Abstract—A transient-detecting very large scale integration (VLSI) pixel is described, suitable for use in a visual-processing, depth-recovery algorithm based upon spike timing. A small array of pixels is coupled to an adaptive system, based upon spike timing dependent plasticity (STDP), that aims to reduce the effect of VLSI process variations on the algorithm’s performance. Results from 0.35 μm CMOS temporal differentiating pixels and STDP circuits show that the system is capable of adapting to substantially reduce the effects of process variations without interrupting the algorithm’s natural processes. The concept is generic to all spike timing driven processing algorithms in a VLSI.

Index Terms—Active pixel, CMOS integrated circuits, focal-plane sensor, neuromorphic analogue very large scale integration (aVLSI), spike timing dependent plasticity (STDP), temporal processing, transistor mismatch.

I. INTRODUCTION

P

ELGROM’s law states that device mismatch is proportional to $1/\sqrt{WL}$ [1], indicating that the problems associated with mismatch will only increase as device sizes approach submicron geometries. When subthreshold currents are required, for power saving or time constant considerations, mismatches have an even greater effect [2]. Some compensation mechanisms have been proposed, which allow the fabrication of, for example, current mirrors, dividers and comparators [3]–[5] with improved matching characteristics while still using minimum size transistors. We introduce an alternative approach that uses the change in the temporal characteristics introduced by the mismatch to effect a correction.

In most systems, a change in transistor parameters has an effect on the system time constants and mismatch affects the temporal characteristics of the circuit. spike timing dependent plasticity (STDP) is a class of neural algorithms driven not by spike-rate correlations, but by individual interspike timings. This class of algorithms has been shown to recover different and potentially richer information from a spike train when compared to rate and population coding methods [6] and has been shown capable of adapting synaptic delays [7]–[9] to improve network performance. STDP is, however, a relatively simple algorithm which has been successfully implemented in analogue very large scale integration (aVLSI) [10]–[17]. If a suitable spiking neural system could be implemented in aVLSI the ability to adapt for delay error could be used to correct timing errors introduced by process mismatch.

Our exemplar system was developed by Wörgötter et al. [18]. It is a spike timing driven algorithm for visual scene analysis whose temporal characteristics will be degraded if implemented in analogue focal-plane VLSI. Although the algorithm is intrinsically interesting, we use it primarily as a background to test the ability of STDP to correct for mismatch. Brief details of the algorithm and its small-scale implementation in aVLSI circuitry are presented in Section II. The transient-detecting pixel circuitry is described in Section III. Section IV presents the STDP circuit and Section V brings pixels and STDP adaptation together. All circuitry is implemented using the AMS 0.35 μm CSI CMOS process.

II. VISION ALGORITHM BASED ON SPIKING NEURONS

This section describes a parallel noise-robust algorithm [18] operating on the neurons of an artificial retina to recover depth information from radial flow fields in a visual scene. This algorithm finds its motivation in the behavior of animals. In different species, varying strategies are observed in order to reduce the optical flow to, if possible, one-dimensional motion. In particular, for airborne animals such as flies and birds, this actually leads to the tendency to fly in straight lines. Ideally, it means that only forward motion exists and that optical flow is reduced to its radial components. Under these circumstances, points in the image plane move radially outwards from the focus of expansion with a velocity proportional to radial distance from the focus of expansion but inversely proportional to their depth in the image. Hence, the depth can be inferred from this image expansion.

The retina in this algorithm, shown in Fig. 1, consists of radially arranged neurons connected only to their nearest neighbors in both directions on the same radius. It functions as a visual system of a moving robot. The robot is constrained to move along the optical axis of the retina. As the robot moves, objects in the field of view are projected on the retinal plane, and neurons are excited as soon as a sufficiently strong brightness transient, i.e., a moving edge, occurs. The time difference between two subsequently activated neurons (e.g., neurons 1 and 2) is used to compute the depth information. The structure of the network is such that all computations remain local, i.e., neurons connect to nearest neighbors, which facilitates parallelization. However, the local nature of all calculations makes the algorithm sensitive to noise, for example that caused by camera jitter. A prediction mechanism is introduced to reduce this sensitivity.
The calculated depth is used to predict when the last-excited neuron (e.g., neuron 3) will fire. Events are accepted as valid if this prediction corresponds with an actual event.

We have based our pixel upon the light-transducing elements described in [19] and [20]. The technique for device mismatch correction is based upon a form of asymmetric Hebbian learning, a neural algorithm implementing STDP.

