PERGAMON

Neural Networks 12 (1999) 163–173

Neural Networks

Contributed Article

Parallel and robust skeletonization built on self-organizing elements

Zsolt Kalmár[a,d], Zsolt Marczell[b], Csaba Szepesvári[c,d], András Lörincz[b,*]

[a]*Department of Informatics, "József Attila" University of Szeged, Szeged, Aradi vrt. tere 1, Hungary H-6720*
[b]*Department of Chemical Physics, Institute of Isotopes of the Hungarian Academy of Sciences, Budapest, P.O. Box 77, Hungary H-1525*
[c]*Research Group on Artificial Intelligence, "József Attila" University of Szeged, Szeged, Aradi vrt. tere 1, Hungary H-6720*
[d]*Associative Computing Ltd. Konkoly Thege M. út 29-33, Budapest 1121*

Received 24 June 1996; accepted 19 May 1998

Abstract

A massively parallel neural architecture is suggested for the approximate computation of the skeleton of a planar shape. Numerical examples demonstrate the robustness of the method. The architecture is constructed from self-organizing elements that allow the extension of the concept of skeletonization to areas remote to image processing.

Keywords: ANN; Skeletonization; Self-organization

1. Introduction

The idea of using the *skeleton of a continuous planar shape* (the loci of centres of those bitangent circles that lie entirely within the shape) for representation, as a kind of primitive of the original, was first proposed by Blum (1967).

Blum's argument was that the skeleton efficiently concentrates the topological information of the original two-dimensional shape. This kind of 'thin' coding is particularly useful for representing amorphous, irregular shapes that cannot be treated by more conventional geometrical methods. The possible applications include the creation of shape primitives, curve segmentation, logging deformation history of deformable objects as well as image preprocessing for shape recognition.

Blum suggested a transformation to elicit the skeleton that has become known as the grassfire transformation. The idea can be sketched as follows: let us imagine that 'a fire is lit' along the boundaries of the shape and monitor how it spreads inwards. By taking the set of points where at least two fire-fronts meet, one obtains a connected pattern, formed of line-segments and arcs. If the fire-fronts propagate isotropically (and some general conditions are satisfied, see, e.g., Serra, 1988), this pattern coincides with the skeleton defined by means of maximal circles[1]. Since the appearance of Blum's paper much effort has been devoted to develop algorithms for the calculation of the skeleton.

Boundary based methods use a representation of the object's boundary as input and make analytical considerations to obtain a representation of the skeleton. Some algorithms calculate the skeleton directly, while others exploit the fact that for polygonal shapes the skeleton is a subgraph of the Voronoi diagram (Kirkpatrick, 1979). If the boundary is replaced with a point set obtained by discrete sampling, then the skeleton of the shape can be approximated by calculating the discrete Voronoi diagram of the point set and extracting the skeleton from it (Schmitt, 1989; Brandt and Algazi, 1992). Another possibility is to use a polygonal (or spline) approximation of the boundary instead of discrete sampling (see, e.g., Martínez-Pérez et al., 1987).

These approaches have many advantages: the resulting pattern is the Euclidean skeleton (of the approximation set) and its connectivity is automatically preserved. Furthermore, the representation of both the boundaries and the Voronoi diagram requires data linearly proportional to the image resolution and the absence of a grid makes the use of the Euclidean metric easy. However, there are some drawbacks as well. For example, shapes with holes are difficult to treat this way and complex and/or noisy shapes may require a very fine polygonal approximation that makes results less accurate.

* Corresponding author. Tel.: +36-1-2754348; Fax: +36-1-2754349; E-mail: lorincz@iserv.iki.kfki.hu; http://iserv.iki.kfki.hu/. Present address: Eötvös Loránd University, Budapest, Múzeum Krt. 8, Hungary H-1068.

[1] Note that some authors allow fire-fronts to propagate both inwards and outwards and define the skeleton as the union of patterns arising in and out of the shape, while others prefer to distinguish between endo- and exoskeletons.

0893–6080/99/$ - see front matter
PII: S0893-6080(98)00119-1

Region based algorithms generally realize an image-to-image transformation that requires data proportional to the square of image resolution, but are more appropriate for parallel implementations. The input of these algorithms is a discretized of the skeleton. In this framework various algorithms are used to elicit the skeleton, such as grassfire propagation and thinning; for surveys see (Smith, 1987; Serra, 1988).

A class of region based algorithms use a distance transform (DT) of the original shape to elicit the skeleton. In a distance transformation to each pixel of the digitized image a number is assigned that measures the distance of the boundary closest to that particular pixel. Skeletonization then can be performed as ridge following on the distance transform surface. Different approximate transformations can be used: the city block DT (Arcelli and di Baja, 1989), the chessboard DT (Arcelli and di Baja, 1985), the hexagonal DT (Meyer, 1988), chamfer and quasi-Euclidean DTs (Montanari, 1968) and the Euclidean distance transformation (EDT) (Ho and Dyer, 1986). The quality of the results and the computational cost depends strongly on the choice of the DT. The city block and the chessboard DTs provide fast computation and skeletal pixels are guaranteed to be connected within a fixed number of steps, however, both DTs give a poor approximation of the EDT. Skeletonization based on a quasi-Euclidean or on a EDT give more appropriate result but the thickness and the connectivity of the skeleton branches must be carefully checked (Shih and Pu, 1990).

Distance transformations have been generalized to *morphological skeleton transformations* that are based on the set-oriented method of mathematical morphology (Serra, 1982), and make use of a set of symmetric or asymmetric structuring elements (Maragos and Schafer, 1986) to form the skeleton on hexagonally or rectangularly sampled binary images.

Skeletonization by means of ridge-following on a distance transform is considered recently as an attractive type of algorithm. The main reasons are the need for simple computation (mask convolutions carried out through a fixed number of iterations, independently of the complexity of the object) and the more accurate result (compared to the computational costs) than that of other algorithms.

In this paper an artificial neural network is proposed as a medium to perform a variation of Blum's grassfire algorithm. The grassfire is modelled as a spreading process of neural activations, in which neurons, located in a hexagonal layer, transmit activation to their nearest neighbours. This spreading process can be regarded as an erosion model (Serra, 1982, 1988), but also as a distance transformation, which, however, is not performed by mask convolutions but rather by *massively parallel* computation. It will be shown that the spreading process is guaranteed to terminate in a fixed number of time-steps, regardless of the complexity of shape, and, moreover, the globality of computation will be demonstrated to induce *rotation invariance*.

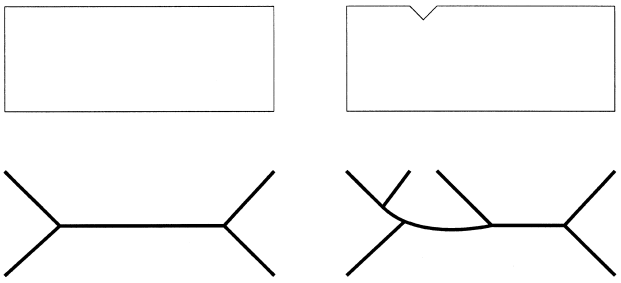

Fig. 1. Rectangles illustrating Marr's criticism. The figure illustrates a stability problem of traditional skeletonization. Let us consider a rectangle (top left) and its skeleton (bottom left). The inclusion of a wedge-shaped perturbation into a side of the rectangle (top right) induces topological changes in its skeleton (bottom right) and thus inhibits the use of noisy boundary data. Note that the deformation of the skeleton is sustained even if the dimensions of the perturbation go to zero.

