THERMAL VISION BASED INTELLIGENT SYSTEM FOR HUMAN DETECTION AND TRACKING IN MOBILE ROBOT CONTROL SYSTEM

by

Ivan T. ĆIRIĆ[a]*, Žarko M. ĆOJBAŠIĆ[a], Danijela D. RISTIĆ-DURRANT[b], Vlastimir D. NIKOLIĆ[a], Milica V. ĆIRIĆ[c], Miloš B. SIMONOVIĆ[a], and Ivan R. PAVLOVIĆ[a]

[a] Faculty of Mechanical Engineering, University of Nis, Nis, Serbia
[b] Institute of Automation, University of Bremen, Bremen, Germany
[c] Faculty of Civil Engineering and Architecture, University of Nis, Nis, Serbia

Original scientific paper
DOI: 10.2298/TSCI16S5553Ć

This paper presents the results of the authors in thermal vision based mobile robot control. The most important segment of the high level control loop of mobile robot platform is an intelligent real-time algorithm for human detection and tracking. Temperature variations across same objects, air flow with different temperature gradients, reflections, person overlap while crossing each other, and many other non-linearities, uncertainty and noise, put challenges in thermal image processing and therefore the need of computationally intelligent algorithms for obtaining the efficient performance from human motion tracking system. The main goal was to enable mobile robot platform or any technical system to recognize the person in indoor environment, localize it and track it with accuracy high enough to allow adequate human-machine interaction. The developed computationally intelligent algorithms enables robust and reliable human detection and tracking based on neural network classifier and autoregressive neural network for time series prediction. Intelligent algorithm used for thermal image segmentation gives accurate inputs for classification.

Key words: *thermal vision, mobile robot platform control, thermal image processing, neural network classification, time series prediction*

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 environ-

* Corresponding author; ciric.ivan@masfak.ni.ac.rs

ment, 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 [1-3]. Most vision-based people recognition systems concern non-mobile 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. Thermal vision helps to overcome some of the problems related to color vision sensors since humans have a distinctive thermal profile compared to non-living objects.

There are many algorithms focusing specifically on the thermal domain for human tracking in robotics. The thermal-based algorithms are inspired in the biological processes of many animals and insects, which are affected by the presence of thermal energy in their environment. Indeed, diverse types of thermo-receptors are found in nature, which aid animals and insects in hunting, feeding and survival. Now, in computer vision, the unifying assumption in most methods is the belief that the objects of interest are warmer than their surroundings. Indeed, some animals can see in total darkness, or even see colors beyond the visual spectrum, that humans have never seen. Thermal infrared video cameras detect relative differences in the amount of thermal energy emitted/reflected from objects in the scene. As long as the thermal properties of a foreground object are slightly different (higher or lower) from the background radiation, the corresponding region in a thermal image appears at a contrast from the environment [4-6].

The system we have developed to solve the person-following task uses computational intelligence for solving different problems of robot vision. The robot will use information provided by a thermal camera to detect human and to track him. The information about position of the person in the scenario will then be sent to a motion controller that controls motors of the mobile robot platform in order to follow target. The thermal vision system will detect the presence of people around the robot by using human detection and tracking algorithm developed in this paper. The developed algorithm has three main parts, image segmentation part, classification part and tracking part. Previous research [5-7] has shown that full potential of intelligent methodologies used for classification of ROI in robot vision can be achieved only if reliable and robust segmentation is done. Variation in temperature across the same objects, air flow with different temperature gradients, person overlap while crossing each other and reflections, put challenges in thermal imaging segmentation. Therefore these problems need to be handled intelligently in order to obtain the efficient performance from motion tracking system.

After segmentation is done neural network classifier determines whether segmented ROI is human or not based on features extracted from binary segmented image and centre of mass is calculated. Aside that, in this paper application of neural networks for human tracking is presented. Vision module of mobile robot consists of segmentation, classification and tracking part and it provides the co-ordinates of a person in scene with respect to the forward direction of the robot so that robot platform can unobtainably track a person.

Mobile robot platform

For development of reliable mobile robot platform that is capable of tracking people based on thermal image National Instruments Robotics Starter Kit 1.0 known as DaNI robot has been combined with FLIR E50 thermal camera [5, 6]. Thermal camera could not be directly connected to on-board controller (single board RIO) of DaNI Robot, nor could the complex segmentation and classification algorithms be implemented directly to low level control algorithm, so additional off-board personal computer was used for thermal vision image processing and high level control. Functional scheme of the thermal vision-based person-following robot platform is shown in fig. 1.

