Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 51(65), Fascicola 1, 2006

An Analog Computing Circuit for SVM classifiers

Lelia Feştilă[1], Lorant Andras Szolga[2], Mihaela Cîrlugea[3], Mihaela Gordan[4]

Abstract – We propose an analog nonlinear current-mode circuit for computing the decision function in a SVM classifier based on radial basis kernels. The validity of design and operation was proved by simulations.

Keywords: SVM, Euclidian distance, analog computing, classifier

I. INTRODUCTION

Support vector machines (SVM) combine many approaches of artificial intelligence and neural networks. Mathematically they are based on statistical learning theory and are especially suited for adaptive object detection and identification with sparse training data [1]. The SVM classification goal is to assign an object to a class that contain similar objects. The SVM is trained with a set of positive and negative labeled examples. Relevant object features are extracted and retained in a set of support vectors.
The automata's response is a decision function that is expressed by a linear combination of kernels whose argument depends on both support and test vectors. The most kernels are based on internal products for polynomial classifications and for the multilayer perceptions [1] [8].

Other categories of classifiers are based on the radial basis functions that depend on the Euclidian distance (ED) between two vectors $\mathbf{X_m}$ and $\mathbf{X}$:

$$D_m\left(X, X_m\right) = \left\|X_m - X\right\| = \left(\sum_{i=1}^{N}\left|X_{mi} - X_i\right|^2\right)^{1/2} \qquad (1)$$

The kernel **K(•,•)** of such a classifier [7] can be linear (proportioal to D_m), a smooth limiter or a Gaussian function:

$$K\left(\left\|X - X_m\right\|, c\right) = ke^{\left(-\frac{\|X - X_m\|^2}{c}\right)} \qquad (2)$$

$\mathbf{X}$ is the input vector and $\mathbf{X_m}$ the reference vector which in SVMs, is called support vector SV. Coefficient **c** is the width of the kernel. The classifier computes the decision function :

$$y = sign\left[\sum y_m \alpha_m K\left(\left\|X - X_m\right\|, c\right) - b\right] \qquad (3)$$

where α_m are synaptic weights and $\mathbf{y_m}$, the labels +1 and –1 for positive and negative SVMs respectively and **b** is the bias term.
The vector X belongs to the class represented by positive defined SVs if its resulted label $y \geq 0$.
Although the number of SVs is usually much smaller than the number of training examples because of large dimensions and large degree of variability in the object class, an excessive amount of computation can appear in both learning and classifying process. The computational time is dominated by the calculation of kernels, so that computing time of large dimension nonlinear functions is needed. Therefore implementing such classifiers in hardware becomes a good alternative to software implementations especially in the case of real-time operation. Analog devices are recommended in such cases because of their run-time performance. They can perform very fast linear and nonlinear mathematical operations that are expensive in digital domain (multiplication, division, square-root, squaring, geometric mean). Current-mode CMOS transliniar circuits have also many other advantages like low voltage, low power aned simple cell structures requiering a small area.
In order to be able to study, analyze and develop different VLSI structures for analog computing with applications in SVM classifiers we created a Pspice environment containing a basic cell library and also some computing blocks for nonlinear functions like ED or different types of kernels.
We designed and simulated an analog computing circuit for parallel SVM image classifiers. The analog hardware is used only to classify unknown images.

[1] Facultatea de Electronică , Telecomunicaţii şi Tehnologia Informaţiei, catedra Bazele electronicii, Str.G.Barițiu 26-28, 400027 Cluj-Napoca, e-mail Lelia.Feştilă@bel.utcluj.ro
[2] Facultatea de Electronică , Telecomunicaţii şi Tehnologia Informaţiei, catedra Bazele electronicii, Str.G.Barițiu 26-28, 400027 Cluj-Napoca, e-mail Szolga.Lorant@bel.utcluj.ro
[3] Facultatea de Electronică , Telecomunicaţii şi Tehnologia Informaţiei, catedra Bazele electronicii, Str.G.Barițiu 26-28, 400027 Cluj-Napoca, e-mail Mihaela.Cirlugea@bel.utcluj.ro
[4] Facultatea de Electronică , Telecomunicaţii şi Tehnologia Informaţiei, catedra Bazele electronicii, Str.G.Barițiu 26-28, 400027 Cluj-Napoca, e-mail Mihaela.Gordan@bel.utcluj.ro

Training is performed separately using a dedicated software [8].

