarez-Santos, V., Pardo, X.M., Iglesias, R., Canedo-Rodriguez, A., Regueiro, C.V., Feature analysis for human recognition and discrimination: Application to a person-following behaviour in a mobile robot, Robotics and Autonomous Systems, 60 (2012), 8, pp. 1021-1036

[2] Böhme, H.-J., Wilhelm, T., Key, J., Schauer, C., Schröter, C., Groß, H.-M., Hempel, T., An approach to multi-modal human–machine interaction for intelligent service robots, Robotics and Autonomous Systems, 44 (2003), 1, pp. 83-96

[3] Ćirić, I., Ćojbašić, Ž., Nikolić, V., Antić, D., Computationally Intelligent System for Thermal Vision People Detection and Tracking in Robotic Applications, Proceedings, 11th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services - TELSIKS 2013, Niš, Serbia, 2013, pp 587-590

[4] H. M. El-Bakry, N. Mastorakis, "Fast Human Motion Tracking by using High Speed Neural Networks", Recent Advances in Signals and Systems ISSN: 1790-5109, 2009.

[5] N. Doulamis, A. Doulamis, T. Varvarigou, "Adaptable Neural Networks for Modeling Recursive Non-linear Systems", 2002.

[6] A. Yeredor, "Yule-Walker Equations Applied to Hessians of the Characteristic Function for Improved AR Estimation", ICASSP 2007.

[7] Emina Petrović, Adrian Leu, Danijela Ristić-Durrant, Vlastimir Nikolić. Stereo Vision-Based Human Tracking for Robotic Follower. Int J Adv Robot Syst, 2013, 10:230. doi: 10.5772/56124

[8] A. Leu, D. Aitenau, A. Graser, "A novel stereo camera based collision warning system for automotive applications" 6th International Symposium on Applied Computational Intelligence and Informatics, May 2011.

[9] V. Nikolić, Ž. Ćojbašić, D. Ristić-Durrant, E. Petrović, S. Matić, I. Ćirić, "Kalman Filter for Robot Vision-Based Human Tracking" 29th Danubia-Adria Symposium, Advances in Experimental Mechanics, Belgrade, Serbia, September 2012.

[10] I. Ćirić, Ž. Ćojbašić, V. Nikolić, M. Ćirić, M. Tomić, E. Petrović, M. Simonović, Neural network prediction of person position for human following mobile robot platform, The 12th International Conference on Accomplishments in Electrical and Mechanical Engineering and Information Technology – DEMI 2015, Bosnia & Herzegovina