MOVING OBJECT RECOGNITION BASED ON RADIAL BASIS FUNCTIONS NETWORKS

Adrian G. Borş *Ioannis Pitas*

Department of Informatics, University of Thessaloniki, Thessaloniki 540 06, Greece

ABSTRACT

We apply a radial basis functions (RBF) neural network for simultaneous optical flow estimation and segmentation in image sequences. The probability of optical flow estimation and moving object segmentation is modeled as a set of energy functions depending on the feature vectors extracted from the image. The energy functions are represented as a metric in the feature space and is used as a classification criterion. The learning algorithm for the RBF network is based on robust estimators. After the learning stage, the network can be used to classify the moving objects of the image and to infer their optical flow. The parameters of the network are used in a multiresolution approach to segment the moving objects with an improved accuracy. The network is applied in real image sequences.

1. THE COST FUNCTION

After the image is partitioned in blocks situated on a rectangular grid, for each block, a feature vector is assigned denoted by $\mathbf{u}_{IJ}$, where I and J are the coordinates of the block. The feature vector contains the motion vector provided by the block matching algorithm, the gray level representing the average gray level and the position vector representing the coordinates of the block. The features are scaled in order to be in the same interval.

Let us denote by $\mathcal{M}_k$ the motion vector and by $\mathcal{S}_k$ the segmentation parameter vector associated with an image region k, drawn from the frames f_{t-1} and f_t. The image is segmented in regions according to the maximum a posteriori probability:

$$P(\mathcal{M}_k, \mathcal{S}_k | f_t, f_{t-1}) = \max_{i=1}^{L} P(\mathcal{M}_i, \mathcal{S}_i | f_t, f_{t-1}), \quad (1)$$

where L is the total number of moving regions. The a posteriori probabilities can be decomposed according to the Bayes rule, and related to the energy functions [1]. The energy associated to a certain moving region is given by:

$$E_j = E(f_t | f_{t-1}, \hat{\mathcal{M}}_j, \hat{\mathcal{S}}_j) + E(\hat{\mathcal{M}}_j | \hat{\mathcal{S}}_j, f_{t-1}) + E(\hat{\mathcal{S}}_j | f_{t-1}).$$

The maximization of the a posteriori probabilities is related to the minimization of the energy functional E_j.

Each energy functional is represented with respect to the feature vectors as a metric in the feature space [2], connecting the feature vectors $\mathbf{u}_{IJ}$ with the vectors characteristic to the moving regions:

$$E_j = \sum_{I,J} (\mathbf{u}_{IJ} - \hat{\mu}_j)^T (\mathbf{u}_{IJ} - \hat{\mu}_j) + H(\hat{\mathcal{M}}_k) \quad (2)$$

where $\hat{\mu}_k = [\hat{\mathcal{M}}_k, \hat{\mathcal{S}}_k]$ is the estimated feature vector, associated with a moving region and $H(\hat{\mathcal{M}}_k)$ is the displaced frame difference, related to the image reconstruction probability $E(f_t | f_{t-1}, \hat{\mathcal{M}}_j, \hat{\mathcal{S}}_j)$.

2. THE LEARNING ALGORITHM

In our approach, the probability function associated to the optical flow and to the segmentation of a moving object is decomposed in a sum of kernels:

$$P(\mathcal{M}_k, \mathcal{S}_k | f_t, f_{t-1}) = \sum_{j=1}^{L} \lambda(k, j) \phi_j(\mathbf{u}_{IJ}) \quad (3)$$

where $\mathbf{u}_{IJ}$ is an input vector, L is the number of kernels and $\phi_j(\mathbf{u}_{IJ})$ is a Gaussian function, chosen as kernel:

$$\phi_j(\mathbf{u}_{IJ}) = \exp\left[-(\mathbf{u}_{IJ} - \hat{\mu}_j)^T \hat{\Sigma}_j^{-1} (\mathbf{u}_{IJ} - \hat{\mu}_j) - H(\hat{\mathcal{M}}_j) \right]$$

where $\hat{\mu}_j$, $\hat{\mathbf{\Sigma}}_j$ are the estimated parameters of the Gaussian function. In order to find the weights for the RBF network we employ a two step learning approach [3].

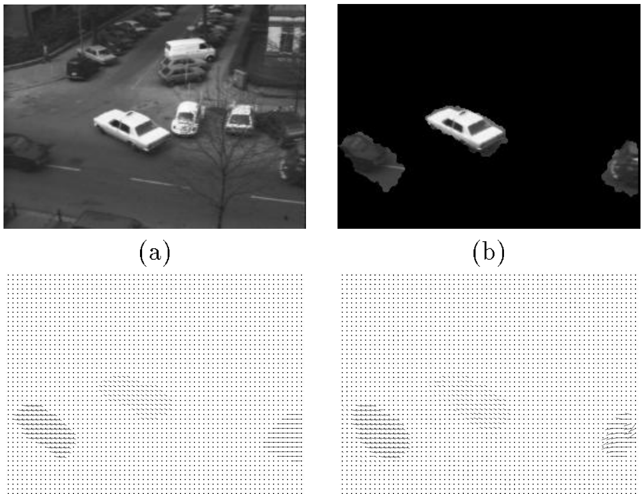
The Median RBF algorithm excludes from training the outlying patterns and it was used for the RBF parameter estimation [3]. The data samples are associated with a basis function according to a minimal energy in (2). After the data samples are ordered, the center is chosen as the marginal median:

$$\hat{\mu}_j = \text{med}\, \{\mathbf{X}_0, \mathbf{X}_2, \ldots, \mathbf{X}_{n-1}\} \quad (4)$$

where $\mathbf{X}_{n-1}$ is the last pattern assigned to the cluster j. For the variance we use the median of the absolute deviation estimator [4]:

$$\hat{\sigma}_{j,h} = \frac{\text{med}\, \{|\mathbf{X}_0 - \hat{\mu}_j|, \ldots, |\mathbf{X}_{n-1} - \hat{\mu}_k|\}}{0.6745} \quad (5)$$

where 0.6745 is a scaling parameter in order to make the estimator, Fisher consistent for the normal distribution.



(a) (b)

(c) (d)

Figure 1: (a) "Hamburg taxi" sequence, (b) The moving objects segmentation, (c) The optical flow processed by the RBF, (d) The optical flow processed by the ICM.

A fast implementation based on histogram modeling was presented in [3]. The backpropagation is used for finding the parameters $\lambda(k, j)$ (3). The hidden units which do not have assigned many data vectors, according to (1), are discarded. After training the network, the network parameters can be used on a different image partition in block sites, in a multiresolution approach.

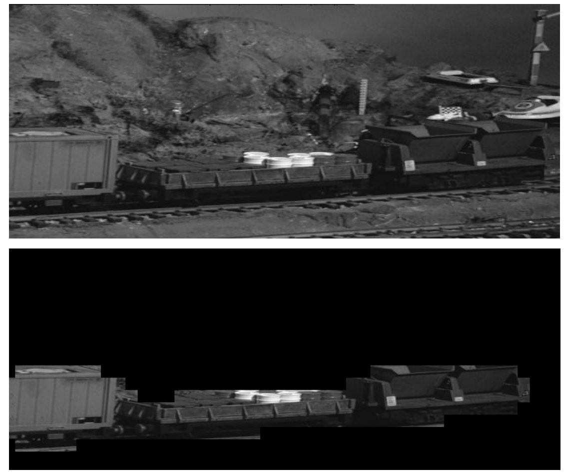
3. SIMULATION RESULTS

Two image sequences are considered for experimental results. The "Hamburg taxi" sequence, Fig. 1 (a), contains three main moving objects. The optical flow processed by the RBF network is represented in Fig. 1 (c). For comparison purposes the Iterated Conditional Modes (ICM) was tested [5]. ICM is minimizing locally a cost function by visiting all the block sites sequentially. This algorithm takes into account the borders of the moving objects, but smoothes only locally, inside a given neighborhood. The results obtained after applying ICM are presented in the Fig. 1. (d). The RBF network was applied in a multiresolution approach, at pixel level in Fig. 1 (b). The results for the "Train" sequence are shown in Fig. 2. A comparison in terms of the classification error, mean absolute error, mean square error and the necessary number of parameters are provided in Table 1. The RBF network provides better estimation and requires lesser parameters then ICM.

4. CONCLUSIONS

The moving object recognition is considered as a pattern recognition application based on the RBF network. The classification criterion takes into account the gray level, the motion and the position.

(a)



(b)

Figure 2: (a) Frame from the "Train" sequence, (b) The optical flow and the moving object segmentation.

Sequen.	Method	Class. error	MAE	MSE	No. of Param.
Hamb. Taxi	RBF	3.02	0.17	0.85	210
	ICM	4.07	0.33	1.15	9216
Train	RBF	3.09	0.54	3.61	36
	ICM	2.84	0.99	4.82	2430

Table 1: Comparison between RBF and ICM.

The RBF learning algorithm is unsupervised. After segmenting the moving scene, the information recorded in the network parameters can be used for image analysis or object based coding.

5. REFERENCES

[1] A. G. Borş, I. Pitas, "Segmentation and estimation of the optical flow," *Proc. Conf. on Computer Analysis of Images and Patterns,* Prague, Czech Republic, pp. 680-685, 6-8 Sep. 1995.

[2] J. Marroquin, S. Mitter, T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *J. Am. Stat. Assoc.,* vol. 82, no. 397, pp. 76-89, 1987.

[3] A. G. Borş, I. Pitas, "Median Radial Basis Functions Network for Optical Flow Processing," *Proc. of IEEE Workshop on Nonlinear Signal and Image Processing,* Neos Marmaras, Greece, pp. 702-705, 20-22 June 1995.

[4] G. Seber, *Multivariate Observations,* J. Wiley, 1986.

[5] M. M. Chan, M. I. Sezan, A. M. Tekalp, "An algorithm for simultaneous motion estimation and scene segmentation," *Proc. IEEE ICASSP,* Adelaide, Australia, pp. V-221 - V-224, 1994.