The presented circuit is destined to SVM classifiers based on radial basis kernels. As we know, the existing hardware implementations are mainly based on polynomial kernels. Radial basis kernels could only be approximately implemented with these types of structures or need extra look-up tables. The proposed circuit computes the ED, the nonlinear kernel and multiplies it by label and weight.

The ED calculator is simpler as the one found in literature [3] containing only 2N+1 cells/template not 3N+1 cells/template as in [3]. Using an extra multiplier for the weight α we could extend the multiplier linearity domain over the one obtained by bias controling of a smooth limiter as in [6].

By simulations we proved the feasability and validity of the algorithm and design and set appropriate parameters for a good functionality of each component and also for the whole operating chain. Chapter 2 describes the algorithm for ED calculation and determination of the decision function while in chapter 3 we present the circuit. Simulation results are given. The circuit functionality was checked on a small dimension image detection problem, presented in chapter 4.

II. ALGORITHM OF THE PROPOSED CLASSIFIER

A radial basis classifier circuit compares an unknown pattern **X** with a given set of **M** predefined patterns $\mathbf{X_m}$ (m= 1,...,M), called support vectors SVs and calculates distance ED denoted $D(X,X_m)$ between **X** and each VS prototype ($\mathbf{X_m}$). The distance (1) is:

$$D_m(X, X_m) = \sqrt{(X_1 - X_{m1})^2 + ... + (X_N - X_{mN})^2} \quad (4)$$

where m=1÷M is the number of SVs,
$X=\{X_1,X_2,...,X_N\}$, $X_m=\{X_{m1},X_{m2},...,X_{mN}\}$, N is the number of features, in our examples, the number of pixels.

The classifier based on ED implements (4) by doing the difference between X and X_m, squaring each result, summing the terms and finally applying the square root. In our circuit kernel K is an inverting smooth limiter with controllable gain and limits. Finally function f is calculated to obtain the decision y:

$$f(\|X - X_m\|) = \sum y_m \alpha_m K_m(\|X - X_m\|) \quad (5)$$

$$y = sign[(\|X - X_m\|) - b] \quad (6)$$

III. ANALOG PARALLEL ARCHITECTURE AND BUILDING BLOCKS CELLS

Fig.1 shows the general block diagram of a classifier with parallel architecture.

