Biol. Cybern. 66, 115–121 (1991)

Biological Cybernetics

The time-into-intensity-mapping network

W. Banzhaf and K. Kyuma

Central Research Laboratory, Mitsubishi Electric Corporation, 1-1, Tsukaguchi Honmachi 8-chome, Amagasaki, 661 Japan

Received June 4, 1991/Accepted June 26, 1991

Abstract. We employ a special combination of different networks in order to process (transient) spatio-temporal patterns. In a first layer, feature analyzing cells translate instantaneous spatial patterns into activities of cells symbolizing the presence of certain feature values. A second layer maps the time sequence of symbols into a spatial activity pattern of the so-called TIM-cells. A third layer recognizes predefined activity patterns. We demonstrate the behaviour of the network using gaussian patterns in $(1+1)$ space-time dimensions.

1 Introduction

The ability to process spatio-temporal patterns is ever crucial for living beings. Patterns which are changing dynamically over time constitute a major part of the information flux to be processed by biological organisms. Moreover, a capacity to recognize and reproduce transient trajectories must have been truly beneficial within the framework of evolution. It was only with these capabilities already developed that oral communication by speech could emerge between members of the early human community.

In this contribution we shall concentrate on the transient character of spatio-temporal patterns. We shall apply a network consisting of three sub-networks to the recognition of predefined spatio-temporal patterns. As we shall see, the network is not only able to recognize transient patterns by means of a special representation of temporal sequences, but also, due to its symmetrical design, to reproduce arbitrary sequences of spatial patterns.

The core of the architecture is a mapping of temporal information into the space domain by the time-intensity-mapping (TIM) subnetwork. It makes use of populations of short term memory cells in order to translate temporal correlations of patterns into spatial correlations. Similar ideas were published in different contexts in the past (Cohen and Grossberg 1987; Hecht-Nielsen 1987; Wang and Arbib 1990). We tried to overcome certain shortcomings in our design.

The temporal information which can be processed by the TIM subnetwork must consist of a sequence of symbols. Spatio-temporal patterns, however, are not usally just a sequence of symbols. Therefore, before entering the TIM-subnet, instantaneous spatial patterns have to be translated into activities of cells representing these symbols. In a previous paper (Banzhaf 1991), each of these cells were grandmother cells, responsible for one out of a finite number of patterns. This strategy was suitable for an unbound pattern space with a small and finite number of patterns. For a bound pattern space with an infinite number of spatial patterns, however, another strategy must be adopted. It consists of applying feature analyzing cells which independently analyze different kinds of features present in the input pattern. The results of this analysis are relayed to corresponding TIM cells. After mapping into firing intensities of TIM-cells, a set of grandmother cells is used to recognize certain intensity distributions among the TIM cells as different spatio-temporal patterns.

The TIM-cells themselves are organized in populations. Entire populations are weakly coupled to individual feature analyzing cells of the first layer. At certain sampling times, the feature analyzing cells are forced to compete for the presented input patterns. Their decision which comes about as the result of their competition is constantly monitored by the array of TIM-populations. The activity of winning feature analyzing cells is sufficient to trigger an individual cell within a corresponding TIM-population. Due to a positive self-coupling this starts an exponentially growing activity of TIM-cells. Activity evolution continues even after the corresponding feature has disappeared from the input. After a TIM-cell reaches a certain activity limiting threshold, its activity rapidly decays.

The cells in the third layer constantly renormalize patterns in the TIM layer and (by the help of their competition dynamics) make guesses as to what spatio-temporal input pattern is presently fed into the network. Consequently the network computes at every sampling time step its expected pattern due to spatial and temporal (intensity) knowledge it may have learned over a long period of time.

2 The network

We consider the network architecture depicted in Fig. 1. Spatial patterns, one at a time, enter the network by arriving at the feature analyzing units of the separate subnets of a first layer.

These sub-networks independently analyze the incoming patterns for the presence of some predefined features or feature combinations. In other words, they measure features and perform competition for highest activity in order to represent the features symbolically. We can think of these cells as being (locally) tuned to certain feature values much like Radial-Basis-Function (RBF) cells (Nowlan 1990) in a corresponding signal space would be. The sub-networks decide winners by performing a (dynamical) competition of its mutually inhibiting cells. We have chosen the network of Haken (Haken 1987, 1988) as a realization of the Winner-Take-All (WTA) function.

