# Circuital implementation of support vector machines

D. Anguita, S. Ridella and S. Rovetta

A circuital implementation of the learning phase of a new model of artificial neural network, the support vector machine, is proposed. It is shown that a Hopfield-type recurrent network can be easily implemented for this purpose.

*Introduction:* Support vector machines (SVMs) are among the newest and most promising models of artificial neural networks. SVMs show remarkable properties in terms of generalisation ability, and their performance is often superior to more traditional architectures such as multilayer perceptrons and radial basis function networks [1, 2]. Learning in SVMs is equivalent to solving a large constrained quadratic programming (CQP) problem, where the number of variables equals the number of training patterns.

Up to now, no hardware implementation of SVM has been proposed; we show in this Letter how a recurrent network can be used to solve the CQP problem associated with an SVM and, therefore, allow a circuital implementation of SVMs.

*Learning in support vector machines:* We will focus here on linear SVMs; in this case the resulting network is equivalent to a perceptron. Extensions to nonlinear SVMs, making use of radial basis functions or sigmoidal functions, do not affect the learning phase, as presented here, and can be found in [1].

Given a set of two-class labelled training patterns $(\underline{x}_1, y_1)$, ..., $(\underline{x}_n, y_n)$, with $\underline{x}_i \in \mathbf{R}^m$ and $y_i = \{+1, -1\}$ , it can be shown that learning in SVM is equivalent to solving the following CQP problem:

$$\min \Phi(\underline{v}) = \frac{1}{2}\underline{v}^T Q \underline{v} + \underline{r}^T \underline{v} \tag{1}$$

$$\sum_{i=1}^{n} y_i v_i = 0 \tag{2}$$

$$0 \le v_i \le u \quad \forall i \tag{3}$$

where $Q$ is a symmetric $n \times n$ matrix with $q_{ij} = y_i y_j \underline{x}_i^T \underline{x}_j$, $r_i = -1\ \forall\, i$ and $u$ is a constant, defined *a priori*. The upper bound $u$ limits the magnitude of the solution $\underline{v}^*$ when the two classes cannot be separated without errors; in fact, in this case, $v_i^* \to +\infty$ for some $i$. The upper bound forces the network to tolerate some errors; as it is decreased, the tolerance to incorrect patterns increases [2].

We write the equality constraint as two inequality constraints, and eqns. 2 and 3 in matrix form [3]:

$$\underline{f}(\underline{v}) = B\underline{v} - \underline{e} \ge 0 \tag{4}$$

where $e_i = \{0, -u\}$. $B$ is a $(2+2n) \times n$ matrix of the form

$$B^T = \lfloor -\underline{y} \quad \underline{y} \quad -I \quad I \rfloor \tag{5}$$

where $I$ is an $n \times n$ identity matrix.

After learning, the patterns corresponding to $v_i \neq 0$ are called 'support vectors'. They are the border patterns that separate the two classes, and from which the weights and the bias of a perceptron can be derived, using the following relations:

$$\underline{w} = \sum_{i=1}^{n_s} y_i v_i \underline{x}_i \tag{6}$$

$$b = \frac{1}{n_s}\sum_{i=1}^{n_s}(y_i - \underline{w}^T \underline{x}_i) \tag{7}$$

where $n_s$ is the number of support vectors.

The perceptron found through the SVM shows remarkable generalisation properties. As an example, its average probability of error does not depend on the dimensionality of the problem, but only on the number of support vectors [1].

*Recurrent neural network:* We propose to use Chua's recurrent network (detailed in [3]) to solve the CQP problem described in the previous Section. Fig. 1 shows an example of this network. The neurons of the network are of two kinds; those indicated by $I$ are simple integrators, while neurons indicated by $g$ implement the penalty function necessary to satisfy eqn. 4:

$$g(v) = \begin{cases} 0 & v > 0 \\ Gv & v \le 0 \end{cases} \tag{8}$$

where $G$ is the conductance of the neuron. For $G \to \infty$, the penalty function ensures that the solution satisfies the constraints as desired. A simple example of neuron implementation is proposed in [3]; more complex but, at the same time, more efficient circuital solutions can be found in [4].

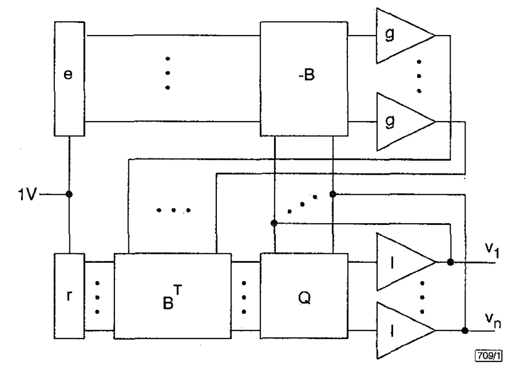


**Fig. 1** *Recurrent neural network implementing the learning phase of support vector machine*

The mapping of the CQP problem on the recurrent network is straightforward, as indicated in Fig. 1. Every non-zero element of the corresponding matrix, or vector, represents a resistive connection. Note that the values must be interpreted as conductances.