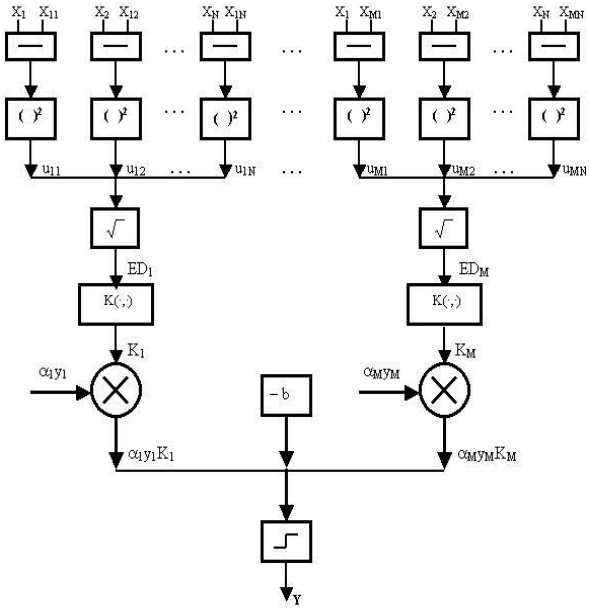


Figure 1. General block diagram of the classifier

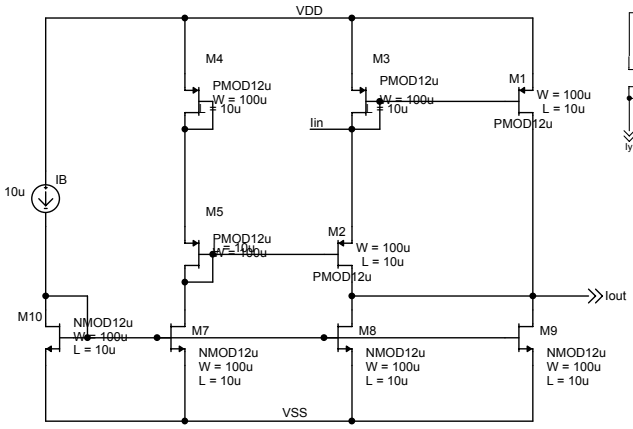
Each path permits the calculation of one term $y_m\alpha_m K_m(||X-X_m||)$ in the sum (5). For parallel current mode architectures these terms are currents and can be simply summed by a direct connection as Fig.1 shows.

The ED calculator contains N substractors, N squarers and a square-rooter/template. The substractors make simply current differences by directly connecting outputs of current mirrors and inverting current mirrors [2].

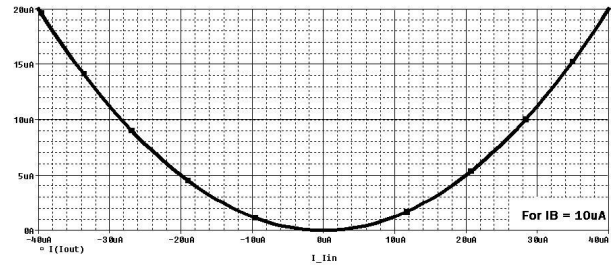
Fig.2.a shows the circuit used for realizing the square of a bidirectional current [10]. Its output is of the form:

$$I_{out} = \frac{i_{in}^2}{8I_B} \quad (7)$$

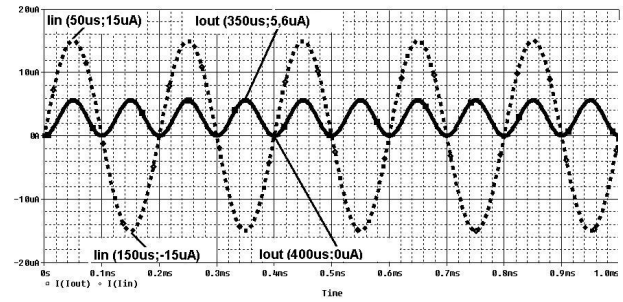
By simulations we proved the validity of the circuit, determined the optimal biasing current I_B and checked the theoretical conditions [2] $\|i_{in}\| \le 4I_B$. Figures 2.b and 2.c show examples of simulation.



a)

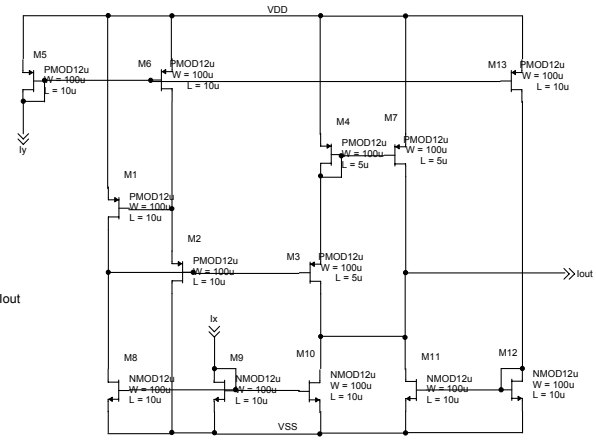


b)