In order to make things more evident, in what follows we shall concentrate on one feature α out of M possible features all of which are analyzed at the same time. Suppose now, there are a number of feature analyzing cells, K, for feature α. Their initial activities d_k^0 are a measure of the corresponding feature values present in the input patterns $\mathbf{q}(t_i)$ at time t_i which are transmitted via connections A_{ik}. Thus, for instance

$$d_k^0(t_i) = \mathrm{N}e^{-(f(t_i)-f_k)^2/2\sigma^2}, \tag{1}$$

where $f(t_i)$ is the presented feature value and f_k, σ, N specify feature value as well as broadness and height of response for a cell k (see Fig. 2). The best matching cell is then found by performing the following dynamics between sampling times t_i and t_{i+1} (using $d_k^0(t_i)$ as initial values):

$$\tau\dot{d}_k(t) = d_k(t)\left(1 + d_k(t)^2 - \sum_{k' \neq k} d_{k'}(t)^2 - \sum_{k'} d_{k'}(t)^2\right). \tag{2}$$

By the lateral inhibition in (2) it is guaranteed that only one cell remains active after the relaxation time t_R, $t_R < t_{i+1} - t_i$, has elapsed. The maximum initial activity allowed by (1) is $d_k^0 = N$ and we set $N = 0.5$. Consequently, due to the competition only one cell out of K will enter the activity range of $0.9 \leqslant d_k(t) \leqslant 1.0$. Somewhere within this range the TIM cell population c_k (cf. Fig. 3) coupled to grandmother cell d_k experiences a threshold excess. Cells within this population which are not yet active are susceptible to the threshold excess. One cell is chosen either systematically or accidently to store the presence of the feature value corresponding to cell k.

Activity c_{kl} of a TIM cell l in population k develops according to the following evolution equation:

$$\dot{c}_{kl}(t) = c_{kl}(t) \cdot [\beta_+ \Theta(c_{kl}(t) - c_+) - \beta_- f(t)] \tag{3}$$

$$f(t) = e^{-\beta_=(t-t_\Theta)}\Theta(t - t_\Theta), \tag{4}$$

where c_+, c_- are threshold activities with t_Θ being the moment in time c_- is reached, $c_{ij}(t_\Theta) = c_-$, and $\beta_+, \beta_-, \beta_=$ are universal growth/decay parameters. $\Theta(x)$ is the Heaviside function

$$\Theta(x) = \begin{cases} 1 & \text{for } x \geqslant 0 \\ 0 & \text{for } x < 0 \end{cases}$$

and $c_+ < c_-$ determine thresholds critical to the onset of growth and decay respectively. This dynamics can be achieved by an activity dependent feed-back connection for each cell c_{kl}. In a way, cells c_{kl} constitute a short term memory of feature values symbolized by high activities of cells d_k.

As we said earlier, everything so far is done in parallel and independently for all feature analyzing sub-networks α. It is only in the last layer of cells that signals are integrated into a unique representation of spatio-temporal patterns passing through the network. This strategy is advantageous in order to avoid combinatorial explosion in the feature analyzing layer. If every cell had to analyze an arbitrary combination of features the number of cells required to cover a certain pattern space would grow dramatically. An independent analysis of single features on the other hand, with an integration of signals at the latest stage possible, can exploit the parallel nature of the network most efficiently.

The cells of the third layer are connected to all TIM populations in Layer II via connections B_{kl}. Again,

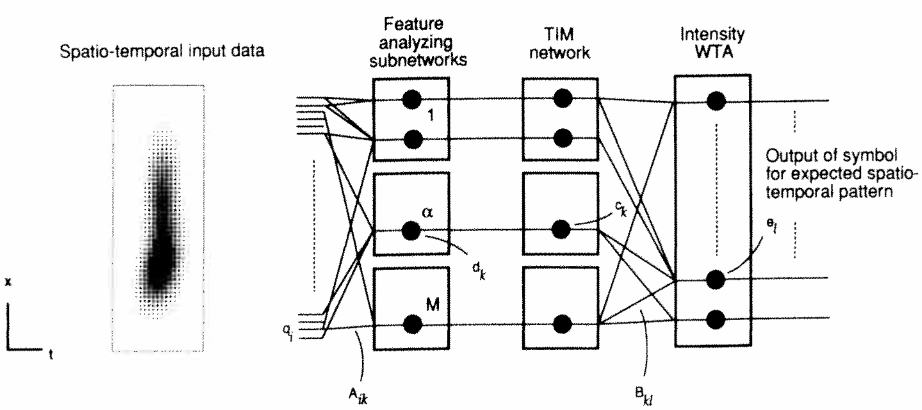


Fig. 1. Network architecture. Spatio-temporal input is independently analyzed by feature analyzing subnetworks $1, \ldots, \alpha, \ldots, M$. Information about present features is passed to independently acting TIM subnetworks. The resulting intensity distribution is finally examined by a WTA network

