Application of artificial neural networks for fusion of data from radar and depth sensors applied for persons’ monitoring

Paweł Mazurek
Warsaw University of Technology, Faculty of Electronics and Information Technology, Institute of Radioelectronics and Multimedia Technology, Nowowiejska 15/19, 00-665 Warsaw, Poland

E-mail: p.mazurek@ire.pw.edu.pl

Abstract. The research reported in this paper is related to the fusion of measurement data from the impulse-radar sensors and infrared depth sensors applied in a system for unobtrusive monitoring of elderly persons. Three methods for data fusion, based on the artificial neural networks – one trained on real-world data, and two trained on synthetic data generated on the basis of two different models of the data – are compared with respect to their capacity of decreasing the uncertainty of position estimation in a series of experiments which involved the tracking of a moving person.

1. Introduction
The life expectancy has been growing in Europe for many years, while the healthy life expectancy has been slightly diminishing since the last decade of the XXth century (cf. http://www.healthy-life-years.eu/). Hence the growing importance of research on new technologies that could be employed in monitoring systems supporting care services for elderly persons. Those systems are expected to detect dangerous events, such as person’s fall, but also to predict those events on the basis of acquired data [1]; it has been shown recently that the analysis of gait may contribute to the fall prevention [2].

There are two emerging categories of monitoring techniques that can be applied for non-intrusive monitoring of elderly persons, viz. radar-based and depth-sensor-based techniques [3]. Radar sensors and depth sensors operate according to different physical principles and, therefore, have specific complementary advantages and disadvantages: radar sensors offer a broad field of view and capacity of the through-the-wall monitoring, but have relatively low accuracy; on the other hand, the depth sensors have higher accuracy, but their coverage is limited, and they cannot detect occluded targets. One may expect that the combination of such sensors, followed by adequate data fusion, could be beneficial. It has been recently shown that the artificial neural networks can be effectively used for the fusion of measurement data acquired by means of several sensors tracking a moving object [4].

Before being used, an artificial neural network has to be trained, what in the case of tracking of an elderly person is particularly challenging task: involving the elderly persons in the experimentation aimed at the acquisition of the reference data could be problematic. One possible solution is the use of the synthetic data for training the neural network, but the effectiveness of this solution may depend on the accuracy of the model of the data. This paper is devoted to the assessment of the uncertainty of the position estimates obtained by means of the neural-network-based method for fusion of the data from the radar sensors and from the depth sensor, trained on the synthetic data.
2. Methodology of experimentation
In the experiments, three methods for data fusion based on the artificial neural networks – one trained on real-world data (hereinafter called NNR method) and two trained on synthetic data generated on the basis of different models of the data (hereinafter called NNS1 method and NNS2 method) – were compared with respect to their capacity of decreasing the uncertainty of estimation of a monitored person’s position. The real-world data and the synthetic data, used for training, were representative of the movement according to the same trajectories; the real-world data used for testing the methods, however, were representative of the movement according to different trajectories. All the methods were based on the feedforward neural network with four input neurons (for two pairs of the x-y coordinates acquired by the radar sensors and the depth sensor), eight neurons in a single hidden layer, and two output neurons (for the fused x-y coordinates). In the experiments, the implementations of the neural networks, available in the MATLAB Deep Learning Toolbox, were used [5].

2.1. Acquisition of real-world data
The raw measurement data for experimentation were acquired by means of sensors configured as shown in figure 1. In the experiments, a person walked along ten straight-line trajectories which together covered all the reference points (see figure 2a), and along three square-shaped trajectories differing in length (see figure 2b); for each trajectory, 20 walks were performed.

2.2. Generation of synthetic data
The synthetic data for training the artificial neural networks were generated on the basis of the reference movement trajectories presented in figure 2a.

In the case of the first set of synthetic data (used for training the neural network of the NNS1 method), the x-y data representative of the measurement data acquired by means of a depth sensor were corrupted with zero-mean white noise with the standard deviation 0.05 m, while the synthetic data representative of the measurement data acquired by means of the impulse-radar sensors were corrupted with zero-mean red noise with the standard deviation 0.1 m. For each straight-line trajectory, \( R = 20 \) realizations of the data were generated for each type of sensor. The data acquisition rate for the depth sensor and for the impulse-radar sensors was set to 10 Hz.