It will be demonstrated that the ANN architecture is built up of elements that can be formed by purely self-organizing means. The construction of a skeletonization algorithm by such purely self-organizing means allows to extend the use of the algorithm beyond the realm of image processing.

The grassfire transformation is often criticized for giving very different results for only slightly differing shapes. This feature can be best illustrated with an example of Marr (1982), in which the skeleton of a rectangle and the skeleton of another shape—a rectangle with a small wedge included in a side—is compared Fig. 1. Skeletonization transforms the smallest disturbances of the boundary into topological differences in the skeleton. A standard technique to treat boundary noises is to use a variation of the multi-resolution approach (Rosenfeld, 1984; Gauch and Pizer, 1993) *and/or* to assign an *importance measure* to each skeletal point. Our algorithm computes a discretized Laplacean importance measure to every skeletal point in a parallel way. It will be argued that there is also a natural way to incorporate a (self-organized) multi-resolution filter system.

The organisation of the paper is the following. Section 2 gives a short description of the algorithm and of the architecture of the system. In Section 3 this is followed by an

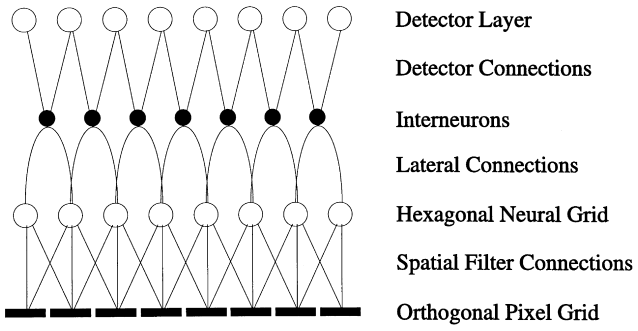

Fig. 2. Layered ANN architecture for skeletonization. Pixels of the input layer are connected to a set of spatial filters with centers arranged into a hexagonal structure (first layer). These filters are connected via lateral connections. Each interneuron (second layer) is assigned to a lateral connection and inputs the difference of activations flowing along the connection times a weight factor. Detector neurons (third layer) collect and integrate the output of their neighbouring interneurons (see text for details).

account of the computer simulations. This is followed by the discussions in Section 4, and conclusions are drawn in Section 5. Finally, the paper is ending with two appendices about the inversion properties of the discretized Laplacean importance measure and about the framework of creating grassfire substances in general input spaces.

2. Algorithm and architecture

2.1. Spatial filter system

Let us imagine a three-layered neural network (see Fig. 2). The functional organization of the layers is the following: the first layer performs a Gaussian blur of the input intensity distribution and, together with the lateral connections that build up the second layer, models the spreading of the fire. The third layer serves the detection of the spreading grassfire.

The first layer consists of n neurons, arranged in a planar hexagonal structure. These neurons are connected to a pixel-discretized image via their feed-forward connections that form Gaussian filters on the input space:

$$q_{ij} = \exp(-d_i^2(j)/d_0^2), \tag{1}$$

where q_{ij} denotes the weight connecting the ith neuron to the jth pixel, $d_i(j)$ is the Euclidean distance between the centre of the receptive field of the neuron i and the pixel j and, finally, d_0 is a scale parameter. The feed-forward connections map the pixel intensity distribution of the image onto the first layer. This is expressed by $a_i = \Sigma_j q_{ij} x_j$, where a_i is the *input activation* of the ith neuron and x_j is the intensity of the jth pixel. The output of the neurons in the first layer will be called *output activation* and will be denoted by σ_i. The connection between a_i and σ_i will be specified later.

The first layer neurons are also connected with each other via lateral connections. The strength of the lateral connections, w_{ik}, is 2/3 between nearest neighbours in the grid and 0 (no connection) between non-nearest neighbours. The w_{ik} numbers can be considered as lateral connections, but they can be considered alternatively as single weights of *interneurons* that are arranged in a second (hexagonal) layer. The input of each interneuron is the difference of two output activations, say $\sigma_k - \sigma_l$, and the output is the following: $w_{kl}(\sigma_k - \sigma_l)$. This value corresponds to the activation that spreads along the connection labelled with the indices k,l. We shall use this latter terminology.

2.2. The model of the grassfire

The grassfire is modelled as the interaction of the spatial filter neurons of the first layer and the interneurons of the second layer. Let us consider the output activation of neuron i, $\sigma_i(t)$, as the amount of material burnt from the beginning of the process up to time t at point $\mathbf{x}_i$. The 'spreading of the fire' is described mathematically as the temporal change of these functions. The time-evolution of σ_i is determined by the interneurons in a set of shunting equations (Grossberg, 1968):

$$\dot{\sigma}_i = (1-\sigma_i)\left(\sum_{\langle k,i\rangle} w_{ik}(\sigma_k - \sigma_i) + \kappa\sigma_i\right), \quad i = 1,\ldots,n, \tag{2}$$

where $\kappa > 0$ is the self-excitation parameter and $\langle\cdot,i\rangle$ stands for the nodes that are connected to node i. (The dot above denotes differentiation with respect to time.) The terms on the right hand side describe the spreading of the fire: the shunting term $(1-\sigma_i)$ warrants that the amount of incandescent material is restricted to 1, the terms of the summation describe the amount of burning material spreading between neighbouring nodes, while the term $\kappa\sigma_i$ is a self-excitatory term that could result in an exponential growth provided that the incandescent material is not bounded.

Since the neurons are arranged in the sites of a hexagonal lattice[2], the $\sigma_i(t)$ functions can be regarded as the discretization of a $\sigma(\mathbf{x},t)$ function of space coordinates and time. Thus Eq. (2) is the discretized version of the following partial differential equation:

$$\dot{\sigma}(\mathbf{x},t) = (1-\sigma)(\Delta\sigma + \kappa\sigma), \tag{3}$$

where Δ denotes the 2D Laplacean. Let us suppose that the initial conditions for σ are such that $\sigma(\mathbf{x},0) \in [0,1]$ holds everywhere. If this is satisfied, the form of Eq. (3) renders the following statements to make.

- $\sigma(\mathbf{x},t) \in [0,1]$ for all $\mathbf{x}$,
- if there exist a point $\mathbf{x}'$ such that $\sigma(\mathbf{x}',0) > 0$ then $\sigma(\mathbf{x},t) \to 1$ as $t \to \infty$ for all $\mathbf{x}$.

To make these points apparent let us image that a $\sigma_D(\mathbf{x},t)$ function satisfies the following equation:

$$\dot{\sigma}_D(\mathbf{x},t) = \Delta\sigma_D + \kappa\sigma_D. \tag{4}$$

At each instant Eq. (4) is bounded below by the diffusion equation (viz., when only $\Delta\sigma_D$ stands on the right side). The initial conditions assure that the solution to the diffusion equation is either zero or positive everywhere, therefore, the solution to Eq. (4) is also non-negative. The shunting term, that distinguishes Eq. (4) and Eq. (3) does not play a role around zero but limits the growth of the activation when it approaches 1.

