An Analog Implementation of Discrete-Time Cellular Neural Networks

Hubert Harrer, *Student Member, IEEE,* Josef A. Nossek, *Senior Member, IEEE,* and Rudolf Stelzl

Abstract— An analog circuit structure for the realization of discrete-time cellular neural networks (DTCNN's) is introduced. The computation is done by a balanced clocked circuit based on the idea of conductance multipliers and operational transconductance amplifiers. The circuit is proposed for a one-neighborhood on a hexagonal grid, but can also be modified to larger neighborhoods and/or other grid topologies. A layout was designed for a standard CMOS process, and the corresponding HSPICE simulation results are given. A test chip containing 16 cells was fabricated, and measurements of the transfer characteristics are provided. The functional behavior is demonstrated for a simple example.

I. INTRODUCTION

A NALOG CMOS circuits for neural networks have recently received growing attention. In contrast to digital solutions, the transistors are not only used as switches. Their nonlinear functionality is applied for implementing complex circuit components (e.g., multipliers) and nonlinear characteristics. This leads to a drastically reduced number of devices and also decreases power consumption. Large networks can be realized if architectures suitable for VLSI are applied. However, this is only possible if the precision requirements are moderate. Tolerances of 1 to 10% have to be accepted in order to achieve a high cell density with simple circuit structures. This is satisfied for discrete-time cellular neural networks as introduced in [1], which represent an efficient architecture for image processing and pattern recognition. They are a combination from general linear threshold networks, where the local cell connectivity and translational invariance of the weights were transferred from cellular neural networks [2]. The system can solve global tasks, although the cells are only locally connected. This is achieved by a propagation of signals like waves traveling from cell to cell. Applications have been found for discrete convolution, connected component detection, hole filling, shadow detection, concentric contouring, increasing and decreasing objects step by step, searching for objects with minimal distance, and oscillation [1]. The simple structure of a single cell and the local connectivity make them well suited for VLSI implementation.

Discrete-time cellular neural networks (DTCNN's) are defined by the algorithm

$$x^c(k) = \sum_{d \in N_r(c)} a_d^c y^d(k) + \sum_{d \in N_r(c)} b_d^c u^d + i^c \qquad (1)$$
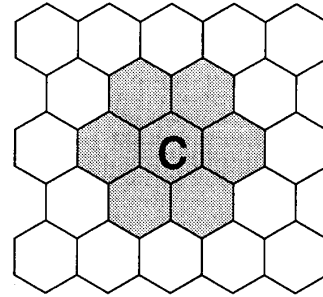
Fig. 1. Neighbor cells of cell c for $r = 1$ are shaded.

$$y^c(k) = f(x^c(k-1)) = \begin{cases} 1 & \text{for } x^c(k-1) > 0 \\ -1 & \text{for } x^c(k-1) < 0. \end{cases} \qquad (2)$$

The output $y^c(k)$ of a cell c is binary and is determined by the sign of $x^c(k-1)$. It is undefined for $x^c(k-1) = 0$, where a random value $+1$ or -1 is assumed by some noise. The value $+1$ denotes a black pixel, -1 a white pixel. The value of x^c is controlled by the inputs and outputs of adjacent cells d within an *r-neighborhood* $N_r(c)$. This is defined as the set of adjacent cells within a distance r including cell c. For $r = 1$ the neighbor cells of cell c are shown in Fig. 1 on a hexagonal grid. The outputs y^d are fed back multiplied by the *feedback parameters* $a_d^c \in \mathbb{R}$, and the constant inputs $u^d \in \mathbb{R}$ are multiplied by the *control parameters* $b_d^c \in \mathbb{R}$. A set of feedback, control coefficients and the thresholds is called a template. Notice that the template coefficients are *translationally invariant*, which means that each cell is influenced identically by its neighbors. This reduces the different weights to a small number that can be implemented by global bus lines in VLSI realizations. The value of $i^c \in \mathbb{R}$ is constant and used for adjusting a threshold. For most applications, it is identical for all cells, which is indicated by the notation $i \in \mathbb{R}$. The system is clocked and only binary values $(-1, +1)$ are weighted by the coefficients of the feedback operator. Before an iteration is started, the initial values $y^c(0)$ have to be defined. They can be considered as a second cell input and their initialization is used for processing an initial pattern.

Abbreviating the notation, the time-independent part of (1) is defined as the cell bias $k^c(u)$, which again is a function of the constant inputs and the threshold values [3]:

$$k^c = k^c(u) = \sum_{d \in N_r(c)} b_d^c u^d + i^c. \qquad (3)$$

If the template coefficients are chosen such that

$$\Delta = \min_{c,k} \left| \sum_{d \in N_r(c)} a_d^c y^d(k) + \sum_{d \in N_r(c)} b_d^c u^d + i^c \right| \qquad (4)$$

is large enough, the algorithm is relatively insensitive to parameter tolerances of a_d^c, b_d^c, or i^c being smaller than Δ. The value of Δ corresponds to the minimum absolute value of the state variable x^c. It can be considered as a safety margin, because any tolerances Δa_d^c, Δb_d^c and Δi^c leading to a state variable tolerance $\Delta x^c(k)$ do not matter if the absolute value of $\Delta x^c(k)$ is smaller than Δ. Then, it cannot change the sign of the state variable, and the correct output value is preserved. For most applications, such a lower bound of the state variable can be given if the continuous inputs are restricted to binary values. This is extremely important in VLSI realization, because small fabrication tolerances do not disturb the system behavior. It is possible to train the network to a desired insensitivity when using the design algorithm described in [4].

Remark: A DTCNN containing a large enough neighborhood can be regarded as a single-layer perceptron with a comparator nonlinearity and feedback connections or as a discrete Hopfield net with self-feedback. However, large neighborhoods are not well suited for simple VLSI implementation.

For the realization of a DTCNN different approaches are possible:

- A low-cost solution is achieved using general-purpose signal processors if the speed requirements can be met. Such a hardware accelerator board for DTCNN's would be identical to that described for CNN's in [5]. Only the software modules changed.

- An optical implementation introduced in [6] has the advantage of a large cell density and a simple realization of larger neighborhoods.

- A digital implementation would lead to a systolic array. Because of the binary output states -1 and $+1$, DTCNN's do not need digital multipliers if the weights or their complementary values are summarized with the constant term k^c. In this case the cell kernel consists of an adder, where the output state is determined by the sign bit.

- An analog realization allows high cell density for simple circuit structures and low power consumption. In the case of analog preprocessing the expensive analog digital conversion is avoided.

Because each paradigm has its own advantages and disadvantages, the solution to be preferred is dictated by the application. Here, it was concentrated on the analog one, which leads naturally to applications such as analog preprocessing for gray-scale scanners, i.e., processing of sensors with continuous outputs.