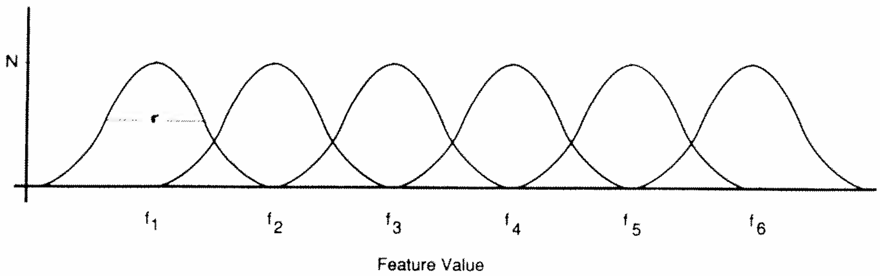


Fig. 2. Activity of feature analyzing cells varies according to (1). σ and N are equal for all cells

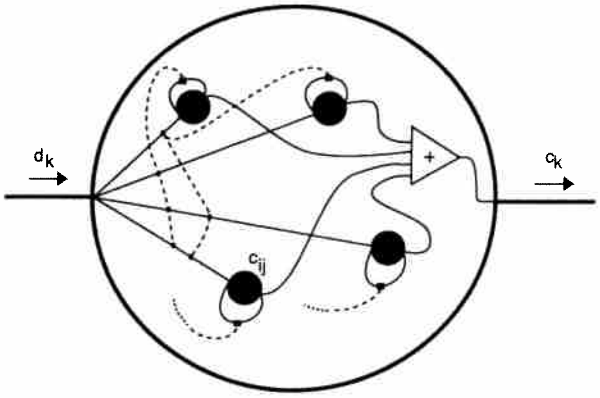


Fig. 3. A population of TIM-cells c_{ij} with total activity c_k, cooperate to allow repetition of spatial patterns. Dashed connections (only at two cells explicitly shown) extend to all other cells within the population and are strongly inhibitory

their activities $e_l(t)$ are governed by competition analogous to (2) with a relaxation time $t_{R'}$. Thus they make use of the partial intensity representations gained by various sub-networks in order to achieve an integrated picture of what is happening. The WTA function implemented in the third layer's lateral connections guarantees that one and only one cell will fire with high activity after $t_{R'}$ has elapsed. The relaxation time of the third layer should be well below the distance between sampling times so that the network can arrive at a conclusion (which is essentially a prediction) about the spatio-temporal input pattern within each sampling period.

Inspection of Fig. 1 reveals that the network is nearly symmetrical allowing for a reverse mode of operation. The reverse mode starts out with the excitation of a cell responsible for a certain spatial pattern which actually represents a TIM-pattern. All other cells should be kept silent during a stimulation. The excited cell generates its corresponding spatial pattern which is relayed to a reversely operating TIM-network. Activities of TIM cells are again growing until they arrive at a certain predefined maximum value. At the moment this threshold is reached by a TIM cell, however, a corresponding spatial pattern is evoked and the TIM cell is reset to "0". Since no two activity values in the initial TIM-pattern were identical, a sequence of spatial patterns is produced as one after the other of the TIM cells reaches the threshold. We will demonstrate this mode of operation below in the latter part of our simulations.

3 Simulation

The patterns to be analyzed are superpositions of 2-dimensional gaussian patterns. They are generated on a 20×60 grid and can be interpreted as artificially generated sonograms. Intensities on 20 frequency channels sampled over 60 time steps are computed by:

$$x' = x \cos \alpha + t \sin \alpha \quad t' = -x \sin \alpha + t \cos \alpha \tag{5}$$

$$I(x', t') = \sum_{\gamma} h_x^{(\gamma)} e^{-[(x' - x_0^{(\gamma)})/\lambda_x^{(\gamma)}]^2} \times h_t^{(\gamma)} e^{-[(t' - t_0^{(\gamma)})/\lambda_t^{(\gamma)}]^2} . \tag{6}$$

Thus we consider one dimension as being spatial, the other as being temporal. The selection of sample patterns is shown in Fig. 4, together with the parameters necessary to generate them.

