ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 357-362

A Neural Approach to a Sensor Fusion Problem

V. Colla*, M. Sgarbi†, L.M. Reyneri◇, A.M. Sabatini*
* Scuola Superiore Sant'Anna
Via Carducci 40, 56127 PISA, ITALY
† Via Aquilaia, 302 - 58054 Scansano (GR) - ITALY
◇ Dipartimento di Elettronica, Politecnico di Torino,
C.so Duca degli Abruzzi 24, 10129 TORINO, ITALY

Abstract.
Our problem concerns the joint interpretation of UltraSonic and InfraRed measurements provided by a composite proximity sensor, in order to extract geometrical and morphological features of a flat target. Neural networks are applied here both for modeling and for classification purposes. Physical knowledge of the sensing modalities helps in simplifying the network structure, in order to minimize its complexity, save computing time and make the system suitable for real–time applications.

1. Introduction

Sensor integration represents an active area of investigation in the field of robotics and industrial automation, as a key feature to improve the perceptual abilities of complex robotic systems. In particular, proximity sensors seem to be quite appealing for their acceptable cost-to-performance ratio, as compared to that of more expensive sensing techniques, e.g., vision or laser range finding.

Among proximity sensors, ultrasonic (US) and infrared (IR) detectors are particularly interesting in real-life applications, as one of their most interesting features is that IR reflecting behaviors have well known characteristics of complementarity [1]; fusion of data provided by different proximity sensors is hence a crucial point to partly overcome their limitations. In [2] a composite sensor has been proposed, which integrates into the same device an US range finder (the couple of emitting and receiving capsules) and an IR detector (the light emitting diode and its coupled phototransistor); in [4] a first attempt was made to accomplish a strategy which uses the spatial information provided by the US for interpreting the signal coming from the IR sensor. The aim was to estimate the spectral reflectivity, what we call the "color", of a flat target, in view of supporting the navigation of a semi autonomous vehicle, as a sort of "label recognizer". In the present paper the same problem is addressed, but a novel approach based on neural networks allows a faster design of the sensor fusion system as well as better performance, both in terms of lower measurement

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 357-362

error, lower classification error probability and fewer measurements needed for making a decision.

The paper is organized as follows: in Sec. 2. some remarks are given on the extraction of the sensor–target relative position from the US measurements; Sec. 3. is devoted to processing the sensorial data by means of a composite neural scheme, and finally some results are presented in Sec. 4..

2. Remarks on the sensing device

The hardware of the composite proximity sensor has already been described in [2]: we just recall that robust analog signal processing techniques have been applied in order to cope with the low-cost, off-the-shelf components employed. In the application discussed here, a linear array of three sensing units is employed, but the IR measurement is extracted only from the central one. The measurement of the peak amplitude of the IR radiation detected by the central sensor is performed by means of an A/D converter integrated within the control unit, i.e. a MOTOROLA HC11 microcontroller; its limited resolution (8 bit) reduces the possibility of exploiting the IR output in a proportional manner when it is small, namely for relatively far targets. An additional logarithmic amplifier can improve detector performance for far objects.

The center-to-center spacing of the three US sensing elements is $h = 0.04$m: the i-th sensor computes a noisy estimate r_i of the target-sensor distance d_i. In conditions of thermally compensated US sensors, each range measurement r_i can be modeled as: $r_i = d_i + n_i$ where n_i is a zero-mean, Gaussian, random variable with variance σ_i^2. Without any loss of generality, we assume that the measurements errors performed by the three sensors are independent (in our case $\sigma_i = 0.7 \cdot 10^{-3}$m).