Simple calculations can justify that the elementary time-independent solution to Eq. (3) has one of the following forms:

$$\sigma_{\text{stat}}(\mathbf{x}) = \int a_k \sin(\mathbf{k}\mathbf{x} + \phi)\delta(\mathbf{k}\mathbf{k}^T - \kappa)d^2\mathbf{k}, \quad \sigma_{\text{stat}} = 0,$$
$$\sigma_{\text{stat}} = 1, \tag{5}$$

where $\delta(\cdot)$ denotes Dirac's delta function. Linear stability analysis shows that the only stable time-independent solution is $\sigma_{\text{stat}} = 1$. (Furthermore, the periodic solution is not in accordance with the condition that $\sigma \in [0,1]$.)

[2] Note that *physically* the neurons need *not* be arranged in a hexagonal lattice since the neighbourhood relations are represented by the interneurons.

The simulations have shown that if the initial conditions are such that $\sigma(0) = 0$ in a half plane and $\sigma(0) = 1$ in the other half then after a small transient, a wave-front with a steady profile is propagated in a direction perpendicular to the border of the two half planes. The speed of propagation, v, is the function of κ. Note that replacing κ with $\alpha\kappa$ in Eq. (3) corresponds to replacing x with $x' = \sqrt{\alpha}x$ and t with $t' = \alpha t$, therefore, the speed of propagation should satisfy $v(\kappa) = v_0/\kappa$, where v_0 is the velocity measured when $\kappa = 1$. This is also shown by the numerical simulations (see Part (a) and (b) of Fig. 4). Together with the constant speed propagation isotropy is another fundamental issue. In the discussion it will be argued that the requirement of isotropy can be satisfied on hexagonal grids in an appropriate limit. In the section dealing with computer simulations the attractive isotropy properties of the hexagonal grid versus the rectangular grid will be demonstrated.

Another important feature of the wave-front propagation is that—as a result of the discretized approximation of the Laplacean—the convex fronts (viewed from the low activation region) propagate slower than their concave counterparts. This effect stabilizes the linear wave-fronts and thus suppresses, e.g., dendritic growth. Just like in the case of isotropy, these shape depending propagation differences disappear as the resolution increases.

2.3. Detectors

The third layer of neurons is the layer of detector neurons; each neuron of the first layer has its own detector (Fig. 2). Each detector receives the output of the six interneurons of its counterpart in the first layer. Let us denote by d_i the sum of the relative output of these interneurons:

$$d_i = \sum_{\langle k,i \rangle} w_{ik}(\sigma_k - \sigma_i), \tag{6}$$

where the summation goes through the nearest sites. Apparently, d_i is zero until the first wave-front arrives when it goes above zero, then returns to zero and optionally becomes negative. Finally, $d_i(t)$ must converge to zero once again when σ saturates at 1 (see Part (c) of Fig. 4). The output of the detectors as a function of time is defined

$$\Sigma_i(t) = \int_0^t \sum_k w_{ik}(\sigma_k(t') - \sigma_i(t'))\mathrm{d}t' = \int_0^t d_i(t')\mathrm{d}t', \tag{7}$$

in this way detectors whose eventual output ($t \to \infty$) is above a threshold can be said to detect the skeleton (see Part (d) of Fig. 4). Clearly, if we let

$$D_i = \int_0^\infty d_i(t')\mathrm{d}t' = \lim_{t\to\infty} \Sigma_i(t), \tag{8}$$

then $D_i = 0$ if only one wavefront passes the site i. This is apparent if we recall that d_i is the approximation of the 2D Laplacean, thus we may write

$$D_i \approx D(\mathbf{x}_i) = \int_0^\infty [\Delta\sigma](\mathbf{x}_i, s)\mathrm{d}s = \mathrm{div} \int_0^\infty [\mathrm{grad}\sigma](\mathbf{x}_i, s)\mathrm{d}s, \tag{9}$$

where we may interchange the integral and the div operator because the integrand is absolutely integrable. Since the wavefront propagation is homogeneous in a neighbourhood of $\mathbf{x}_i$ (remember that we have assumed that only one wavefront passes $\mathbf{x}_i$) we get that $\int_0^\infty [\mathrm{grad}\sigma](\mathbf{x}, s)\mathrm{d}s$ is constant in a neighbourhood of $\mathbf{x}_i$, and thus D_i is zero. On the other hand, nearby skeletal points the wavefront propagation is non-homogeneous and thus $D_i > 0$. As the wavefront profile approaches the Heaviside-function the distribution sign ($D(\mathbf{x})$) converges to the skeleton (i.e., it becomes 1 at skeletal points and zero everywhere else).

We notice here that in any case we deal with images that correspond to bounded regions in space, therefore we never

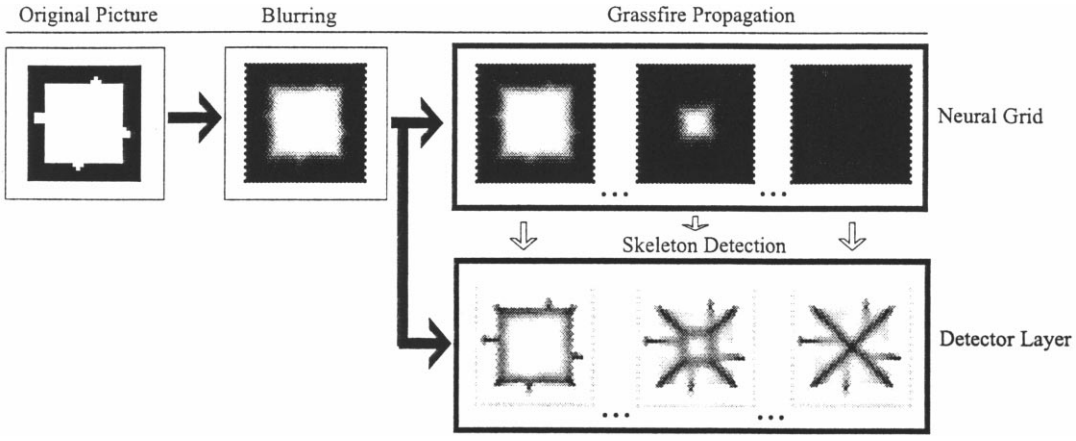

Fig. 3. The skeletonization process. The figure shows the various steps of the skeletonization process. The bolt arrows show the flow of execution of various steps. First, the original picture is blurred to reduce the effect of noise and the resulting picture is fed into the neural skeletonization process. This process has two subprocesses that work in parallel: the grassfire and the skeleton subprocesses: the figure shows three scenes of both processes: the initial, a middle (in time), and the final pictures.