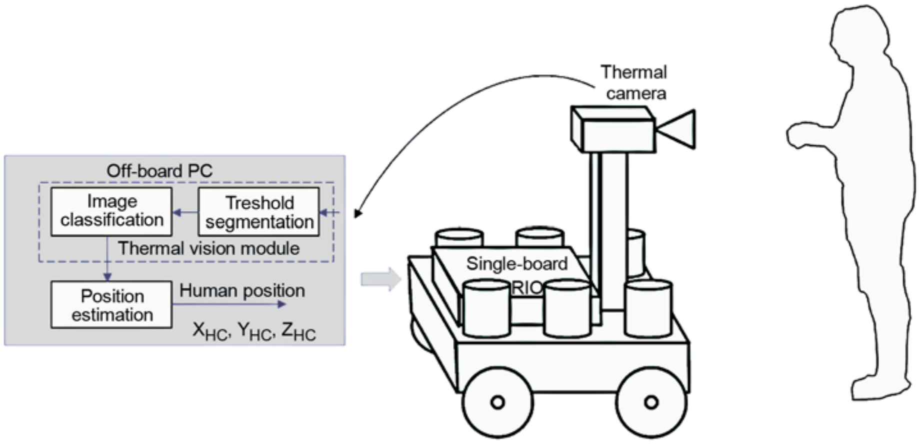

Figure 1. Principle layout of thermal vision-based person-following robot platform

Acquired thermal image is segmented and segmented objects are then classified in thermal vision module in order to detect human. Segmentation and classification algorithms are two main parts of thermal vision module, a part of high level control system of mobile robot platform. After person was detected, simple estimation algorithm that uses average human height as a reference estimates relative position for the following task. Co-ordinates are then sent to low level controller that uses conventional PID controllers for wheel velocities.

Thermal image processing

Typical image processing module consists of four main parts, image acquisition, image segmentation, image understanding, and application specific feature extraction (fig. 2). The vision module developed in this paper consists of the same four parts, but each part is specific since we were using thermal image. Image segmentation aside main image segmentation algorithm has image pre-processing and post-processing of segmented image. Image understanding is done in order to detect human in the scene and human co-ordinates can be extracted. The core of the image understanding is neural network classifier that detects human based on extracted features from segmented objects.

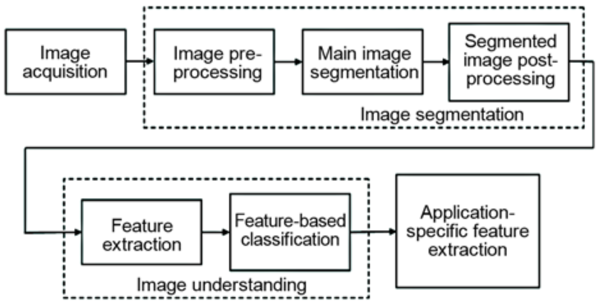

Figure 2. Computer vision block-diagram

Thermographic cameras detect radiation in the infrared range of the electromagnetic spectrum (roughly 9-14 μm) and produce images of that radiation, called thermograms. Since infrared radiation is emitted by all objects above absolute zero according to the black body radiation law, thermography makes it possible to see one's environment with or without visible illumination. The amount of radiation emitted by an object increases with temperature; therefore, thermography allows one to see variations in temperature. When viewed through a thermal imaging camera, warm objects stand out well against cooler backgrounds; humans and other warm-blooded animals become easily visible against the environment, day or night.

The robot was operated in an unconstrained indoor environment (a corridor and a hallway at our Faculty in Nis). Persons taking part in the experiments were asked to walk in front of the robot while the robot was in stationary position and while it performed person following (using information from the implemented tracker). In our experiments the visible range on the image was equivalent to the temperature range 17.3-38.2 °C and one of predefined six color palettes has been chosen. The thermal camera can detect infrared radiation and convert this information into an image where each pixel corresponds to a temperature value. The blue color corresponds to low temperature and high temperature corresponds to yellow color that is a combination of red and green color in RGB image [5, 6], see fig. 3.

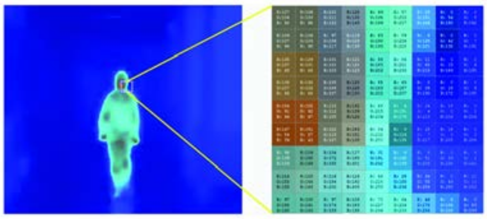

Figure 3. Thermal image and pixel values for selected area *(for color image see journal web-site)*

The segmentation algorithm that was used in this research [5, 6] was based on thresholding method and the problem was boiled down to determining possible region of interests in real world scenarios, where there are obstacles in front of the person, two persons are overlapping when moving close to one another and reflexion is possible to occur. For adequate segmentation algorithm also high noise that occurs in thermal image must be taken into consideration.