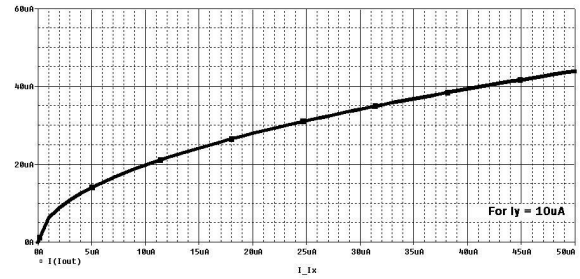


c)

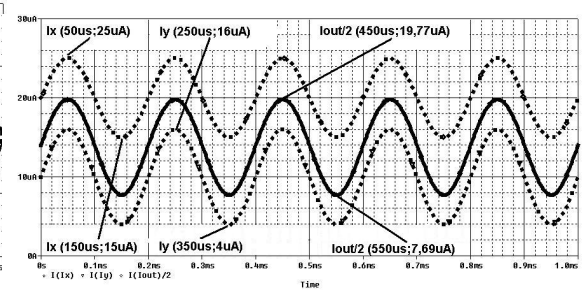
Figure 2. Squaring block
- a) circuit schematic;
- b) in/out transfer characteristic;
- c) time diagram.



a)



b)



c)

Figure 3. Geometric mean block
- a) circuit schematic;
- b) in/out transfer characteristic;
- c) time diagram.

Fig.3.a represents the schematic of a geometric mean circuit [3]. One can calculate [2] and prove by simulations that the output current of this circuit is:

$$I_{out} = 2\sqrt{I_x I_y} \qquad (10)$$

The simulations are shown in Figures 3.b and 3.c for input currents I_X, I_Y(dots) and $I_{out}/2$ the half of output current that is the geometric mean of I_X and I_Y. In the classifier schematic we use this block for square-rooting the sum of squares and set $I_x = I_B$.

The smooth limiter function close to a sigmoidal form is realized with a differential CMOS inverting amplifier (DA).

Fig.4.a and Fig.4.b shows the input-output characteristics for different biasing currents I_B. Because input signals supplied by the square-rooter are positive only the fourth quadrant will be used.

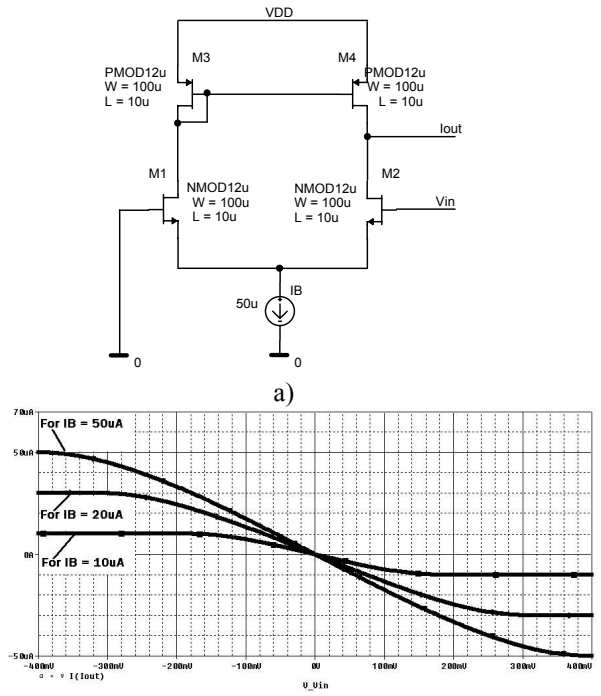


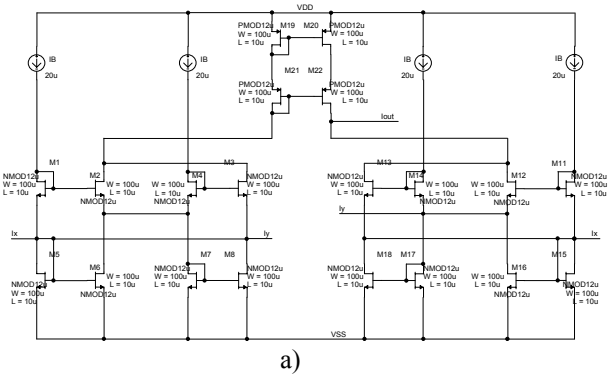
Figure 4. Differential amplifier
a) circuit schematic;
b) in/out transfer characteristic.

