ALPHA-TRIMMED MEAN RADIAL BASIS FUNCTIONS AND THEIR APPLICATION IN OBJECT MODELING

Adrian G. Borş *Ioannis Pitas*

Department of Informatics
University of Thessaloniki
Thessaloniki 54006, Greece

ABSTRACT

In this paper we use Radial Basis Function (RBF) networks for object modeling in images. An object is composed from a set of overlapping ellipsoids and has assigned an output unit in the RBF network. Each basis function can be geometrically represented by an ellipsoid. We introduce a new robust statistics based algorithm for training radial basis function networks. This algorithm relies on α-trimmed mean statistics. The use of the proposed algorithm in estimating ellipse parameters is analyzed.

1. INTRODUCTION

Radial basis function neural network consists of a two layer feed-forward structure employed for functional approximation and classification proposes. When used in pattern classification an RBF network successfully approximates the Bayesian classifier [1, 2]. In this case, the underlying probability functions are decomposed in a sum of kernel functions with localized support. The functions, implemented by the hidden units, are usually chosen as Gaussian. The intersection of an Gaussian with an hyperplane is geometrically an ellipsoid. Objects in images can be seen as composed from a set of overlapping elliptic discs in 2-D [3], or ellipsoids in 3-D.

A classical RBF network training algorithm relies on the learning vector quantization [1]. This learning algorithm represents the adaptive version of the moments approach, used for estimating the ellipse parameters [4]. However, the moments method is likely to provide biased estimates in the case when the ellipses are overlapping or when they are embedded in noise. In [2, 5] marginal median and median of the absolute deviation [6] have been proposed as robust estimators for finding the RBF's hidden unit parameters. In this paper we analyze the α-trimmed mean RBF algorithm. The previous two algorithms are particular cases of the new approach. α-trimmed mean has been extensively used in image filtering [6, 7]. After ranking the data samples assigned to a basis function, a certain number of them is eliminated from estimation. The number of data samples to be eliminated depends on the distribution [8, 9]. We analyze the estimation of the ellipse parameters when using α-trimmed mean statistics. The proposed algorithm is applied for modeling artificial generated patterns and for segmenting a set of microscope images.

2. THE RADIAL BASIS FUNCTION MODEL

The RBF network has been employed in many applications including curve modeling [1] or moving object recognition [5]. The output unit models a complex function composed of a mixture of overlapping kernels. Each kernel implements an unnormalized Gaussian function :

$$\phi_j(\mathbf{X}) = \exp\left[-(\mathbf{X}-\mu_j)'\,\mathbf{\Sigma}_j^{-1}\,(\mathbf{X}-\hat{\mu}_j)\right] \quad (1)$$

where $\mathbf{X}$ represents the input feature vector, and $\mu_j, \mathbf{\Sigma}_j$, the Gaussian center vector and covariance matrix. The output represents a weighted sum of hidden units scaled to the interval $(0,1)$ by the sigmoidal function :

$$Y_k(\mathbf{X}) = \frac{1}{1+\exp\left[-\sum_{j=1}^{L}\lambda_{kj}\phi_j(\mathbf{X})\right]} \quad (2)$$

where λ_{kj} are output weights, L is the number of hidden units and $k = 1, \ldots, M$ are the output units.

The parameters of the RBF network are found by training. A classical algorithm for training RBF networks consists of assigning a group of data samples to a basis function according to the smallest Euclidean distance between every data sample and the basis function center :

$$||\mathbf{X}-\mu_k|| = \max_{j=1}^{L}||\mathbf{X}-\mu_j||. \quad (3)$$

The initial values of the center vectors are taken randomly in the input range. From the data distribution, assigned to a basis function, the parameters of

(1) are calculated. In [1] a variant of the learning vector quantization, which corresponds to classical statistics assumptions, has been used. Marginal median and median of absolute deviation have been considered for calculating the hidden unit parameters in [2, 5]. This algorithm has been called median RBF (MRBF).