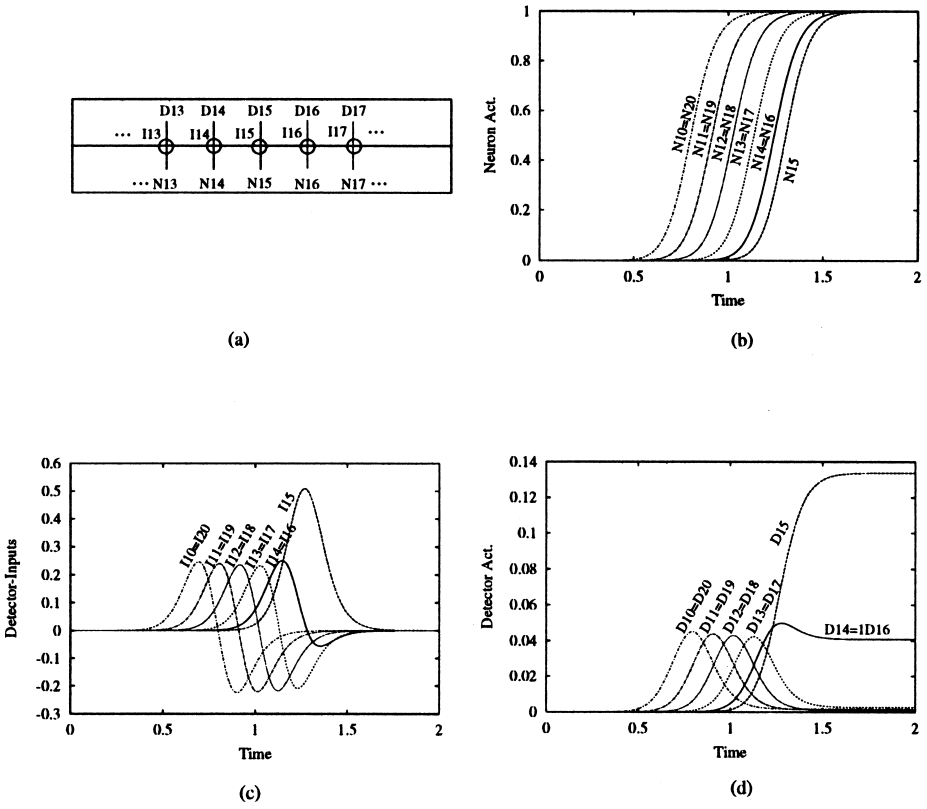

Fig. 4. Neuron activities during skeletonization. Part (a) of the figure illustrates the neural network used in this experiment. Neurons are labelled by N⟨n⟩, where n is the index of the neuron, detectors are labelled similarly by D⟨n⟩, and inputs to detectors are labelled by I⟨n⟩. For easy of visualization the network's structure was chosen to be a rectangular grid of size 5×30. Fire propagation started from the left and right sides of the grid. The activities of the selected fire-propagating neurons (N10–N_20) are shown in Part (b) as a function of time, while Parts (c) and (d) show the inputs to and outputs of the detectors, d_i and Σ_i, respectively. Part (d) illustrates that a detector which corresponds to a point of the skeleton will develop high activities, neighbouring detectors will develop above zero, but much smaller activities, while the rest of the neurons will develop approximately zero activities. A monotonous function of the distance transform can be obtained if the time instant when the input to the detectors first reaches zero from above is identified with the distance from the object boundary (see Part (c)).

have to wait infinitely long until the activations saturate. There is always a time τ that for $t \geq \tau$ the system can be considered saturated ($\sigma_i \approx 1$, $\dot{\sigma}_i \approx 0$). In this sense the sign ($\Sigma_i(\tau)$) distribution is a discrete approximation of the skeleton.

The course of simulation is the following. First the initial conditions are specified. The methods to be employed strongly depend on the result one wishes to reach, e.g., whether the exo- or the endoskeleton or both are required. The actual calculations may include edge-detection, blurring and other simple or complicated image-processing steps. This is the point in the calculation when the a_i input activations are used to calculate the initial conditions for the σ_i output activations. Until $t = 0$ the σ_i activations are clamped to $\sigma_i(0)$ and then the system is released, activation wave-fronts start and propagate from the boundaries in and/or outside of the shape and the detector signals are continuously calculated. An example for the course of simulation can be followed on Fig. 3. The computational complexity of the algorithm is dominated by the cost of the diffusion like process whose cost is $O(n)$, where n is the length of the larger side of the picture to be processed. However, since an analog VLSI implementation is planned the algorithm is hoped to run in real-time even for very large images (Oláh and Lörincz, 1995).

2.4. Relation to distance transforms

In order to have a better understanding of how the actual process works let us now briefly examine how the present algorithm is related to distance transforms. To do this let us consider a lattice site, $\mathbf{x}_i$, and suppose that the activation at $\mathbf{x}_i$ and the nearest sites is zero. Then we may distinguish between two cases: there is not a second wave-front to come until the activation at $\mathbf{x}_i$ and around it saturates or there are subsequent wave-fronts.

In the first case, when a single wave-front passes, the input to the ith detector, d_i, first starts to grow to reach its maximum, then decreases, crosses zero, sinks to its minimum and finally—as the activations converge to 1—returns to zero. The exact time course of this process depends on the profile of the wave-front, however, in our case this is not that important (see Fig. 4 for an example). What is important is the time of first return of d_i to zero (zero-crossing), because it can be considered as the time instant in which the

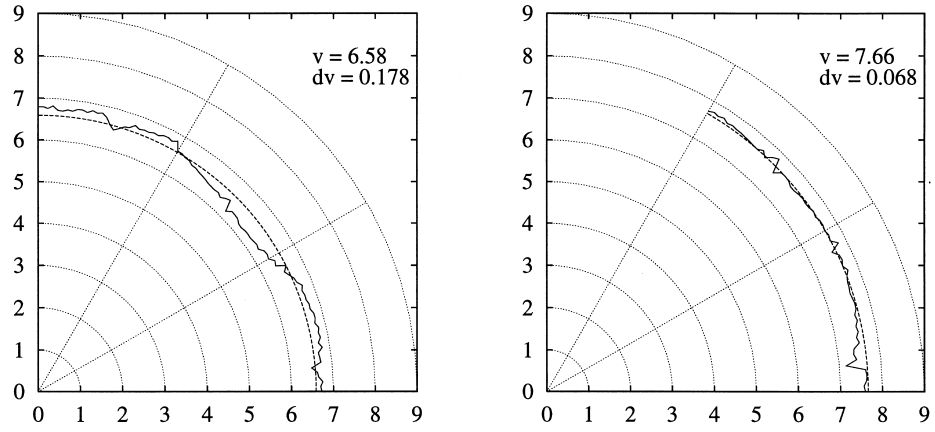

Fig. 5. Isotropy of the wave-front speed. When using spatial discretization the isotropy of the wave-front propagation is an important issue. The speed of a travelling plane wave-front is plotted as a function of the angle between the propagation vector and a lattice vector on a polar diagram. The left figure corresponds to an orthogonal spatial filter arrangement and the right figure to a hexagonal one. The average value (dotted line) and the variance of the speed is indicated in both cases in the figure. Note that the hexagonal first layer give better results. Data points are calculated in steps of 1°.

wave-front just passes the site $\mathbf{x}_i$. Since the propagation is isotropic and after a short transient it is at constant speed, knowing the starting time of propagation, the speed of propagation and the time of the zero-crossing at $\mathbf{x}_i$, a trivial calculation yields the distance of the site $\mathbf{x}_i$ to the nearest site $\mathbf{x}_k$ where the fire was ignited ($\sigma_k(0) > 0$). In other words, a distance transform can be carried out by measuring the zero-crossing times at each pixel.