The multiplication by $\pm\alpha$, resulted from a software training, is performed by a multiplier shown in Fig.5. It is a four quadrant multiplier used to realize product $\boldsymbol{\alpha}_m \mathbf{y}_m \mathbf{K}$. One can calculate that the output is [2]:

$$I_{out} = \frac{I_x I_y}{2I_B} \qquad (11)$$

The performed simulations in Fig.5.b and c proved this result.

We chosed this variant with a separate multiplying block instead to simply control or set the biasing current of the DA beacause for certain applications the weights $\boldsymbol{\alpha}$ can vary in a large domain and labels and $\boldsymbol{y}$ can be positive or negative. By using a multiplier the domain of permitted $\boldsymbol{\alpha}$ values was considerable increased because of the good linearity domain of the multiplier (Fig.6.b)



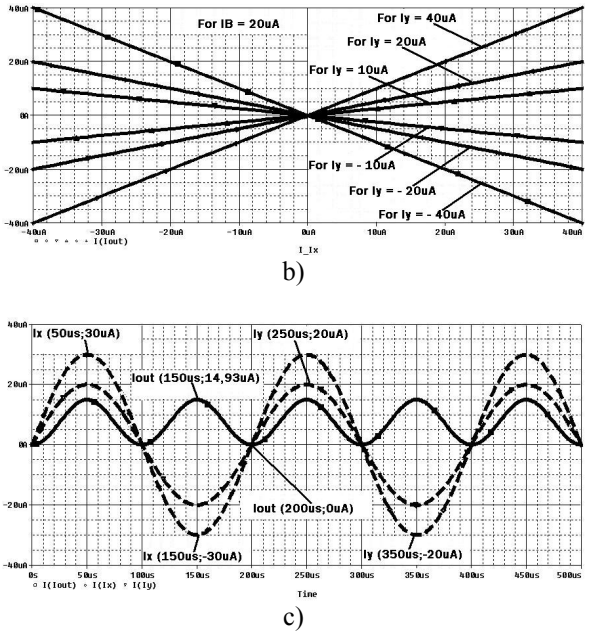


Figure 5. Four-quadrant current multiplier
a) circuit schematic;
b) in/out transfer characteristic;
c) time diagram.

If the distances $\|X-X_m\|$ are calculated in parallel as Fig.1 shows, the circuit is fast and does not need a sumator accumulator. For area reasons euclidian distances can be also sequentially calculated, but with a lower run-time.[4]

Four paths of the whole simulated structure of the classifier are shown in Fig.6.