Since the problem of segmentation of thermal image needs involving some intelligent and heuristic optimization algorithm the authors have decided to implement genetic algorithm (GA) for this purpose [5, 6]. The segmentation results are shown in fig. 4.

Figure 4. (a) thermal image, (b) segmentation results with manualy determined threshold, (c) fuzzy segmentation results, (d) genetic algorithm optimized threshold *(for color image see journal web-site)*

Neural network classification

Feature-based classification (fig. 5) is most common for image understanding. From segmented objects we can extract many features, like height, width, colour, or same shape descriptors, that can help us in classification.

For human detection in this research neural network classifier is developed, and its architecture is shown in fig. 6. It is a two layer feed forward neural network where neurons in

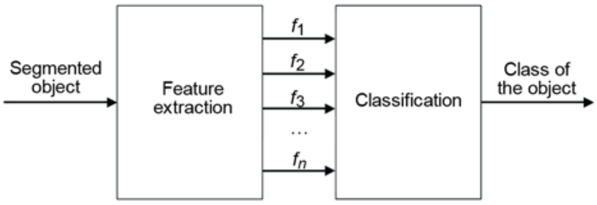

Figure 5. Block-diagram of feature-based classification

hidden layer have sigmoid activation function, while in output layer neurons have *softmax* function. Based on authors expert knowledge, 5 parameters are used as classifier inputs, object height, *h*, width, *w*, and first 3 Hu moments of inertia [8] as defined by eq. (1).

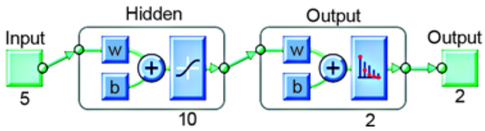

Figure 6. Neural network classifier for human detection

$$I_1 = \eta_{02} + \eta_{20}$$

$$I_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}{}^2$$

$$I_3 = (\eta_{30} - \eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \qquad (1)$$

Training, testing and validation is done with more than 3000 manually labeled objects, where 70% randomly selected data sets were used for training, 15% randomly selected data sets were used for testing, and 15% randomly selected data sets were used for validation. For training, the back propagation scaled conjugate gradient algorithm that updates weight and bias states according to Levenberg-Marquardt optimization was used, while the mean squared error was used as a performance measure during training. Training, testing and validation results are shown in fig. 7 via confusion matrices.

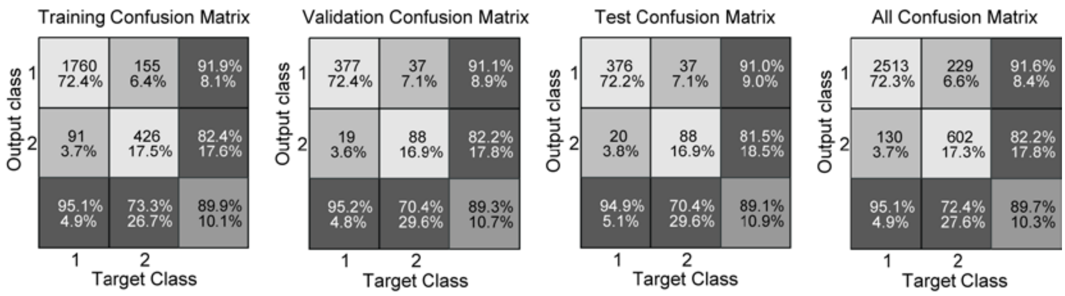

Figure 7. Neural network classification results

Neural network training was finished after 90 iterations and it shows good classification results with accuracy higher than 89%. It is possible to achieve higher accuracy by inserting additional neurons into hidden layer, or by advanced training set selection. Classificator output is probability that analized object whose features we have extracted belongs to a class named *human* and *non-human*.

Neural network for time series prediction

According to capability to learn about its environment through an iterative process of adjustments applied to its synaptic weights and thresholds, non-linear autoregressive neural network (NAR) is the most straightforward network in time series prediction [9, 10]. The NAR method is derived from the linearly autoregressive (AR) method to solve many real life applications, in which many non-linearities usually appear.

In order to acquire training data and later to validate recurrent network, indoor human motion was captured by presented mobile robotic system. The frame rate of the used camera is 60 frames per second (fps) and each thermal video frame was processed in order to extract information for reconstruction of human motion with respect to camera co-ordinate system. The state of the person is presented in Cartesian co-ordinates X, Y, Z, and three dimensional velocity. The purpose of training network is to forecast the human trajectory and its X, Y, and Z co-ordinates 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 fig. 8.

