Universidad de Sevilla
Departamento de Electrónica y Electromagnetismo

Estudio, Diseño, Implementación y Test de Retinas Visuales VLSI Sensoras de Contraste Espacial y Temporal

Memoria presentada por:

Juan Antonio Leñero Bardallo

para optar al grado de Doctor por la Universidad de Sevilla

Sevilla, 2010

Estudio, Diseño, Implementación y Test de Retinas Visuales VLSI Sensoras de Contraste Espacial y Temporal

Memoria presentada por:

Juan Antonio Leñero Bardallo

para optar al grado de Doctor por la Universidad de Sevilla
Sevilla, 2010

Directores:

Dr. Bernabé Linares Barranco

Dra. Teresa Serrano Gotarredona

Tutora por la Universidad de Sevilla:

Dra. Adoración Rueda Rueda

Departamento de Electrónica y Electromagnestismo
Universidad de Sevilla

A mis padres

AGRADECIMIENTOS

Llegados a este punto, me gustaría agradecer a mis directores de tesis, Bernabé Linares y Teresa Serrano, su ayuda, tiempo y dedicación durante el tiempo de elaboración de la tesis. Sin sus valiosos consejos, la redacción de este documento no hubiera sido posible.

También me gustaría agradecer a mis padres y hermanos su apoyo y comprensión durante estos años.

He de mencionar al programa JAE (Junta para la Ampliación de Estudios) del CSIC para la formación de investigadores, gracias al cual he obtenido financiación para mi formación como doctor, así como la oportunidad de realizar una estancia de dos meses en el Departamento de Informática de la Universidad de Oslo, tutelado por el profesor Philipp Häfliger.

Son varias las personas que han colaborado en la realización de la tesis. He de citar al grupo de Arquitectura de Computadores (ATC), de la Universidad de Sevilla, por las PCBs y el software suministrados para el test de los sensores diseñados; a Philipp Häfliger por la placa de test y los montajes para lentes suministrados; a Tobi Delbrück por el software jAER y su retina de contraste espacial sumnistrada; y a los revisores anónimos que han contribuido a mejorar la calidad de los artículos que hemos publicado fruto del trabajo de esta tesis.

Por último, me gustaría citar a mis compañeros del CNM en Sevilla, con los que he trabajado e intercambiado conocimientos durante la elaboración de mi tesis. También, gracias a ellos, mi estancia en el CNM durante estos cuatro años ha sido francamente grata.

Índice de Contenidos

Contents

SISTEMAS DE VISIÓN AER SENSIBLES AL CONTRASTE

1 Introducción

1.1 Sistemas de Visión Convencionales y Biológicos

En los últimos años hemos podido apreciar un crecimiento exponencial en las ventas de sistemas de visión. Dicha tendencia no parece detenerse y es, en parte, debida al auge de los sistemas de telefonía móvil. Probablemente, cualquier persona está familiarizada con conceptos como el número de Megapíxeles o la resolución de una cámara. Ello da una idea de la importancia que tales dispositivos tienen hoy en día.

Los sistemas de visión convencionales están basados en frames. Un frame no es más que una matriz bidimensional en la que cada elemento representa a un píxel y almacena una información relacionada con el nivel luminosidad en la región espacial que ocupa dicho píxel. Tales matrices se generan de forma periódica y continua a lo largo del tiempo. Ello conlleva una serie de limitaciones importantes. La primera de ellas es que hay tiempos muertos mientras se lee la información asociada a cada píxel para generar el frame. Los cambios que ocurran en la escena durante esos intervalos de tiempo no serán detectados. Además, los sistemas de visión convencionales no suelen realizar un procesamiento de la información. Simplemente transmiten valores de intensidad luminosa, que puede ser muy redundantes en el caso de que no haya cambios en la escena. Otras limitaciones que encontramos son el hecho de todos los píxeles tienen la misma ganancia (lo cual conlleva problemas en el caso de haya regiones con niveles de luminosidad distintos) y bajo rango dinámico de operación.

Si comparamos las prestaciones de los sistemas de visión convencionales con los biológicos, vemos que los segundos tienen claras ventajas sobre los primeros [4]. Aunque existen sistemas de visión artificial que ofrecen muy buenas prestaciones, normalmente su funcionamiento se restringe a condiciones de iluminación muy específicas y controladas. La retina humana es capaz de ofrecer un funcionamiento excelente en las más diversas condiciones de luminosidad. Además está optimizada para extraer y enviar la información relevante de la escena al cerebro.

Por ello, parece razonable tratar de imitar el funcionamiento de los sistemas biológicos para aprovechar sus ventajas inherentes. Éste es el objetivo de la disciplina conocida como ingeniería neuromórfica. El auge de los sistemas VLSI (Very Large Scale Integration) posibilita incluir grandes poblaciones de píxeles en un chip. Los sistemas de visión biológicos son extraordinariamente complejos y están estructurados en capas de neuronas. Las neuronas de cada capa están masiva-

mente interconectadas con las siguientes y se encarga de realizar un determinado procesamiento de la información. Al contrario que los sistemas de visión convencionales, la información no se transmite en forma de frames de forma periódica. La transmisión de información es continua a lo largo del tiempo y ello posibilita que sea mucho más rápida. Además, la información transmitida no es redundante. Sólo se transmiten los cambios espacio-temporales [4].

Si deseamos implementar tales sistemas en silicio, es necesario usar medios de comunicación eficientes entre las distintas capas de neuronas. Ello supone una limitación cuando se desean interconectar grandes poblaciones de neuronas porque el número de conexiones que pueden realizarse en silicio es bastante limitado. En ese sentido, el método AER (Address Event Representation, [1]-[2]) es altamente eficiente para interconectar grandes poblaciones de neuronas.

Tal como mostraremos más adelante, las retinas AER tienen mayor ancho de banda, menor latencia, mayor rango dinámico y menor consumo que los sistemas basados en frames. Ello justifica, la realización de esta tesis.

El objetivo de la tesis consistió en diseñar dos sensores AER de visión. El primero de ellos es capaz de detectar el contraste espacial y el segundo, el temporal.

1.2 Extracción del Contraste Espacial

Desde que el primer sensor AER, fuera diseñado por Mahowald y sus colegas, [6], en 1992, han aparecido varios sensores AER capaces de detectar el contraste temporal. Es de especial relevancia el trabajo [44] presentado por K. Boahen y A. Andreou en 1996. Aunque las salidas no eran AER y estaban basadas en frames, los autores implementaron un sistema neuromórfico capaz de detectar el contraste espacio-temporal. La retina era capaz de captar imágenes con aceptable calidad, presentaba un consumo muy reducido y realizaba un AGC local, lo cual dotaba al sensor de un gran rango dinámico. Sin embargo, tenía una serie de limitaciones, como el el alto mismatch. En particular, el mismatch ha sido y es uno de los campos de batalla de los sensores de contraste espacial. Posteriormente, los autores implementaron una retina similar a la enterior con salida AER [21]-[22]. Sus principales limitaciones seguían siendo el elevado mismatch junto con el consumo innecesario de ancho de banda cuando no se detectaba contraste espacial.

Una de las tareas a realizar durante este tiempo consistió en la implementación de un sensor sensible al contraste espacial [23] que solventara las limitaciones antes comentadas. En concreto, partiendo del trabajo de Boahen, se diseñó un nuevo sensor AER robusto frente al mismatch y que sólo consumía ancho de banda cuando se detectaba contraste espacial.

Para hacer el sensor robusto frente al mismatch, se incorporó y diseñó un nuevo sistema de calibrado. Además, el sensor incorpora un modo opcional de funcionamiento (time-to-first spike), basado en las teorías de Thorpe [11]. Este método de funcionamiento combina las ventajas del procesamiento basado en frames y el de los sistemas AER. En él, tras un reset global los píxeles que detectan información más relevante disparan primero. Luego, tras un periodo de escaneo, todos los píxeles son reseteados y el ciclo comienza de nuevo. El cerebro humano realiza un procesamiento de la información similar cuando por ejemplo, giramos la cabeza de forma repentina y miramos hacia un lugar distinto.

1.3 Extracción del Contraste Temporal

La detección del movimiento es una de las tareas más importantes que se realiza en el cerebro. Gracias a ello, se puede adquirir información muy relevante de nuestro entorno. Por ejemplo, permite agrupar objetos, estimar profundidades o estructuras tridimensionales.

Los sensores AER de contraste espacial se desarrollaron con posterioridad a los de contraste temporal. El primer trabajo destacable en esta línea fue presentado por Mallik y sus colegas en 2005 [7]-[8]. Era un dispositivo basado en frames que incorporaba algunas de las ventajas de los sistemas AER. Básicamente computaba los cambios relativos en iluminación entre frames consecutivos. Los eventos se generaban en instantes de tiempo discretos. El sensor combinaba las ventajas típicas de los sensores APS (Active Pixel Sensor) convencionales con algunas de los sistemas AER.

Posteriormente, hay que destacar el trabajo de Lichsteiner y Delbrück [10]-[9]. Los autores fueron capaces de implementar un sensor de contraste temporal puramente AER con algunas características bastante destacables: alto ancho de banda (>2*KHz*), baja latencia ($15\mu s$ @1*Klux*), amplio rango dinámico (120dB), bajo consumo ($24mW$) y buena resolución (128 x 128 píxeles).

El segundo objetivo de la tesis consistió en implementar un sensor AER de contraste temporal que mejorara algunas de las prestaciones del sensor antes comentado. En particular, la nueva retina fue diseñada para poder detectar movimiento a muy alta velocidad y con un tiempo de respuesta bastante bajo. Además, el tamaño del píxel se redujo con respecto a la retina de Lichsteiner y sus colaboradores.

2 Corrientes de Polarización Programables (The I-Pot System)

2.1 Introducción

Los circuitos analógicos requieren en general de una serie de corrientes de polarización para su funcionamiento. Dichas corrientes son normalmente proporcionadas por un circuito externo, comúnmente situado en la periferia del chip y destinado a tal fin. Sin embargo, a veces, es deseable ajustar los valores de estas corrientes de polarización para poder compensar variaciones del proceso y el mismatch de los componentes del circuito. Otras veces, cuando se testa un chip, simplemente se desea tener la posibilidad de experimentar con valores distintos de polarización. Para ello, una solución trivial es disponer de un pin externo por cada corriente de polarización del circuito. Sin embargo, el número de pines disponibles es limitado.

Con la intención de generar bias totalmente programables y solventar la limitación antes comentada, se ideó el *I-Pot Estocástico* [35]. Los *I-Pots* son fuentes de corrientes programables mediante una palabra digital de control. Cada *I-Pot* es una celda que recibe una corriente de referencia a la entrada y proporciona la corriente deseada con precisión por debajo de los pico-amperios. El número de pines externos para programar, caracterizar y alimentar los *I-Pots* es tres, independientemente del número de celdas programables que se incorporen al chip. En un proceso de $0.35\mu m$ estándar, el área de un *I-Pot* es de $130\mu m \times 68\mu m$, lo que equivale a una tercera parte del pad al que reemplaza.

2.2 Arquitectura de los *I-Pots*

Los *I-Pots* están basados en estructuras en escalera construidas con transistores MOS [32]-[28]. En la Figura 1, se muestra una estructura en escalera genérica construida con transistores MOS y configurada en modo de fuente de corriente. Todos los transistores tienen un tamaño proporcional a uno unitario de dimensiones *W/L*. El factor de proporcionalidad puede ser *1, N-1, N/(N-1).* De esta forma, las corrientes de cada rama I_i tienen un factor de atenuación de valor *N* con respecto a la rama previa. El *I-Pot Estocástico* usa dos de estas estructuras en escalera, tal como podemos ver en la Figura 2, donde se presenta el diagrama de bloques completo de un *I-Pod*. La

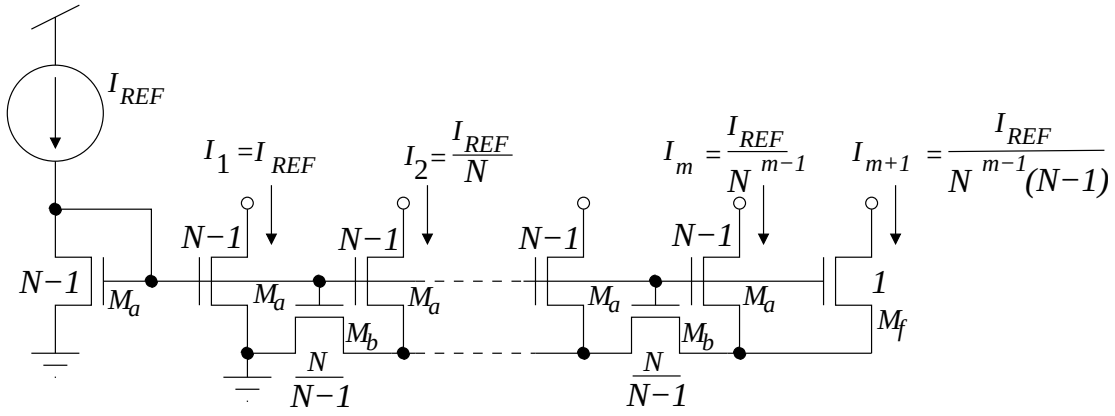


Fig 1: Estructura en escalera genérica con *N* ramas. Se indica el valor de la corriente de cada rama.

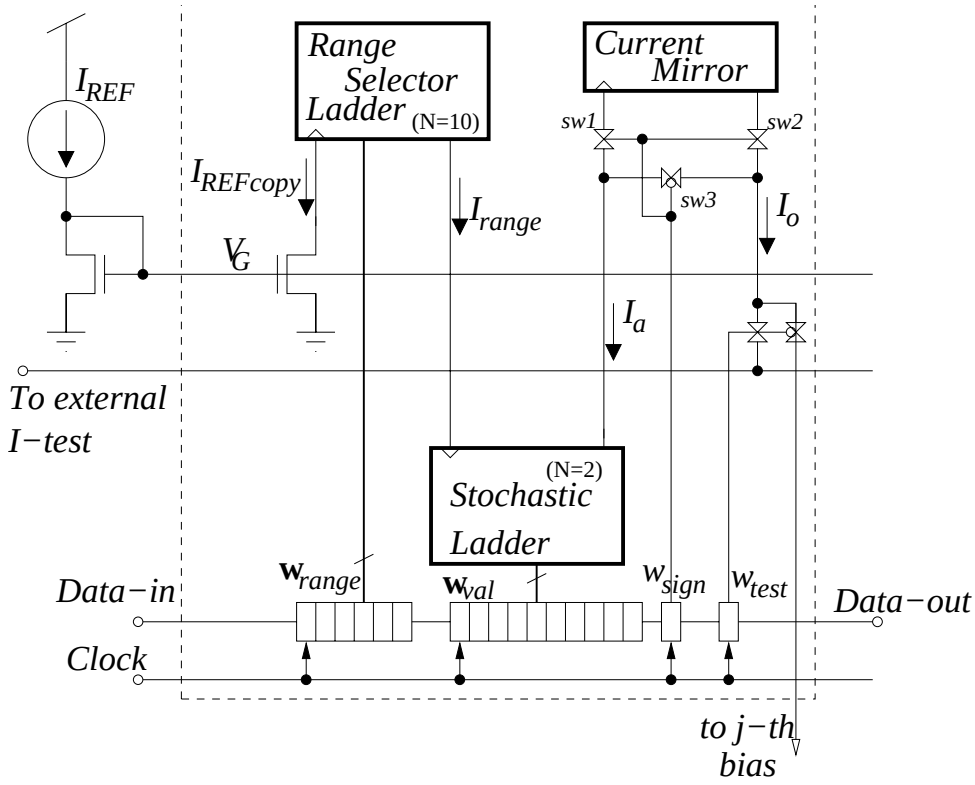


Fig 2: Diagrama de las distintas partes de cada *I-Pot* o celda programable.

primera de ellas (Range Selector Ladder) con un valor de atenuación de *N=10* selecciona el rango de operación. Tiene 6 ramas de salida y la corriente de referencia de entrada I_{REF} puede tomar un valor a la salida comprendido entre I_{REF} *e* $I_{REF}/10^6$. La segunda de ellas (Stochastic Ladder) con un valor de atenuación de *N=2* permite hacer combinaciones binarias de sus ramas para generar un valor preciso de intensidad a su salida.

En nuestro diseño, la segunda estructura en escalera hace uso de transistores de tamaño pequeño. De esta forma, se consiguen valores muy dispares que cubren de forma homogénea todo

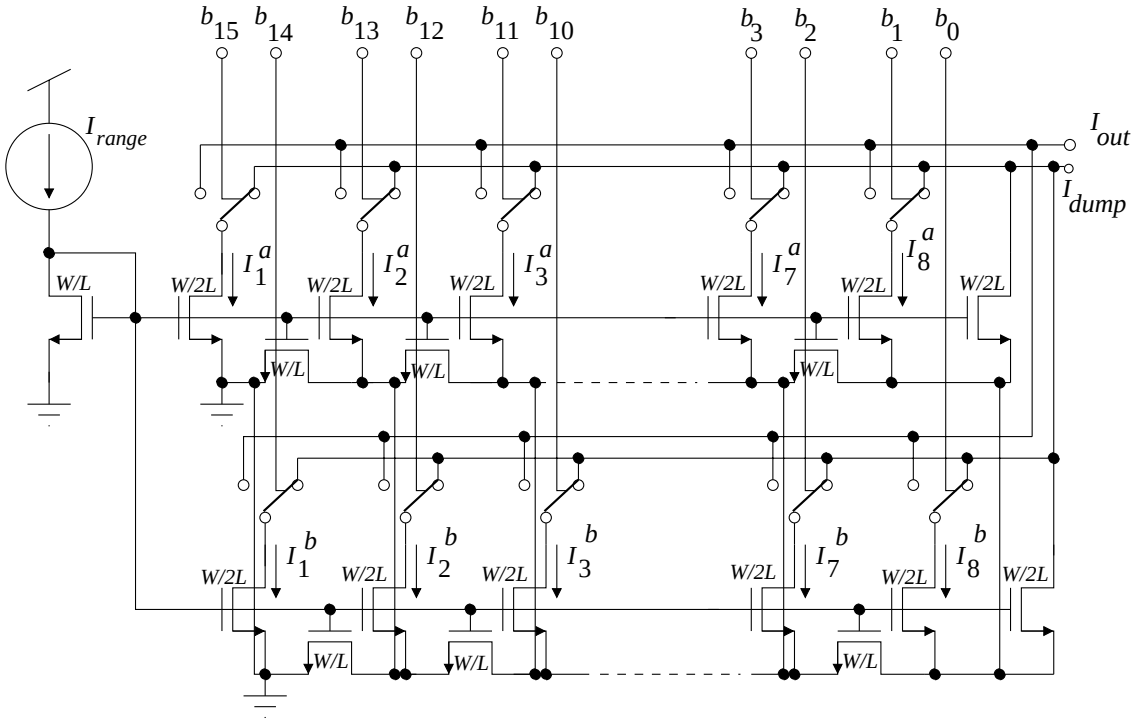


Fig 3: Esquemático de una estructura en escalera con factor de atenuación *N=2* y ramas de salida duplicadas.

el rango de operación. Además las ramas de salida de esta escalera están duplicadas tal como puede verse en la figura Figura 3. De esta forma se consigue un rango continuo de valores posibles de salida sin que haya grandes discontinuidades entre los valores de corriente que se pueden obtener con cada *I-Pot*.

Volviendo al esquema completo de cada *I-Pot* de la Figura 2, vemos que cada celda de corriente es alimentada por una corriente externa de referencia I_{REF} a través de un espejo PMOS. Dicha corriente es la entrada a la primera escalera que selecciona el rango de operación. La salida de este bloque es la entrada de la escalera con rango de atenuación *N=2* que permite hacer un ajuste fino de la corriente de salida. Finalmente a la derecha, tenemos una circuitería que permite seleccionar el signo de la corriente y decidir si la salida se conecta a un pin externo de test o bien alimenta al circuito que corresponda. Todo se programa mediante una palabra digital de control. Con $\boldsymbol{w}_{range}$ se selecciona el rango de operación, con $\boldsymbol{w}_{val}$ se selecciona uno de los posibles valores dentro del rango elegido, con $\boldsymbol{w}_{sign}$ se elige el signo de la salida y con $\boldsymbol{w}_{test}$ se decide si la corriente de salida va a un pin externo de test.

La única desventaja de usar los *I-P*ots es que cada uno necesita ser caracterizado una vez que el chip ha sido fabricado. Sin embargo, el proceso es sencillo y sólo necesita realizarse una vez. Para ello, sólo es necesario usar un ordenador que vaya cargando los registros con todos los posibles valores y un amperímetro conectado al pin de salida. El procedimiento detallado para caracterizar los *I-Pots* es el siguiente:

1. Cada *I-Pot* necesita ser caracterizado independientemente. Por tanto, sólo el bit $\boldsymbol{w}_{test}$ de uno de los *I-Pots* puede estar activo. Así, sólo uno de ellos está conectado a la línea externa I_{test}.
2. Se barren los dos signos del *I-Pot* activo.

3. Para el *I-Pot* y el signo seleccionado, se barren todos los rangos de corriente mediante la palabra digital $\boldsymbol{w}_{range}$.
4. Para el valor elegido de $\boldsymbol{w}_{range}$, se barren las 20 ramas de salida, se miden cada uno de los posibles valores de salida a través del pin externo I_{test} y se almacenan en el ordenador.

Una vez que se ha completado el proceso de medida del *I-Pot*, tendremos almacenados en el ordenador un total de 2 signos x 6 rangos x 16 ramas=192 valores de corriente. Para cada signo y para cada rango, podemos generar 2^{16} combinaciones. Finalmente, con un sencillo programa de ordenador podemos obtener los valores óptimos de $\boldsymbol{w}_{range}$ y $\boldsymbol{w}_{val}$ que minimizan el error cuando se desea obtener una determinada corriente de salida.

2.3 Resultados experimentales

Para caracterizar los *I-Pots,* se fabricaron un conjunto de ellos en tecnología AMS de $0.35\mu m$. La escalera con factor de atenuación *N=2* se implementó con 16 ramas de salida. El tamaño de los transistores unitarios de esta escalera era $W = 1\mu m$ y $L = 0.7\mu m$. Se utilizó una $I_{REF} = 100\mu A$.

Si obtenemos todos los posibles valores de salida de un *I-Pot* y los ordenamos, podemos definir el incremento entre dos valores consecutivos, Δ_{rel}, como

$$\Delta_{rel} = 2\frac{|I_n - I_{n+1}|}{|I_n + I_{n+1}|} \tag{1}$$

Y conociendo el valor de Δ_{rel}, podemos expresar la precisión de cada valor de salida en bits:

$$1/2^{n_{bits}} = \Delta_{rel} \Leftrightarrow n_{bits} = -\log_2(\Delta_{rel}) = -(\ln\Delta_{rel})/(\ln 2) \tag{2}$$

En la Figura 4, se han representado los incrementos entre los valores medidos y ordenados de intensidad para cada uno de los 6 rangos de operación. Puede verse que la mayor precisión se alcanza en las zonas centrales de los rangos de operación, mientras que los peores resultados se obtienen en los extremos de los rangos. Para el valor más pequeño del rango de operación, la medidas presentan estriaciones. Ello puede deberse a errores debidos al instrumento de medida. En la Figura 5, se presentan la precisión en bits de todos los valores posibles que pueden obtenerse con todos los rangos de operación. Vemos que hay zonas en las que se alcanza una resolución superior a los 13 bits. En las regiones correspondientes a los extremos de los rangos, se alcanza una precisión de 8 bits. Obviando los extremos de los 6 rangos superpuestos, la precisión mínima que se alcanza es de 8.1 bits aproximadamente.

2.4 Conclusiones sobre los *I-Pots*

Los *I-Pots* son una poderosa herramienta para polarizar y testar circuitos neuromórficos. Permiten generar un número arbitrario de corrientes de polarización programables con una precisión por debajo de los pico-amperios. Tan sólo tres pines de salida son necesarios para incorporar-

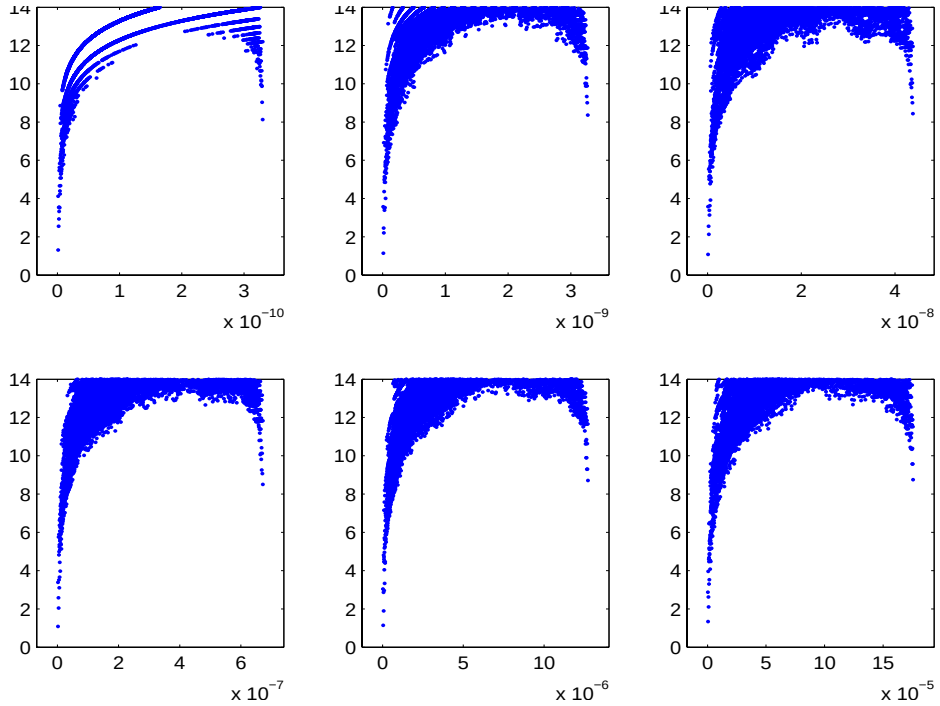


Fig 4: Incrementos relativos entre valores consecutivos ordenados y expresados en bits. El eje de ordenadas indica la precisión en bits, mientras que el de abcisas indica el valor de la corriente.

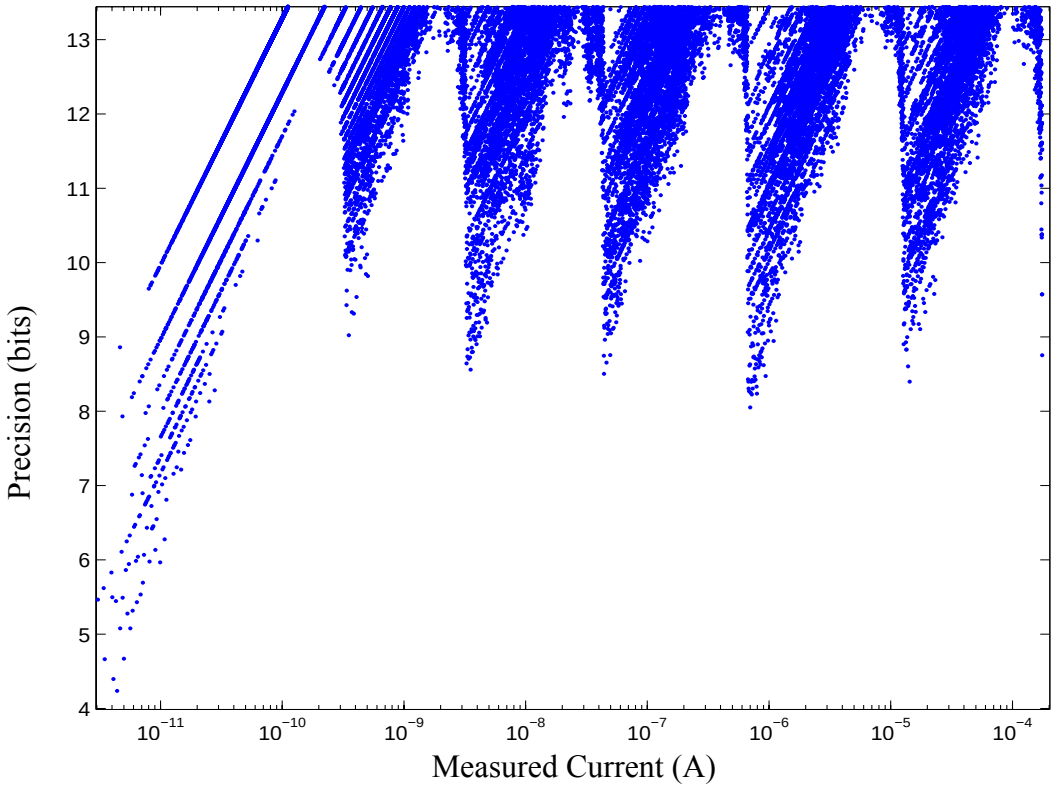


Fig 5: Incrementos relativos entre valores de corriente cuando se ordenan todos los valores obtenidos en cada rango. El eje de abcisas indica el valor de la corriente y el eje de ordenadas indica la precisión en bits.

los en un chip. La única desventaja es que deben ser caracterizados una vez que se haya fabricado el chip. El área de un *I-Pot* es de $130\mu m \times 68\mu m$.

Los *I-Pots* fueron testados por separado y usados para polarizar y testar las retinas que se describirán a continuación en este documento. Han demostrado ser muy útiles en el test de circuitos neuromórficos donde se necesitan números elevados de corrientes de polarización y el número de pines externos disponibles es limitado.

3 El Nuevo Método de Calibración (Newcalib)

3.1 Introducción

Durante los últimos 20 años, una gran variedad de sistemas VLSI neuromórficos han sido publicados. Normalmente, tales sistemas están compuestos por una gran número de píxeles que forman un array. La tendencia actual es reducir al máximo el tamaño de los píxeles y el consumo. Para ello, corrientes del orden de nano amperios o menores tienen que ser usadas. Esto da lugar necesariamente a un alto mismatch. A pesar de los sistemas VLSI neuromórficos, tienen un gran potencial, aún no están suficientemente maduros para ser comercializados. Una de las razones de que esto ocurra, es el alto mismatch que presentan. En concreto, una de las grandes limitaciones de las retinas de contraste espacial implementadas hasta la fecha ha sido el mismatch que empeoraba la calidad de las imágenes captadas.

Si queremos reducir el mismatch sin aumentar el tamaño de los transistores, la única solución viable es la calibración. Existen técnicas de calibración para redes neuromórficas [14]-[27], basadas en DACs compactos realizados con estructuras en escalera programables [28]. El problema de tales estructuras, era que la precisión estaba limitada a un punto (a un valor concreto para el cual se calibraba). En nuestro caso, hemos diseñado un nuevo sistema de calibración multi punto [13], mucho más compacto que una estructura en escalera, que permite escalar las corrientes, una vez que el sistema ha sido calibrado, sin que la precisión del calibrado se degrade de forma notoria. Tal sistema fue testado e incorporado en los píxeles de la retina de contraste espacial para compensar su mismatch.

3.2 Nuevo Sistema de Calibrado (Newcalib)

El nuevo método de calibración [13] está basado en transistores MOS cuya longitud es programable mediante una palabra digital de control, $\boldsymbol{w}_{cal}$. En la Figura 6, podemos ver los esquemáticos de la nueva estructura de calibrado y su símbolo. El circuito de Figura 6(a) entre los terminales D, G y S, se comporta exactamente igual que un transistor MOS digitalmente ajustable mediante la palabra digital $\boldsymbol{w}_{cal} = \{b_{N-1}b_{N-2}\ldots b_2 b_1 b_0\}$ Y su longitud equivalente viene dada por $L_{eq} = (W/2) \times g(\boldsymbol{w}_{cal})$, donde

$$g(\boldsymbol{w}_{cal}) = \sum_{i=0}^{N-1} \frac{b_i}{2^{N-1-i}} \tag{3}$$

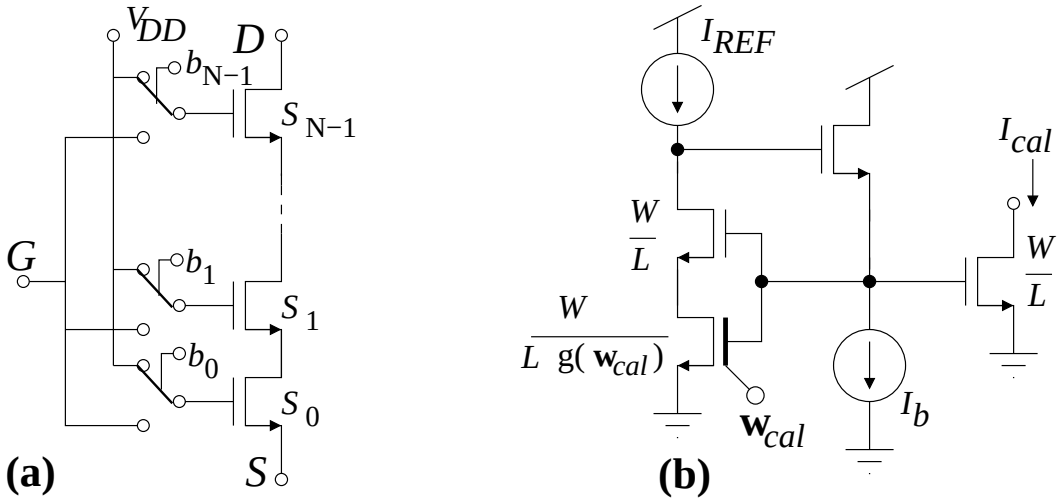


Fig 6: (a) Esquemáticos del transistor MOS con longitud controlada digitalemente. (b) Posible aplicación para calibrar una fuente de corriente.

La palabra digital es almacenada en una memoria estática de tipo RAM cuando el circuito empieza a funcionar. La longitud equivalente del transistor puede ser ajustada entre 0 y $2-(1/2^{N-1})$ en pasos de $1/2^{N-1}$. Este transistor puede ser usado como parte de un espejo de corriente, tal como es mostrado en la Figura 6(b), para proporcionar una corriente de calibración $I_{cal} = I_{REF} \times (g(\boldsymbol{w}_{cal}) + 1)$. En la Figura 7, se muestran los resultados de una simulación Montecalo con 100 iteraciones distintas del circuito de la Figura 6(b). Se utilizó una palabra digital de 4 bits. En el eje de abcisas, se fue variando el valor de la palabra de calibrado, $\boldsymbol{w}_{cal}$, y en el eje de ordenadas, se representa el valor de la intensidad de calibrado resultante. El circuito se polarizó con una corriente de valor $I_{REF} = 3nA$ y tenía transistores unitarios de tamaño $1\mu/4\mu m$. La simulación se realizó usando modelos de un proceso CMOS estándar de $0.35\mu m$.

3.3 Uso de Circuitos Translineales para el Calibrado

El circuito de la Figura 6(b), tiene el grave inconveniente de que sólo permite calibrarlo en un punto. Es decir, si escalamos la corriente de referencia I_{REF} el calibrado no se conserva y la corriente I_{cal} no se escala en la misma proporción. Para conseguir esto, introducimos bucles translineales. De esta forma, se consigue mantener fijas algunas corrientes, incluida la del transistor con longitud programable, y se escalan otras. Esto se muestra en la Figura 8. La circuiteria encerrada dentro de las líneas discontinuas se replica dentro de cada píxel, mientras que el resto aparece una vez en la periferia del chip. El grupo de transistores que va de M_1 a M_4 forma un bucle translineal, lo cual implica que $I_{oi} = I_1 I_2 / I_{i3}$. La corriente local I_{i3} se copia desde la corriente periférica I_3, mediante un espejo de corriente local con una un transistor de longitud programable. Para conseguir un rango de calibración con factor 2, se incluyen dos transistores en serie en el espejo de corriente de salida. Uno de tamaño fijo $W/2L$ y el otro de tamaño calibrable. Consecuentemente, $I_{i3} = I_3/(2+g(\boldsymbol{w}_{cal_i}))$ y

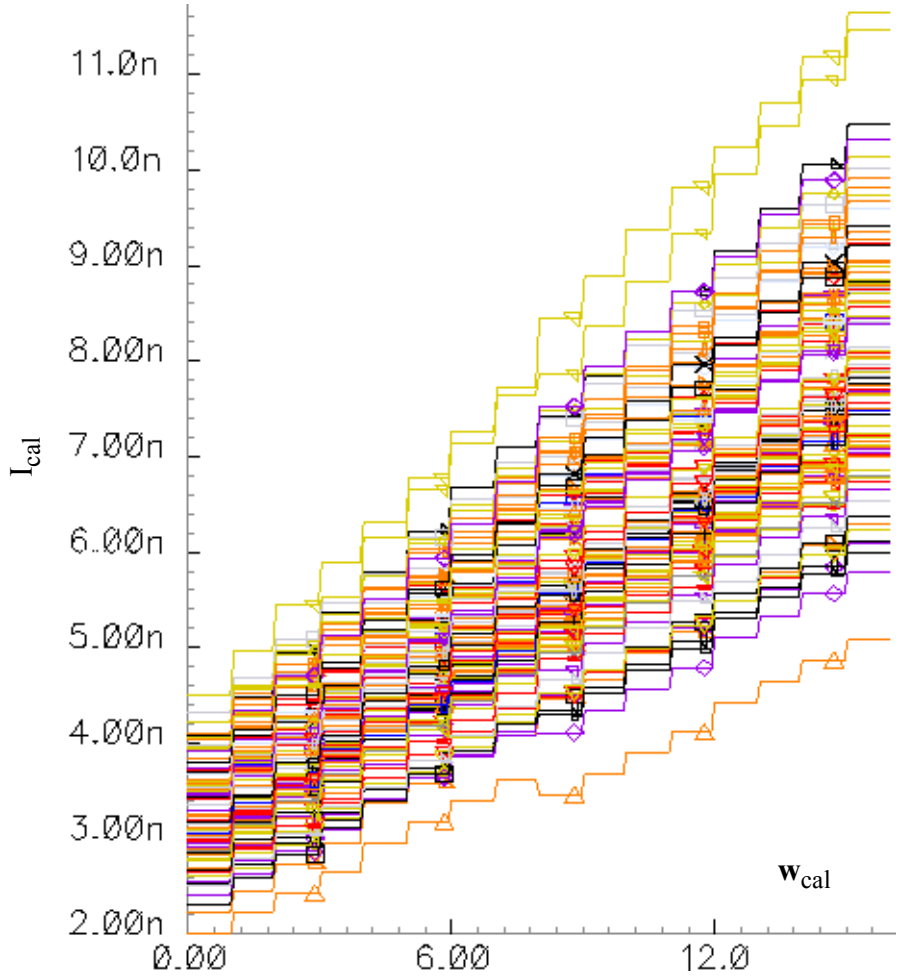


Fig 7: Simulación Montecarlo (con 100 iteraciones) del circuito de la Figura 6(b), usando un transistor MOS de longitud ajustable con una palabra digital de 4 bits.

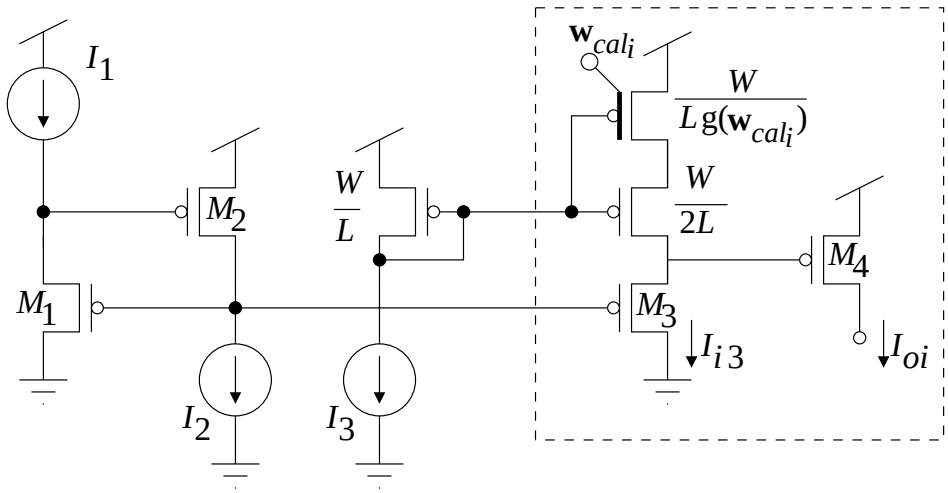


Fig 8: Circuito translineal usado para permitir ajustar el rango del circuito de calibrado.

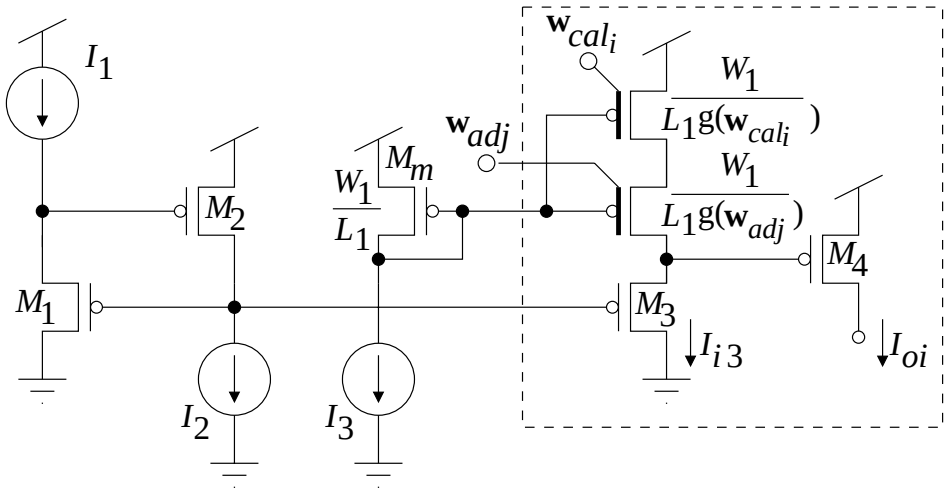


Fig 9: Primera estrategia para optimizar el rango de calibración.

$$I_{oi} = \frac{I_1 I_2}{I_3}(2 + g(\boldsymbol{w}_{cal_i})) \tag{4}$$

Con este circuito, uno puede mantener (después de calibrar) constante las corrientes I_3 (e I_{i3}) e I_1, mientras I_2 se ajusta globalmente para escalar por arriba o por abajo las corrientes locales I_{oi}.

3.4 Estrategias para el Calibrado

Basándonos en las estructuras para el calibrado con bucles translineales, se propusieron y testaron dos posibles esquemas de calibrado con bucles translineales. En la Figura 9 y la Figura 10, se pueden ver las dos estrategias usadas para implementar el calibrado.

Con la primera estrategia de calibrado, mostrada en la Figura 9, dos transistores con longitud ajustable mediante una palabra digital son usados. Uno de ellos es ajustado localmente, como en Figura 8, pero el otro es ajustado globalmente y su longitud se varía de la misma forma para todos los píxeles del array. De esta forma, todas las puertas de los segmentos del transistor son controladas desde la periferia (véase Figura 9). Como resultado,

$$I_{oi} = \frac{I_1 I_2}{I_3}(g(\boldsymbol{w}_{adj}) + g(\boldsymbol{w}_{cal_i})) \tag{5}$$

Con la segunda estrategia de calibrado, mostrada en la Figura 10. El bucle translineal se ha replicado dos veces, de forma que se tienen dos bucles translineales en paralelo. Uno de ellos usa la palabra de calibración local, $\boldsymbol{w}_{cal_i}$. El otro es ajustado de forma global a todo el array y sólo el transistor de salida M_x es replicado una vez por píxel. Esto permite usar un transistor de salida con un tamaño mayor y consecuentemente menos mismatch. El propósito del circuito translineal calibrable es compensar el mismatch de M_x.

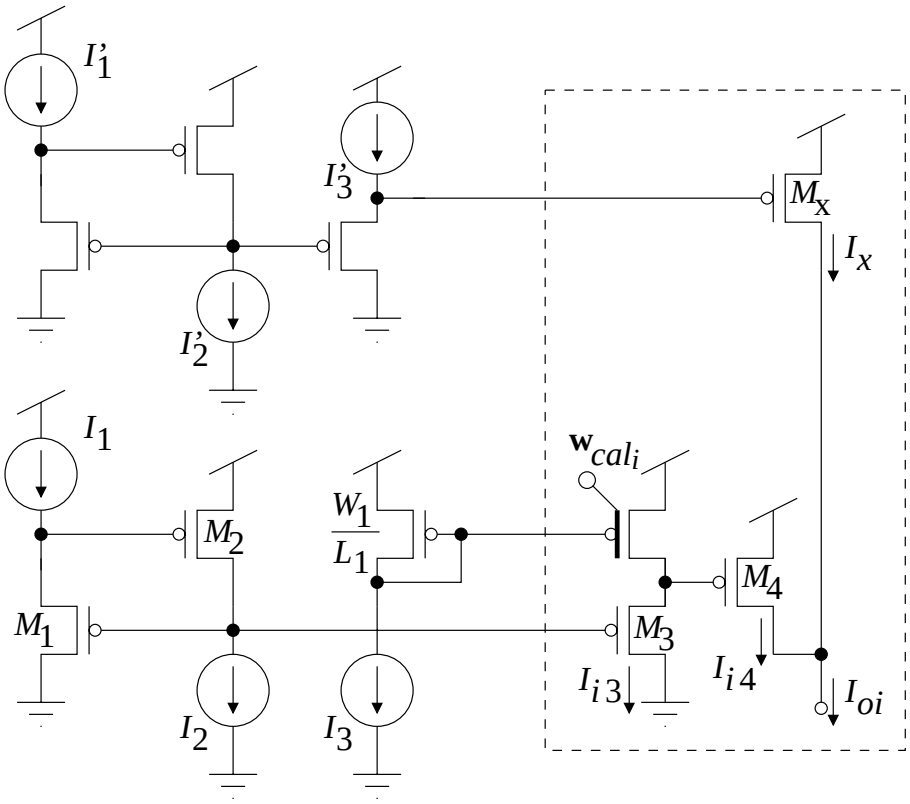


Fig 10: Segunda estrategia para optimizar el rango de calibración.

3.5 Resultados Experimentales

Para caracterizar los nuevos sistemas de calibrado, un chip prototipo se fabricó en un proceso CMOS de $0.35\mu m$. Se fabricaron 20 DACs de corriente de 5 bits cada uno. Diez de ellos usaban la estrategia de calibración de la Figura 9, y los otros diez usaban la estrategia de calibración de la Figura 10.

Cada uno de los 10 primeros DACs usaba 5 réplicas del circuito de la Figura 9, una por cada bit. La corriente nominal de salida de cada uno (I_{oi}) se ajustaba para que fuera escalada variando la palabra digital. Consecuentemente, en la periferia, necesitamos 5 grupos de fuentes de corriente $\{I_1, I_2, I_3\}$ y 5 grupos de transistores $\{M_1, M_2, M_m\}$, uno por cada bit. Sin embargo, esos 5 grupos de fuentes de corriente periféricas y transistores son compartidos por todos los 10 DACs.

Cada uno de los segundos 10 DACs usa 5 réplicas del circuito de la Figura 10. De nuevo, para cada uno de los 10 DACs, la circuitería se replicó 5 veces (una por cada bit), y la circuitería periférica (fuera de las líneas con trazo discontinuo de la Figura 10) es compartida, por cada bit, por los 10 DACs.

El área del layout del circuito dentro de la línea discontinua es $18 \times 14\mu m^2$ para la Figura 9 y $17 \times 14\mu m^2$ para el de la Figura 10.

La Figura 11(a) muestra las corrientes de salida medidas de forma experimental para las 10 réplicas del circuito de la Figura 9 cuando se fija $w_{adj} = 0$. Las corrientes periféricas de salida

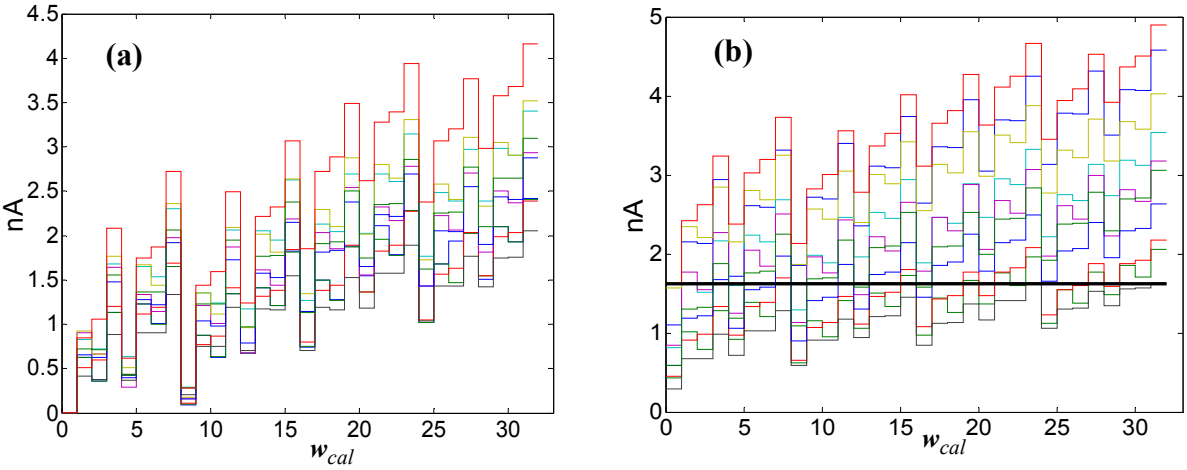


Fig 11: Corrientes de salida obtenidas experimentalmente para el circuito de la Figura 9 (a) para $w_{adj}=0$, (b) para el valor óptimo de w_{adj}. La línea horizontal en (b) es el valor objetivo, el cual es tocado o cortado por las 10 curvas.

eran $I_1 = I_2 = I_3 = 10nA$, y todas las palabras de calibrado w_{cal_i} (i = 1,... 10) fueron barridas de forma simultánea de 0 a 31. Después de repetir estas medidas para todos los valores posibles de w_{adj}, el valor óptimo para w_{adj} corresponde con la situación donde el valor más alto de la izquierda es más cercano al valor más bajo de la derecha. Esta situación se muestra en la Figura 11(b). En este punto podemos obtener las palabras de calibrado óptimas w_{cal_i} que dan la mínima variación. La máxima desviación de las corrientes de salida se obtiene bajo esas circunstancias es $|\Delta I_{oi}|_{max} = 0.57nA$, lo que corresponde al 5.7%, de la corriente nominal $I_b = 10nA$. Si ésta fuera la fuente de corriente controlada por el bit más significativo de un DAC de corriente (con un rango máximo de $20nA$), esto limitaría la precisión del DAC a $-\ln(|\Delta I_{oi}|_{max}/2I_b)/\ln 2 = 5.13\text{bits}$. Para evaluar cómo la calibración se degrada cuando cambian las condiciones de polarización, se barrió I_2 en la Figura 9 entre $100pA$ y $1\mu A$. La máxima desviación entre las 10 fuentes de corriente calibradas se muestra con un trazo con círculos en la Figura 12. El trazo con triángulos corresponde a medidas obtenidas antes de calibrar ($w_{cal_i} = 0$, para todas la i). Podemos ver que las 10 muestras mantienen una precisión de 4 bits para corrientes por encima de $3nA$. El eje de abcisas es un promedio de I_{oi} entre las 10 muestras. También se muestra en la Figura 12 (trazo con cruces) la precisión que resulta después de calibrar obtenida mediante simulaciones. Nótese que es muy optimista, excepto para el punto para el que se calibró ($10nA$). La razón es que en este circuito en particular Figura 9 el calibrado se degrada es debido al mismatch en el factor de forma del transistor. El mismatch de este parámetro no estaba modelado en nuestro simulador.

De forma similar, la Figura 13 muestra la precisión medida antes y después de calibrar de 10 fuentes calibrables de corriente que siguen el esquema de calibración explicado en Figura 10. Ahora el mismatch antes de calibrar era menor que en la Figura 12. Esto es porque ahora el área usada por el transistor de longitud ajustable mediante una palabra digital M_b en la Figura 9, está disponible para el transistor M_x de la Figura 10, con lo que puede ser de mayor tamaño. Con la

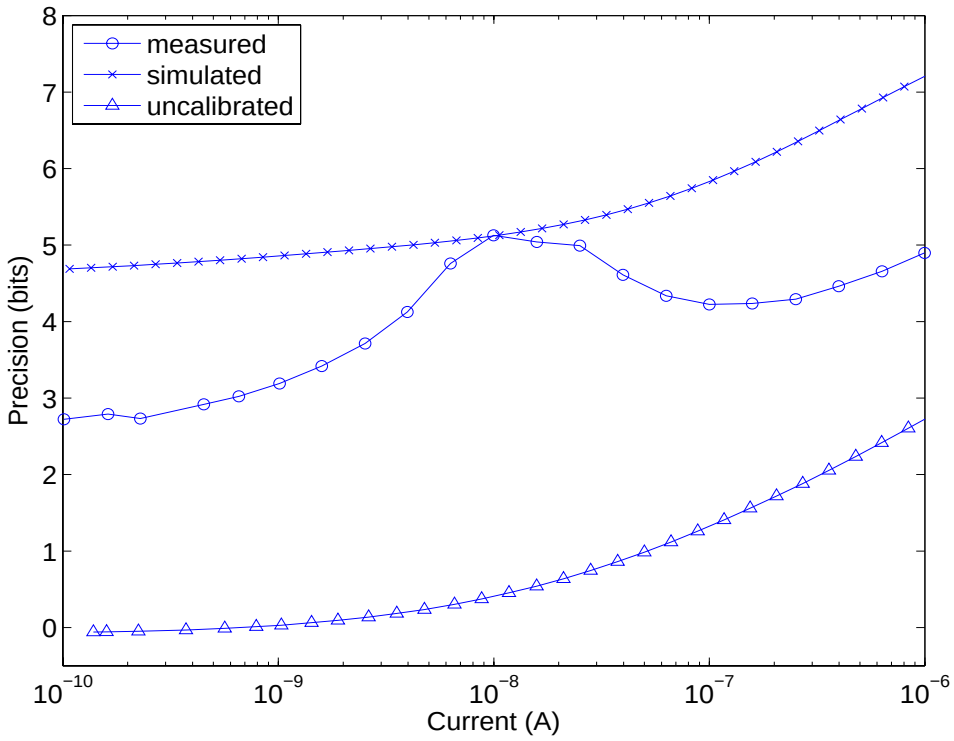


Fig 12: Medidas de precisión de la fuente calibrable y configurable con el esquema propuesto en la Figura 9. Trazo con círculos: precisión medida tras el calibrado (con valor óptimo de w_{calj} para cada una de las fuentes de calibrado). Las fuentes de corriente se calibraron a $10nA$. Curva con triángulos: precisión medida antes de calibrar ($w_{cal_i} = 0$ para todas las fuentes de corriente). Curva con cruces: precisión tras calibrar, obtenida con el simulador.

estructura de la Figura 10 obtenemos mucha más precisión en el punto de calibrado (8.30 bits a $10nA$), pero se degrada rápidamente, especialmente para corrientes altas. La precisión obtenida mediante simulación, es ligeramente pesimista en el punto de calibrado (7.63 bits a $10nA$), pero se degrada de forma más optimista cuando las corrientes se alejan del punto de calibrado.

3.6 Conclusiones sobre el Sistema de Calibrado

Parte del trabajo realizado durante la tesis consistió en implementar un sistema de calibrado compacto y eficiente que pueda ser empleado en sistemas neuromórficos. El sistema de calibrado se testó de forma individual con un chip específico para ello. Existen dos estrategias para implementar el calibrado. Como posteriormente explicaremos, la retina de contraste temporal hace uso de la segunda estrategia de calibrado (Figura 10).

4 La Retina AER de Contraste Espacial

4.1 Introducción

Las retinas están entre los primeros bloques AER que fueron publicados [1]-[2]. En concreto, las retinas de contraste espacial son de especial interés porque proporcionan un flujo de

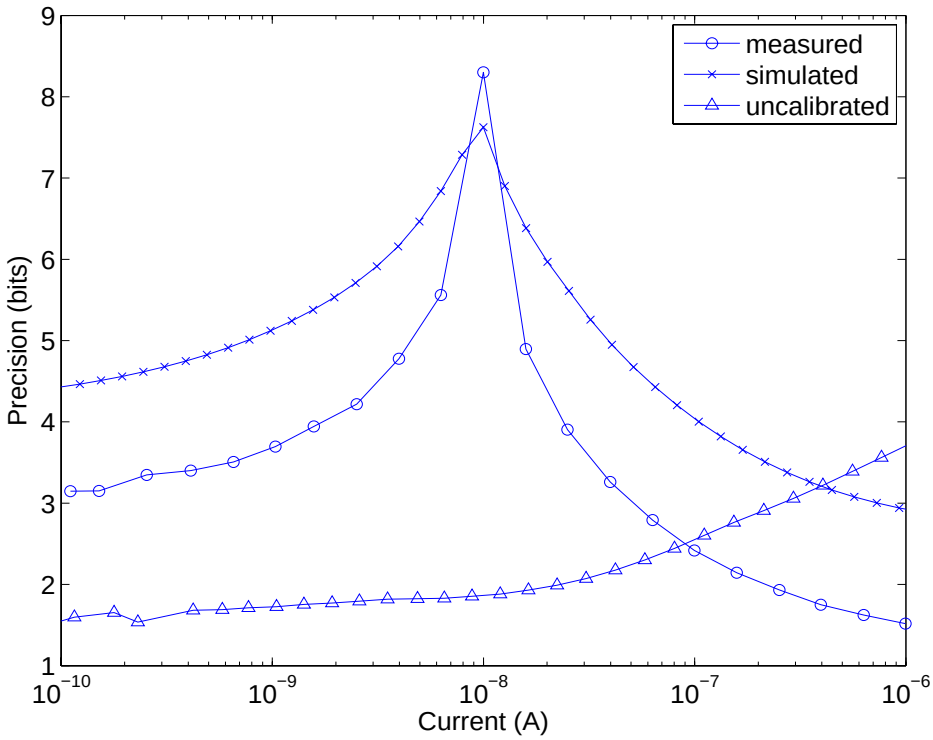


Fig 13: Medidas de precisión de la fuente calibrable y configurable con el esquema propuesto en la Figura 10. Trazo con círculos: precisión medida tras el calibrado. Las fuentes de corriente se calibraron a 10*nA*. Curva con triángulos: precisión medida antes de calibrar ($w_{cal_i} = 0$ para todas las fuentes de corriente). Curva con cruces: precisión tras calibrar, obtenida con el simulador.

datos comprimido, preservando la información necesaria para reconocer objetos. Este tipo de sensores pueden codificar la información de dos formas: La primera de ellas es generando una secuencia de pulsos con una frecuencia proporcional al valor de la señal que se desea transmitir (como luminancia o contraste). La segunda está basada en el modo TFS (Time to First Spike). En este modo de funcionamiento, existe un reset global tras el cual los píxeles que detectan contraste emiten un único pulso. La información se codifica como el tiempo transcurrido entre el reset y el tiempo que tarda cada píxel en generar un pulso. Lógicamente los píxeles que detecten más contraste dispararán primero. A pesar de que el modo TFS está sujeto a las restricciones de los sistemas basados en frames, es un mecanismo altamente eficiente para comprimir la información. Además, el tiempo de frame puede ser ajustado de forma dinámica mediante las etapas de posteriores de procesamiento.