The system block diagram is shown in Fig. 2. The spike firing circuit comprises three leaky integrate-and-fire (LIF) neurons [13]. These LIF neurons integrate output currents received from transient-detecting pixels (E), cancel the dark current, and fire when the accumulated signal reaches a threshold. The spikes from neurons 1 and 2 are passed to the prediction network which generates another spike, spike 3p (i.e., “predicted,” not actual), at a time determined by the firing times of the input spikes. The timing of spike 3p can then be compared to that of spike 3 by the spike confirmation block. In the absence of process variations, Spikes 3 and 3p will be coincident. Spike 3 is, therefore, accepted as genuine only if it arrives within the time window of spike 3p that represents the acceptable tolerance on the prediction. The adaptive STDP network is also shown. It aims to “adapt out” the process-induced mismatch between the actual firing time \( t(3) \) and its predicted time \( t(3p) \).

### III. Transient-Detecting Pixel

In this section, we present a compact integrated circuit suitable for use in a focal-plane imaging processing pixel. Its main function is the enhancement of local, temporal brightness transients.

#### A. Circuit Description

The pixel circuit is shown in Fig. 3. It combines an adaptive photoreceptor with a rectifying differentiating element and consists of a photodiode \( D \) in series with a transistor \( M_{B} \), in source-follower configuration. A negative feedback loop is created between the source and the gate of \( M_{B} \). The feedback loop consists of a high-gain inverting amplifier in common-source configuration \((M_{ON}, M_{OFF}, M_{AB}, C = 500 \ \text{fF})\) and a capacitive gain stage \((C_1 = 1 \ \text{pF}, C_2 = 50 \ \text{fF})\). We employed Nwell-over-Psubstrate photodiodes as opposed to Ndiff-over-Psubstrate photodiodes used in [19] and [20]. This increases the pixel circuit response to illumination changes because the parasitic capacitance is reduced by a factor of about 10 [21]. The capacitors occupy a large part of the pixel area. Hence, for a \( 20 \ \mu m \times 20 \ \mu m \) photodiode in our pixel circuit, the fill factor is about 10%.

The class-AB differentiator brings two benefits. First, it eliminates the “dead” band inherent in its class-B counterpart [20], which can slow down the response when a voltage signal is present within the band. Second, it provides some current to counterbalance leakage currents at the differentiator node, reducing asymmetry between positive and negative transients.

An additional capacitive gain stage in the feedback loop gives enhancement to the response [19]. The voltage variations on the differentiator node \( V_{\text{diff}} \) are amplified with respect to the variations of \( V_{B} \), by the capacitive divider ratio \( A_C = (C_1 + C_2)/C_2 = 21 \) and so are the transient currents \( I_{ON} \) and \( I_{OFF} \).

The resistive element, constructed from two oppositely directed diodes in series, ensures that \( V_{\text{diff}} \) eventually adapts to a dc value close to \( V_{B} \). In ideal circumstances, the current is limited to a low value by the reverse-biased diode, such that the element exhibits a symmetric, saturating, sigmoidal current-voltage characteristic.

A full circuit analysis of the steady state and transient responses of the pixel is presented in the Appendix.

#### B. Experimental Setup

The power supply voltage \( V_{dd} \) was set to 3.3 V and the bias voltages \( V_{13}, V_{12}, \) and \( V_{13} \) were fixed for all measurements such that \( M_p \) and \( M_n \) were operated slightly above threshold. As the pixel provides output in the form of current \((I_{ON} + I_{OFF})\),
Fig. 3. Circuit diagram of the transient-detecting pixel. Positive or ON transients and negative or OFF transients, although they can be assigned to separate terminals, are combined on a single output node and fed to a neuron. The circuit has a class-AB differentiator and a capacitive gain stage, which eliminates the "dead" band that can slow down the response and amplifies changes of the sustained voltage, respectively.

A 1 MΩ resistor was connected between the output node and ground, enabling voltage readout for an oscilloscope. As a result, the output time constant was massively increased and does not reflect the natural on-chip response of the pixel.

The entire chip was illuminated through a diffuser with a radiance-modulated light-emitting diode (LED), operating around an emission wavelength of 590 nm, while shielding the chip from ambient light (Fig. 4). A forward voltage was applied to the LED with a series resistor, such that the current and, thus, the radiance of the LED were approximately linear with the applied voltage in a certain range. A radiance-voltage calibration was performed using a digital light meter.