![Figure 1. Experimental setup; the crosses indicate the reference points.](image1)

![Figure 2. Movement scenarios considered in the experiments: a) straight-line trajectories, b) squares-shaped trajectories.](image2)
In the case of the second set of synthetic data (used for training the neural network of the NNS2 method), the silhouette of a moving person was modelled by means of an ellipse, with semi axes of 0.1 m and 0.25 m, whose center followed a sine-shaped trajectory oscillating around the reference trajectory: the amplitude of the sine-shaped trajectory, modelling the sway of a body during a single step, was 0.04 m, while the half of the period of that trajectory, modelling the length of a single step, was 0.5 m. The coordinates of the points lying on the circumference of the ellipse were corrupted with zero-mean white noise with the standard deviation 0.1 m. To take into account the fact that the depth sensors “can see” only one side of the body, the x-y data representative of the measurement data acquired by means of this sensor were calculated as the mean values of the coordinates of half of the points of the ellipse – the points being closest to the x-axis. The synthetic data representative of the measurement data acquired by means of the impulse-radar sensors were generated as in the case of the first set of synthetic data, but were smoothed using the moving-average filter.

All the trajectories used for training the networks, are presented in figure 3.

2.3. Criteria for performance evaluation
The experiments were aimed at assessing the uncertainty of the estimation of the person’s position with respect to the assumed reference trajectory. The position estimates, obtained for each type of sensor and for each method for data fusion, were treated separately.

For each of the \( R \) sequences of the data representative of the person’s position in the x- and y-dimension:

\[
\left\{ \bar{x}_n^{(r)} \mid n = 1, \ldots, N^{(r)} \right\} \quad \text{and} \quad \left\{ \bar{y}_n^{(r)} \mid n = 1, \ldots, N^{(r)} \right\} \quad \text{for} \quad r = 1, \ldots, R
\]

the corresponding sequence of the absolute errors of the position estimation has been computed in the following way:

\[
\left\{ \Delta x_n^{(r)} = x_n^{(r)} - \bar{x}_n^{(r)} \mid n = 1, \ldots, N^{(r)} \right\} \quad \text{and} \quad \left\{ \Delta y_n^{(r)} = y_n^{(r)} - \bar{y}_n^{(r)} \mid n = 1, \ldots, N^{(r)} \right\} \quad \text{for} \quad r = 1, \ldots, R
\]

where \( x_n^{(r)} \) and \( y_n^{(r)} \) denote the reference value for \( \bar{x}_n^{(r)} \) and \( \bar{y}_n^{(r)} \).

![Figure 3](image-url)

**Figure 3.** Data used for training the artificial neural networks: the set of real-world data (left column), the first set of synthetic data (middle column) and the second set of synthetic data (right column).
The sequences of the absolute errors in the x-dimension and in the y-dimension have been then used to determine $R$ sequences of the position errors $\{\Delta \hat{r}_{n}^{(r)}\}$:

$$
\left\{\Delta \hat{r}_{n}^{(r)} = \sqrt{\left(\Delta x_{n}^{(r)}\right)^2 + \left(\Delta y_{n}^{(r)}\right)^2}\right\}_{n=1,\ldots,N^{(r)}} \text{ for } r=1,\ldots,R
$$

The qualitative assessment of the performance of the methods for data fusion was based on the inspection of the empirical cumulative distribution functions $F(\xi)$ [6]. $F(\xi)$ is an estimate of the cumulative distribution function characterizing the position errors, defined as:

$$
F(\xi) = \frac{1}{M} \sum_{m=1}^{M} I(\Delta \hat{r}_{m} \leq \xi)
$$

where $\{\Delta \hat{r}_{m} | m=1,\ldots,M\}$ is a sequence composed of all the sequences of the position errors $\{\Delta \hat{r}_{n}^{(r)}\}$, and $I(\cdot)$ is the indicator function which takes the value 1 if the condition inside the brackets is met, and 0 otherwise.

The quantitative assessment of the performance of all the methods for data fusion was based on an indicator called $A_{\text{ECDF}}$ and defined as the area under the empirical cumulative distribution functions $F(\xi)$ in the interval $\xi \in [0,1]$ m, taking the values from the interval [0,1]. The higher the value of $A_{\text{ECDF}}$, the better the performance of the data fusion. Furthermore, this type of assessment was based on the following statistics of the distribution of position errors: the mean position error (abbreviated to MEAE), the median position error (abbreviated to MEDAE), the maximum position error (abbreviated to MAXE) and the standard deviation of the position errors (abbreviated to STDE).