The US device is able to estimate the distance d of a planar surface and its orientation θ with respect to the sensing device, namely the angle between the target surface and the line joining the sensor center and the incidence point by exploiting the following relations [3]:

$$d = r_2 \qquad\qquad \theta = \arcsin\left(\frac{r_3 - r_1}{2h}\right) \tag{1}$$

Thus, the US position estimator output can be represented as $\mathbf{z}_{US} = \mathbf{v}_{US} + \epsilon_{US}$ where $\mathbf{v}_{US} = \begin{bmatrix} d & \theta \end{bmatrix}^T$ and ϵ_{US} is a random vector composed of two additive, zero-mean, jointly Gaussian noise components.

3. Neural processing of the sensorial data

The problem to be addressed in order to interpret the data coming from the IR subsystem is twofold: first of all a model of the IR output signal v_{IR} has to be developed to express its dependence on $\mathbf{v}_{US}$ and on the target color. Afterwards, a methodology for recognizing the target color from the sensing device outputs has to be developed. These issues have already been discussed in [4], where the IR analytical model has been empirically developed on the basis of some results available in literature and of the knowledge of the physics of the sensor.

In the present work, a neural approach is adopted both for the modeling and for the classification problem: anyway the physics of the IR sensing device is not neglected, but it is exploited in order to simplify the network structure. By means of an exhaustive test of the IR output, we point out a nearly linear dependence of the IR measurement v_{IR} on $1/d^2$. For 5 cm $\leq d \leq$ 25 cm, the output values for the diverse colors differ for a constant multiplying factor and this suggests us to associate a proportionality coefficient α ($0 < \alpha \leq 1$) to each color; the value 1 corresponds to what we call a "WHITE" target. Moreover, the IR output features a monotonically decreasing behavior with increasing values of $|\theta|$, which is related to the limited width of the radiation lobe and to the consequent loss of energy when the received signal comes back from a surface which is not orthogonal. These considerations allow us to model the IR output by separating the contributions of d, θ and color to the IR output signal, provided that the analysis is restricted to the region 5 cm $\leq d \leq$ 25 cm and $-45° \leq \theta \leq 45°$. For distances and angles outside of these ranges, there is more dependence of the two parameters on each other. The model function has the following expression:

$$m_{IR}(\mathbf{v}_{US}, \alpha) = \alpha(\text{color}) \cdot m_d(d) \cdot m_\theta(\theta) \tag{2}$$

where $m_d(d)$ is a function which takes into account the dependence of the IR output on the target distance and the detector sensitivity, while $m_\theta(\theta)$ is a function whose value lies in $[0, 1]$ and expresses the dependence on the target orientation. Because of the introduction of α, $m_d(d)$ and $m_\theta(\theta)$ need to be estimated just for the WHITE target. As shown in fig. 1.a, for mapping each of these functions we adopt a 2–layers feed-forward neural network with sigmoidal activation functions in the hidden layer and a linear output layer [5], trained on a great number of measurements performed on WHITE targets at different values of distance and angle, within the considered ranges.

The integration of the data provided by the US and IR subsystem consists of using the US measurements concerning d and θ to make a prediction of $z_W(d, \theta) = m_d(d) \cdot m_\theta(\theta)$ in the case of a WHITE target: the ratio $\hat{\alpha} = v_{IR}/z_W$ is then an estimate of α (from (2)) and is fed to the classifier. The goodness of this approximation depends on different factors: the accuracy of the model (2), the internal noise of the IR sensor and the noise in the US measurements.

A complex and oversized structure of the network estimating $m_d(d)$ and $m_\theta(\theta)$ is not a viable solution, because the noise sources in the sensing devices could not be eliminated, and, moreover, the risk of overfitting and loss of generalization would increase. In practice the network would be prone to training the noise more than the real signal.

A more practical solution is to make the classification process take into account the aforementioned uncertainties in a probabilistic sense, as was also

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 357-362

made in our previous work [4], where possible inconsistency of the IR and US data were taken into account in an heuristic fashion. Therefore we have decided to use smaller neural networks which have the major advantage of "smoothing" the learned function, canceling out the effects of noise.