In the other case, when subsequent wave-fronts arrive after the first one, the zero-crossing of d_i may be absent, but we may define the 'instant of overlap' with the time when d_i finally returns to zero. Let us consider, e.g., two wave-fronts that move in opposite directions. Assume that the wave-fronts have linear profile with the same slopes of opposite signs. Let us further suppose that the wave-fronts are launched from the same distance from $\mathbf{x}_i$ in opposite directions. In the particular case of linear wave-fronts the d_i values thus reach zero when the fronts half way meet. In this case D_i is no longer zero, but positive. It is so because when linear fronts overlap the section of the plane, from which activation flows into $\mathbf{x}_i$ is larger than the section into which activation may flow out from $\mathbf{x}_i$. In other words, D_i is positive if and only if $\mathbf{x}_i$ is a skeletal point[3].

3. Computer simulations

3.1. Characteristics of grassfire propagation

The numerical simulations were carried out with 64×64 pixel images. The 64^2 dimensional input space was discretized by 4096 neurons. The d_0 scale parameter in Eq. (1) was set to 3 and the κ self-excitation parameter was set to 13 (unless otherwise stated). All skeleton figures were thresholded at zero.

It has already been mentioned that the isotropy of propagation is an important feature. We have measured the speed of a relaxed linear wave-front as it was propagated in different directions on either an orthogonal or a hexagonal lattice. The results are depicted in Fig. 5: the propagation on the orthogonal grid is faster in the direction of lattice vectors. The deviation from the average is 17%. On the hexagonal grid the result is better, the deviation from the average is 6%, which is comparable to the error of other DTs. A weak point of many skeletonization algorithms working with pixel discretization is *rotation invariance*. In our case rotation invariance relies on the isotropic propagation of wave-fronts.

We have argued that if $\mathbf{x}_i$ is a skeletal point then D_i is positive. Let us suppose that $\mathbf{x}_i$ is designated by the overlap of *exactly* two wave-fronts[4]. During the saturation of activations around site $\mathbf{x}_i$ the self-excitation ($\kappa\sigma_i$) competes with the transferred excitation ($\Sigma_j\ w_{ik}(\sigma_k - \sigma_i)$) to drive σ_i towards 1. Since the self-excitation only depends on the actual level of σ_i, D_i will be largest for wave-fronts advancing in opposite directions and will monotonically decrease with the angle included by the fronts. If the wave-front profiles are increasing linearly from 0 to 1 then D_i goes with $\cos^2(\alpha/2)$ (α is the included angle). For the more complicated fronts arising from Eq. (2), Fig. 6 shows how D_i depends on α. The calculations have been carried out for 12 different angles, starting from 0 with an increment of 15° (the 13th angle, 180°, was not computed). The value of D_i has been measured at a site far from the vertex of the angle formed by the two fronts. According to our previous expectations the plot of the curve was found to be monotonically decreasing. For relatively small angles (α is up to 30°) it is rather flat, but at approximately 30° it becomes steeper and then falls more or less steadily to zero.

3.2. Robustness

Since the algorithm is intended for analogue VLSI implementation (Oláh and Lörincz, 1995), it has been examined

[3] Provided, of course, that the fire was initialized along the boundaries.

[4] Such an $\mathbf{x}_i$ point is called an ordinary skeletal point. The set of non-ordinary skeletal points is of zero measure.

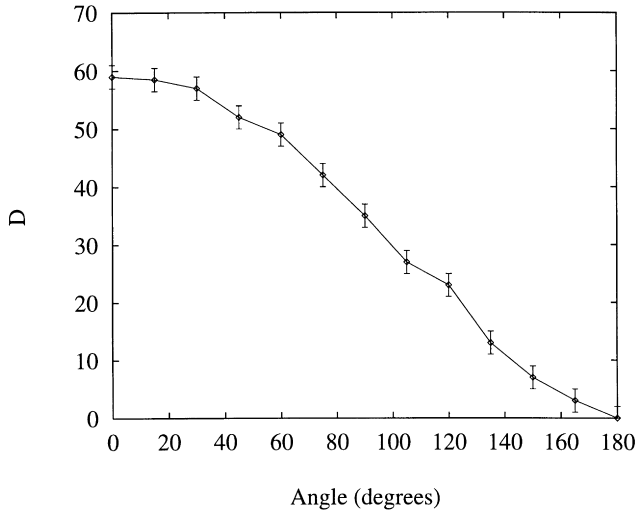

Fig. 6. Angular dependence of detected activation. In ordinary skeleton points the Σ activation is an injective function of the angle included between the folding wave-fronts. The figure shows the plot of this functional relation in case of the particular dynamics described in Section 2. Data were obtained by a software implementation of the ANN architecture.

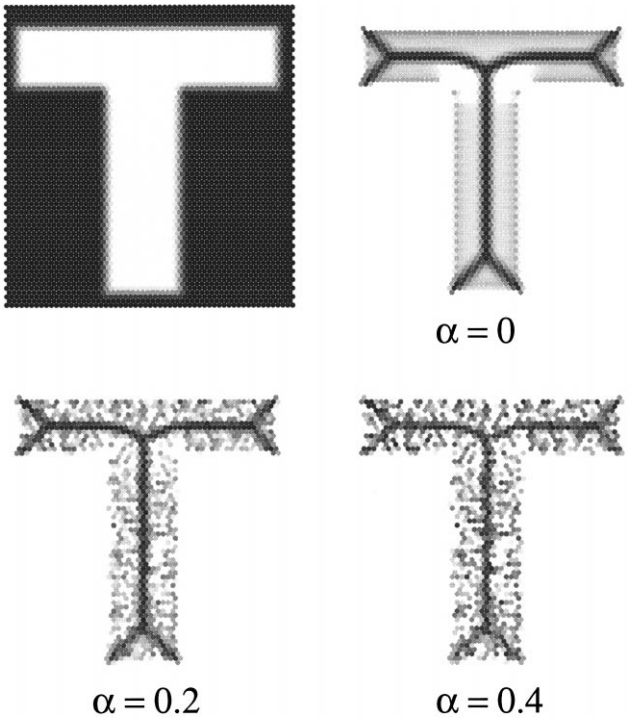

Fig. 7. Robustness of the ANN architecture. The robustness of the ANN architecture was tested by adding noise to the value of feed-forward and lateral connections. The noisy lateral connections were calculated according to $w'_{ij} = w_{ij}\,(1 + \alpha n)$, where w_{ij}, $\alpha_1 n$ are respectively a lateral connection in the noiseless case, a parameter between 0 and 1, and a random number drawn from [−1,1] with uniform distribution. The top left figure shows the initial distribution of activation, the successive three figures shows the generalized skeleton for the parameter values $\alpha = 0$, 0.2, 0.4. The superposition of noise of a similar amplitude on the q_{ij} feed-forward connections had no significant effect.

how noise in the connection strengths and in the filter positions affects the performance of the system. One expects that the success of the algorithm relies on the accurate reproduction of the theoretical values of all parameters involved in the system. It turns out, however, that the by far most important feature is the uniformity of the lateral connections. Simulations show that if one perturbs the position of the neurons in the first layer with a random error of 15% of the lattice constant the performance of the system is till unaffected. Conversely, the modification of lateral connections has a more characteristic effect. In Fig. 7 the lateral connections were modified according to $w'_{ik} = w_{ik}\,(1 + \alpha p)$, where α is a parameter, w_{ik} is the undisturbed connection and p is a random number drawn from a uniform distribution on [−1,1]. The top left panel shows the initial output activation of the second layer. The other three panels are the skeletonization outputs when $\alpha = 0$, 0.2, 0.4 successively.