II. NETWORK STRUCTURE

For the algorithm (1) and (2), the network structure of a single cell is given in Fig. 2. The variables are replaced by voltages and the iteration steps by discrete time instances.
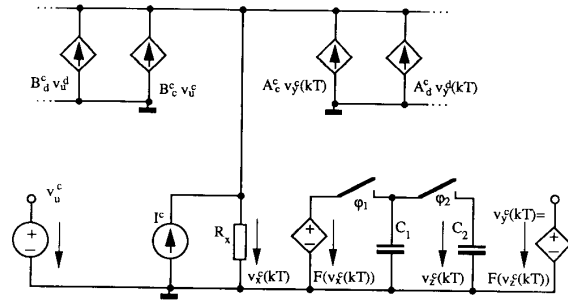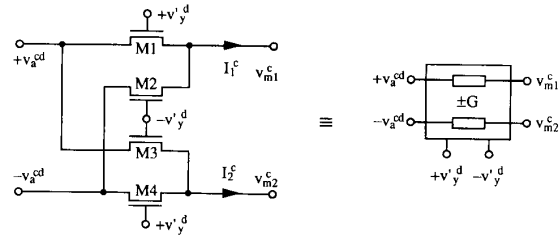

Fig. 2. Network structure of a single cell.

TABLE I
NORMALIZATION OF THE VARIABLES AND NETWORK ELEMENTS

$u^c = \dfrac{v_u^c}{V_{\text{sat}}}$	$a_d^c = A_d^c R_x$
$x^c(k) = \dfrac{v_x^c(kT)}{V_{\text{sat}}}$	$b_d^c = B_d^c R_x$
$z^c(k) = \dfrac{v_z^c(kT)}{V_{\text{sat}}}$	$i^c = I^c \dfrac{R_x}{V_{\text{sat}}}$
$y^c(k) = \dfrac{v_y^c(kT)}{V_{\text{sat}}}$	$\tau = \dfrac{kT}{T} = k$

M1, M2, M3, M4: $L = 100\mu$ $W = 4\mu$

Fig. 3. MOS implementation of conductance multipliers.

The next iteration step of the system is obtained after one period of the dual nonoverlapping clock signals φ_1 and φ_2. Neglecting parasitic transients, the time instance kT can be chosen arbitrarily during the "high" phase of φ_1. The multiplications are done by linear voltage-controlled current sources. The total current sum is transformed into a voltage $v_x^c(kT)$ by the resistor R_x. The voltage-controlled voltage sources $F(v_x^c(kT))$ and $F(v_z^c(kT))$ have the nonlinear characteristic

$$F(v) = \begin{cases} V_{\text{sat}} & \text{if } v > 0 \\ -V_{\text{sat}} & \text{if } v < 0. \end{cases} \qquad (5)$$

The capacitor C_1 is charged during the "high" phase of φ_1 to $+V_{\text{sat}}$ or $-V_{\text{sat}}$. During the "high" phase of φ_2 (φ_1 "low"), the capacitors C_1 and C_2 are connected in parallel. The voltage $v_z^c(kT)$ settles after the transient to

$$v_z^c(kT) = \frac{C_1 F(v_x^c((k-1)T)) + C_2 v_z^c((k-1)T)}{C_1 + C_2}. \qquad (6)$$
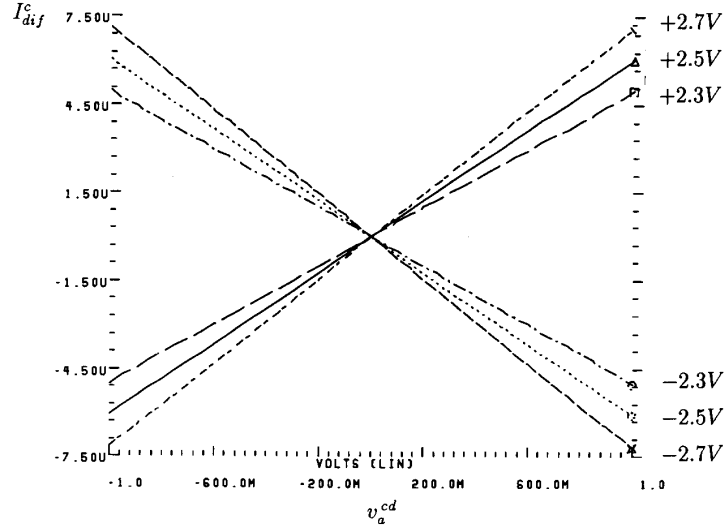
Fig. 4. Difference current I^c_{dif} for $\hat{v}^d_y = \pm 2.3$ V, ± 2.5 V, and ± 2.7 V.

Because of $|F(v^c_x)| = V_{\mathrm{sat}}$ and $|v^c_z| \leq V_{\mathrm{sat}}$ $\forall$ k, it follows, for $C_1 > C_2$,

$$v^c_y(kT) = F(v^c_x((k-1)T))$$

$$= \begin{cases} V_{\mathrm{sat}} & \text{if } v^c_x((k-1)T) > 0 \\ -V_{\mathrm{sat}} & \text{if } v^c_x((k-1)T) < 0. \end{cases} \quad (7)$$

Then we obtain

$$v^c_y(kT) = F\left(R_x\left(\sum_{d \in N_r(c)} A^c_d v^d_y((k-1)T)\right.\right.$$

$$\left.\left. + \sum_{d \in N_r(c)} B^c_d v^d_u + I^c\right)\right). \quad (8)$$

Applying the normalization in Table I with an arbitrary reference voltage V_{sat}, the algorithm (1), (2) is obtained.

Remark: Notice that the outputs $v^c_y(kT)$ have to be constant during the high phase of φ_1, because they are fed back by the voltage-controlled current sources. This is achieved by using two capacitors, C_1 and C_2, where the new output value is charged in C_1 while the previous output value is stored in C_2.

The nonlinear voltage source $F(v^c_x(kT))$ ensures that C_1 is always charged to $+V_{\mathrm{sat}}$ or $-V_{\mathrm{sat}}$; therefore, it always determines the sign of $v^c_z((k+1)T)$. The sign of the voltage $v^c_z(kT)$ is identical to the output state in the subsequent iteration and is extracted by the voltage source $F(v^c_z(kT))$.

III. CONDUCTANCE MULTIPLIER

If accuracy requirements are moderate. an efficient MOS circuit for the linear voltage-controlled current sources, which was proposed in [7], is used consisting of only four transistors. It performs a four quadrant multiplication if balanced signals are provided and the transistors are operated in *strong inversion.* Their MOS implementation is shown in Fig. 3.