Las salidas de las retinas de contraste espacial publicadas previamente [43]-[4]-[14] son normalmente una corriente de contraste por píxel $I_{cont}(x, y)$. Dicha salida se computa como el cociente entre la foto corriente obtenida al sensar la intensidad luminosa $I_{ph}(x, y)$ y el promedio espacial de la foto corriente en una determinada región alrededor del píxel. El promedio se obtiene normalmente mediante algún tipo de red difusiva

$$I_{cont}(x,y) = I_{ref}\frac{I_{ph}(x,y)}{I_{avg}(x,y)} \tag{6}$$

donde I_{ref} es una corriente global que escala el resultado. La intensidad de salida es siempre positiva (salida unipolar sin signo) y en el caso de que no haya contraste e $I_{ph}(x,y) = I_{avg}(x,y)$, la intensidad de salida de cada píxel será igual a I_{ref}. Esto supone una importante limitación porque los píxeles consumirán ancho de banda de forma innecesaria. Además contradice o viola las ventajas de los sistemas AER, donde sólo la información relevante es transmitida.

En otro trabajo previo [14], el contraste se computaba mediante la fórmula de Weber,

$$I_{cont}(x,y) = I_{ref}\left(\frac{I_{ph}(x,y)}{I_{avg}(x,y)} - 1\right) \tag{7}$$

lo cual solventaba el problema del consumo de ancho de banda en ausencia de contraste y además permitía generar una salida bipolar y detectar el signo del contraste. Para ello, una etapa de post-procesamiento, que anulaba los eventos generados en ausencia de contraste, era usada. Sin embargo, esto hacía que la velocidad de respuesta de la retina se redujera un factor 10 aproximadamente.

En esta sección, vamos a presentar una nueva retina AER sensible al contraste espacial [23], en la que el contraste se computa de la misma forma que en la eq. (7). El diseño se basa en el propuesto por Boahen [43], el cual ha sido mejorado para solucionar algunas limitaciones como el elevado mismatch, la dependencia de las salidas con el nivel de luminosidad ambiente, y la dificultad de ajuste del dispositivo. También añade nuevas prestaciones. La salida es bipolar. Además, un mecanismo de umbralización ajustable se ha incorporado, de forma que los píxeles no generan eventos, a menos que su salida supere un cierto umbral. Por último, la retina también incluye un mecanismo global de reset que permite operar en modo TFS (Time to First Spike) con independencia del nivel de luminosidad ambiente. Un prototipo de 32 x 32 píxeles en tecnología CMOS de $0.35\mu m$ fue fabricado y testado.

4.2 Píxel de Contraste Espacial de Boahen

El sensor de contraste espacial descrito en esta tesis solventa todas las limitaciones encontradas en el diseño de Boahen [44]. El comportamiento del circuito mostrado en la Figura 14 queda aproximadamente descrito mediante las siguientes ecuaciones:

$$\begin{aligned} I_h(x,y) &= I_{ph}(x,y) + a\nabla^2 I_c(x,y) \\ I_c(x,y) &= I_u - b\nabla^2 I_h(x,y) \end{aligned} \tag{8}$$

Obtener una expresión para la salida de cada píxel a partir de las ecuaciones anteriores no es fácil ni intuitivo. La corriente de salida de cada píxel depende de todos los píxeles de la retina.

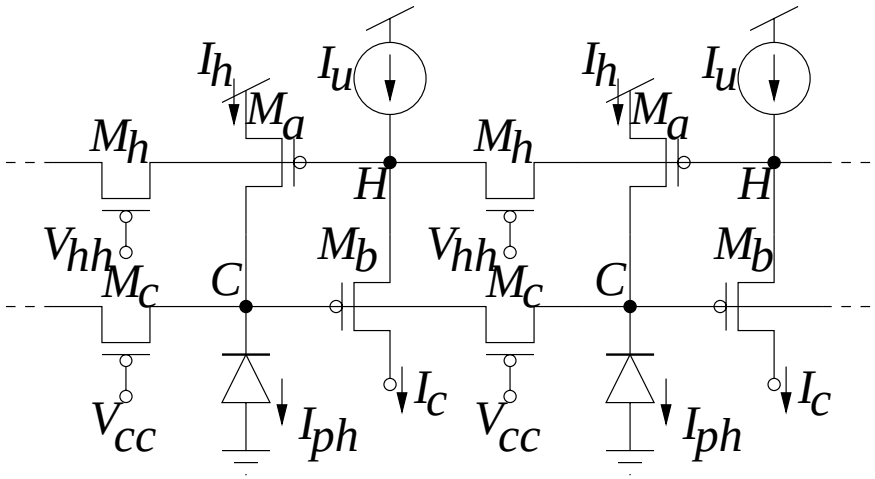


Fig 14: Circuito original de Boahen para el cómputo del contraste espacial.

Una aproximación propuesta por Boahen para la salida de cada píxel [44] a partir de las ecuaciones (8) es

$$I_{cont}(x,y) = I_{ref}\frac{I_{ph}(x,y)}{\langle I_{ph}\rangle + I_{ph}(x,y)} \tag{9}$$

donde $\langle I_{ph}\rangle$ es el promedio de la foto corriente en el vecindario próximo al píxel. Obviamente, los píxeles más cercanos tendrán mayor influencia en la salida. Ajustando I_{cc}, el alcance de los píxeles que afectan a la salida puede ser controlado.

El circuito original de Boahen tiene una serie de limitaciones. En primer lugar, el mismatch era comparable a la propia señal de interés de salida. En segundo lugar, la señal de salida se cambiaba para el mismo estímulo cuando se variaban las condiciones de iluminación. Tercero, los voltajes de control V_{cc} y V_{hh} en la Figura 14 tenían un rango de ajuste muy estrecho y crítico. En la nueva versión de la retina, todos estos inconvenientes han sido corregidos.

4.3 Nuevo Diseño del Píxel

La Figura 15 muestra un diagrama de bloques de las distintas partes que componen el píxel, indicando cuáles son las señales intercambiadas entre los distintos bloques. El píxel contiene tres partes principales: (1) la parte de foto sensado y cálculo del contraste, incluyendo calibración, la cual proporciona la corriente de salida que computa el contraste I_{cont} independientemente del nivel de iluminación; (2) el bloque de integración y disparo, que incluye la circuitería para ajustar el periodo refractario, umbralización, y el modo TFS de disparo; (3) y la circuitería de comunicación del píxel que manda los eventos hacia la periferia e interacciona con ella. Una explicación detallada de todos los elementos del píxel puede encontrarse en [23]. A continuación, vamos a describir en detalle el primero de los bloques.

La Figura 16 muestra cómo el circuito de Boahen para el cómputo del contraste ha sido modificado para incluir un esquema de polarización basado en corrientes, para controlar las seña-

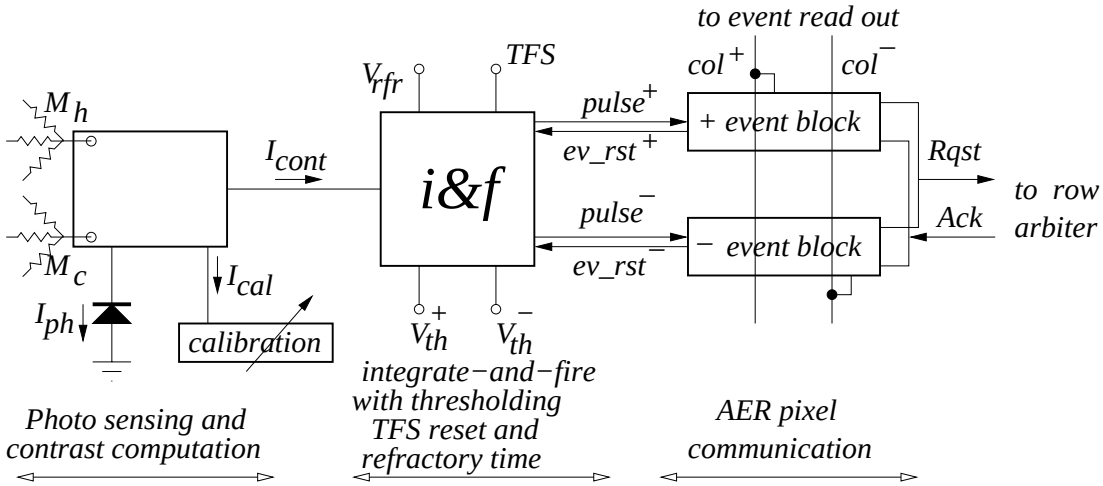


Fig 15: Diagrama de bloques con las partes principales que forman el píxel.

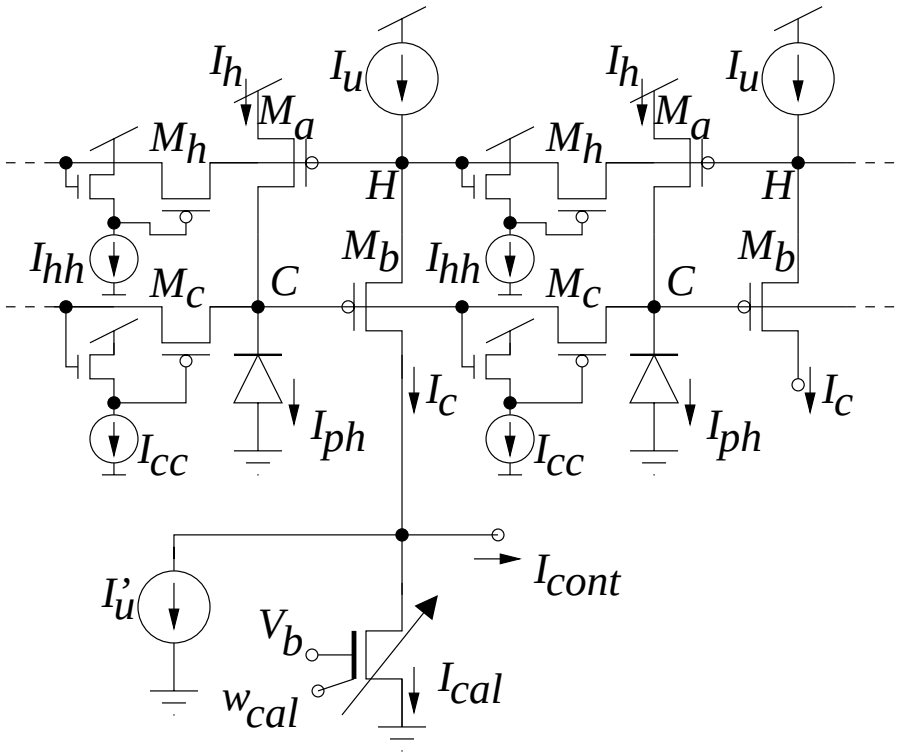


Fig 16: Detalle del bloque de sensado de la foto corriente y computación del contraste.

les originales de control de las redes difusivas V_{cc} y V_{hh}, que pueden verse en el circuito original de la Figura 14. De esta forma, los voltajes en las puertas V_{cc} y V_{hh} tienden a seguir las excursiones de tensión en los nodos 'C' y 'H'.

La primera ventaja de usar este esquema es que las tensiones de los nodos del circuito se adaptarán a las condiciones de iluminación ambiente. Por ejemplo, si todas las corrientes de los

foto diodos se escalan por arriba o por debajo por el mismo factor, el voltaje de todos los nodos 'H' seguirá los cambios de forma logarítmica. Como I_u es constante, el voltaje del nodo 'C' seguirá, por tanto, el mismo desplazamiento en tensión. Como las corrientes de polarización I_{hh} e I_{cc} son constantes y no varían, los voltajes en la puertas de los transistores M_h y M_c seguirán la misma variación global de voltaje, adaptándose ellos mismos a los cambios globales de iluminación.

La segunda ventaja de este esquema es que atenúa el mismatch. Después de hacer un estudio concienzudo del mismatch e identificar las principales fuentes de mismatch del circuito, se descubrió que el transistor M_a y la corriente I_u eran las fuentes dominantes. Esto puede ser entendido si tenemos en cuenta lo siguiente. Las variaciones en la corriente I_u afectan directamente a la componente en DC de I_c, que será calibrada con I_{cal}. El mismatch de M_b es menos crítico porque las variaciones de tensión en el nodo 'C' afectan a la red difusiva inferior y al cómputo de la corriente promedio I_h de la eq. (8). Por tanto, el efecto de su variabilidad es atenuado al hacer el cómputo de la corriente promedio. Sin embargo, el mismatch de M_a (las variaciones de V_{gs} de M_a) cambian directamente el voltaje en la fuente de M_b, afectando directamente a la ganancia del contraste de salida (el coeficiente 'b' en la eq. (8)), cuyo efecto no es directamente calibrado por I_{cal}. Consecuentemente, M_a necesita ser dimensionado para minimizar el mismatch. El efecto de I_u es compensado mediante calibración, y el efecto de M_a será atenuado por el nuevo esquema de polarización. Nótese, que el mismatch en todos los transistores M_a introducirá variaciones aleatorias en los nodos 'H' y 'C'. Tales variaciones se transformarán en corrientes laterales aleatorias a través de los transistores M_h y M_c. Las corrientes aleatorias a través de M_h se añadirán a la corriente de salida I_c y pueden ser compensadas mediante calibración. Sin embargo, las corrientes aleatorias a través de los transistores M_c afectan si fueron generadas por los foto diodos. Gracias al nuevo esquema de polarización, un incremento en 'C' incrementará el voltaje de la puerta del nuevo transistor NMOS, incrementando su voltaje en la fuente, y a su vez, incrementando el voltaje en la puerta de M_c, lo cual reduce la corriente lateral. Un efecto similar ocurrirá con los transistores M_h.

Finalmente, la tercera ventaja es un mecanismo más robusto de polarización de los transistores laterales. En el esquema original, las tensiones V_{cc} y V_{hh} tenían un rango de ajuste muy pequeño y crítico (aproximadamente $100mV$ o menos). Ahora, las corrientes I_{cc} e I_{hh} pueden ser variadas varias décadas, percibiendo aún sus efectos.

4.4 Resultados Experimentales

Se fabricó un prototipo de 32 x 32 píxeles de la retina AER de contraste espacial con signo. El sensor se diseño y fabricó en una tecnología CMOS de 0.35µm con 4 metales y dos tipos de polisilicio, con una alimentación de $V_{DD} = 3{,}3V$. La Tabla 1 resume las especificaciones del chip. La Figura 17 muestra una micro-fotografía del dado, de tamaño 2.5 x 2.6 mm^2. El chip completo, excepto el anillo de pads, está cubierto con la capa de metal más alta dejando aperturas para que los foto diodos puedan captar la luz. La Figura 17 también muestra el layout de un píxel indi-

Tabla 1: Especificaciones del chip.

tecnología	CMOS 0.35μm 4M 2P
tensión de alimentación	3.3V
tamaño del chip	2.5 x 2.6 mm^2
tamaño del array	32 x 32
tamaño del píxel	80 x 80 μm^2
fill factor	2.0%
eficiencia cuántica del fotodiodo	0.34 @ 450nm
complejidad del píxel	131 transistores + 2 caps
consumo de corriente	65μA @ 10keps

vidual resaltando sus principales partes. El layout de cada píxel es simétrico y está especulado con respecto al de sus vecinos más próximos. De esta forma, las líneas digitales ruidosas son compartidas entre los vecinos, al mismo tiempo que las líneas de alimentación, y las líneas sensibles al ruido. También, las líneas más sensible a las perturbaciones son separadas de las ruidosas. El área del píxel es $80 \times 80 \mu m^2$, incluyendo el rutado. El píxel estaba compuesto de 131 transistores y 2 capacidades (la capacidad del circuito de integración y disparo y la capacidad del circuito que genera el reset y controla el período refractario).

La retina se sometió a múltiples experimentos y pruebas en condiciones muy diversas de iluminación para evaluar sus prestaciones. La Tabla 2 resume sus características más significativas. Para una explicación detallada de cómo se tomaron las medidas, véase [23]. Las características más significativas del sensor son: muy bajo FPN ($\sim 0.6\%$), bajo consumo (0.66-66mW), baja

Tabla 2: Especificaciones del sensor de contraste espacial

Funcionalidad	Dependencia con el nivel de luminsidad	Latencia	Rango dinámico	FPN	Fill Factor	Tamaño del píxel μm^2	Proceso de fabricación	Consumo
Contraste espacial a número de eventos	NO	0.1-10ms	>100dB	0.6%	2%	80x80	0.35 μm 4M 2P	0.66-66mW

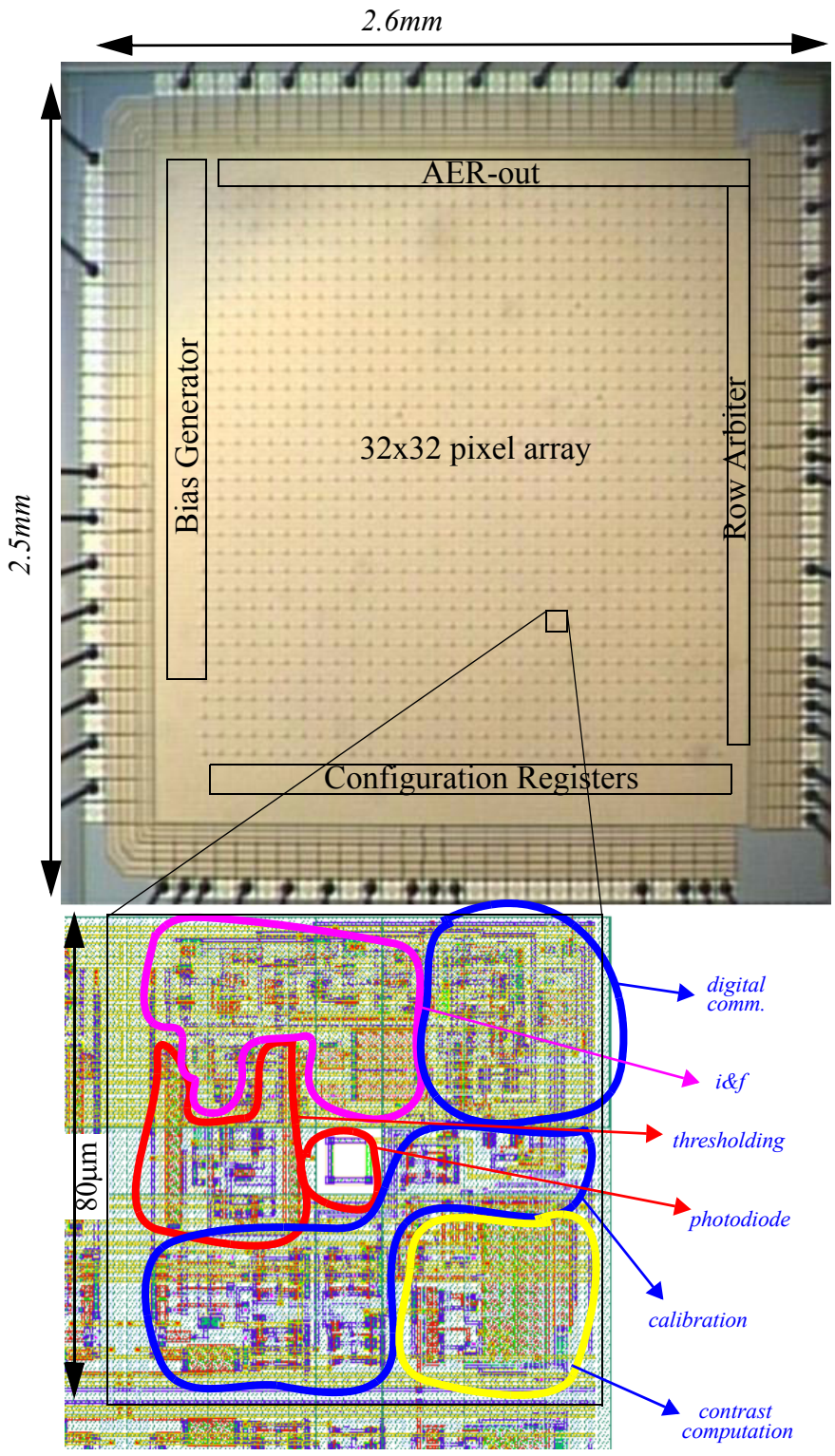


Fig 17: Micro fotografía del dado de 2.5*mm* x 2.6*mm*, y zoom del layout del píxel de 80µ*m* x 80µ*m* , indicando la localización de sus partes principales.

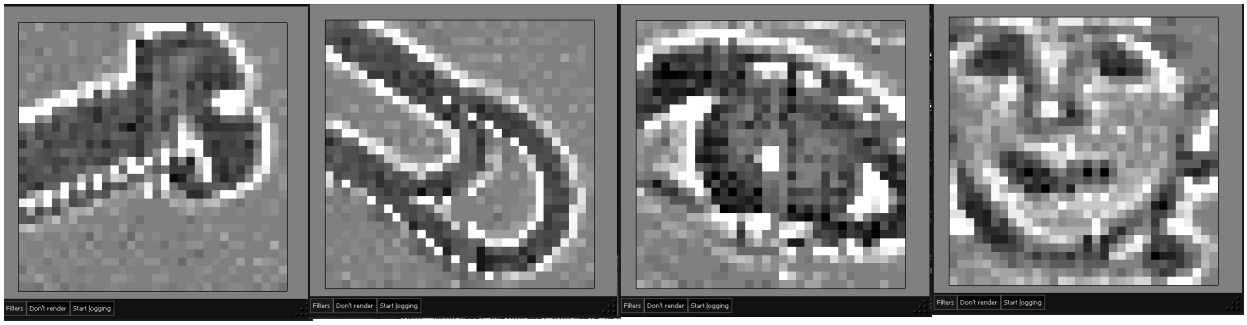


Fig 18:. Elementos naturales. De izquierda a derecha: tornillo, clip, ojo y cara de una niña.

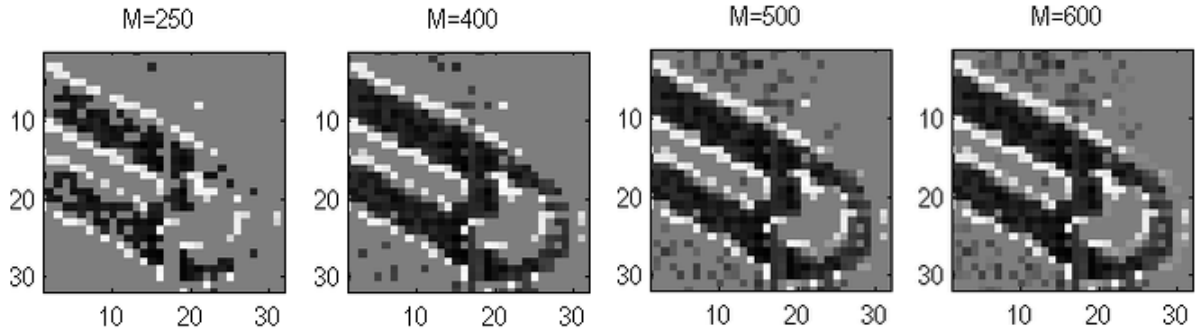


Fig 19: Capturas de un clip en modo TFS con distinto número de eventos capturados, *M*.

latencia (0.1-10*ms*) y elevado rango dinámico (>100dB). En la Figura 18, pueden verse imágenes captadas con la retina. Pese a la baja resolución, muestran el correcto funcionamiento del sensor. Por último, la retina incorpora un modo de funcionamiento TFS opcional. En la Figura 19, se muestra una imagen captada empleando este modo de operación. Se muestra la imagen capturada para distintos números de eventos capturados tras el reset global. Puede apreciarse que con un número de eventos, *M*=250, es posible distinguir la imagen original. El modo TFS es también útil para eliminar el ruido. Nótese que los píxeles con información relevante disparan primero. Si se escoge la duración del frame de forma adecuada, los píxeles que simplemente detectan ruido, no dispararán nunca.

4.5 Conclusiones sobre la Retina AER de Contraste Espacial

Parte del trabajo de la tesis consitió en testar una retina AER de contraste espacial. En esta sección se han descrito la nueva arquitectura del chip y se han resumido los resultados experimentales que se obtuvieron al caracterizar el sensor. La nueva retina hace uso del esquema de calibrado explicado en la Figura 9, obteniendo unos niveles de ruido bastante bajos en comparación con el sensor previo propuesto por Boahen [44]. Otras ventajas y mejoras de la nueva implementación han sido explicadas en este apartado de la tesis.

5 La Retina AER de Contraste Temporal

5.1 Introducción

La detección de movimiento es una de las tareas más importantes que el cerebro humano puede realizar. Gracias a ello, se puede conseguir mucha información relevante sobre nuestro entorno. La detección del contraste posibilita el reconocimiento de estructuras tridimensionales y permite agrupar los objetos que podemos ver. Tal es la importancia de la detección del contraste espacial, que en el cerebro humano hay un canal especializado en ello.

Recientemente, varios sensores específicos de contraste temporal han sido publicados. El primer sensor específico de contraste fue presentado por Mallik y sus colegas [7] en 2005 y estaba basado en el trabajo previo de V. Gruev [48] que describía un sensor óptico basado en las variaciones temporales entre frames consecutivos. La retina de Mallik modificaba el tradicional pixel activo (APS) para detectar un cambio absoluto y cuantizado en iluminación. Era un dispositivo basado en frames con algunas de las ventajas de los sistemas AER. Recientemente, varios sensores de contraste espacial han sido publicados [9]-[24]. No son dispositivos basados en frames y tienen varias ventajas sobre los sistemas convencionales. Sus características más importantes son alto rango dinámico, capacidad de detectar objetos a muy alta velocidad y muy bajo consumo de potencia.

Las retinas AER son inherentemente más rápidas que los dispositivos basados en frames. La principal razón es que pueden detectar los cambios en iluminación de forma continua a lo largo del tiempo y no hay tiempos muertos entre frames donde los cambios que ocurran no pueden ser detectados.

Parte del trabajo de la tesis consistió en la implementación de un sensor AER de contraste espacial. Está basado en el trabajo previo de Lichsteiner y sus colegas [9]. Sus píxeles responden de forma asíncrona a los cambios relativos de intensidad. El sensor directamente codifica los cambios de luminosidad en la escena reduciendo la redundancia en los datos y el consumo de ancho de banda. La nueva retina tiene características mejoradas y píxeles con menor tamaño. Como discutiremos más adelante, es especialmente adecuada para aplicaciones de seguimiento a alta velocidad y vigilancia.

5.2 Arquitectura del Píxel y Layout

La Figura 20 muestra el diagrama de bloques de las distintas partes del píxel. Hay cuatro bloques principales: 1) El fotoreceptor, 2) Dos amplificadores para incrementar la sensibilidad y disminuir el tiempo de respuesta, 3) Un integrador para detectar las variaciones temporales en iluminación, 4) Bloque de comunicación con la periferia.

Los dos amplificadores están conectados a un circuito externo que controla la ganancia de forma automática (AGC). Éste se encarga de fijar la ganancia de los amplificadores a un valor constante que no depende de la iluminación.

A continuación, vamos a describir cada uno de los bloques que forman el píxel y el circuito externo que controla la ganancia de los amplificadores.

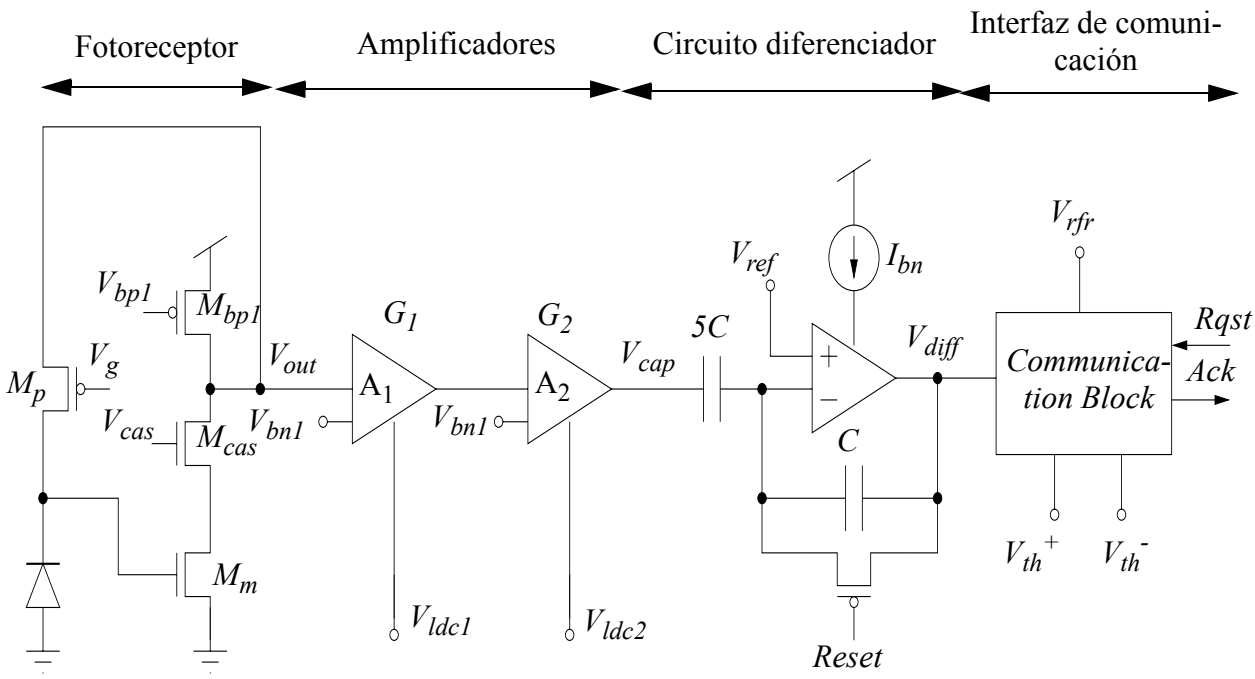


Fig 20: Diagrama de bloques de las distintas partes que componen el píxel.

E. Fotoreceptor

La primera etapa es el fotoreceptor. Este circuito responde logaritmicamente a los cambios en la iluminación local. Hay un offset en la etapa de salida, V_{out}, que será amplificador en la siguiente etapa de amplificación. Esto no es un problema porque será eliminado con el circuito diferenciador, el cual sólo responde a variaciones en la iluminación local. Ajustando, V_g, V_{bp1} y V_{cas}, podemos variar la velocidad de respuesta del fotoreceptor. El voltaje a la salida viene dado por la siguiente expresión,

$$V_{out} = \kappa_p V_g + nU_T \ln\left(\frac{I_{ph}}{I_{sp}}\right) = K + nU_T \ln\left(\frac{I_{ph}}{I_{sp}}\right) \tag{10}$$

F. Amplificadores

Para incrementar la sensibilidad y el tiempo de respuesta de la retina, dos amplificadores se añadieron después del fotoreceptor. Ambos trabajan en la región de saturación. De esta forma, la ganancia total de cada píxel puede ser incrementada con respecto a la del diseño previo propuesto por Lichsteiner y sus colaboradores [9]. En ese diseño, la ganancia total del píxel era igual al cociente entre las dos capacidades del circuito derivador. Para conseguir una ganancia total de 20 con un buen macheo entre píxeles, grandes capacidades eran necesarias. Si usamos amplificadores, podemos decrementar la ganancia del circuito diferenciador. Esto significa que podemos

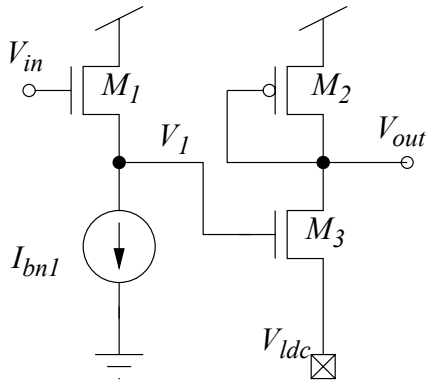


Fig 21: Esquemático de los amplificadores.

reducir el tamaño de las capacidades y conseguir así mejor fill factor y mayor ganancia al mismo tiempo. Mayor ganancia implica mayor velocidad y mayor sensibilidad al contraste. Para aplicaciones de alta velocidad, alta ganancia y sensibilidad son deseables porque los cambios en la iluminación local ocurren muy rápido. Como discutiremos con posterioridad, alta ganancia también implica bajo latencia. El único inconveniente de usar amplificadores es un aumento del consumo.

Si comparamos el nuevo diseño con el de Lichsteiner [9], la total ganancia ha sido incrementada y el área del píxel se redujo a la mitad.

La Figura 21 muestra los esquemáticos de un amplificador. Los transistores M_2 y M_3 trabajan en inversión fuerte. El valor de V_{ldc} es fijado por el circuito externo de control de la ganancia. De esta forma, los niveles de DC a la salida no dependen de la iluminación. M_1 es un seguidor de fuente que introduce un desplazamiento en el voltaje en V_{in} que puede ser ajustado variando V_{bn1}, por tanto $V_1 = V_{in} - \Delta V$. Si M_2 y M_3 están en inversión fuerte,

$$\left(\frac{W_2}{L_2}\right)\beta_{n2}(V_1 - V_{ldc} - V_{Th})^2 = \left(\frac{W_3}{L_3}\right)\beta_{p3}(V_{dd} - V_{out} - |V_{Tp}|)^2 \tag{11}$$

Resolviendo para V_{out}, y sabiendo que $V_1 = V_{in} - \Delta V$,

$$V_{out} = V_{dd} - |V_{Th}| - \sqrt{\frac{(W_4/L_4)\beta_4}{(W_3/L_3)\beta_3}}(V_{in} - \Delta V - V_{Tp} - V_{ldc}) \tag{12}$$

Donde ΔV se fija a un valor constante. Un valor recomendable es $\Delta V \approx 0.7\text{V}$. En cada píxel, hay dos amplificadores (A_1 y A_2), y la ganancia total a la salida de la etapa de amplificación es

$$G_T = G_1 G_2 = \sqrt{\frac{(W_{2_1}/L_{2_1})\beta_{2_1}}{(W_{3_1}/L_{3_1})\beta_{3_1}}}\sqrt{\frac{(W_{2_2}/L_{2_2})\beta_{2_2}}{(W_{3_2}/L_{3_2})\beta_{3_2}}} \tag{13}$$

En nuestro caso particular, ambos amplificadores se diseñaron para tener la misma ganancia $G_1 \approx G_2$ y $\Delta V_{cap} = G_1 G_2 \Delta V_{out}$. La componente DC a la salida es eliminada por el circuito diferenciador. Simplemente hemos de preocuparnos de que la ganancia de los amplificadores no sature la salida de alguno de ellos. El mismatch en la ganancia de los amplificadores es relativamente bajo (comparable al mismatch en los comparadores de la etapa de integración y disparo) y no supone una limitación.

G. Circuito derivador

Esta etapa responde a cambios relativos en el voltaje de la señal de entrada con una ganancia igual a $G_3 = 5C/C$. El mismatch de las capacidades es pequeño. El circuito diferenciador también elimina la componente de DC en la señal de entrada. Podemos expresar las variaciones de voltaje a la salida como

$$\Delta V_{diff} = -G_3 \Delta V_{cap} = -G_1 G_2 G_3 \Delta V_{out} \tag{14}$$

Y de la ecuación (10),

$$\Delta V_{out} = n U_T \ln\left(\frac{I_{ph}(\Delta t + t)}{I_{ph}(t)}\right) \tag{15}$$

La ganancia total a la salida de esta etapa es $G_t = G_1 G_2 G_3$.

H. Circuito de comunicación con la periferia

La Figura 22(a) muestra los esquemáticos del circuito de comunicación con la periferia. Los píxeles pueden detectar el signo del contraste temporal. Para ello, hay dos comparadores con diferentes umbrales ajustables V_{high} y V_{low} que generan eventos positivos o negativos (contraste con signo) cuando V_{diff} está por encima del umbral positivo o por debajo del negativo, respectivamente. También se incorporó un circuito refractario (véase Figura 22(b)) que permite controlar el tiempo entre eventos consecutivos que un píxel individual puede generar. Esto es especialmente útil cuando hay píxel con alta actividad y se desea controlar la actividad máxima en el bus AER. Finalmente, hay dos bloques idénticos de comunicación con la periferia, tal como se ve en la Figura 22(c), los circuitos de comunicación están inspirados en las técnicas de Boahen para mandar eventos al exterior mediante un bus paralelo [45]. Cuando se generan eventos con signo, cada píxel necesita proporcionar dos señales por columna y por signo del evento, *col+* y *col-*. Este concepto ya fue testado e implementado en la retina de contraste espacial [23] y diseños previos [27] que generaban eventos con signo.

5.3 Circuito Externo de Control de la Ganancia

Este bloque está situado en la periferia del chip. En la Figura 23(a), podemos ver los esquemáticos se la circuitería de control de la ganancia de los amplificadores. Sus funciones son mantener constantes los niveles de DC a la salida de los amplificadores de la Figura 21, así como su

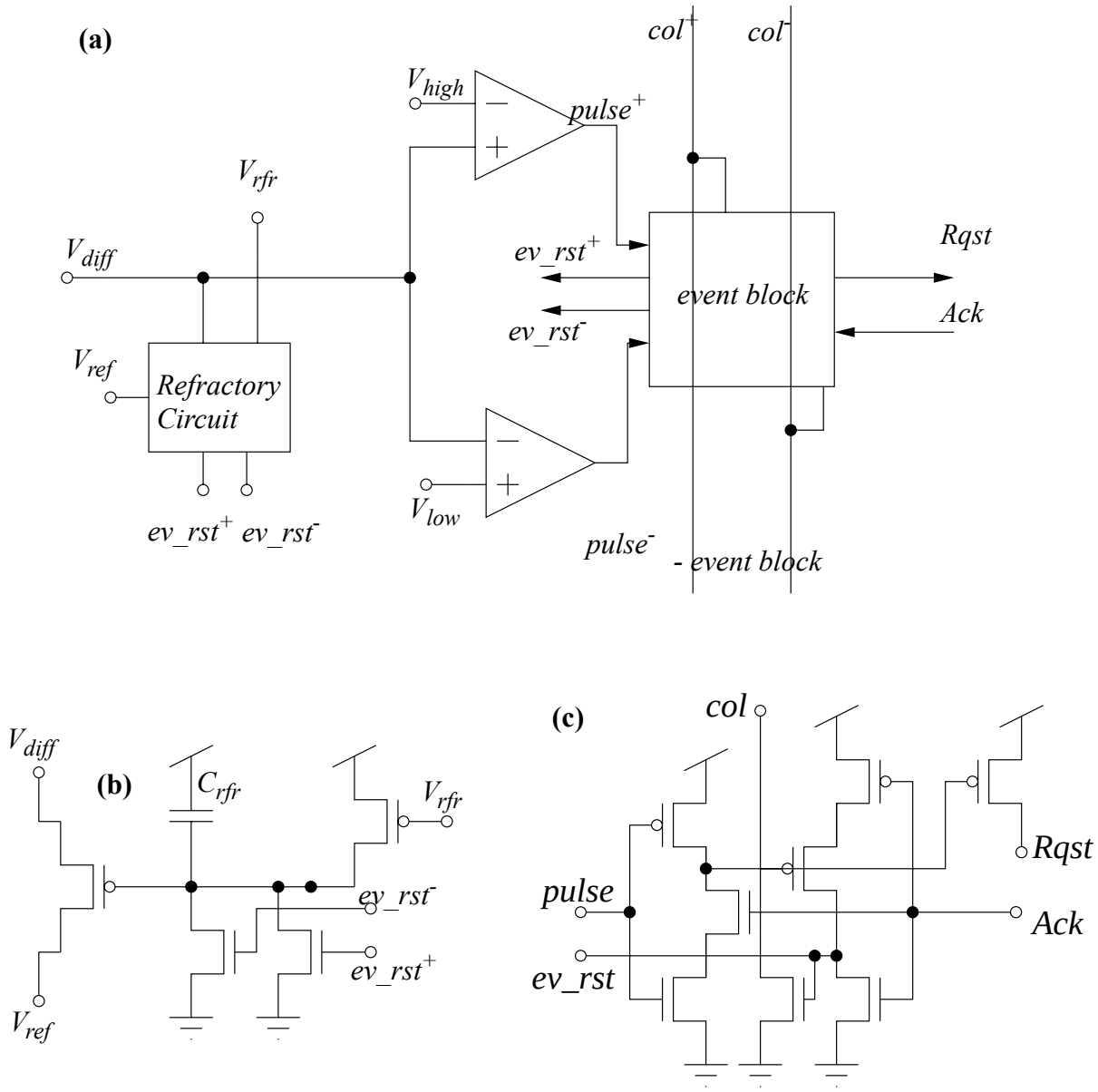


Fig 22: (a) Neurona de integración y disparo con módulo AER de intercomunicación con la periferia. (b) Detalle del circuito refractario. (c) Detalle del bloque AER de comunicación con la periferia dentro del píxel.

ganancia, independientemente de los valores de luminancia del chip. Esto se consigue ajustando los voltajes V_{ldc1} y V_{ldc2} cuando varía la iluminación del chip. El bloque de control de la ganancia tiene amplificadores idénticos a los de la Figura 21. Las salidas de esos amplificadores, A_1 y A_2, son fijadas a unos valores constantes de tensión que podemos ajustar, V_{odc1} y V_{odc2}. Para ello, las salidas de los amplificadores, A_1 y A_2, están conectadas a dos amplificadores operacionales con

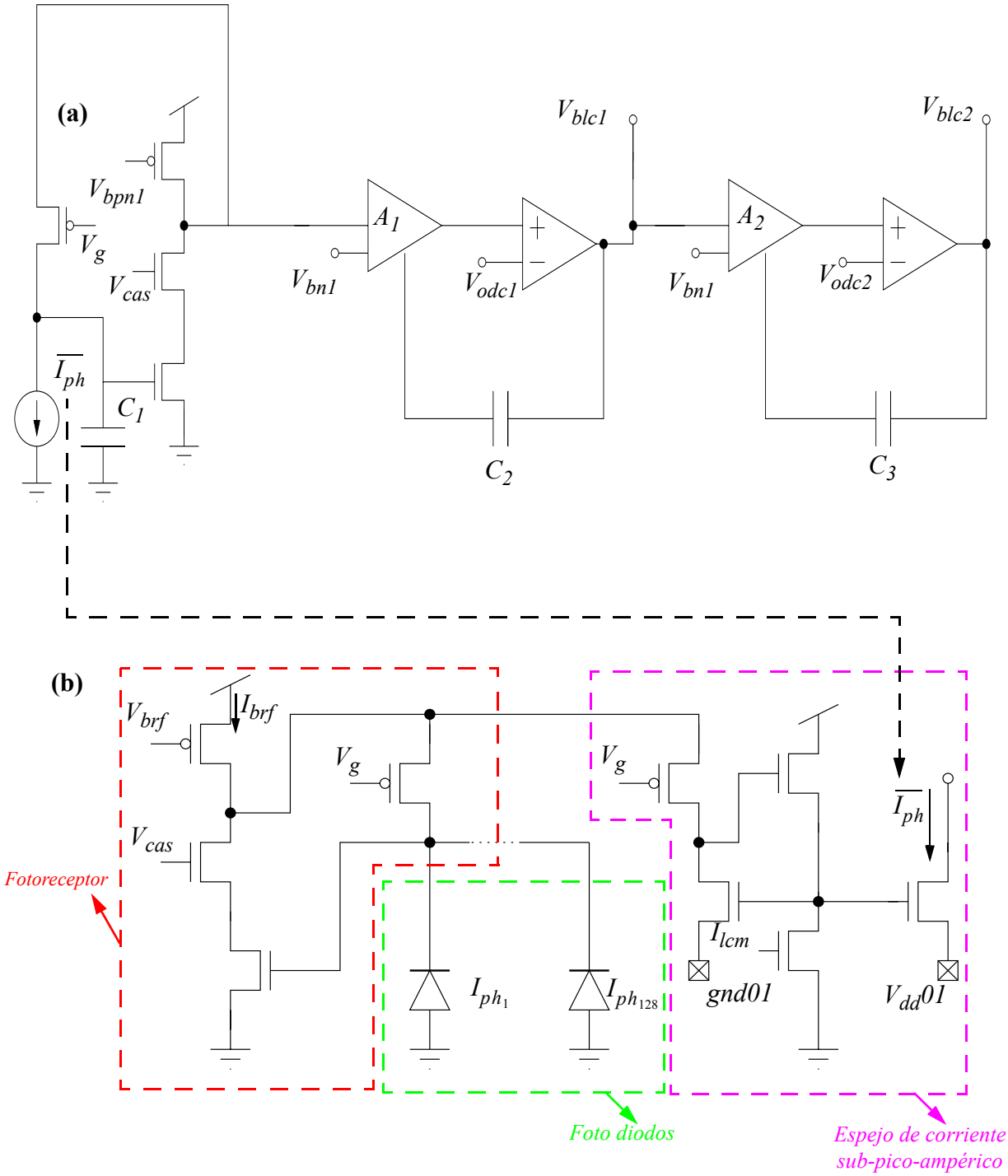


Fig 23: (a) Bloque de control automático de la ganancia. (b) Detalle de la circuitería usada para sensar la iluminación promedio del chip. Se han señalado sus partes principales.

Tabla 3: Especificaciones del Chip.

tecnología	CMOS 0.35μm 4M 2P
tensión de alimentación	3.3V
tamaño del chip	5.54 x 5.69 mm^2
tamaño del array	128 x 128
tamaño del píxel	35.7 x 35.5 μm^2
fill factor	8%
eficiencia cuántica del fotodiodo	0.34 @ 450nm
complejidad del píxel	58 transistors + 3 caps
consumo de corriente	44mA @ 10keps

realimentación negativa que se encargan de fijar los valores de tensión. De esta forma, los valores de V_{ldc1} y V_{ldc2} serán idénticos en la etapa de amplificación y el bloque de control de la ganancia. La corriente de entrada al bloque AGC es la foto corriente promedio sensada por 128 foto diodos situados alrededor de la matriz de píxeles. En la Figura 23(b) podemos ver el circuito que sensa el valor promedio de la foto corriente. Tiene 128 foto diodos y un fotoreceptor idéntico al de la Figura 20. Cada uno está conectado a una capacidad para hacer que el circuito de control de la ganancia sea insensible a cambios muy rápidos en los niveles de luminosidad. Este circuito sólo debe detectar cambios lentos en la iluminación promedio del chip. El valor de esta capacidad fue elegido para que la constante de tiempo fuera del orden de los milisegundos. Este tiempo es suficiente para adaptarnos a un cambio global en la iluminación. Por ejemplo, cuando salimos fuera de una habitación y entramos en una zona muy iluminada con luz solar. Un tiempo de respuesta excesivamente bajo, podría dar lugar a problemas de inestabilidad en el circuito. Finalmente, hay un espejo sub-pico-ampérico [36] que copia la foto corriente sensada.

5.4 Resultados experimentales

Se fabricó un sensor AER de contraste temporal (con signo) de 128 x 128 píxeles. La retina fue diseñada y fabricada en una tecnología CMOS de 0.35 μ m con cuatro metales y doble polisilicio con una tensión de alimentación $V_{DD} = 3.3\,V$. La Tabla 3 resume las especificaciones del chip. La Figura 24 muestra una micro-fotografía del dado, de tamaño $5.689x5.538mm^2$. El chip completo, excepto el anillo de pads, está cubierto con la capa de metal más alta, dejando aberturas encima de los foto diodos. La Figura 24 muestra el layout de un píxel individual resaltando sus partes principales. El layout de cada píxel es una especulación simétrica de los píxeles de su alrededor. De esta forma, las líneas digitales ruidosas son compartidas entre vecinos, al mismo tiempo que las líneas de alimentación y las líneas sensibles al ruido. Al mismo tiempo, las líneas sensibles al ruido se separan de las ruidosas. El área del píxel es $35.5x35.7\mu m^2$, incluyendo el rutado. El píxel está formado por 58 transistores y 3 capacidades (las dos del divisor capacitivo y la del circuito refractario).

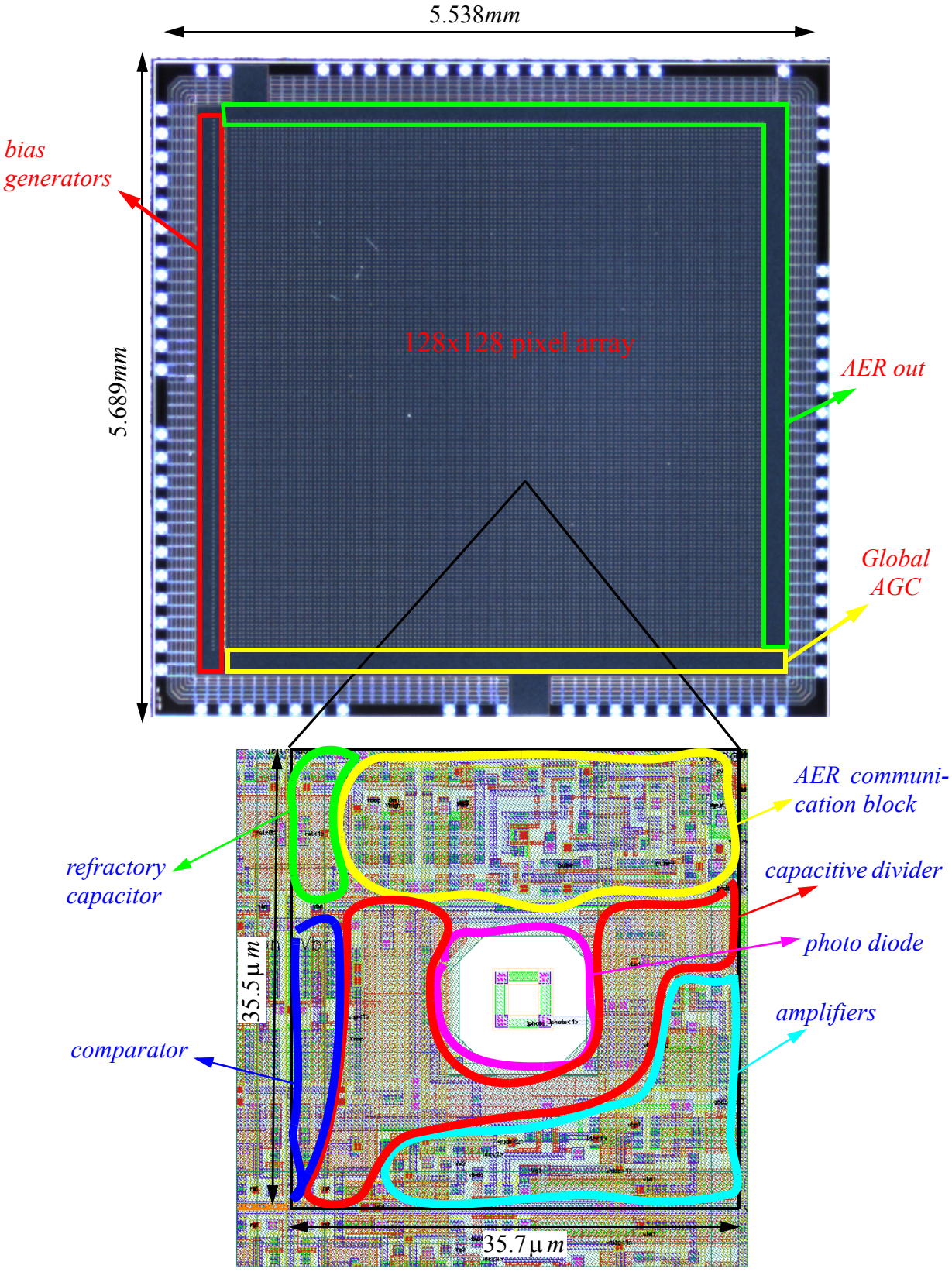
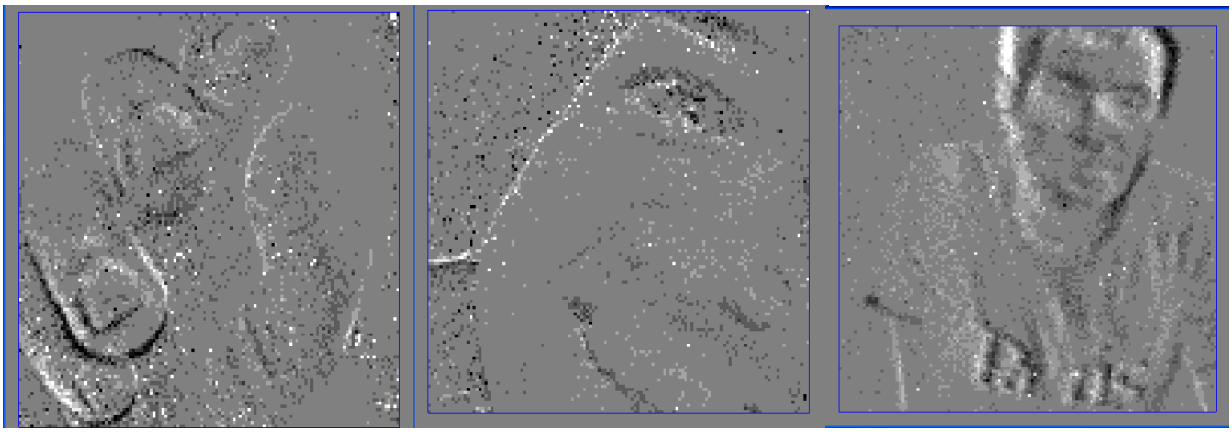


Fig 24: Micro-fotografía del dado 5.538*mm* x 5.689*mm*, y detalle del píxel de 35.7µ*m* x 35.5µ*m*, indicando la localización de sus partes principales.

Tabla 4.: Especificaciones del Sensor de Contraste Temporal.

Funcionalidad	Dependencia con el nivel de luminsidad	Latencia/ Ancho de banda	Rango Dinámico	FPN	Fill Factor	Tamaño del píxel μm^2	Proceso de Fabricación	Potencia
Contraste Temporal a número de eventos	NO	$3.6\mu s$ / $8KHz$	>100dB	2.8%	8%	35x35	$0.35\mu m$ 4M 2P	132-185mW



7.8*Keps* in 32.8*ms* 20*Keps* in 30*ms* 1.8*Keps* in 32.8*ms*

Fig 25: Algunas capturas de escenas naturales. De izquierda a derecha: Mano, cara y hombre. La objetivo usado era de *16mm.*

La retina se sometió a múltiples experimentos y pruebas en condiciones muy diversas de iluminación para evaluar sus prestaciones. La Tabla 4 resume sus características más significativas. Entre éstas cabe destacar las siguientes: muy baja latencia ($3.6\mu s$ @$25Klux$), buen fill factor (8%), elevado ancho de banda (>$8KHz$) y elevado rango dinámico (>100dB). En la Figura 25 , pueden verse imágenes captadas con la retina.

5.5 Conclusiones sobre la Retina AER de Contraste Temporal

Parte del trabajo de la tesis consitió en testar una retina AER de contraste temporal. En esta sección se han descrito la nueva arquitectura del chip y se han resumido los resultados experimentales que se obtuvieron al caracterizar el sensor. La retina hace uso de una etapa de amplificación tras el fotoreceptor logarítmico. De esta forma, se aumenta la sensibilidad, el ancho de banda y se reduce la latencia y el tamaño del píxel, con respecto al diseño previo propuesto por Lichsteiner y sus colegas [9].

6 Conclusiones y Líneas Futuras de Investigación

En el presente documento se han descrito dos sensores AER distintos capaz de detectar el contraste espacial y temporal, además del signo. El primero de ellos es un sensor de contraste espacial que solventa las limitaciones de otros sensores del mismo tipo publicados previamente. Dicho sensor, está basado en el trabajo previo de Boahen [21]-[22] y resuelve alguna de sus limitaciones. Se elimina el consumo innecesario de ancho de banda cuando no hay contraste, proporciona salida unipolar, el cómputo del contraste no depende de la iluminación del chip, se reduce el consumo de potencia y el FPN se ha reducido notablemente. El mismatch era una de las limitaciones que ha afectado tradicionalmente a los sistemas neuromórficos y en particular a las retinas de contraste espacial basadas en el cómputo del contraste mediante el cociente de intensidades. Por esta razón, se desarrolló un nuevo sistemas de calibrado compacto para redes neuromórficas [13]. El nuevo sistema tiene un tamaño menor que otros implementados previamente [14]-[27]. Además, permite escalar las corrientes de calibrado sin que el calibrado se degrade. El circuito de calibrado se probó de forma individual y los resultados obtenidos fueron satisfactorios. Posteriormente, se incorporó a la retina de contraste espacial para solventar los problemas de mismatch, reduciendo el valor del FPN de forma notable.

El segundo de los sensores presentados en la tesis es un sensor de contraste temporal. Esta retina es especialmente útil para el seguimiento de objetos a alta velocidad y aplicaciones de vigilancia. Está basado en el sensor propuesto por Lichsteiner y sus colegas [10] y mejora algunas de sus características como el ancho de banda, el tiempo de latencia y el fill factor. Para ello, hace uso de amplificadores tras el fotoreceptor. De esta forma, se aumenta la ganancia y se reduce el tamaño del píxel, ya que la ganancia del circuito derivador no necesita ser tan alta, y, por tanto, las capacidades del mismo pueden hacerse más pequeñas. En el circuito propuesto por Lichsteiner la ganancia del píxel estaba basada únicamente en el cociente entre capacidades del circuito diferenciador. Por esta razón, grandes capacidades eran requeridas para obtener una sensibilidad al contraste aceptable.

En cuanto al trabajo futuro, después de haber testado exhaustivamente los chips fabricados, se pueden proponer varias mejoras. En primer lugar, el calibrado de la retina podría hacerse en el propio chip de forma automática. Hasta ahora el calibrado se había hecho off-line según el procedimiento descrito en [13]. Sería perfectamente factible, implementar un controlador en el propio chip que lo realice de forma automático. Simplemente, sería necesario iluminar el chip de forma uniforme e implementar un algoritmo que escoja las palabras de calibrado que minimizan las frecuencias de salida de cada píxel.

En cuanto a la retina de contraste espacial, el fill factor obtenido (2%) es relativamente bajo. Podría mejorarse si se reduce el número de bits de las palabras de calibrado (3 ó 4 bits sería suficiente) porque el valor del FPN obtenido (0.6%) es muy bajo. Ello posibilitaría implementar una retina de contraste espacial de mayor resolución.