3. THE ALPHA-TRIMMED MEAN RBF TRAINING ALGORITHM

We propose a new robust statistics based training algorithm. Let us order the data samples assigned to a specific basis center when using (3) and denote them according to their rank as $X_{(k)}$ for $k = 0, \ldots, N_k$, where N_k is the number of data samples assigned to the kth basis function. A general description of many robust statistics algorithms consists of assigning a weight to every data sample with respect to its rank and calculating the location estimate as :

$$\hat{\mu}_k = \frac{\sum_{i=0}^{N_k} W_i \, \mathbf{X}_{(i)}}{\sum_{i=0}^{N_k} W_i} \tag{4}$$

where W_i is the weight assigned to the data sample with respect to its rank. If $W_i = 1$ for $i = 0, \ldots, N_k$ then the location is computed by averaging. If $W_i = 0$ for $i = 0, \ldots, \frac{N_k}{2} - 1$, $i = \frac{N_k}{2} + 1, \ldots, N_k$ and $W_i = 1$ for $i = \frac{N_k}{2}$ we obtain the marginal median estimator. In other robust statistics algorithms, W_i is replaced by a function which decreases with respect to the distance of the ordered sample $X_{(i)}$ from the central data sample $X_{(\frac{N_k}{2})}$. In this study we propose the α-trimmed mean algorithm [6] which assigns $W_i = 1$ for $i = \alpha N_k, \ldots, N_k - \alpha N_k$ in (4) and $W_i = 0$ for the rest of data sample, where α is the percentage of data samples to be eliminated from the estimation. This algorithm is called α-trimmed mean RBF and estimates the center of the basis function as :

$$\hat{\mu}_k = \frac{\sum_{i=\alpha N_k}^{N_k - \alpha N_k} \mathbf{X}_{(i)}}{N_k - 2\alpha N_k} \tag{5}$$

The α-trimmed mean algorithm has been shown as a good choice for the long and medium tail data distributions. This algorithm has been extensively used for image filtering [6, 7].

The parameter α is chosen according to the data distribution. The following measure is used for estimating the tail of the data distribution [8, 9] :

$$Q = \frac{U[0.5] - L[0.5]}{U[0.05] - L[0.05]} \tag{6}$$

where $U[\beta]$, $L[\beta]$ represent the average of the upper and respectively lower β percentage of data samples. The number of data samples to be trimmed away relies directly on the value of Q :

$$\hat{\alpha} = \frac{1 - Q}{2}. \tag{7}$$

When the distribution is long tailed, the amount of data samples to be trimmed is large, and when the distribution tail is short, the amount of data samples to be trimmed is small.

For the second order statistics parameters, we order the data samples according to their Mahalanobis distance from the estimate $\hat{\mu}_k$:

$$M_{(0)} = \min_{j=0}^{L} [(\mathbf{X} - \hat{\mu}_j)' \; \hat{\mathbf{\Sigma}}_j^{-1} \; (\mathbf{X} - \hat{\mu}_j)] \tag{8}$$

After ordering the data samples $M_{(0)} < \ldots < M_{(N_k)}$, the estimate of the covariance matrix is chosen as :

$$\hat{\mathbf{\Sigma}}_j = \frac{\sum_{i=0}^{N_k - 2\alpha_M N_k} (\mathbf{X}_{(i),M} - \hat{\mu}_j)' \, (\mathbf{X}_{(i),M} - \hat{\mu}_j)}{N_k - 2\alpha_M N_k} \tag{9}$$

where $\mathbf{X}_{(i),M}$ denotes the ith ordered data sample according to the Mahalanobis distance (8). The covariance matrix in (8) is initially estimated based on classical second order statistics, without trimming any data sample away from the given distribution, i.e. by considering $\alpha_M = 0$ in (9). The number of data samples to be trimmed is calculated using :

$$\hat{\alpha}_M = 1 - \frac{U[0.5]}{U[0.05]} \tag{10}$$

Both formulas used in (5,9) can be calculated from data histograms. For the output parameters we employ the backpropagation algorithm as in [2]. The mixture of Gaussian functions can approximate complex functions.

If we drop the exp and the sign under the exponential in the expression (1) we obtain the equation of an ellipsoid. We can observe that for $\mathbf{X} \in \mathbb{R}^2$, if we intersect the Gaussian function with a plane we obtain an ellipse. With ellipses we can model complex 2-D objects [3]. The mean of the Gaussian function from (1) corresponds to the center, the variance to the width and the cross-correlation to the orientation of the ellipse.