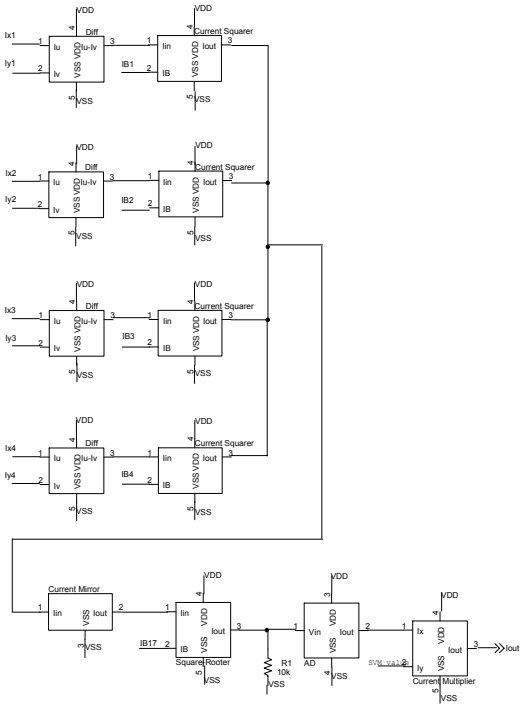


Figure 6. A part of the simulated circuit

IV. EXPERIMENTAL RESULTS

We took an example of a binary pattern classifier. Each template is a 4x4 black and white image. For a black quadrant, the input current is 25μA and for a white one 15μA. For each pair **(X,X_m)** we used 16 circuit paths, each of them containing modules for difference and power. The 16 output currents are summed and enter the square-root-circuit (Fig.6). A current mirror is also needed for adapting impedances. The SVM was trained separately by soft for identifying a diagonal in a 4x4 image. Some experiences needed to determine the proper number of positive and negative training vectors. The best result was obtained for 28 positive and 29 negative examples. The learning process resulted in 22 support vectors, their labels and weights Fig.7 and Table I.

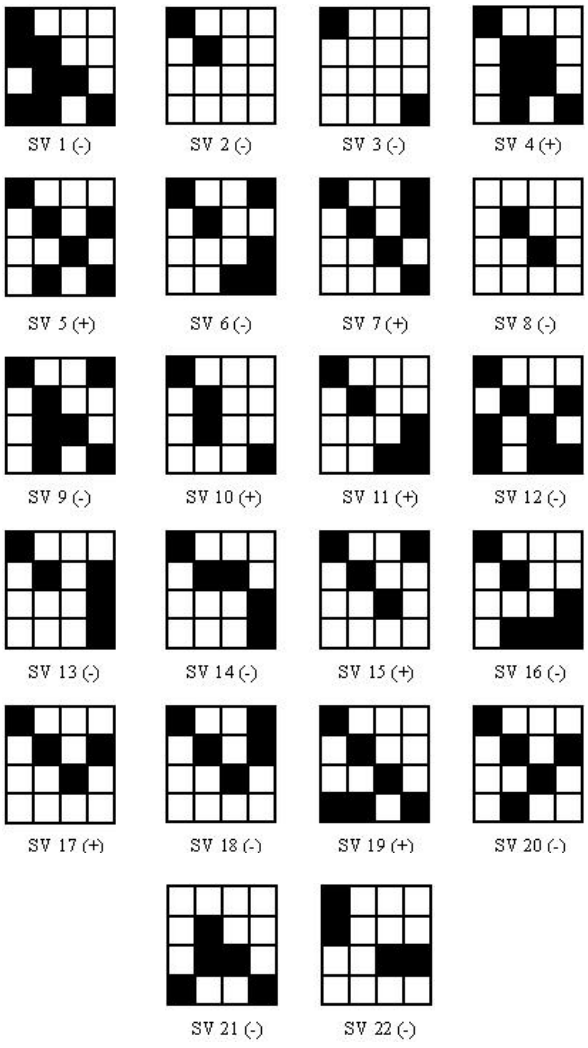


Figure 7. Predefined pattern class (SVs)

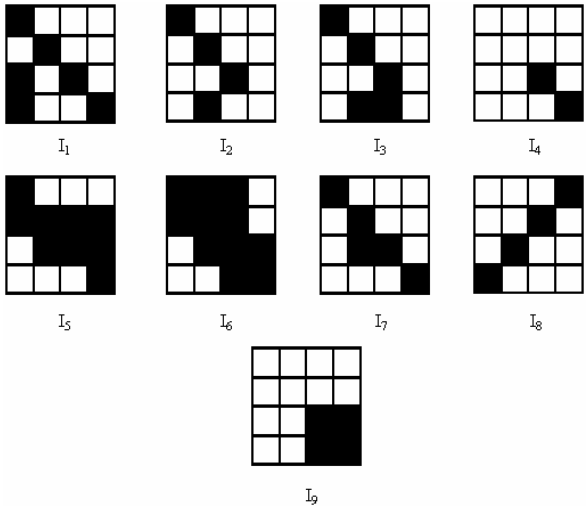


Figure 8. Patterns to be classified

As example nine images $I_1 \ldots I_9$ to be classified are given in Fig.8.