Finalmente, la retina de contraste temporal hace uso de amplificadores en inversión fuerte. Ello hace que el consumo sea ligeramente superior al del diseño de Lichsteiner [10]. En el futuro, sería posible implementar un diseño similar con transistores operando en inversión débil para reducir el consumo del chip.

Bibliografía

[1] M. Sivilotti, "Wiring considerations in analog VLSI systems with application to field-programmable networks," *Ph.D. dissertation, Calif. Inst.Technol.*, Pasadena, 1991.

[2] M. Mahowald, "VLSI analogs of neural visual processing: a synthesis of form and function" Ph.D. dissertation, California Institute of Technology, Pasadena, 1991.

[3] T. Delbrück, Library essentials, Analog VLSI and neural systems by Carver Mead, Addison Wesley, 1986. *The Neuromorphic Engineer,* 1(1):11, 2004. http://ine-web.org/research/newsletters/index.html.

[4] Kwabena A. Boahen, "Retinomorphic Vision Systems: Reverse Engineering the Vertebrate Retina", Ph. D. dissertation, California Institute of Technology, Pasadena, 1996.

[5] G. M. Shepherd, *The Synaptic Organization of the Brain,* 3rd ed. Oxford, U.K.: Oxford University Press, 1990.

[6] M. Mahowald. *An Analog VLSI System for Stereoscopic Vision,* Kluwer, Boston, MA, 1994.

[7] U. Mallik, et al., *"Temporal Change Threshold Detection Imager",* in *ISDD Dig. of Tech. Papers, San Francisco, 2005, pp362-363.*

[8] Y. Chi et al., *"CMOS Camera with In-pixel Temporal Change Detection and ADC",* IEEE Journal of Solid State Circuits, vol. 42, pp, 2187-2196, OCT 2007.

[9] Patrick Lichsteiner et all, *"A 128x128 120dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor"*, IEEE Journal of Solid State Circuits, vol. 43, pp. 566-576, 2008.

[10] P. Lichsteiner, et al., "*A 128x128 120dB 30mW Asynchronous Vision Sensor that Responds to Relative Intensity Change*", in ISSCC Dig. of Tech. Papers, San Francisco, 2006, pp. 508-509 (27.9)

[11] S. Thorpe, D. Fize, C. Marlot, *"Speed of Processing in the Human Visual System",* Nature 381:520-2, 1996.

[12] X. G. Qi, X.; Harris J., *"A Time-to-first-spike CMOS imager",* in Proc. of the 2004 IEEE International Symposium on Circuits and Systems (ISCAS04), Vancouver, Canada, 2004, pp. 824-827.

[13] J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, *"A Calibration Technique for Very Low Current and Compact Tunable Neuromorphic Cells. Application to*

5-bit 20nA DACs", IEEE Transactions on Circuits and Systems, Part-II: Brief Papers, vol.55, No. 6, pp. 522-526, June 2008.

[14] J. Costas-Santos, T. Serrano-Gotarredona, R. Serrano-Gotarredona and B. Linares-Barranco, *"A Spatial Contrast Retina with On-chip Calibration for Neuromorphic Spike-Based AER Vision Systems",* IEEE Trans. Circuits and Systems, Part-I: Regular Papers, vol. 54, No. 7, pp. 1444-1458, July 2007.

[15] M. Barbaro, P.-Y. Burgi, A. Mortara, P. Nussbaum, and F. Heitger. *A 100100 Pixel Silicon Retina for Gradient Extraction with Steering Filter Capabilities and Temporal Output Coding.* IEEE Journal of Solid-State Circuits, 37(2):160-172, Feb. 2002.

[16] P. F. Ruedi, et al., *"A 128x128 Pixel 120-dB Dynamic-range Vision-Sensor Chip for Image Contrast and Orientation Extraction",* IEEE Journal of Solid-State Circuits, 38:2325-2333, Dec. 2003.

[17] P. F. Ruedi, et al., *"An SoC Combining a 132dB QVGA Pixel Array and a 32b DSP/MCU Processor for Vision Applications",* in IEEE ISSCC Dig. of Tech. Papers, 2009, pp.46-47, 47a.

[18] Honghao Ji and Pamela A. Abshire, "*Fundamentals of Silicon-Based Phototransduction"* in *"CMOS Imagers from Phototransduction to Image Processing"*, Kluwer Academic Publishers.

[19] E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, *"A Biomorphic Digital Image Sensor",* IEEE J. Solid-State Circuits, vol.38, pp.281-294, 2003.

[20] S. Chen, and A. Bermak, *"Arbitrated Time-to-First Spike CMOS Image Sensor with On-Chip Histogram Equalization",* IEEE Transactions VLSI Systems, vol. 15, No. 3, pp 346-357, March 2007.

[21] K. A. Zaghloul and K. Boahen, *"Optic Nerve Signals in a Neuromorphic Chip: Part 1",* IEEE Transactions on Biomedical Engineering, vol 51, pp. 657-666, 2004.

[22] K. A. Zaghloul and K. Boahen, *"Optic Nerve Signals in a Neuromorphic Chip: Part 2",* IEEE Transactions on Biomedical Engineering, vol 51, pp. 667-675, 2004.

[23] J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, *"A 5-Decade Dynamic Range Ambient-Light-Independent Calibrated Signed-Spatial-Contrast AER Retina with 0.1ms Latency and Optional Time-to-First-Spike Mode",* Transactions on Circuits and Systems, Part I, Under revision.

[24] C. Posch, et al., *"High DR, Low Date-rate Imaging Based on an Asynchronous, Self-triggered Address-event PWM Array with Pixel-level Temporal Redundancy Suppression"*, in 2010, in ISCAS 2010.

[25] R. R. Harrison, J.A. Bragg, P. Hasler, B.A. Minch, and S.P. DeWeerth, *"A CMOS programmable analog memory-cell array using floating-gate circuits," IEEE Trans. on*

Circuits and Systems, Part II, vol. 48, No. 1, pp. 4-11, Jan. 2001.

[26] Y. L. Wong, M. H. Cohen, and P. A. Abshire, *"128x128 floating gate imager with self-adapting fixed pattern noise reduction," Proc. of the IEEE 2005 Int. Symp. on Circuits and Systems* (ISCAS'05), vol. 5, pp. 5314-5317, 2005.

[27] R. Serrano-Gotarredona, T. Serrano-Gotarredona, A. Acosta-Jimenez, and B. Linares-Barranco, *"A Neuromorphic Cortical-Layer Microchip for Spike-Based Event Processing Vision Systems," IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 53, No. 12, pp. 2548-2566, Dec. 2006.

[28] B. Linares-Barranco, T. Serrano-Gotarredona, and R. Serrano-Gotarredona, *"Compact Low-Power Calibration Mini-DACs for Neural Massive Arrays with Programmable Weights," IEEE Trans. on Neural Networks*, vol. 14, No. 5, pp. 1207-1216, September 2003.

[29] C. A. Laber, C. F. Rahim, S. F. Dreyer, G. T. Uehara, P. T. Kwok, and P. R. Gray, *"Design Considerations for a high-Performance 3mm CMOS Analog Standard-Cell Library," IEEE J. Solid-State Circuits*, vol. SC-22, No. 2, pp. 181-189, April 1987.

[30] T. Delbrück and A. Van Shaik, *"Bias Current Generators with Wide Dynamic Range," Int. Journal of Analog Integrated Circuits and Signal Processing*, No. 43, pp. 247-268, June 2005.

[31] R. R. Harrison, J.A. Bragg, P. Hasler, B.A. Minch, and S.P. DeWeerth, *"A CMOS programmable analog memory-cell array using floating-gate circuits," IEEE Trans. on Circuits and Systems, Part II*, vol. 48, No. 1, pp. 4-11, Jan. 2001.

[32] K. Bult and G.J.G.M. Geelen, *"An inherently linear and compact MOST-only current division technique," IEEE J. Solid-State Circuits,* vol. 27, No. 12, pp. 1730-1735, Dec. 1992.

[33] T. Serrano-Gotarredona and B. Linares-Barranco, *"CMOS Mismatch Model valid from Weak to Strong Inversion"*, *Proc. of the 2003 European Solid State Circuits Conference*, (ESSCIRC'03), pp. 627-630, September 2003.

[34] P.R. Gray, P.J. Hurst, S.H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits,* 4th Edition, John Wiley, 2001.

[35] Rafael Serrano-Gotarredona, Luis Camuñas-Mesa, Teresa Serrano-Gotarredona, Juan A. Leñero-Bardallo and Bernabé Linares-Barranco, *"The Stochastic I-Pot: A Circuit Block for Programming Bias Currents", IEEE Transaction on Circuits and Systems-II: Brief Papers,* vol 19, No. 7, pp. 1196-1219, July 2008.

[36] B. Linares-Barranco and T. Serrano-Gotarredona, *"On the Design and Characterization of Femtoampere Current-Mode Circuits", IEEE Journal of Solid-State Circuits,* vol. 38, No. 8, pp. 1353-1363, August 2003.

[37] M. Azadmehr, J. Abrahamsen, and P. Häfliger, *"A Foveated AER Imager Chip",* in *Proc. of*

the IEEE International Symposium on Circuits and Systems (ISCAS2005), pp. 2751-2754, Kobe, Japan, 2005.

[38] R.J. Vogelstein, U. Mallik, E. Culurciello, R. Etienne-Cummings, and G. Cauwenberghs, "Spatial acuity modulation of an address-event imager," *Proc. of the 2004 11th IEEE International Conference on Electronics, Circuits and Systems (*ICECS 2004), pp. 207-210, Dec. 2004.

[39] J. Kramer, *"An Integrated Optical Transient Sensor", IEEE Transactions on Circuits and Systems, Part II: Analog and Digital Signal Processing,* v. 49, No. 9, pp. 612-628, Sep. 2002.

[40] M. Arias-Estrada, D. Poussart, and M. Tremblay, *"Motion Vision Sensor Architecture with Asynchronous Self-Signaling Pixels", Proceedings of the 7th International Workshop on Computer Architecture for Machine Perception (CAMP07),* pp. 75-83, 1997.

[41] K. Boahen, *"Retinomorphic Chips that see Quadruple Images", Proceedings of International Conference of Microelectronics for Neural, Fuzzy and Bio-inspired Systems* (Microneuro99), pp. 12-20, Granada, Spain, 1999.

[42] S. Thorpe, et al., *"SpikeNet: Real-time Visual Processing in the Human Visual System", Neurocomputing 58-60: 857-64, 2004.*

[43] K. Boahen and A. Andreou, *"A Contrast-Sensitive Retina with Reciprocal Synapses",* in J. E. Moddy (Ed.), *Advances in neural information processing,* vol. 4, pp- 764-772, San Mateo CA, 1992. Morgan Kaufman.

[44] A. G. Andreou and K. Boahen, *"Translinear Circuits in Subthreshold CMOS", Analog Integrated Circuits and Signal Processing,* Kluwer, no. 9 , pp. 141-166, Apr. 1996.

[45] K. Boahen, *"Point-to-Point Connectivity Between Neuromorphic Chips Using Address Events", IEEE Transactions on Circuits and Systems Part-II,* vol. 47, No. 5, pp. 416-434, May 2000.

[46] R. Berner, T. Delbruck, A. Civit-Balcells and A. Linares-Barranco,"A 5 Meps $100 USB2.0 Address-Event Monitor-Sequencer Interface", *IEEE International Symposium on Circuits and Systems,2007, ISCAS 2007.*

[47] jAER Open Source Project 2007 [Online]. Available: http://jaer.wiki.sourceforge.net.

[48] V. Gruev and Etienne-Cummings, R. A. pipelined temporal difference imager. *IEEE J. Solid State Circuits,* 39:538-543, 2004.

[49] M. A. Mahowald, *"VLSI Analogs of Neuronal Visual Processing: A Synthesis of Form and Function",* PhD, Computation adn Neural Systems, Caltech, Pasadena, California, 1992.

[50] D. A. Baylor, T. D. Lamb, and K.W. Yau, *"Responses of Retinal Rods to Single Photons",* J. Physiology., 288:613-634, 1979.

[51] D. A. Baylor, B. J. Nunn, and J. L. Schnapf. *"The Photo Current, Noise, and Spectral Sen-*

sitivity of Rods of the Monkey Macaca Fascicularis". J. Physiology, 357:575-607, 1984.

ANEXOS

University of Sevilla
Department of Electronics and Electromagnetism

Ph.D. Dissertation

Study, Design, Implementation, and Test of VLSI Retinae Sensitive to Spatial and Temporal Contrast

By Juan Antonio Leñero Bardallo

Advisors:

Dr. Bernabé Linares Barranco

Dr. Teresa Serrano Gotarredona

Seville, 2010

To my parents

ACKNOWLEDGEMENTS

This dissertation is the result of my work at the Sevilla Microelectronics Institute during four years. I would like to be thankful to my thesis advisors, Bernabé Linares and Teresa Serrano. They have helped me and they have conducted my work during all the time I have been elaborating my thesis. Without their help, this document would not have been possible.

All my work at the Sevilla Microelectronics Institute during this time has been funded by a JAE scholarship. Thanks to this program, I also had the opportunity of spending two months in the University of Oslo, researching at the Department of Informatics with Dr. Philipp Häfliger.

I also would like to mention my parents and brothers for their support and care during all this time.

This work was made with the collaboration of several people. I have to mention the ATC group of the University of Sevilla for their useful boards and software that I used to test the sensors, Philipp Häfliger for the supplied Caviar board and the lens mount, Tobi Delbrück for the jAER software and the supplied temporal contrast sensor, and the anonymous reviewers that have contributed to improved the quality of this document and our publications.

Finally, I have to thank my colleagues at the Sevilla Microelectronics Institute. I have learnt a lot of valuable skills from them. With their company, all this time in Sevilla has been really enjoyable.

OUTLINE

Nowadays, vision systems are commonplace and widely used to perform plenty of different tasks. For example, we are recorded when we park our car in a car place or when we walk though an airport. Everybody is used to such kind of devices. There is also an huge potential market, specially related to cell phones. Almost all of them have cameras nowadays. All these imagers are based on frames (frame-based devices). With regard to artificial vision, we usually speak about machine vision systems (systems made up by cameras and actuators) that also perform specific tasks like surveillance or remote control. However, such devices work in very controlled environments in a repetitive way.

Frame-based devices have inherent limitations. Their bandwidth is limited by the sample rate. They are always transmitting data periodically that can be very redundant if there is no changes in the scene (frames are always transmitted if the system is working). They do not process the information and a lot of redundant data is involved. Their dynamic range is reduced because AGC (Automatic Gain Control) is global to the pixel matrix. Each pixel has the same gain and the same integration time. The capacity to integrate charge is also limited. In unsupervised environments, bandwidth and dynamic range can limit the performance of the whole vision system.

On the other hand, biological systems are frameless. They are massively parallel and data driven. Such systems are structured in layers. Each layer performs one specific task and sends the relevant information to the next layers. Eyes are optimized to extract the relevant information from the visual scene. If we compare conventional systems to bio-inspired ones, the advantages of the second ones are obvious.

Neuromorphic engineering is based on the belief that emulating such biological systems, we could develop electronic devices with similar features. The rise of VLSI systems leads to systems with pixels with an acceptable fill factor and level of complexity.

In this thesis, we will describe two AER (Address Event Representation) vision systems that imitate two specific retinal functionalities. The first one detects the spatial contrast. The second one is sensitive to temporal contrast. Both of them are spike-based devices and they have two separate channels to process the negative and positive contrast (signed output). Detection of spatial and temporal contrast are two very important tasks that the human retina can perform. With the information of the spatial contrast, we can detect edges and shapes. With the detection of the temporal contrast, we can detect movement, relative depths and group objects.

The first sensor (described in chapter 4) overcomes some of the drawbacks of prior AER spatial contrast imagers. It uses a new calibration system to reduce mismatch, it also has signed

output, and a new thresholding mechanism that allows removal of residual mismatch after calibration. It also has an optional TFS (time-to-first spike) mode that combines the advantages of frame-based systems and AER ones. Its main features are very low FPN (0.6% of the maximum possible output frequency when the sensor detects spatial contrast), good response time (0.1-10*ms*), high sensitivity and very low power consumption (0.66-6.6*mW*). A test prototype of 32 x 32 pixels was fabricated.

The second sensor (described in chapter 5) is specially suitable for high speed applications. It has very low latency (down to $3.5\mu s$) and high pixel bandwidth (up to *8KHz*). It uses amplifiers after the logarithmic photoreceptor to reduce the retina response time and increase its sensitivity to temporal contrast. This prototype has a fill factor of 8% and has 128 x 128 pixels. It also has two independent channels to process positive and negative events independently.

The thesis is structured into 6 chapters. Chapter 1 describes current trends in imager design and compares the features of conventional imagers to the ones of bio-inspired sensors. It also explains the state of the art and establishes a comparison between the more relevant AER imagers presented during the last few years. Chapter 2 presents the *Stochastic I-Pot*, a powerful circuit to generate precisely programmable bias currents. Chapter 3 explains thoroughly the new calibration system and its possible applications to reduce mismatch effects in large neuromorphic arrays. Chapter 4 presents the new AER spatial contrast vision sensor. The complete circuitry is explained. Detailed experimental results are also provided. Chapter 5 describes the new AER temporal contrast vision sensor. Circuitry and experimental results are displayed. Finally, in chapter 6, we draw some conclusions about the present and the possible future work.

CHAPTER 1

Introduction

1.1 Conventional Vision Systems Based on Frames

In the last ten years, sales of digital cameras have grown incredibly. The huge potential market and the raise of cell phones has driven a very fast development in electronic imager design. This has also led to improved and cheaper imagers, with very small pixels. This tendency is continuing. Nowadays, we usually speak about the number of megapixels as a figure of merit.

All these conventional vision systems are based on frames [52]. A frame is an 2D matrix with elements that represent the intensities and color of the visual scene under observation during a period of time. Video data consists of a stream of frames taken with a fixed frequency. Frame-based systems are commonplace and sometimes when we speak about vision systems they are taken for granted. Today's CCDs and CMOS imagers are frame-based devices.

The main advantage of these devices is that they use simple and small size pixels. A small pixel implies high resolution, large fill factor and low cost. Such advantages are inherent to frame-based devices. Another point to consider is that frames are compatible with all the commercial frame-based displays.

However, frames have some drawbacks to take into account. Pixel bandwidth is limited by the frame rate. In conventional systems, there is a trade-off among system bandwidth, frame rate, sensor sensitivity, and the scan period. The frame rate ($1/T_{frame}$) is the number of frames transmitted per unit of time. System bandwidth is the maximum amount of data that our system can process per unit of time. The exposition time, $T_{expositon}$, is the amount of time that pixels are collecting charge to generate one frame ($T_{\exp osition} \leq T_{frame}$) and we usually can adjust it. Sensitivity is the minimum detectable illumination level during the scan period. Frame-based systems accumulate charge (in a capacitor or a photogate) during the scan period. After this, charge at the photogate or the capacitor is read and resetted. If the scan period is too low, the amount of charge integrated will be very low. For this reason, typically conventional systems usually need to work in environments with high and controled illumination conditions to achieve high speed without loosing sensitivity. If we increase frame rate, the amount of output data will be increased. We cannot increase the frame rate arbitrarily because system bandwidth limits the amount of data that the

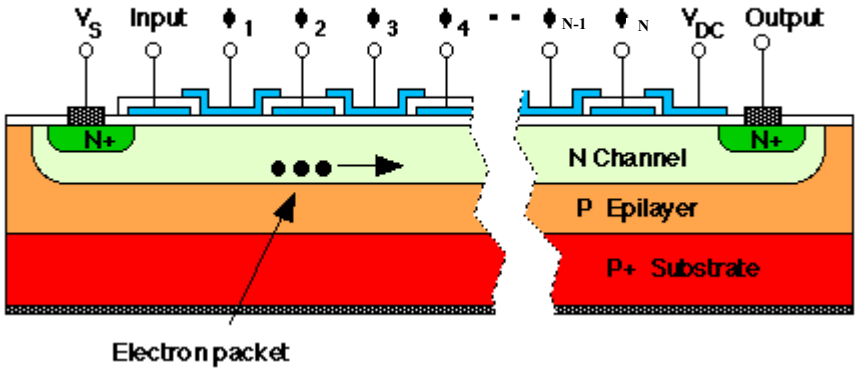


Fig 1.1: Structure of a CCD device with *N* photogates. We can notice how the charge packet is transferred from one side to the another one.

sensor can process. Temporal changes in the scene cannot be detected in the time between two different frames.

Another disadvantage is redundancy. Pixels are continuously scanning and sending information, even if there is no new information to transmit. This leads to a huge load of redundant data. Finally, AGC (Automatic Gain Control) is usually global in frame-based systems. Thus, gain control is not properly adjusted when there is intrascene variations of illumination. For this reason, it is very important to create a good scene illumination in machine vision. In conclusion, limited bandwidth, identical gain for all the pixels and redundancy are assumed for conventional vision systems.

1.1.1 Charge-Coupled Devices

During several years the predominant technology for optical imaging applications has been the charge-coupled device (CCD) [18]. CCD technology revolutionized the field of digital imaging by enabling diverse applications in consumer electronics, scientific imaging and computer vision. Its advantages are high sensitivity, high resolution, large dynamic range, and large array size. CCD is a very mature and developed technology. CCD imagers provide images with excellent quality and are still the dominant imaging technology.

A Charge-Coupled Device (CCD) is an image sensor (see Figure 1.1), consisting of an integrated circuit containing an array of linked, or coupled, MOS capacitors sensitive to light. MOS capacitors (also called photogates) are operated in deep depletion on a continuous insulator layer over the semiconductor substrate. Nowadays, pinned photodiodes are being used instead photogates. Under the control of an external circuit, each capacitor can transfer its electric charge to one of its neighbors as can be seen in Figure 1.1.

There are three conventional architectures: full-frame, frame-transfer and interline. They basically differ on their approach to the problem of shuttering or how the charge is transferred from the photogates.

In a full-frame device, all of the image area is active and there is no electronic shutter. A mechanical shutter must be added to this type of sensor or the image will smear (accumulated charge loss) as the device is clocked or read out. It is a useful choice when light has to be collected accurately.

Frame-transfer architectures are made with an opaque mask (typically aluminium). Half of the pixel area is covered with this mask. The charge is quickly transferred from the photogates to this opaque region with little smear percent. Once the charge has been transferred, it can be read out slowly and the active area can be integrating a new frame at the same time. The main drawback of this architecture is that it requires twice the silicon area than a full-frame device.

In the interline architecture there is a mask over each column of the image sensor, so only one pixel shift has to occur to transfer charge from the image area to the storage area. The disadvantage of the scheme is that fill factor is reduced 50%. Adding special microlenses over the opaque mask, light can be directed away of the opaque region to the active area increasing the fill factor. With this approach, charge can be transferred in the order of microseconds with a very low smear. This is the dominant architecture that we can find in consumer CCD cameras.

The choice of architecture comes down to one of utility. If the application cannot tolerate an expensive, failure-prone, power-intensive mechanical shutter, an interline device is the right choice. Consumer snap-shot cameras have used interline devices. On the other hand, for those applications that require the best possible light collection and issues of money, power and time are less important, the full-frame device is the right choice. Astronomers tend to prefer full-frame devices. The frame-transfer falls in between and was a common choice before the fill-factor issue of interline devices was addressed. Today, frame-transfer is usually chosen when an interline architecture is not available, such as in a back-illuminated device.

1.1.2 The CMOS Sensor

CMOS sensors, also known as Active Pixel Sensors (APS), have been studied since the 1980s as an alternative to CCD technology. APS technology is now beginning to enable new applications in digital imaging by offering improved performance relative to CCD technology in the areas of low cost, low-power operation, high speed and ease of integration.

An APS is an imager in which every pixel includes at least one active transistor. Transistors may operate as both amplifier and buffer to isolate photogenerated charge from the large capacitance of the common output line. Any photodetector can be used for this approach: photo diodes or photogates [18]. APS pixels usually operate in charge integration mode, but may operate with photo transistors as well. Pixels are implemented with circuitry shared by pixels in the same row, and other circuitry by the same column. At the end of each column there is an integrated amplifier to read out the selected pixel.

Figure 1.2 shows an APS with a photo diode. Light causes an accumulation of charge on the capacitance of the photo diode photo current that depends on light intensity. Transistor, M_1, is a switch that resets the device. M_2 is a source follower that allows to observe the pixel voltage without removing the accumulated charge. M_3 acts as a switch to select one single pixel of one column to be read.

The primary motivation for developing CMOS-based imagers is the ease of integration and the ability to incorporate signal conditioning circuitry and even image processing circuitry into the same device. APS systems offer compact size, light weight, low cost, and low power consumption. Despite these advantages, CCD still remains the dominant imaging technology [18]. The main reason is that CCD is a very mature technology with specialized and optimized fabrication processes that provide images with excellent quality, higher sensitivity and higher dynamic range.

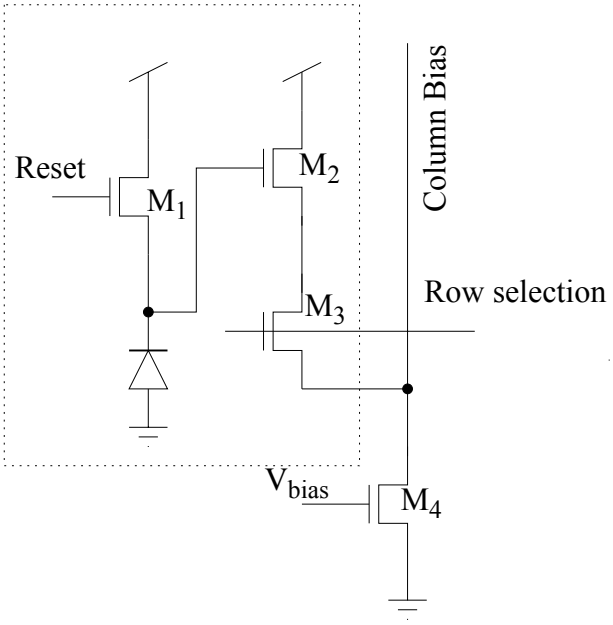


Fig 1.2: APS photocircuit using a photo diode. The components inside the dashed-line box constitute the pixel [18].

1.1.3 Color Cameras

Digital color cameras generally use a Bayer mask over the CCD or the APS as is depicted in Figure 1.3. Each square of four pixels has one filter red, one blue, and two green (the human eye is more sensitive to green than either red or blue). The result of this is that luminance information is collected at every pixel, but the color resolution is lower than the luminance resolution.

Better color separation can be reached by three-CCD devices (3CCD) and a dichroic beam splitter prism, that splits the image into red, green and blue components. Each of the three CCDs is arranged to respond to a particular color. Some semi-professional digital video camcorders (and most professional camcorders) use this technique. Another advantage of 3CCD over a Bayer mask device is higher quantum efficiency (and therefore higher light sensitivity for a given aperture size). This is because in a 3CCD device most of the light entering the aperture is captured by a sensor, while a Bayer mask absorbs a high proportion (about 2/3) of the light falling on each CCD pixel.

The Foveon X3 sensor [55] presents another approach to obtain color images. It is a CMOS image sensor for digital cameras, designed by Foveon, Inc. It uses an array of photosites, each of which consists of three vertically stacked photodiodes, that are organized in a two-dimensional grid (as is shown in Figure 1.4). Each of the three stacked photodiodes responds to different wavelengths of light, because they have a different spectral sensitivity curve. This is due to that fact that different wavelengths of light penetrate silicon to different depths. The signals from the three photodiodes are then processed, resulting in data that provides the three additive primary colors, red, green, and blue.

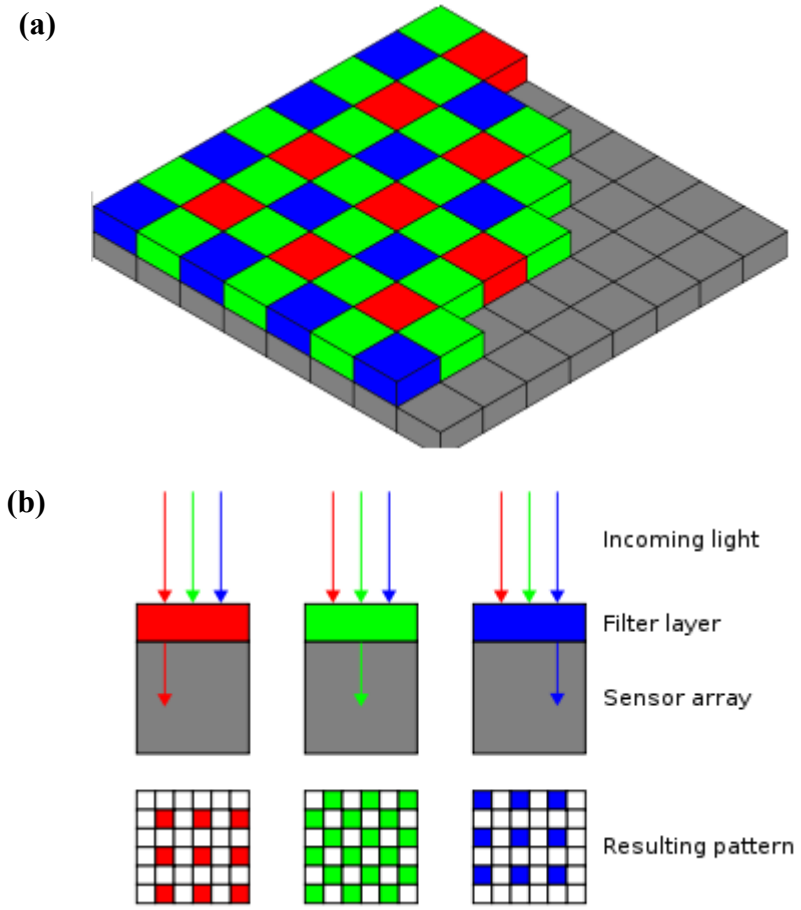


Fig 1.3: (a) The Bayer arrangement of color filters on the pixel array of an image sensor. (b) Cross-section of the sensor.

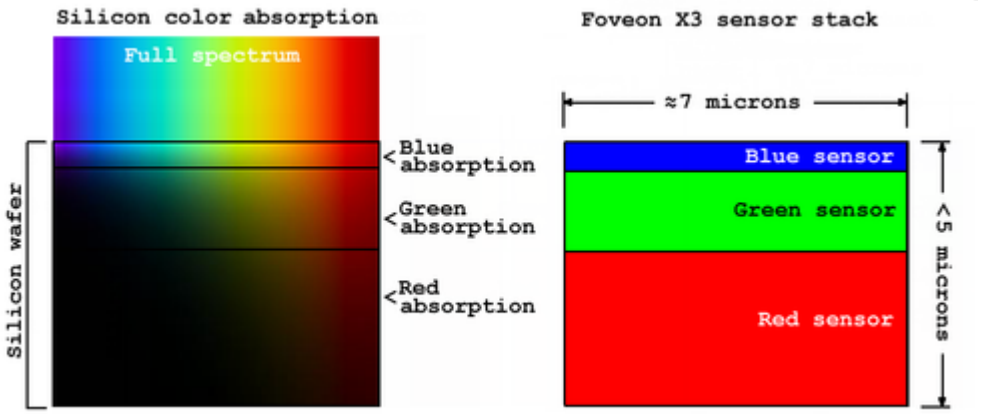


Fig 1.4: Color absorption in silicon and the Foveon X3 sensor.

1.2 Bio-inspired Systems and Neuromorphic Engineering

If we compare biological vision systems to frame-based systems, the first ones have clear advantages over the artificial ones. Biological vision systems are very complex and have amazing capabilities [4]. The visual system contains about 100 million photoreceptors. It has a dynamic range of more than 10 decades [4] that allows to detect dim starlight (it can sense the absorption of one single photon per integration time) to direct sunlight (it saturates at a 100 photons per integration time) [52]-[53]. Under ideal conditions, humans can detect a 0.5-percent change in mean intensity [56]. The human retina can resolve 1 arc minute (1/60 degree) [57] at the fovea, a small specialized region in the center of the retina where the cones are extremely small and are packed densely. Our eyes are also sensitive to temporal changes in the image. The integration time of the primate visual system is on the orther of 0.1 second [58]. Our eyes are sensitive to the wavelength of light. The human eye has several 100 million photoreceptors and the density of photoreceptors in the center of the visual fields as high as 150.000 photoreceptors per square-millimeter [4]. Photoreceptor pitch goes down to about $2.6\mu m$.

Looking at the state of the art, machine vision has only certain features that are better when they are used in very controlled environments. Biological vision systems perform efficiently real time tasks under very different conditions and are quite different to frame-based systems. Biological systems are not based on frames. They are continually detecting information. They are also massively parallel and data driven. Such systems present low redundancy, high speed response and high dynamic range. Human eyes has also high power efficiency. However, their most important feature is the possibility of extraction of relevant information from the visual scenes and transmit it to the brain in an accurate way. They are estructurated in layers of neurons connected between them. Each layer perform a certain task like spatial or temporal contrast detection. To extract the relevant information of the scene, a spatio-temporal filtering operation is done. This way, temporal and spatial redundancy is removed. Figure 1.5 shows an example of the operation of spatial filtering done in the human eye before transmitting the visual data to the brain. On the left side, there is on image captured with a conventional system. On the left side, there is the same image after filtering it with a high pass filter. We can notice that edges and shapes have been enhanced. At the same time, all the redundant data that was not necessary to distinguish shapes and edges was removed. Developing bio-inspired systems that performs any kind of processing the bandwidth usage will be optimized.

Figure 1.6 illustrates how frame-based and bio-inspired ones performs when they detect a very fast rotating stimulus and some advantages of bio-inspired systems. We have plotted with points the digital recorded output of one of our temporal contrast retinas when it was stimulated with a high speed rotating dot. Only pixels that detected temporal contrast generated outputs. In biological systems there are not scan periods and pixels are continuously detecting changes in the scene. If pixel bandwidth is enough high, the rotating dot will be detected. As we mentioned previously, biological systems usually process the information. In this particular case, the sensor was able to detect the temporal contrast (temporal filtering) and its sign (each color represents one sign of the spatial contrast). We have also plotted how a frame-based system would perform. It would be generating frames periodically. Such frames would sense the accumulated photo charge of all the pixels of the retina without processing the information. In this particular case, T_{frame} is higher than the stimulus period. This would be a serious limitation and we would not detect the

(a)

(b)

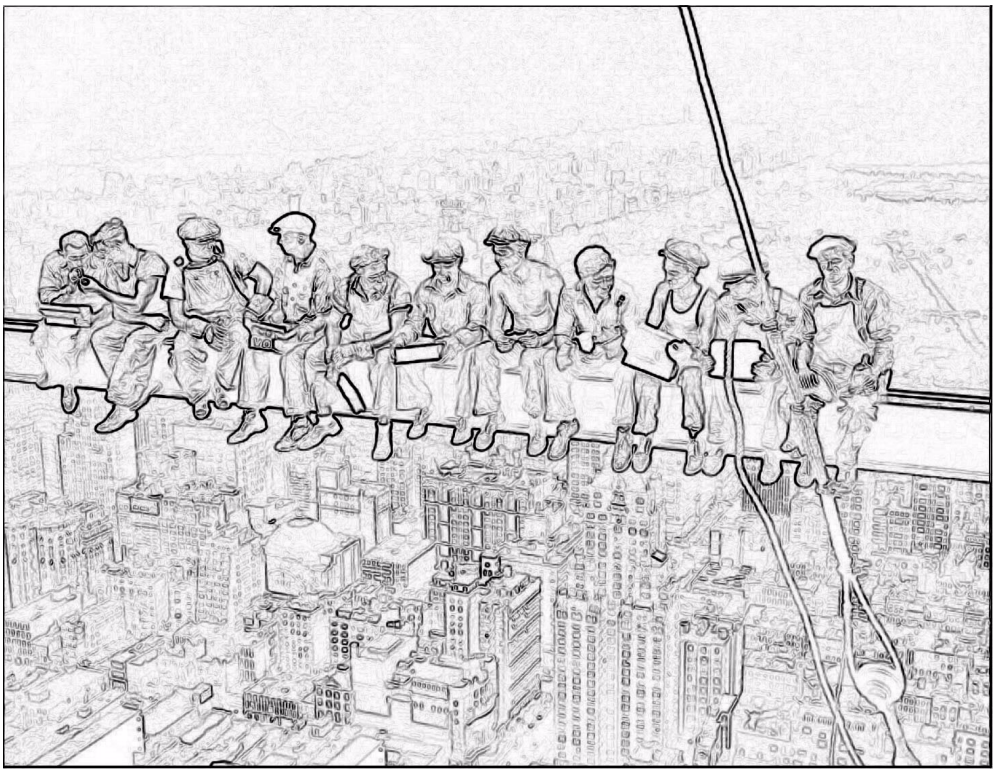


Fig 1.5: Example of the importance of spatial filtering to detect shapes and edges and remove redundancy. (a) Original image captured with a frame-based system. (b) Resulting image after high pass spatial filtering by using Matlab.

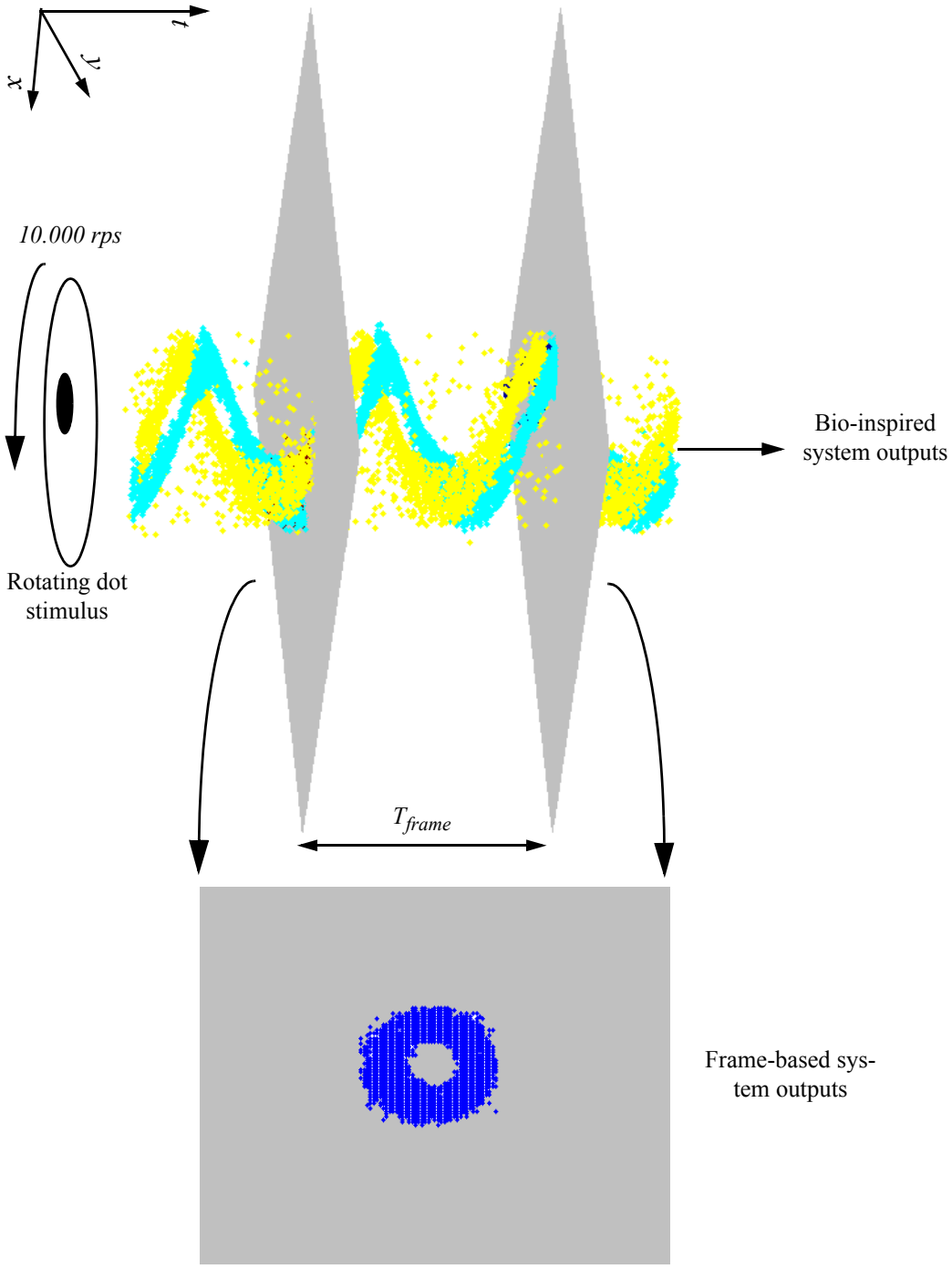


Fig 1.6: Comparison between a conventional system based on frames and a bio-inspired one tested in our laboratories. The first one (blue points) does not process the information and cannot detect the stimulus because its period is lower than T_{frame}. All the matrix pixels transmit the sensed luminance level. The second one (yellow and cyan points) extracts the temporal contrast (with sign) and can detect the rotating dot. In this case, only regions with contrast transmit information. A temporal filtering operation was done by the sensor.

Fig 1.7: (a) Image with very different levels of illumination taken with a conventional camera with a global AGC. (b) Same image after processing it with an algorithm that implements a local AGC.

dot movement. Each frame would detect a circle, but not the dot positions. All the frames would be identical.

Another advantage of bio-inspired systems is the local AGC (Automatic Gain Control). The operation range of a typical imager is about 10 decades. Pixel gain has to be adjust to the illumination level to have always the same sensitivity and response in the same way to the relative illumination variations. Conventional systems have a global AGC mechanism. The average chip luminance level is sensed and all the pixel gains are adjusted in the same way. However, this mechanism does not offer good results when there are regions in the pixels matrix with very different levels of illumination. On the contrary, biological systems have a local AGC. Each individual pixel senses the local photo current and its sensitivity is adjusted according this sensed illumination value. Thus, it is possible to capture images where there are very different levels of illumination without loosing sensitivity in any region of the scene. Figure 1.7(a) shows a image taken with a conventional camera with a global AGC. In the image, there were regions with very different levels of illumination. We can notice that it is not possible to distinguish the objects and shapes beyond the window. Figure 1.7(b) shows the some picture after using an algorithm that sets each pixel luminance value taking into account only its neighboring pixels (local AGC) and not all the matrix pixels. The algorithm increases the level of contrast of the scene. In this case, it is possible to distinguish the objects beyond the window.

For the previously mentioned reasons, it makes sense to try to develop devices that emulate the biological systems. Neuromorphic engineering is a discipline that seeks to implement neural structures in silicon. Primate brains are structured in layers of neurons. In each layer, neurons are connected to a large number of neurons ($\sim 10^4$) from the next layer [5]. There are massive connections that allows to implement efficient processing of the information between layers. In early neuromorphic engineering, it soon became clear that one of the major differences between brains

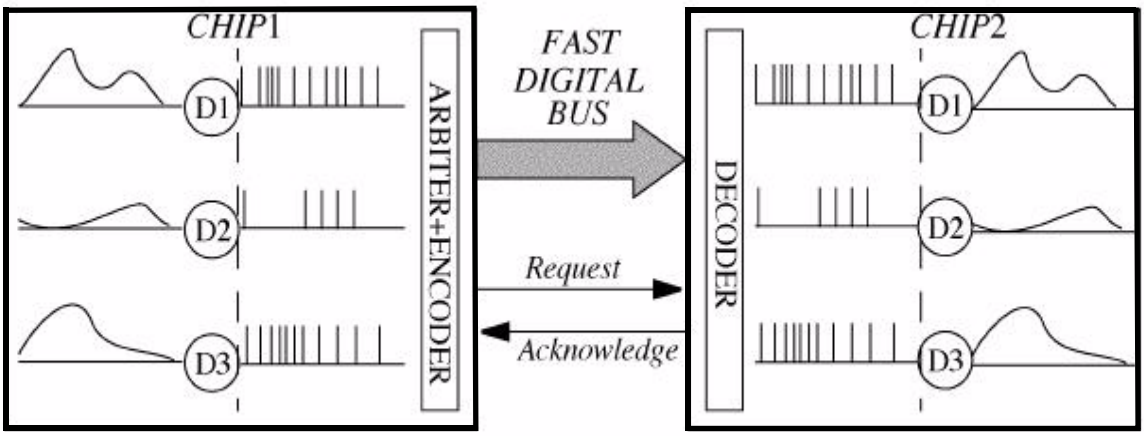


Fig 1.8: AER interchip point to point communication scheme.

and computers was the connectivity between processing units. In order to try to develop multichip systems connected between them massively in a accurate way, efficient mechanisms of interconnection and specific communication protocols are required. Nowadays, VLSI circuit technologies allows to fabricate on a single chip thousands of neurons. However, it is not viable and possible to connect all of them physically. It has also sense to implement layers in different chips leading to multichip systems. In such systems, each single chip (layer) run in parallel and the single components are more simple.

AER interchip communication was originally proposed by Mahowald and Silvilotti [1]-[2]. Biological neurons of a layer communicate with neurons of the next layer by sending information with spikes. In Fig 1.8, we can see a scheme of an AER communication between two different chips that implement two different layers of neurons. One chip initiates the communication sending a Request Signal (RQST) to the other chip. If the receiver is ready to receive information, it sends an Acknowledge (ACK) signal to the first chip. Then the first chip sends the information with the address of the neuron that has to receive the spike. This way, different neurons in different chips can communicate without being physically connected. Arbitration mechanisms are necessary in the transmitter to avoid collisions between neurons in the fast digital bus.

Multi-chip systems are usually made up with flexible systems that can be built independently to meet specific needs. They also allow to combine the effort of different groups avoiding the chip interconnection problems. AER can be considered a real time communication system because the total activity of a population of sensory or computational units is not close to the bandwidth of the communication channel.

1.3 The Biological Retina

Fig 1.9 shows a cross-section of the human eye and schematic of the different cells that compound the retina. Light enters the eye passing through the cornea. Around the lens there is a colored ring of muscle called the iris. The iris blocks light, forming an adjustable aperture in the center where light can enter the eye through the lens. This hole, called the pupil, is adjusted by the iris muscles to maintain a relatively constant level of light entering the eye. The light passing

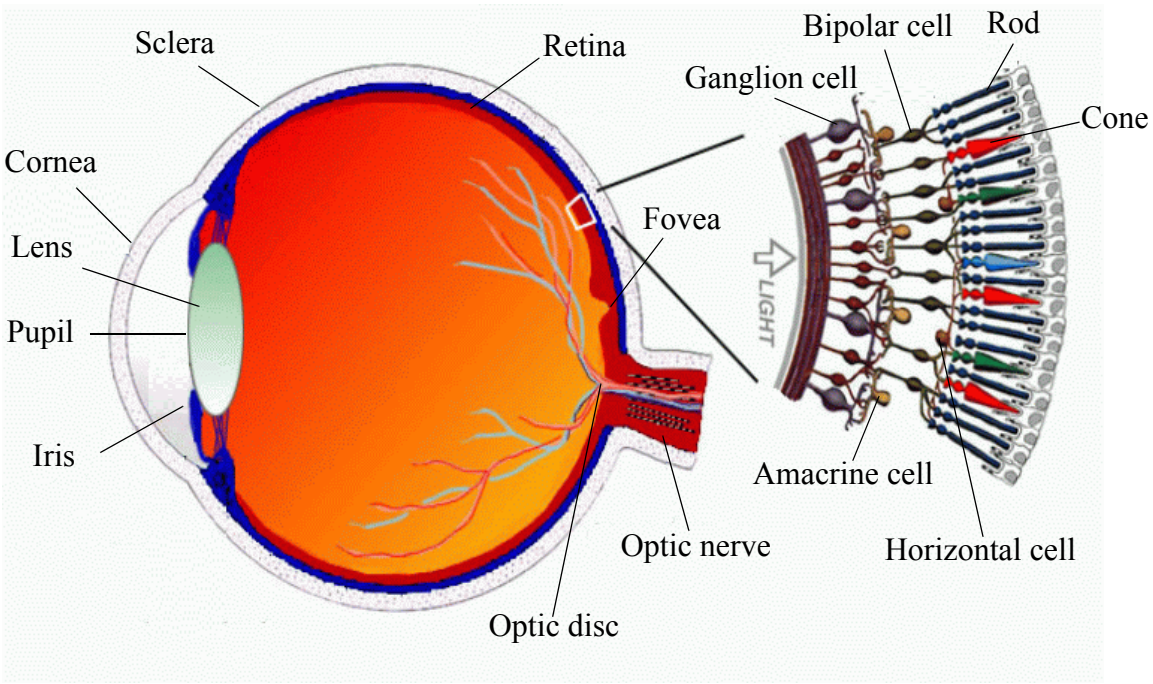


Fig 1.9: Human eye and schematic of the retina.

through the pupil is focused by the cornea and the lens. While the cornea has a fixed curvature, the lens can be stretched to adjust the focus to the distance of the object which the eye is looking at. The innermost tissue layer on the back portion of the eyeball is called the retina. All the individual components through which light travels within the eye before reaching the retina are transparent, minimizing attenuation of the light. As we can see in Fig 1.9, the retina contains different kind of cells: rods, cones, and neurons. Its main parts are the fovea, the optical disc and the optical nerve. The fovea is a small specialized region in the center of the retina where the cones are extremely small and packed densely (there are no rods in this region). The optical disc is a point on the retina where the axons of the ganglion cells leave the eye to form the optic nerve. There are no photosensitive cells at the optic disc.

The retina is structured in two layers: the outer plexiform layer (OPL) and the inner plexiform layer (IPL). The OPL is the more pheriferical layer and the IPL is the more central one.

The retina contain several kinds of cells and photoreceptors. There are two different kinds of photoreceptors [4], called rods and cones. They operate under different conditions of illumination. Rods operate in dim light and they cannot distinguish colors. Cones are less sensitive and operate under bright light and are sensitive to the wavelength of light due to the presence of three different types of cone pigments, with peak absorbances at 420*nm* (violet), 530*nm* (green-yellow) and 560nm (blue-green). In comparison, rods are tuned to 500*nm* (blue-green).

Neural responses go from the cones and rods to a series of linking cells, called bipolar, horizontal, amacrine cells, and ganglion cells. Bipolar cells are neurons placed between the OPL and the IPL and have a bipolar structure. Horizontal cells play an inhibitory role. Primates have two horizontal-cell types with no color selectivity. Amacrine cells are also inhibitory. Ganglion cells are the only output channel of the retina. These cells communicate by sending trains of impulses.

All these cells are connected creating a neural network that provide a mechanism for spatial and temporal filtering that allows to emphasize edges and temporal changes in the visual field.

This network enhances the relevant information and does not send to the brain redundant or useless information.

We can conclude that biological vision is not just about measuring or sensing the power and the light spectral composition of the visual scene. Biological systems can also process the information and enhance the useful information to send to the brain and avoid redundancy.

1.4 AER Vision Sensors

1.4.1 Adaptive Temporal and Spatial Filtering

The retina has evolved sophisticated filtering and adaptation mechanisms to reduce redundancy and to improve coding efficiency [4]. Some of these mechanisms are local automatic gain control, bandpass spatio-temporal spatio-temporal filtering and highpass temporal and spatial filtering. Bandpass spatio-temporal filtering in the first stage of the retina (OPL) attenuates signals that do not occur at a fine spatial or temporal scale, removing redundant transmission of low-frequency signals and eliminating noisy high frequency signals. Highpass temporal and spatial filtering in the second stage of the retina (IPL) attenuates signals that do not occur at a fine spatial scale and temporal scale, eliminating the redundant signals passed by the OPL, which responds strongly to low temporal frequencies that occur at high spatial frequencies (sustained response to static edge) or to low spatial frequencies that occur at high temporal frequencies (blurring of rapidly moving edge).

Boahen developed a simple dynamic model of the retina [4] to study the spatio-temporal behavior. It was a physical model made up resistors, capacitors, and transconductances. It was based on the neurocircuitry of the vertebrate retina and included several major synaptic interactions in the outer plexiform layer (OPL). This model was very useful to understand and study the spatio-temporal filtering in the retina, but it did not have useful applications because it has so many practical difficulties.

1.4.2 Spatial Contrast Extraction

There have been several attemps to implement a spatial contrast sensitive retina. The first functional retina able to detect contrast [46] was proposed by K. Boahen and A. Andreou. That retina could measure the spatio-temporal contrast. It was based on the model [4] previously proposed by Boahen to study the spatio-temporal filtering in the human retina. The spatio-temporal retina had a resolution of 90 x 92 pixels and a dynamic range of 5 decades. It mains drawbacks were mismatch and bandwidth consumption when there was no contrast. It could not detect the sign of the contrast. The first version was a neuromorphic system with a frame-based output. A further version of the retina [21]-[22] had AER output and was an event-driven system. Nevertheless, it also had some drawbacks like high mismatch and bandwidth consumption when there was no contrast.

After this approach, we have to highlight the work of Barbaro and Ruedi et al. [15]-[16]. They designed two devices. The first one was based on current mode and the second one was a improved version based on voltage mode. It was close to frame-devices because events were transmitted after a global reset. However, there was no information loss because information were immediately available after resetting the device and spikes were sent in order. The device had high dynamic range (120dB) and very low contrast mismatch (2%). It could detect spatial contrast and gradient direction. The main limitations of this retina were the power consumption ($300mW$) and temporal redundancy. The device was not able to detect the temporal contrast and its temporal

resolution was limited by the frame rate. There is also an improved version of this retina [17], it has higher resolution (320 x 240), 132dB of dynamic range, and very low contrast mismatch (0.5%).

The contrast retina presented in this thesis [23] and explained in chapter 4 is an improved version of Boahen's retina. It solves the main limitations of Boahen's approach and has some additional features like calibration, thresholding, time-to-first spike mode and bipolar output.

1.4.3 Temporal Contrast Extraction

Motion provides important information. This information is useful for the perception of relative depths, 3D structures, and for grouping objects that we see. The motion pathway is also independent of other visual field pathways and plays a very important role.

The first attempt to implement an AER temporal contrast sensor was done by Mallik et al. in 2005 [7]-[8] and was based on the prior work of V. Gruev et al. [50] describing a pipelined temporal difference imager. Some previous reported retinae were also able to detect temporal contrast, but they had low bandwidth and they were not specific devices for temporal contrast detection. Mallik's retina modified the traditional active pixel sensor (APS) to detect a quantize absolute change in illumination. It was a frame-based device. Temporal contrast was computed from the absolute difference between frames, so events times were discrete and not continuous. It had the typical advantages of APS sensors: good fill factor (17%), pixels with reduced size and very low fixed pattern noise (0.5%). Its bandwidth was limited by the frame rate and its dynamic range was relatively low (51dB) compared to other AER sensors.

The first temporal contrast AER sensor (also known as the optical transient sensor) was designed and proposed by Jörg Kramer [61]-[62]. It was a combination of Tobi Delbück's adaptive photoreceptor [63] with a rectifying and thresholding differentiating element [64]. The outstanding sensor proposed by Lichsteiner and Delbrück [10]-[9] was based on the Kramer's optical transient sensor and has high dynamic range (120dB), low latency ($15\mu s$), good fill factor (8.1%), and low noise (2.1% contrast). That sensor calculates the temporal derivative of the relative illumination. It is an asynchronous sensor with 128 x 128 pixels. Its bandwidth was limited by the photoreceptor bandwidth.

In chapter 5, we present a new AER temporal sensor based on that work. It has lower latency, higher bandwidth and better fill factor. This sensor is specially useful for high speed applications and surveillance. It uses amplifiers at the photoreptor stage to improve the sensitivity and the speed response. It also has a more efficient scheme of arbitration [47] that makes it faster. Pixel size has been reduced 50%.

1.4.4 Time-to-First Spike Imagers

Vision Systems are not based on frames, however the data that arrive immediately after the end of a rapid eye movement could be considered as a frame. Sometimes, it is useful to stop the data transmission and then restart it without data loss (bandwidth consumption control). The first time-to-first spike sensors [7]-[12] are based on S. Thorpe [11] ways of processing the visual information like it is processed in the human visual system. After a global reset, the most active pixels (pixels with more relevant information) will spike first. After a certain period of time, we will only receive redundant information and noise. If we stop the data transmission, we will improve the quality of the image and the bandwidth consumption will be reduced.

The new AER spatial contrast extraction retina presented in this thesis has an optional time-to-first spike mode that allows to combine the advantages of the frame-based processing in an

AER system: fast response and adjustable frame time. The sensor has a global reset. After presetting the retina, it will stop sending information to the AER bus. Thus, frame time can be set autonomically by adjusting the global preset period.

1.4.5 Luminance Sensors

These sensors [19]-[20] (also called octopus retinae) detect the light intensity and generate a train of spikes with a frequency proportional to the local luminance. Although these sensors just sense the illumination level, they have some advantages over frame-based devices, like continuous detection. However, spike frequency depends on the light and dynamic range is reduced in comparison to other kinds of AER retinae. For low illumination levels, the output frequency can be very slow.

1.5 New Developed AER Sensors

The goal of this work was to develop two different vision sensors that can detect spatial and temporal contrast. Both sensors improve the state of the art and are insensitive to illumination conditions and propose new ideas that can be used in further AER sensors.

Mismatch is a problem that has plagued prior AER retinae. For this reason, we have also proposed a new calibration mechanism, that has been validated experimentally and used in the design of the contrast retina.

The first sensor (described in chapter 4) can detect spatial contrast. It is a 32 x 32 pixels prototype that has some advantages over prior designs. It is insensitive to illumination conditions, it has a new thresholding mechanism, an optional time-to-first spike mode and calibration.

One of the main disadvantages of prior AER contrast retinae was high mismatch. For this reason, we developed a new compact calibration system (described in chapter 3) for large neuromorphic arrays [13]. This calibration technique has been successfully used to calibrate the mismatch of the contrast retina and offers higher resolutions that other calibrations techniques used in prior designs [14].

The second sensor (described in chapter 5) detects temporal contrast. It is a 128 x 128 pixels retina with high bandwidth and very low latency. It is specially useful for high speed applications and surveillance.

The thesis has five chapters. Chapter 2 presents the *Stochastic I-Pot*, a compact and powerful circuitry to polarize and test circuits. Chapter 3 describes the new proposed compact calibration system (Newcalib) for neuromorphic arrays. In chapter 4 we present the new spatial contrast AER retina. We explain thoughtfully the pixel desing and we show experimental results. In chapter 4, we present the temporal contrast AER retina. Explanations about the pixel design and experimental results are also provided. Finally, in chapter 6, we draw some conclusions and possible future work is proposed.