4. ELLIPSE PARAMETERS ESTIMATION

Let us consider the equation of an ellipse in the analitic form :

$$(\mathbf{X}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{X}-\mu)=1 \qquad (11)$$

where μ denotes the center vector of the ellipse and $\mathbf{\Sigma}^{-1}$ its width and orientation. We denote the components of this matrix as :

$$\mathbf{\Sigma}^{-1}=\begin{pmatrix} a & c \\ c & b \end{pmatrix} \qquad (12)$$

The method based on moments for estimating the parameters of an ellipse is described in [4]. We consider an elliptic disc as a uniform distributed statistics inside of an elliptic shape. We can calculate the ellipsis center based on the first order moments :

$$E[\hat{\mu}_x]=\frac{\int_{x_{inf}}^{x_{sup}}\int_{y_{inf}}^{y_{sup}} x\, dydx}{\int_{x_{inf}}^{x_{sup}}\int_{y_{inf}}^{y_{sup}} dydx} \qquad (13)$$

where $x_{inf}, x_{sup}, y_{inf}, y_{sup}$ denote the geometrical limits of the ellipse and the denominator denotes the area left after truncation.

In order to evaluate the limits of integration we calculate the tangents to the ellipse, which are parallel to the x axis [4] :

$$x_t=\mu_x+\frac{\mathbf{\Sigma}^{-1}\mathbf{v}}{\sqrt{\mathbf{v}'\mathbf{\Sigma}^{-1}\mathbf{v}}} \qquad (14)$$

where $\mathbf{v}$ is the vector normal to the ellipse and parallel with the x axis, and μ_x is the x component of the center vector. After choosing two vectors $\mathbf{v}_1=(1\quad 0)'$ and $\mathbf{v}_2=(-1\quad 0)'$ we derive the limits on the x axis. In the case of trimming, let us denote with x_α the interval from the ellipse eliminated by α-trimming. The integration limits corresponding to the truncated ellipse are :

$$x_{sup}, x_{inf}=\mu_x\pm\left(x_\alpha-\sqrt{\frac{b}{ab-c^2}}\right) \qquad (15)$$

The integration limits on the y axis are derived from the equation (11) :

$$y_{sup}, y_{inf}=\mu_y-\frac{c(x-\mu_x)\pm\sqrt{b-(ab-c^2)(x-\mu_x)^2}}{b} \qquad (16)$$

The truncated ellipse area can be expressed, after changing the variable, and considering (15,16) :

$$\int_{x_{inf}}^{x_{sup}}\int_{y_{inf}}^{y_{sup}} dydx=$$

$$\frac{2}{\sqrt{ab-c^2}}\int_{-1+x_\alpha\sqrt{(ab-c^2)/b}}^{1-x_\alpha\sqrt{(ab-c^2)/b}}\sqrt{1-z^2}dz \qquad (17)$$

For the numerator from (13), similarly with the above derivation, we obtain :

$$\int_{x_{inf}}^{x_{sup}}\int_{y_{inf}}^{y_{sup}} x\, dydx=$$

$$\frac{2\mu_x}{\sqrt{ab-c^2}}\int_{-1+x_\alpha\sqrt{(ab-c^2)/b}}^{1-x_\alpha\sqrt{(ab-c^2)/b}}\sqrt{1-z^2}dz+$$

$$\frac{2\sqrt{b}}{ab-c^2}\int_{-1+x_\alpha\sqrt{(ab-c^2)/b}}^{1-x_\alpha\sqrt{(ab-c^2)/b}} z\sqrt{1-z^2}dz \qquad (18)$$

The second integral in (18) is zero and from (13,17,18) the result of the estimation is :

$$E[\hat{\mu}_x]=\mu_x. \qquad (19)$$

This result proves that the estimation of the ellipse center by trimming is unbiased in the case of perfect ellipses, $\forall\ \alpha\in[0,\frac{1}{2}]$.