Fig. 8 illustrates the performance of the algorithm with another type of noise: when the boundary data are noisy. The top left panel again shows the initial condition. The skeleton (top right panel) contains a strong middle branch but there are also discrete pixel with high D_i along the contour. The bottom left panel shows the distribution of $\sigma(0)$ in case of the same input but with different (greater) scale parameter ($d_0 = 6$ in Eq. (1)). The bottom right panel shows the skeleton at this resolution.

4. Discussion

We have presented the sketch of an algorithm that realizes an image-to-image transformation: from a pixel discretized shape, after a Guassian blur, an approximation of the shape's skeleton is computed. The algorithm is implemented on a multi-layered neural network. The output of the network allows different interpretations: we have argued that the calculation of the d_i interneural sums can be made an analog of distance transformation. However, in this case the distance transform is not obtained as a result of successive convolutions but rather as the outcome of a global and parallel calculation of wave-propagation.

Another interpretation of the output was formulated with the help of the D_i integrals. We have seen that the final distribution of sign (D_i) is closely related to the skeleton. The pattern received after the discretized Laplacean operator weighted grassfire mechanism reflects the compromise imposed by the finite resolution of the grid: the propagation is not perfectly isotropic and the resulting pattern, a set of skeleton branches, is not guaranteed to be thin. Notice that the requirements are conflicting; the isotropy improves if the transition zone of the fronts (where the activation is neither 0 nor 1) is widened, and conversely, the skeletal lines approach more and more the ideal one pixel width as the transition zone of the fronts shrinks. The width of the transition zone can be arbitrarily set by the parameter κ. It may be worth noting that if κ is increased the width of the transition lines shrinks as $1/\sqrt{\kappa}$. If at the same time the pixel-size is

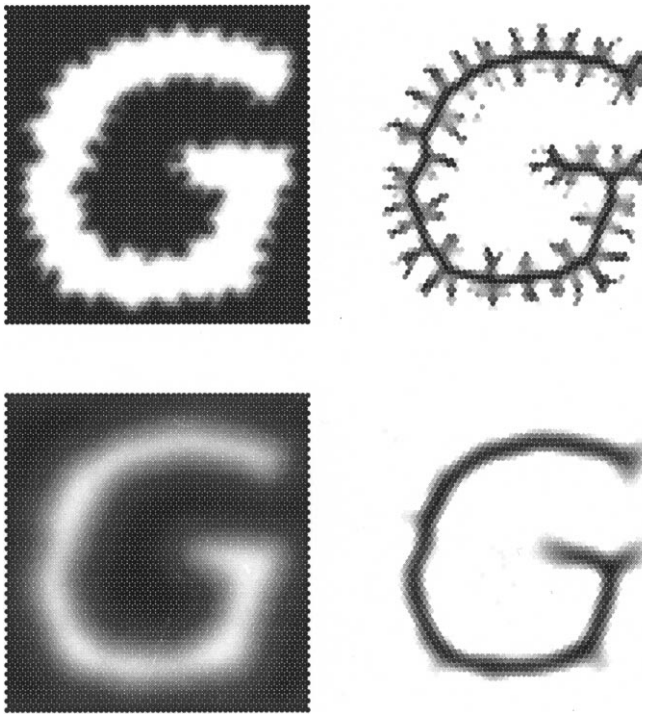

Fig. 8. Skeletonization of shapes with noisy boundaries. The boundary noises are traditionally filtered by a multi-resolution filter system. The top left panel shows the initial activation distribution arising due to a noisy input. The skeleton is (top right) noisy. The bottom left panel shows the initial activation distribution arising due to the same input but on filter system with lower resolution. Correspondingly the skeleton (bottom right) improves, the effect of boundary noises is largely reduced. The easily tuneable filter sizes and the use of a multi-filter approach are natural options of the suggested ANN architecture.

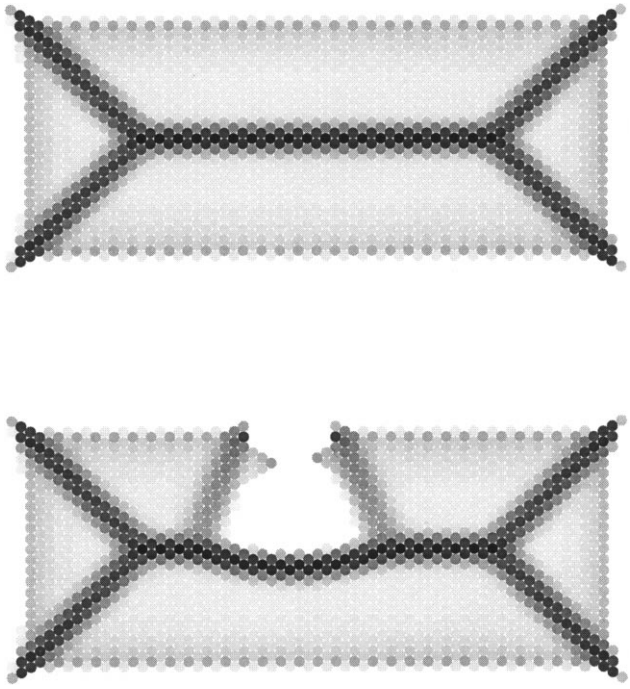

Fig. 9. Generalized skeleton of the Marr rectangle. The introduction of the skeleton activation function solves Marr's problem. The top figure shows the generalized skeleton of a rectangle. The activation of the pixel (Σ) is shown in greyscale. The bottom figure shows the generalized skeleton of a rectangle deformed in a way seen in Fig. 1. Note that an appropriate thresholding of the generalized skeleton may either restore the traditional skeleton or minimize noise effects. Both figures have been calculated by a software implementation of the ANN architecture.

decreased with, say $1/\kappa$, then the transition zone increases compared to the pixel-size (to enhance rotation invariance) but it decreases compared to the size of the image (to improve skeleton representation). Also, isotropy is a function of wavefront thickness measured in pixel units: the broader the wavefront the more isotropic the propagation is. Thus this continuous valued grassfire transformation represents a resolution based compromise or computation power based compromise towards perfect skeletonization.

At the points where two wave-fronts overlap D_i is found to be proportional to the angle included by the fronts, i.e., the relation can be inverted: at every point $\mathbf{x}_i$ it can be determined upon D_i: (i) whether $\mathbf{x}_i$ is a skeletal point, and if so; (ii) at what angle the fire-fronts have met. This inversion only fails in non-ordinary skeletal points, which, however, are negligible. In some cases the value of D_i along the skeleton branches is enough for the reconstruction. In Appendix A we illustrate how simple skeletons can be reconstructed.