Three types of feature analyzing cells (see (1)) are employed which analyze the spatial dimension of patterns at every sampling time $t = 1, \ldots, 60$. Specifically, they analyze height, width and center of assumed gaussian spatial patterns. Although this is an approximation, it is the kind of knowledge feature analyzing cells must possess in order to process input effectively. We assume that these feature analyzing cells have divided their respective feature space equally as can be achieved using a competitive learning rule (Banzhaf and Haken 1990). For our simulation we have simply set up 20 feature analyzing cells in each feature between a minimum and a maximum value with equal response width and maximum response distance (cf. Fig. 2).

By independently measuring height, width and center of spatial gaussians with 20 cells each, we can effectively code for the appearance of $20^3 = 8000$ different spatial gaussians. At each time step the feature analyzing cells first measure their degree of matching with the input pattern (see Fig. 5 for one instance). They then perform a competition which results in 3 winning cells representing the best matched cell in height, width and center for spatial gaussian patterns. Winners are able to trigger TIM cells as described in the preceding section.

The parameters of our network are:

$$c_+ = 0.2 \qquad c_- = 1.0$$

$$\beta_+ = 0.007 \qquad \beta_- = 0.9 \qquad \beta_z = 0.09 .$$

Figure 6 shows the combined intensity patterns constituting the input to the third layer WTA for the prototype

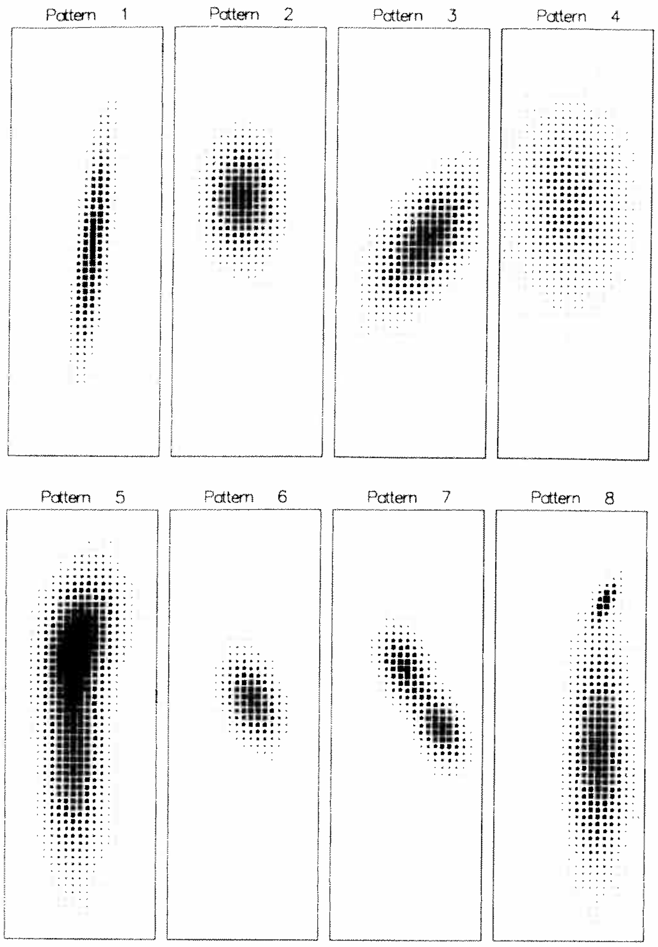


Fig. 4. Sample patterns used in the simulation. Parameters $(h_x\ x_0\ \lambda_x\ h_t\ t_0\ \lambda_t\ \alpha)$ are respectively for these patterns:
1: (1 15 2 1 30 15 0.1); 2: (1 10 5 1 25 8 0); 3: (1 28 5 1 18 10 0.6);
4: (0.75 10 8 0.75 25 12 0); 5: (1 15 5 1 15 9 0.25), (1 10 4 1 35 15 0);
6: (1 2 3.5 1 30 6 −0.35); 7: (1 2 3 1 25 5 −0.35), (1 4 3 1 35 5 −0.35);
8: (1 20 1 1 5 3 0.45), (1 15 4 1 35 15 0)