C. Test Results

For steady-state measurements, the irradiance range was expanded by inserting neutral density (ND) filters with attenuation factors of 10, 100, and 1000 between the diffuser and the chip. The voltage $V_{fb}$ at the feedback node of the transient-detecting pixel shows the predicted logarithmic behavior over a large irradiance range, as shown in Fig. 5. The dark line amidst the data points shows a linear-logarithmic relationship between $V_{fb}$ and illuminance.

For transient measurements, the LED was modulated using a square wave signal from a waveform generator which was buffered and low-pass filtered. The low-pass filtering was necessary to reduce artifacts due to the discretization of the waveform generator’s output signal. The time constant of the low-pass filter was set to 1 ms. The circuit was allowed to reach a steady state at a default illuminance of 1 lux. Saturation and adaptation of the pixel can be observed in the $V_{amp}$ and $V_{diff}$ traces for a contrast step of 100 lux, as shown in Fig. 6. Since the transient currents $I_{ON}$ and $I_{OFF}$ of the pixel circuit are very small (a few hundred nA), a large irradiance step (100:1) is necessary in order to display a reasonably large $V_{out}$ waveform (a few hundred mV) on the oscilloscope. However, when testing the combined pixel/STDP circuit as will be described in Section V, we used a smaller irradiance step (15:1).

IV. SPIKE TIMING DRIVEN CIRCUITS

In these concept-proving experiments, the output from three transient-detecting pixels drives three LIF neurons. As an edge passes a pixel the corresponding neuron fires a spike. These spikes are used directly by the STDP adaptive circuit. The two main components making up the spike timing driven circuit are the predictive and the adaptive circuits. The adaptive circuit is designed to minimize the effects of process mismatch in the prediction.
The results described in this section are from a replicated copy of the circuitry with voltage controlled inputs acting as artificial pixels. This allowed initial testing of the spiking circuitry in isolation from the pixels.

A. Prediction Circuitry

Wörgötter et al. [18] showed that knowledge of a pixel array layout can be used to infer when an edge will pass a particular pixel from the time of previous edges. Three pixels suffice to demonstrate the success of our approach and their layout was chosen such that the time between the first and second spikes is equal to the time between the second and third. Therefore the predicted time of arrival for the third spike, spike 3p, is \( t(3p) = t(2) + [t(2) - t(1)] \).

The prediction spike is generated by comparing a voltage across a capacitor to a reference voltage as shown in Fig. 7. The charging and discharging currents connected to \( C1 \) are designed to be identical so that the time to discharge the capacitor is equal to the charging time. The sequence of events is as follows:

- \( t(1) \): Spike 1 is fired and P6 and N6 act as a switch which sets the initial value of \( V_{ramp} \) to \( V_{rampth} \).
- \( t(1)-t(2) \): P5 is on and C1 is charged through P3 and P4 which supplies a mirrored version of the current set by \( V_h \).
- \( t(2) \): P5 turns off while N5 becomes active.
- \( t(2)-t(3p) \): C1 is discharged through N3 and N4. If the current source and sink are matched the time taken to discharge C1 should match the charging time.
- \( t(3p) \): \( V_{ramp} \) is compared to \( V_{rampth} \) using a differential pair. When \( V_{ramp} \) crosses it, neuron 3p is “fired.” This ends the discharge period.

The control signals \( t1-2 \) and \( t2-3p \) are generated by spike triggered SR Latches, e.g., set connected to spike 2 and reset connected to spike 1 results in \( t1 \).

The accuracy of the prediction depends upon how well the charging and discharging currents are matched. We are primarily interested in gain mismatch and therefore transistors P1, P4, N2, and N4 are used to minimize the current change caused by the early effect. Fig. 8(a) shows the result if the currents are not matched. First, the charging current is smaller than the discharging current and then the situation is reversed. The currents will not be perfectly matched unless post-fabrication trimming is performed on the circuit. This is both costly and time consuming. Instead, we use the relationship between the actual and predicted third spike to correct for any mismatch.

It should also be noticed that the addition of a second power supply pin, \( V_{dd_L\_mirr} \), to the circuit in Fig. 7 allows mismatch to be forced on to the circuit. This facilitates testing of the adaptive network under a wide range of induced prediction errors, which would not be possible with the limited number of chips received. Under most test conditions the two pins are shorted together.

In future chips with larger pixel arrays the prediction for different layout configurations can be achieved by ratioing the charging and discharging currents to alter the time between spikes 2 and 3p.

B. Adaptive Circuitry

To correct the prediction error a method must be found either to adjust one, or both, of the currents until they are matched or to compensate for the effect of the difference in the currents. Changing the voltage across C1 at the time of the switching between charging and discharging, \( t(2) \), will alter the discharge time. The voltage can be increased or decreased, which has a direct effect on the discharge time and therefore compensates for the current difference. This method was chosen and examples are shown in Fig. 8(b).