Another feature of D_i is that it can be interpreted as an importance measure. Let us recall the example of Marr for the low noise tolerance of skeletonization. In Fig. 9 we depicted the distribution of D_i for the corrupted rectangle of the Marr example. It is noteworthy how the redundant branch fades with the distance from the vertex. In Fig. 8 we have seen, however, that the Laplacean importance measure is not satisfying in all cases. In these cases the initial blurring of the image generally helps. Note that this implementation of skeletonization allows the easy readjustment of the filter sizes as well.

Finally let us mention that most parts of this neural network implementation is built up of elements than can be formed in a self-organized way: the discretization grid as well as a multi-resolution grid system, including the inter-discretization point nearest neighbour connections may be developed in a self-organisation fashion. A short explanation of this possibility is given in Appendix B; here let us rather report the motivations. Skeletonization is widely considered as an image processing method and image processing rarely needs self-organizing methods. Let us consider, however, that skeletonization is not limited to 2D image processing and can be regarded as a general data compressing tool, or, more importantly, just like in the case of biological shapes, skeletonization is a tool that transforms slight shape differences into topological ones that may promote recognition or classification. The first two layers of the neural network may realize a representation of a many dimensional input space. The spatial filters and the interneurons represent the neighbourhood relations of this space and if they can be formed in a self-organizing way then the

skeletonization is possible regardless of that the input space in fact portrays a two-dimensional pixel image or not.

As a particular example consider the problem of characterization of a plug flow reactor with distributed sensory and control systems and designed to control complex reactions (Rojmuckarin et al., 1993). The plug flow reactor model permits control by the chemical and heat fluxes added through the side wall of the reactor. The sensory systems may detect temperature, pressure, spectral information and/or concentration along the reactor in a distributed fashion. The geometry representation may be built up by transients (local extended objects) as they proceed along the reactor. The dimensionality of the representation is not well defined: it is at least 1, since the flow is one-dimensional. Additional dimensions may arise, however, from competing chemical reactions. Now, a given steady state manifests itself through a combination of sensory activities on the geometry representing system allowing skeletonization and sensitive monitoring of small shifts of this steady state.

5. Conclusion

An artificial neural network architecture has been suggested to model the grassfire transformation. The network is made up of three neural layers, from which the first two are connected recurrently while the second is connected to the third in a feed-forward fashion. The grassfire propagation is modelled through the dynamic interaction of the first two layers. We have argued that this part of the algorithm can be considered as a fully parallel implementation of a distance transformation. Based on this distance transform the third layer designates the skeletal points and computes an importance measure.

Computer simulations were presented to illustrate the properties of the wave-front propagation, the usability of the output as an importance measure, the rotation invariance of the skeletonization and the robustness of the architecture. This robustness together with the fully parallel architecture and the exclusive use of simple operations on analog values (addition, multiplication, time integration, thresholding, etc.) renders probable on analog VLSI implementation of the algorithm. In the appendix we shall argue that the output of the algorithm provides the basis for reconstruction of the original shape.

Finally, it has been suggested the skeletonization might be used as a general data compressing tool. Since the neural network including the multi-grid structure we have used to implement the grassfire transformation can be wired in a self-organizing way and is capable to represent the geometry of an unknown input space, it is possible to extend the concept of the skeleton into areas remote to image processing. It was argued that this option may, eventually, lead to applications beyond image processing.

Appendix A Inversion

Skeletonization is a method for building a representation of geometrical data. A new data representation is generally expected to comply with requirements that the original representation could not. The value of such a method is partly judged upon the extent these requirements are satisfied. Another important issue is to decide what sort of information loss, if any, has occurred during the transition from one representation to another. To investigate this latter point the examination of reconstructability is a promising method.

The skeleton alone is not enough for the reconstruction of the original shape but if each skeletal point is labelled with the collision time of the wave-fronts that designated it, we obtain a full representation. There are cases, however, when less information is enough for the reconstruction.

The reconstruction of simple (planar, not self-intersecting) polygons is much simpler than that of ordinary curved shapes. It is known from elementary geometry that the reconstruction is possible if the location of the vertices, the size of the vertical angles and the direction of the bisectors are known. We shall show how the above mentioned data can be peeled out of the output of the algorithm described in the article. There are two important steps of this demonstration: (1) we will argue that the vertices of the polygon and the end-points of the skeleton branches are determined by each other in a unique way; and (2) that the direction of the bisectors and the angles can be approximated from the output of our detectors in an appropriate neighbourhood of the vertices.

The skeleton is generally defined with the help of closed balls[5] (see, e.g., Serra, 1988). In this case the vertices of the polygon are contact points of its skeleton[6]. It should be demonstrated that there are no other contact points than the vertices. It is known that the contact points of a given set A can be obtained as the set-difference of the adherence of the set, $\bar{A}$, and A itself. Therefore, if the union of the set of vertices and the skeleton is set-equal to the adherence of the skeleton then the set of vertices is demonstrated to be set-equal to the set of contact points. The demonstration may go along the verification of mutual involvement.

The forward direction is simple: a vertex cannot be in the skeleton since there can be no maximal ball drawn around it, on the other hand, in every suitably small neighbourhood of the vertex the points of the bisector are in the skeleton, therefore every vertex is a contact point of the skeleton. But the points and the contact points of a set are necessarily in the adherence of it, i.e., the union of the skeleton and the vertices is a subset of the adherence of the skeleton.

The backward direction is more tedious to prove. However, it is enough to show that the union of the skeleton and the set of vertices is a closed set, since all closed sets that

[5] In this article we restrict ourselves to the skeletonization of closed sets.

[6] Recall that a point x is called a contact point of a set A, if $x \notin A$ but in each neighbourhood of x there is a point, x', such that $x' \in A$.

contain the skeleton as a subset contain their adherence sets as well. This can be proved by recalling that closed sets contain the limit points of convergent series of its own points.

Once the vertices of the polygon are designated by the end-points of its skeleton the direction of the bisectors can be obtained easily. It is clear that every vertex has a neighbourhood, in which there are no other points of the polygon than those lying on one of the two sides that originate from the vertex. In such a neighbourhood the skeleton is a line-segment that lies on the bisector. Moreover, the importance measure along this line-segment is a one-to-one and, therefore, an invertible function of the angle included by the two sides. The grassfire and the detector system on a discretizing system approximates that function.

Appendix B Self-organization

It has been mentioned in the Discussion that most elements of the neural implementation can be formed in a self-organizing (SO) way. In this appendix we briefly summarize what is meant on geometry representation and how the SO formation of neural networks is performed. This topic is explained in detail in (Szepesvári and Lörincz, 1996).

The problem of geometry representation is the following: there is an external world that we only perceive through a set of transducers. Any information we obtain is transmitted by these transducers. The objects of the external world are mapped into the output spaces of the transducers. The product of these output spaces will be called the pattern space. For example, let us image that a region of the 3D space is monitored with two pixel cameras[7]. The cameras see the same region from different points of view. Each pixel can be regarded as a individual transducer; the pattern space is the product of the discretized intensity of the pixel. The dimension of the pattern space is the number of the transducers; here the number of the pixels. The task of geometry representation is two reveal the neighbourhood relations of the external space.