Table 1.1 [57] summarizes the more important functionalities and performance figures of prior AER sensors. Three types of functionalities are considered: sensing pixel luminance, sensing pixel temporal contrast, and sensing pixel spatial contrast with respect a given neighborghood. For (spike) signal coding, three methods are used: signal to frequency (rate) coding, signal to number of events (NE) coding, and signal to time-to-first spike (TFS) coding. When using rate-coding (as in [21]-[22]-[14]-[23]), a current that carries the information of interest (luminance, contrast) is fed to an integrate-and-fire circuit whose spike frequency is controlled by the current. For NE coding (as in [7]-[9]-[24]), every time the light sensed by a pixel changes by a relative

amount, a new spike is generated. In TFS coding (as in [20]-[16]-[23]), the information signal is also fed to an integrate-and-fire circuit, but the integrators are (periodically and) globally reset and only fire one spike between consecutive resets. This way, the information is coded as the time between the global reset and pixel spike time. For a luminance retina [19]-[20], the photo current is the one to be integrated, resulting in timings strongly dependent on ambient light. Consequently, the dynamic range of light sensing capability is directly transformed into the latency variation of the output. This is a severe practical restriction, labelled in Table 1.1 as "Light to Time Restriction". For contrast computations (either temporal or spatial), light difference is normalized to average light, so that contrast is (by definition) independently of ambient light. Consequently, these retinae should not suffer for "Light to Time Restriction". This is the case for all contrast retinae in Table 1.1, except for [19]-[20]-[7]-[16]. [19] and [20] are luminance retinas. In [16], for each frame there are two separate steps in time. The first one uses a Light to Time integration (which lasts between $0.5\mu s$ - $150ms$ depending on the ambient light) to obtain a voltage representation of pixel contrast. The second step transforms these voltages into a TFS representation requiring an ambient-light-independence time of about $2ms$. In [7], pixels use an modified active pixel sensor (APS) and pixels can only detect absolute changes in illumination, so there is a also a light to time restriction. Actually, the dynamic range is reduced in comparison to the other devices.

Finally, the features and functionalities of our two new sensors are presented. The most significant features of each one have been highlighted.

Table 1.1: Comparison among AER vision sensor devices [57]-[23].

	Cul03[19]	Chen07[20]	Mallik05 [7]	Licht07[9]	Posch10 [24]	Zagh04 [21]-[22]	Ruedi09 [16]	Costas07 [14]	Spatial Contrast Retina [23]	Temporal Contrast Retina
Function-ality	Lumi-nance to Frequency	Luminance to TFS	Temporal Contrast to Num-ber of Events	Temporal Contrast to Number of Events	Temporal Contrast to Number of Events	Spatial and Tem-poral Contrast to Fre-quency	Spatial Contrast Magni-tude and Direction to TFS	Spatial Contrast to Fre-quency	Spatial Contrast to Fre-quency or **TFS**	Temporal Contrast to Number of Events
Light to Time Restriction	YES	YES	YES	NO	NO	NO	YES	NO	NO	NO
Latency	120 μs - 125s	10 μs -1s	<5ms	15 μs @1 Klux chip illumina-tion 2Meps	3.5 μs @1 klux	Not Reported	2ms-150ms	Not Reported	0.1ms-10ms	**3.5 μs @ 25Klux**
Dynamic Range	120dB	>100dB	51dB	120dB	125dB @30FPS	50dB	110dB	100dB	>100dB	>100dB
Spatial Contrast Computa-tion	N/A	N/A	N/A	N/A	N/A	difusive grid neigh-bourhood	4 nearest pixels (up, right, left, bot-tom)	diffu-sive grid neigh-bour-hood (adjust-able up to 10 pixels)	diffu-sive grid neigh-bour-hood (adjust-able up to 10 pixels)	N/A
Temporal Contrast Computa-tion	N/A	N/A	Temporal frame-dif-ference intensity charge detection	Temporal derivative of the rel-ative luminance	Temporal derivative of the rel-ative luminance	Temporal deriva-tive of the relative luminance	N/A	N/A	N/A	Temporal derivative of the rela-tive lumi-nance
FPN	4%	4.6%	0.5% of full scale, 2.1% TD charge	2.5%	SNR_{typ} >56dB	1-2dec	1.7%	6.6%	**0.6%**	2.8%
Fill Factor	14%	33%	17%	8.1%	10%	14%	20%	3%	2%	8%
Pixel Size μm	32x30	17x17	25x25	40x40	30x30	34x40	14x14	58x56	80x80	35x35
Fabrica-tion Pro-cess	0.6 μm 3M CMOS	AMIS 0.35 μm 5M 1P CMOS	0.5 μm 3M 2P	0.35 μm 4M 2P	180nm 4M 2P MIM	0.35 μm 2M 2P	180nm 1P6M	0.35 μm 4M 2P	0.35 μm 4M 2P	0.35 μm 4M 2P
Power	3-71mW	N/A	30mW@ 5V (50fps)	24mW	50-175 mW	63mW	300mW	33 μW - 10mW	**0.66-6.6mW**	132-185 mW

CHAPTER 2

Biasing for Neuromorphic Arrays: The I-Pot System[1]

2.1 Introduction

Analog circuits require, in general, a set of bias currents which are usually provided by a current reference circuit allocated somewhere at the periphery of the chip together with some current scaling and distribution circuitry [29], [30]-[31]-[32]. Many times, circuit designers would like to be able to fine tune some of the designed bias currents in order to compensate for process variations and components mismatches. In other occasions, when experimenting with new circuits, it is desirable to have large degrees of freedom to play with all available bias currents and test the circuits for wide ranges of operating conditions. Under these situations, the safest solution is to provide one external pin for each bias current, making accessible a transistor gate. This allows to connect all available gates to external potentiometers (or a *pot-box*) so that one can experiment freely in the lab with all combinations of biases. In practice, this approach is limited to a reduced number of freely adjustable biases, since the number of pins of a chip cannot grow arbitrarily. To overcome this problem, Hasler et al. introduced the E-Pot circuit block [33], which exploited floating gate and tunneling/injection circuit techniques to program non-erasable analog voltages onto gates of biasing transistors. However, floating gate and tunneling/injection techniques are still complicated to use successfully in standard CMOS processes.

In order to generate programable bias currents we introduce an alternative approach, based on the use of what we call the '*Stochastic I-Pot*' concept [37]. The '*Stochastic I-Pot*' is a digitally programmable current source which, from a reference current, can provide any desired current with high precision and down to pico-amperes. Each I-Pot cell includes a digital register to select a current range, a current value for this range, and a sign. I-Pot cells can be chained so that any arbitrary number of current biases can be generated and independently programmed. The number of external pins that a chip needs for characterizing and programming the chain of I-Pots is three, independent of the number of I-Pots assembled. Consequently, designers can include any arbitrary number of programmable current biases, with the only restriction of area consumption. In our par-

1. This chapter of the thesis is based on a previous reported paper ([37]).

ticular implementation, in a $0.35\mu m$ process, the I-Pot cell area is $130\mu m \times 68\mu m$, one third the area the pad it replaces.

We used the *Stochastic I-Pot* to generate all the bias currents that we needed to provide to the contrast retinae.

2.2 I-pots Circuit Architecture

The '*Stochastic I-Pot'* exploits the use of current mode ladder structures applied to MOS transistors [34]-[28]. A generic MOS ladder structure, configured as current source, is shown in Figure 2.1. All transistors are proportional to a unit size ratio *W/L* by either a factor 1, *N*-1, or *N*/(*N*-1). This way, branch currents I_i have an attenuation ratio of *N* from each branch to the next one. In the '*Stochastic I-Pot*' circuit, we use two of these ladder structures (as we can see in the circuit diagram of complete I-Pot cell in Figure 2.2)

The first ladder structure (Range Selector Ladder), with attenuation ratio around *N*=10, selects an operating current range. This range ladder has 6 output branches, so that the output current can be selected between the input reference current I_{REF} and around $I_{REF}/10^6$.

For the second ladder structure (Stochastic Ladder), we use an attenuation ratio of $N = 2$. This allows for selecting any binary combination of current branches, in the same way a current DAC would do. However, we do not want to have a high precision (like 8 bit) current DAC within each I-Pot, because they would require extremely large transistor sizes and would most probably not provide such precision for very small currents down to pico-amperes.

In our approach we use ladder structures with attenuation ratio $N = 2$, with a large number of branches, but with small transistor sizes so that we intentionally provide large mismatches between the current branches. By having a large number of branches, each with large mismatch, we achieve a good coverage of possible output current combinations. Consider a ladder with input current $300pA$, attenuation ratio $N = 2$, and 8 branches. The unit transistor size is $W/L = 1/0.7\mu m$. For a transistor of this size, fabricated in the AMS 0.35µm CMOS process, and driving a current of $\sim 700pA$, results in a current mismatch standard deviation of around $\sigma \approx 35\%$, as can be seen by interpolating the measurements shown in Figure 2.3, [28]-[35]. Figure 2.4(a) shows the output current obtained as function of the 8-bit digital word $\boldsymbol{w}_{val}$ that controls

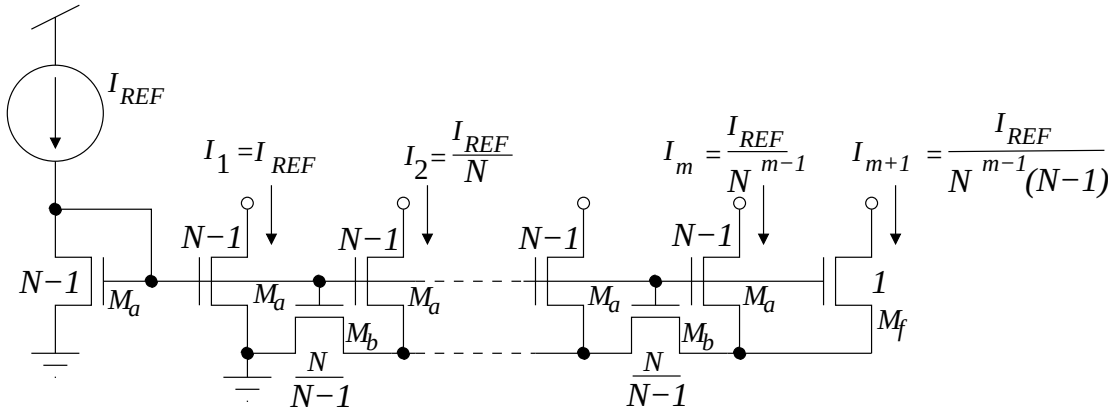


Fig 2.1: Generic ladder structure with *N* branches. The value of the current in each branch is indicated.

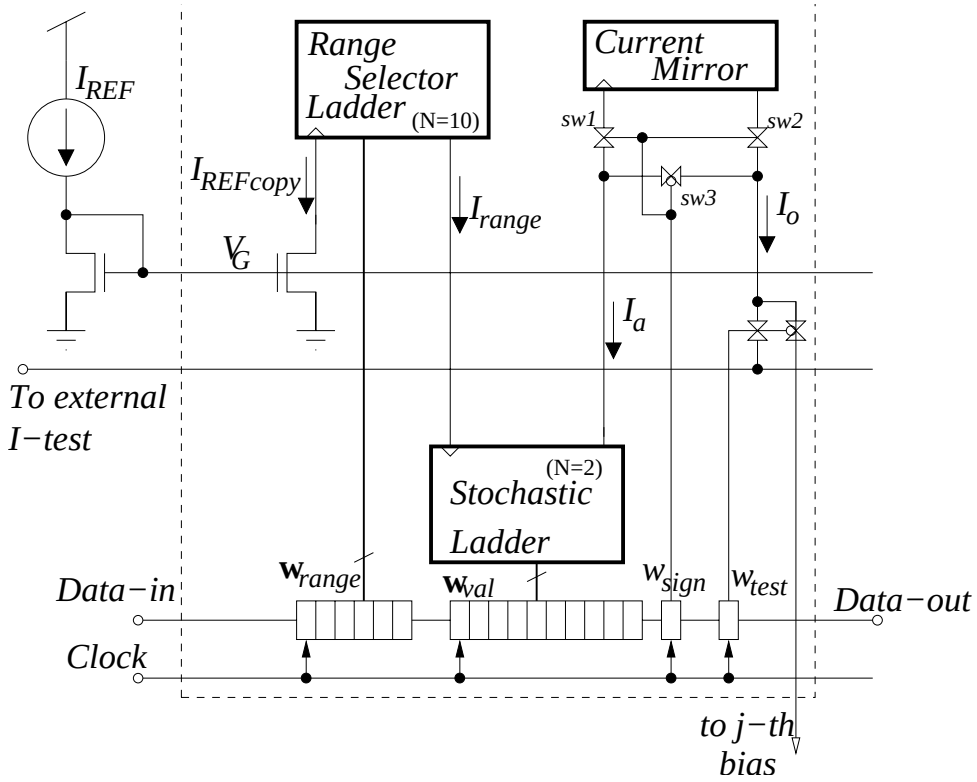


Fig 2.2: Circuit diagram of complete I-Pot cell.

the combination of branches as is shown in Figure 2.2. As can be seen, this characteristic differs dramatically from a conventional stair-case that a high-precision 8-bit DAC would provide. However, suppose we introduce a look-up table between the digital word we provide and the one physically applied to the ladder structure, so that the output currents become ordered. The result is shown in Figure 2.4(b) for the same ladder and bias reference current. Now we see a monotonic increasing dependence between the digital control word $\boldsymbol{w}_{ord}$ and the output current. It does not matter that this relationship is not perfectly linear. Our objective is simply to provide a bias current as close as possible to a desired value. Such objective is limited by the intervals between consecutive current values in Figure 2.4(b). To characterize these intervals, we show in Figure 2.4(c) the relative difference between the consecutive values in Figure 2.4(b),

$$\Delta_{rel} = 2\left|\frac{I_n - I_{n+1}}{I_n + I_{n+1}}\right| \tag{2.1}$$

As can be seen, we have errors below 10% for the whole range except the first 1/20 of the range, and below 1% for last 1/4 of the range. But this was a 'lucky' particular case. It is perfectly possible to find situations where the mismatch plays against us and we obtain an unfortunate extremely large maximum gap. This is for example, the case illustrated in Figure 2.4(d-f), for another I-Pot, exactly equal to the one used for Figure 2.4(a-c). Such situations occur for example, when the maximum branch current I_1 in Figure 2.1 results much larger than the sum of all the others. This produces an extremely large gap in the center of the characteristic. In Figure 2.4(e) the largest gap was of $92.3pA$ for a total range of $591.3pA$. Consequently, as can be seen in

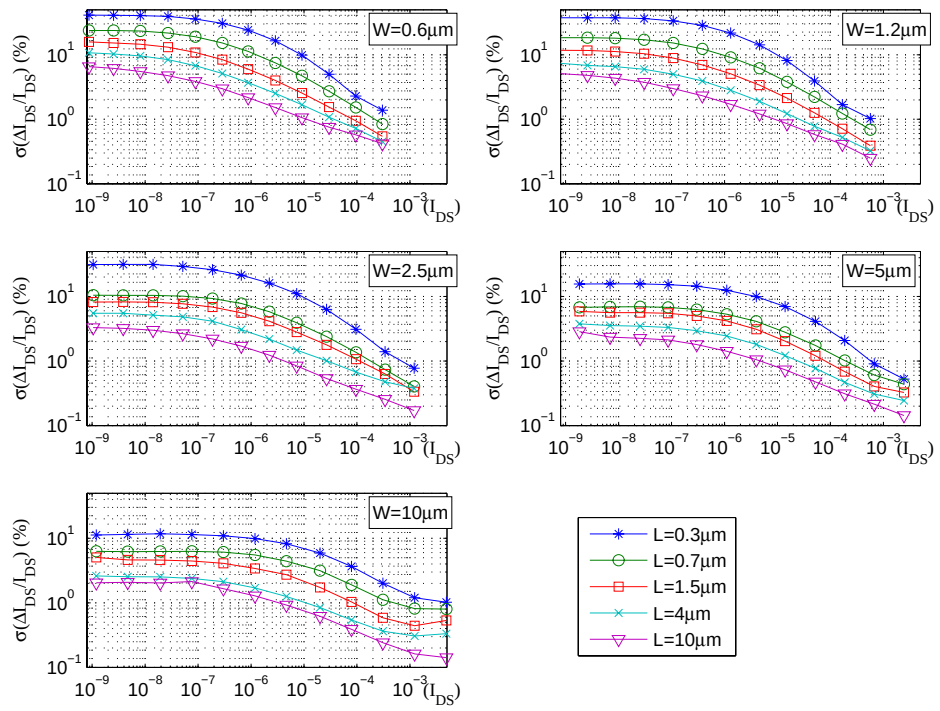


Fig 2.3: Current mismatch standard deviation measurements, for NMOS transistors in a 0.35µm CMOS process [28]-[35]. Vertical axes represent standard deviation in %, and horizontal axes operating current. Measurements are taken for 30 different transistor sizes, by sweeping width {10, 5, 2.5, 1.2, 0.6}µm and length {10, 4, 1.5, 0.7, 0.3}µm.

Figure 2.4(f), current values in the range between 250 to 350pA cannot be generated with a precision better than 30%.

One can think of several solutions to solve this problem. After playing with a few of them and testing them with statistical simulations, we found out that the most reliable solution is to duplicate the output branches of the ladder with attenuation ratio $N = 2$. This not only guarantees there will be no large gaps between consecutive current values, but at the same time reduces dramatically the value of the largest gap found, for the same transistor sizes and input reference current.

Figure 2.4(g-i) shows the same situation for the circuit in Figure 2.4(a-c), but where now the output branches are duplicated. Now there are 16 bits to select output branches combinations but the maximum current gap in Figure 2.4(h) is now reduced to $0.20pA$, excluding the first and last $100pA$ intervals of the range. The relative error is shown in Figure 2.4(i). As can be seen it is less than 0.001 for currents between $137pA$ and $1214pA$. Figure 2.5 shows the complete circuit schematics of a ladder with attenuation ratio $N = 2$ and duplicated output branches.

Looking at the complete circuit diagram of a stochastic I-Pot cell in Figure 2.2. Each stochastic I-Pot receives a copy $I_{REFcopy}$ of a common reference current I_{REF}, which is the input to a PMOS $N = 10$ range ladder structure with 6 output branches. The common reference current can be generated by any bandgap type circuit [36] with temperature compensation, or provided off-chip through an extra pin. The digital word $\boldsymbol{w}_{range}$ selects just one of the range ladder outputs (not a combination of them). This current I_{range} sets the coarse mode range of the I-Pot

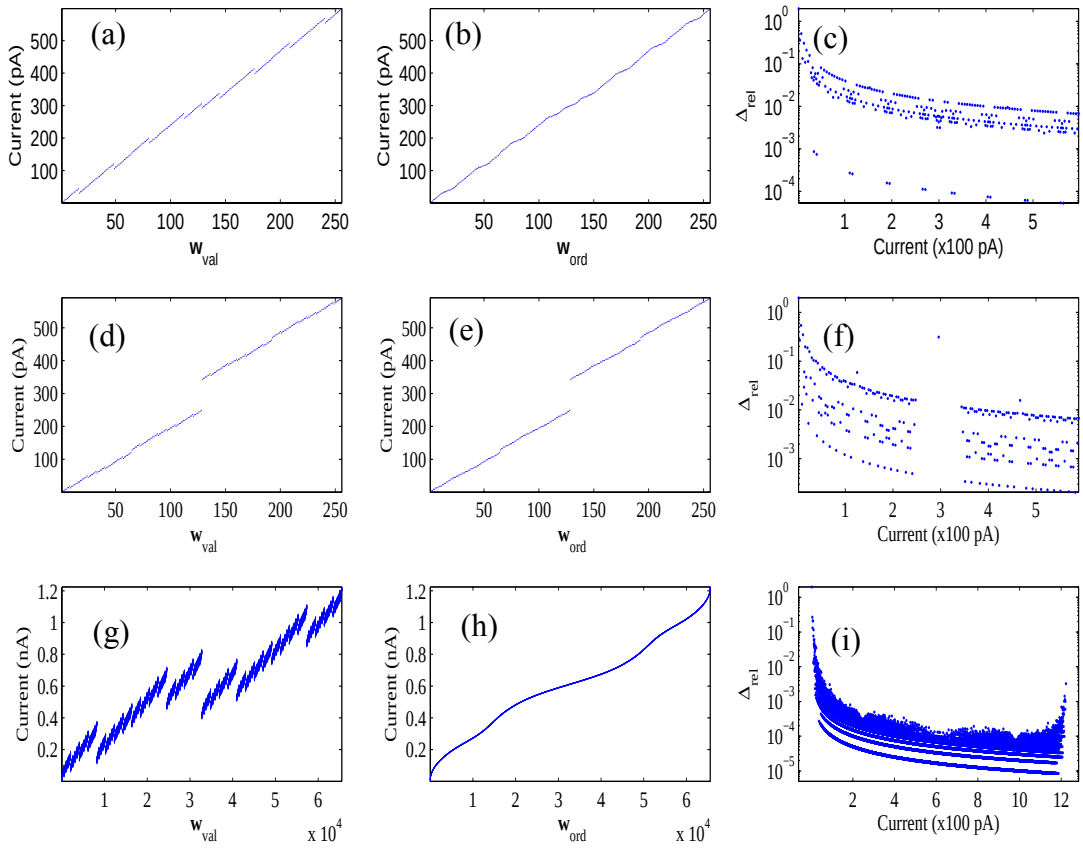


Fig 2.4: (a-c) Illustration of mismatch effects in a MOS ladder with ratio *N=2.* (d-f) Same, but for an 'unlucky' ladder example. (g-i) Illustration of a ladder structure with duplicated output branches (a,d,g) I-Pot output currents versus digital word w_{val}, (b,e,h) same output currents after ordering versus w_{ord}, (c,f,i) relative difference between consecutive values.

output current. Current I_{range} is now fed to the input of an NMOS $N = 2$ ladder structure with duplicated output branches. Let us call this ladder a "stochastic ladder". In our particular case we implement 16 (2×8) duplicated branches. The particular combination of output current branches is controlled by the digital word in register w_{val}. The output of this ladder I_a can be optionally sign inverted by transmission gates $sw1 - sw3$, controlled by the state of an extra register w_{sign}, which inserts or not a PMOS current mirror in the output current path. Finally, the signed output current I_o is directed to either its destination bias point, or to a chip output pin I_{test} for characterization purposes, depending on the state of register w_{test}. All registers, w_{range}, w_{val}, w_{sign} and w_{test} are shift registers, connected in series, and clocked by the same clock signal. I-Pot cells can be chained directly in series by chaining their shift registers sharing all the same clock signal. All I-Pot cells share also the same gate line V_G and test line I_{test} .

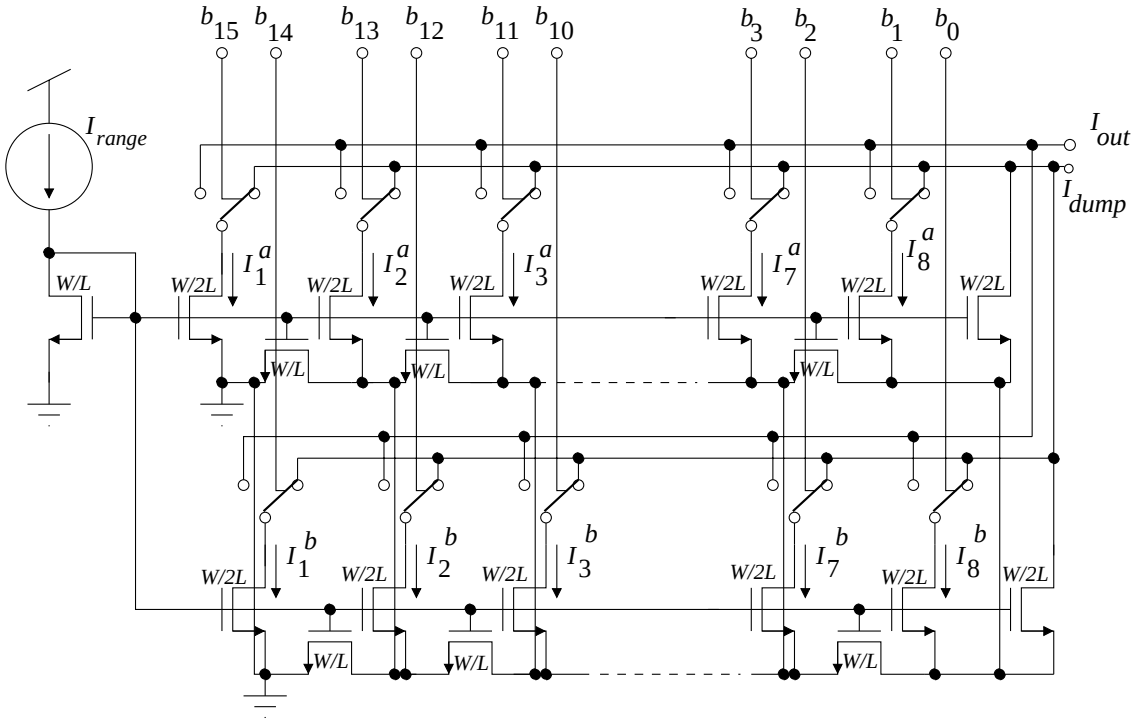


Fig 2.5: Circuit schematic of ladder structure with attenuation ratio *N=2* and duplicated output branches.

The main drawback of the present I-Pot approach is that each I-Pot of each fabricated chip needs to be characterized individually. The good news are that this is quite easy to do by using a host computer that loads the chained shift registers, while at the same time we require the use of just one single external current metering instrument.

The procedure for characterizing the I-Pots of a chip is as follows:

1.- Each I-Pot has to be characterized individually. Consequently, the $\boldsymbol{w}_{test}$ bit of only one single I-Pot has to be set to 'active'. All others must be disabled. This way, only one single I-Pot output is connected to external line I_{test}.

2.- Sweep the two signs for the active I-Pot.

3.- For the selected I-Pot and sign, sweep all current ranges through digital word $\boldsymbol{w}_{range}$.

4.- For the selected $\boldsymbol{w}_{range}$, sweep all 20 output current branches, measure the selected branch with the external current meter through pin I_{test}, and store it in a file in the computer.

After completing the measurements for one I-Pot, we will have stored in the computer a total of 2 signs x 6 ranges x 16 branches = 192 current values. For each sign and range, we can now produce all 2^{16} possible combinations, order them, and find the maximum gap. What now remains to do, is writing a computer program that given a desired I-Pot value will return the optimum $\boldsymbol{w}_{val}$ and $\boldsymbol{w}_{range}$ that gives the minimum error. The total amount of time required to caracterize one individual I-Pot was about 1 hour.

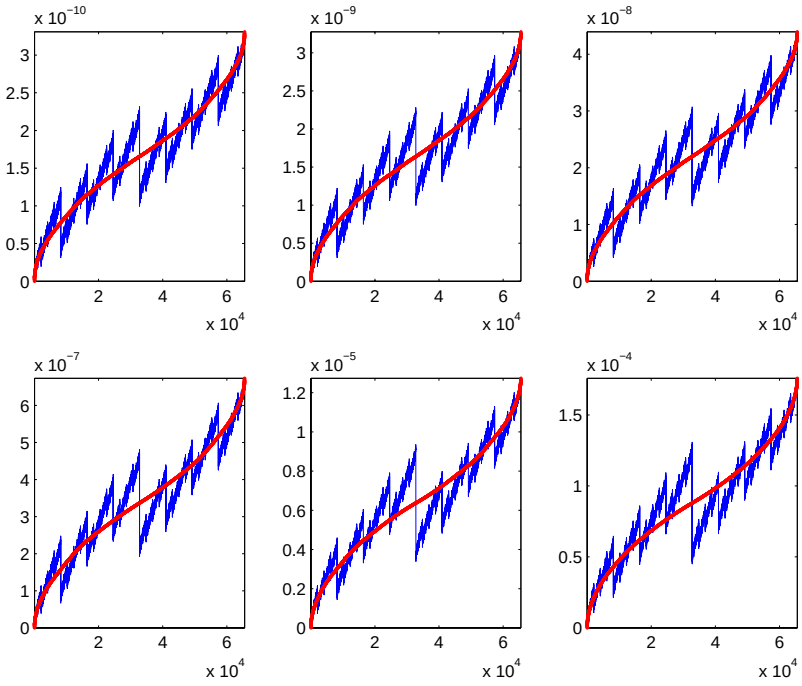


Fig 2.6: Measured currents of one of the fabricated I-Pots for all six current ranges. Continuos lines show the same values after ordering. Vertical axes are measured currents, horizontal axes are code values for the 2^{16}=65536 currents of each range.

2.3 Experimental Results

Here, we present some experimental results after characterizing the I-Pots. I-Pots were fabricated in the AMS $0.35\mu m$ CMOS process. The stochastic ladder uses 16 duplicated output branches, and the unit transistor of the ladder structure has a size of $W = 1\mu m$, $L = 0.7\mu m$. This ladder was intentionally made with a small unit transistor to increase mismatch, and therefore increase its stochasticity. The range ladder used 6 output branches, and the input current to the range ladder was set to $100\mu A$. The range ladder was designed according to Figure 2.1 with an attenuation ratio of $N = 11$, approximately.

After measuring one of the fabricated I-Pots we obtain the currents shown in Figure 2.6. Each of the six subgraphs corresponds to one of the available ranges provided by the range ladder. In each subgraph we have added the ordered version of the measured current values. This ordered version is drawn with the continuous line. The minimum and maximum current values provided by each range are as follows. First range [0.4*mA*, 176*mA*], second range [30*nA*, 12.7*mA*], third range [1.9*nA*, 673*nA*], fourth range [126*pA*, 44.0*nA*], fifth range [9.9*pA*, 3.28*nA*], and sixth range [1.2*pA*, 331*pA*].

Figure 2.7 shows these values, expressed in bits[2] as function of currents. Striation effects can be observed, specially as current decreases, due to quantization effects in the data acquisition instrument.

---

2. $1/2^{n_{bits}} = \Delta_{rel} \Leftrightarrow n_{bits} = -\log_2(\Delta_{rel}) = -(\ln\Delta_{rel})/(\ln 2)$

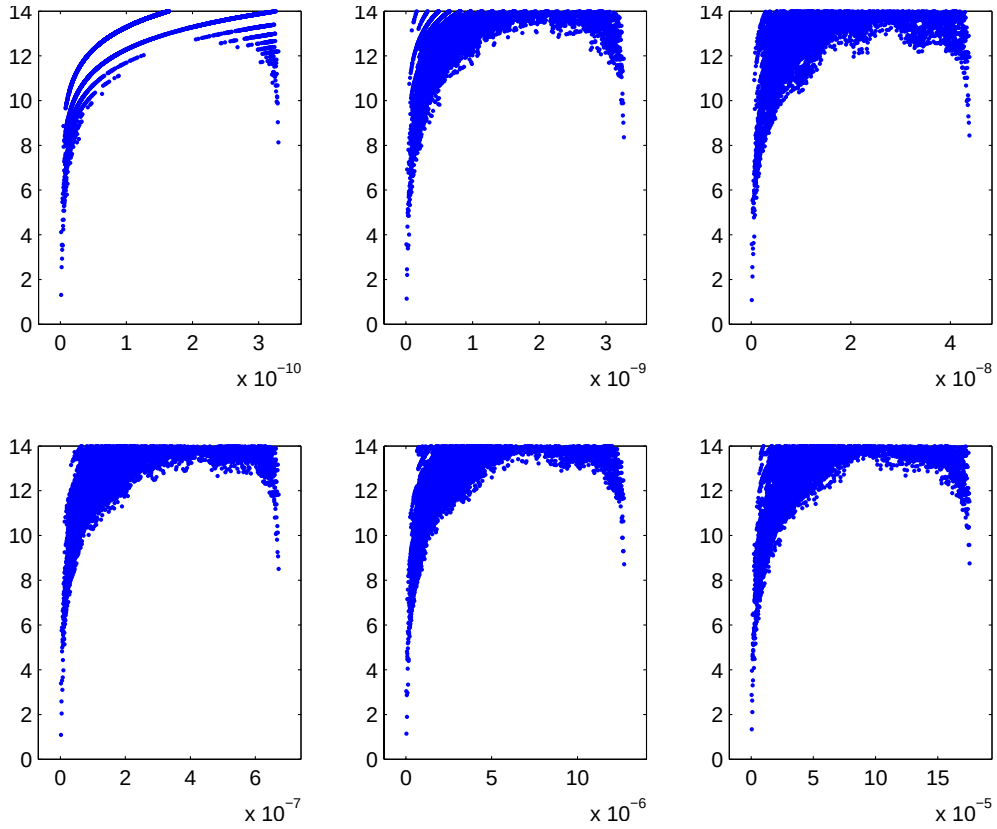


Fig 2.7: Computed relative increments between ordered consecutive values expresed in bits. Vertical axes are resulting bits, while horizontal axes are measured currents.

We can see that the maximum resolution is obtained for the central parts of the ranges, reaching values as high as 13-bits. On the other hand, for the external parts of the ranges the resolution decreases dramatically down to values as low as 1-bit. This is because in the central part of the ranges there is a higher density of redundant values than on the extremes of the ranges. However, if the ranges overlap, we can increase the density of redundant values and improve the resolution. The procedure is as follows. Let us take all the measured values of all six ranges and order them as one unique set of current values. Each current value is uniquely defined by its range and its 16 bit word within that range. Let us now compute the difference of consecutive values, as defined by eq. (2.1) on page 63, and express them in bits The result is shown in Figure 2.8. We can still see very well the regions of the six ranges with their respective maximum resolution central parts of up to 13-bits. However, the resolution of the extreme regions of the ranges has improved to values of above 8-bits. Neglecting the two extreme regions of the whole merged six ranges, the worst resolution is obtained for currents around $0.7mA$, yielding a resolution of 8.05-bits. The first current value showing a resolution above 8 bits is $19.6pA$, and the largest one $176mA$.

2.3.1 Temperature Effects

Ideally speaking, since the I-Pot operation is based on current splitting techniques, the effect of temperature should be minor as long as the current division ratios between ladder branches remain unaltered. Unfortunately, this is not completely true, and those ratios do depend slightly on temperature. We have observed that those ratios can vary between 0.1% and around 1% for tem-

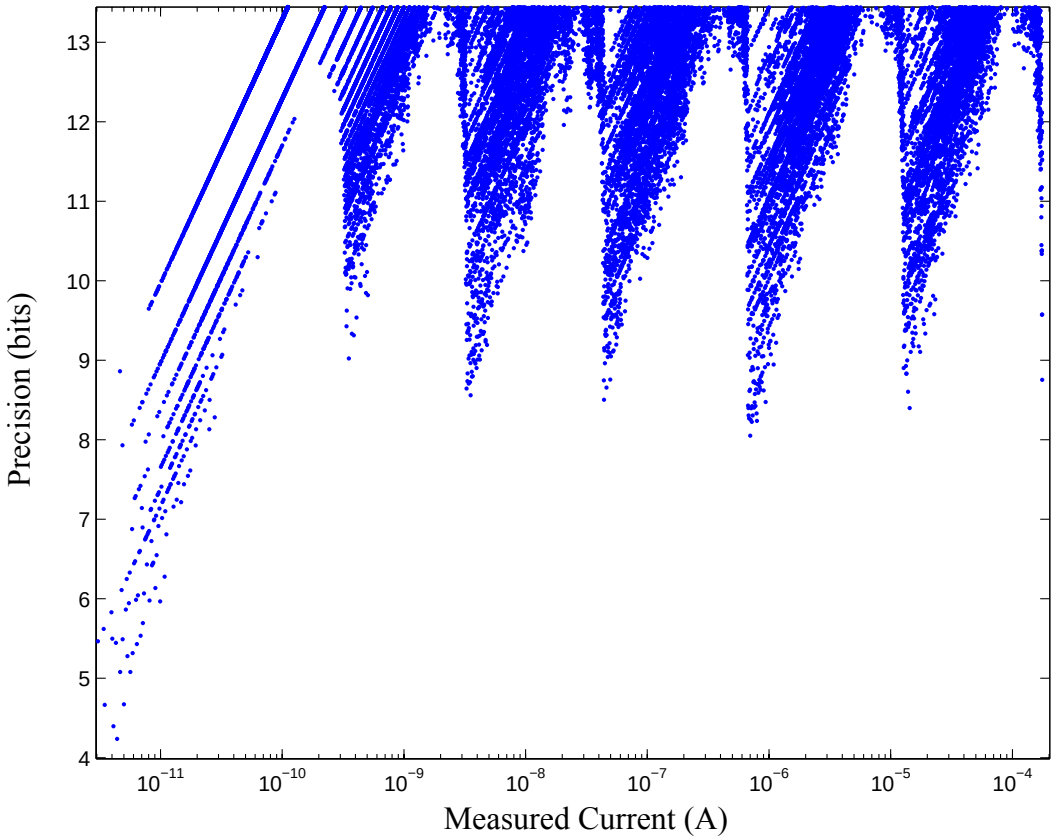


Fig 2.8: Computed relative increments, expressed in bits, when considering a measured values if all six ranges as a unique set. Vertical axis represents resulting bits, while horizontal axis represents absolute measured current.

perature variations between 0º-40ºC for currents above tens of nano-amps. For lower currents the effect of temperature is more drastic, since leakage currents increase rapidly with temperature. Figure 2.9 shows the error (expressed in bits) between the I-Pot current expected when characterized at 20ºC, and the one obtained when doing new measurements, sweeping temperature between 0ºC and 40ºC. Each data point is obtained by taking the maximum error within 1/30-th of the 7-decade current range.

2.4 Conclusions

In this chapter, we have presented the *Stochastic I-Pot.* It is a versatile and powerful circuit for generating digitally controlled precise bias currents has been presented. 8-bit precision has been verified from currents as low as 19.6*pA* up to values of 176*mA* (almost 7 decades). Temperature degrades precision gracefully, and is more severe for smaller currents. This circuit is specially handy for experimenting with current-mode circuits operating in weak inversion, where mismatch is high and operating range extends over several decades. It is also very useful for experimenting with new circuits, where it might be desirable to include a large number of fine-tunable current biases for trimming gains and offsets. The only drawback is that each I-pot needs to be characterized individually using an external off-chip current metering instrument.

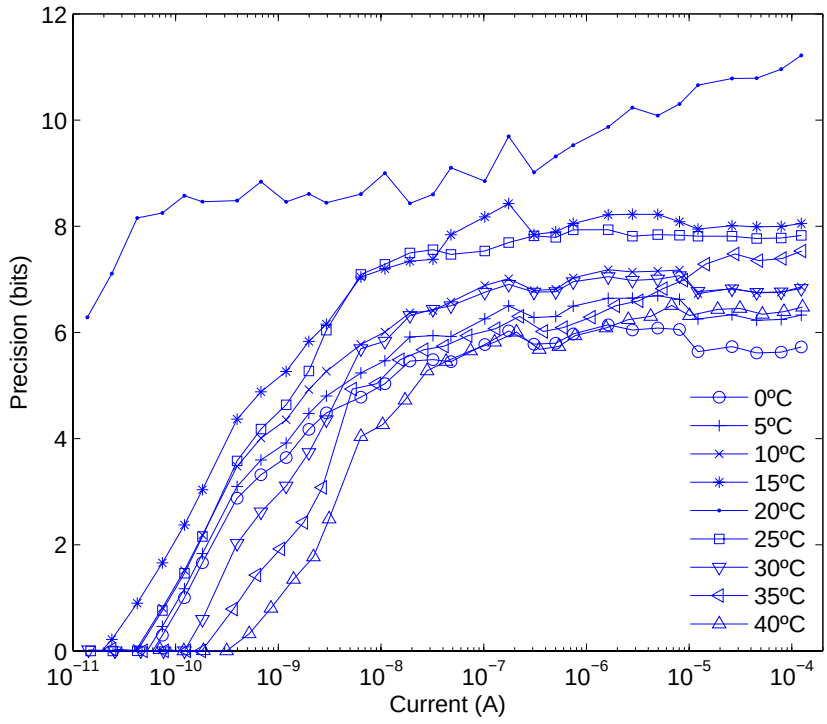


Fig 2.9:Effect of temperature on precision.

I-pots were successfully used to test the two retinae presented in the thesis. In both cases, we set $I_{ref} = 17\mu A$ to achieve more precision with lower currents values.

CHAPTER 3

The New Calibration System[1]

3.1 Introduction

Over the last 20 years, a vast amount of neuromorphic VLSI systems have been reported ([6]-[22], just to mention a few) which usually consist of large arrays of special processing pixels. Since pixels have to be of reduced size and power consumption, analog design techniques are used with transistors of small size operating with *nano* amps or less. This yields necessarily high mismatch. Although reported neuromorphic VLSI systems have revealed interesting, powerful, and fast information sensing and processing capabilities, they still have not evolved clearly to specific marketable products. One of the main reasons for this is the unavoidable excessive mismatch that plagues most of the so far reported neuromorphic chips [6]-[22].

To keep mismatch low without increasing transistor sizes nor operating currents, the only known solution is calibration. Some researchers have reported calibration techniques based on floating-gate MOS transistors [25]-[26] in standard CMOS processes. However, these techniques require a special know-how, which makes difficult the path towards marketable products. Recently, some neuromorphic systems with in-pixel RAM based calibration techniques have been reported [14]-[27], which exploit the use of compact current DACs made with calibrable MOS ladder structures [28]. The drawback of this approach is that it uses a one-point calibration principle, which limits the final precision to 3-bit for practical transistor currents and sizes. In this thesis, we present another way of implementing a digitally adjustable MOS, [13], much more compact than a MOS ladder structure, which makes viable a multi-point calibration. Also, we introduce some extra translinear circuitry which enables to sweep operating currents of the calibrated current sources without requiring recalibration. This way, circuit precision degrades smoothly when changing operating currents.

By using a calibration system, the performance of a system can be improved. However, adding some calibration circuitry has some drawbacks. Obviously, the circuit has to been calibrated at least once. This implies time and it is specially important when electronics circuits are fabricated massively. Calibration time has to be reduced as much as possible. Another issue is area. Calibra-

1. This chapter of the thesis is based on the previous reported paper, ([13]).

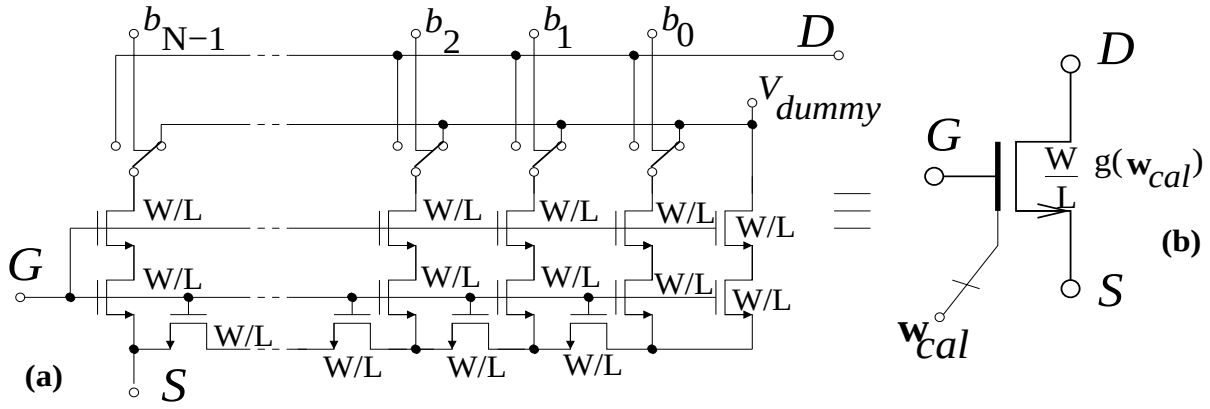


Fig 3.1: Prior reported digitally controlled width MOS transistor. (a) Circuit schematics and (b) symbol

tion circuitry inside pixels has to be simple and reduced. If not, we will degrade the fill factor and we will increase the total area of the circuit considerably.

Another point to consider is how to estimate the maximum admissible mismatch. This is not trivial, because the values of some currents (i.e. diode photo current chapters) are not always known and can vary significantly. Once the mismatch is known, the calibration system can be designed and we have to decide if it makes sense to add it to the pixel. Sometimes, it could be better to increase the transistors area or to use higher bias currents to decrease mismatch effects.

When we design a calibration system for a large neuromorphic array, the number of transistors has to be as low as possible. We cannot try to oversize the system. It is also desirable to implement calibration systems whose effects do not degrade too fast when bias currents are scaled or there is a change in illumination. We would like to calibrate the system in certain average conditions and then preserve the calibration when bias are scaled or there are different settings in the visual scene. It would be also desirable to do the calibration on chip without external circuitry just one time.

One of the targets of the thesis was to design a contrast sensitive retina based on the previous work of Boahen. After studying and simulating the circuit, we checked that one of the main drawbacks was mismatch. For this reason, we tried to develop new calibration techniques that could be used in generic neuromorphic VLSI systems. Not only does calibration improve mismatch, it also makes possible to work with lower bias currents and reduced consumption or to have a bipolar output (as we will discuss in chapter 4) in large arrays of pixels.

In this chapter, we will explain the new calibration system and its more important features. We will present some experimental results. Its specific use in the contrast retina will be discussed in the next chapter.

3.2 MOS Transistor with Digitally Adjustable Length

Previously reported in-pixel RAM based calibration circuits [14]-[27] were based on the use of MOS ladder structures [28]. Figure 3.1 shows the circuit of an equivalent *N*-bit digitally adjustable MOS transistor. This circuit is composed of $3N+1$ unit transistors of size *W/L* and *N* switches. Because of the current splitting principle in MOS ladder structures [32] each vertical branch drives a current half of the previous one, except for the last one. The circuit in Figure 3.1(a) between terminals *D, G* and *S* behaves (at DC) identically to a MOS transistor, as in Figure

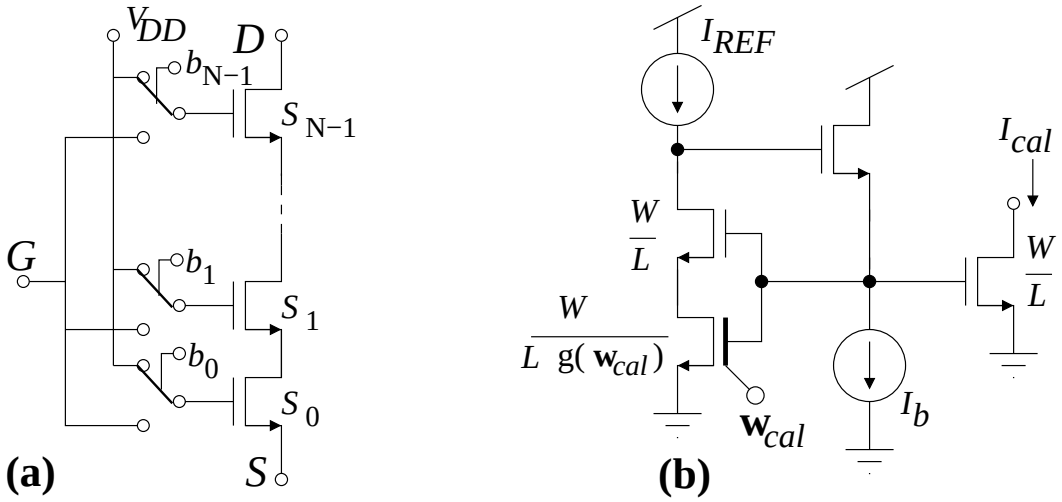


Fig 3.2: (a) Schematics of the digitally controlled length MOS transistor. (b) Application to a calibration current source.

3.1(b), of length L and whose equivalent width is digitally adjustable through the N-bit digital word $\boldsymbol{w}_{cal} = \{b_{N-1}b_{N-2}\ldots b_2 b_1 b_0\}$. The equivalent width is given by $W_{eq} = (W/2) \times g(\boldsymbol{w}_{cal})$, where

$$g(\boldsymbol{w}_{cal}) = \sum_{i=0}^{N-1} \frac{b_i}{2^{N-1-i}} \tag{3.1}$$

Digital word $\boldsymbol{w}_{cal}$ is stored on in-pixel static RAM cells at startup. The optimum calibration words are obtained from a calibration procedure followed the first time the chip is used [14]-[27]. From eq. (3.1) we can see that the equivalent width can be adjusted between 0 and $2-(1/2^{N-1})$ in steps of $1/2^{N-1}$.

To develop calibration circuitry for large neuromorphic arrays, we developed a new approach to digitally adjust the size of a MOS transistor using a more compact circuitry [13]. Figure 3.2(a) shows the schematics of the proposed circuit. There are N transistor segments between terminals D and S. Each segment is either enabled by connecting its gate to terminal G, or disabled by connecting its gate to V_{DD}. Transistor sizes can be for example $S_{N-1} = W/L$, $S_{N-2} = 2W/L$, ... and $S_0 = (2^{N-1})W/L$. This can be implemented physically by using unit transistors of size W/L (one for S_{N-1}, two in parallel for S_{N-2}, ... 2^{N-1} in parallel for S_0). This way, each segment would be equivalent to a transistor of size $S_i = W/(L/2^{N-i-1})$. Consequently, the digitally adjustable transistor in Figure 3.2(a) would be equivalent to one of width W and digitally adjustable length $L_{eq} = L \times g(\boldsymbol{w}_{cal}) = L \times \sum_{i=0}^{N-1} \frac{b_i}{2^{N-1-i}}$. This transistor can be used as part of a current mirror[1], as shown in Figure 3.2(b), to provide a calibration current

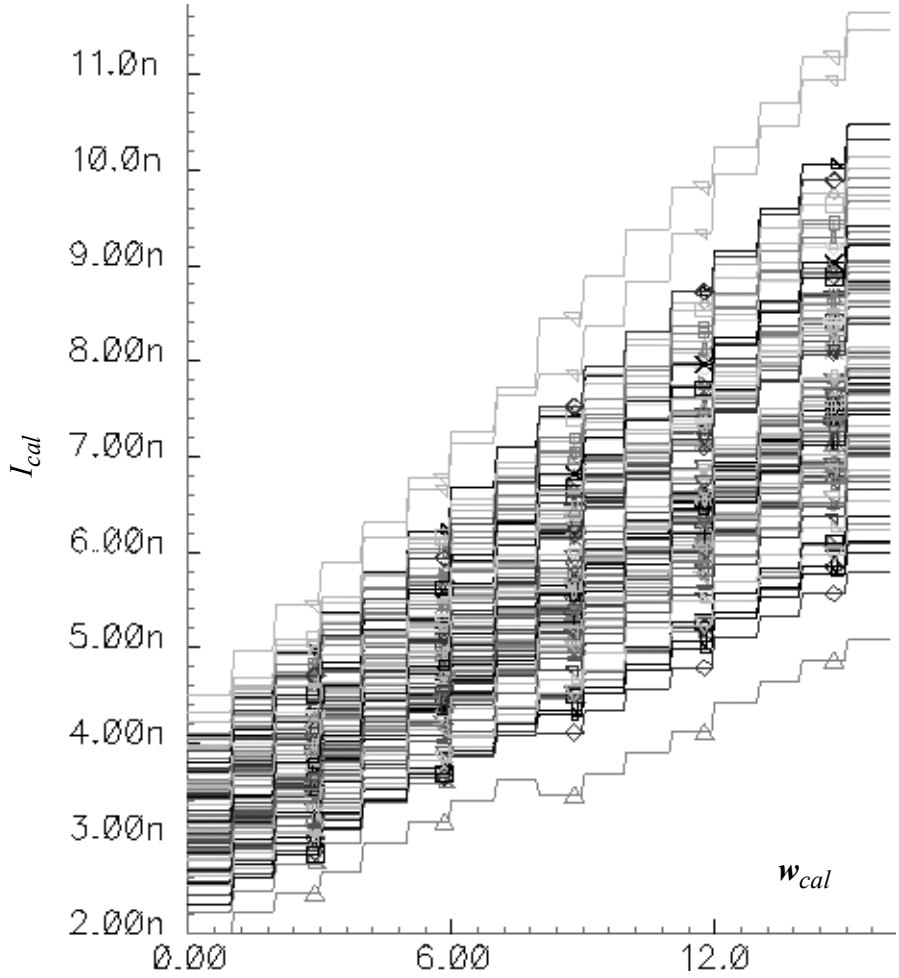


Fig 3.3: Montecarlo simulation (with 100 iterations) of the circuit in Figure 3.2(b), using a 4-bit digitally controlled length MOS.

$I_{cal} = I_{REF} \times (g(\boldsymbol{w}_{cal}) + 1) = I_{REF} \times \left(\sum_{i=0}^{N-1} \frac{b_i}{2^{N-1-i}} + 1 \right)$. Figure 3.3 shows the simulated stairs of I_{cal} as function of $\boldsymbol{w}_{cal}$ (using a 4-bit digitally-controlled-length MOS) with $I_{REF} = 3nA$, using unit MOS sizes of $1\mu/4\mu m$, and models for a $0.35\mu m$ standard CMOS process. Figure 3.4(a) shows I_{cal} as function of I_{REF} before calibration, with $\boldsymbol{w}_{cal} = 15$ for each of the 100 simulated Montecarlo iterations. The mismatch at $I_{REF} = 3nA$ is $\Delta I_{cal}/I_{REF} = 110\%$ and at $I_{REF} = 1pA$ is 130%. Using the results in Figure 3.3 (for $I_{REF} = 3nA$), one can compute for each Montecarlo iteration the optimum value of $\boldsymbol{w}_{cal}$ for minimum spread at I_{cal}. Once setting this optimum set of values for $\boldsymbol{w}_{cal}$, the resulting I_{cal} as function of I_{REF} is shown in Figure 3.4(b). Now the mismatch at $I_{REF} = 3nA$ has been reduced to 4% (4.6 bits).

1. Here we use a sub-pico-ampere current mirror topology [36], since we want to use eventually I_{REF} values down to the *pico ampere* range [27].

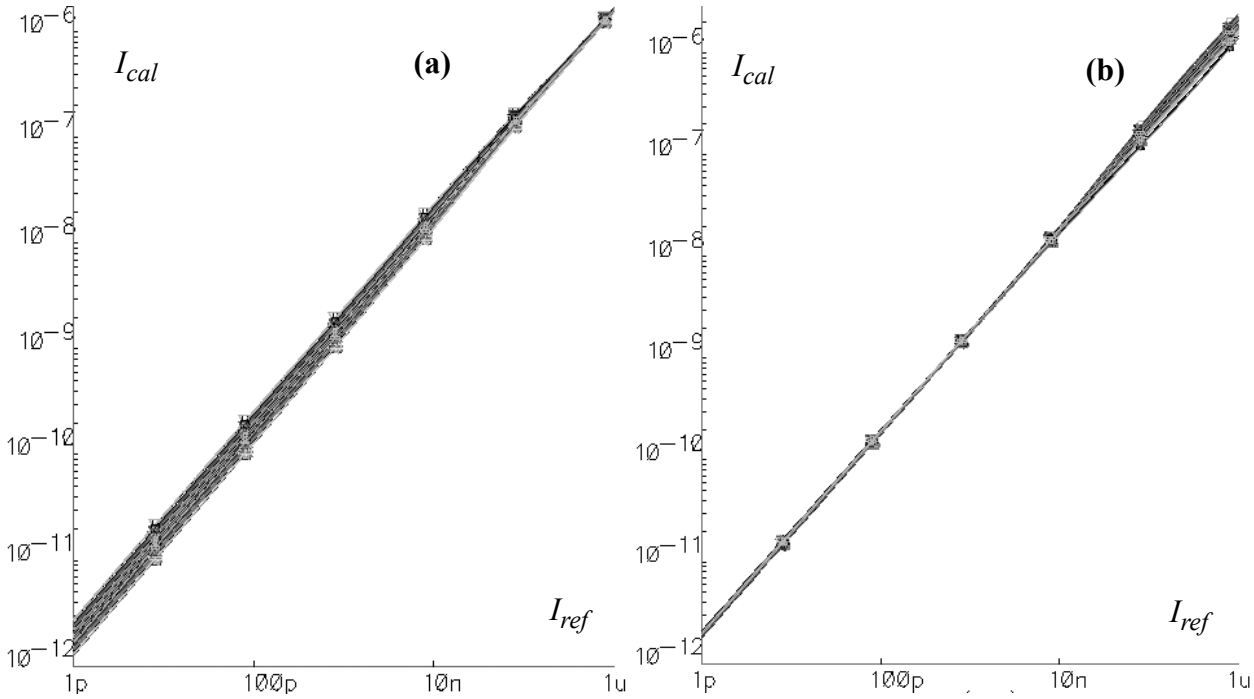


Fig 3.4: Montecarlo simulation results for the circuit in Figure 3.2(b) when sweeping I_{ref}. (a) Before calibration with w_{cal}=15 for all Montecarlo iterations. (b) After calibration with optimum w_{cal} for each iteration.

From a practical point of view, it is not efficient to follow the previous unit transistor based sizing strategy. Note that transistor of size S_{N-1} is the most critical for mismatch (since this segment contributes the largest *L*). However, it uses only one unit transistor, while transistor of size S_0 is the least critical for mismatch and uses 2^{N-1} units. In practice it is more efficient to use one single transistor for each MOS segment and adjust its size to have a similar effect. Furthermore, from a statistical point of view, since mismatch plays an important role, the steps of the final stair cases will not be all equal. The maximum step heights will limit the final calibration capability. Therefore, it is important to minimize this maximum possible step height. To do this, the nominal stair case should be designed with some intentional 'down-steps', so that when mismatch introduces random variations the extra redundancy compensates for eventual large up-steps. Figure 3.5, for example, shows Montecarlo simulation results of a 5-bit structure that uses one single transistor per segment and has intentional down-steps. Simulated transistor sizes were {2/3, 2/1.8, 2/1.8, 2/1, 2/0.7}.