The change in the voltage across C1 is achieved using the excitatory/inhibitory synapse described in [22] and is shown in Fig. 9 along with the output waveform. The circuit is designed such that if \( V_w \) is greater than \( V_{synth} \) the synapse is excitatory. Should \( V_w \) be smaller the synapse is inhibitory. \( I_{inj/\text{out}} \) is only connected to \( V_{ramp} \) at \( t(2) \) and the value of \( V_w \) will determine the amount and the direction that the peak voltage will change. The width of spike 2 and therefore the time \( I_{inj/\text{out}} \) is connected to \( V_{ramp} \) will be constant for a given set of environmental conditions. Should change \( V_w \) should change accordingly.

Transistors N1 and N6 are both small W/L transistors cascaded by the wide transistors N2 and N7. This maintains N1 and N6 in the linear region over a wide range as described in [12] and [13]. N1 is biased by \( V_{synth} \) into the linear region. The value of \( V_{cas} \) determines the minimum value of \( V_w \) that N6 enters the linear region and will set the maximum amount of current that can be injected/removed from C1. It also determines the gradient of the \( I-V \) curve which has a direct effect on the amount the output current changes with a step change in \( V_w \). The smaller the range of required current the more accurate an output current can be selected.
Fig. 8. Pre- and postadaptation waveforms. These oscilloscope traces show the spike 3, spike 3p and Vramp signals during testing. At $t(1)$ Vramp is set to $V_{ramp_{th}}$ and charging begins. At $t(2)$, the current direction is switched to start the discharge. Spike 3p is generated when $V_{ramp}$ returns to $V_{ramp_{th}}$. The timescale shown is 5 ms/div. (a) Preadaptation and the error in spike 3p is clear. (b) The signals postadaptation where the error has been reduced. The alteration to $V_{ramp}$ can also be seen.

Fig. 9. Excitatory/inhibitory synapse circuit and $I-V$ graph. The synapse is excitatory if $V_{w}$ is greater than $V_{synth}$ and inhibitory if $V_{synth}$ is greater. (a) Shows the circuit configuration. (b) Is the synapse output current with $V_{synth} = 1.8$ V. $I_{in/out}$ is zero when $V_{w} = V_{synth}$. 

(a) (b)
Most variants of STDP maximize the synaptic weight-change when spikes are near-simultaneous. In contrast, our system is designed to apply weight changes which result in the spikes becoming more coincident.

The setting of $V_{W}$ must be controlled in order that the appropriate value is set to give coincidence of spikes 3 and 3p. The circuit shown in Fig. 10 provides this control by implementing the weight-change curve shown in Fig. 11.

The circuit in Fig. 10 has two distinct parts. MOSFETs N1-3 and P1-3 control reductions in $V_{W}$ (depression) through N7-9, while increases (potentiation) are achieved through transistors N4-6 and P4-9. As the potentiation and depression circuits are mirror-images, with transistor polarities changed, we will present a detailed description of only the depression mechanism.

When spike 3 occurs C1 is discharged through N1 and then charges slowly to $V_{dd}$ through P1. This results in a pulse at the gate of N8, whose width is determined by $V_{b2}$. This pulse defines the window within which weight change occurs. At $t(3) - t(3p)$ until the voltage reaches $V_{n}$. This defines the maximum weight change. If spike 3p occurs within the time window, $C_{w}$ is discharged through N7, N8, and N9 by an amount related to the voltage across C2. When the weight voltage is decreased, the peak of $V_{TDP}$ decreases. This reduces the discharge time and spike 3p will occur at a time $t(3p)$ that is closer to $t(3)$.

Fig. 10. Weight adaptation circuit. $V_{n}$ is increased if spike 3p occurs before spike 3 and decreased if neuron 3 fires first. Weight change only occurs if spikes 3 and 3p happen within a specified time window. The weight change ($\Delta W$) is "capped" by $V_{n}$ and $V_{p}$ to prevent $V_{n}$ moving from one supply rail to the other. Three circuit blocks are highlighted. A: The circuit setting the time window. B: The circuit setting the amount of weight change. C: Combining the two together to create the negative part of the $\Delta W$ graph.

Fig. 11. Weight adaptation algorithm. The weight change algorithm implemented by the circuitry shown in Fig. 10. The change in $\Delta W$ was measured in simulation. The inset picture is the range $-500 \mu s - 500 \mu s$. The null point around $t(3p) - t(3)$ occurs when charge injection due to switching dominates.