A modified operation mode for the multiplication of the binary outputs $y^d \in \{-1, 1\}$, where only two transistors
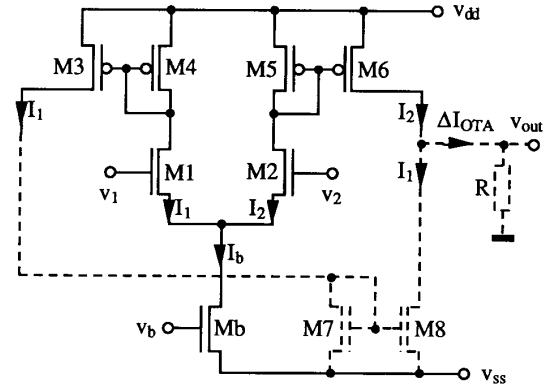

Fig. 5. Circuit structure of an operational transconductance amplifier.

operate in strong inversion at the same time, is described below. Assuming large *sign symmetric* gate potentials $+\hat{v}^d_y$ and $-\hat{v}^d_y$, defined by

$$\hat{v}^d_y = v^d_y + v_{tb}, \quad (9)$$

two transistors, e.g., M1 and M4, are operated in strong inversion, while M2 and M3 are turned off. The drain source currents of M1 and M4 can be separated in a linear term, I^c_L, and a nonlinear term, I^c_N. For M1, the linear part, I^c_{L1}, is described by

$$I^c_{L1} = G\left(v^{cd}_a - v^c_{m1}\right) \quad (10)$$

with

$$G = \frac{W}{L}\mu C_{\mathrm{ox}}\left(\hat{v}^d_y - v_{tb}\right) = \frac{W}{L}\mu C_{\mathrm{OX}}v^d_y, \quad (11)$$

where v_{tb} is the transistor voltage with respect to bulk; L and W are the transistor length and width; μ is the effective
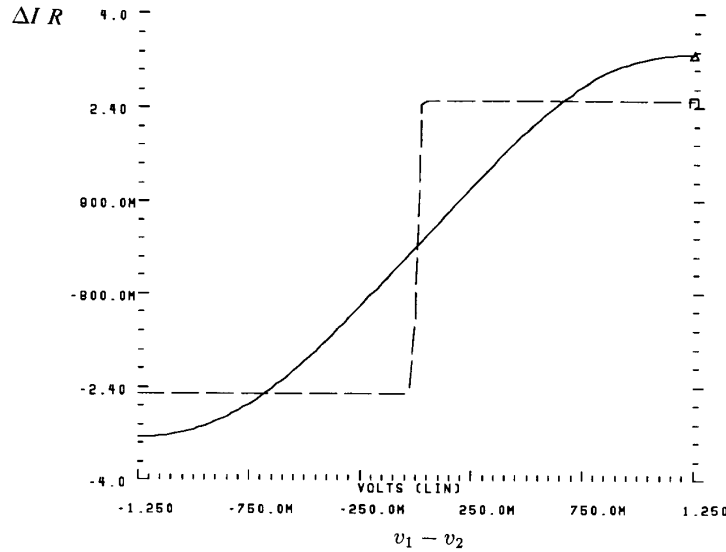
Fig. 6. The dc transfer characteristic of the voltage follower and the comparator.

mobility; and C_{ox} is the oxide capacitance per unit area. Similar considerations can be invoked for I_2^c and lead to

$$I_{L2}^c = G\left(-v_a^{cd} - v_{m2}^c\right). \tag{12}$$

It is shown by a Taylor series expansion (see the Appendix) that the difference of the nonlinear parts $I_{N1}^c - I_{N2}^c$ can be neglected if the conditions

$$\left|v_a^{cd}\right| < 1V \quad \text{and} \quad v_{m1}^c = v_{m2}^c$$

are satisfied. For the difference current I_{dif}^c it follows that

$$I_{\text{dif}}^c \approx I_{\text{difL}}^c = I_{L1}^c - I_{L2}^c = 2Gv_a^{cd}. \tag{13}$$

The circuit is approximated by two linear conductances of equal value G for identical transistor lengths and widths. When changing the sign of the gate potentials $+\hat{v}_y^d$ and $-\hat{v}_y^d$, M1 and M4 are turned off. Now, M2 and M3 are operated in strong inversion. The circuit is described by two negative conductances $-G$ because of the crossed-over inputs $+v_a^{cd}$ and $-v_a^{cd}$. Hence, the gate potentials control the sign of the conductance G. This corresponds to a multiplication by ± 1 as required for the weighted outputs of DTCNN's with

$$A_d^c = 2\frac{W}{L}\mu C_{\text{ox}}v_a^{cd}, \tag{14}$$

where the multiplied voltage $\hat{v}_y^d$ is reduced by the threshold voltage v_{tb}. However, this was taken into consideration by the transformation in (9).

The simulation results are summarized in Fig. 4 for $v_{m1}^c = v_{m2}^c = 0$ V, where the ratio of the transistor lengths and widths was chosen as $L/W = 100 \ \mu\text{m}/4 \ \mu\text{m} = 25$. The power consumption amounts to 0.02 mW for a weight voltage $v_a^{cd} = 1$ V.

IV. OPERATIONAL TRANSCONDUCTANCE AMPLIFIERS

Operational transconductance amplifiers (OTA's) transform a differential input voltage, $v_2 - v_1$, into a differential output current, $I_2 - I_1$. The transformation is linear within a certain range, and goes into saturation if the linear range is left. Their circuit structure is illustrated in Fig. 5. For a detailed analysis, refer to [8] and [9]. The analog implementation of DTCNN's uses the OTA for two different tasks:

- A linear transfer characteristic realizes an analog memory if the gate of M2 is connected to ground and a capacitor is applied to the gate of M1. Then the difference of the output currents, ΔI_{OTA}, depends linearly on the capacitor voltage v_1:

$$\Delta I_{\text{OTA}} = I_2 - I_1 = g_m v_1, \tag{15}$$

where g_m denotes the conductance of the amplifier.

- The OTA is also used as a comparator for the nonlinear voltage-controlled voltage source $F(v_x^c(kT))$ in Fig. 2. Here, a steplike transfer characteristic is required. The current I_1 is mirrored by M7 and M8 to obtain a single ended output current $\Delta I_{\text{OTA}} = I_2 - I_1$. This is transformed into a voltage by a linear resistor and is indicated by the dashed part in Fig. 5.

The simulated dc transfer characteristic is given in Fig. 6 with the corresponding transistor sizes in Table II.

Operational transconductance amplifiers have the following desirable properties:

- low power consumption,
- small chip area,
- fast transient behavior,
- high flexibility in setting the conductance, g_m, and the saturation current.

The magnitude of the currents I_1 and I_2 can be increased or decreased by changing the W/L ratios of the current mirrors.