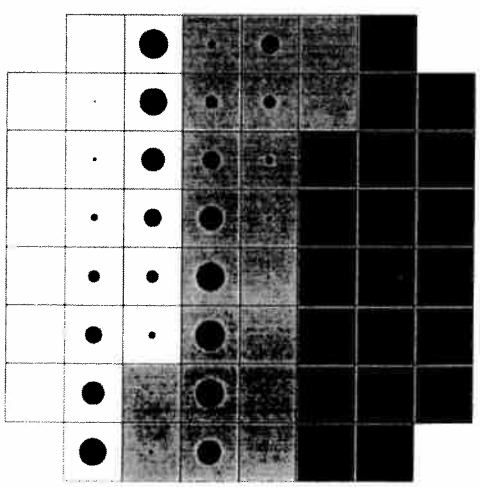


Fig. 5. One instance of an initial matching measurement by gaussian feature analyzing cells. The first 20 cells (*white background*) code for the height, the next 20 cells (*light gray background*) code for width, the last 20 cells (*dark gray background*) code for center of an assumed spatial gaussian in the input. Initial firing intensities are shown as size of black dots for $t = 30$ of pattern 5

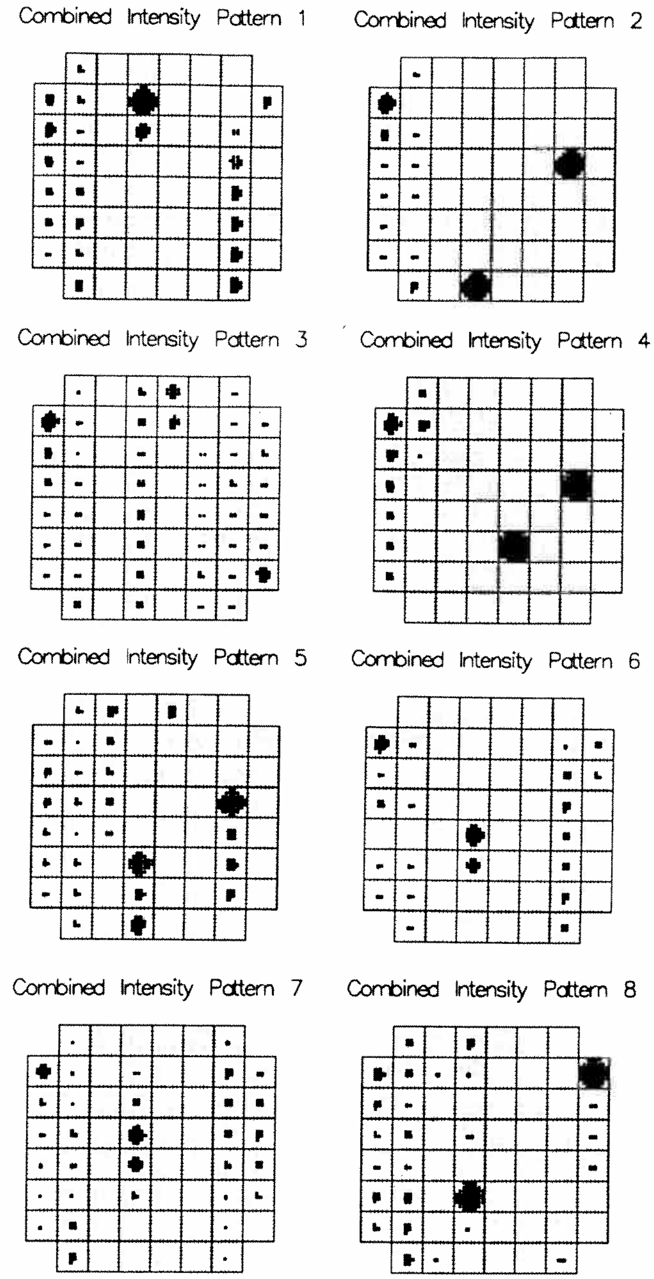


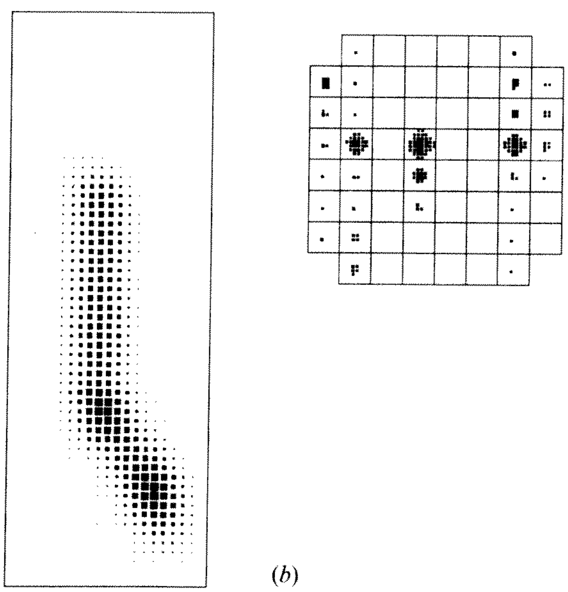
Fig. 6. Combined intensity patterns, i.e. input to the third layer WTA for patterns 1 to 8 of Fig. 4

type patterns. 60 TIM cells have been employed for every feature in order to allow an arbitrary repetition of features in time. Table 1 compares the respective overlaps of input patterns and time-intensity patterns with each other. We can see immediately that a discrimination based on overlap is much easier with time dependent patterns mapped to intensities.

A typical recognition example is shown in Fig. 7. Figure 7a shows one of the prototype patterns in which a delay was inserted after $t = 20$. In fact, from $t = 20$ to $t = 40$ the signal was just identical with the original signal of $t = 20$. Then the development resumed with a delay of 20 time-steps. As can be observed from a

Table 1. Average off-diagonal overlap between patterns before and after time-intensity mapping

Pattern Nr.	Off-diagonal Overlap	
	2-D patterns	TIM patterns
1	0.577	0.211
2	0.669	0.328
3	0.637	0.359
4	0.700	0.248
5	0.634	0.375
6	0.676	0.394
7	0.622	0.342
8	0.501	0.234



(*a*) (*b*)

Fig. 7. **a** Distorted pattern sample. Between $t = 20$ and $t = 40$ the same spatial pattern was repeated. Development of the original pattern resumed after this delay of $t = 20$ time steps. **b** Resulting TIM pattern

comparision of Fig. 7b and Fig. 6, the resulting TIM-pattern does not differ greatly from the prototype TIM-pattern. Table 2 summarizes these results in terms of respective overlaps of the distorted pattern with all the prototype patterns.