Neural processing implements two major functions, as shown in fig. 1.a: two networks (A and B) are trained to predict a value of z_W, from which $\hat{\alpha}$ can be estimated; a third network (CLASSIFIER) is used to classify the values of $\hat{\alpha}$ (and therefore the detected target color) into a number of color classes.

Networks A and B are two Multi-Layer Perceptrons with one hidden layer [5] trained with the available experimental data (samples of v_{IR} for different values of d_{US}, θ_{US} and v_{IR}). Network CLASSIFIER is a single layer WRBF network trained to classify the different colors. The classifier is a *one-hot decoder* with as many outputs as classes and samples measured with different target colors are used for its training. After training, the network output delivers a continuous value in each component of the output vector, which can be interpreted as being proportional to the class probability and used for attributing the input pattern to one of the classes. Fig. 1.a summarize the complete processing of the sensorial data.

4. Numerical results

To evaluate the influence of color and location of the reflecting surface on

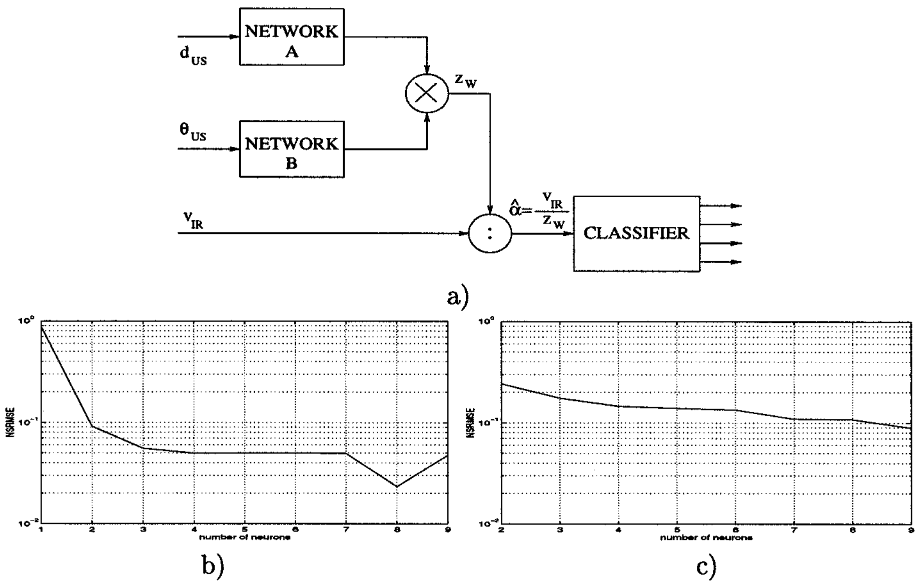

Figure 1: a) Block diagram of the processing performed on the data. b) NSRMSE vs. the number of neurons in the hidden layer for network A. c) NSRMSE vs. of the number of neurons in the hidden layer for network B.

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 357-362

the amplitude of the IR sensor output, extensive calibration tests have been carried out by means of a load frame INSTRON Mod. 4464, for imparting precisely controlled displacements to the target, and of a MITUTOYO rotator for varying the sensor–target relative orientation. The targets are standard A4 thin cards in 4 different colors: WHITE, BROWN, GREEN and BLACK.

The collected data has been used for networks training and validation. Networks A and B in Fig. 1.a approximate the functions $m_d(d)$ and $m_\theta(\theta)$ respectively. In particular, network A is trained with a set of measurements performed for $\theta = 0$ and d varying in the range $[5\text{cm}, 25\text{cm}]$, while network B is trained with the ratios between the outputs measured (at a fixed distance) for θ variable and the value obtained for $\theta = 0$.