TABLE II
TRANSISTOR DIMENSIONS (W/L) IN μm

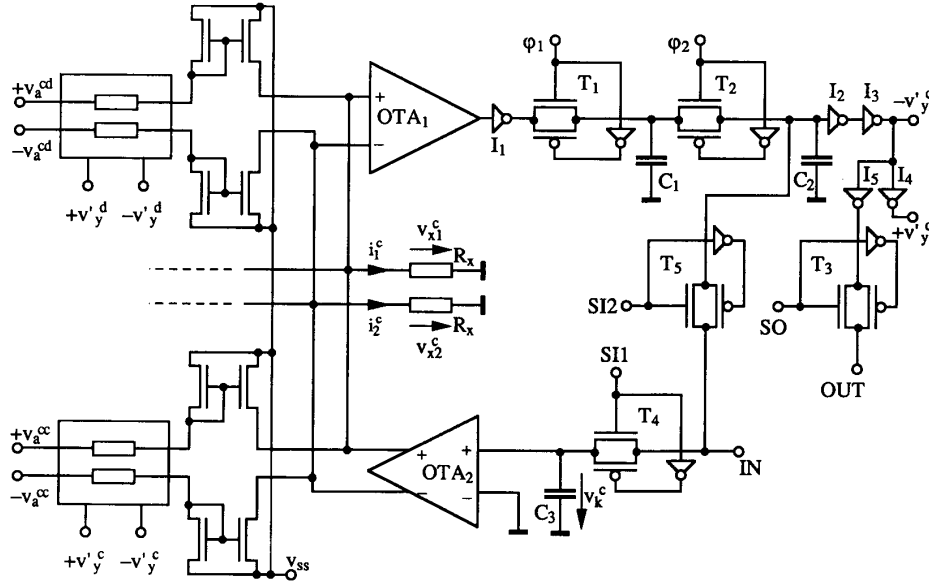|	Mb	M1	M2	M3	M4	M5	M6	M7	M8
Analog memory	8/4	4/10	4/10	9/4	8/4	8/4	9/4		
Comparator	4/4	45/4	45/4	4/4	4/4	4/4	4/4	4/4	4/4

Fig. 7. Realization of a single cell.

V. CIRCUIT STRUCTURE

Before fixing the circuit structure, the grid topology and the neighborhood size have to be determined. We have chosen a 1-neighborhood on the hexagonal grid, where each cell is connected only to its six nearest neighbors. To simplify the circuit architecture, the constant term

$$v_k^c = R\left(\sum_{d \in N_r(c)} B_d^c v_u^d + I^c\right), \tag{16}$$

which corresponds to the constant cell bias in (3), is assumed to be processed on a separate chip [10]. This has the advantage of reducing the interconnections between adjacent cells, as only the outputs are fed back. Then, the circuit for a single cell is shown in Fig. 7.

The capacitance C_3 is used as an analog memory for v_k^c that is different for each cell in general. It is charged by a bus IN connected to all cell inputs in the same row of the network. The column selection is done by the transmission gate T_4 using the bus SI1. The operational transconductance amplifier OTA2 provides differential output currents. If the transconductance of the OTA is determined as $g_m = 1/R$, a differential output current

$$\Delta I_{OTA2} = g_m v_k^c = \frac{1}{R} v_k^c = \sum_{d \in N_r(c)} B_d^c v_u^d + I^c \tag{17}$$

is obtained. The transconductance, g_m, and the saturated difference current can be adjusted by the bias voltage, v_b, of the OTA.
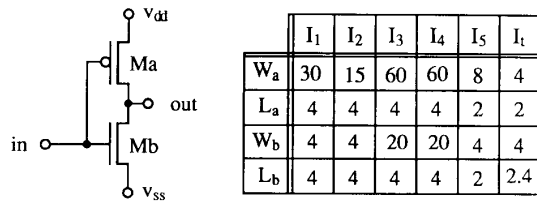
Because the conductance multipliers require equal output potentials, $v_{m1}^c = v_{m2}^c$, they are strictly decoupled by current mirrors with transistor lengths and widths of 4 μm. The differences of their gate source voltages are negligible when using identical potentials, v_{ss}, for all current mirrors. Simulations have shown that the linear behavior was preserved.

All output currents are summed differentially and transformed into voltages $v_{x1}^c(kT)$ and $v_{x2}^c(kT)$ by linear resistors R_x. They are realized by two complementary MOS transistors connected in parallel. Operated in the triode range, their behavior is like that of a linear resistor, whose value depends on the gate potentials and on the transistor lengths and widths. Notice that different W/L ratios have to be used for the compensation of the nonlinearities because of the different mobilities of the p and n transistors ($W_p = 4\ \mu$m, $L_p = 40\ \mu$m, $W_n = 4\ \mu$m, $L_n = 100\ \mu$m).

The difference voltage

$$v_{x1}^c(kT) - v_{x2}^c(kT) = R_x(i_1^c(kT) - i_2^c(kT)) \tag{18}$$

is applied to the OTA$_1$, which has a comparator characteristic realizing the nonlinear function in (5). The accuracy requirements for the resistors R_x are low, because only the sign of the voltage $v_{x1}^c(kT) - v_{x2}^c(kT)$ is of importance.

Fig. 8. Inverter stage and transistor dimensions in μm.

TABLE III
CIRCUIT SIZES AND TECHNICAL DATA OF THE DESIGNED CELL

Conductance multiplier	$140~\mu\mathrm{m} \times 40~\mu\mathrm{m} = 0.0056~\mathrm{mm}^2$
Analog memory	$60~\mu\mathrm{m} \times 50~\mu\mathrm{m} = 0.0030~\mathrm{mm}^2$
Comparator	$90~\mu\mathrm{m} \times 50~\mu\mathrm{m} = 0.0045~\mathrm{mm}^2$
Capacitors	$290~\mu\mathrm{m} \times 30~\mu\mathrm{m} = 0.0087~\mathrm{mm}^2$
Digital part	$180~\mu\mathrm{m} \times 70~\mu\mathrm{m} = 0.0126~\mathrm{mm}^2$
Total cell size	$290~\mu\mathrm{m} \times 275~\mu\mathrm{m} = 0.0795~\mathrm{mm}^2$
Number of cells per cm^2	ca. 1200
Number of transistors per cell	106
Maximum clock rate	3.3 MHz
Power consumption of a single cell	0.9 mW

However, there are crucial restrictions on the operation range of $v_{x1}^c(kT)$ and $v_{x2}^c(kT)$. If their values are too close to the v_{dd} or v_{ss} of OTA$_1$, the comparator does not work correctly, because its transistors leave saturation and turn off. This occurs if the resistors R_x are chosen too large. The allowed range of the difference voltage $v_{x1}^c(kT) - v_{x2}^c(kT)$ is maximum if the differential bias currents for $v_a^{cd} = 0$ V of all the conductance multipliers cancel approximately the differential bias currents of OTA$_2$ for $v_k^c = 0$ V. This cancellation is achieved by use of complementary current mirrors for the OTA (p mirrors) and the controlled conductances (n mirrors). Then $v_{x1}^c(kT) = 0$ V and $v_{x2}^c(kT) = 0$ V is obtained for zero weights $\left(v_a^{cd} = 0 \text{ V.} \, v_k^c = 0 \text{ V}\right)$. The operating potentials v_{x1}^c and v_{x2}^c of the comparator are also adjusted by changing the ground potential of the resistors R_x.