Our experiments on different kinds of distortion have shown the following general tendencies:

(i) Distortions in time are generally easily removed by time-intensity-mapping.

(ii) Multiplicative noise imposed on the patterns sometimes can be removed quite radically by the TIM-process. In most cases, however, TIM-patterns are not as resistant against noise than the original patterns are.

(iii) Global changes in the intensity of patterns are removed by the normalization operation, even for the original space-time patterns. Local variations of intensities, however, are more effectively removed by TIM-patterns.

(iv) Although width invariance was not built into the system, a tolerance against width variation of $\pm$ 10% was observed.

(v) Similarly, a tolerance against the maximum of the spatial gaussian detected in the first layer of the network of $\pm$ 10% was observed.

Table 2. Resulting overlaps for the distorted patterns of Fig. 7. In TIM representation, the pattern is recognized correctly as pattern 7

Pattern Nr.	Overlap with distorted pattern	
	2-D	TIM
1	0.565	0.010
2	0.560	0.400
3	0.463	0.309
4	0.545	0.370
5	0.723	0.346
6	0.425	0.428
7	0.362	0.800
8	0.411	0.136

Since the network was basically constructed to remove distortions in time, we concentrated in the statistical part of our simulations on this kind of pattern variation. Figure 8 shows a comparison of the original 2 dimensional patterns and TIM patterns with respect to time delays. Time delays between 5 and 40 time steps for the patterns of length $T = 60$ were applied and the resulting overlap to prototype patterns was calculated. One observes a rapid deterioration for the original patterns. Since other original patterns ususally show overlap of around 0.5 . . . 0.8, time delays greater than 10 will inevitably lead to erroneous classification. On