We evaluate the α-trimmed mean algorithm for estimating the parameters of the ellipse covariance matrix (width and orientation). The equation of the ellipse, resulted after trimming α_M of the data samples and after ranking them according to the Mahalanobis distance with respect to the ellipse center (8), is :

$$(\mathbf{X}-\mu)'\mathbf{\Sigma}^{-1}(\mathbf{X}-\mu)=1-\alpha_M \qquad (20)$$

The estimation of the ellipse width is given by normalized second order moments [4] :

$$E[\sigma_x^2]=\frac{\int_{x_{M,inf}}^{x_{M,sup}}\int_{y_{M,inf}}^{y_{M,sup}}(x-\mu_x)^2\, dydx}{\int_{x_{M,inf}}^{x_{M,sup}}\int_{y_{M,inf}}^{y_{M,sup}} dydx} \qquad (21)$$

where $x_{M,inf}, x_{M,sup}, y_{M,inf}, y_{M,sup}$ are the extreme points of the ellipse after trimming.

Following similar derivations as for (15) we obtain the integration limits on x :

$$x_{M,sup}, x_{M,inf}=\mu_x\pm\frac{\sqrt{b(1-\alpha_M)}}{ab-c^2} \qquad (22)$$

Similarly with (16) we derive the integration limits on y. We obtain the area of the ellipse in this case as :

$$\int_{x_{M,inf}}^{x_{M,sup}}\int_{y_{M,inf}}^{y_{M,sup}} dydx=\frac{(1-\alpha_M)\pi}{\sqrt{ab-c^2}} \qquad (23)$$

For the numerator in expression (21) we obtain :

$$\int_{x_{M,inf}}^{x_{M,sup}}\int_{y_{M,inf}}^{y_{M,sup}}(x-\mu_x)^2\, dydx=$$

$$\frac{(1-\alpha_M)^2 b}{\sqrt{(ab-c^2)^3}}\int_{-1}^{1} z^2\sqrt{1-z^2}dz=\frac{(1-\alpha_M)^2\pi b}{4\sqrt{(ab-c^2)^3}} \qquad (24)$$

From (23) and (24) we derive the estimate for the variance :

$$E[\hat{\sigma}_x^2] = \frac{(1-\alpha_M)b}{4(ab-c^2)} \quad (25)$$

We get a similar expression for the cross-correlation. In the expression (25), if we consider $\alpha_M = 0$ we obtain the same result as in [4] for estimating the normalized second order moment of an ellipse. In order to correct this bias, when estimating ellipsis covariance matrix we use the following expression, instead of (9) :

$$\hat{\mathbf{\Sigma}}_j = \frac{\sum_{i=0}^{N_k - 2\alpha_M N_k} (\mathbf{X}_{(i),M} - \hat{\mu}_j)' (\mathbf{X}_{(i),M} - \hat{\mu}_j)}{(1-\alpha_M)(N_k - 2\alpha_M N_k)}. \quad (26)$$

5. SIMULATION RESULTS

We have tested the α-trimmed mean RBF algorithm on artificial generated patterns. We consider three overlapping discs in a 512×512 image, without noise and with additive uniform distributed noise as shown in Figure 1 (a). We apply the classical learning algorithm for RBF, MRBF [2] and α-trimmed mean RBF algorithm. The results are displayed in Figures 1 (b), (c) and (d), respectively. We consider the coordinates of the white pixels as inputs of the network. The amount of uniform distributed noise accounts for 4 % of the total number of pixels. The noise effect is considered as switching the value of the pattern and background. As it can be observed in Figure 1, the initial pattern is deteriorated. The comparison measures are the average of the Euclidean distances between the centers of the discs and those of the estimated model, and the total number of erroneous pixels in the reconstructed model with respect to the noise-free image. From Table 1 we can observe that α-trimmed mean RBF algorithm provides better results than the other two algorithms in modeling the object. As it can be observed in Figure 1 the pattern corrupted by noise is better modeled by MRBF and α-trimmed mean RBF when compared to the classical statistics based approach.