The output voltage of the comparator is applied to the digital inverter I_1. It consists of two complementary transistors, whose W/L ratios are determined to obtain a symmetric transfer characteristic. Inserting this inverter stage accelerates the transient behavior. The capacitor C_1 is charged during the "high" phase of φ_1 (φ_2 "low") by the transmission gate T_1. During the "high" phase of φ_2 (φ_1 "low") the transmission gate T_2 connects the capacitors C_1 and C_2 in parallel. The voltage of the capacitor C_2 is amplified by two inverting stages, I_2 and I_3. The inverters I_3 and I_4 generate the gate potentials $\pm \hat{v}_y^c$ for the conductance multipliers. They are connected to all adjacent cells within the 1-neighborhood and have to drive the capacitive loads of the conductance multipliers. The power consumption of the multipliers is mainly drawn from the seven weight voltages $\pm v_a^{cd}$ being identical for all cells; these voltages are applied by global bus lines.

The output data transfer is done by the inverter I_5 and the bus OUT, which is common for a single row of cells in the network. The column selection is performed by the transmission gate T_3 using the column bus SO. Chips can be interconnected in a simple way, because the output values are only binary. A special hardware support for transferring output states of boundary cells would be useful. The initial state of C_2 is set by the bus IN, with T_5 selected by the bus SI2.

Fig. 8 shows the circuit structure for the inverter stages with the corresponding transistor lengths and widths in μm, where I_t denotes the inverter of the transmission gates.

VI. LAYOUT AND SIMULATION RESULTS

The layout of the described circuit structure was designed for a 1.5 μm single-poly, double-metal process. It is illustrated

in Fig. 9, where the second metal layer is omitted. The large transistors of the seven controlled conductances can be recognized clearly. The conductances are neighbored by current mirrors providing the differential currents for the comparator. The summing current line is not visible, because it is implemented in metal 2, as are the power supplies and global bus lines.

The capacitors are placed above the four bottom conductances. They are realized by a layer of polysilicon over diffusion area, where the thin oxide layer is used as dielectric. The OTA with the capacitor C_3, storing the constant cell bias k^c, is in the middle on the left side. The comparator is put below the single controlled conductance at the top right corner. It is followed by the digital part, mainly consisting of inverter stages and transmission gates. It is emphasized that the capacitor $C_1 = 1$ pF at the right border is much larger than the adjacent $C_2 = 0.5$ pF to guarantee correct functional behavior when discharging C_2. The cell contains all connecting wires as global bus lines or power supplies. A network is simply built up by placing the cells to a hexagonal structure. While the lower and upper neighbors are centered at the cell borders, the left and right neighbors are shifted to a hexagon.

The results given in this section are based on HSPICE level 6 transistor simulations of the designed layout. Technical data are summarized in Table III.

Simulation results have shown that the maximum clock rate is restricted by the controlled conductances. Because the transistors are very long, it takes some time until the output currents have reached their final values, when the voltages $\pm \hat{v}_y^d$ change their signs. This settling time depends also on the weights $\pm v_a^{cd}$, where extremely small voltages are critical if the difference voltage of the comparator gets too low. The maximum clock rate has to make sure that also for small weight coefficients the sign of the state variable is decided correctly. The value in Table III corresponds to a weight coefficient $v_a^{cd} = 0.05$ V, which generates a differential input voltage of 0.03 V for the comparator. The power consumption of a single cell was simulated for a clock frequency of 1 MHz.

VII. MEASURED RESULTS

For experimental purposes, a chip was fabricated containing 4 by 4 cells. The cell core is illustrated in Fig. 10. The chip size is determined by the 84 input and output buffers, because
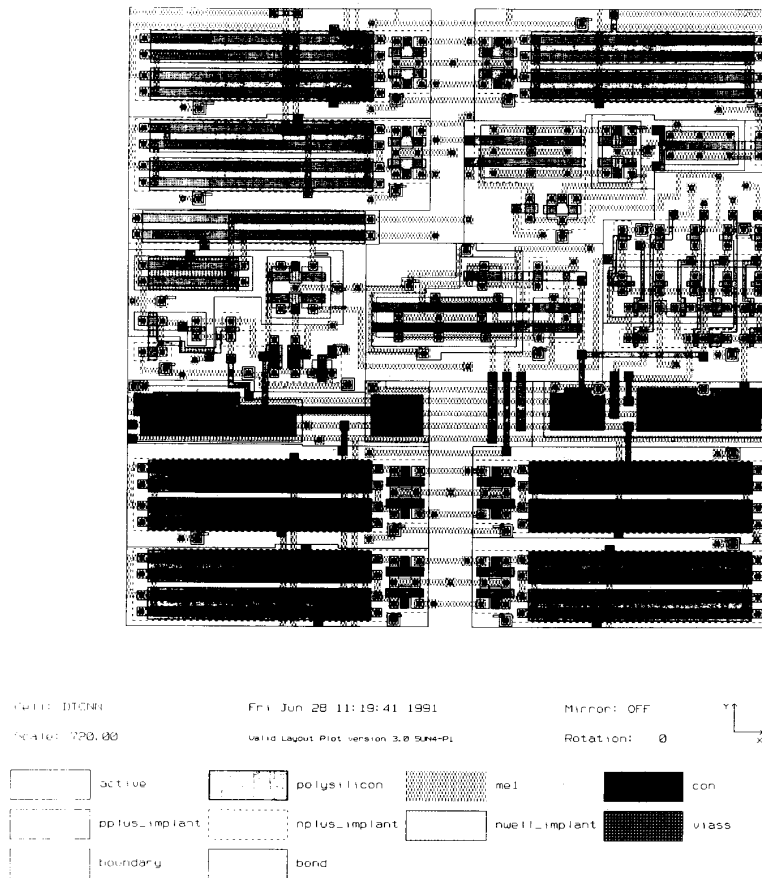
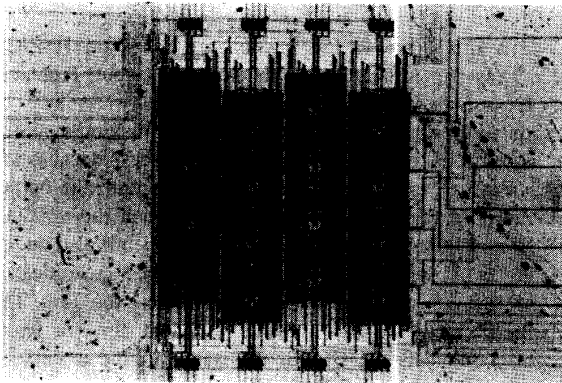Fig. 9. Layout of a single cell without metal 2.