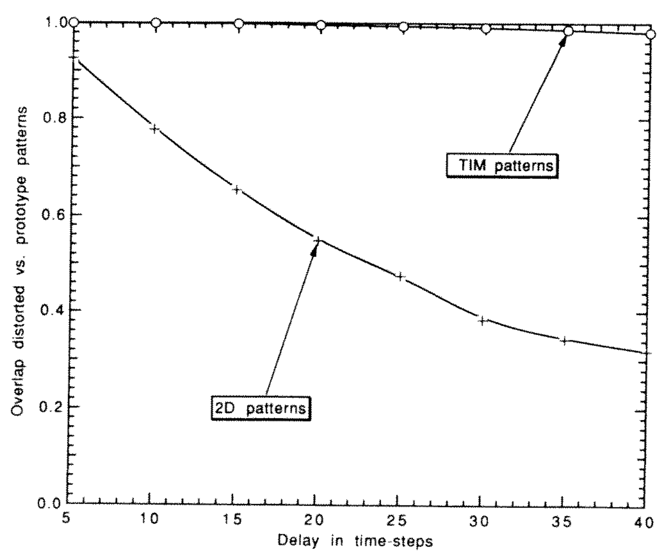


Fig. 8. Comparison of overlaps between distorted and original patterns for 2-dimensional and TIM patterns. 100 randomly chosen local time delays were used for each data point

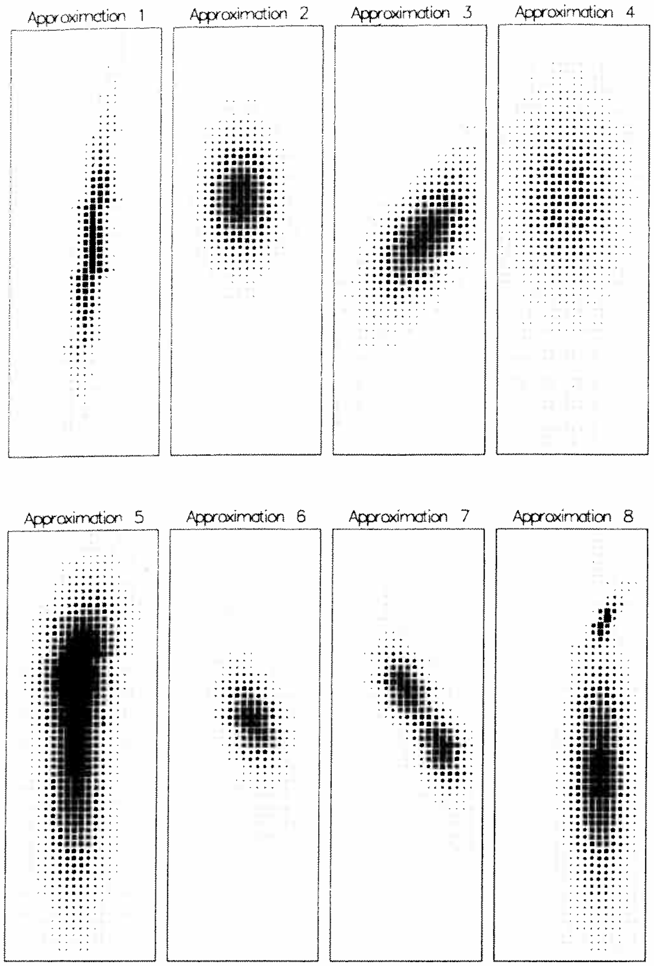


Fig. 9. Reconstructed patterns, lacking some smoothness due to the reconstruction from spatial gaussians

Table 3. Average quadratic reproduction errors per component for patterns 1–8 if number and resolution of feature detecting cells is lowered

Number of Feature detecting cells	Height	Width	Center	All
20	0.0017	0.0017	0.0017	0.0017
18	0.0007	0.0018	0.0065	0.0628
16	0.0016	0.0040	0.0194	0.0212
14	0.0039	0.0072	0.0410	0.0410
12	0.0082	0.0136	0.0648	0.0561
10	0.0156	0.0160	0.0873	0.0644
8	0.0196	0.0212	0.0967	0.0670
6	0.0275	0.0288	0.1052	0.0678
4	0.0549	0.0532	0.0987	0.0672
2	0.0549	0.0532	0.0987	0.0672

the other hand, TIM-patterns are very resistant against time delays as shown by the 2% figure.

In the remaining part of this section we want to address the reproduction abilities of the TIM network. After setting all connections for the reverse mode of operation of the former case, we can observe the formation of patterns in time which bear a close similarity to the original patterns. Figure 9 shows the reproduced patterns generated in the reverse mode which correspond to original patterns of Fig. 4. Although the smoothness of the original patterns could not be reached, the overall ability is evident.