3.3 Translinear Circuits for Tuning

The calibration technique shown in Figure 3.2 requires to recalibrate all circuits when there is a global change in the operating current I_{REF}. In practice, it is desirable to allow a change in the operating current I_{REF} without requiring recalibration. Note that, all transistors introduce mismatching and calibration compensates for the combination of all mismatches of all transistors. The mismatch introduced by each transistor is dependent on its operation current and bias conditions. To have calibration less sensitive to bias conditions one should use topologies that change

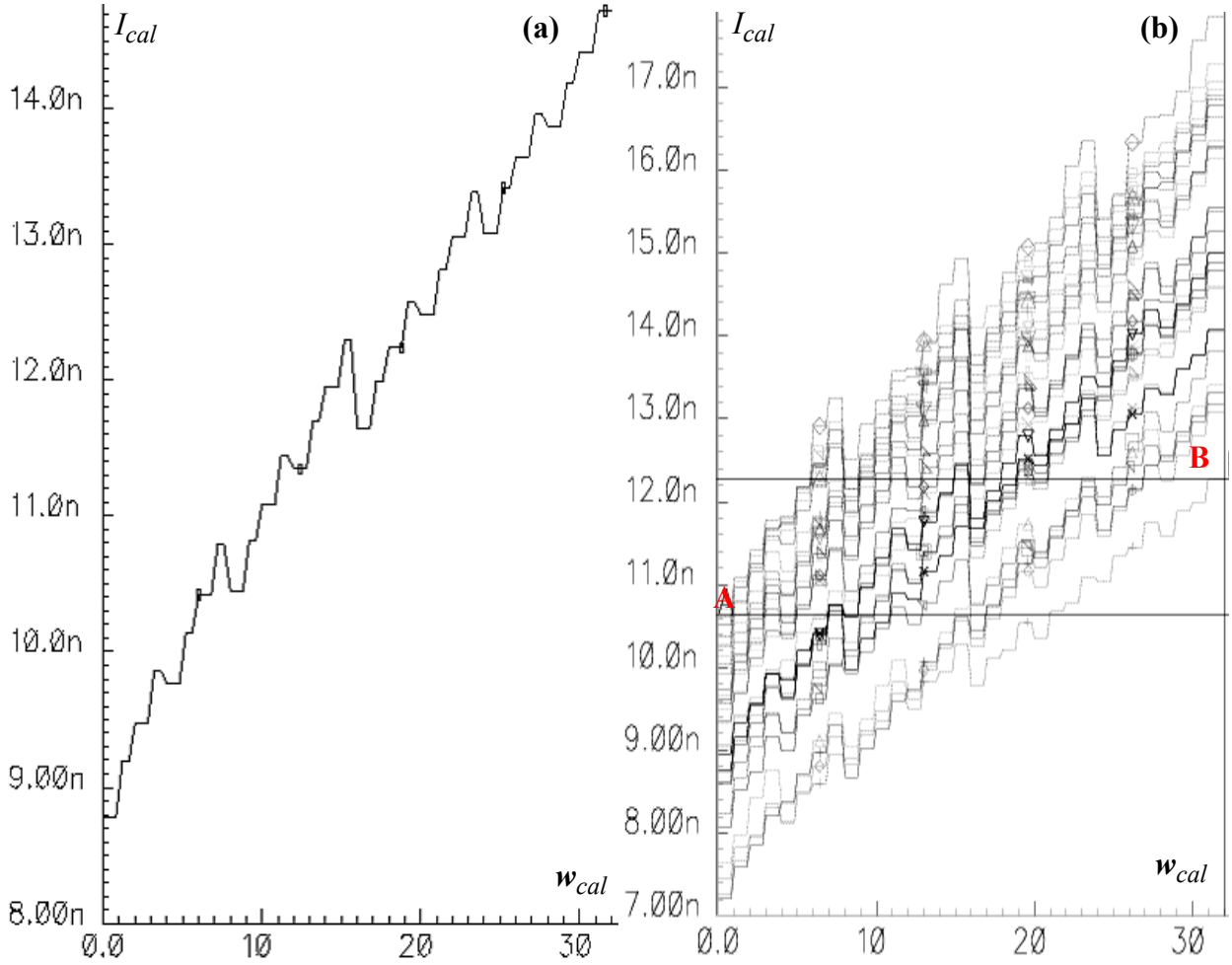


Fig 3.5: Example simulation of a 5-bit digitally controlled length MOS with one transistor per segment and intentional down-steps. (a) Nominal mismatch-less simulation. (b) Monte Carlo simulation with 100 iterations.

bias conditions for as few transistors as possible. To achieve this we introduce tunable translinear circuits, which will allow us to keep fixed bias currents for some transistors, including the digitally-controlled-length ones. This is shown in Figure 3.6. The circuitry comprised by broken lines is replicated once per pixel, but the rest is implemented only once at the periphery. Transistors M_1 to M_4 form a translinear loop, thus $I_{oi} = I_1 I_2 / I_{i3}$. Local current I_{i3} is mirrored from the peripheral global current I_3, through a current mirror with a local digitally-controlled-length MOS. To achieve a factor 2 calibration range, we include two transistors in series for this current mirror output. One of fixed size $W/2L$ and the other calibrable. Consequently, $I_{i3} = I_3/(2 + g(w_{cal_i}))$ and

$$I_{oi} = \frac{I_1 I_2}{I_3}(2 + g(w_{cal_i})) \quad \text{, with } g(w_{cal_i}) = \sum_{i=0}^{N-1} \frac{b_i}{2^{N-1-i}} \tag{3.2}$$

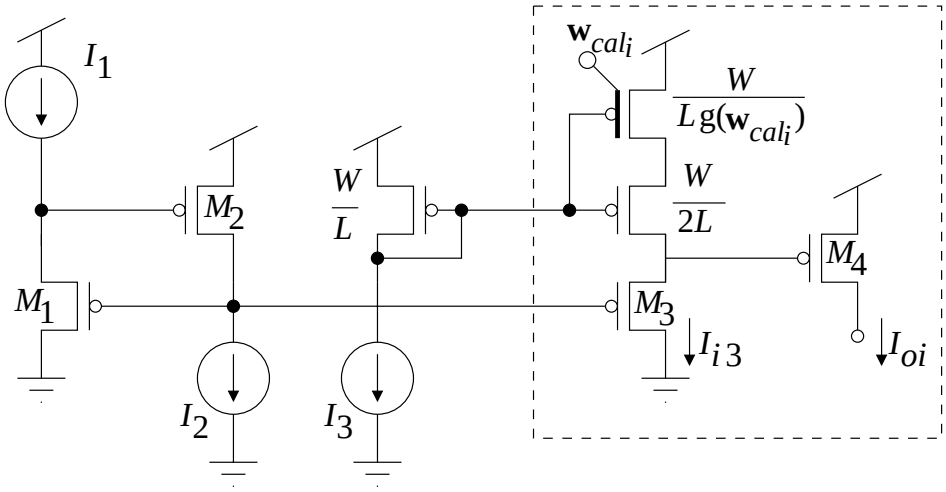


Fig 3.6: Translinear circuit for tuning operating range of calibration circuit.

With this circuit, one can maintain (after calibration) constant currents I_3 (and I_{i3}) and I_1, while tuning I_2 globally to scale up or down all local currents I_{oi}.

3.4 How to Optimize the Calibration Ranges.

For calibration, the goal is to find the optimal horizontal line that cuts through all stairs and produces the minimum dispersion among all stairs. Note in Figure 3.5(b) points 'A' (top value of left side) and 'B' (bottom of right side). If 'B' is below 'A' the maximum dispersion after calibration will be high, because there will be no horizontal line cutting all stairs. If 'A' is below 'B' it is possible to find for each stair a value close enough to the desired horizontal line cutting all stairs. For optimum calibration it is desired that 'A' be close to 'B', so that final calibration words may spread over the whole range. The resulting relative position of points 'A' and 'B' depends on the resulting mismatch distribution of the array and the resulting process corner of the sample. One can design the nominal case to have 'A' as close as possible to 'B', but then many fabricated samples will result with 'A' higher than 'B' yielding poor calibration capability. On the other hand, if one designs the nominal case for 'A' conservatively lower than 'B', then many samples will not take advantage of all their bits for calibration, resulting in reduced calibration capability. Consequently, in practice, it will be desirable to be capable to adjust the relative positions of points 'A' and 'B' electronically. For this, we have implemented two different global optimization strategies. In the first strategy, shown in Figure 3.7, two digitally-controlled-length transistors are used. One of them is adjusted locally, as in Figure 3.6, but the other is adjusted globally. Thus all gates of its transistor segments (see Figure 3.2(a)) are shared by all pixels and controlled from the periphery. As a result,

$$I_{oi} = \frac{I_1 I_2}{I_3}(g(\boldsymbol{w}_{adj}) + g(\boldsymbol{w}_{cal_i})) \tag{3.3}$$

Figure 3.8 shows the resulting simulated staircases for 3 different values of global control word $\boldsymbol{w}_{adj}$. For one extreme ($\boldsymbol{w}_{adj} = 31$, as in Figure 3.8(a)) 'A' is above 'B', and the array has very

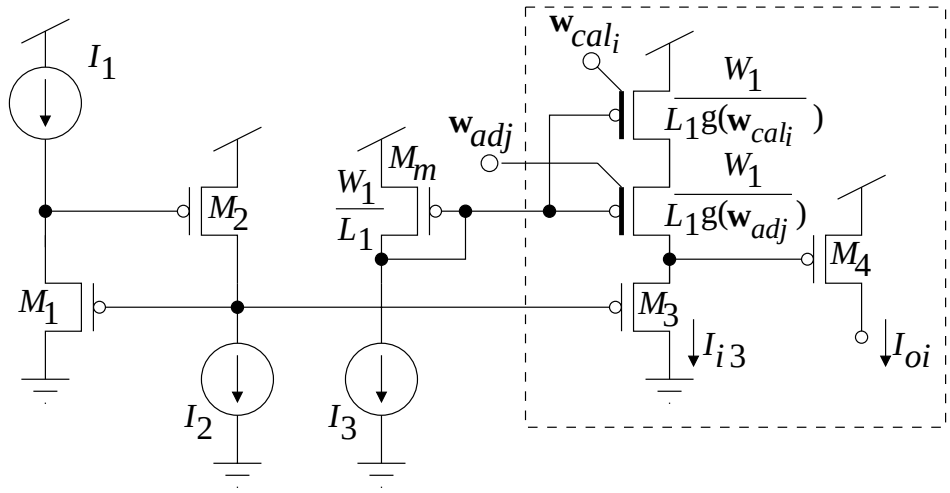


Fig 3.7: First strategy for optimizing calibration range.

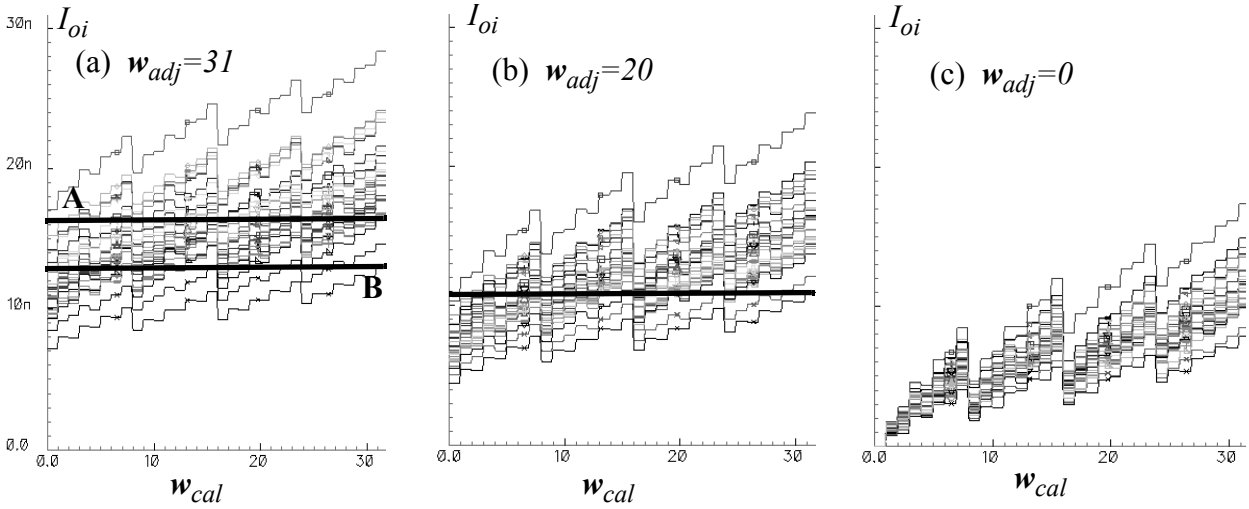


Fig 3.8: Simulation results for first strategy. (a) Simulated stairs for w_{adj}=31, (b) for w_{adj}=16, (c) and w_{adj}=0. Vertical scale is the same for the three graphs.

poor calibration capability. For the other extreme ($w_{adj} = 0$, as in Figure 3.8(c)), 'A' is at the bottom and the horizontal lines cut only a reduced range of the stairs, thus reducing significantly the available number of bits for calibration. The optimum solution is an intermediate one, in this case $w_{adj} = 20$ as in Figure 3.8(b), which sets points 'A' and 'B' to be close. The optimum value of w_{adj} is sample dependent.

The second global optimization strategy is shown in Figure 3.9. Here the translinear circuit has been replicated twice, so that there are two of such translinear circuits in parallel. One of them uses local calibration through local digital control word w_{cal_i}. The other is adjusted globally and only the output transistor M_x of its translinear set is replicated once per pixel. This allows for a larger size of this transistor, and consequently less mismatch. The purpose of the locally cali-

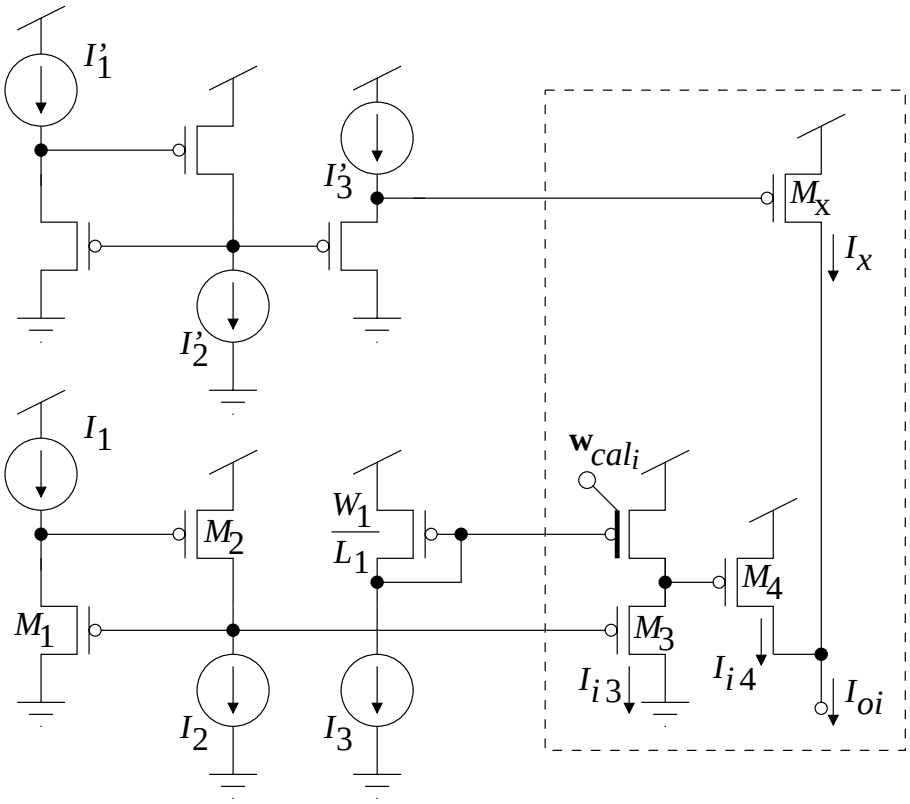


Fig 3.9: Second strategy for optimizing calibration range.

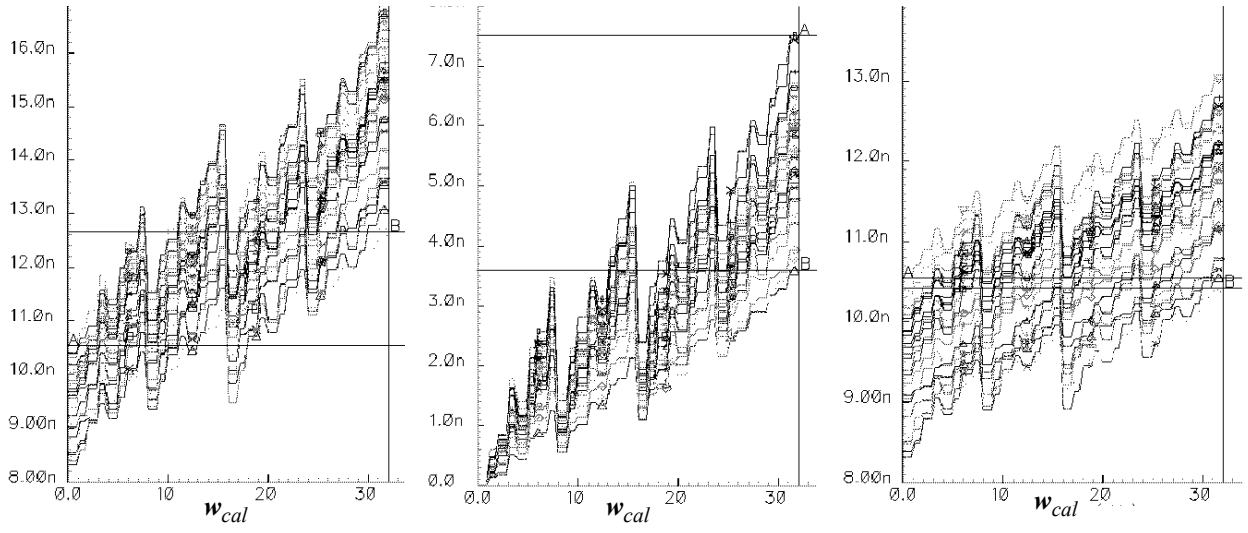


Fig 3.10: Simulation results for second strategy. (a) For all bias currents equal to 10*nA*, (b) detail of the bottom calibrate subcircuit I_{i4}, and (c) results for turning bias currents I_i down to 4.5*nA*.

brated translinear circuit is to compensate for the mismatch at M_x. Figure 3.10 shows simulation results for this circuit. In Figure 3.10(a) all peripheral bias currents I_k and I_k' were set to $10nA$.

The result is 'A' lower than 'B' and a reduced range for the calibration words. Figure 3.10(b) shows the contribution of only the bottom locally adjustable subcircuit (I_{i4} in Figure 3.9). Note that for $w_{cal_i} = 0$ the bottom circuit does not add current to I_{oi}. Consequently, in Figure 3.10(a), for $w_{cal_i} = 0$ the mismatch is produced only by the upper M_x transistors. Note that this left part of the stairs will be fixed if peripheral currents I_k' are maintained fixed. The tuning strategy consists now in scaling peripheral currents I_k until obtaining the optimum situation shown in Figure 3.10(c). In this case, we have set all $I_k = 4.5nA$. After finding the optimum calibration words, the resulting operating point can be scaled by adjusting simultaneously only peripheral currents I_2 and I_2'.

3.5 Experimental Results

In order to characterize the new calibration system, a test prototype microchip was fabricated in a standard $0.35\mu m$ CMOS process. Twenty 5-bit current DACs were fabricated. Ten of them used the first calibration range optimization strategy (Figure 3.7), and the other ten used the second one (Figure 3.9).

Each of the first ten DACs uses five replicas of the circuit in Figure 3.7, one for each bit. The nominal output currents of each (I_{oi}) were adjusted to be binarily scaled. Consequently, at the periphery, we need five groups of current sources $\{I_1, I_2, I_3\}$ and five groups of transistors $\{M_1, M_2, M_m\}$, one for each bit. However, these five groups of peripheral current sources and transistors are shared by all ten DACs.

Each of the second ten DACs uses five replicas of the circuit in Figure 3.9. Again, for each of the ten DACs, the circuitry is replicated five times (one per bit), and the peripheral circuitry (outside broken lines in Figure 3.9) is shared, per bit, by all ten DACs.

The area of the circuit layout inside broken lines is $18 \times 14\mu m^2$ for Figure 3.7 and $7 \times 14\mu m^2$ for Figure 3.9.

Figure 3.11(a) shows the experimentally measured output currents for ten replicas of the circuit in Figure 3.7, when setting $w_{adj} = 0$. Peripheral bias currents were made equal to $I_1 = I_2 = I_3 = 10nA$, and all calibration words w_{cal_i} (i = 1, ... 10) were swept simultaneously from 0 to 31. After repeating this measurement for all possible w_{adj} values, the optimum value for w_{adj} corresponds to the situation where the top left value is closest to the bottom right one. This case is shown in Figure 3.11(b). At this point we can obtain the ten optimum calibration words w_{cal_i} that render the minimum variation. The maximum output current spread obtained under these circumstances is $|\Delta I_{oi}|_{max} = 0.57nA$, which corresponds to 5.7%, at a nominal current of $I_b = 10nA$. If this were the current source controlled by the most significant bit of a current DAC (with $20nA$ maximum range), it would limit the DAC precision to $-\ln(|\Delta I_{oi}|_{max}/2I_b)/\ln 2 = 5.13 \text{bits}$. To verify how calibration degrades when changing bias conditions, we swept I_2 in Figure 3.7 between $100pA$ and $1\mu A$. The maximum current spread among all 10 calibrated current sources is shown in the trace with circles in Figure 3.12. The trace with triangles are measurements obtained before calibration ($w_{cal_i} = 0$, for all i). We can see the 10 samples maintain a precision of 4 bits for currents above $3nA$. Horizontal axis is average of

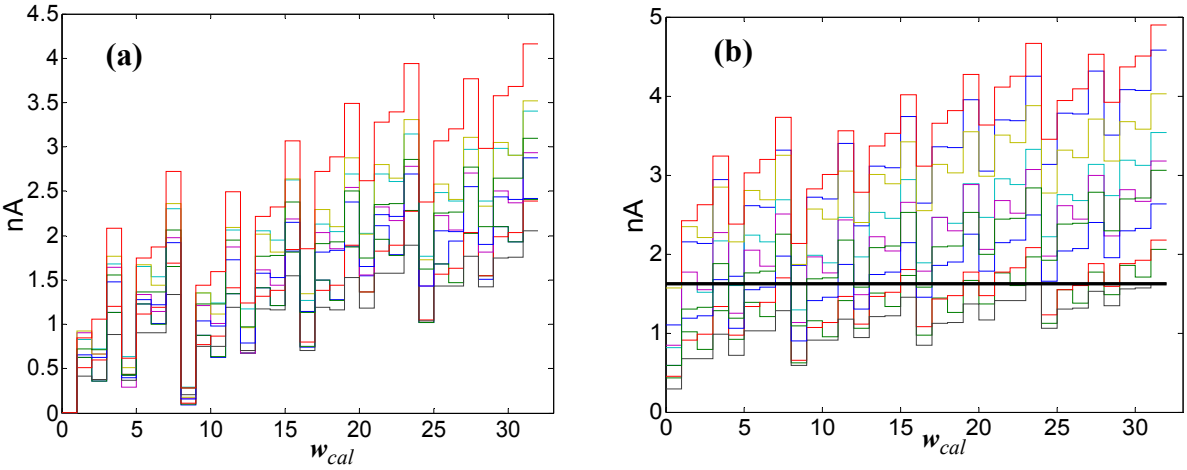


Fig 3.11: Experimentally measured output currents for the circuit in Figure 3.7. (a) for $w_{adj}=0$, (b) for optimum w_{adj}. The horizontal line in (b) is the target value, which is cut/ touched by all 10 traces.

I_{oi} among all ten samples. We also show in Figure 3.12 the resulting precision after calibration obtained through simulations, shown with crosses. Note that it is over optimistic, except for the point at which calibration was done ($10nA$). The reason is that in this particular circuit (Figure 3.7) calibration degrades because of mismatch in MOS transistor slope factor. Mismatch in this parameter is not modelled in our simulator.

In a similar way, Figure 3.13 shows the measured precision before and after calibration of ten calibrable and tunable current sources that follow the approach depicted in Figure 3.9. Note that now the mismatch before calibration is less than in Figure 3.12. This is because now the area used by digitally-controlled-length transistor M_b in Figure 3.7, is available for transistor M_x in Figure 3.9, which can be made larger. With the structure of Figure 3.9 we obtain a much better precision at the calibration point (8.30 bits at $10nA$), but degrades rapidly, specially for high currents. The precision after calibration obtained by simulation, is slightly pessimistic at the calibration point (7.63 bits at $10nA$), but it degrades optimistically as operating current departs from the calibration point.

Figure 3.12 and Figure 3.13 show the matching precision among 10 current sources calibrated at $10nA$. Now we use 5 of these sources, calibrated at $\{10, 5, 2.5, 1.25, 0.625\}nA$, to build a 5-bit current DAC. The matching precision obtained among the 10 fabricated DACs is shown in Figure 3.14 (for the tuning scheme of Figure 3.7) and in Figure 3.15 (for the tuning scheme of Figure 3.9). The DACs were calibrated at 16ºC, and the figures also illustrate the DACs behavior when temperature is changed between 0ºC and 40ºC. We can see that the effect of temperature is not severe for the lower current range, while for higher currents the DACs are almost insensitive to temperature variations.

3.6 Discussion

In this chapter, we have described a new method to calibrate large neuromorphic arrays. The approach is illustrated for current sources operating in the nano ampere range. Two tuning schemes are proposed for sweeping the operating range over four decades. The first one achieves

less precision at the calibration point but degrades more gracefully as operating current is increased. The second one achieves higher precision at the calibration point but precision degrades more as current increases. Test prototypes were fabricated and extensively tested and characterized. As an example application, current DACs of 5-bit resolution and 20*nA* range have been fabricated and characterized.

As it will be explained in the next chapter, the calibration procedure was used to reduce mismatch in the spatial contrast retina. In this case, the scheme of Figure 3.9 was chosen because we were interested in getting the highest precision at a fixed calibrated point without large variations of bias currents.

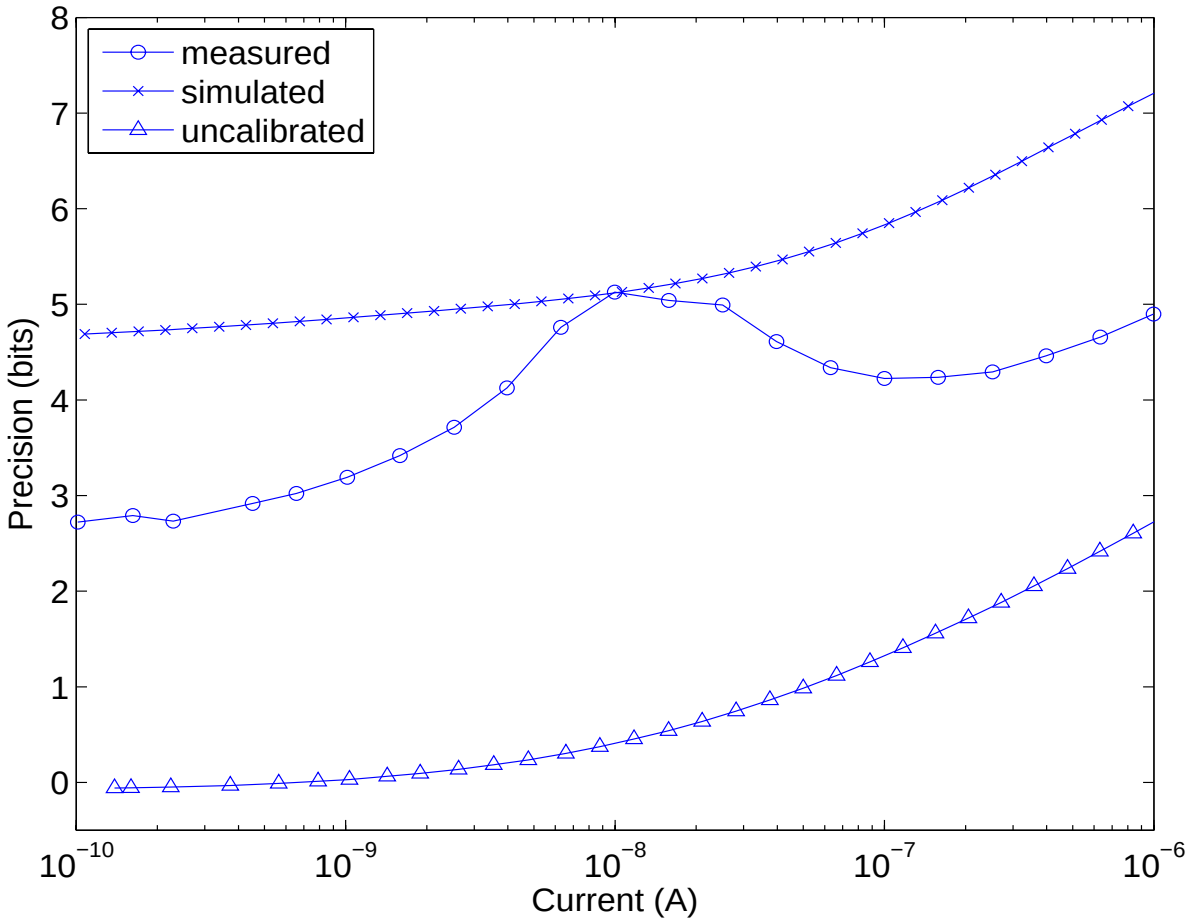


Fig 3.12: Measured precision of calibrate and tunable current source with the approach of Figure 3.7. Trace with circles: measured precision after calibration (with optimum w_{cal_j} for each of the ten current sources). Current sources were calibrated at 10*nA*. Trace with triangles: measured precision before calibration ($w_{cal_i} = 0$ for all current sources). Trace with crosses: precision after calibration, obtained through simulations.

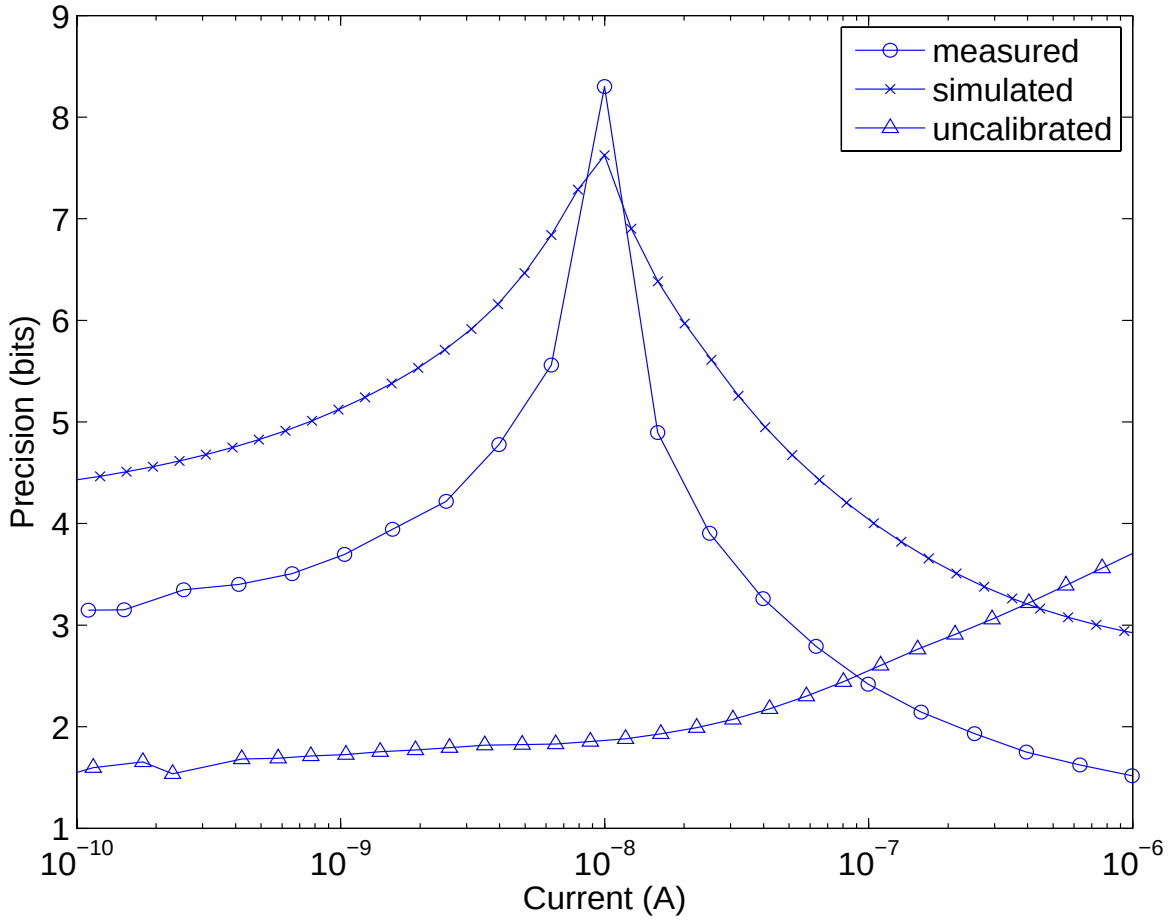


Fig 3.13: Measured precision of calibrable and tunable current source with the approach of Figure 3.9. Trace with circles: measured precision after calibration. Current sources were calibrated at 10*nA*. Trace with triangles: measured precision before calibration ($w_{cal_i} = 0$ for all current sources). Trace with crosses: precision after calibration, obtained through simulations.

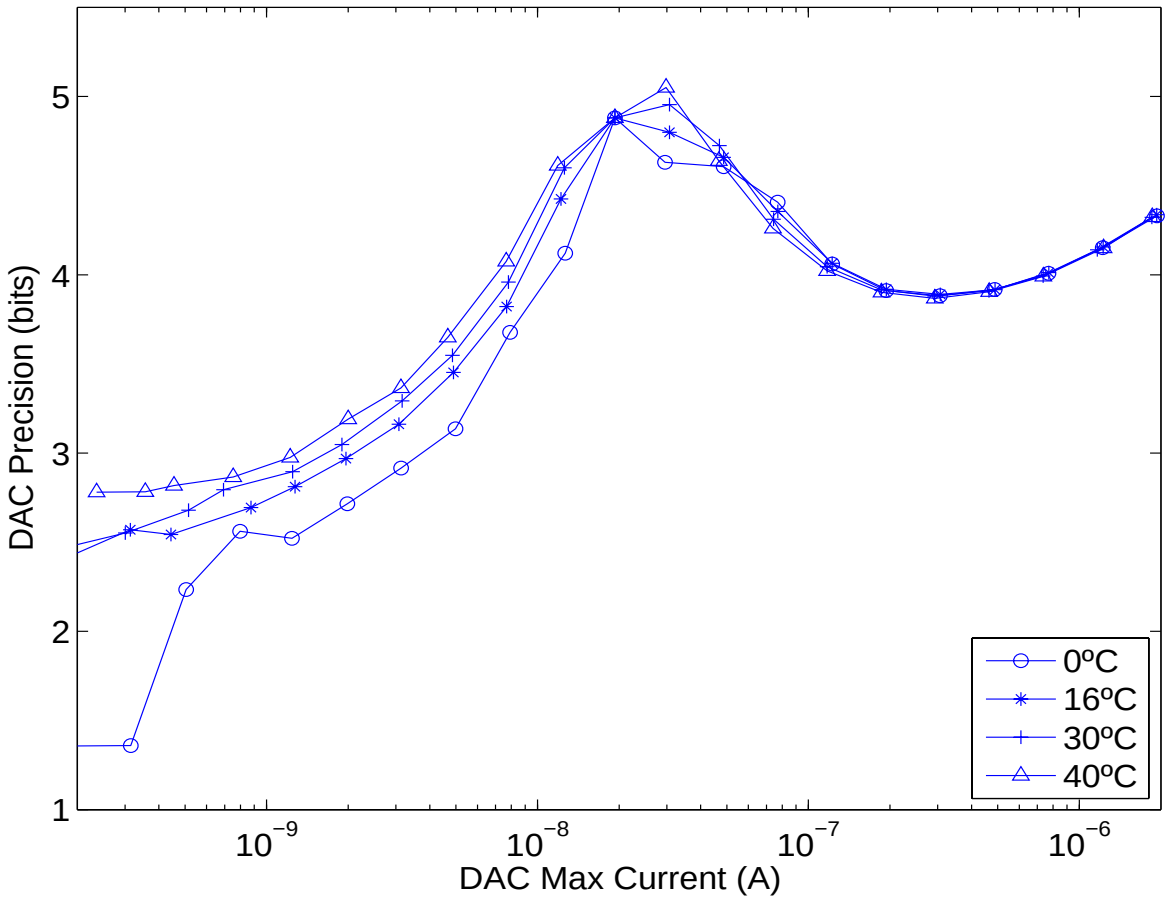


Fig 3.14: Measured Precision for the 5-bit DAC samples that use the first tuning strategy of Figure 3.7. DACs were calibrated with MSB at 10*nA* and at 16ºC. After calibration, precision is characterized sweeping operating current for different temperatures.

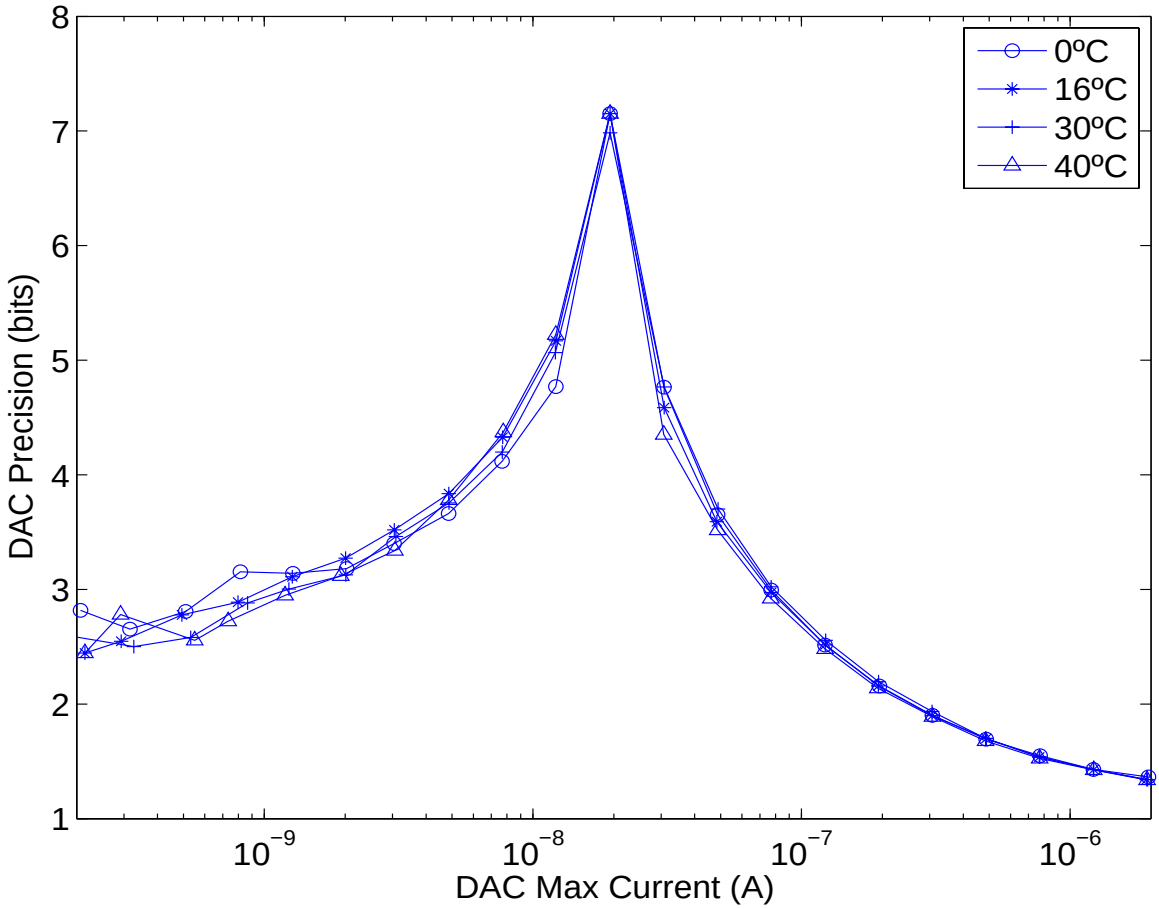


Fig 3.15: Measured precision for the ten 5-bit DAC samples that use the second tuning strategy of Figure 3.9. DACs were calibrated with MSB at 10*nA* and 16ºC. After calibration, precision is characterized sweeping operating current for different temperatures.

CHAPTER 4

The AER Spatial Contrast Sensor[1]

4.1 Introduction

Visual sensors (retinae) are among the first AER modules to be reported since the introduction of the technology [1]-[2]. Spatial contrast AER retinae are of special interest since they provided highly compressed data flow without reducing the relevant information required for performing recognition. As we explained in chapter 1, a variety of AER visual sensors can be found in the literature, such as simple luminance to frequency transformation sensors [19], Time-to-First-Spike (TFS) coding sensors [20]-[16], foveated sensors [39]-[40], more elaborate transient detectors [9]-[24]-[41], motion sensing and computation systems [42]-[43], and spatial and temporal filtering sensors that adapt to illumination and spatio-temporal contrast [21]-[22].

Spike based visual sensors can code their output signals using rate coding or TFS coding. When using rate coding, each pixel is autonomous and continuously generates spikes at a frequency proportional to the signal to transmit (such as luminance or contrast). Under such circumstances, there are no video frames, so that sensing and processing is continuous and frame-free. When using TFS conding, a global system-wide reset is provided and each pixel encodes its signal by the time between this reset and the time of the only spike it generates. Sensing and processing is frame-constraint. However, TFS is a highly compress coding scheme (each pixel generates at the most one spike per frame) and frame time can be dynamically adjusted to an optimum minimum by subsequent processing stages. TFS and subsequent ideas were originally proposed by Thorpe based on neuro physiological and psycophysical experiments [11], and have evolved to practical very high speed image processing software techniques [44].

Computing contrast on the focal plane significantly reduces data flow, while relevant information for shape and object recognition is preserved.

Previously reported plain spatial contrast retinae [45]-[4]-[14] compute a contrast current per pixel $I_{cont}(x, y)$ as the ratio between pixel's locally sensed light intensity $I_{ph}(x, y)$ and a spe-

1. This chapter of the thesis is based on a previous reported paper ([23]). I would like to thank the anonymous reviewers for improving the quality of the paper.

cially weighted average of its surrounding neighborhood computed with some kind of diffusive network

$$I_{cont}(x,y) = I_{ref}\frac{I_{ph}(x,y)}{I_{avg}(x,y)} \tag{4.1}$$

where I_{ref} is a global scaling current. Since this is always positive, let us call it "unipolar" contrast computation, with contrast being computed as the ratio between two photo currents. This yielded circuits where no subtraction operation was required. This was crucial to maintain mismatch (and precision) at reasonable levels. Note that for computing I_{avg} and I_{cont} circuits have to handle directly photo currents, which can be as low as pico-amperes or less. Performing a simple mirroring operation introduces mismatches with errors in the order of 100% [14]. This can be overcome by increasing transistor area, but then leakage currents may become comparable to the available photo currents. Consequently, while handling photo currents, it is desirable to keep complexity at a minimum. Therefore, from a circuit point of view, the way of computing contrast as in eq. (4.1) was very convenient. However, this presents an important drawback: when there is no contrast ($I_{avg} = I_{ph}$) then $I_{cont} \neq 0$. In an AER circuit this means that a pixel sensing no contrast will be sending out information (events) and consuming communication bandwidth on the AER channels. This is contrary to the advantages of AER (where it is expected that only information relevant events will be transmitted) and contrary to the advantages of computing contrast at the focal plane (so that only contrast relevant pixels need to send information). In another work [14], although spatial contrast was computed by eq. (4.1) in the retina, a post-processing with AER (convolution) modules was added to effectively compute the Weber Contrast[2] as the signed quantity

$$I_{cont}(x,y) = I_{ref}\left(\frac{I_{ph}(x,y)}{I_{avg}(x,y)} - 1\right) \tag{4.2}$$

This reduced significantly the data flow (from about 400keps[3] to about 10keps), but also at the expense of reducing the speed response of a pixel by a factor of about 10.

In this chapter, we present a new spatial contrast retina design [23], where the contrast computation follows eq. (4.2). The design is based on the original contrast computation circuit by Boahen [45], which has been improved to overcome its inherent limitations on mismatch, ambient light dependence, and critical controllability. It has also new functionalities. We have signed output. Furthermore, and adjustable thresholding mechanism has been included, such that pixels remain silent until they sense an absolute contrast above the adjustable threshold. The retina also includes an optional global reset mechanism for operation in ambient-light-independent Time-to-First-Spike Contrast Computation Mode. A 32 x 32 pixel test prototype was fabricated in $0.35\mu m$ CMOS technology.

2. Weber Contrast is defined as $WC = (I-I_{avg})/I_{avg}$ for a pixel photo current with respect to its neighbourhood average photo current, or as $WC = (I_1-I_2)/(I_1+I_2)$ between two adjacent pixels or regions. Both expressions are equivalent by making $I = I_1$ and $I_{avg} = (I_1+I_2)/2$.
3. keps stands for "kilo events per second".

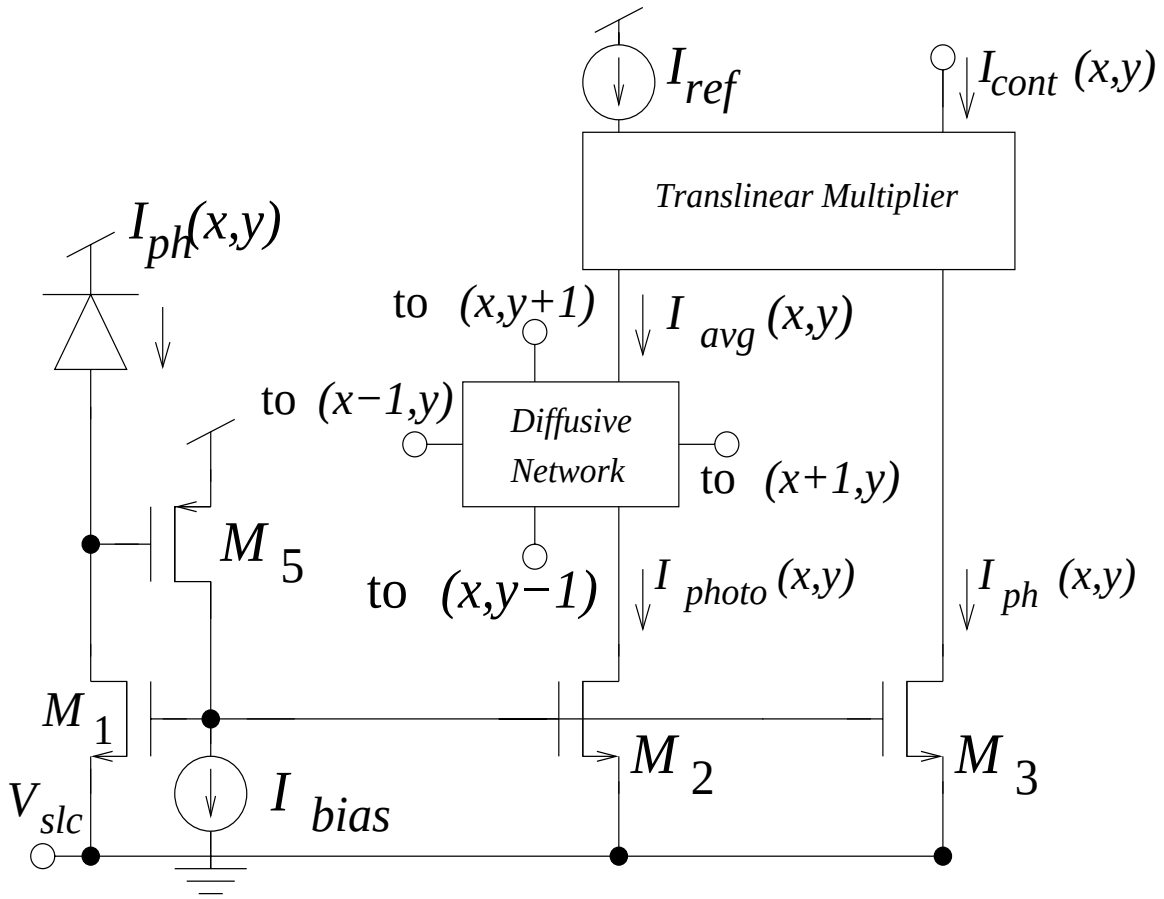


Fig 4.1: Block diagram of pixel in prior unipolar contrast retina.

4.2 Prior AER Mismatch-Calibrated Unipolar Spatial Contrast AER Retina

The AER contrast vision sensor is based on a previous design [14] presented by Costas et al. in 2007. Figure 4.1 shows the basic schematic of the contrast computation circuit used in the previous unipolar spatial contrast retina. A p+/nwell photo diode senses current $I_{ph}(x, y)$ that is replicated twice using a sub-pico-ampere current mirror [38]. The first replica is used in a cascoded diffusive network [46], which implements the discrete approximation of the 2D Laplacian equation

$$I_{ph}(x, y) = \left(1 - \lambda_x \frac{\partial^2}{\partial x^2} - \lambda_y \frac{\partial^2}{\partial y^2}\right) I_{avg}(x, y) \tag{4.3}$$

This equation provides a good spatial average of I_{ph} over neighboring pixels, such that closer pixels contribute more to this average than distant pixels. The second replica of the photo current is fed together with $I_{avg}(x, y)$ to a translinear circuit computing the ratio between both, scaled by

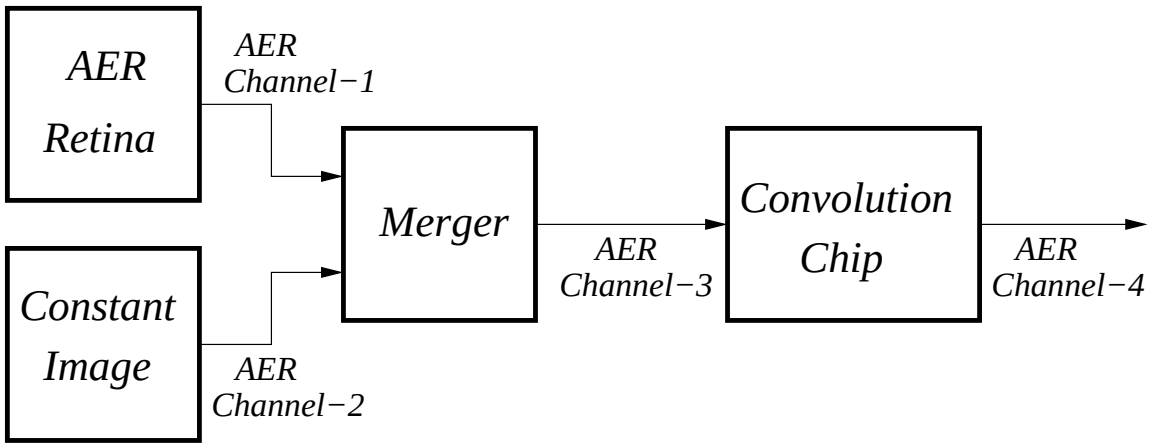


Fig 4.2: Setup used to convert the unsigned AER retina output with DC level to a signed AER stream with no DC level.

reference current I_{ref}. The resulting current $I_{cont}(x, y)$ is thus proportional to a unipolar contrast (as in eq. (4.2)) and is fed to an integrate-and-fire neuron generating periodic spikes with a frequency proportional to $I_{cont}(x, y)$. Scaling current I_{ref} is made locally trimmable for each pixel in order to compensate for all mismatch contributions from the photo diode, current mirror, diffusive network, translinear circuit, and integrate-and-fire neuron. As a result, inter-pixel mismatch contrast computation could be reduced from about $\sigma \approx 60\%$ to $\sigma \approx 6\%$ using 5-bit pixel registers to control $I_{ref}(x, y)$. Pixel complexity was kept relatively simple (104 transistors + 1 capacitor) thanks to the unipolar nature of the contrast computation, and the whole pixel could be fit into an area of $58\mu m \times 56\mu m$ in a $0.35\mu m$ CMOS process. The main drawback is that pixels with no contrast would generate output events at a constant rate proportional to I_{ref}. To overcome this, the 4-AER-module system shown in Figure 4.2 was assembled to compute effectively a bipolar contrast as in eq. (4.2). A uniform image AER flow with negative sign bit was merged with the retina AER flow and fed to an AER convolution chip [27] configured to operate as an array of signed (bipolar) integrators. As a result, the background DC component in eq. (4.1) was removed, yielding a computation equivalent to that in eq. (4.2). However, as a backside effect, the effective firing rate of a pixel at the output channel was reduced by a factor of 8, thus diminishing its speed response. In the design presented in this chapter, this is solved by performing all the bipolar contrast computation at the sensor chip using an improved version of Boahen's original biharmonic contrast computation circuit.

4.3 Boahen Spatial Contrast Pixel

In the spatial contrast sensor presented in this thesis all these drawbacks are solved by performing all the signed-spatial-contrast computation at the sensor chip using an improved version of Boahen's pixel circuit [46]. The continuous approximation of Boahen's pixel circuit, shown in Figure 4.3 solves approximately the following equations [46]

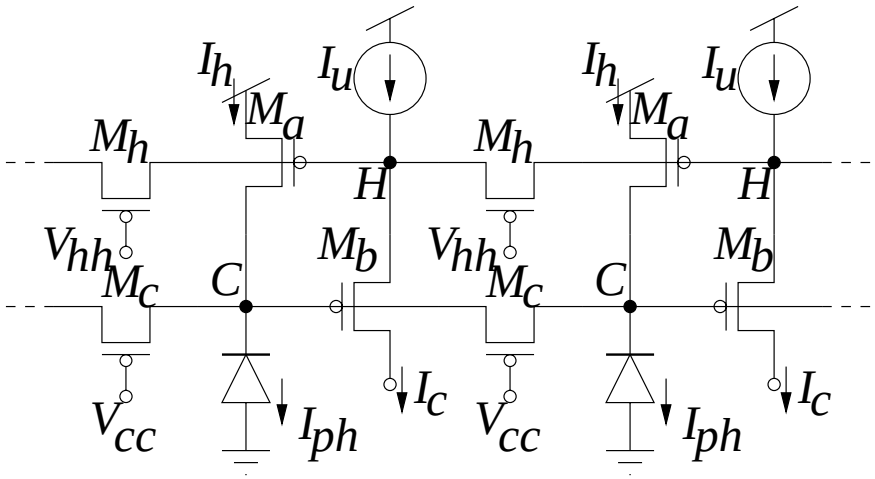


Fig 4.3: Boahen original contrast computation circuit.

$$\begin{aligned} I_h(x,y) &= I_{ph}(x,y) + a\nabla^2 I_c(x,y) \\ I_c(x,y) &= I_u - b\nabla^2 I_h(x,y) \end{aligned} \tag{4.4}$$

Solving for I_h results in the biharmonic equation used in computer vision to find an optimally smooth interpolating function of the stimulus I_{ph}. Consequently, the output $I_c(x,y)$ is the second order spatial derivative of the interpolation I_h according to the bottom eq. (4.4). Since the interpolation is a spatially integrated version of the stimulus, I_c can be interpreted as a version of a first order derivative of the stimulus, therefore, spatial contrast. This can also be understood with the help of Figure 4.4. The top trace shows a step stimulus I_{ph} and its spatial average (I_{avg} or I_h). The center trace shows the contrast computation as I_{avg}/I_{ph} , and the bottom trace shows the contrast computation as the second order spatial derivative of I_h . Both are equivalent. According to the bottom eq. (4.4), I_c includes a DC term I_u .

Obtaining an expression for each pixel output from equations (4.4) is not trivial and intuitive. Each pixel output current depends on all the pixels of the retina. An approximation given by Boahen [46] for each pixel output from equations (4.4) is

$$I_c(x,y) = I_{ref}\frac{I_{ph}(x,y)}{\langle I_{ph}\rangle + I_{ph}(x,y)} \tag{4.5}$$

where $\langle I_{ph}\rangle$ is the average photo current in the pixel's surrounding neighborhood. Obviously, closer neighborgs will have higher influence on the output and far pixels will not have a noticeable effect. By adjusting I_{cc}, neighbors influence can be controlled.

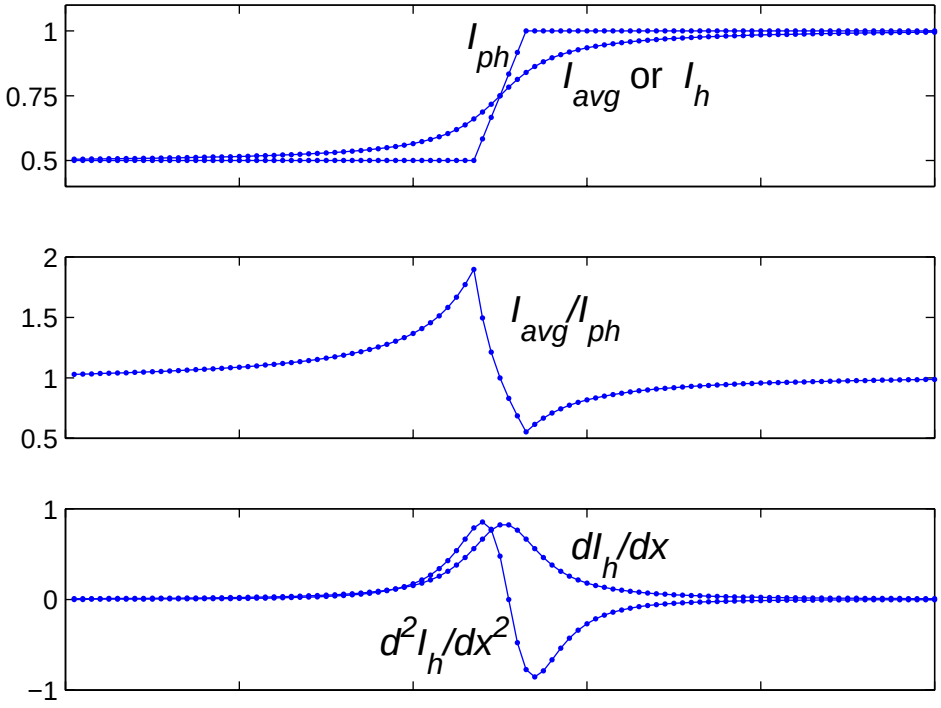


Fig 4.4: Interpretation of spatial contrast computations.

The original circuit implementation of this model suffered from a series of drawbacks. First, mismatch was comparable to output signal. Second, output signal would degrade for the same contrast stimulus when changing lighting conditions. Third, bias voltages V_{cc} and V_{hh} in Figure 4.3 had very narrow and critical tuning range. All three drawbacks have been improved with the present implementation.

4.4 Improved Signal-Spatial-Contrast Pixel

Figure 4.5 provides an overall block diagram, indicating the signals interchanged between blocks. The pixel contains three main parts: (1) the photo sensing and contrast computation part, (2) the calibration part, which provides the ambient light independent contrast current I_{cont}; (3) the integrate-and-fire part, which includes refractory circuitry, thresholding, and TFS mode; (4) and the pixel AER communication circuitry that sends out events to the periphery. Let us now describe each one.