Fig. 12. Weight change. The difference in spike timing alters the synaptic weight and the prediction becomes more accurate. The timescale shown is 5 ms/div.
The weight change is “capped” by \( V_H \) and \( V_P \) to prevent \( V_W \) moving from one supply rail to the other. The window width is determined empirically from a priori knowledge as to when two spikes may be regarded as unrelated events.

When choosing the bias voltages \( Vb1, Vb3, V_H, \) and \( V_P \), there is a trade-off to be made between speed of convergence and the width of the null point, determined by \( Vb1 \) and \( Vb3 \), around \( t(3p) - t(3) = 0 \) which determines how close spikes must be before the time difference has a negligible effect on the weight. This results in a tolerance associated with the prediction. The maximum weight change set by \( V_H \) and \( V_P \) has an effect on how many spike pairs are required to correct a particular prediction.

If the adaptation windows, set by \( Vb2 \) and \( Vb4 \), are too small, large mismatch errors will be judged to be uncorrelated spikes and no adaptation will occur. This leads to a further trade-off between the circuit’s ability to adapt and the maximum allowable density of edges in the scene. A larger window will allow a greater range of errors to be corrected but, for the circuitry to function, no two edges can occur within that interval.

When spike 3p is a good prediction of spike 3 no further weight change occurs. It is assumed that correct edges will vastly outnumber incorrect edges and that if one does fall within the adaptation time window it will only cause a small deviation of weight which will be corrected by subsequent correct spikes. The adaptation network can therefore be left active after this “self-calibration” and will continue to compensate for any circuit drift caused by changes in environmental conditions or charge leakage from the capacitor. The latter is particularly important as many compensation techniques suffer from problems in the permanent storage of calibration values.

The weight change is defined as \( t(3p) - t(3) \). Each graph has the preadaptation results above the postadaptation ones. The results were obtained with \( V_{dd_{nirr}} \) varying between 3.28 and 3.31 V. Results were when (a) \( t(2) - t(1) = 1 \) ms and (b) \( t(2) - t(1) = 5 \) ms.

C. Test Results

When the circuitry was fabricated, using the AMS 0.35 \( \mu \)m CSI process, a copy of the predictive and adaptive circuit was laid out with voltage controlled inputs to allow testing in isolation from the pixels. This was also the circuit that was tested with varying \( V_{dd_{nirr}} \), see Fig. 7.

The oscilloscope traces in Fig. 8 are taken from one of the chips when the input neurons were fired with a 5 ms interneural delay. It can be seen that the STDP adaptation improves the prediction significantly. Fig. 12 shows part of the adaptive process. As the weight increases the resultant improvement in the prediction can be seen.

The circuit was further tested by activating the three input neurons with 1 and 5 ms interneural delays and measuring the prediction error, \( t(3p) - t(3) \) before and after adaptation. This was done three times on eight different chips with \( V_{dd_{nirr}} \) varying between 3.28 and 3.31 V. Fig. 13 shows the pre-
Fig. 15. Pre- and postadaptation results. These results show the spike 1, spike 2, spike 3 and spike 3p measured during experiments when \( t(2) - t(1) = 5 \text{ ms} \). The timescale is 5 ms/div. (a) Is preadaptation and the error in spike 3p is clear. (b) Shows the signals postadaptation where the error has been reduced.

postadaptation error when the interneural delay is (a) 1 ms and (b) 5 ms. These results compare favorably with the simulation results reported in [23].

V. SMALL ADAPTIVE PIXEL ARRAY

The combined pixel/STDP circuit was tested by stimulating it with moving light patterns generated by an array of flashing LEDs (10 lux) as shown in Fig. 14. The LED test patterns were set in such a way that transient-detecting pixels were illuminated chronologically. The optical part of the imaging system was a surveillance camera lens with a focal length of 6 mm and an \( f \)-number of 1.2. The LED array was placed at a distance of 15 cm from the lens. This experiments were repeated 10 times over five chips and conducted under the ambient lighting condition of fluorescent office lighting (150 lux).

Fig. 15 shows an example of spike waveforms observed in the experiments before and after STDP adaptation. It is important to note that each pixel is able to detect the LED test pattern (high-frequency signal) while rejecting the ambient light (low-frequency signal).