Appendix B.1 Adaptive vector quantization

Since the dimensionality of the pattern space is typically very high, an obvious thing to do is to reduce the dimensionality. This can be done with the help of correlations in the transducer signals, e.g., by a process called adaptive vector quantization. Let us image that we have a fixed number of prototype vectors in the pattern space. To reduce the dimensionality of the incoming data we would like to assign every pattern to the prototype vector that matches to it best (the number of prototype vectors is much lower than the dimensionality of the pattern space). The function that measures the error of this assignment is called a similarity function. The similarity function is monotonically decreasing and larger it is, the pattern is closer to the prototype vector. The usual choices for it are the inner product or the Euclidean distance of the two vectors. Each prototype vector can be implemented by an artificial neuron that is connected to every pixel of the pattern space via feed-forward connections. The strength of these connections encode the components of the prototype vector.

Prototype vectors, i.e., the feed-forward connections, can be formed adaptively in a process known as learning. During learning, patterns are presented subsequently and for each the best matching prototype vector is determined and slightly moved in the direction of the pattern. In neural networks terms this algorithm is called a ‘winner-takes-all’ (WTA) method because in each iteration only the best matching prototype vector is modified (Grossberg, 1976, 1987).

It is proved that under certain circumstances the lateral connections (the edges of the graph) that make a correct geometry representation can be formed adaptively with the help of the competitive Hebb rule, which for each pattern strengthens the connection between the two neurons with highest similarity (Martinetz, 1993). Moreover, the formation of lateral connections and the construction of the embedding can be done simultaneously if the embedding (i.e., the prototype vectors) is updated according to the WTA rule and the connections according to the competitive Hebb rule. The conditions we have mentioned, together with some technical restrictions, include the important requirement that the similarity function preserve the overlap (Szepesvári, 1993; Szepesvári et al., 1994; Szepesvári and Lörincz, 1996). Similar ideas but for feature space, not for sensory space, have been developed by Martinetz and Schulten (Martinetz, 1993; Martinetz and Schulten, 1994).

Self-organization also allows the development of multi-resolution discretizing systems by using the same WTA method and competitive Hebb rule with slightly modified updating of the feedforward connection (e.q. 4.15 of Kohonen, 1984). Details are published elsewhere (Rozgonyi et al., 1996; Szepesvári and Lörincz, 1996).

[7] Those who find that this example is too obvious may consider the other one we have mentioned, when the external world is the inside of a chemical reactor and the transducers are pressure- and thermometers.

References

Arcelli, C., & di Baja, G.S. (1985). A width-independent fast thinning algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI7*, 463–474.

Arcelli, C., & di Baja, G.S. (1989). A one-ass two-operation process to detect the skeletal pixels on the 4-distance transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*, 411–414.

Blum, H. (1967). A transformation for extracting new descriptions of shape. In: Wathen & Dunn (Eds.), *Models for the perception of speech and visual form* (pp. 362–380). Cambridge, MA: MIT Press.

Brandt, J.W., & Algazi, V.R. (1992). Continuous skeletons computation by voronoi diagram. *CVGIP: Image Understanding, 55*, 329–338.

Gauch, J.M., & Pizer, S.M. (1993). Multiresolution analysis of ridges and valleys in grey-scale images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*, 635–646.

Grossberg, S.A. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences*, *60*, 758–765.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological Cybernetics*, *23*, 121–134.

Grossberg, S. (1987). From interaction activation to adaptive resonance theory. *Cognitive Science*, *11*, 23–63.

Ho, S.B., & Dyer, C.R. (1986). Shape smoothing using medial axis transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-8*, 512–520.

Kirkpatrick, D.G. (1979). Efficient computation of continuous skeletons. In: *IEEE 20th Annual Symposium on Foundations of Computer Science*, pp. 18–27.

Kohonen, T. (1984). *Self organisation and associative memory*. Springer-Verlag, Berlin.

Maragos, P.A., & Schafer, R.W. (1986). Morphological skeleton representation and coding of binary images. *IEEE Transactions on Acoustics Speech and Signal Processing*, *34*, 1228–1244.

Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freemans.

Martinetz, T. (1993). Competitive Hebbian learning rule forms perfectly topology preserving maps. In: *Proceedings of ICANN* (pp. 427–434). Berlin: Springer Verlag.

Martinetz, T., & Schulten, K. (1994). Topology representing networks. *Neural Networks*, *7 (3)*, 507–522.

Martínez-Pérez, M.P., Jimenéz, J., & Navalón, J.L. (1987). A thinning algorithm based on contours. *Comput. Vision Graphics Image Process*, *39*, 186–201.

Meyer, F. (1988). Skeletons in digital spaces. In: Serra, J. (Ed.), *Image analysis and mathematical morphology* (Vol. 2: Theoretical Advances, ch. 12, pp. 257–296). London: Academic Press.

Montanari, U. (1968). A method of obtaining skeletons using a quasi-euclidean distance. *J. Assoc. Comput. Mach.*, *15*, 600–624.

Oláh, M., & Lörincz, A. (1995). Analog VLSI implementation of grassfire transformation for generalized skeleton formation. In: *Proceedings of International Conference on Artificial Neural Networks*, Industrial Conference, Paris, 1995.

Rojmuckarin, A., Floudas, C., Rabitz, H., & Yelter, R.A. (1993). Optimal control of a plug flow reactor with a complex reaction mechanism. *Journal of Physical Chemistry*, *97*, 11689–11695.

Rosenfeld, A. (1984). *Multiresolution image processing and analysis*. Berlin: Springer-Verlag.

Rozgonyi, T., Balázs, L., Fomin, T., & Lörincz, A. (1996). Self-organized formation of a set of scaling filters and their neighbouring connections. *Biological Cybernetics*, *75*, 37–47.

Schmitt, M. (1989). Some examples of algorithms analysis in computational geometry by means of mathematical morphology techniques. In: *Geometry and robotics*. Berlin: Springer-Verlag.

Serra, J. (1982). *Image analysis and mathematical morphology*. London: Academic Press.

Serra, J. (1988). *Image analysis and mathematical morphology*, *vol. 2: theoretical advances*. London: Academic Press.

Shih, F.Y., & Pu, C.C. (1990). Medial axis transformation with single-pixel and connectivity preservation using euclidean distance computation. In: *Proceedings of the 10th International Conference on Pattern Recognition* (Vol. 1, pp. 723–725). Atlantic City, NJ.

Smith, R.W. (1987). Computer processing of line images: A survey. *Pattern Recognition*, *20*, 7–15.

Szepesvári, C. (1993). The role of local connections in competitive neural networks: collective learning and representation of geometry. Master's Thesis, Attila Jozsef University of Szeged (in Hungarian).

Szepesvári, Cs., & Lörincz, A. (1996). Approximate geometry representations and sensory fusion. *Neurocomputing*, *July 12 (2) 3*, 267–287.

Szepesvári, C., Balázs, L., & Lörincz, A. (1994). Topology learning solved by extended objects: a neural network model. *Neural Computation*, *6*, 441–458.