A. Compact Calibration Circuit

As we explained previously, one of the main drawbacks of prior AER contrast retinae was high mismatch. To overcome that problem, we reduce pixel mismatch by introducing calibration. One dominant source of mismatch is the DC component I_u in eq. (4.4). Since this current is set

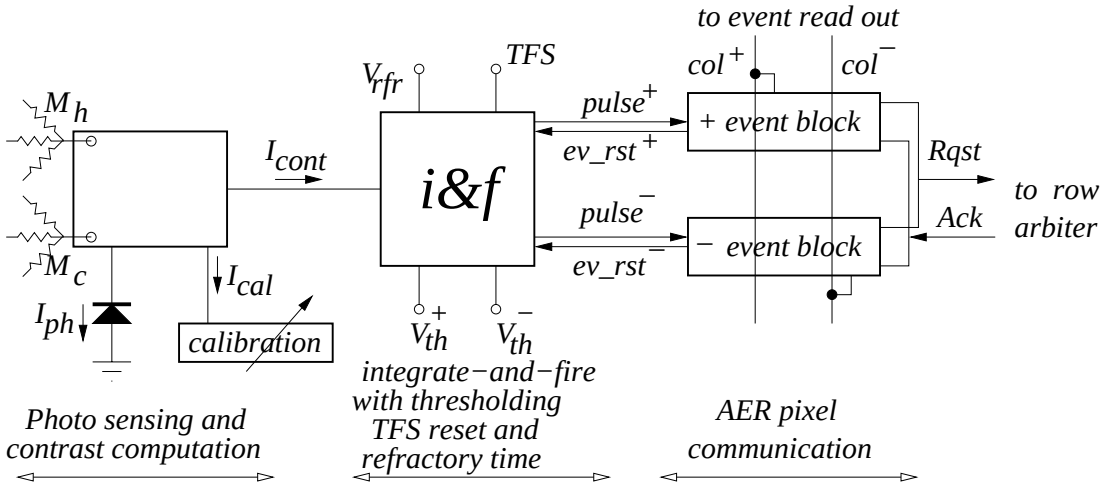


Fig 4.5: Overall block diagram of the contrast pixel.

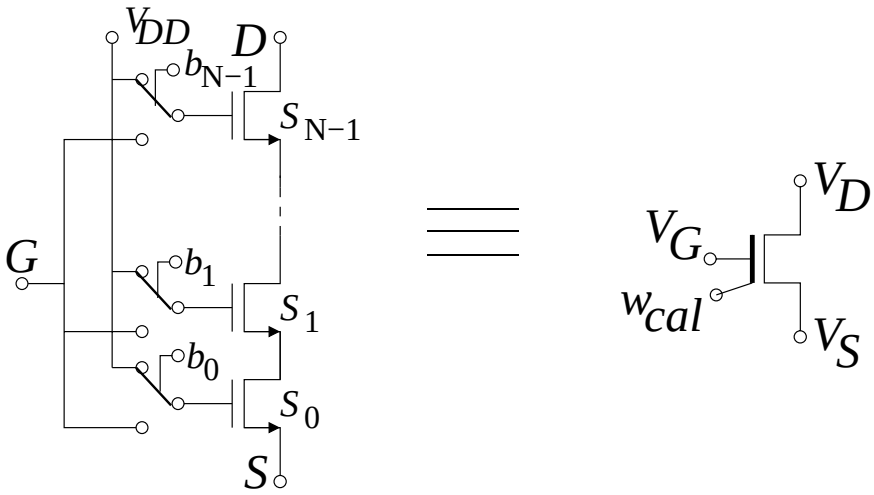


Fig 4.6: Digitally controlled length MOS used for calibration and its symbol.

constant, independent of lighting conditions, we can directly subtract it with a trimmable current source. The output current will thus be directly the bipolar contrast current we were looking for. To implement the trimmable current source, we follow the explained compact circuit based on series transistors association [13] previously explained. Figure 4.6 shows the basic principle behind this circuit. Each switched MOS operates as a segment of an effective longer MOS whose length is controlled digitally by switching individual segments from ohmic to saturation, and vice versa. The key consists in making each segment to contribute approximately as a power of 2 to the total length. The digital control word $\boldsymbol{w}_{cal} = \{b_{N-1}\ldots b_1 b_0\}$ sets the state of the switches. As a result, the effective length is digitally controlled as in a digital-to-analog conversion. On the right of Figure 4.6 we show the symbol of a digi-MOS (digitally-controlled MOS) which we use to represent the circuit on the left.

To calibrate the circuit, we use the second approach described in chapter 3, [13]. This scheme has higher precision at the calibration point but calibration degrades faster when bias cur-

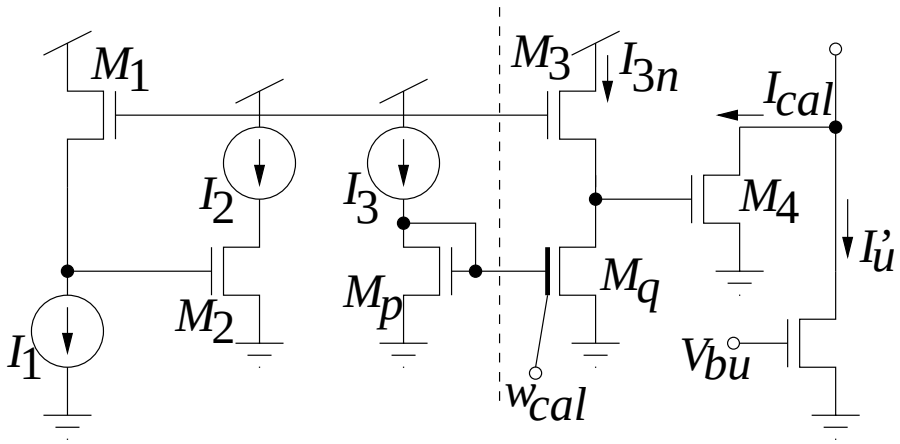


Fig 4.7: Translinear tuning circuit used to calibrate the contrast retina.

rents are scaled. In our particular case, we were interesting in achieving the highest precision as the bias current I_u value will not be scaled significantly after calibration.

Figure 4.7 shows the circuitry used to subtract the DC component I_u of the contrast current. Transistors to the left of the dashed line are shared by all pixels and are located at the chip periphery, while those to the right are replicated for each pixel. Current I_u' sets the subtracting DC level (while also introducing mismatch), while I_{cal} is adjusted to a tuning range covering the interpixel mismatch. Transistors M_{1-4} form a translinear loop [46], thus $I_{cal} = I_1 I_2 / I_{3n}$. And I_{3n} is a mirrored version of I_3 by transistors M_p and M_q. Transistor M_q is digi-MOS of Figure 4.6. Consequently, I_{cal} is proportional to the pixel calibration word $\boldsymbol{w}_{cal}(x, y)$, which is stored on in-pixel static RAM latches loaded at start-up. Note that current I_{cal} could have been generated directly by current mirror M_p-M_q. However, in this case, if one wants to scale $\{I_u, I_{cal}, I_u'\}$ globally (to adjust the retina output frequency range) the circuit would change the current through the calibration branch containing M_q. On the contrary, with the circuit in Figure 4.7 one can scale $\{I_u, I_{cal}, I_u'\}$ while keeping the calibration branch current I_{3n} (and I_3) constant, and scale through peripheral currents I_1 and/or I_2. This way, calibration degrades less when tuning the output frequency range.

B. Photo Sensing and Contrast Computation

Figure 4.8 shows how Boahen's contrast computation circuit has been modified to include a current biasing scheme for controlling the original voltages V_{cc} and V_{hh} in Figure 4.3. This way, gate voltages V_{cc} and V_{hh} tend to follow voltage excursions at nodes 'C' and 'H'.

The first advantage of this is that biasing will adapt to ambient light conditions. For example, if all photodiode currents are scaled up/down by the same factor, the voltage at all nodes 'H' will follow it logarithmically. Since I_u is constant, the voltage at node 'C' will thus also follow the same shift. Since bias currents I_{hh} and I_{cc} are kept constant, the gate voltages of transistors

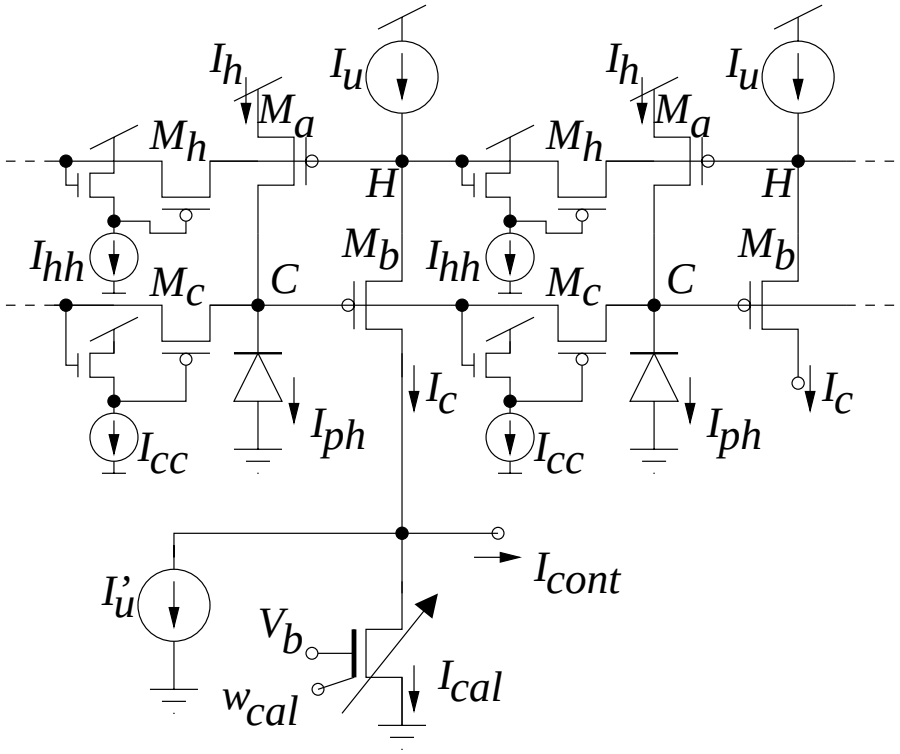


Fig 4.8: Detail of photo sensing and contrast computation circuit.

M_h and M_c will thus follow also this same global voltage shift, adapting themselves to the global light change.

The second advantage of this current biasing scheme is that it attenuates mismatch. After doing careful mismatch analysis and identifying the main sources of mismatch for this circuit, one can find out that transistor M_a and current I_u are the dominant sources of mismatch. This can be understood as follows. Mismatch in I_u goes directly into the DC offset of I_c, which will be calibrated by I_{cal}. Mismatch of M_b is less critical because its inter-pixel gate voltage (node 'C') variability affects the bottom diffusive grid and the computation of the average current I_h. Thus its variability impact is attenuated by the average computation. However, M_a mismatch (V_{gs} variation of M_a) changes directly the source voltage of M_b, affecting directly the gain of contrast output (coefficient 'b' in eq. (4.4)), whose effect is not directly calibrated by I_{cal}. Consequently, M_a needs to be sized to minimize mismatch. The effect of I_u will be compensated by calibration, and the effect of M_a will be attenuated by the current biasing scheme. Note that mismatch in all M_a transistors will introduce random voltage variations at nodes 'H' and 'C'. These variations will be transformed into random lateral currents through transistors M_h and M_c. The random currents through M_h will be collected by output current I_c and can be compensated by calibration. However, random currents through M_c transistors operate as if they were generated by the photo diodes. Thanks to the current biasing scheme, an increase in 'C' will increase the

gate voltage of the new bottom NMOS transistor, increasing its source voltage, thus increasing the gate voltage of M_c, which will reduce the lateral random current. A similar effect will be happening for transistors M_h.

Finally, the third advantage is a more robust means for biasing the lateral transistors. In the original scheme, voltages V_{cc} and V_{hh} suffered from a very narrow and critical tuning range (about $100mV$ or less). Now, bias currents I_{cc} and I_{hh} can be tuned over several decades, while still perceiving their effect.

C. Integrate-and-Fire

Figure 4.9(a) shows the integrate-and-fire block. Input contrast current I_{cont} is integrated on capacitor C_{int}. Two comparators detect whether the capacitor voltage V_{cap} reaches an upper (V_{high}) or lower (V_{low}) threshold, triggering the generation of a positive (*pulse+*) or negative (*pulse-*) event, respectively. To accelerate the comparisons, both comparators activate a positive feedback loop (from V_{cap} to V_{dd04} for a positive event, or from V_{cap} to V_{gn04} for a negative event).

After event generation, capacitor C_{int} is reset to the central voltage V_{ref}. This is done by the reset circuit shown in Figure 4.9(b). This reset mechanism includes a refractory timing circuit that inhibits the pixel from generating subsequent events before refractory capacitor C_{rfr} has been discharged by the DC current source MOS controlled by V_{rfr}. The reset circuit also includes the global TFS (Time-to-First-Spike) mode reset signal, which resets all pixel capacitors C_{int} simultaneously. Note that this signal inhibits the positive feedback loops in Figure 4.9(a). This allows resetting quickly those pixels generating an event when TFS becomes active.

Figure 4.9(c) shows the minimum contrast thresholding circuit. A comparator detects whether capacitor voltage is above or below V_{ref} and turns on either a positive (I_{low}) or negative (I_{high}) threshold current, which I_{cont} needs to exceed for producing an event. Figure 4.10 shows the resulting relationship between integrate-and-fire circuit output frequency f_{out} and the input signed contrast current I_{cont} while bias voltages V_{th}^{+} and V_{th}^{-} are set to generate threshold currents I_{high} and I_{low}, respectively. Naturally, threshold transistors would also introduce mismatch. Consequently, they were layed out with a large area of $2/20\mu m$.

Figure 4.9(d) shows the two-stage comparators used in Figure 4.9(d). At stand by they are biased at low current through V_{b1} and V_{b2}. However, during event generation its bias current is increased. This increase starts when signals *pulse* starts to depart from its resting voltage and stops after the pixel event reset signal *ev_rst* returns to its resting level. The comparator within the thresholding circuit in Figure 4.9(d) does not have this feature, since this comparator only needs to detect whether the so far accumulated contrast for the pixel is positive or negative, which is a slow process compared to the event generation timings.

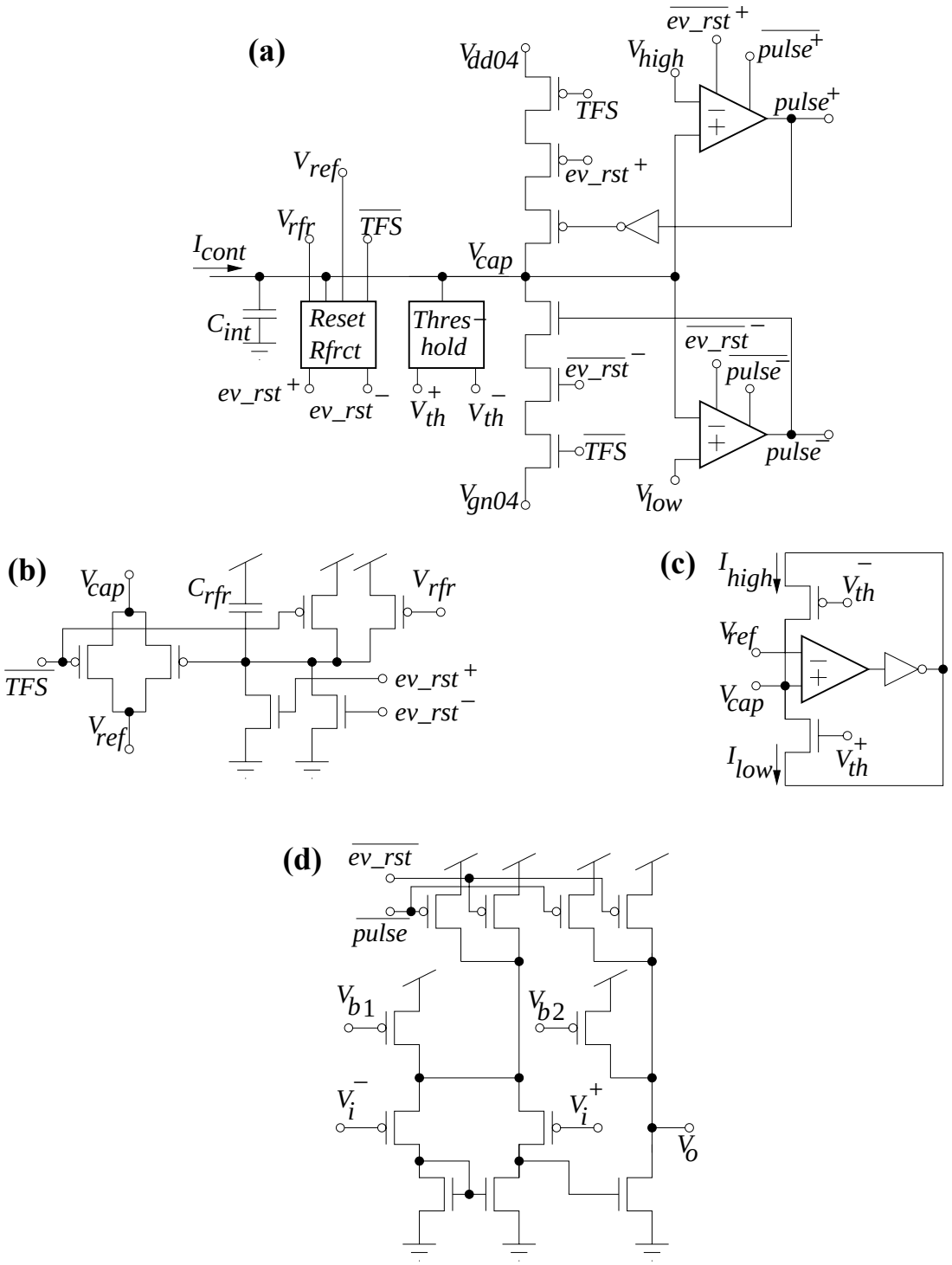


Fig 4.9: (a) Integrate and fire neuron. (b) Reset and refractory circuit. (c) Thresholding circuit. (d) Detail of comparators.

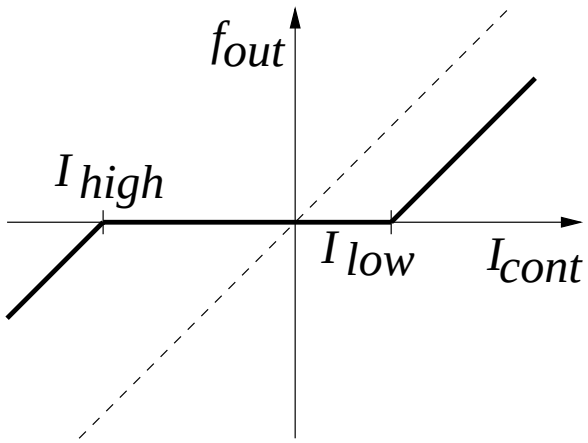


Fig 4.10: Effect of contrast thresholding on the relationship between pixel output frequency and contrast current.

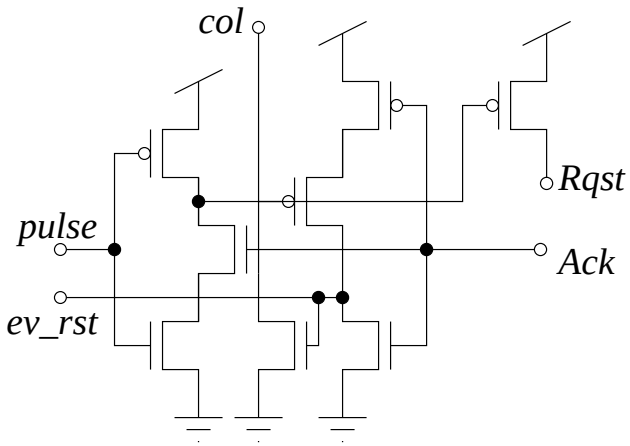


Fig 4.11: AER pixel communication blocks.

D. Communication Block Circuit

Finally, the AER pixel communication part in Figure 4.5 contains two identical "event block" circuits, which are shown in Figure 4.11 These are standard AER pixel communication circuits taken from Boahen's row parallel event read-out technique [47]. When generating signed

Table 4.1: Chip Specifications.

technology	CMOS 0.35μm 4M 2P
power supply	3.3V
chip size	2.5 x 2.6 mm^2
array size	32 x 32
pixel size	80 x 80 μm^2
fill factor	2.0%
photodiode quantum efficiency	0.34 @ 450nm
pixel complexity	131 transistors + 2 caps
current consumption	65μA @ 10keps

events, each pixel needs to provide two column event signals *col+* and *col-*. This concept was already implemented and tested in prior designs [27] that required signed events.

4.5 Experimental Results

A 32 x 32 pixel test prototype AER signed spatial contrast retina chip was designed and fabricated in a double poly 4-metal 0.35μm CMOS process with a power supply of $V_{DD} = 3{,}3V$. Table 4.1 summarizes the chip specifications. Figure 4.12 shows a micro photograph of the die, of size 2.5 x 2.6 mm^2. The whole chip, except the pad ring, is covered with the top metal layer leaving openings for the photo diode sensors. Figure 4.12 also shows the layout of a single pixel highlighting its components. Each pixel layout is a symmetrical reflection of its neighboring pixels. This way noisy digital lines are shared among neighbors, as well as power supplies, and noise sensitive bias lines among other capabilities [52]. At the same time, noise sensitive lines are separated from noisy ones. Pixel area is $80 \times 80\mu m^2$, including routing. The pixel was made up of 131 transistors and 2 capacitors (the capacitance of the integrate-and-fire circuit and the capacitance of the reset and the refractory circuit).

Figure 4.13 shows the experimental setup to characterize the temporal contrast sensor. The chip was mounted on the *Caviar Board* (a specific PCB designed to test generic AER devices). This board was connected to a USB port. The digital words to program the bias currents of each *I-pot* [37] were sent through this port. The outputs of the *Caviar Board* were sent to an AER bus. This bus was connected to a Mapper. The Mapper was a programmable device that was used to filter or remove the activity of some pixels for some specific measurements (latency characterization), without affecting the activity of the rest of the pixels. The output bus of the Mapper was connected the board *USB2AER,* [52]*,* which sends the AER data from the bus to one of the USB computer ports. This information was processed by jAER [49] that allows to see real time data. The *USB2AER* board was specially useful for real-time monitoring. The input bus to this board was replicated and sent to a *Datalogger*. This board has an internal memory that can store the data generated by 524.000 events and was used to save efficiently on the PC the AER information transmitted from the AER bus.

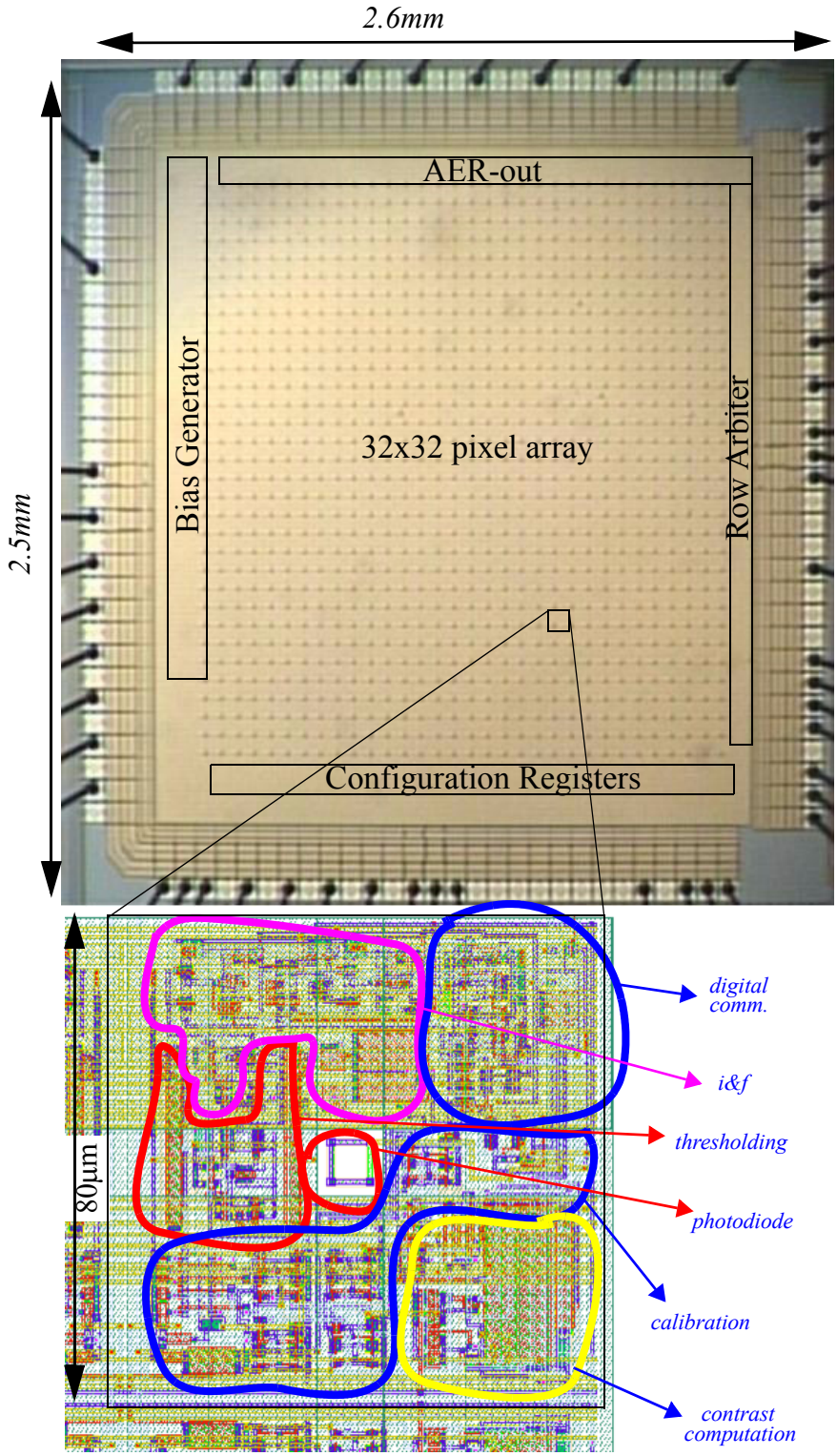
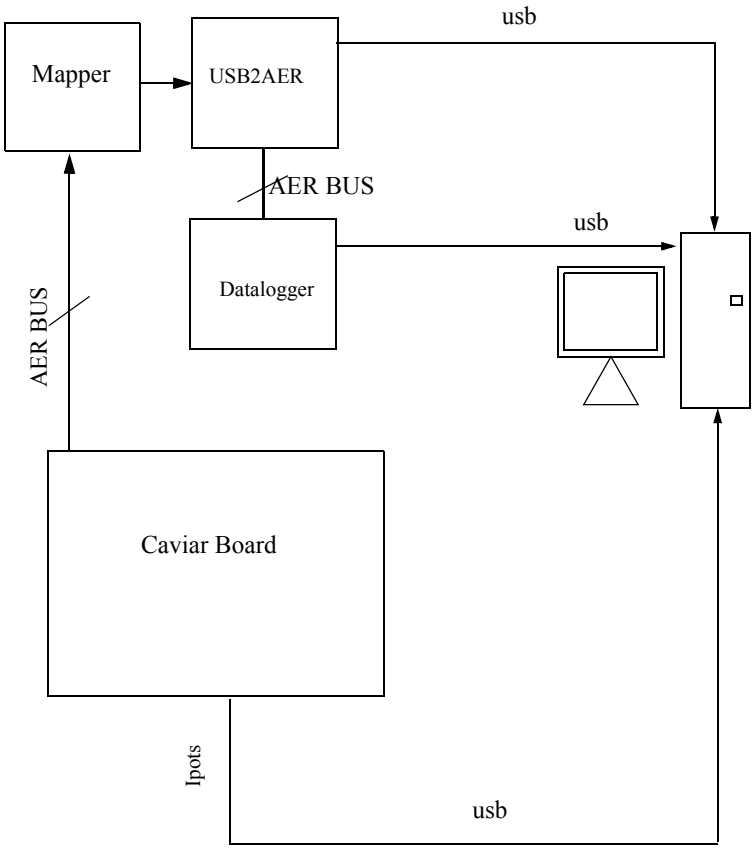


Fig 4.12: Microphotograph of 2.5*mm* x 2.6*mm* die, and zoom out of 80µ*m* x 80µ*m* pixel (layout) indicating the location of its components.



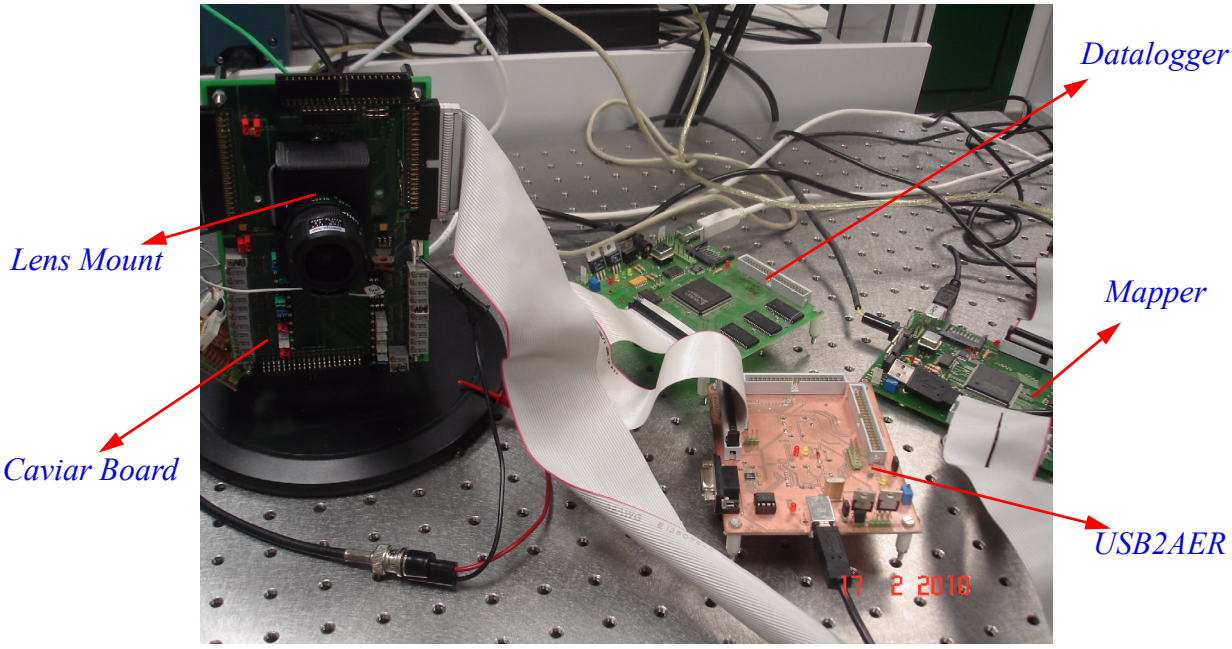


Fig 4.13: Experimental setup and photo of the retina and main boards used to test the imager.

4.5.1 Pixel Frequency Range

In order to control the pixel output frequency, one of the corner pixels had its integrating capacitor node connected to a low-input-capacitance analog buffer for monitoring purposes. Pixel integrating capacitors have a capacitance of about $C_{int} \approx 118fF$ (obtained from the layout extractor), while the corner pixel with monitoring buffer has a total capacitance of about $C_{mntr} \approx 196fF$ (estimated from layout extraction and simulation). Figure 4.14 shows recorded waveforms (for positive and negative currents) for this capacitor when turning off horizontal interactions among neighboring pixels (by turning off transistors M_h and M_c in Figure 4.8), and for a typical value of $I_u \approx 100pA$. By changing I_u (with $I_u' = I_{cal} = 0$) or I_u' (while $I_u = I_{cal} = 0$), pixel oscillation frequency could be tuned between 1.2*Hz* and 5*KHz*. For the maximum frequency the arbitrating periphery inserts varying delays. This is because all pixels are also firing with maximum frequency (even higher than the pixel we are observing which has slightly higher integrating capacitance) and are collapsing the arbiter. Consequently, in a practical situation where only a small percentage of the pixels would fire with maximum frequency, they would be able to fire with a higher than 5*KHz* maximum frequency.

4.5.2 Calibration

Although the retina can work without calibration, in order to use have bipolar output and reduce mismatch, the first requirement is to calibrate it. For this, the retina was exposed to a uniform stimulus, while biased for normal operation conditions but without thresholding. In our case, normal operation conditions are $I_u = 150pA$, $V_{ref} = 1{,}65V$, $V_{high} = 2{,}8V$, $V_{low} = 0{,}45V$, $I_{hh} = 10pA$, $I_{cc} = 5pA$. Also, before calibration, we set $I_{cal} = I_u' = 0$. Under these conditions, retina output events are recorded, from which one can obtain the firing frequency of each pixel. Next, we set current $I_u' = 80pA$ so that the pixel with minimum frequency has a frequency close to zero (or slightly negative). Under these conditions the resulting histogram of pixel frequencies distributions is shown in Figure 4.15.(a) After this, the calibration circuit biases (I_1, I_2, I_3 in Figure 4.7) were set for optimum coverage of this distribution, and for each pixel the optimum calibration word $\boldsymbol{w}_{cal}(x, y)$ was found. This is computed off-line by optimally combining biases $\{I_1, I_2, I_3\}$ and calibration words $\boldsymbol{w}_{cal}(x, y)$, following a pre-established optimization criterion. In our case, we allowed for a few outliers in order to minimize the residual standard deviation. One could also target to minimize the spread among the most extreme pixels at the expense of a higher standard deviation. After this process, the histogram of calibrated pixel frequencies obtained is shown in Figure 4.15(b). The residual inter-pixel standard deviation is 26Hz. As we will explain later, maximum contrast frequency for these biases is $\pm 4400Hz$. Consequently, post-calibration residual mismatch is 0.6%. Figure 4.16 shows the calibration ladders obtained for the 1024 pixels that compounds the retina. We have also plotted with red asterisks the optimum output frequencies after calibration versus the optimum calibration word for each pixel. Finally, in Figure 4.17 we have plotted the output frequencies before calibration (with a bipolar output subtracting the current I_u) and after calibration. Output frequencies were calculated counting the pixels spikes within a time interval. *Datalogger* [48] was used to save the number of spikes generated for each pixel during the interval.

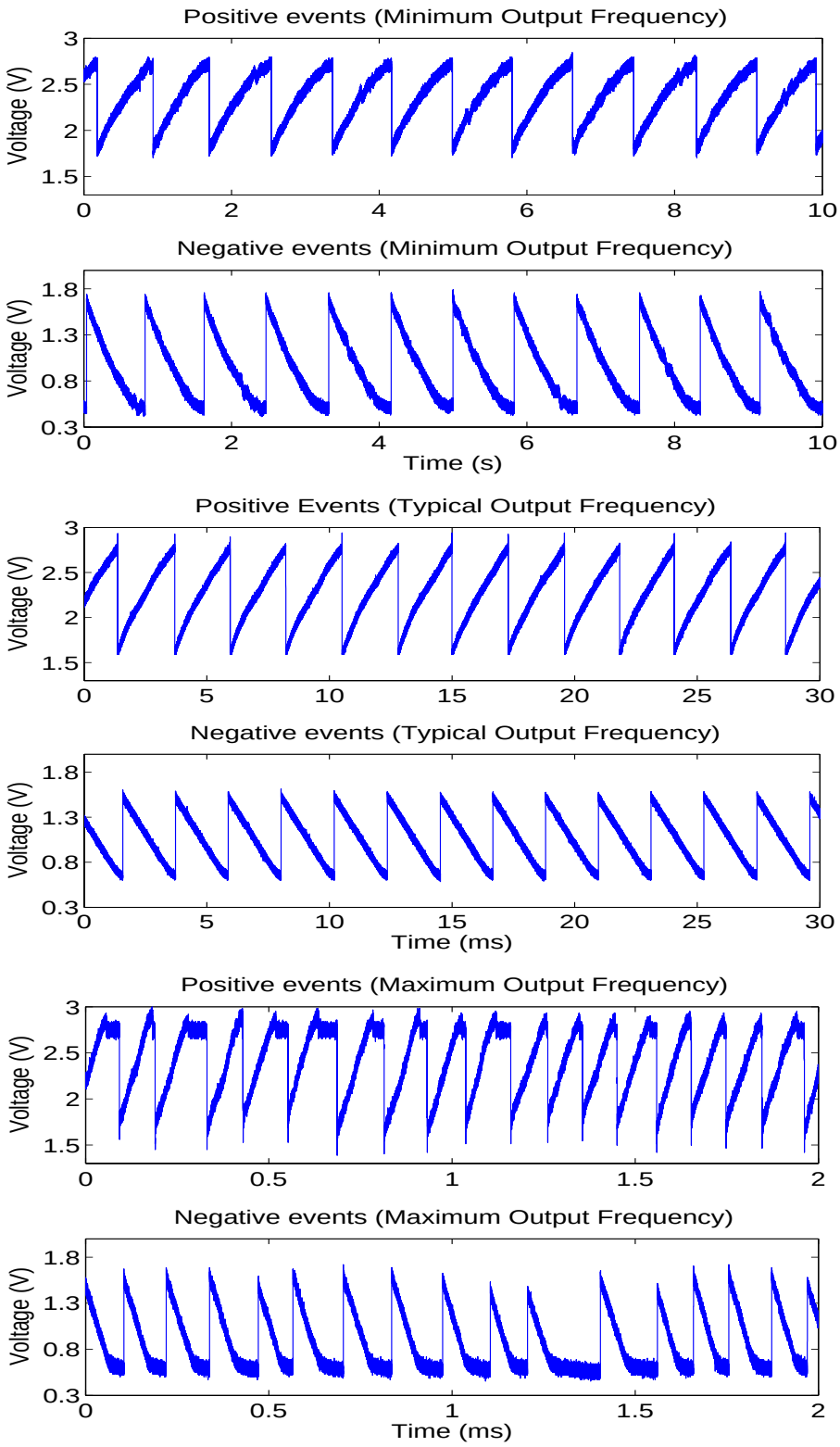


Fig 4.14: Recorded waveforms at the integrating capacitor. Minimal oscillation frequency (1.2 Hz), Typical oscillation frequency (466 Hz), and maximum oscillation frequency (5 KhZ).

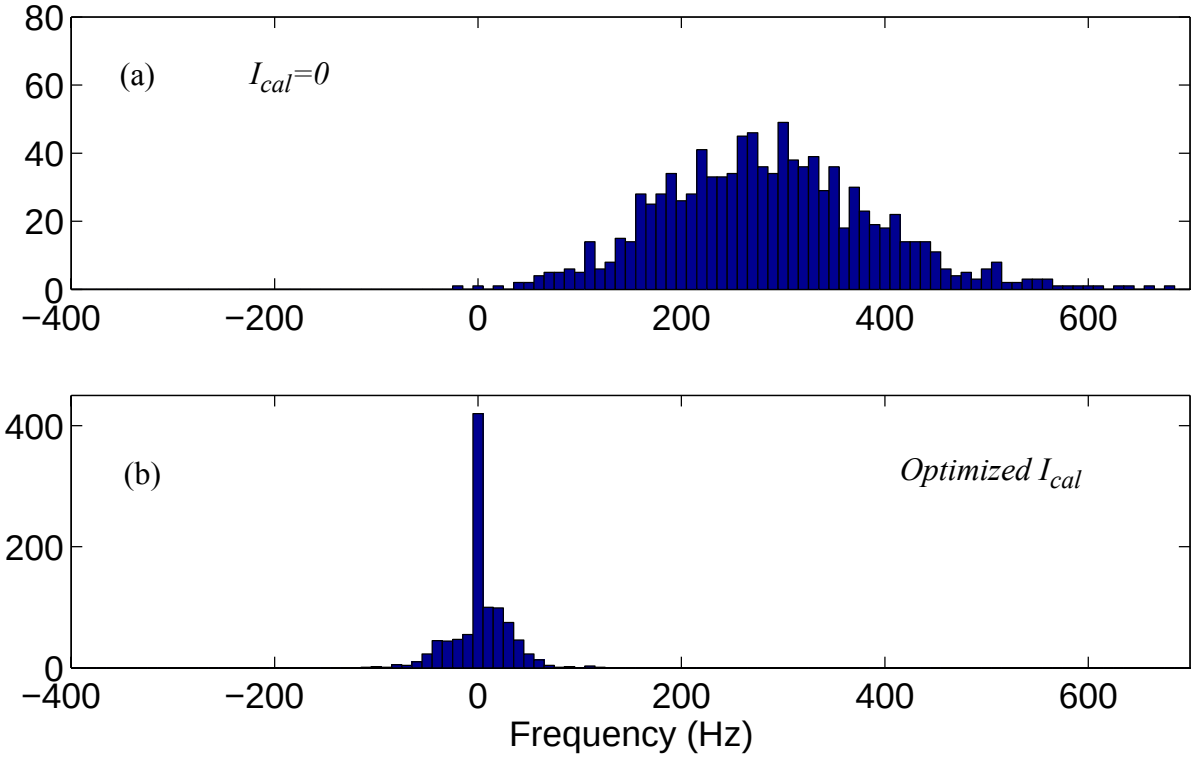


Fig 4.15: Histograms of retina pixels frequencies distribution before and after calibration at the same illumination level.

Figure 4.18 shows how the standard deviation of the post-calibration residual mismatch changes with illumination level. The figures shows three superimposed graphs. Each correspond to performing calibration at different illumination levels {50, 5, 1, and 0.25 *Klux*}.The worst case situation corresponds to calibrating at about 1k-lux and using the retina at very high light conditions, resulting in a standard deviation of almost 140*Hz* (3%). On the other hand, the optimum situation corresponds to calibrating at 15k-lux, which results in a standard deviation of less than *80Hz* (1.8%) over the entire 5 decade range.

The calibration process is all done off-line. However, it is conceivable to implement it fully on-chip (through, for example, a described state machine), since it only requires to expose the chip to uniform illumination (one can simply remove the optics), compare the pixel frequencies (for which not even a precise clock reference is required), and compute an optimum set of calibration weights.

4.5.3 Contrast Step Response

Figure 4.19 illustrates the retina response to a luminance step of different contrast levels, while thresholding is turned off. Input stimulus is printed paper, providing a static image with a half dark and a half grey side. The half gray side intensity is adjusted between 100% (white) and 30% (darkest gray). Table 4.2 indicates the relationship of the luminance steps, with the ratio of photo currents between the gray and black parts, and the resulting Weber Contrast, defined as

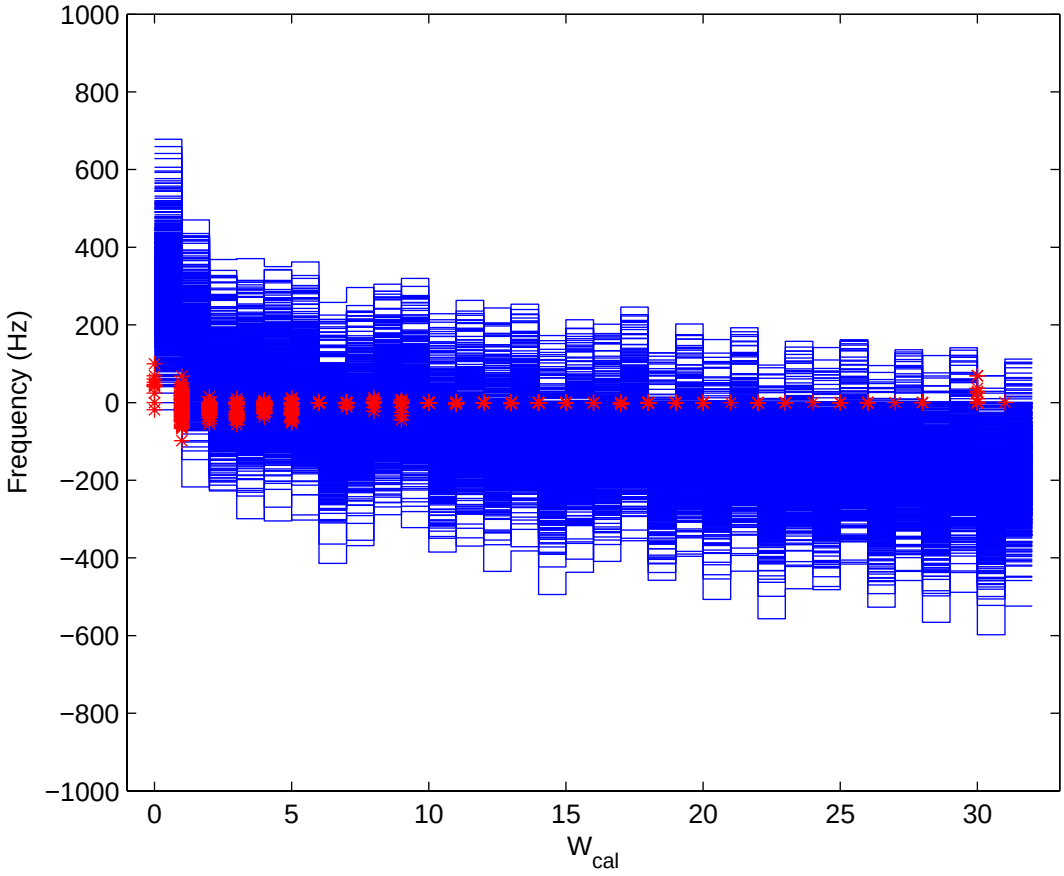


Fig 4.16: Calibration ladders and optimum calibrate words for all the pixels of the retina.

$$WC = \frac{I_{light} - I_{dark}}{I_{light} + I_{dark}} \tag{4.6}$$

The left column in Figure 4.19 shows this input stimulus image. The center column in Figure 4.19 shows the retina output response before calibration, while the right column shows the retina response after calibration. Central gray level is zero pixel frequency. Brighter pixels are firing positively signed events, while darker pixels are firing negatively signed events. Absolute maxi-

Table 4.2: Measured luminance steps and Weber contrast.

luminance step	100% to 0%	70% to 0%	50% to 0%	30% to 0%	10% to 0%	0% to 0%
I_{light}/I_{dark}	9	6	3.6	2.4	1.5	1
Weber Contrast (WC)	0.80	0.72	0.56	0.41	0.20	0

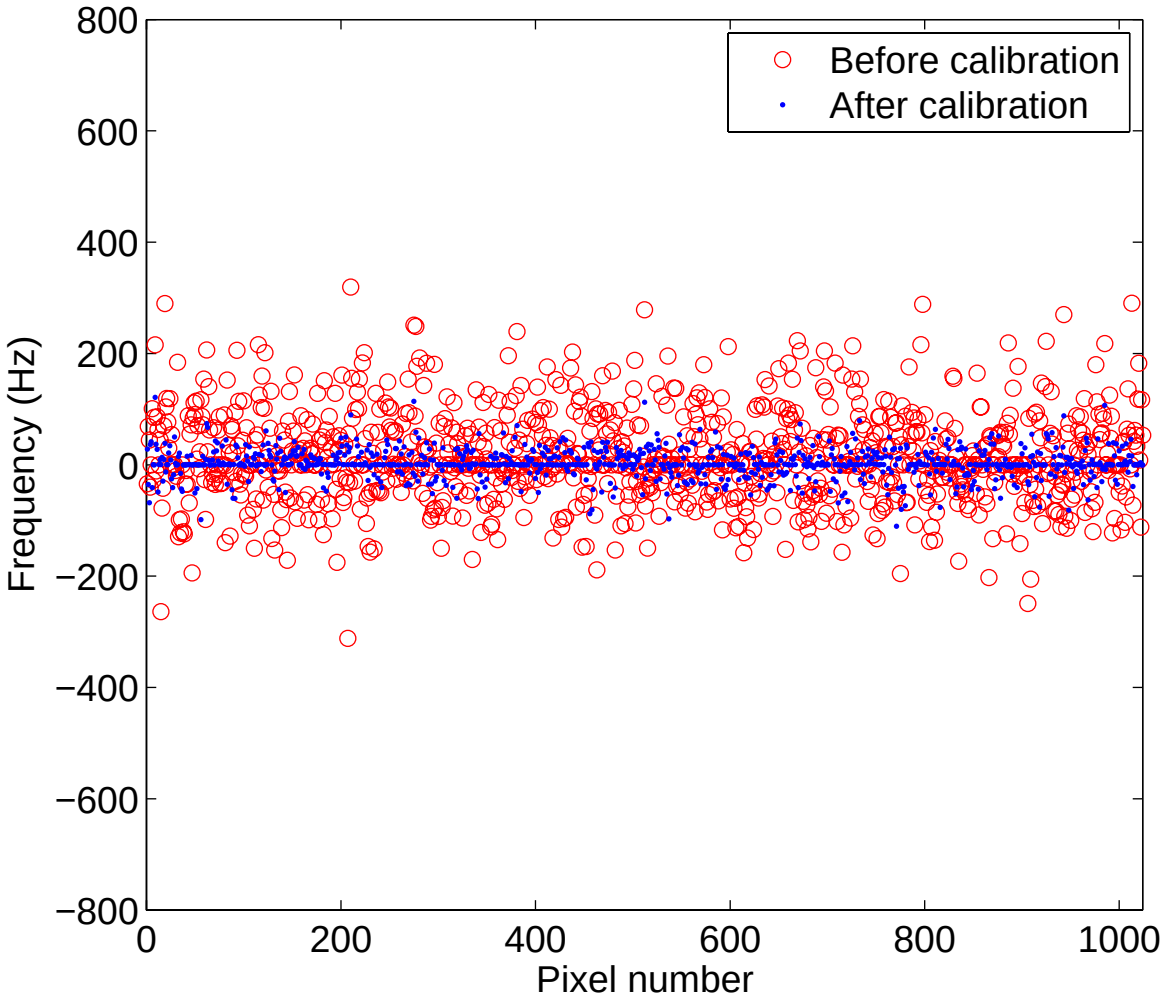


Fig 4.17: Output frequencies before and after calibration.

mum pixel frequency was 250*Hz*. Biasing conditions in Figure 4.19 were $I_u = 150pA$, $I_u' = 150pA$, $V_{high} = 2.9\text{V}$, $V_{low} = 0.4\text{V}$, and $V_{ref} = 1{,}65V$.

4.5.4 Contrast Sensitivity

An important characterization for a spatial contrast retina is its contrast sensitivity what is the output event rate for a given input contrast stimulus. We have characterized spatial contrast sensitivity for the positive event branch and the negative event branch (see Figure 4.5) separately, since they have separate circuitry. Usually, under normal operation, the retina will be biased to have the same sensitivity for positive and negative events. However, there might be situations where one would prefer to set different contrast sensitivities for positive and negative events, and this retina offers this possibility. To characterize pixel contrast sensitivity, a gray level step stimulus (as shown in Figure 4.19) of different contrast values, was used. Pixels frequencies of the two columns with the highest activity (the ones just on the left and right of the stimulus center) were recorded. This process was repeated for different bias values for V_{high} and V_{low}, with

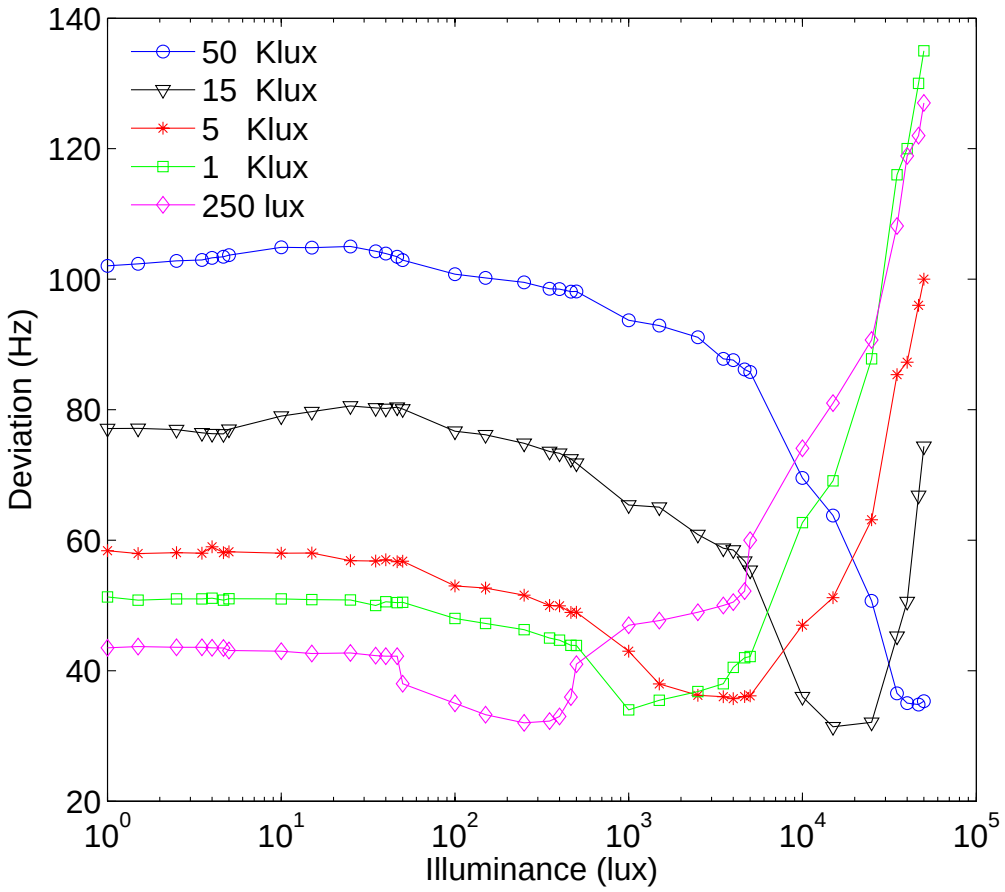


Fig 4.18: Effect of ambient illumination on post-calibration residual mismatch standard deviation. Three curves are shown, each for calibration at the given illumination level.

V_{ref} = 1.65V . The results are shown in Figure 4.20(a). The measured maximum contrast sensitivity was 4400*Hz/WC* (*Hz* per Weber Contrast) for $V_{high} - V_{ref} = V_{ref} - V_{low} = 0.15\text{V}$. Error bars indicate inter-pixel variability.

To show the sensitivity dependence with illumination, the maximum output frequency for a Weber Contrast of $WC = 0.8$ was measured (for both signs of contrast) with different illumination levels. As shown in Figure 4.20(b), sensitivity degrades slightly when illumination decreases. Sensitivity remains almost constant over the first two decades, and approximately doubles over the second two decades.

4.5.5 Contrast Thresholding

In Figure 4.21, the typical pixel output when the visual field is swept with a grey level bar stimulus of $WC = 0.8$ is shown. The x-axis indicates bar position in row number units. The pixel output spike frequency reaches the maximum value when the stimulus is at the pixel´s row. This value depends on the width of the sweeping bar. Several outputs using different bar widths have been plotted for the same pixel. The bar width is expressed in projected pixel units. The maximum frequency is proportional to the stimulus width. In both cases, the following voltages were used:

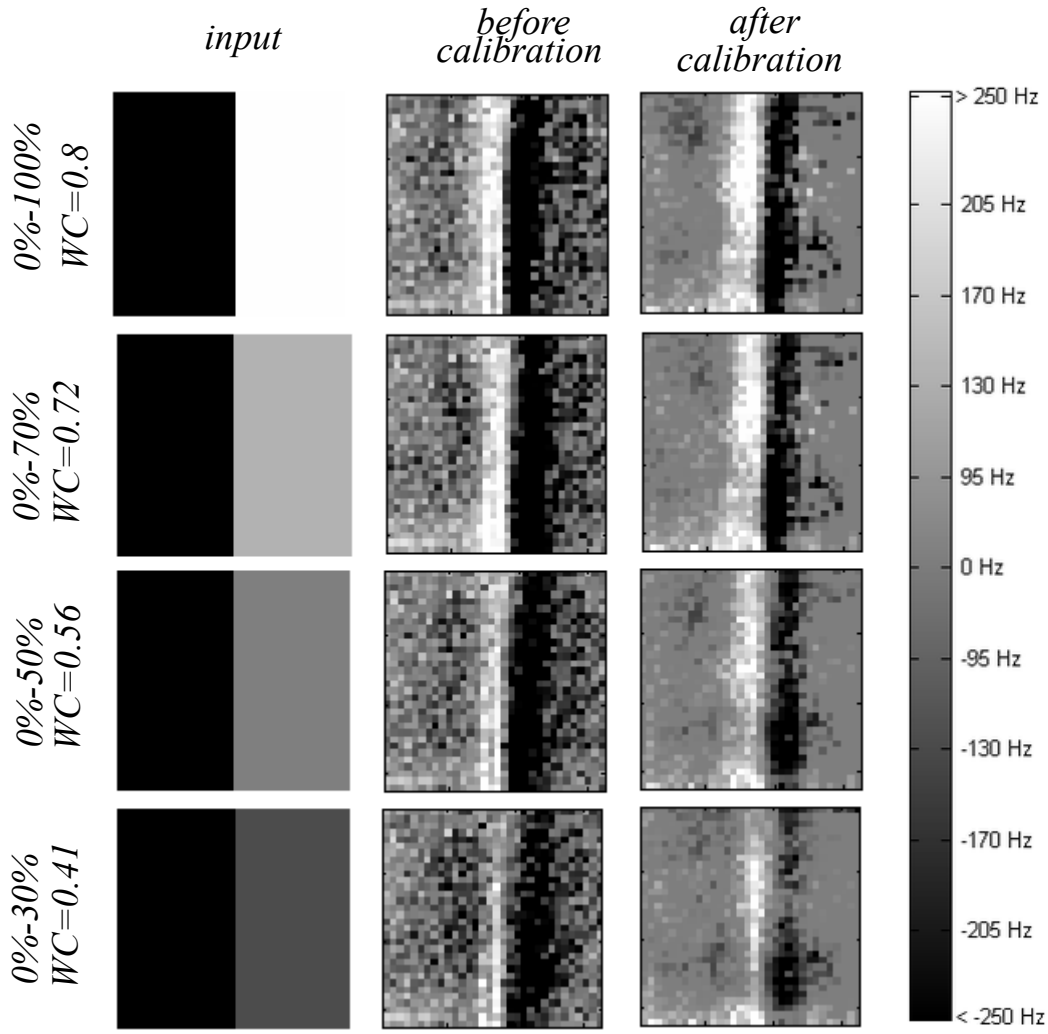


Fig 4.19: Retina response to a luminance step changing Weber Contrast. Left column is input stimulus. Centre column is output response before calibration, and right column is output response after calibration.

$V_{high} = 2.9V$, $V_{low} = 1.4V$ and $V_{ref} = 1.65V$. With these settings, $V_{high} - V_{ref} > V_{ref} - V_{low}$, so negative events were enhanced.

It is also possible to fully inhibit positive or negative events by setting either I_{high} or I_{low} (see Figure 4.9(c)) to sufficiently large values. Asymmetrical thresholds ($I_{low} \neq I_{high}$) can also be used. Therefore, positive and negative events can be inhibited independently. In Figure 4.22, the effect of thresholding is shown. First, the visual field was swept with a 100% contrast bar for different thresholds. Figure 4.22(a) shows the output frequency for pixel (17,11) when setting symmetric thresholds. Figure 4.22(b) shows the same pixel results but when setting only threshold values to inhibit positive events. The negative output frequency remains constant.