As an index for evaluating the networks performances, we adopt the *Normalized Square Root Mean Square Error* (NSRMSE):

$$\text{NSRMSE} = \frac{1}{\sigma_y}\sqrt{\sum_{i=1}^{M}[\hat{f}(x_i) - y_i]^2} \qquad (3)$$

where M is the number of training points, while (x_i, y_i) are the sample points, σ_y is the standard deviation of the output data y_i and $\hat{f}$ is the function estimate obtained from the network. The training is performed in 10000 epochs. Fig. 1.b and Fig. 1.c show the NSRMSE as a function of the number of neurons in the hidden layer for networks A and B respectively; it is clear that three and four neurons (for networks A and B, respectively) are sufficient to achieve a reasonable error.

Moreover, we evaluate the values of α reported in Tab. 1.a for the different colors from the mean ratio of the experimental data and the model for $\alpha = 1$.

We test the system performances with different numbers of neurons for networks A and B and with a WRBF network as classifier, because elliptic decision boundaries are needed in our case. Tab.1 summarize the results obtained with both training and validation sets composed of 14400 measurements.

COLOR	α
brown	0.90
green	0.81
black	0.15

a)

N_A	N_B	$P_c\%$	$P_e\%$
3	2	96.2	1.9
5	5	96.8	0.8
6	5	96.9	0.5
7	5	97.3	0.6
7	6	97.4	0.6

b)

Table 1: a) Measured values of α for the different target colors. b) Performances of different networks: N_A and N_B are the number of neurons in the hidden layer of netwoks A and B respectively, P_e and P_c are the probabilities of error and of correct decision.

It is worthwhile to remember that, in order to save computational time by reducing the number of neurons in networks A and B, a moderate decrease on

ESANN'1998 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 22-23-24 April 1998, D-Facto public., ISBN 2-9600049-8-1, pp. 357-362

the probabilities of correct decisions could also be accepted, provided that the increment in the error probability is negligible; in fact a missed detection can occur when the US and IR measurement are not perfectly consistent and the problem can be overcome by simply performing a second pair of measurements.

5. Conclusions

We have proposed a neural–based strategy for fusing the information provided by two different sensing modalities, UltraSonic range finder and InfraRed detectors: the IR output signal amplitude is interpreted in a proportional manner, in order to exploit the IR output dependency on the target surface reflectivity. Neural networks have been used for both modeling and classification purposes, but some knowledge on the sensor physics has allowed some simplifications.

Acknowledgments

The composite sensor system has been designed and developed within the TIDE/OMNI project TP 1097. The authors gratefully acknowledge Dr. Vincenzo Genovese for his contribution to the hardware design.

References

[1] A.M. Flynn: "Combining Sonar and Infrared Sensors for Mobile Robot Navigation", *Int. Jour. Robotic Research*, Vol. 7, pp. 5-14, N. 6, 1988.

[2] A.M. Sabatini, V. Genovese, E. Guglielmelli, A. Mantuano, G. Ratti, P. Dario: "A Low-Cost, Composite Sensor Array Combining Ultrasonic and Infrared Proximity Sensors", *Proc. 1995 IEEE/RSJ Int. Conf. Intell. Rob. Syst.*, Vol. 3, pp. 120-125, August 5-9, 1995, Pittsburgh, PA, USA.

[3] Y. Nagashima, S. Yuta: "Ultrasonic sensing for a mobile robot to recognize an environment – Measuring the normal direction of walls," *Proc. 1992 IEEE/RSJ Int. Conf. Intell. Robots. Syst.*, Vol. 2, pp. 805-812, Raleigh, NC, July 7–10, 1992.

[4] A.M. Sabatini, V. Colla: "A Composite Proximity Sensor for Target Location and Color Estimation ," *Proc. 5-the Int. Symp. Measur. Cont. in Robotics*, pp. 124–129, 6-8 May, 1996, Brussels, Belgium.

[5] S. Haykin: "Neural Networks: A Comprehensive Foundation," *Mc Millan College Publishing Company*, New York, 1994.

[6] L.M. Reyneri: "Weighted Radial Basis Functions for Improved Pattern Recognition and Signal Processing," *Neural Processing Letters* Vol. 2, No. 3, pp. 2–6, May 1995.