Fig. 10. Photo of the fabricated chip.

all reference voltages are applied by pads and most cells have each two test lines for measuring the state potentials v_{x1}^c and v_{x2}^c. The test lines can be switched off by transmission gates, which is of importance for dynamic testing, as the large input capacitance of the output buffers can disturb the measurements.

Fig. 11 gives the characteristic of a single conductance multiplier for a positive and a negative cell output. The linear range is stretched from -0.7 V to 0.7 V and the zero crossing is accurate. Differences between this and the simulation in Fig. 12 are only a slightly increased gain and a larger saturation voltage.

The linear superposition of the differential currents of two conductance multipliers is shown in Fig. 13. The remaining weights were set to zero. The weight voltages have a triangular waveform with a maximum amplitude of 1 V (division: 0.5 V per cm for both channels). The second trace on the screen shows the difference voltage $v_{x1}^c - v_{x2}^c$. It is almost zero on the left picture, because the two multipliers are operated with opposite sign voltages and, therefore, compensate each other. The right picture shows the result of the superposition when operating both multipliers with identical weight voltages. Notice that the gain in the linear range is approximately twice as large as for a single multiplier.

The characteristic of the comparator is illustrated in Fig. 14. The first channel shows the differential input voltage $v_{x1}^c - v_{x2}^c$ (division: 0.5 V per cm) of the comparator. The second gives the output signal v_y^c behind the analog output buffers. As the output range of the buffers is restricted to ± 1.2 V, the absolute values of the outputs are lower than v_{dd}. The switching of the
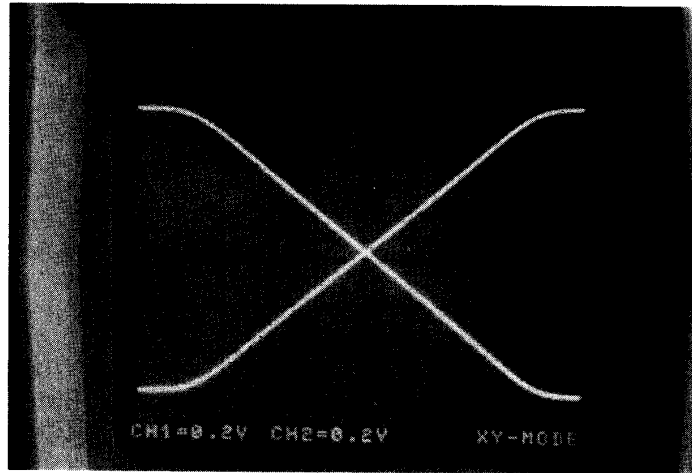
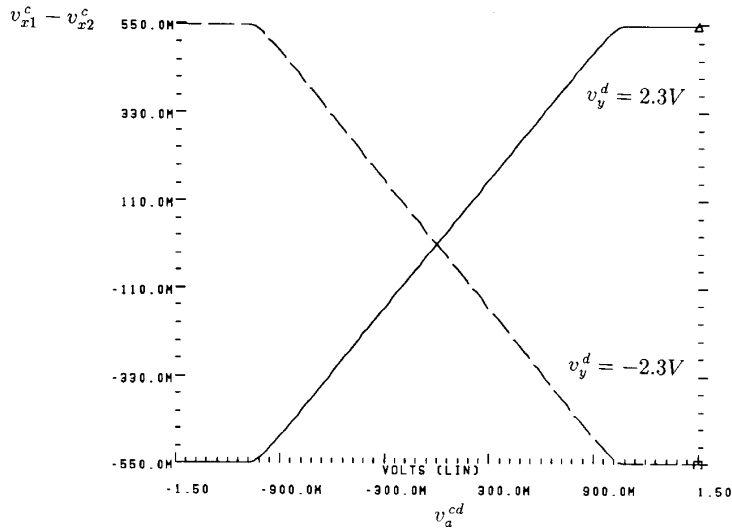Fig. 11.   Measured transfer characteristic of a single conductance multiplier.

Fig. 12.   HSPICE simulation of a single conductance multiplier.

comparator is exactly performed at the zero crossing of the input voltage.

VIII. APPLICATION

The 4 by 4 size of the test chip is sufficient to demonstrate functional behavior of all applications described in [1]. It will be shown exemplarily for the connected component detector, which is a very simple example because of the small number of feedback coefficients. For pattern recognition and data compression, the connected components can be detected if the pattern is applied to the initial state. Black objects are compressed to the size of one pixel and moved to the right boundary, where black and white pixels are alternating dependent on the number of connected components. This can

be done using the template

$$a = \boxed{+1}\,\boxed{+1}\,\boxed{-1} \qquad b = \boxed{0}\,\boxed{0}\,\boxed{0} \qquad i = 0. \qquad (19)$$

The time-dependent development is illustrated in Fig. 15 for the given initial pattern. After four iterations, a convergent output pattern is obtained. A transition from black to white is allowed only if the left neighbor of a black cell is white and the right cell is black. A transition from white to black occurs only if a white cell has a black neighboring cell on the left side and a white neighbor on the right side. In all other combinations, the output state of a cell does not change.