The main advantage of thresholding is to remove the residual mismatch after calibration. Pixels usually spike with a low residual output frequency after calibration. Positive and negative thresholds can be set to remove these undesirable outputs after calibration. Figure 4.22(c-e) show

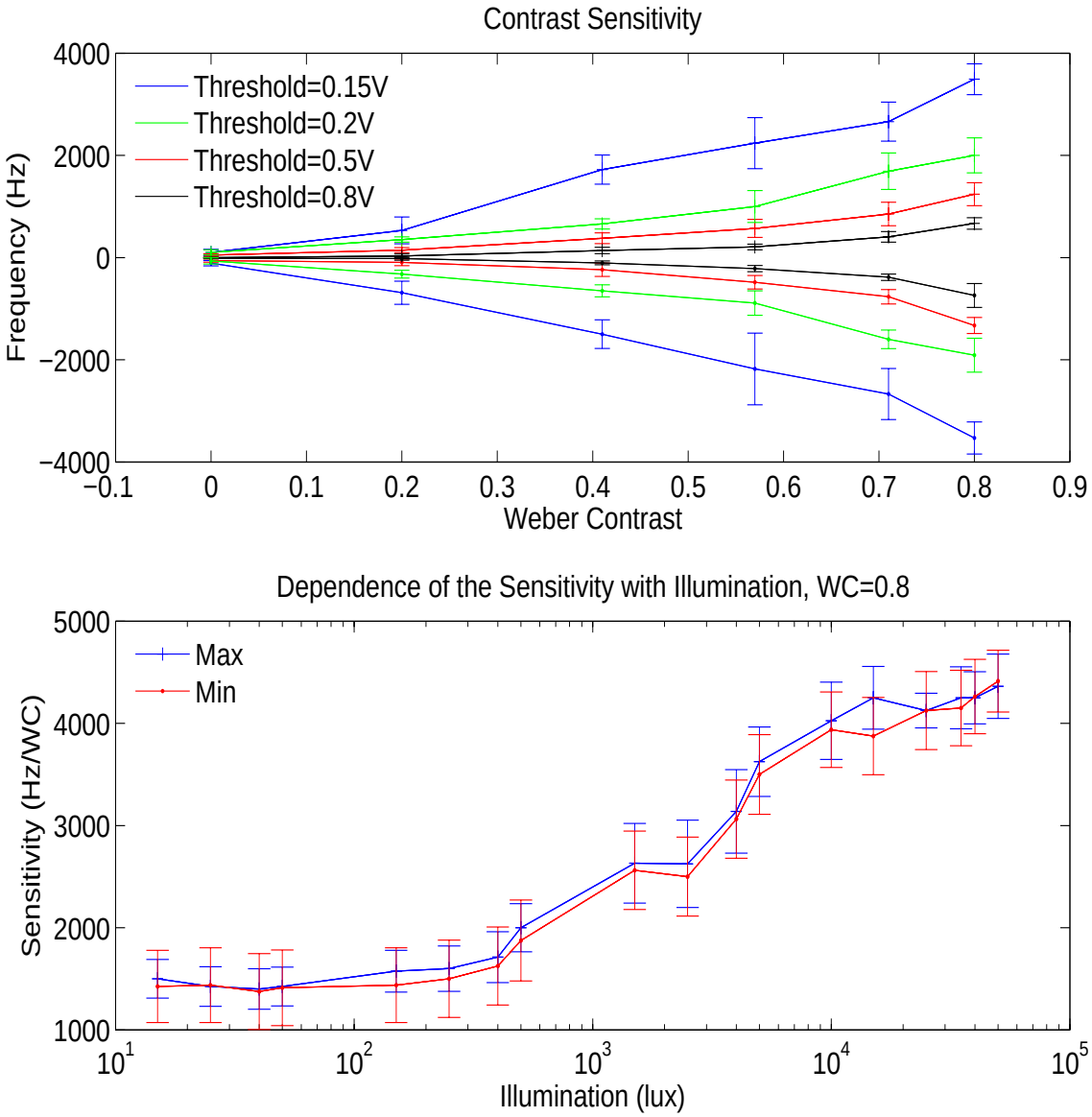


Fig 4.20: Contrast sensitivity measurements. A stimulus step (as in Figure 4.19) was applied and max and min frequencies were recorded. Top panel shows mas and min frequencies for different stimulus step contrasts and different threshold values. Bottom panel shows how the maximum and minimum frequencies depends on illumination (*WC=0.8*).

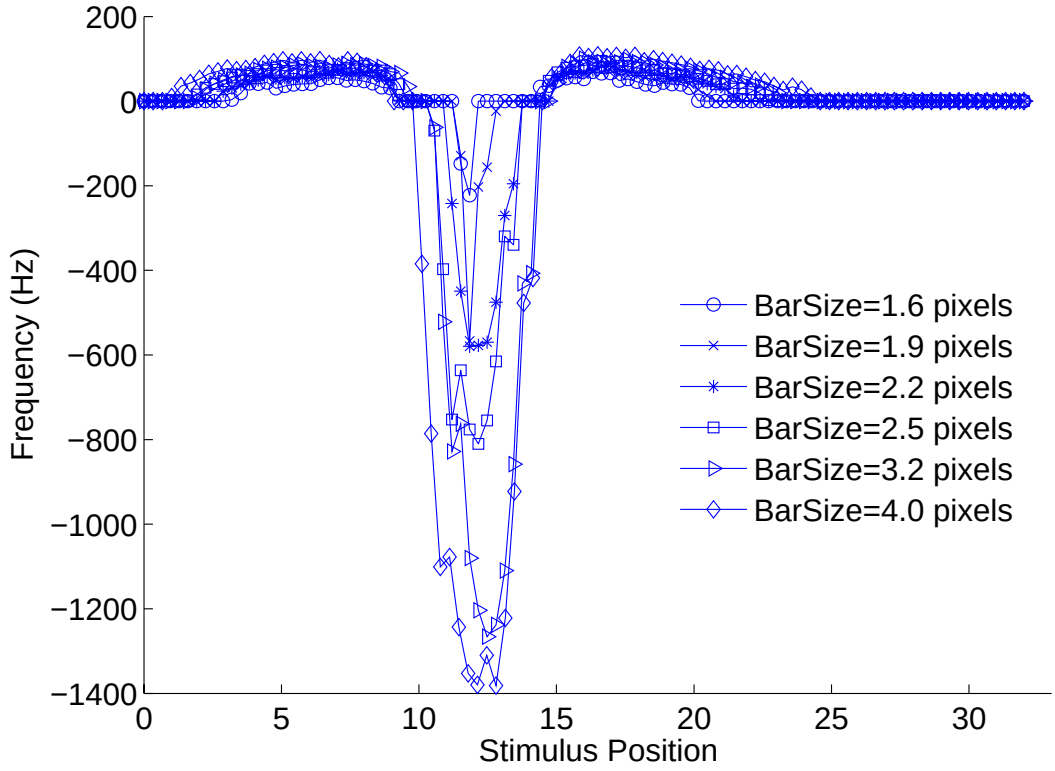


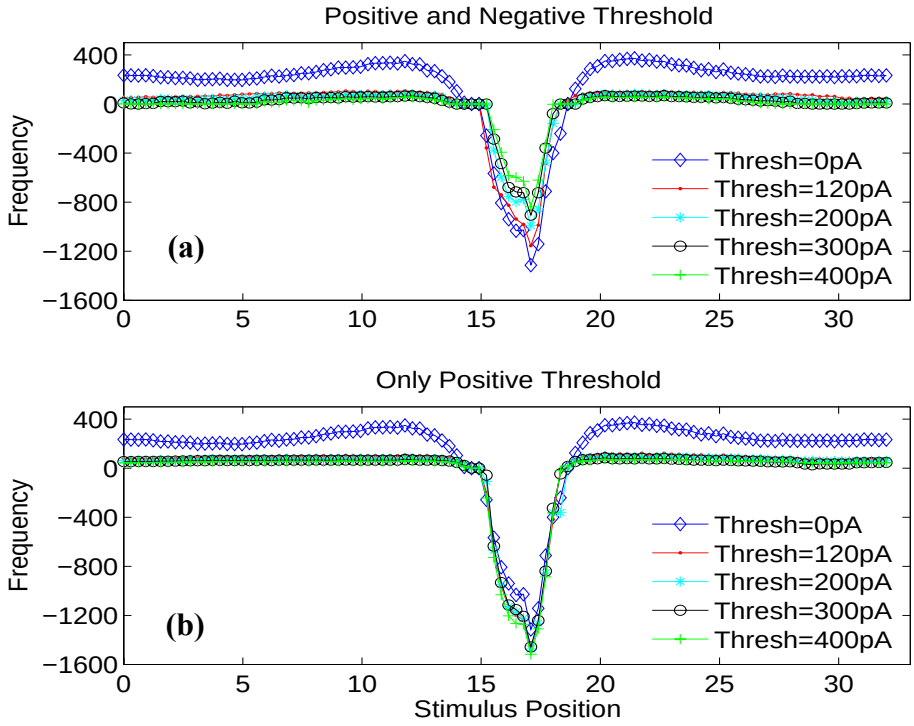
Fig 4.21: Typical pixel's output when the retina is stimulated with a 100% contrast bar of different widths.

some snapshots captured with the contrast retina. Central gray color indicates zero output (no contrast). Positive events range from this gray to black and negative events range from this gray to white. The three snapshots were taken for different values of the positive and negative thresholds. For the three cases, $I_u = 150pA$. In Figure 4.22(c) a positive threshold current of $1nA$ was set to inhibit positive events completely after calibration. I_{low} was $150pA$. In Figure 4.22(d) a symmetric threshold of $80pA$ was set after calibration. In Figure 4.22(e) the retina output without neither calibration nor thresholding is shown. Above each snapshot the sum of all pixels' frequencies f_{total} is indicated. We can see, by comparing (d) and (e), that calibration reduces event flow (communication bandwidth) while enhancing contrast gain.

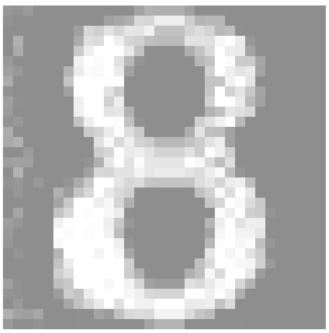
4.5.6 Latency Characterization

Sensor latency is the time it takes from the occurrence of the illumination change at the photo diode to the output of an address event. Sensors with quick answers to stimuli are suitable to be used as part of a feedback controller. Latency depends on the pixel, the asynchronous bus arbitrer and the periphery.

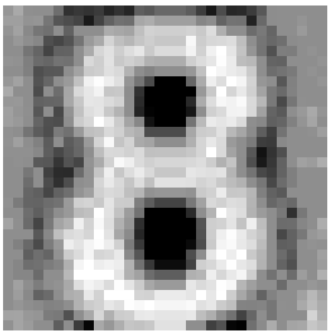
To characterize the retina latency we proceeded as follows. We stimulated a LED with a step signal to turn it ON, focused it over a central region of the sensor array, and recorded the time delay between the step signal and the first event *Rqst* coming out of the pixel under test from that region. The Mapper was programed to allow spike only the pixel under test of the stimulated



(c) f_{total}=1.3 x 10^5 *Hz*



(c) f_{total}=1.8 x 10^5 *Hz*



(e) f_{total}=3.3 x 10^5 *Hz*

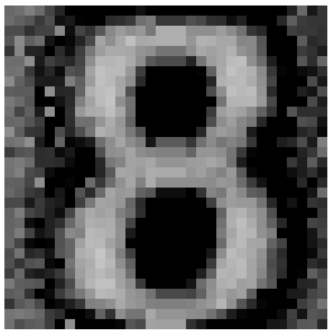


Fig 4.22: Effect of thresholding. (a) Bar is swept for different symmetric thresholds. (b) No threshold for negative events, and positive event thresholds are changed. (c) Events captured for calibrated retina when all positive events are inhibited by setting a high positive threshold. (d) Events captured for calibrated retina with symmetric threshold. (e) Events captured for uncalibrated retina.

region. An oscilloscope working in single mode was used to measure the latency. The measurements were repeated by inserting different neutral density filters to attenuate light intensity from about 50k-lux down to 2 lux. The resulting latencies are shown in Figure 4.23. The measurement was repeated by focusing the LED over different regions of the pixel array. The bars in Figure 4.23 show the spread obtained when repeating measurements. As can be seen, latency changes

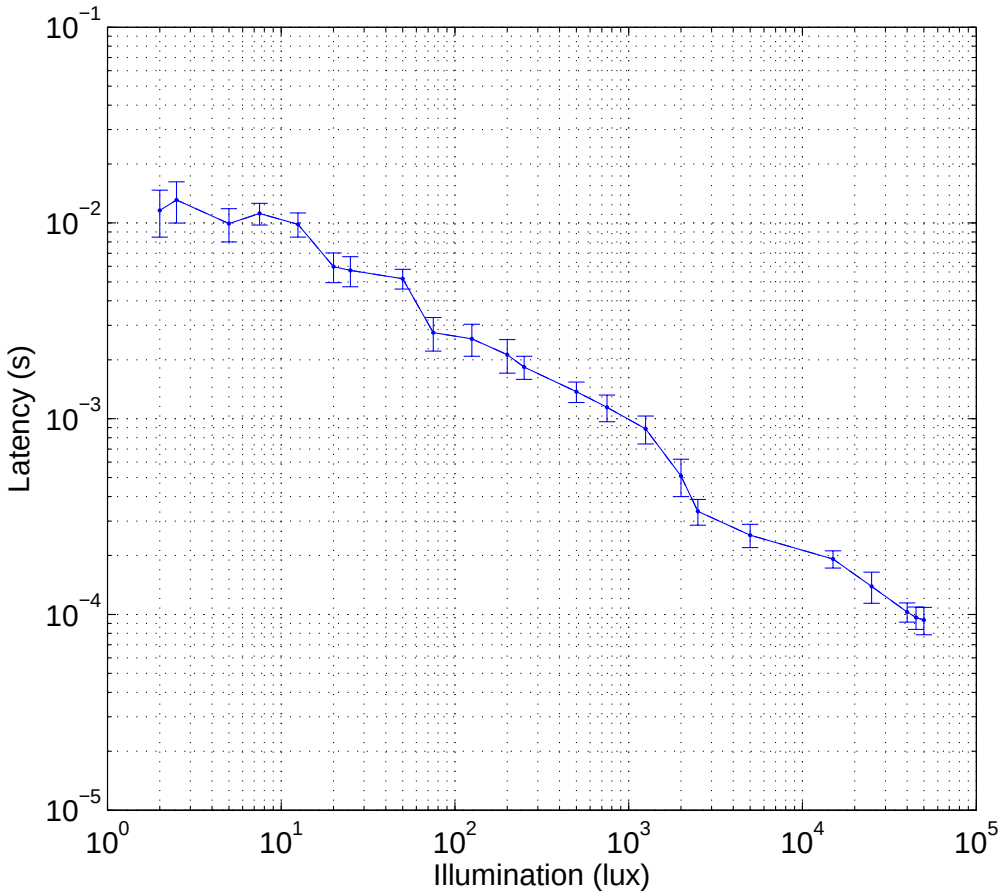


Fig 4.23: Latency measurements under changing illumination conditions.

from about 10*ms* down to about 0.1*ms* when illumination varies over almost 5 decades. This means that latency is dominated by the photo sensing circuits. However, latency does not scale proportionally to light, and consequently this retina does not suffer from the severe Light-to-Time restriction listed in Table 1.1.

4.5.7 Natural Scenes

Although the retina resolution is rather low (32 x 32 pixels) for observing natural scenes, Figure 4.24 shows some captured images when observing natural elements, which give a first order feeling of how an up-scaled retina would respond under a natural scene. jAER [49] was used to convert AER data into a sequence of frames for real-time visualization and monitorization. We used a lens objective of *16mm*.

4.5.8 TFS Output Mode

As mentioned on page 112, the integrate-and-fire circuit of the retina pixel can be configured to operate in TFS mode. In this mode, the refractory period of the retina has to be set to its largest possible value (by connecting voltage V_{rfr} to V_{dd}) to guarantee that each pixel will fire at the most one single event. Then a periodic reset pulse has to be provided for global signal $\overline{TFS}$.

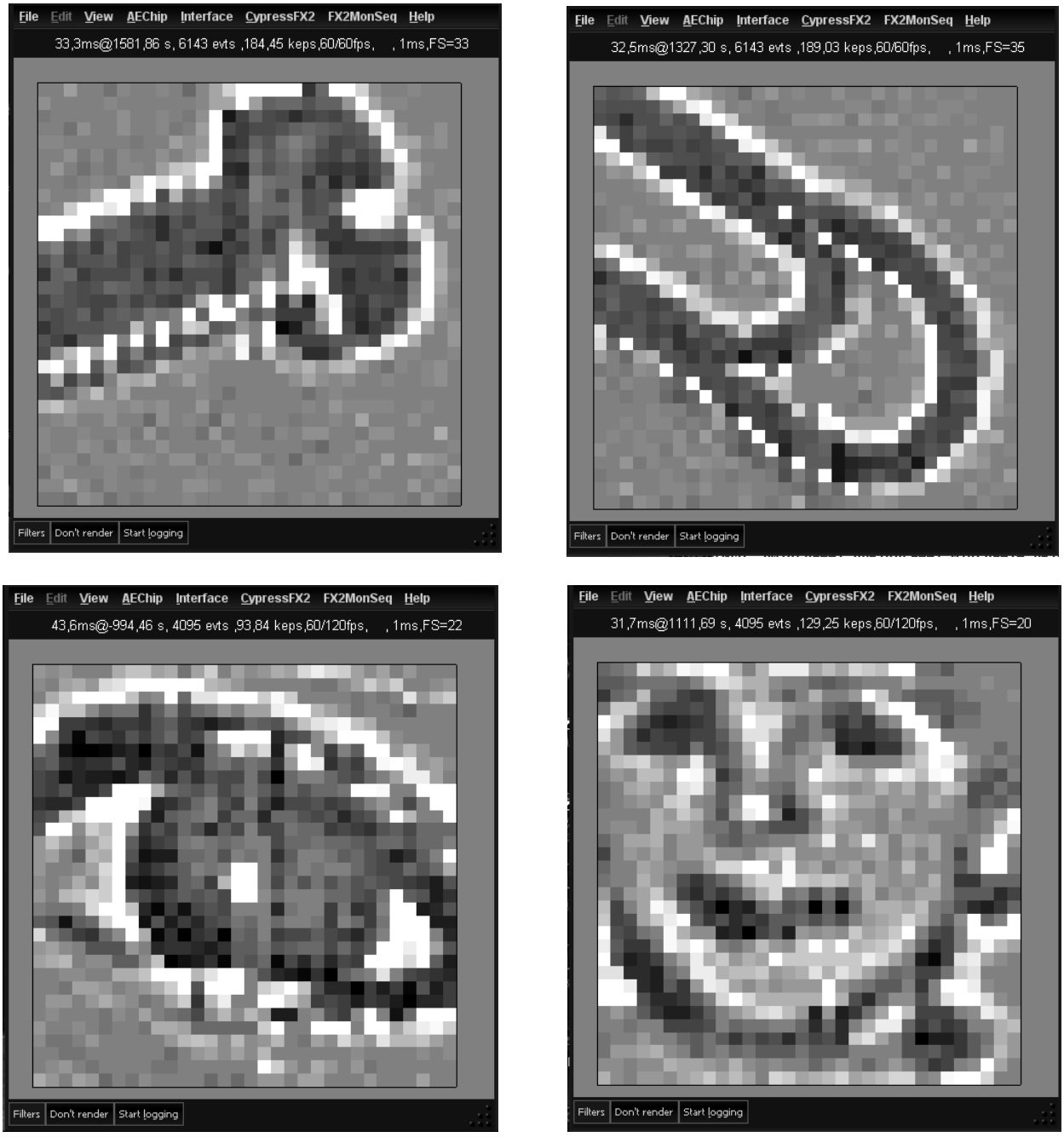


Fig 4.24: Natural elements. From left to right: screw, paper clip, eye and child face. We used a lens objective of *16mm*.

This can be done in several ways. One trivial option is to reset at a fixed preset frequency. However, another more efficient option is by counting the output events. Since output events are coming out in decreasing order of pixel contrast, high contrast pixels (either positive or negative) come out first. These are the pixels carrying more relevant information, for example, for a recognition application. Consequently, one could add a simple counter at the *Rqst* line and have it generating a reset pulse for $\overline{TFS}$ after each M events. This way, a dynamic "frame time" T_{frame} would

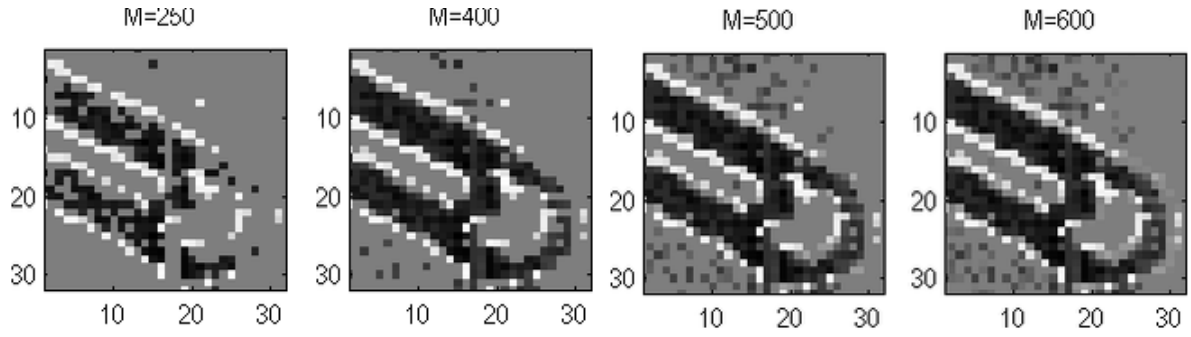


Fig 4.25: Paper clip snapshots in TFS mode for different number of captured events, *M*.

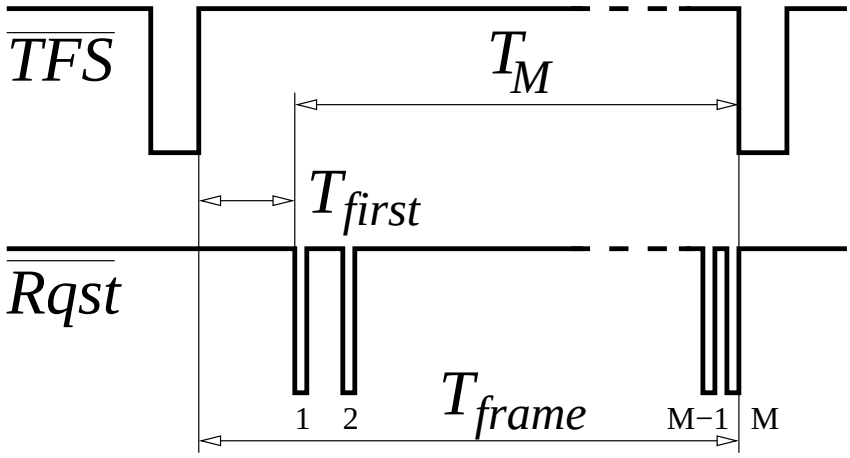


Fig 4.26: Time line of the Global Reset and the Request signal.

be produced which self adjusts to the contrast level of the scene, independent of ambient light. High contrast scenes would self-tune to faster frames, while low contrast scenes would self-tune to slower frames for the same amount of contrast information. Other more sophisticated options could use a post processing event based system for performing a given recognition and provide the reset pulse once a recognition has been achieved, or reset after a preset time if no recognition was possible. In what follows we count a fixed number of events *M*. Figure 4.25 illustrates the effect of changing *M* when observing the paper clip of Figure 4.24. Note that setting *M* to low values also removes background noise.

The TFS output mode is also insensitive to illumination (in first order), since it operates directly on I_{cont} within the integrate-and-fire circuit (see Figure 4.9(a-b)). To show this, several snapshots of the paper clip of Figure 4.24 were taken under different illumination conditions. As shown in Figure 4.26, T_{frame} is the sum of T_{first} (the time the retina needs to generate the first spike after the reset) and T_M (the time between the first and *M-th* spike). ,Figure 4.27 shows the value of T_{frame} for different values of *M* and illumination levels. T_{frame} is almost independent on illumination and is approximately constant for a given *M*. Figure 4.27 also shows the value of T_{first} versus illumination. In principle, T_{first} should not depend on ambient light because this reset is performed within the integrate-and-fire circuit (see Figure 4.9(a)) and not the photo sensing cir-

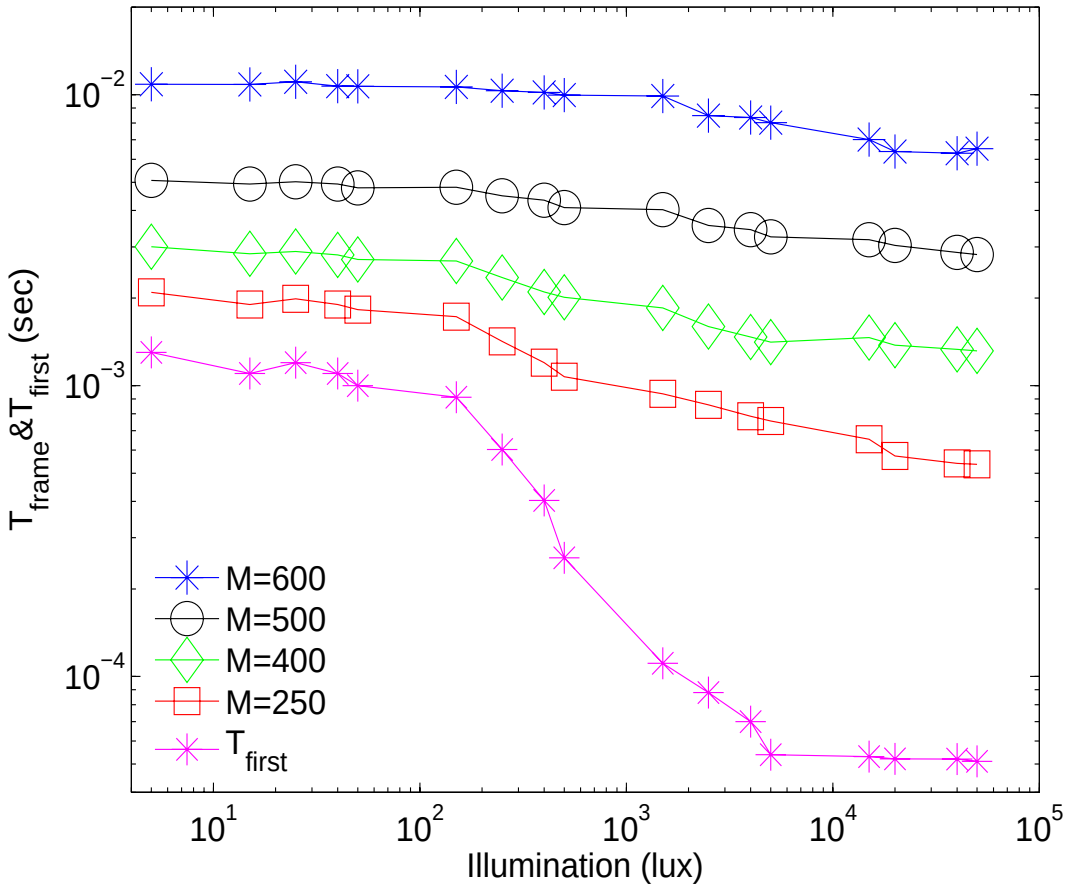


Fig 4.27: Effect of Illumination on T_{frame} and T_{first}

cuit (Figure 4.8). However, Figure 4.27 reveals a slow-down process when decreasing ambient light (between 5k-lux and 200 lux, approximately). This is probably due to switching crosstalk between the integrate-and-fire and photo sensing circuits, which introduces a switching transient in the latter that cannot be prevented when the photo currents are too small. Such problem can be attenuated in future designs by improving decoupling between the two stages, for example, through cascoding techniques.

4.5.9 Power Consumption

Chip power consumption has been characterized. Supply voltage is $3.3V$. In principle, it would depend on both static bias conditions and output event rate. However, in practice, it is dominated by the latter, because of the high consumption of digital pads communicating output events. Static power dissipation is negligible, since pixel current biases are set to relatively low values. Typical bias settings are $I_u = 150pA$, $I_{low} = 50pA$ and $I_{high} = 50pA$. This results in a pixel static current consumption of $15nA$. At very low output event rate (1keps) we measured a chip current consumption of $40\mu A$ ($130\mu W$). Figure 4.28 shows the measured current consumption of the chip as a function of output event rate. As can be seen, for normal operation regimes (between 100keps and 1Meps) current consumption varies between $200\mu A$ and $2mA$ ($660\mu W$ - $6{,}6mW$).

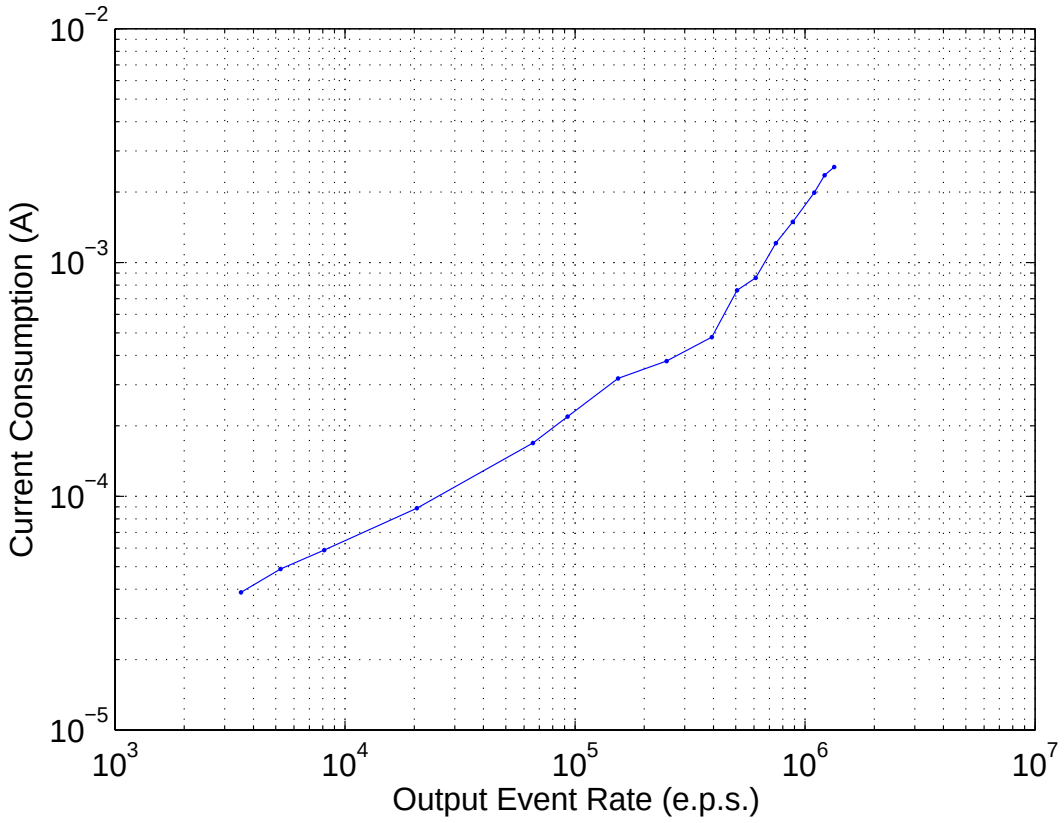


Fig 4.28: Chip total current consumption as function of total output event rate

Pixel output frequency (or TFS timing) range is directly controlled by bias current I_u (see Figure 4.8 and Figure 4.9(a)). Therefore, I_u controls also the overall power consumption and the speed-power trade-off.

4.6 Discussion

In this chapter, a new AER signed spatial contrast retina has been presented. It uses an improved and calibrated version of Boahen's contrast circuit. The design avoids the problem of AER communication bandwidth consumption present in prior designs. Furthermore, it also includes a thresholding mechanism, so that only pixels sensing spatial contrast above a given threshold generate events. A calibration scheme is included to partially compensate for pixel mis-match. An optional TFS coding scheme is also available. Extensive experimental results from a test prototype of 32 x 32 pixels, fabricated in a 0.35μm CMOS technology, are provided. Table 4.3 summarizes the imager main specifications.

An interesting advantage of this contrast retina is its fast response time as well as low communication throughput, compared to commercial video cameras rendering full frames every 30-40*ms*. Information throughput is reduced because only relevant contrast information is provided. Regarding speed response, for example when operating in rate coded mode, since active pixels fire at frequencies in the range of 1-5*KHz*, they would all update its state within fractions of

one mili second, independent of ambient light. In TFS mode, the first front of relevant events ($M = 250$ in Figure 4.27) is available in less than $1ms$. If the stimulus changes, the retina latency depends on lighting conditions, ranging from about $100ms$ at sun light (50k-lux) to $10ms$ at moon light (2 lux), with $1ms$ for indoor ambient light (1 k-lux).

Consequently, the complexity of developing spike based AER spatial contrast retinae, as opposed to conventional frame-scanned video cameras, could justified by its higher speed response for a very wide range of illumination conditions, while maintaining the information throughput low and ambient light independent. Although information throughput is low, relevant (contrast) information is preserved, which results in significant processing performance improvement for subsequent stages.

Table 4.3: Sensor Specifications.

Functionality	Light to time restriction	Latency	Dynamic Range	FPN	Fill Factor	Pixel Size μm^2	Fabrication Process	Power
Spatial Contrast to Number of Events	NO	0.1-10ms	>100dB	0.6%	2%	80x80	$0.35\mu m$ 4M 2P	0.66-66mW

CHAPTER 5

The AER Temporal Contrast Vision Sensor

5.1 Introduction

Motion detection is one of the most important tasks that the brain can perform. By detecting movement, plenty of relevant information about our environment is received. Temporal contrast detection is useful for the perception of relative depths, 3D structures, and for grouping objects that we can see. The motion perception pathway is independent of other visual pathways [58].

AER imagers are inherently faster than frame-based sensors. The main reason is that they can detect illumination changes continuously. There are not dead times between frames where changes can not be detected.

Several AER temporal contrast vision sensors have been reported previously. The first attempt to design a temporal contrast sensor was done by Mallik et al. [7] in 2005 and was based on the prior work of V. Gruev et al. [48] describing a pipeline temporal difference imager. Mallik's retina modified the traditional active pixel sensor (APS) to detect a quantized absolute change in illumination. It was a frame-based device with some advantages of AER systems. Several motion detection sensors have been reported [9]-[24] recently. They are not frame-based devices and have several advantages over conventional vision systems based on frames. Their most important features are high dynamic range, very high speed tracking and very low power consumption.The first specific temporal contrast sensor was proposed by Jörg Kramer [59]-[60]. It was a combination of Tobi Delbück's adaptive photoreceptor [61] with a rectifying and thresholding differentiating element [62]. After this, Lichsteiner et al. proposed an improved temporal contrast sensor [10]-[9] based on the Kramer's optical transient sensor. It has high dynamic range (120dB), low latency ($15\mu s$), good fill factor (8.1%), and low noise (2.1% contrast). Its bandwidth was limited by the photoreceptor bandwidth.

In this chapter, we present a new AER temporal vision sensor. It is based on the prior work of Lichtsteiner et al. [9]. Its pixels respond asynchronously to relative changes in intensity. The sensor directly encodes scene reflectance changes reducing data redundancy and bandwidth consumption. The new imager has improved features and pixels have smaller size. As we will discuss later, it is specially useful for very high speed tracking and surveillance applications.

In order to increase the sensitivity and the speed response of the retina, two amplifiers were added after the photoreptor stage. Both of them operate in strong inversion. By this way, the total

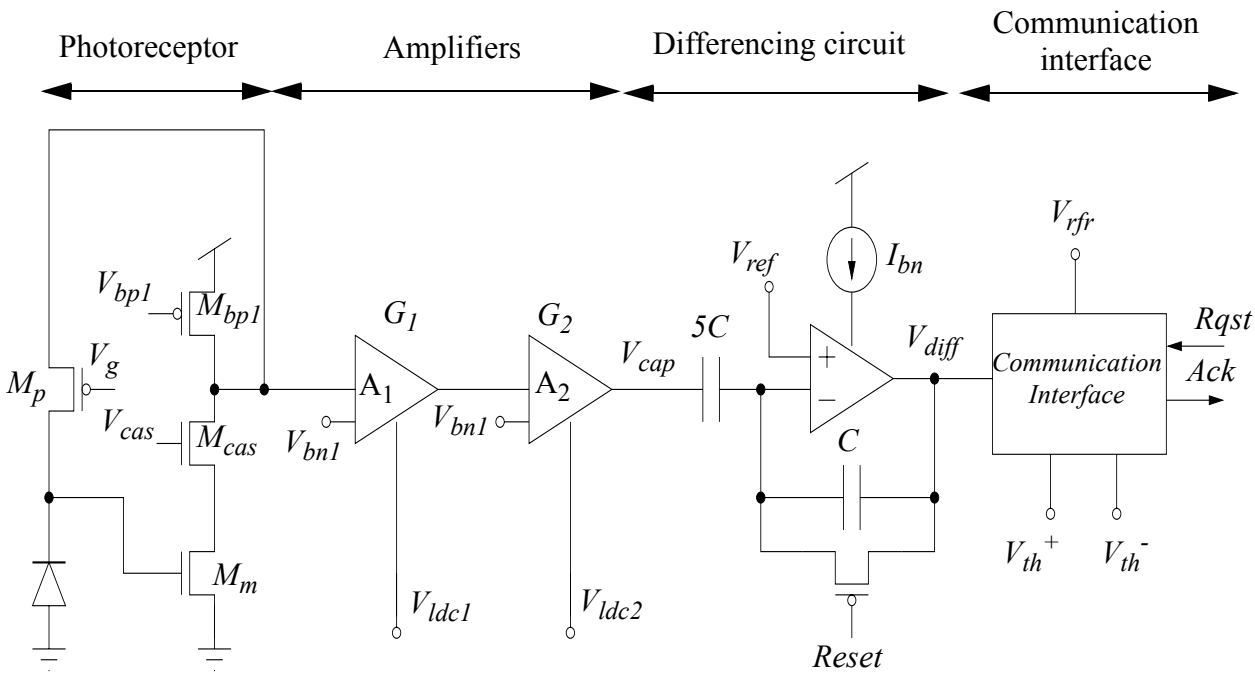


Fig 5.1: Overall block diagram of the contrast pixel

gain of the pixel can be increased over the prior design proposed by Lichsteiner and his colleagues. In that design, the total pixel gain was equal to the ratio between the two capacitances at the differencing stage. To achieve a total gain of 20 with good matching, large capacitors were required. If we use amplifiers, we can decrease the gain of the differencing circuit. This means that we can reduce the size of the capacitors and achieve a higher fill factor and higher gain at the same time. More gain also implies more bandwidth and more sensitivity to temporal contrast stimuli. For high speed applications, high gain and sensitivity are desirable because changes in local illuminance occur very fast. As we will discuss later, high gain also implies low latency response. The only drawback of using amplifiers is more power consumption. However, as we will discuss later, the total chip consumption has reasonable values in comparison to prior designs.

The new imager was fabricated and tested. A 128 x 128 pixels prototype was designed in a $0.35\mu m$ CMOS technology. Extensive experimental results are provided showing the feasibility of the design. Examples, of high speed tracking applications are also displayed in the chapter.

5.2 The Pixel Circuit

Figure 5.1 shows the overall diagram of the temporal contrast pixel. The pixel contains four main parts: 1) The photoreceptor, 2) Two amplifiers to increase sensitivity and speed response, 3) A differencing circuit to detect temporal luminance variations, 4) Communication interface with the periphery.

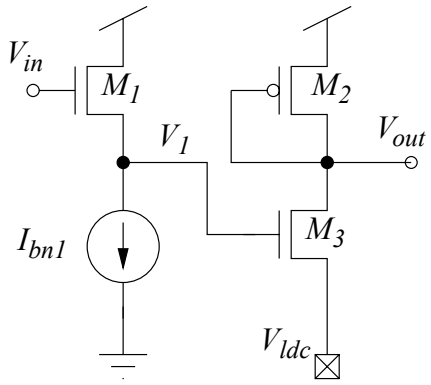


Fig 5.2: Schematic of the amplifiers.

The two amplifiers are connected to a external circuit (AGC control) that sets the gain of amplifiers to a value that does not depend on illumination.

Let us describe all these blocks that compound the pixel circuit and the external AGC circuit.

A. Photoreceptor

The first stage is the photoreptor circuit. That circuit responds logaritmically to changes in local illumination. There is an offset at the output of this stage, V_{out}, that will be amplified later. This is not a problem because it will be removed with the differencing circuit than only responds to variations in the local illumination. By adjusting V_g, V_{bp1} and V_{cas}, we can vary the speed of the photoreptor. The voltage at the output is given by the following expression,

$$V_{out} = \kappa_p V_g + n U_T \ln\left(\frac{I_{ph}}{I_{sp}}\right) = K + n U_T \ln\left(\frac{I_{ph}}{I_{sp}}\right) \tag{5.1}$$

B. Pre-amplifiers

To increase the sensitivity and the speed response of the retina, two amplifiers were added after the photoreptor stage. Both of them operate in saturation region. If we compare the new design to the prior design of Lichsteiner [9], the total gain has been increased and the pixel area has been reduced to the half. As we will discuss in section 5.4.1, FPN is similar because the ratio between gain and interpixel dispersion is approximately the same.

Figure 5.2 displays the schematics of one amplifier. Transistors M_2 and M_3 operate in strong inversion. The value of V_{ldc} is set by an external AGC circuit described in section 5.3. By this way, DC levels at the output do not depend on illumination. In Figure 5.1, we can see how amplifiers are connected to the AGC circuit. I_{bn1} is a current source that generates a bias current exter-

nally. M_1 is a source follower that introduces a voltage shift in V_{in} that can adjusted varying I_{bn1}, so $V_1 = V_{in} - \Delta V$. If M_2 and M_3 are above threshold saturation region,

$$\left(\frac{W_2}{L_2}\right)\beta_{n2}(V_1 - V_{ldc} - V_{Th})^2 = \left(\frac{W_3}{L_3}\right)\beta_{p3}(V_{dd} - V_{out} - |V_{Tp}|)^2 \tag{5.2}$$

Solving for V_{out}, and knowing that $V_1 = V_{in} - \Delta V$,

$$V_{out} = V_{dd} - |V_{Tp}| - \sqrt{\frac{(W_2/L_2)\beta_2}{(W_3/L_3)\beta_3}}(V_{in} - \Delta V - V_{Th} - V_{ldc}) \tag{5.3}$$

Where ΔV can be set by adjusting I_{bn1}. For a good operation, $\Delta V \approx 0.7V$. There are two amplifiers (A_1 and A_2) in each pixel, so the total gain at the output of the amplification stage is

$$G_T = G_1 G_2 = \sqrt{\frac{(W_{2_1}/L_{2_1})\beta_{2_1}}{(W_{3_1}/L_{3_1})\beta_{3_1}}}\sqrt{\frac{(W_{2_2}/L_{2_2})\beta_{2_2}}{(W_{3_2}/L_{3_2})\beta_{3_2}}} \tag{5.4}$$

In our particular case, both amplifiers were designed to have the same gain $G_1 \approx G_2 \approx 4$ and $\Delta V_{cap} = G_1 G_2 \Delta V_{out}$. The DC component at the output will be removed with the differencing circuit. Gain mismatch is low and mismatch because amplifiers transistors operate in strong inversion and there is a good matching between capacitors. Mismatch at the DC component does not affect the pixels output. Transistor sizes were chosen to operate in strong inversion with reduced currents (100nA-400nA). Thus, power consumption was reduced. Amplifiers are connected to the AGC block. That block is placed in the periphery and senses the average photo current. This block sets the values of I_{ldc1} and I_{ldc2}, according the global illumination level. This way, DC levels at the amplifiers output are controlled and kept within the amplifiers dynamic range.

The transistor aspect ratios were $\left(\frac{W_{2_1}}{L_{2_1}}\right) = \frac{0{,}6\mu m}{16\mu m}$, $\left(\frac{W_{3_1}}{L_{3_1}}\right) = \frac{2{,}4\mu m}{2{,}4\mu m}$, $\left(\frac{W_{2_2}}{W_{2_2}}\right) = \frac{0{,}6\mu m}{8\mu m}$ and $\left(\frac{W_{3_2}}{L_{3_2}}\right) = \frac{1{,}2\mu m}{1{,}2\mu m}$.

C. Differencing Circuit

This stage responds to relative changes at the voltage input with a gain equal to $G_3 = 5C/C$. There is a good match between the two capacitors. The differencing circuit also removes the DC component at its input. We can express the variations of voltage at the output as

$$\Delta V_{diff} = -G_3 \Delta V_{cap} = -G_1 G_2 G_3 \Delta V_{out} \tag{5.5}$$

And from eq. (5.1),

$$\Delta V_{out} = nU_T \ln\left(\frac{I_{ph}(\Delta t + t)}{I_{ph}(t)}\right) \tag{5.6}$$

The total pixel from the input to the output of this stage is $G_t = G_1 G_2 G_3$.

D. Communication Block Circuit

Figure 5.3(a) shows the schematics of the communication circuit block. Pixels can detect the sign of the temporal contrast. There are two comparators with different adjustable thresholds V_{high} and V_{low} that generate positive or negative events (signed contrast) when V_{diff} is above the positive threshold or below the negative threshold respectively. There is also a refractory circuit (see Figure 5.3(b)) that allows to control the time between consecutive events that one individual pixel can generate. This is specially useful when there are pixels with very high activity and we want to control the maximum activity in the AER bus. Finally there are two identical communication blocks, see Figure 5.3(c), that interact with the arbiter and the periphery. These are standard AER pixel communication circuits taken from Boahen's row parallel event read-out technique [45]. When generating signed events, each pixel needs to provide two column event signals *col+* and *col-*. This concept was already implemented and tested in the AER spatial contrast retina [23] and prior designs [27] that required signed events.

5.3 Automatic Gain Control Block

This block is placed at the chip periphery and allows to control the DC levels at the output of the two amplifiers adapting them to the global chip illuminance level. In Figure 5.4(a), we can see the schematics of the AGC circuitry. Its function is to maintain constant the DC levels at the output of the amplifiers of Figure 5.1, A_1 and A_2, by adjusting the voltages V_{ldc1} and V_{ldc2}. All pixel amplifiers, A_1 and A_2, are connected to the external nodes, V_{ldc1} and V_{ldc2}, respectively. The AGC control block has two identical amplifiers to the ones of Figure 5.2. The outputs of these amplifiers are set to a constant voltage, V_{odc1} and V_{odc2}, that we can adjust. To set these constant voltages, there is one operational amplifier connected to the output of each amplifier. Both of them have negative feedback to set up the voltages V_{ldc1} and V_{ldc2} that will be identical at the amplification stage and the AGC block.

The input of the circuit of Figure 5.4(a) is the summed photo current sensed by 128 photo diodes placed around the pixels matrix, by this way voltages V_{ldc1} and V_{ldc2} can be set according to the overall chip illumination level. In Figure 5.4(b) we can see the circuit that senses the average photo current level. It has 128 photo diodes (distributed along the chip periphery) connected to one photoreceptor identical to the one of Figure 5.1. In this circuit, there is also a femtoampere current mirror [36] that copies the average sensed photo current. Finally, the output of circuit of Figure 5.4(b) is connected to the input of the AGC as is indicated in Figure 5.4. There is a capacitor at this input node to make the AGC block insensitive to very quick changes at the luminance level. We are only interested in slow global changes in the average chip illumination. The value of

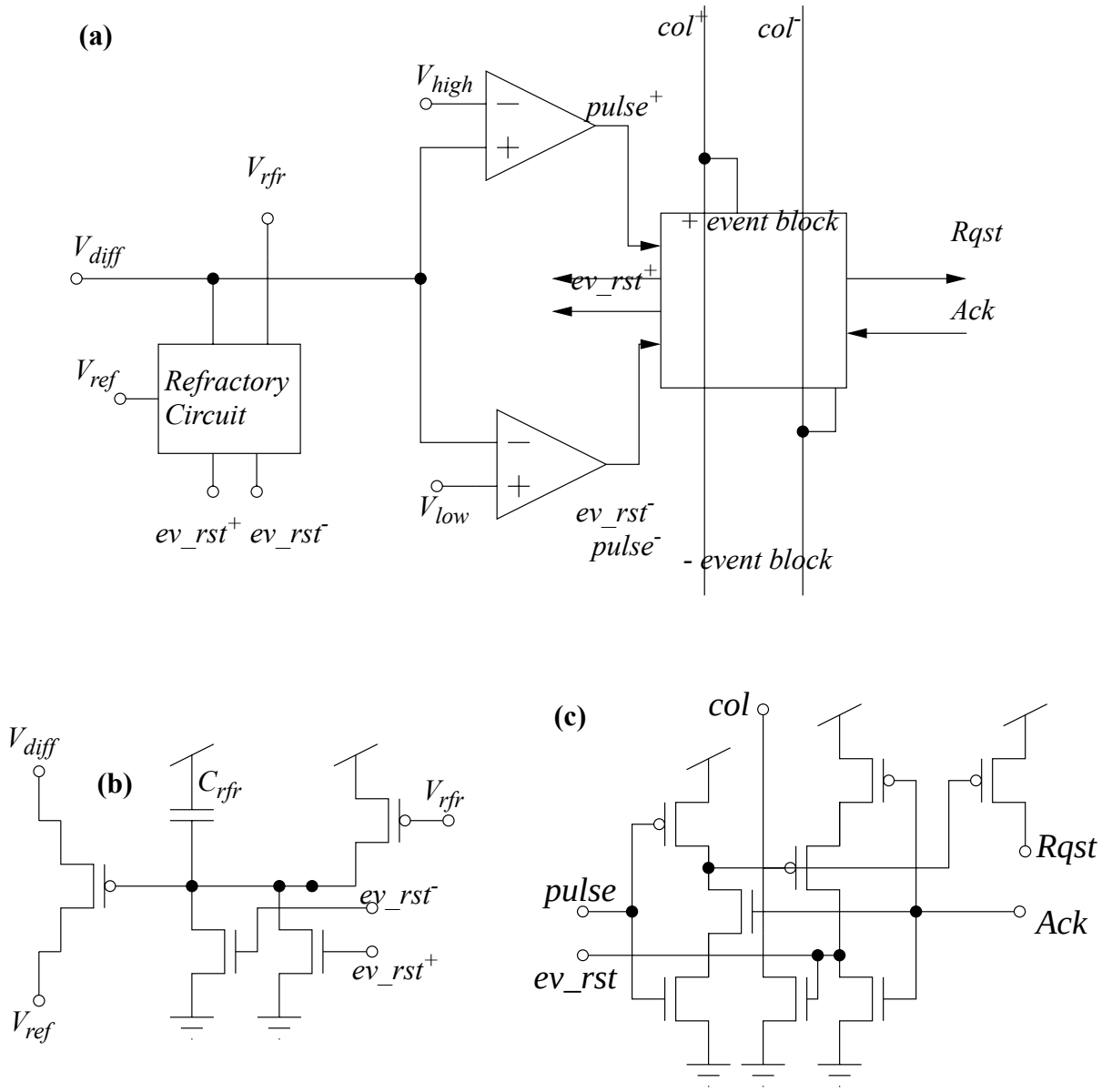


Fig 5.3: (a) Integrate and fire neuron with AER pixel intercommunication modules. (b) Detail of the refractory circuit. (c) Detail of AER pixel communication blocks.

this capacitor was chosen to have a time constant of milliseconds. This is enough to detect a global change in illumination. For example, when we go out of a house, and the average photo current value, $\overline{I_{ph}}$, changes significantly.

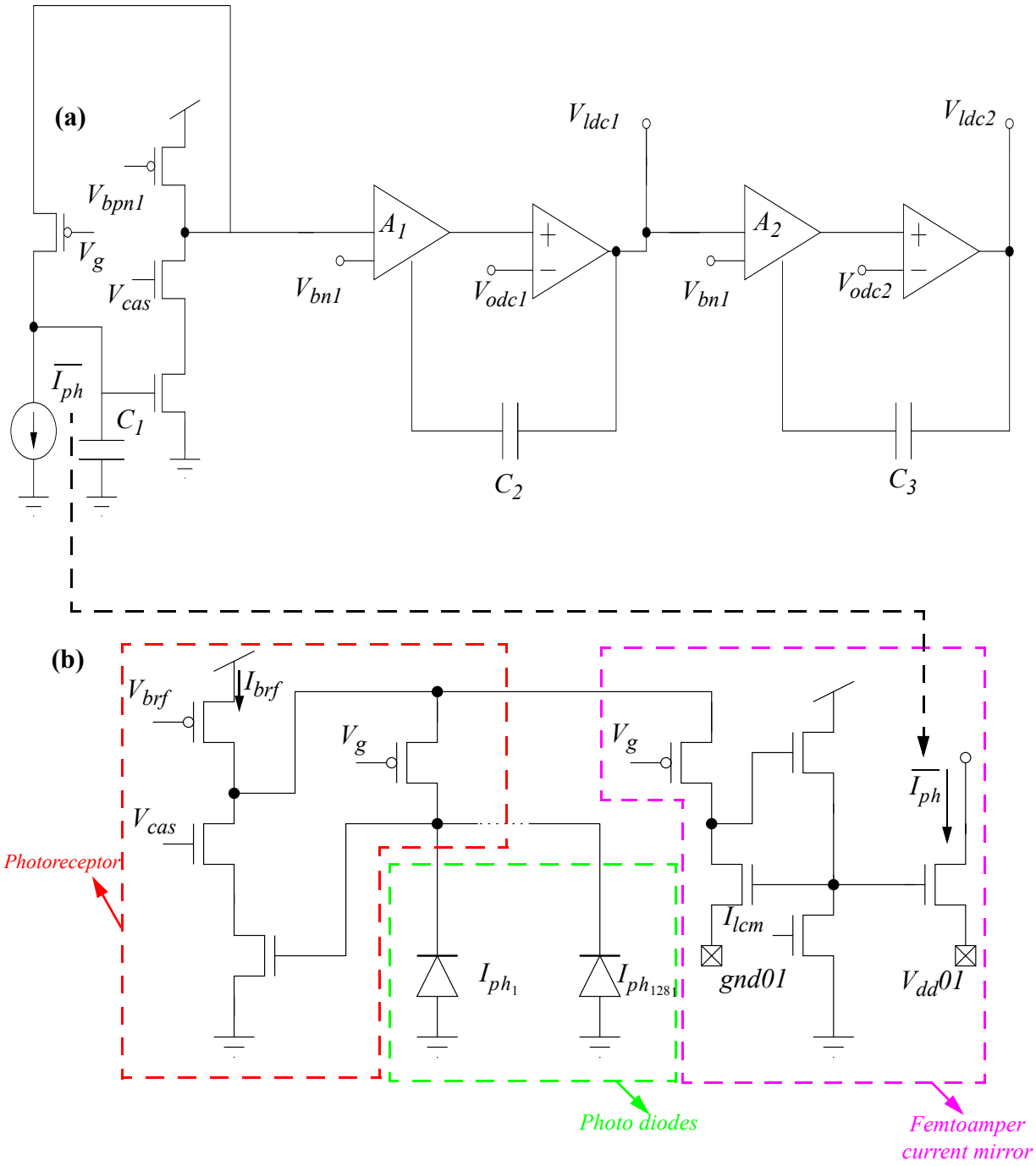


Fig 5.4: (a) AGC circuit block. (b) Detail of the circuitry used to sense the average chip illuminance, $\overline{I_{ph}}$, highlighting its main parts.

Table 5.1: Chip Specifications.

technology	CMOS 0.35μm 4M 2P
power supply	3.3V
chip size	5.54 x 5.69 mm^2
array size	128 x 128
pixel size	35.7 x 35.5 μm^2
fill factor	8%
photodiode quantum efficiency	0.34 @ 450nm
pixel complexity	58 transistors + 3 caps
current consumption	44mA @ 10keps

5.4 Experimental Results

A 128 x 128 pixel AER signed temporal retina chip was designed and fabricated in a double poly 4-metal 0.35μm CMOS process with a power supply of $V_{DD} = 3.3\,V$. Table 5.1 summarizes the chip specifications. Figure 5.5 shows a micro photograph of the die, of size $5.689x5.538mm^2$. The whole chip, except the pad ring, is covered with the top metal layer leaving openings for the photo diode sensors. Figure 5.5 also shows the layout of a single pixel highlighting its main components. Each pixel layout is a symmetrical reflection of its neighboring pixels. This way noisy digital lines are shared among neighbors, as well as power supplies, and noise sensitive bias lines. At the same time, noise sensitive lines are separated from noisy ones. Pixel area is $35.5x35.7\mu m^2$, including routing. The pixel was made up of 58 transistors and 3 capacitors (the two of the capacitive divider of the differencing circuit and the capacitance of the refractory circuit).