We have applied the proposed algorithm on a stack of 3-D microscope images representing the inner structure of a tooth. In Figures 2 (a), (d) two frames of this stack of images are shown. The graylevel and the pixel coordinates have been used as inputs in the network. The graylevel accounts for the object segmentation while the coordinates account for the object localization. Both, the blood vessels and the background are modeled by RBF or α-trimmed mean RBF functions. The result of the segmentation using the RBF, based on moving average and classical variance is shown in Figures 2 (b), (e). In Figures 2 (c), (f) the result provided by α-trimmed mean RBF algorithm is displayed. The blood vessels are quite well segmented in the results provided by the α-trimmed mean RBF algorithm.

6. CONCLUSION

In this study we introduce a new algorithm for estimating the parameters of hidden units in RBF networks. The proposed algorithm employs the α-trimmed mean statistics for calculating the hidden unit parameters. The number of data samples trimmed away in the proposed algorithm depends on the distribution. Ellipse parameter estimation using the α-trimmed mean RBF algorithm is analyzed. The proposed algorithm is applied in image modeling and segmentation.

7. REFERENCES

[1] L. Xu, A. Krzyzak, E. Oja, "Rival penalized competitive learning for clustering analysis, RBF net, and curve detection," *IEEE Trans. on Neural Networks,* vol. 4, no. 4, pp. 636-649, 1993.

[2] A. G. Borş, I. Pitas, "Median radial basis function neural network," *IEEE Trans. on Neural Networks,* vol. 7 , no. 6 , pp. 1351-1364 , 1996.

[3] T. Kimoto, Y. Yasuda, "Shape description and representation by ellipsoids," *Signal Proc.: Image Comm.,* vol. 9, no. 3, pp. 275-290, Mar. 1997.

[4] R.M. Haralick, L.G. Shapiro. *Computer and Robot Vision,* vol. I, Reading, MA:Addison-Wesley, 1992.

[5] A. G. Borş, I. Pitas, "Moving scene segmentation using median radial basis function network," *Proc. IEEE Inter. Symp. on Circuits and Systems (ISCAS'97),* vol. I, Hong Kong, pp. 529-532, 1997.

[6] I. Pitas, A. N. Venetsanopoulos. *Nonlinear Digital Filters, Principles and Applications.* Norwell, MA: Kluwer, 1989.

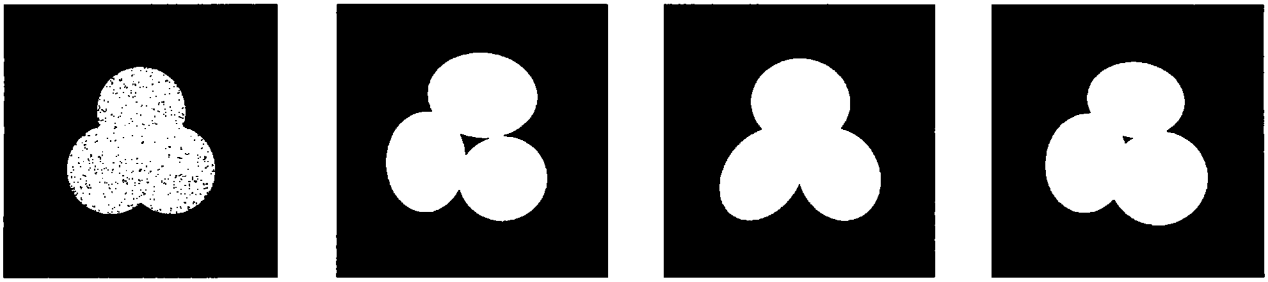
[7] S.R. Peterson, Y.-H. Lee, S.A. Kassam, "Some statistical properties of Alpha-Trimmed mean and standard type M Filters," *IEEE Trans. on Acoust., Speech, Sig. Proc.,* vol. 36, no. 5, pp. 707-713, 1988.

[8] P. Prescott, "Selection of trimming proportions for robust adaptive trimmed means," *J. American Stat. Assoc.,* no. 361, vol. 73, pp. 133-140, 1978.

[9] D. M. Titterington, "Estimation of correlation coefficients by ellipsoidal trimming," *Applied Statistics,* no. 3, vol. 27, pp. 227-234, 1978.