After the generation of spikes had been verified the prediction delay was measured before and after adaptation. For this test \( V_{dd,norr} \) was kept at 3.3 V and the interneural delay times used were 1 and 5 ms. Fig. 16 shows the resultant spread of the errors. It can be seen that adaptation improves the prediction substantially. When these results are compared with those taken from the voltage controlled circuit when \( V_{dd,norr} \) was 3.3 V, Fig. 17, it can be seen that the pixel does not affect the spread of error. The pixel driven, pre-adaptation results have a less smooth distribution, as these results were taken from 5 chips rather than the 8 for the voltage driven circuit. The mean prediction error and the standard deviation was measured for each distribution and can be found in Table I.

The similarity of the numbers in Table I between the pixel and voltage driven circuitry confirms that pixel activation does not introduce additional error. Small differences should be expected as the results are being collected from a small sample of die.

The error when the interneural time delay is 5 ms will always be greater than the 1 ms error, because the prediction error is related to the mismatch in current and charging time. The mean values shown in Table I can be seen to have a positive bias. Transistors N1 and N3 in Fig. 7 will be affected by the early effect and this would cause a small positive bias but simulation results indicate it should not be as big as the value seen. We believe the majority of the bias is caused by connecting the drains of the current mirror, \( V_{dd} \) and \( V_{dd,norr} \), externally to the chip rather than with a direct metal connection. A voltage drop of a few mVs on the \( V_{dd} \) line would explain this systematic offset. This chip is therefore not suitable for determining the level of mismatch in this process but it does illustrate the effectiveness of an adaptive, STDP approach for correcting mismatch. The adaptation does move the mean back toward zero and the standard deviation is greatly reduced.

Two different types of error are included in the prediction: an offset error and a gain error. The offset error is introduced by the noninfinite gain of the comparator in Fig. 7 and the early effect, whereas the gain error is due to mismatch in the \( gm \) of the current mirror transistors. Steps were taken to minimize offset errors and Monte Carlo simulations indicated that differences in gain would be the dominant source of prediction error.

This particular approach for prediction correction will correct for both types of error but only for one time interval at a time as it essentially applies an offset correction. Therefore if a new time delay is introduced the adaptive circuit must move to a new weight value. It is the long term goal of this work to implement the correction by adjusting the current sources in the current sink/source circuit (Fig. 7) so that the prediction can be correct over a range of time intervals. It is also worth noting that should a systematic error occur during spike generation at the pixels the prediction will adapt to accept that error. A random error will have no effect.

The STDP adaptation network has reduced the effect of process mismatch by “pulling together” spikes which are effectively coincident in real time, but have been spread on silicon by process imperfections.

VI. CONCLUSION

We have shown that STDP can be used to correct for process mismatch in an analogue VLSI system, using the signals that are naturally present in a spike-driven processing algorithm and without the need for an explicit calibration. We have demonstrated the success of the technique in the context of a neuro-morphic, spike-time driven vision-processing algorithm, but the results have potential implications for all spike-timing processes and algorithms on silicon.
Fig. 16. Distribution of prediction errors from pixel activated circuit. The error measurements were taken with $V_{dd, nirr} = 3.3$ V. The current integrated on the LIF neurons came from the transient-detecting pixels. Results (a) were from a simulation when $t(2) - t(1) = 1$ ms and (b) $t(2) - t(1) = 5$ ms.

Fig. 17. Distribution of prediction errors from voltage controlled circuit. These results were recorded when $V_{dd, nirr} = 3.3$ V. Results (a) were from a simulation when $t(2) - t(1) = 1$ ms and (b) $t(2) - t(1) = 5$ ms.
While the STDP circuits are not small, they will shrink as process geometries shrink and the problem of mismatch becomes ever more bothersome. Furthermore, pixel-array applications are particularly area-sensitive and the technique may have immediate applications in tasks where the spike-processing circuits are themselves relatively large and the “overhead” of the STDP augmentation less of a concern.

This paper’s primary conclusion is, however, that STDP can be used to adapt system characteristics, in this case temporal characteristics “on the fly” to improve system performance. We have demonstrated that capability in 0.35 \( \mu m \) CMOS aVLSI.