Table 3 gives the quadratic error of the reproduction, averaged over all dimensions and all the different patterns for various resolutions of the network's feature detecting cells. One observes that good resolution in the detection of the center of the spatial gaussian is more crucial for a good approximation than a good resolution for either the width or height of the gaussian.

4 Discussion

We have shown a network capable of removing distortions of patterns in time. The network's abilities were demonstrated using artificially constructed patterns resembling sonograms. It was shown that by using the Time-Intensity-Mapping method correlations between distorted and original patterns were by far higher than those generated using the prototype patterns directly.

Mapping time into intensity, which is effectively employing a short term memory with appropriate read-out, uses the constant and homogeneous activity growth or decay rate of the TIM-cells (memory) to generate patterns corresponding to transient spatio-temporal patterns. Even after the momentaneous spatial patterns have disappeared from the input, TIM-cells continue to hold their effects dynamically. By applying constantly normalizing WTA cells, patterns can be discerned according to the spatial distribution of activities they generated in the TIM populations. We introduced populations of cells in order to allow for an arbitrary repetition of spatial signals in the input.

Recognition of transient patterns is most obvious if a continuous stream of input data is used, as was already demonstrated in Banzhaf (1991). There, the prediction capabilities of the network were shown by following the expectations of the network about the spatio-temporal patterns at the input using selected words of the English language.

In this contribution we did not address the question of choosing a suitable time frame for the recognition process. As the network can output not only its present expectation but also the present degree of matching between stored prototype (TIM) patterns and newly generated ones, this could provide a natural signal for framing the data-stream. Local maxima of the degree of matching would signal a successful identification of a pattern which could reset the TIM cells to a quiet state.

Another problem only briefly mentioned was that of learning. Once the general architecture is set up, synapses could be subjected to learning rules. For fixing the features which are detected in the first layer we propose an unsupervised (competitive) learning rule which should be able to equally distribute the feature analyzing cells among the different feature. As far as the connections between the TIM layer and the third layer WTA are concerned, we imagine a supervised learning rule telling the network which TIM pattern it should devote a grandmother cell to. Connections within the TIM layer should be kept fixed for the entire learning process.

An appropriate extension to the network is to use an entire hierarchy of networks of the kind WTA $\Rightarrow$ TIM $\Rightarrow$ WTA $\Rightarrow$ TIM $\Rightarrow \cdots$. The first WTA recognizes "instantaneous" spatial patterns, the second WTA recognizes spatial patterns relayed through a very short TIM process, the third and subsequent ones recognize TIM patterns of longer and longer time-scales. In this way, patterns can be recognized and reconstructed based on their internal time wise structure.

Acknowledgement. Discussions with our colleagues T. Ishii, T. Iwamoto and E. Lange are gratefully acknowledged. We thank Dr. J. Bell for checking the English.

References

Banzhaf W (1991) Processing spatio-temporal patterns by mapping time into intensity. In: Proc IJCNN Seattle 1991 II:871–877

Banzhaf W, Haken H (1990) Learning in a competitive network. Neural Networks 3:421–435

Cohen MA, Grossberg S (1987) Masking fields: a massively parallel architecture for learning, recognizing, and predicting multiple groupings of patterned data. Appl Opt 26:1866–1891

Haken H (1987) Synergetic computers for pattern recognition and associative memory. In: Haken H (ed) Computational systems, natural and artificial. Proceedings of the Elmau International Symposium on Synergetics 1987. Springer, Berlin Heidelberg New York, pp 1–21

Haken H (1988) Nonequilibrium phase transitions in pattern recognition and associative memory. Z Phys B-Condensed Matter 70:121–123

Hecht-Nielsen R (1987) Nearest matched filter classification of spatiotemporal patterns. Appl Opt 26:1892–1899

Nowlan SJ (1990) Maximum likelihood competitive learning. In: Touretzky DS (ed) Advances in neural information processing systems, vol 2. Morgan Kaufmann, San Mateo, Calif pp 574–582

Wang D, Arbib M (1990) Complex temporal pattern sequence learning based on short-term memory. Proc IEEE 78:1536–1543

Dr. W Banzhaf
Central Research Laboratory
Mitsubishi Electric Corporation
1-1 Tsukaguchi Honmachi 8-chome
Amagasaki
661 Japan