The number of connections is greatly reduced, with respect to Chua's general version, thanks to the particular structure of eqn. 4. As an example, the connection matrix $B$ has only $4n$ non-zero entries, compared to a total of $2(n^2+n)$ elements.

The evolution of the recurrent network is described by the following equation:

$$C_i \frac{dv_i}{dt} = -\frac{\partial \Phi}{\partial v_i} - \sum_{j=1}^{2n+2} g(f_j(\underline{v}))\frac{\partial f_j}{\partial v_i} \tag{9}$$

where $C_i$ is the self-capacity of the integrators. Its convergence has recently been proved [5], showing that the stable points of eqn. 9 satisfy the Kuhn-Tucker conditions of the associated optimisation problem.

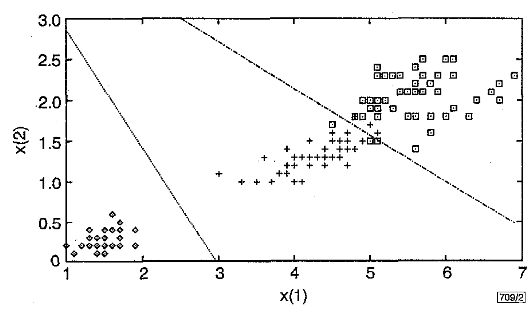


**Fig. 2** *Linear separators found by two SVMs on IRIS data-set*

·········· SVM1
– · – · SVM2

*Experimental results:* We have used, for our experiment, the well-known three-class IRIS dataset (Fig. 2). The first two classes can be linearly separated without errors, while there is no exact linear separation between the second and the third. The experiments were performed with a conductance $G = 10^3\Omega^{-1}$ and the upper bound for the non-separable case was set to $u = 15$. We chose $v_i = 0\ \forall\, i$ as the starting point for the recurrent network; this is both a feasible point of the CQP problem and an unstable state of the network ($\nabla\Phi(0) \neq 0$), and therefore it can always be used as a starting point for the solution search.

Two SVMs were implemented: SVM1 finds a linear separator between the first two classes, while SVM2 deals with the second and third classes. In both cases, as predicted by theory, the SVMs find discriminating functions with good generalisation properties. Note, for example, that the separator found by SVM1 is maximally distant from each class.

*Conclusions:* We have proposed a circuital implementation of the learning phase of support vector machines through the use of a recurrent network. This is the first attempt to realise a hardware implementation of SVMs. Our experiment confirms that the equilibrium point of the network coincides, as predicted by theory, with the solution of the associated optimisation problem.


28 May 1998
*Electronics Letters Online No: 19981092*

D. Anguita, S. Ridella and S. Rovetta (*University of Genova, Department of Biophysical and Electronic Engineering, Via Opera Pia 11A, 16145 Genova, Italy*)

**References**

1 VAPNIK, V.N.: 'The nature of statistical learning theory' (Springer-Verlag, 1995)
2 VAPNIK, V.N., GOLOWICH, S., and SMOLA, A.: 'Support vector method for function approximation, regression estimation, and signal processing', *Adv. Neural Inf. Process. Syst.*, 1997, **9**, pp. 281–287
3 KENNEDY, M.P., and CHUA, L.O.: 'Neural networks for nonlinear programming', *IEEE Trans. Circuits Syst.*, 1988, **CAS-35**, (5), pp. 554–562
4 CHEN, J., SHANBLATT, M.A., and MAA, C.: 'Improved neural networks for linear and nonlinear programming', *Int. J. Neural Syst.*, 1992, **2**, (4), pp. 331–339
5 MAA, C., and SHANBLATT, M.A.: 'Linear and quadratic programming neural network analysis', *IEEE Trans. Neural Netw.*, 1992, **3**, (4), pp. 580–594

## 10Gbit/s transmission over 1500km with semiconductor optical amplifiers

G. Onishchukov, V. Lokhnygin, A. Shipulin and P. Riedel

The authors experimentally demonstrate, in a fibre-loop setup, 10Gbit/s return-to-zero propagation at 1.3μm over more than 1500km in a standard fibre using an optimised in-line semiconductor optical amplifier. System margins have been determined.

*Introduction:* Recently, the performance of 1.3μm optical fibre transmission lines with multiple quantum well (MQW) semiconductor optical amplifiers (SOAs) as in-line amplifiers has been intensively investigated. A 10Gbit/s return-to-zero (RZ) transmission over 420km was demonstrated in a straight-line test bed [1] and over 210km in the field trial [2]. It was also shown that the propagation distance could be increased by using an optimised SOA [3, 4].

The aim of the current work was the experimental investigation of the performance of such an optimised MQW-SOA in a recirculating fibre-loop setup. A considerable increase in the propagation distance has been demonstrated.