### Appendix

#### Pixel Circuit Analysis

A. Adapted Steady State

The transistor \( M_{fb} \) is operated in saturation and, for typical irradiance, in weak inversion. When the circuit is fully adapted to background illuminance \( V_{fb} \approx V_{diff} \) and neglecting early effects, we obtain

\[
V_{fb} = \kappa_{fb}^{-1} \left( V_{ph} + V_T \log \left( \frac{I_{ph}}{I_{fb0}} \right) \right)
\]

where \( I_{fb0} \) is the current-scaling parameter and \( \kappa_{fb} \) the subthreshold slope factor of \( M_{fb} \) and \( V_T \) denotes the thermal voltage \( kT/q \), given by the absolute temperature \( T \), the Boltzmann constant \( k \) and the elementary charge \( q \). The voltage \( V_{ph} \) is set by the bias current through the inverting amplifier. Assuming a bias voltage \( V_t \) that puts \( M_n \) and \( M_p \) into weak inversion and again neglecting Early effects we obtain

\[
V_{ph} = \kappa_{n}^{-1} \left( \kappa_{n}(V_{dk} - V_t) + V_T \log \left( \frac{I_{ph0}}{I_{ph0}} \right) \right)
\]

where \( I_{ph0} \) and \( I_{ph} \) are the current-scaling parameters of \( M_n \) and \( M_p \), respectively, \( \kappa_n \) and \( \kappa_p \) are the corresponding subthreshold slope factors and \( V_{dk} \) is the potential of the positive power rail. In this approximation, \( V_{ph} \) is independent of \( I_{ph} \). The actual dependence is due to the Early effects of \( M_n \) and \( M_p \), i.e., to the limited gain of the inverting amplifier.

Neglecting the leakage currents of \( M_{ON} \) and \( M_{OFF} \) and assuming weak inversion and saturation for these transistors, we can compute the steady-state ON and OFF currents as

\[
I_{ON} = I_{ON0}e^{-\frac{\kappa_{ON}V_{amp}+V_{AB}-V_{fb0}}{V_T}}
\]

where \( I_{ON0} \) and \( I_{OFF0} \) are the current-scaling parameters of \( M_{ON} \) and \( M_{OFF} \), respectively, and \( \kappa_{ON} \) and \( \kappa_{OFF} \) are the corresponding subthreshold slope factors. The bias voltage \( V_{AB} \) for the class-AB differentiator can be expressed as

\[
V_{AB} = \kappa_{AB}^{-1}V_T \log \left( \frac{I_{AB}}{I_{AD0}} \right)
\]

where \( I_{AD0} \) is the current-scaling parameter and \( \kappa_{AB} \) the subthreshold slope factor of \( M_{AB} \).

Leakage currents in the transient pathway affect, primarily, the diodes to the substrate, i.e., the source and drain diodes of \( M_{ON} \) and the well-to-substrate diode of \( M_{OFF} \). The source diode of \( M_{OFF} \) does not contribute a leakage current because it is shorted and the leakage current of the drain diode of \( M_{OFF} \) is small. Furthermore, the parasitic currents at the drain diodes of \( M_{ON} \) and \( M_{OFF} \) do not influence the channel current, but add to the current drawn from subsequent devices. The source leakage current of \( M_{ON} \) and the well leakage current of \( M_{OFF} \), however, result in a parasitic current \( I_{par} \) from the differentiator node into the substrate, which has to be balanced by an increased current through \( M_{ON} \) and, therefore, an increased gate voltage of \( V_{fb} \) with respect to \( V_t \). According to Kirchoff’s current law, \( I_{par} = I_{ON} + I_{OFF} \). Since \( I_{par} \) is typically large compared to \( I_{ON} \) and \( I_{OFF} \), it dominates those currents \( (I_{par} \approx I_{ON} \gg I_{OFF}) \). Therefore

\[
V_{amp} = \kappa_{n}^{-1} \left( V_{fb} + V_T \log \left( \frac{I_{par}}{I_{ON0}} \right) \right) - \kappa_{AB}^{-1}V_T \log \left( \frac{I_{AB}}{I_{AD0}} \right)
\]

If \( M_{ON} \) and \( M_{OFF} \) are implemented in the vicinity of the photodiode on the same silicon substrate, photo-induced minority carriers also contribute to the leakage and dominate it for large irradiance, such that it becomes roughly proportional to the photocurrent, i.e., \( I_{par} = \delta I_{ph} \) where \( \delta \) denotes the ratio of the electrons collected by the photodiode. The dependence of \( I_{par} \) on \( I_{ph} \) can be determined from the slope of the \( V_{amp} \) versus \( V_{ph} \) characteristic.

B. Transient

In the following, we will make a transient analysis of the circuit without considering parasitic capacitances and the effects of leakage currents in the transistors. The analysis assumes that the circuit variables have reached an equilibrium state before a transient change in the photocurrent is applied, and that no adaptation occurs in the considered time window.