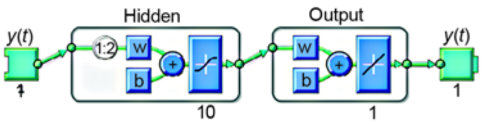

Figure 8. NAR NN topology

The system is able to detect and track multiple persons. The performance of the tracking system here depends heavily on the intensity of interaction between persons. The tracker tends to easily lose the track in such cases but it recovers quickly from tracking failures [13, 14]. Future work would include incorporating adaptive segmentation algorithm and action recognition as well as other sensors in thermal vision-based mobile robot platform.

Conclusions

This paper presented scientific achievements of the authors in the field of mobile robotics and people tracking system thermal vision. The system uses a robust and fast image segmentation method and neural network based classifier. Based on features extracted from the processed thermal image classifier determines whether the segmented object is human or not. The results indicate good detection performance and consistent tracking in the case of single persons independently from current light conditions and in situations where no skin color is visible.

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 where experimental results showed us that NAR neural networks give good results, NN made predictions accurately and reliably an it 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. The system is also able to detect and track multiple persons.

Acknowledgment

This research was supported by the Ministry of Education, Science and Technological Development of Republic of Serbia through the project TR35005, and by German-Serbian bilateral project of the Ministry of Education, Science and Technological Development of Republic of Serbia and DAAD named "Robust Action Recognition for Human-Robot Synergy".

References

[1] Alvarez-Santos, V., *et al.*, 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] Boehme, H.-J., *et al.*, An Approach to Multi-Modal Human-Machine Interaction for Intelligent Service Robots, *Robotics and Autonomous Systems*, *44* (2003), 1, pp. 83-96
[3] Borja, R., *et al.*, Integration of Service Robots in the Smart Home by Means of UPnP: A Surveillance Robot Case Study, *Robotics and Autonomous Systems*, *61* (2013), 2, pp. 153-160
[4] Fernandez-Caballero, A., *et al.*, Optical Flow or Image Subtraction in Human Detection from Infrared Camera on Mobile Robot, *Robotics and Autonomous Systems*, *58* (2010), 12, pp. 1273-1281
[5] Ćirić, I., *et. al*, 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, Nis, Serbia, 2013, pp. 587-590
[6] Ćirić, I., *et al.*, Intelligent Optimal Control of Thermal Vision-Based Person-Following Robot Platform, *Thermal Science*, *18* (2014), 3, pp. 957-966
[7] Ćirić I., *et al.*, Intelligent Control of DaNi Robot Based on Robot Vision and Object Recognition, *Facta Universitatis, Series: Automatic Control And Robotics, 11* (2012), 2, pp. 129-140
[8] Huang, Z., Leng, J., Analysis of Hu's Moment Invariants on Image Scaling and Rotation, *Proceedings*, 2nd International Conference on Computer Engineering and Technology (ICCET), 2010, 7, pp. 476-480
[9] Treptow, A., *et al.*, Real-Time People Tracking for Mobile Robots Using Thermal Vision, *Robotics and Autonomous Systems*, *54* (2006), 9, pp. 729-739
[10] El-Bakry, H. M., Mastorakis, N., Fast Human Motion Tracking by Using High Speed Neural Networks, *Proceedings*, 9th WSEAS International Conference on Signal, Speech and Image Processing, and 9th WSEAS International Conference on Multimedia, Internet & Video Technologies, Budapest, Hungary, 2009, pp. 221-240
[11] Doulamis, N., *et al.*, Adaptable Neural Networks for Modeling Recursive Non-Linear Systems, *Proceedings*, 14th International Conference on Digital Signal Processing, Santorini, Greece, 2002, 2, pp. 1191-1194
[12] Yeredor A., Yule-Walker Equations Applied to Hessians of the Characteristic Function for Improved AR Estimation, *Proceedings*, IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP, Honolulu, Hawai, 2007, 3, pp. 1025-1028
[13] Petrović, E., *et. al.*, Stereo Vision-Based Human Tracking for Robotic Follower, *International Journal of Advanced Robotic Systems*, 2013, 10:230. doi: 10.5772/56124
[14] Ćirić, I., *et. al.*, Thermal Vision Based Intelligent System for Human Detection and Tracking in Mobile Robot Control System, *Proceedings*, 17th Symposium on Thermal Science and Engineering of Serbia, Sokobanja, Serbia, 2015, pp. 730-740

Paper submitted: April 1, 2016
Paper revised: August 8, 2016
Paper accepted: August 10, 2016