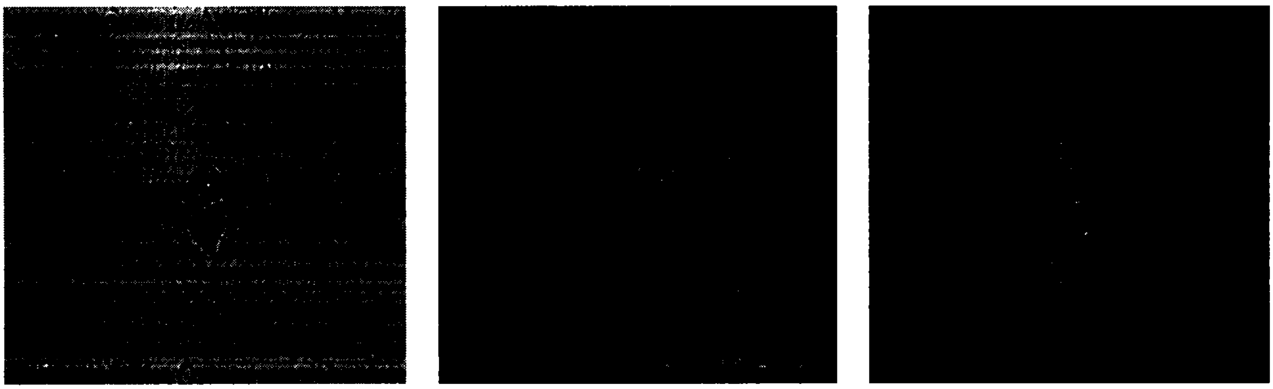
Algorithm	Noise-free Model		Noisy Model (4 % noise)	
	Center Location Estimation	Modeling Error (%)	Center Location Estimation	Modeling Error (%)
RBF	134.23	5.23	620.47	28.11
MRBF	162.40	13.72	339.17	16.22
α-trimmed **RBF**	129.49	4.84	356.32	15.18

Table 1. Numerical comparison when modeling an artificial pattern.

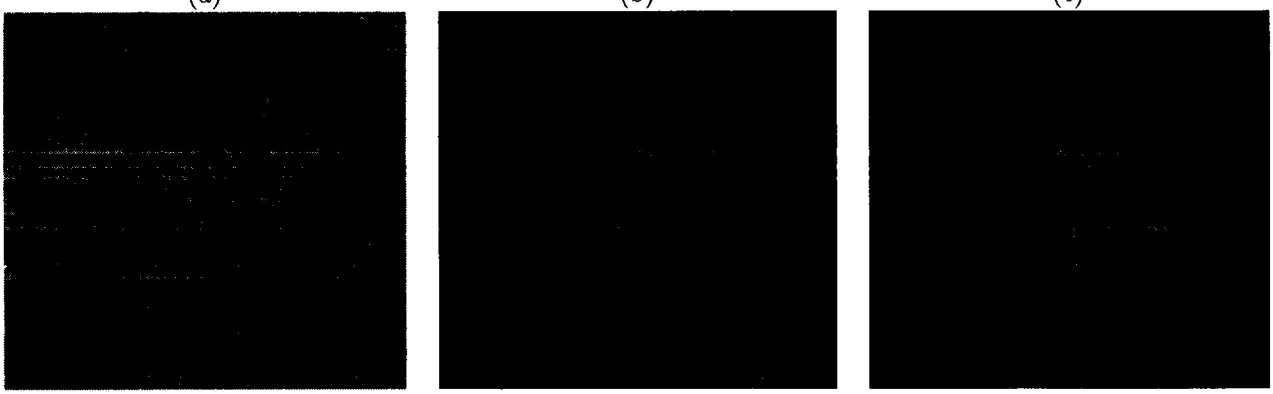


(a) Artificial generated image corrupted by noise. (b) Modeling provided by classical RBF algorithm (c) Modeling provided by MRBF (d) Modeling provided by α-trimmed mean RBF.

Figure 1. Comparative algorithm performance in modeling an artificial pattern.



(a) (b) (c)



(d) (e) (f)

Figure 2. Segmentation of blood vessels in a set of microscope images representing inner structure of a tooth : (a), (d) original microscope images ; (b), (e) segmentation provided by classical RBF; (c), (f) segmentation provided by α-trimmed mean algorithm.