Because the cell bias k^c is zero for all cells in this application, the decision of the comparator depends only on the conductance multipliers. For the weight voltages $v_a^c = v_a^l = 0.3$ V and $v_a^r = -0.3$ V can be chosen, where v_a^c corresponds
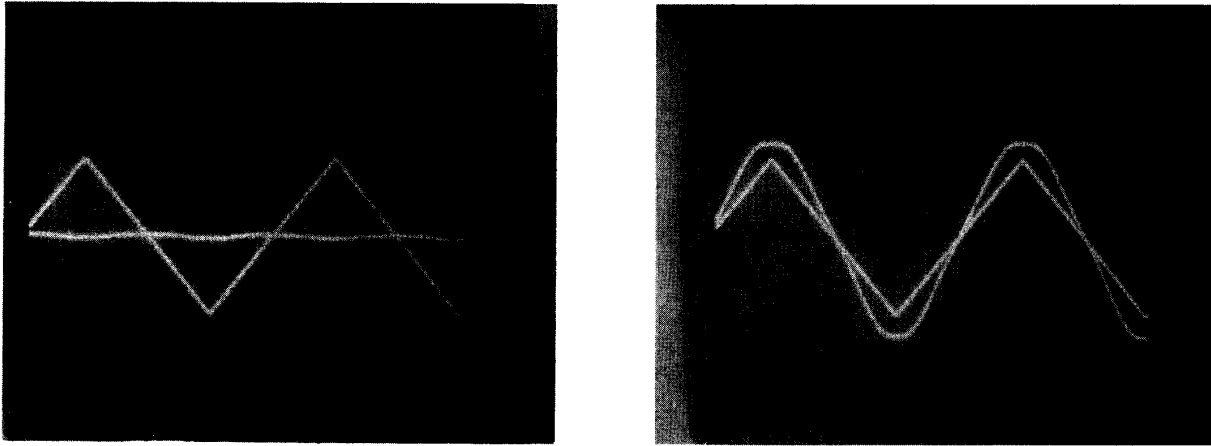
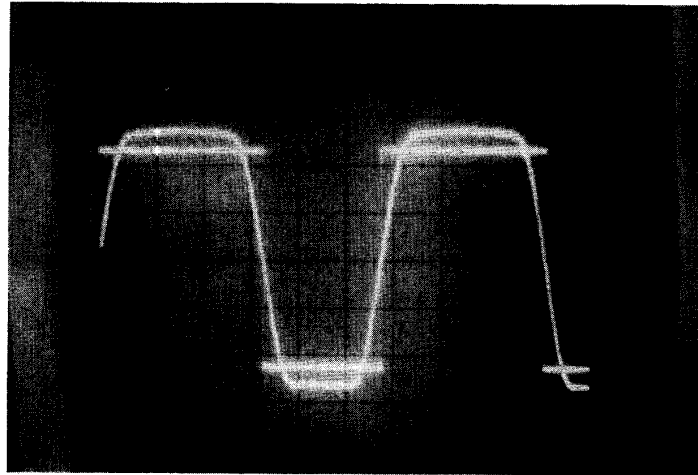Fig. 13.   Linear superposition of two conductance multipliers.

Fig. 14.   Measured comparator voltage.

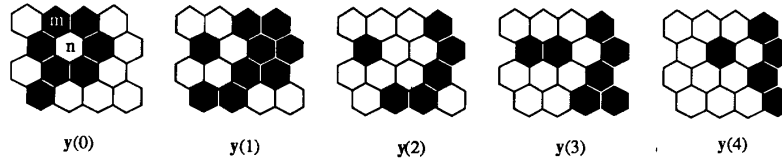y(0)   y(1)   y(2)   y(3)   ·   y(4)

Fig. 15.   Time-dependent development of the output pattern.

to the weight of the self-feedback, v_a^l to the left neighbor, and v_a^r to the right neighbor. The output values of missing neighbors at the boundaries of the cell grid are fixed to v_{ss}, which corresponds to a white ring of surrounding cells. Now, the initial values are loaded by the pattern to be processed. It defines the output $y(0)$. During the high phase of φ_1 in the first iteration, the subsequent output state is computed and stored in C_1, while the initial value is still stored in C_2. The difference voltage of cell n in Fig. 15 is negative, because the differential output currents from the multipliers of the left and right neighbors compensate each other and the negative value

from the self-feedback dominates. In contrast, the sign of the differential state voltage $v_{x1}^m - v_{x2}^m$ of cell m is determined by the left neighbor, as the self-feedback and the right neighbor cancel each other (opposite sign feedback coefficients). Hence, its output voltage is negative in the following iteration. During the high phase of φ_2, the capacitor C_1 discharges C_2. For cell n, this does not lead to a changing output state, but the capacitor voltage of C_2 is now negative for cell m. Notice that the discharging is not influenced from the comparator, because φ_1 is low. After a convergent output pattern is reached, the output values can be read out.

Remark: As there is an odd number of feedback coefficients with identical absolute values and no cell bias, the state variables always have an absolute value greater or equal one. This corresponds to the value of Δ in (4). Transferred to the chip circuit, this means that the minimum absolute difference voltage at the comparator input amounts to approximated 0.25 V and smaller tolerances cannot cause any erroneous output states.

IX. CONCLUSION

An efficient architecture has been proposed for an analog realization of discrete-time cellular neural networks, which takes advantage of

- binary outputs, allowing a simple interconnection of several chips;
- the parameter insensitivity of (4), which makes the network robust against fabrication tolerances if the templates are designed appropriately;
- simple control of the propagation speed over a large range merely by changing the clock rate, which also simplifies the testing of a chip;
- completely differential processing of the analog signals, which automatically suppresses dc disturbances.

The circuit is based on a modified operation mode of conductance multipliers that perform an analog multiplication by ±1, providing differential output currents. This allows the use of sign symmetric gate voltages for enhancement transistors. Operational transconductance amplifiers have been used as linear amplifiers and comparators.

A layout was designed for a 4 by 4 testing chip fixed to a hexagonal grid topology and 1-neighborhood templates. It was realized on a standard CMOS process. The small cell size allows a high density (1200 cells per cm^2). Measurements of the transfer characteristics and the superposition of several weights have shown that the performance of the circuit is sufficient for the DTCNN architecture.

APPENDIX
NONLINEAR CURRENT PARTS OF
CONTROLLED CONDUCTANCES

Assuming the accurate strong inversion model [7] for M1 and M4 (M2 and M3 are turned off), we obtain for the current I_1^c

$$I_1^c = c\left[\left(\acute{v}_y^d - v_{tb} + \gamma\sqrt{\Phi_b - v_b}\right)\left(v_a^{cd} - v_{m1}^c\right) - \frac{1}{2}\right.$$
$$\cdot \left(\left(v_a^{cd}\right)^2 - \left(v_{m1}^c\right)^2\right) -$$
$$\left. \frac{2}{3}\gamma\left(\left(\Phi_b - v_b + v_a^{cd}\right)^{3/2} - \left(\Phi_b - v_b + v_{m1}^c\right)^{3/2}\right)\right]$$

(A1)

with

$$c = \frac{W}{L}\mu C_{ox}$$ (A2)

where γ is the body effect coefficient, Φ_b is the surface potential, and v_b is the bulk potential.