TABLE I
TRAINING WEIGHTS AND LABELS

SV	$\alpha_i y_i$	SV	$\alpha_i y_i$
SV1	- 3,57	SV12	- 8,78
SV2	- 17,79	SV13	- 6,28
SV3	- 8,18	SV14	- 9,08
SV4	8,15	SV15	37,97
SV5	4,83	SV16	- 5,81
SV6	- 11,17	SV17	32,86
SV7	19,62	SV18	- 42,89
SV8	- 12,56	SV19	9,23
SV9	- 16,11	SV20	- 8,53
SV10	13,7	SV21	- 3,58
SV11	28,11	SV22	- 0,12

TABLE II
"ED" AND "DA" OUTPUT CURRENTS

Difference	ED [μA]	DA [μA]
1	12,97	- 22,42
2	15,48	- 26,43
3	17,74	- 29,91
4	20,02	- 33,24
5	22,33	- 36,41
6	24,42	- 39,09
7	26,34	- 41,37
8	28,42	- 43,3
9	29,82	- 44,94
10	31,41	- 46,32

Table II shows the measured currents at the outputs of the ED and DA blocks in function of the number of differences between colours of template pixels.
We had to use a current multiplier to obtain $\alpha_i y_i \mathbf{I}_k (y_i = \pm 1)$. The two inputs are current $\mathbf{I}_k$ from the DA output and a current $\mathbf{I}_\alpha$ proportional to $\alpha_i y_i$. To

choose the right current multiplier, first of all we observed and analyzed the limits of these currents obtained for SVM coefficients $\alpha_i y_i$ and from the AD output I_k.

From TABLE I and TABLE II one can observe that:

$I_\alpha \sim \alpha_i y_i \in [-45,14; 46,15]$ and

$I_K \in [-43,3\mu A; -22,42\mu A]$, so a current multiplier which can accept both negative and positive currents for the input is needed. The four-quadrant current multiplier from Fig.5.a was ideal for us and established the best working conditions for this circuit in our implementation. For a good evaluation of this conditions we considered current domains:

$I_\alpha \in [-50\mu A; 50\mu A]$ and $I_K \in [-50\mu A; 0\mu A]$.

With a simple correlation between the input currents of the multiplier and the above currents, we set:

$I_x = I_K$ and $I_y = I_\alpha$.

One can prove that the proper operating conditions for this multiplier circuit are:

$(I_x + I_y) \in [-4I_B; 4I_B]$ and $(I_x - I_y) \in [-4I_B; 4I_B]$.

In our case $(I_K + I_\alpha), (I_K - I_\alpha) \in [-100\mu A; 50\mu A]$. With this last condition we can set the minimum value for I_B so that the circuit should work properly, value which was set to $I_B = 25\mu A$.

The threshold **b**, depends on the scale factor, different scale factors have been analyzed, by considering different reference currents I_B of multiplier.

TABLE III
EXPERIMENTAL RESULTS

Image	I_{out} [µA]		Classified	Diff	Threshold b [µA]
	for I_B= 25µA	for I_B= 50µA			
I_1	5,78	2,98	+	2	3,2 for I_B=25µA
I_2	3,93	1,96	+	2	
I_3	4,19	2,09	+	3	
I_4	1,533	0,76	–	2	
I_5	3,01	1,5	–	5	
I_6	3,22	1,61	+	7	
I_7	7,18	3,59	+	1	
I_8	- 1,56	- 0,78	–	8	
I_9	3,16	1,58	–	4	

Table III shows two final results: we used two values for the I_B currents. The output currents $\mathbf{I_{out}}$ corresponds to function ***f***, relation (5), for each image to be classified. With I_B=50µA we can choose the threshold value much better, because the interval between the smallest score for a positive classified image and the bigger one for a negative image is smaller than for I_B=25µA. So we set the threshold b=1,6µA, for the circuit having I_B=50µA. To have a bigger threshold current we are going to use the I_B=25µA case, multiplying the previously obtained threshold by 2. The final threshold value in this case is b=3,2µA.