3. Results and discussion

The real-world trajectories used for testing are presented in figure 4, while in figure 5 – the trajectories obtained by means of the fusion of radar data and depth data. The empirical cumulative distribution functions characterizing the position errors are shown in figure 6, and the values of the uncertainty indicators, calculated on the basis of the position errors, are provided in table 1.

The analysis of the presented results is leading to the following conclusions:

- It can be seen that the data fusion has alleviated the problems occurring if only single-type sensors are used for monitoring, viz. the dispersion of the estimates have decreased, and the blind spots have disappeared.
- The estimates of the trajectories, obtained by means of the NNS1 method, are subject to bias introduced by the depth sensor – the depth sensor “can see” only one side of the body; this phenomenon was not taken into account in the generation of the first set of synthetic data, and therefore the bias has not decreased.
- The estimates of the trajectories, obtained by means of the NNS2 method, are more accurate than the estimates obtained by means of the NNS1 method – the bias introduced by the depth sensor has decreased since during the generation of the second set of synthetic data more realistic model of the person’s movement was used. However, it can be noticed that the trajectories are slightly distorted around the point (4,1) – in this area the person was on the edge of the field of view of the depth sensor, and the resulting deformation of the silhouette of a person caused the corruption of the estimates of the person’s position.
- The estimates of the movement trajectories, obtained by means of the NNR method, seem to be the most accurate – the bias and the distortion of the trajectories, introduced by the depth sensor, have decreased.
- The values of the uncertainty indicators, calculated on the basis of the fused data, are better than their values calculated on the basis of the radar data only or of the depth data only. Moreover, the values obtained for the NNR method and for the NNS2 method are very similar, and notably better than the values obtained for the NNS1 method.
During the experiments, various configurations of the artificial neural network were considered, and the following observations have been made:

- As long as the networks are trained sufficiently long, the number of neurons in the hidden layer varying from 6 to 10 does not influence the results of the data fusion.
- The best results are obtained for the networks trained with the Levenberg-Marquardt algorithm.
- All types of sigmoid transfer function – such as logistic function, hyperbolic tangent function or Elliot sigmoid function – may be effectively used, and they lead to similar results.
4. Conclusions
The novelty of the research results, presented in this paper, consists in identification of the effectiveness of neural-network-based methods for fusion of data from the impulse-radar sensors and from the depth sensors, when applied for unobtrusive monitoring of elderly persons in their home environment.

In the experiments, three methods for data fusion, based on the artificial neural networks, were compared with respect to their capacity of decreasing the uncertainty of estimation of a monitored person’s position: one method (NNR) was based on the network trained on the real-world data, and two methods (NNS1 and NNS2) were based on the network trained on the synthetic data, but the sets of synthetic data were generated according to two different models of the data.

The experiments have demonstrated that the artificial neural networks trained on the synthetic data may be effectively used for fusing the data from the radar sensors and from the depth sensors – and yield as good results as the networks trained on the real-world data – but the model of the data, used for generation of the synthetic data, has to take into account all the relevant characteristics of the person’s movement, and characteristics of the sensors used for monitoring.

Acknowledgments
This work has been accomplished partially within the project PL12-0001 financially supported by EEA Grants – Norway Grants, and partially within the project 2017/25/N/ST7/00411 financially supported by the National Science Centre, Poland.

References
[1] Chaccour K, Darazi R, Hassani A H E and Andres E 2017 IEEE Sens. J. 17 pp 812–22
[2] Stone E, Skubic M, Rantz M, Abbott C and Miller S 2015 Gait Posture 41 pp 57–62
[3] Wagner J, Mazurek P, Miekina A, Morawski R Z, Jacobsen F F, Sudmann T T, Borsheim I T, Øvsthus K and Ciamulski T 2017 Measurement 111 pp 420–31
[4] Sun B, Jiang C and Li M 2016 Sensors 16 p 1823
[5] Mathworks, Matlab Documentation: Deep Learning Toolbox. Available: https://www.mathworks.com/help/deeplearning/index.html (2018-12-14)
[6] van der Vaart A W 1998 Asymptotic Statistics (Cambridge: Cambridge University Press)