*Experimental setup:* The experimental setup was similar to that used in [5]. An 8:1 time multiplexer (a three-cascade Mach-Zehnder interferometer) was used to generate the most critical bit patterns by blocking appropriate multiplexer arms. 37.6km of standard singlemode fibre with mean zero dispersion wavelength of 1307.2nm was used in the fibre loop. An optimised, polarisation insensitive MQW-SOA was supplied by Philips Optoelectronics. It has an 18.5dB fibre-to-fibre gain and a noise figure of 7.6dB at 1310nm with peak gain of 21dB at 1285nm and 1.2dBm total ASE output power when driven at 300mA current. As the in-line optical filter, a 0.8nm bandwidth interference filter was used, which was superior to a 3nm filter and allowed us to abolish an additional receiver filter [6]. A 7.5GHz low-pass RF filter was used after the lightwave converter (HP 11982A).

Because of a strong amplitude pattern effect (inter-symbol interference), the usual Q-factor measurements (overlapping of all possible bits) did not provide any reasonable results for BER estimation due to an appreciable deviation of the noise distribution from Gaussian. To overcome this problem, we measured the Q-factor for each bit in a pattern separately, overlapping it with a long set of zeros. As the system parameter, the worst case (usually Q-factor of the fourth pulse in the '...00001111...' bit sequence) was considered.

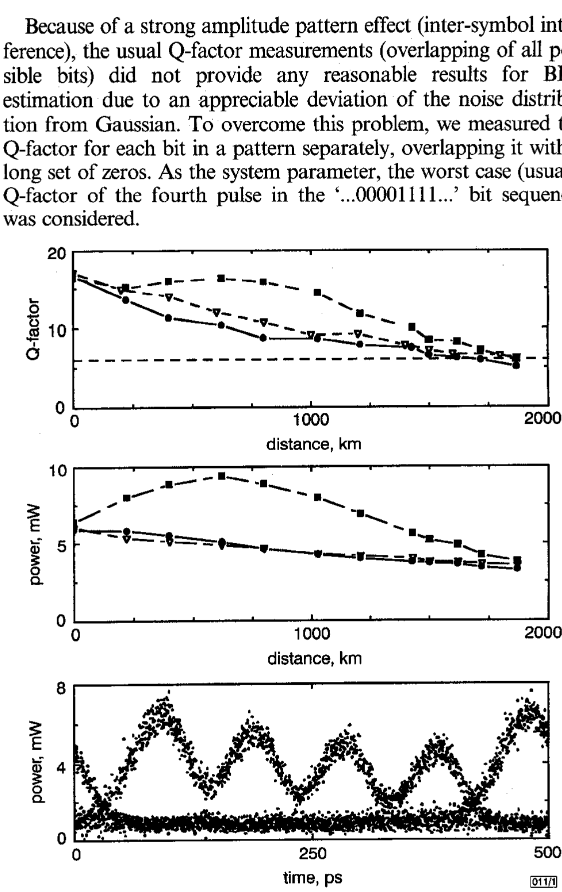


**Fig. 1** *Distance dependencies of Q-factors and pulse amplitudes at SOA output and eye-diagram at 1500km*

– – ■ – – first pulse in '...00001111...' bit sequence
——●—— fourth pulse in '...00001111...' bit sequence
– – ▽ – – regular '...00001111...' bit sequence

*Results and discussion:* Our results showed that the performance of the system under investigation is essentially governed by the dynamics of the gain saturation and recovery, and most of the specific features could be understood by taking into account amplitude pattern and spectral walk-off effects [7]. Fibre Kerr nonlinearity (self-phase modulation) does not really affect the system performance because it is much weaker than the phase modulation provided by gain saturation of the SOA. One reason for this behaviour is the large (~5dB) losses of the elements between the SOA output and the fibre (isolator, bandpass filter, tap coupler).

Better results have been obtained by using 30–40ps optical pulses rather than 20ps pulses as is usually assumed for 10Gbit/s soliton transmission lines. The recovery time of the SOA towards the small signal gain typically amounted to ~400ps. This was concluded from measurements of the time profile of ASE in a long set of zeros. Thus, the SOA gain recovery time is much longer than the pulse duration, and the gain saturation in SOA is determined by the pulse energy and SOA saturation energy, rather than by the corresponding power parameters. Compared with 20ps pulses of the same energy, longer 30–40ps pulses exhibit a smaller spectral walk-off; thus they are less affected by the in-line filter and yield better results.

The distance dependence of the amplitudes and Q-factors for the first and fourth pulses in a '...00001111...' bit sequence and for a regular 10Gbit/s pulse train are shown in Fig. 1. The low ASE power of the SOA used permitted us to work with a small positive (~0.1dB) net gain in the line. This, together with the initial shift of the signal wavelength to the blue edge of the filter, leads to the initial growth of the first signal pulse after a long set of zeros (in our case, four zeros). Simultaneously, the amplitudes of the other pulses in a set of '1's continuously decrease with distance due to the incomplete gain recovery between the subsequent pulses. Since the stronger pulse has a larger red shift of its frequency at long