The absolute value of the gain of the inverting amplifier is determined by the Early effects of \( M_n \) and \( M_p \), and is given by

\[
A = -\frac{\partial V_{amp}}{\partial V_{ph}} = \kappa_n V_T \left( \frac{1}{nE + V_T} \right)^{-1}
\]
where \( V_{OE} \) and \( V_{PE} \) are the Early voltages of \( M_{th} \) and \( M_{th'} \), respectively. Differentiating (1) and substituting \( V_{amp} = -A_{amp} V_{ph} \) yields
\[
\frac{I_{ph}}{I_{ph}} = \frac{V_{th}}{V_T} + A_{amp} \frac{V_{amp}}{V_T} \quad (8)
\]
Differentiating (3) and (4), respectively, gives
\[
\frac{I_{ON}}{I_{OFF}} = \frac{V_{amp} + V_{AB}}{V_T} - \frac{V_{th}}{V_T} \quad (9)
\]
\[
\frac{I_{OFF}}{I_{ON}} = - \frac{V_{amp}}{V_T} + \frac{V_{th}}{V_T} \quad (10)
\]
If leakage currents in the differentiator stage are neglected, the capacitor current is given by
\[
I_{ON} = I_{OFF} = CV_{th} \quad (11)
\]
It follows from (1) that:
\[
I_{ON} = I_{OFF} = C \left( \frac{V_{ph}}{V_T} + V_T \frac{d}{dt} \ln \left( \frac{I_{ph}(t)}{I_{th}} \right) \right)
\]
\[
= C \left( \frac{V_{th}}{V_T} + \frac{I_{ph}}{I_{th}} \right) \quad (12)
\]
In the closed-loop domain, where the feedback loop is activated, the \( V_{ph} \) term can be neglected if the loop gain is much larger than unity (\( V_{th} \gg V_{ph} \)). The difference in the transient currents is then proportional to the temporal derivative of the logarithm of the photocurrent, i.e., to the relative transient of the photocurrent.

ACKNOWLEDGMENT

The authors would like to thank P. Fleury for his assistance during chip layout and with tapeout. They would also like to thank A. Bofill, H. Chen, and S. Collins for their helpful discussions. The technical support of Austrianmicrosystems and Europractice is gratefully acknowledged.

REFERENCES


Katherine Cameron (S'04) was born in Edinburgh, U.K., in 1979. She received the M.Eng. (Hons.) degree in electronics from the University of Edinburgh, Edinburgh, U.K., in 2002. She is currently working toward the Ph.D. degree in neuromorphic engineering at the University of Edinburgh. Her current research interests include bio-inspired engineering solutions to analogue computation imperfections, mixed-signal VLSI design, and neural computation.
Vasin Boonsohak (S’03) was born in Bangkok, Thailand, in 1975. He received the B.Eng. (Hons.) degree in electrical engineering from Mahanakorn University of Technology, Bangkok, Thailand, in 1996 and the M.Sc. degree in analogue and digital integrated circuit design from Imperial College, London, U.K., in 1998. He is currently working toward the Ph.D. degree in electrical engineering at the University of Edinburgh, Edinburgh, U.K. Since 1996, he has been with the Department of Electronics Engineering, Mahanakorn University of Technology, Bangkok, Thailand. His current research interests include mixed-mode VLSI systems for image processing, image sensors and asynchronous digital interfaces for interchip connectivity.

Alan Murray (SM’94) was born in Edinburgh, U.K., in 1953. He received the B.Sc. degree (Hons.) in physics and the Ph.D. degree in solid state physics from the University of Edinburgh, Edinburgh, U.K., in 1975 and 1978, respectively. He worked for three years as a Research Physicist and for three years as an Integrated Circuit Design Engineer. In 1984, he was appointed a Lecturer in electrical engineering at the University of Edinburgh. He became a Reader in 1991 and Professor of Neural Electronics in 1994. He is interested in all aspects of neural computation and hardware issues and applications have been his primary research interest since 1985. In 1986, he developed the “pulse stream” method for neural integration. His interests have since widened to include all aspects of neural computation, particularly hardware-compatible learning schemes, probabilistic neural computation and neural forms that utilize the temporal- and noisy characteristics of analogue VLSI—as well as applications of hardware neural networks. He is also developing a new interest in the interface between silicon and neurobiology, along with colleagues in Biomedical Sciences and in Glasgow University. He has over 200 publications, including an undergraduate textbook and research texts on neural VLSI, applications of neural networks and noise in neural training (with Peter Edwards).

Dr. Murray is a Fellow of the Institution of Electrical Engineers (IEE) and the Royal Society of Edinburgh, and a Member of INNS.
Addendum