For the current I_2^c, v_a^{cd} is replaced by $-v_a^{cd}$ and v_{m1}^c by v_{m2}^c. Computing the difference current I_{dif}^c, it follows, for $v_{m1}^c = v_{m2}^c$,

$$I_{dif}^c = I_1^c - I_2^c$$
$$= c\left[\left(\acute{v}_y^d - v_{tb} + \gamma\sqrt{\Phi_b - v_b}\right)2v_a^{cd}\right.$$
$$- \frac{2}{3}\gamma\left(\left(\Phi_b - v_b + v_a^{cd}\right)^{3/2}\right.$$
$$\left.\left. - \left(\Phi_b - v_b - v_a^{cd}\right)^{3/2}\right)\right].$$

(A3)

Now, the difference current is developed into a Taylor series expansion for $v_a^{cd} = 0$ V:

$$I_{dif}^c = \left.\frac{dI_{dif}^c}{dv_a^{cd}}\right|_{v_a^{cd}=0} v_a^{cd} + \frac{1}{2!}\left.\frac{d^2 I_{dif}^c}{d(v_a^{cd})^2}\right|_{v_a^{cd}=0}\left(v_a^{cd}\right)^2$$
$$+ \frac{1}{3!}\left.\frac{d^3 I_{dif}^c}{d(v_a^{cd})^3}\right|_{v_a^{cd}=0}\left(v_a^{cd}\right)^3 + \cdots$$

(A4)

The first coefficient, corresponding to the linear part in (13), is the derivative of I_{dif}^c with respect to v_a^{cd}:

$$\left.\frac{dI_{dif}^c}{dv_a^{cd}}\right|_{v_a^{cd}=0} = 2c\left(\acute{v}_y^d - v_{tb}\right).$$ (A5)

For $n > 1$, we obtain

$$\frac{1}{n!}\left.\frac{d^n I_{dif}^c}{d(v_a^{cd})^n}\right|_{v_a^{cd}=0} = \frac{1}{n!}\left(-\frac{2}{3}\right)c\gamma(\Phi_b - v_b)^{3/2-n}$$
$$\cdot (1 - (-1)^n)\prod_{i=0}^{n-1}\left(\frac{3}{2} - i\right).$$ (A6)

The Taylor series has only odd, strictly decreasing coefficients. Inserting typical values leads to

$$I_{dif}^c = 5.6\frac{\mu A}{V}v_a^{cd} + 0.021\frac{\mu A}{V^3}\left(v_a^{cd}\right)^3 + 0.00038\frac{\mu A}{V^5}\left(v_a^{cd}\right)^5$$
$$+ 0.000014\frac{\mu A}{V^7}\left(v_a^{cd}\right)^7 + \cdots.$$

It is easy to see that the nonlinear part can be neglected for $\left|v_a^{cd}\right| < 1$ V.

REFERENCES

[1] H. Harrer and J. A. Nossek, "Discrete-time cellular neural networks," *Int. J. Circuit Theory Appl.*, to be published.
[2] L. O. Chua and L. Yang, "Cellular neural networks: Theory," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 1257–1272, 1988.
[3] J. A. Nossek, G. Seiler, T. Roska, and L. O. Chua, "Cellular neural networks: Theory and circuit design," *Int. J. Circuit Theory Appl.*, to be published.
[4] H. Harrer, J. A. Nossek, and F. Zou, "A learning algorithm for time-discrete cellular neural networks," *Proc. Int. Joint Conf. Neural Networks, IJCNN 91* (Singapore), Nov. 1991, pp. 717–722.
[5] T. Roska *et al.*, "A hardware accelerator board for cellular neural networks: CNN-HAC," in *Proc. First IEEE Int. Workshop Cellular Neural Networks Appl., CNNA-90* (Budapest), 1990, pp. 160–168.
[6] N. Frühauf and E. Lüder, "Realization of CNNs by optical parallel processing with spatial light valves," *Proc. First IEEE Int. Workshop Cellular Neural Networks Appl., CNNA-90* (Budapest), 1990, pp. 281–290.
[7] Y. Tsividis, M. Banu, and J. Khoury, "Continuous-time MOSFET-C filters in VLSI," *IEEE Trans. Circuits Syst.*, vol. CAS-33, pp. 125–140, 1986.
[8] C. Mead, *Analog VLSI and Neural Systems.* Reading, MA: Addison-Wesley, 1989.

[9] E. Vittoz, "Analog VLSI implementation of neural networks," *J. D'Électronique,* pp. 224–250, 1989 (Lausanne).

[10] H. Harrer, J. A. Nossek, G. Seiler, and R. Stelzl, "An analog CMOS compatible convolution circuit for analog neural networks," in *Proc. Micro Neuro-91* (Munich), Oct. 1991, pp. 231–241.

[11] K. Halonen, V. Porra, T. Roska, and L. Chua, "VLSI implementation of a reconfigurable cellular neural network containing local logic (CNNL)," in *Proc. First IEEE Int. Workshop Cellular Neural Networks Appl., CNNA-90* (Budapest), 1990, pp. 206–215.

[12] J. E. Varrientos, J. Ramirez-Angulo, and E. Sanchez-Sinencio, "Cellular neural network implementations: A current mode approach," in *Proc. First IEEE Int. Workshop Cellular Neural Networks Appl., CNNA-90* (Budapest), 1990, pp. 216–225.

[13] K. Slot, "Determination of cellular neural networks parameters for feature detection of two-dimensional images," in *Proc. First IEEE Int. Workshop Cellular Neural Networks Appl., CNNA-90* (Budapest), 1990, pp. 82–91.

Josef A. Nossek (S'72–M'74–SM'81) received the Dipl.-Ing. and Dr. degrees, both in electrical engineering, from the Technical University of Vienna, Austria, in 1974 and 1980, respectively.

In 1974 he joined Siemens AG, Munich, Germany, where he was engaged in the design of passive networks, monolithic filters (analog and digital), electromechanical and microwave filters. Since 1982 he has been head of a group of laboratories designing digital radio systems within the Transmission Systems Department. From 1987 to 1989 he was head of the Radio Systems Design Department and since April 1989 he has been Professor for Circuit Theory and Design at the Technical University of Munich. His research interests include real-time signal processing, neural networks, and dedicated VLSI architectures.

Dr. Nossek has published more than 50 papers in scientific and technical journals and conference proceedings. He holds a number of patents. In 1988 he received the ITG price.

Hubert Harrer (S'90) received the Dipl.-Ing. in electrical engineering from the Technical University of Munich in 1989. Since August 1989 he has been working toward the Dr. degree at the Institute for Network Theory and Circuit Design. His research interests include neural networks and analog VLSI implementation.
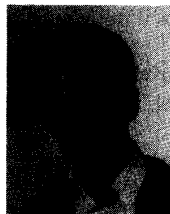
Rudolf Stelzl received the Dipl.-Ing. in electrical engineering from the Technical University of Munich in 1991.

Since June 1991 he has been with the Siemens AG, Munich. His interests include neural networks and analog VLSI implementations.