In order to have a comparative view over the SVM classifiers and a classifier based only on ED, the circuit has also been used for calculating ED between each input vector under test and the pattern corresponding to a pure diagonal image. The decision in the ED method is only based on the minimum distance that in fact corresponds to the minimum number of different pixels in the compared images.

Therefore the ED could be the same for two images that belong to different classes. As example $\mathbf{I_1,I_2,I_4}$ have the same ED and are classified in the same class by an ED classifier. SVM classifier takes better decisions, as example, although $\mathbf{I_1,I_2,I_4}$ have the same ED, $\mathbf{I_4}$ is negative classified and $\mathbf{I_1,I_2}$ positive being closer to the diagonal. One can see in Table III that the decision functions of the SVM classifier for these images are very different. The image can be detected correctly by a well trained SVM classifier even if dots or other spots exists.

V. CONCLUSION

We have designed a current-mode analog computing CMOS circuit for SVM parallel classifiers based on radial basis functions. It consists of an ED computing stage and a circuit for realizing the kernel required by the nonlinear transformation with controllable weights. Simulations proved the validity of this circuit and of the SVM classifying application.

Using the SVM algorithm, the classifying task is by far improved in comparison to the method based only by appreciating the ED between two patterns [5][6].

REFERENCES

[1] R. Genov, S. Chakrabartty, G. Cauwenberghs – *Silicon SVM with on Line Learning*, J.J.of Pattern Recognition and Artificial Intelligent Systems, Darjon Korea, pp 43-48, 2002

[2] Lelia Feştilă – *Analog Integrated Circuits Translinear Networks*, Ed. UTPRES, Cluj N, 2003

[3] Bin – Du Lin, Chnen Yan Chen, Yu-Ying Teao – A Modular Current – Mode Classifier Circuit for Template Matching Application, IEEE Trans. On CSS- II, vol 47, No 2, 2000 pp 145-150

[4] A. Fazakas, M. Gordan, L. Feştilă, Laura Kovacs – *Considerations Regarding Implementation of SVM Classification Functions on FPGA*, Proceedings on International Symposium for Design and Technology of electronic Packages, Cluj-Napoca, 2005, pp. 164-167

[5] Sebastian Berindean – *Studiul şi proiectarea unor circuite neliniare VLSI utilizate în sistemele integrate de calcul analogi*c, Proiect de diplomă, UT Cluj-Napoca, 2005

[6] Mihaela Cîrlugea, Albert Fazakas, Robert Groza: Analog Clasiifier Based on Euclidian Distance, Inter-Ing 2005, Targu Mures, Romania 2005, pp 685-690

[7] A Chihocki, R Unbehauen: *Neural Networks for Optimization and Signal Processing*, J. W. Wiley, Stuttgart, 1993

[8] Mihaela Gordan, *Doctoral Thesis*, *Applications of the fuzzy technics and the machines with vector support in the image processing* Aplicatii ale tehnicilor fuzzy si masinilor cu vectori suport in prelucrarea imaginilor. Teza de doctorat UTCluj, 2004

[9] K.J. Wiegering – *Analysis and Synthesis of MOS Translinear circuits*, Kluwer Academic Bublication,1993

[10] Liong-Hung Lu, Chung-Yu, Wu-Thu, - *Design of the CMOS Current Mode Semnal-Purpose Analog Processor*, Proc of the 1997 IEEE Int. Symp. On C&S, Hong-Kong (549-552)