Figure 5.6 shows the experimental setup to characterize the temporal contrast sensor. The chip was mounted on the *Caviar Board* (a specific PCB designed to test generic AER devices). This board was connected to a USB port. The digital words to program the bias currents of each *I-pot* [35] were sent through this port. The outputs of the *Caviar Board* were sent to an AER bus. This bus was connected to a Mapper. The Mapper was a programmable device that was used to filter or remove the activity of some pixels for some specific measurements (bandwidth and latency characterization). The output bus of the Mapper was connected the board *USB2AER* [46], which sends the AER data from the bus to one of the USB computer ports. This information was processed by a jAER [47] that allows to see real time images. The *USB2AER* board was specially useful for real-time monitoring. The input bus to this board was replicated and sent to a *Datalogger*. This board has an internal memory that can store the data generated by 524,000 events and was used to save efficiently on the PC the AER information transmitted from the AER bus. To

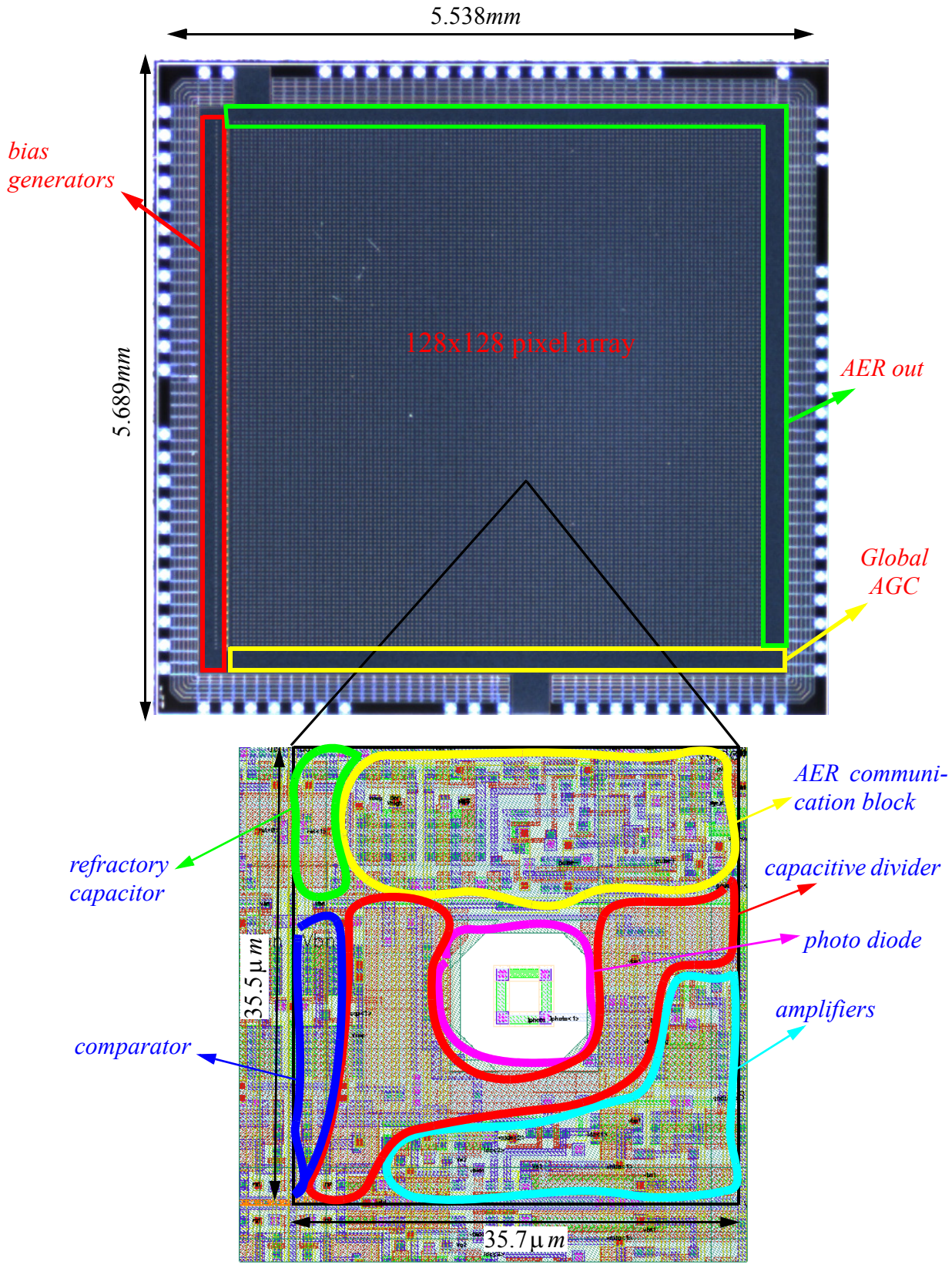


Fig 5.5: Microphotograph of 5.538*mm* x 5.689*mm* die, and zoom out of 35.7µ*m* x 35.5µ*m* pixel (layout) indicating the location of its main components.

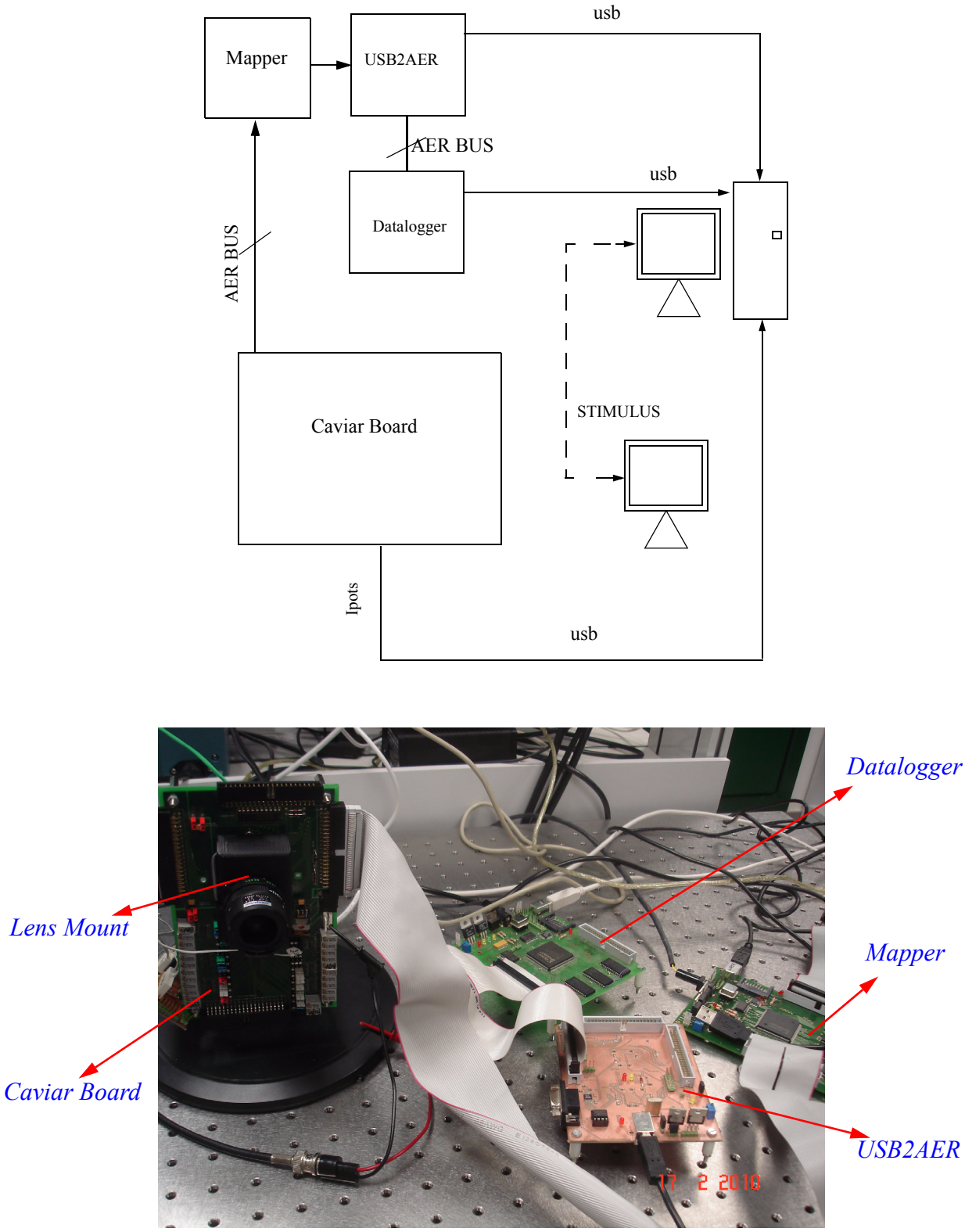


Fig 5.6: Experimental setup and photo of the retina and main boards used to test the imager.

generate programable stimuli, a monitor can be placed over the retina visual field. The stimuli can be generated with a Matlab routine.

The nominal bias settings to test the retina were $V_{ref} = 1.5\text{V}$, $V_g = 600mV$, $V_{cas} = 1.8\text{V}$, $V_{bpn_1} = 2V$, $I_{bn} = 100nA$, $I_{bldc} = 100nA$, $I_{brf} = 10nA$ and $I_{bn_1} = 200nA$. As we will discuss in section 5.4.4, the bias current I_{bn_1} can be adjusted to control the pixels bandwidth and the amplifiers power consumption.Voltages were generated with external potenciometers and bias currents were generated by using the I-Pots described in chapter 2.

5.4.1 Uniformity of Response

One important feature of imager sensors [10]-[9] is the FPN that characterizes the uniformity of the response. In our particular case, we were interested in knowing how was the pixel contrast threshold variation, θ_{ev}, to generate one event. Ideally, all the pixels should generate the same number of events when they are stimulated with the same stimulus. For one individual pixel, we can define the temporal contrast as

$$\theta = 2 \cdot \frac{I(t+\Delta t) - I(t)}{I(t) + I(t+\Delta t)} \tag{5.7}$$

Where $I(t)$ is pixel photo current value at one instant, *t*. If both photo currents are very similar, $I(t) \approx I(t+\Delta t)$, then we can approximate the temporal contrast by

$$\theta \approx \frac{I(t+\Delta t) - I(t)}{I(t)} = \frac{\Delta t}{I(t)} \tag{5.8}$$

where $\Delta t = I(t+\Delta t) - I(t)$.

From eq. (5.1), we know that a given change in the pixel photo current will produce the following change in V_{out},

$$\Delta V_{out} = nU_t \ln\left(\frac{I(t+\Delta t)}{I(t)}\right) \tag{5.9}$$

Combining eq. (5.5) and eq. (5.9), and knowing that the total gain at the output of the differenciator is $G_t = G_1 G_2 G_3$ (where G_1, G_2 and G_3 are the gains of the amplifiers and differentiator, respectively), we can express the change at the differencing circuit output ΔV_{diff} as

$$\Delta V_{diff} = -G_t U_t n \ln\left(\frac{I(\Delta t + t)}{I(t)}\right) = -G_t U_t n \ln(\theta + 1) \tag{5.10}$$

And usually $\theta \ll 1$, so ΔV_{diff} is approximately given by the following expression

$$\Delta V_{diff} = -G_t U_t n \theta \tag{5.11}$$

After the differencing circuit, there are two comparators that generate positive or negative events (see Figure 5.3(a)) depending on the value of V_{diff} and the positive and negative thresholds, V_{low} and V_{high}. If we set symmetric thresholds, one positive or negative event will be generated when there is a change in V_{diff}, ΔV_{diff_ev}, and

$$-\Delta V_{diff\_ev} = V_{high} - V_{ref} \text{ for a positive event} \tag{5.12}$$

$$-\Delta V_{diff\_ev} = V_{ref} - V_{low} \text{ for a negative event} \tag{5.13}$$

Combining eq. (5.12), (5.13) and (5.11), we can obtain an expression for the minimum temporal contrast, θ_{ev}, necessary to produce one single event (either positive or negative)

$$\theta_{ev} = \pm\frac{\left|\Delta V_{diff\_ev}\right|}{nG_t U_t} \tag{5.14}$$

And when there is a slow temporal change in the photo current, if the output voltage variation is higher than ΔV_{diff_ev}, the pixel will generate $N = \frac{\Delta V_{diff}}{\Delta V_{diff\_ev}}$ events. Hence,

$$\theta_{ev} = \pm\frac{\left|\Delta V_{diff\_ev}\right|}{nG_t U_t} = \frac{\ln((I(\Delta t + t))/I(t))}{N} \tag{5.15}$$

In order to characterize the FPN, we stimulated the retina pixels with a black bar with linear gradient edges that was moving at constant speed through the visual field of the sensor [9]. The bar was passed through the visual field 30 times. After this, the average number of events generated by each pixel, $\overline{N}$, was recorded. Ideally, if all the pixels are stimulated with the same stimulus, they should spike the same number of times. The goal of the experiment was to quantify the pixel deviation. The effect of the refractory period was reduced to the minimum. The experiment was also repeated for different positive and negative symmetric thresholds, $\left|\Delta V_{diff\_ev}\right| = [0.2, 0.24, 0.33, 0.33, 0.41, 0.46, 0.51]V$. The stimulus was generated with a LCD monitor. Its contrast and speed was controlled with a computer program. The number of events does not depend on the speed of the stimulus. It only depends on its contrast. In our particular case, the ratio between the photo currents measured at the brightest and the darkest grey levels of the stimulus was $I_{bright}/I_{dark} \approx 4$. We determined this ratio by measuring the reflectance of the brightest and the darkest grey levels. Let us define the log contrast as

$$C = \ln\left(\frac{I_{bright}}{I_{dark}}\right) \tag{5.16}$$

Therefore, an approximation for the contrast event threshold of eq. (5.15) is

$$\theta_{ev} = \frac{C}{\overline{N}} \tag{5.17}$$

Figure 5.7 shows the histograms with the average number of positive and negative events generated for each pixel of the retina. These values of the contrast event thresholds were calculated using eq. (5.17) and are indicated on each graph. For the highest threshold, there is an average of 3.6 positive and 3.6 negative events per positive and negative edge. For the lowest threshold, there is an average of 16.1 events per pixel (either positive or negative events). For lower thresholds, the average number of events and the deviation is higher.

5.4.2 Pixel Gain Measurements

From the experimental results of Figure 5.7, we can obtain the average gain of each pixel. Combining eq. (5.15) and eq. (5.11), we can express the total gain (number of events per log contrast of the stimulus) as

$$G_t = \frac{\overline{N} \cdot \Delta V_{diff\_\mathrm{ev}}}{n \cdot U_t \cdot C} \tag{5.18}$$

Figure 5.8 displays the average measured individual pixel gain values and the number of pixels which have each gain value. The estimated gain values are plotted for different values of the voltage thresholds. For all the pixels, the average gain value is $G_t \approx 65$. Ideally speaking, gain should not depend on the contrast threshold. We can see that for the highest threshold ($|\Delta V_{diff\_\mathrm{ev}}| = 0.51\mathrm{V}$, the total gain is approximately 55. The reason is that there are pixels that do not spike sometimes when the stimulus pass through the visual field and the threshold is too high. When the threshold is too low, $|\Delta V_{diff\_\mathrm{ev}}| = 0.2\mathrm{V}$, outliers that spike with very high frequency appear. For this reason, the total gain is slightly higher ($G_t \approx 70$). For the rest of the threshold values, the average gain is quite similar, $G_t \approx 65$.

5.4.3 Contrast Threshold Deviation

Let us know study the deviation in the contrast threshold from the measurements shown in Figure 5.7. If there is a variation of ΔV_{diff} higher than the contrast threshold, from equations (5.10), (5.12) and (5.13), we can obtain the number of spikes that the pixel will generate

$$N = \frac{\Delta V_{diff}}{\Delta V_{diff\_\mathrm{ev}}} = \frac{\ln(\theta + 1)}{\theta_{ev}} \tag{5.19}$$

Because of pixel mismatch not all generate the same number of events, but there will be a statistical distribution where each pixel will generate a number of events given by

$$N + \Delta N = \frac{\ln(\theta + 1)}{\theta_{ev} + \Delta\theta_{ev}} \approx \frac{\ln(\theta + 1)}{\theta_{ev}}\left(1 - \frac{\Delta\theta_{ev}}{\theta_{ev}}\right) = N\left(1 - \frac{\Delta\theta_{ev}}{\theta_{ev}}\right) \tag{5.20}$$

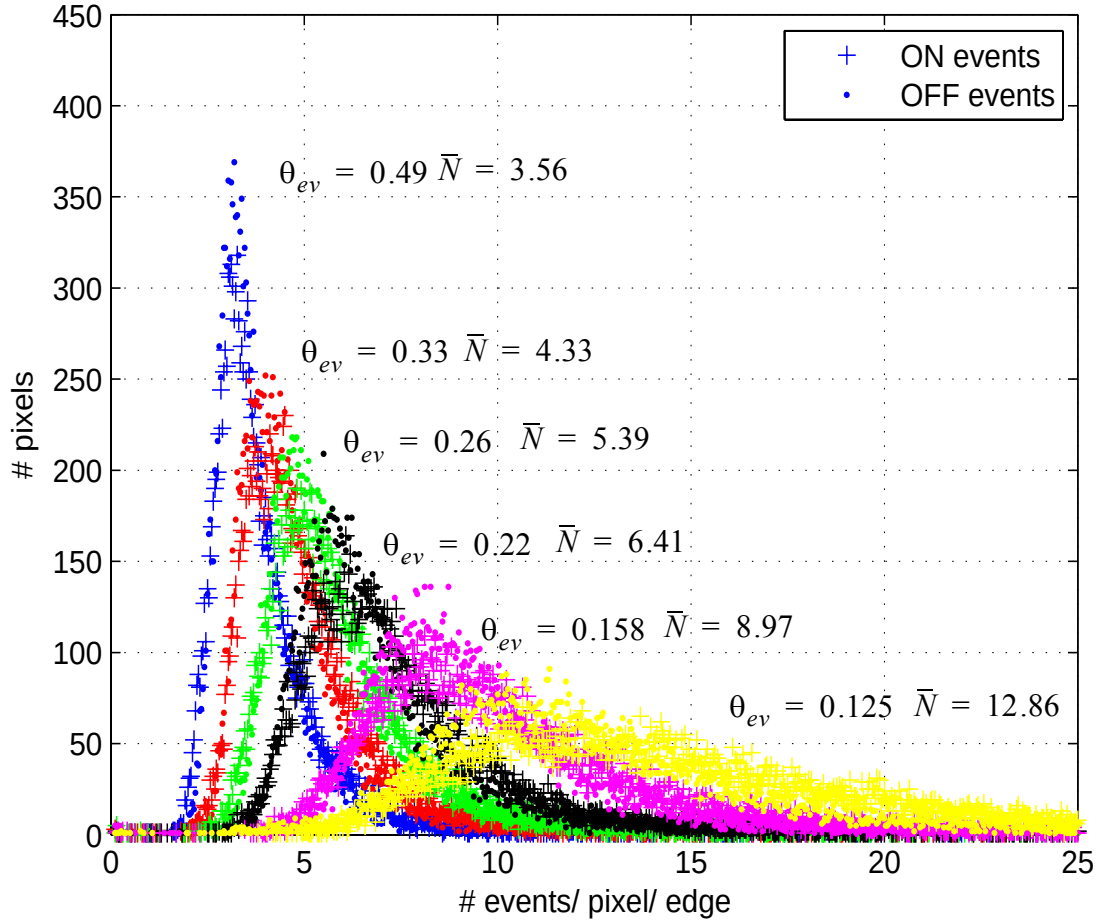


Fig 5.7: Pixel histograms with the average number of events recorded per pass of the stimulus bar for 30 repetitions. The experiment was also repeated for different values of the event threshold.

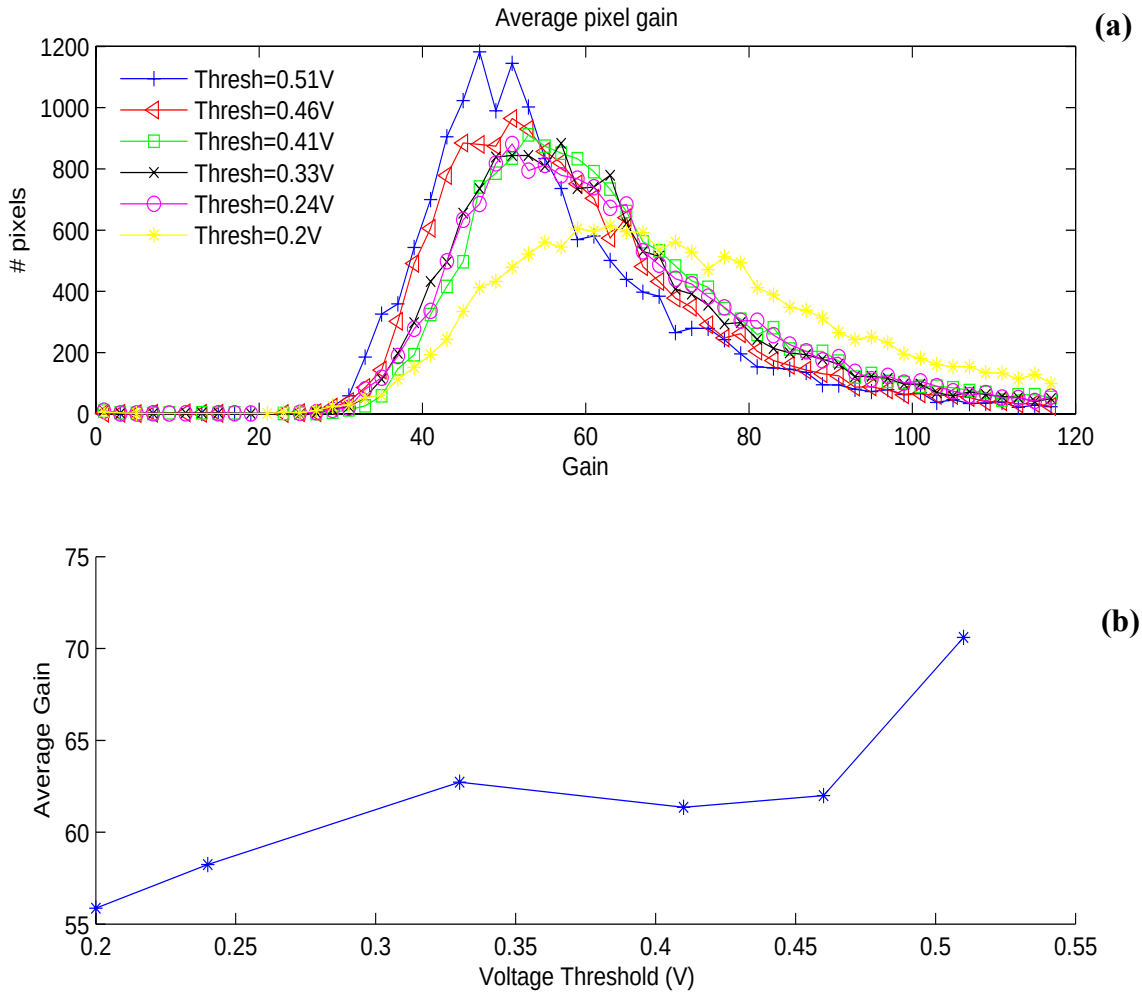


Fig 5.8: (a) Distributions in the pixels gains obtained from the measurements of Figure 5.7 and eq. (5.18). (b) Average gain values for each voltage threshold.

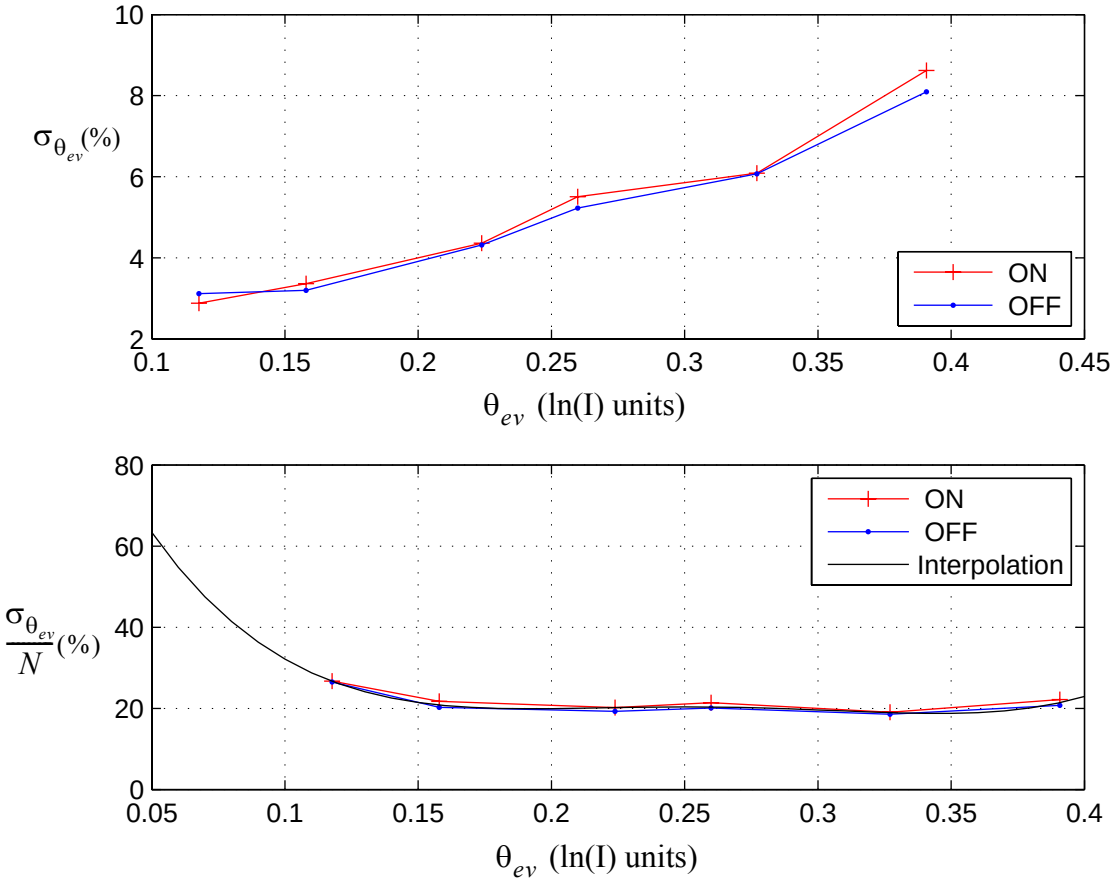


Fig 5.9: Top panel shows the contrast deviation expressed in % as a function of the contrast deviation. Bottom panel shows the ration between the contrast deviation and the average number of events (also in %).

Consequently,

$$\frac{\Delta N}{N} = -\frac{\Delta \theta_{ev}}{\theta_{ev}} \tag{5.21}$$

and

$$\sigma_{\theta_{ev}} = \frac{\theta_{ev}}{N}\sigma(\Delta N) \tag{5.22}$$

Top panel of Figure 5.9 shows the standard deviation of measured contrast threshold in % (computed from eq. (5.22) and the measurements of Figure 5.7) plotted as a function of contrast thresh-

old. At the bottom panel, we have plotted $\frac{\sigma_{\theta_{ev}}}{N}$ as a function of contrast threshold. The minimum experimental mismatch deviation (FPN) was about 2.8%.

The main sources of mismatch are the transistor threshold voltage, ΔV_{diff}, of comparators and the interpixel mismatch of the gain, G_t. If we assume independent statistical distributions, the contrast threshold can be expressed as

$$\sigma_{\theta_{ev}}^2 = \sigma_{\theta_{ev}}^2\Big|_{G_t} + \sigma_{\theta_{ev}}^2\Big|_{\Delta V_{diff}} \tag{5.23}$$

And using the chain rule, we can obtain equations for the two sources of mismatch from eq. (5.15)

$$\sigma_{\theta_{ev}}\Big|_{G_t} = -\frac{\Delta V_{diff}}{nU_t G_t} = \frac{\sigma_{G_t}}{G_t}\theta_{ev} \tag{5.24}$$

$$\sigma_{\theta_{ev}}\Big|_{\Delta V_{diff}} = \frac{\sigma_{\Delta V_{diff}}}{nU_t G_t} \tag{5.25}$$

And the total threshold deviation can be expressed as follows

$$\sigma_{\theta_{ev}}^2 = \theta_{ev}^2\sigma_A^2 + \sigma_o^2 \text{ , with } \sigma_A = \frac{\sigma_{\theta_{ev}}\Big|_{G_t}}{G_t} \text{ and } \sigma_o = \frac{\sigma_{\Delta V_{diff}}}{nU_t G_t} \tag{5.26}$$

In Figure 5.10, we have plotted the probability density function of the contrast threshold to produce one output event, θ_{ev}, assuming that has a Gaussian distribution. Let us denote as θ_{evn} the mean value of θ_{ev} for all the retina pixels. Looking at Figure 5.10, if we want the 99% of the pixels to spike when there is a contrast stimulus, the minimum detectable contrast will be

$$\theta_{min} = \theta_{evn} + 3\sigma_{\theta_{ev}} = 6\sigma_{\theta_{ev}} \tag{5.27}$$

We have to highlight that if there were pixels with a value of $\theta_{ev} = 0$, they would be always spiking, even if there is no contrast. Therefore, the minimum detectable threshold will be

$$\theta_{ev} = 3\sigma_{\theta_{ev}} \tag{5.28}$$

From equations (5.27) and (5.28), we can state that,

$$\theta_{min} = 2\theta_{ev} \tag{5.29}$$

And substituting equations (5.27) and (5.29) in (5.26), we can deduce the value of the minimum detectable contrast is

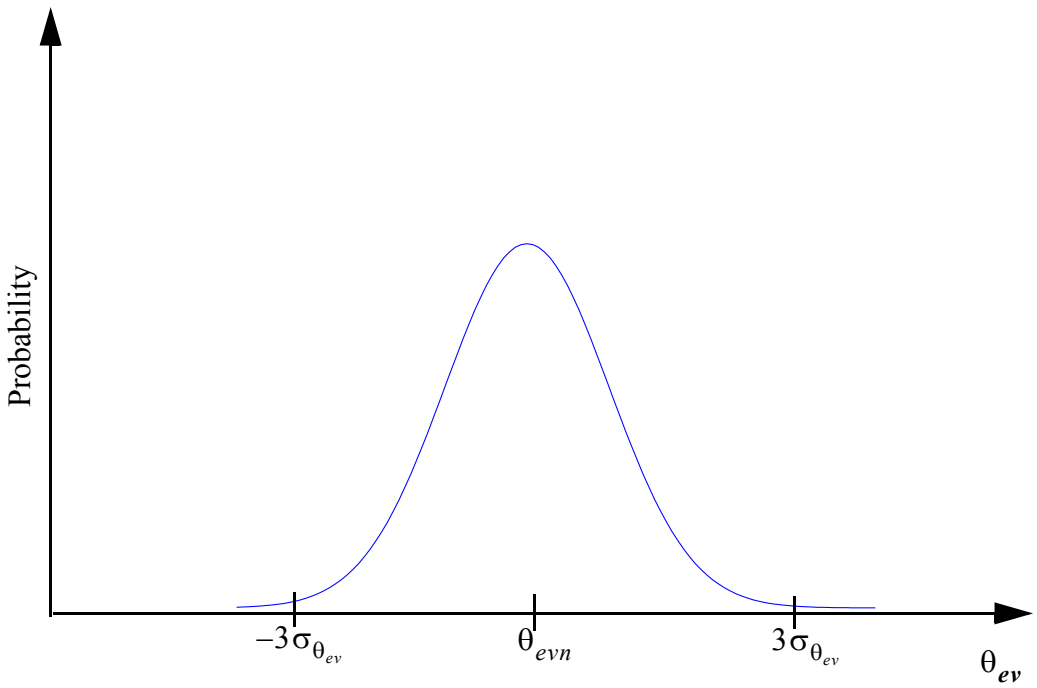


Fig 5.10: Probability density function of the contrast threshold, θ_{ev} .

$$\theta_{min} = 2\theta_{ev} = \frac{2\sigma_o}{\sqrt{\frac{1}{9} - \sigma_A^2}} \tag{5.30}$$

In our particular case, we can compute the values of σ_o and σ_{G_t} from the experimental results of Figure 5.9 and substitute them in eq (5.26) to obtain the minimum detectable contrast θ_{min}. Note that in (5.26), for low values of θ_{ev}, $\sigma_{\theta_{ev}} \approx \sigma_o$, because mismatch at the offset of comparators is the main source of mismatch. Therefore, $\sigma_{\theta_{ev}}$ would be approximately equals to the minimum value of the curve at top panel of Figure 5.8. For higher values of θ_{ev}, $\sigma_{\theta_{ev}} \approx \theta_{ev}\sigma_A$. Thus, $\sigma_{\theta_{ev}}$ is approximately the slope of the curve at top panel of Figure 5.8. In this case, amplifiers gain is the main source of mismatch. Looking at Figure 5.8, $\sigma_A \approx 0.16$ and $\sigma_o \approx 0.025$. Therefore, the minimum contrast that the 99% of the pixels can detect is

$$\theta_{min} = \frac{2\sigma_o}{\sqrt{\frac{1}{9} - \sigma_A^2}} \approx 16.9\% \tag{5.31}$$

And the comparators offset is given by

$$\sigma_{V_{offset}} = \sigma(\Delta V_{diff}) = G_t U_t n \sigma_{\theta_{ev}}\big|_{\sigma_{G_t}=0} = G_t U_t n \sigma_o = 30mV \tag{5.32}$$

According to eq. (5.27), ΔV_{diff} has to satisfy $|\Delta V_{diff}| > 6 \cdot 30 = 180mV$. On the contrary, some pixels will be always spiking. Experimentally, we could only set values of $|\Delta V_{diff}|$ above 0.2V. For lower values of voltage thresholds, the event rate was too high (there were pixels spiking all the time with $\theta_{ev} = 0$) and the arbiter was not able to send out all the events properly.

5.4.4 Bandwidth Measurements

One of the more impressive features of the sensor is the pixel bandwidth. If we calculate the transfer functions of the main subsystems that compound the pixel (photoreceptor, amplifiers and differencing circuit), we can understand the dynamic response of retina pixels. Let us study all these blocks separately to analyze their influence on the dynamic response and obtain their transfer functions. The overall transfer function will be equal to the product of these individual transfer functions.

As we will show, the photoreceptor limits the pixel bandwidth. In Figure 5.11(a), the photoreceptor and its main parasitic capacitances are shown. If we calculate the transfer function between the input and output nodes (V_{in} and V_{out}), taking into account the main parasitic capacitances, we obtain a second order low-pass filter with a transfer function given by the following expression:

$$H_1(s) = -\frac{g_{m_m} g_{m_{cas}}}{C_d C_1 s^2 + g_{m_p} C_2 s + g_{m_{cas}} g_{m_p}} = -\frac{g_{m_{cas}} g_{m_m}}{C_d C_1} \cdot \frac{1}{s^2 + \frac{g_{m_p} C_2}{C_1 C_d} s + \frac{g_{m_p} g_{m_{cas}}}{C_1 C_d}}, \tag{5.33}$$

$$\text{with } C_1 = C_{gs_p} + C_{gd_{bp1}} + C_{gd_{cas}} \tag{5.34}$$

$$\text{and } C_2 = C_{gs_{cas}} + C_{ds_m} + C_d \tag{5.35}$$

If we calculate the poles of eq. (5.33),

$$s_{12} = -\frac{g_{m_p} C_2}{2 C_1 C_d} \pm \frac{1}{2} \cdot \sqrt{g_{m_p}^2 C_2^2 - 4 g_{m_p} g_{m_{cas}} C_1 C_d} \tag{5.36}$$

Under typical operation conditions, $g_{m_p} < g_{m_{cas}}$ because transistor M_g operates in weak inversion and transistor M_{cas} operates in strong inversion. Hence, $g_{m_p}^2 C_2^2 < 4 g_{m_p} g_{m_{cas}} C_1 C_d$ and poles are complex conjugates. However, there might be situations were poles were real number, depending of the chip illumination and bias settings. A generic second-order low-pass filter with complex conjugates poles can be expressed as,

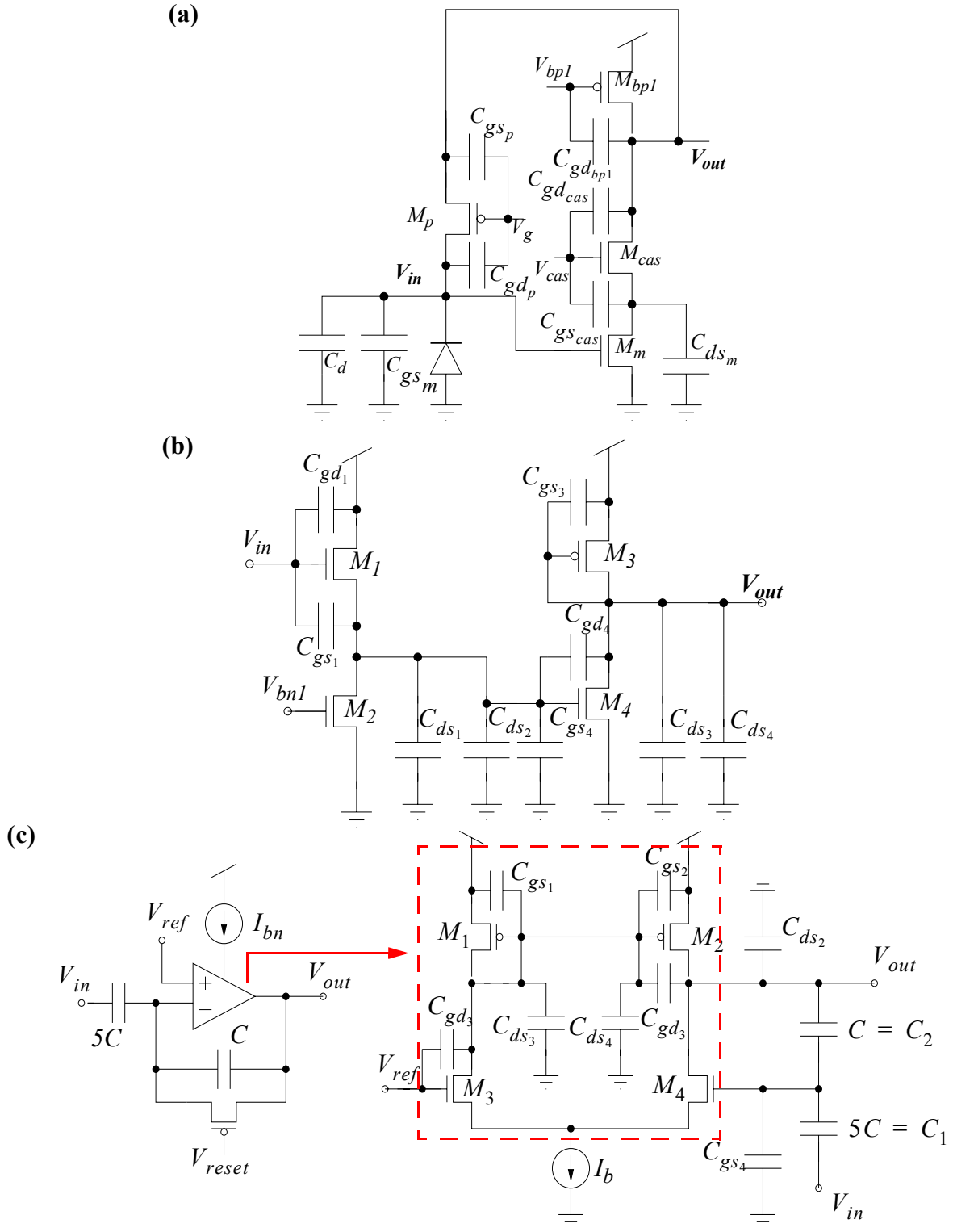


Fig 5.11: Schematics of the main pixel blocks used to calculate the pixel transfer function. The main parasitic capacitances have been highlighted in each schematic. (a) Photoreceptor, (b) Amplifiers, and (c) Differencing circuit with circuitry detail including the amplifier and its main parasitic capacitances.

$$H(s) = \frac{k}{s^2 + \frac{w_o}{Q}s + w_o^2} \tag{5.37}$$

We can obtain the filter cut-off frequency, assuming complex poles,

$$w_{o_1} = \sqrt{\frac{g_{m_{cas}} g_{m_p}}{C_d C_1}} \tag{5.38}$$

Studying eq. (5.38), we can state that pixel bandwidth depends of transistor M_m conductance. Transistor M_m works in weak inversion. Conductance changes with illumination because, $g_{m_p} \approx \frac{I_{ph}}{U_T}$. For this reason, pixel bandwidth is proportional to $\sqrt{I_{ph}}$. For low luminance levels, g_{m_p} has low values and pixel bandwidth is lower. We also have to highlight that C_d, the photo diode parasitic capacitance has a higher value than other parasitic capacitances. For this two reasons, this stage limits pixel bandwidth. Figure 5.12 displays the bode diagram of the photoreceptor transfer function. We can also see how bandwidth changes when we vary illumination conditions. The phase is approximately flat at the frequencies of interest and equal to -90º because the sign of the output voltage is inverted as is depicted in eq. (5.33).

In the amplification stage, we find two identical amplifiers. In Figure 5.11(b), there we can see the schematics of one of them. The main parasitic capacitances are indicated in the figure. The current source of Figure 5.2 has been replaced by transistor M_2 in Figure 5.11(b). This transistor is part of a current mirror that feeds the amplifier. The transfer function of each amplifier is approximately given by,

$$H(s) = -\frac{-s^2 + \frac{C_{gs_1} g_{m_4} - C_{gd_4} g_{m_1}}{C_{gs_1} C_{gd_4}} s + \frac{g_{m_p} g_{m_p}}{C_{gs_1} C_{gd_4}}}{s^2 + \frac{g_{m_3} C_3 + g_{m_4} C_{gd_m} + g_{m_1} C_4}{C_5^2} s + \frac{g_{m_1} g_{m_3}}{C_5^2}} \tag{5.39}$$

$$\text{With } C_5^2 = C_{gd_4}(C_1 + C_{gs_1}) + C_2(C_{gs_1} + C_{gd_4} + C_1), \tag{5.40}$$

$$C_4 = C_{gd_4} + C_2 \tag{5.41}$$

$$C_3 = C_{gs_1} + C_{gd_4} + C_1 \tag{5.42}$$

$$C_2 = C_{ds_3} + C_{ds_4} \tag{5.43}$$

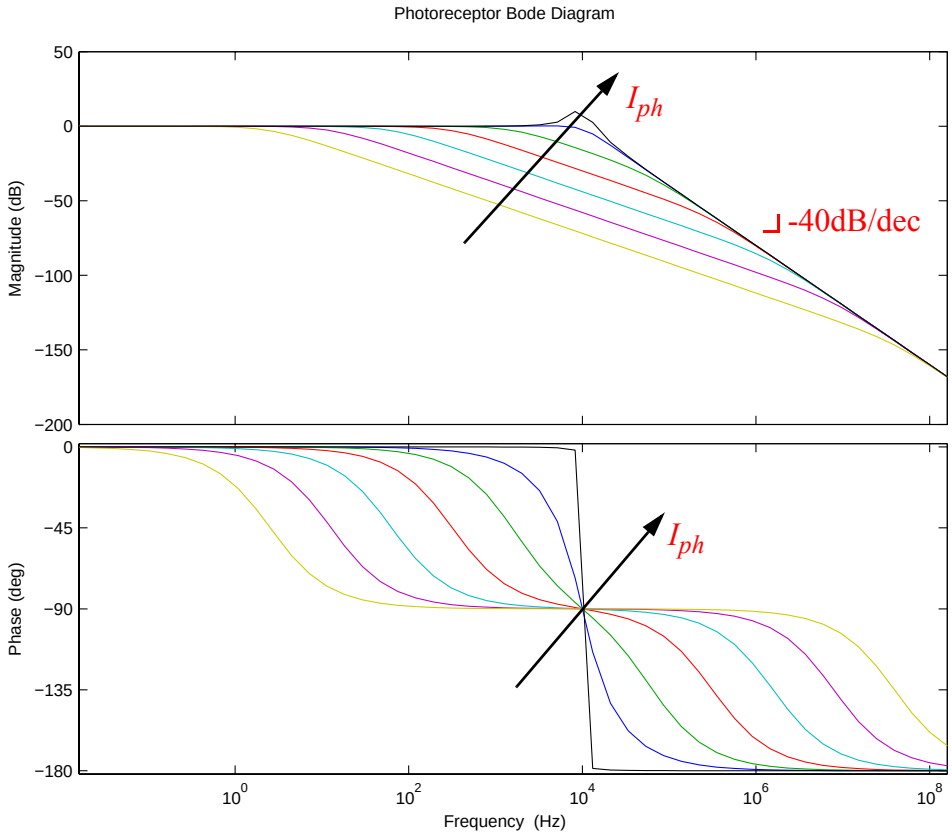


Fig 5.12: Bode diagram of the photoreceptor transfer function. It is shown how illumination affects pixel bandwidth.

$$C_1 = C_{ds_1} + C_{ds_2} + C_{gs_4} \tag{5.44}$$

Let us denote $\tau = g_{m_3} C_3 + g_{m_4} C_{gd_m} + g_{m_1} C_4$, the poles of eq. (5.39) are equal to the following expression

$$s_{12} = -\frac{\tau}{C_5^2} \pm \frac{1}{2C_5^2} \cdot \sqrt{\tau^2 - 4 g_{m_3} g_{m_1} C_5^2} \tag{5.45}$$

Assuming that $g_{m_1} \approx g_{m_3} \approx g_{m_4}$, poles are complex conjugates because $\tau^2 < 4 g_{m_3} g_{m_1} C_5^2$ (C_5 is the dominant capacitance). Figure 5.13 shows the bode diagram of the transfer function of each amplifier. It is approximately a low-pass filter. Amplifiers have a constant gain within the band pass. For higher frequencies, input signals are attenuated. In this case, the cut-off frequency is

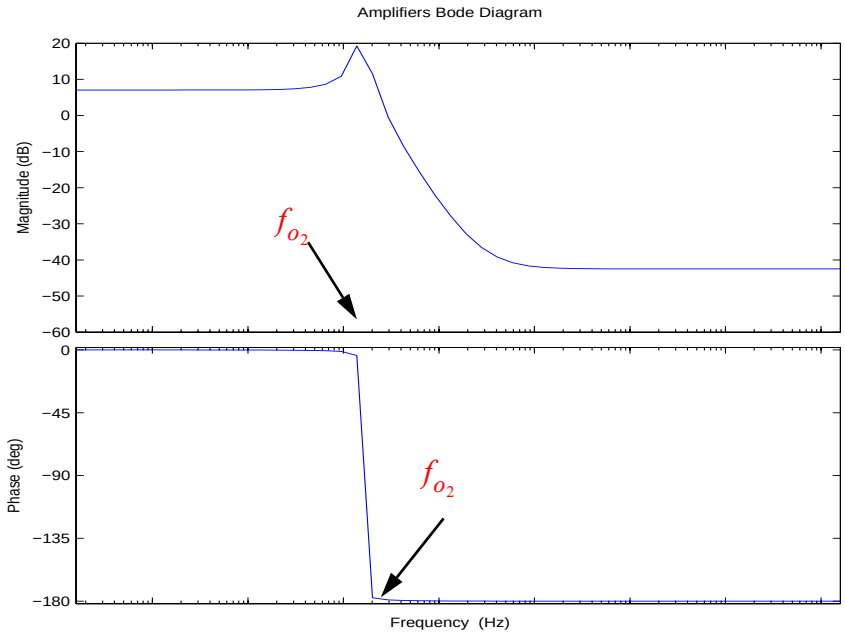


Fig 5.13: Bode diagram of amplifiers transfer function.

$$w_{o_2} = \sqrt{\frac{g_{m_3} g_{m_1}}{C_{gd_4}(C_1 + C_{gs_1}) + C_2(C_{gs_1} + C_{gd_4} + C_1)}} = \sqrt{\frac{g_{m_3} g_{m_1}}{C_5^2}} \tag{5.46}$$

Amplifiers have been designed to have a bandwidth higher than the photoreceptor stage, $w_{o_2} > w_{o_1}$. C_5^2 in eq. (5.46) is lower than the product of capacitances $C_d C_1$ in eq. (5.38). Numerator of eq. (5.46) is also lower than numerator of eq. (5.38) because amplifiers transistors M_1 y M_3 work in strong inversion. Therefore, we can assume that this stage only amplifies the input signal coming from the photoreceptor.

Figure 5.11(c) displays the differencing circuit and its main parasitic capacitances. The amplifier symbol of Figure 5.1 has been replaced by its schematics. In this case, the transfer function is given by the following expression

$$H_3(s) = -\frac{C_3}{C_4} \cdot \frac{-s^2 - \frac{g_{m_1}}{C_3} s + \frac{g_{m_1} g_{m_4}}{C_2 C_3}}{s^2 + \frac{g_{m_1}}{C_4} s + \frac{g_{m_1} g_{m_4}}{C_1 C_4}} \approx \frac{C_1}{C_2} \cdot \frac{s + \frac{g_{m_1}}{C_3}}{s - \frac{g_{m_1}}{C_4}} \tag{5.47}$$

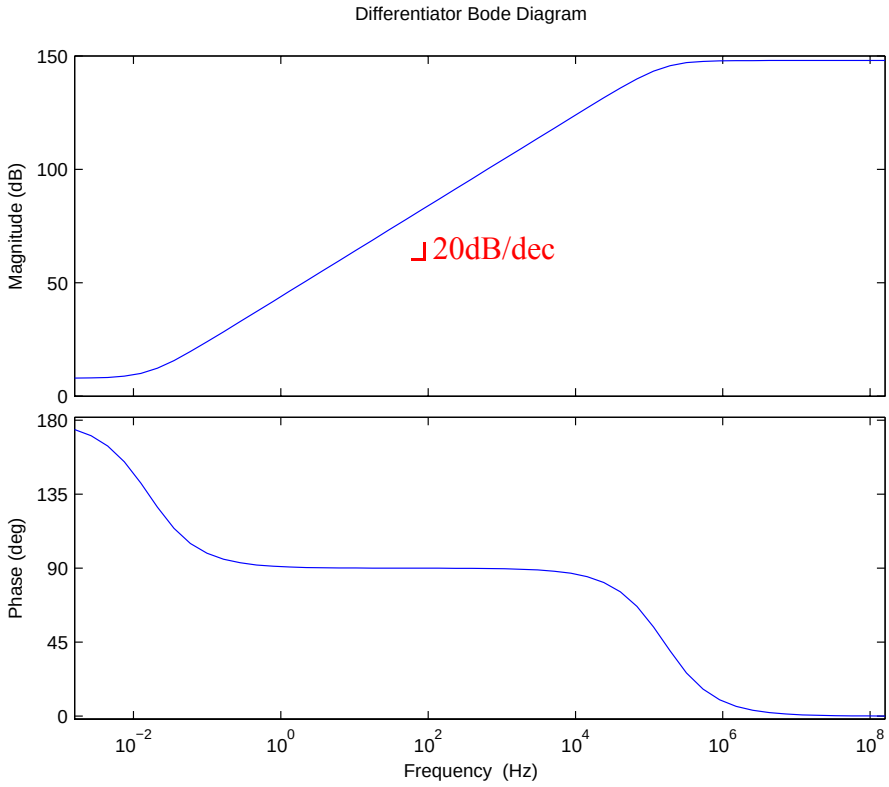


Fig 5.14: Bode diagram of the differencing circuit transfer function.

C_1 and C_2 are the external capacitors of the differencing circuit. The ratio between them is equal to the gain of this stage. In our case, $G_3 = C_1/C_2 \approx 5$ and $C_3 > C_4$. Hence, $\frac{g_{m_1}}{C_3} < \frac{g_{m_1}}{C_4}$ In Figure 5.14, we can see the bode diagram of its transfer function. The system work as a differentiator from DC to frequencies higher than w_{o_1}. This stage does not limit the pixel bandwidth and also introduces and additional gain.

In conclusion, the logarithmic photoreceptor limits pixel bandwidth. This depends on illumination because the conductance of transistor M_p in Figure 5.1 is proportional to illumination, $g_{m_p} = \frac{I_{ph}}{U_T}$. We can consider pixels as a second-order systems. The total transfer function is given by the product of the three transfer functions, $H(s) = H_1(s)H_2(s)H_3(s)$. It can be approximated as $H(s) \approx G_1 G_2 G_3 H_1(s) = G_t H_1(s)$, where G_1 and G_2 are the amplifiers gain and G_3 is the differencing circuit gain.

The photoreceptor has been chosen to improve the pixel bandwidth. In the prior design of Lichsteiner [9], there was a Miller capacitance between the input and the output node. With the new configuration, this limitation has been solved.

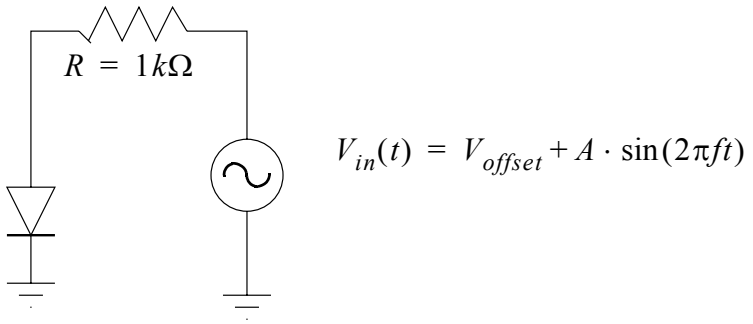


Fig 5.15: LED modulated with a sinosoidal wave used to take the bandwidth measurements.

To take the bandwidth measurements, a group of pixels of the center of the matrix were stimulated with a LED. The diode was modulated with a sinosoidal signal (see Figure 5.15). The signal has an offset level to keep the LED always ON, so $V_{offset} - A > V_{\gamma}$. We were varying the modulation frequency and counting the number of positive and negative events generate per stimulus cycle during a scan period of 20s. In the band pass, this ratio remains constant. For frequencies higher than the cut-off frequency, this ratio becomes smaller. In Figure 5.16, there are some transfer functions of several pixels. The *x-axis* represents the stimulus frequency and the *y-axis* the number of events (positive and negative) per cycle of the sinewave. Measurements were taken at 1*Klux* with indoor light. Some pixels even exhibit a small resonant peak about 2*KHz*. There is a mismatch between the pixels bandwidth. All the pixel transfer functions have a bandwidth higher than *5KHz*. Some pixels display a bandwidth even higher than *10KHz*. The average pixel bandwidth was about *8KHz*. This gives an idea about how fast can be the temporal changes that the retina can detect. We set symmetric positive and negative thresholds to take the measurements.

Figure 5.17 depicts how the pixel bandwidth depends on illumination. As we explained previously, pixel bandwidth is proportional to photo current when transistor M_p is working in weak inversion. In Figure 5.17, initially, illumination was 1*Klux* (for higher values of illumination, the shape of the pixel transfer functions does not change significantly and pixel bandwidth remains constant) and pixel bandwidth was about 6*KHz*. Then, the event transfer function was measured several times by using neutral density filters to decrease the illumination level. Bandwidth is proportional to illumination. For low illumination levels (10 lux), the pixel still has a bandwidth of 300*Hz* that is enough to detect very quick changes of illumination in the scene. In Figure 5.17, top panel shows experimental event transfer functions for different illumination levels. Pixels have approximately a first-order response. Fitting measured data with first-order transfer functions, we achieve a good matching for all the illumination values. We have plotted in red the first-order transfer functions that fit the data of the curves measured with lowest and highest illumination levels. For the rest of the curves, there is also a good matching with first-order transfer functions. Bottom panel displays how pixel bandwidth depends on illumination. We have plotted pixel bandwidth as a function of illumination. The linear fitting of the measurements is also displayed. For illumination levels higher than 1*Klux*, pixel bandwidth remains constant and does not change significantly. Ideally speaking, amplifiers gain should be the same within the band-pass. However, there are small variations within the band-pass of the number of events per cycle for each value of illumination. These variation could be due to a non perfect tune of the external voltage V_{ldc}.

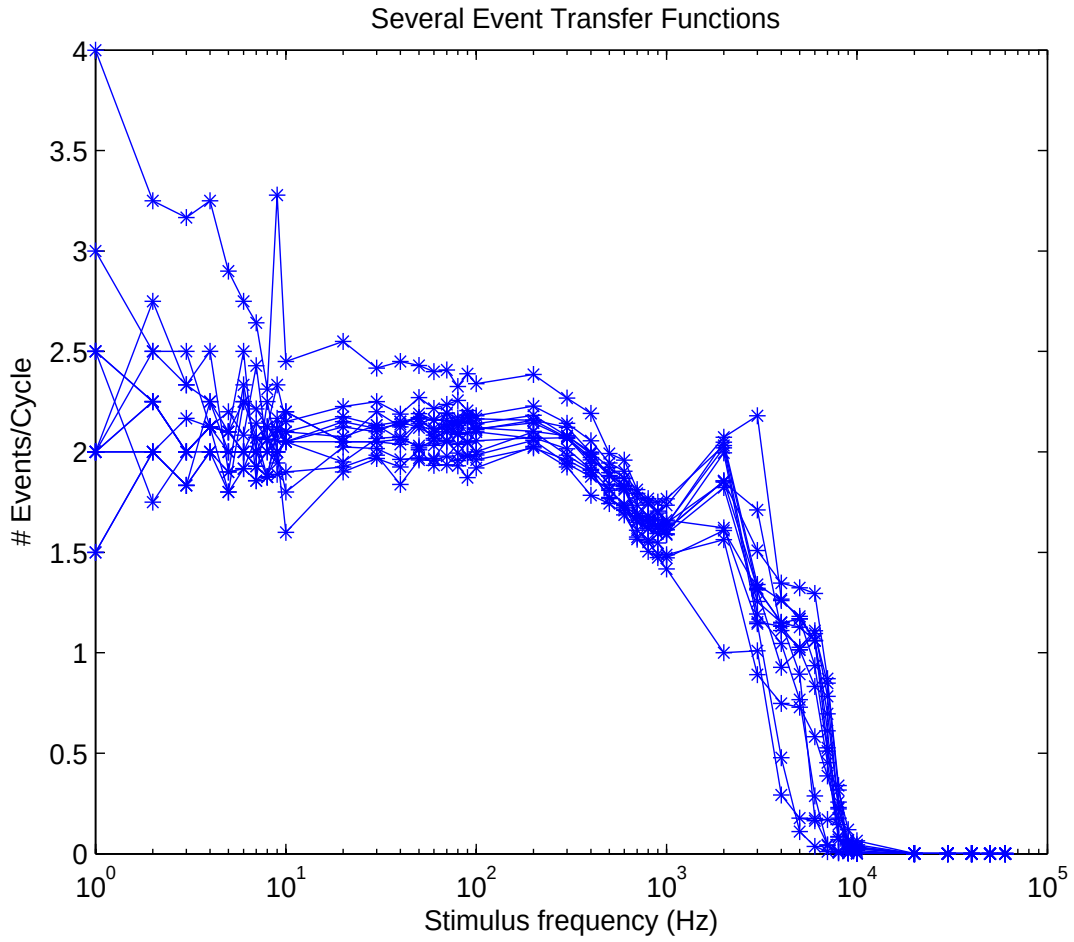


Fig 5.16: Several event transfer functions obtained when a group of pixel of the center of the retina were stimulated with a LED, as is shown in Figure 5.15.

It is possible to exert some control over the pixels bandwidth and the power consumption. Obviously, the more bandwidth the more consumption. For some applications, it is not necessary to have a pixel bandwidth higher than $1KHz$. Figure 5.18 illustrates how the pixel bandwidth changes when we vary the bias current I_{bn_1} that controls the gain of the amplifiers. For bias values higher than $400nA$, pixel bandwidth remains constant. For a value of $I_{bn_1} = 80nA$, we still have a pixel bandwidth of $800Hz$.

It is possible to control the bandwidth varying other parameters. For example, varying V_g or V_{bp_1} in the photoreptor. However, it is desirable to keep constant these bias and control the power consumption and the pixel bandwidth varying the gain of the amplifiers. Its amplifiers work in saturation region and chip consumption can be reduced significantly by varying amplifiers bias currents. Bias settings at the photoreceptor has a more restrictive tuning range and bad tuning of them could affect the pixel response.

In order to characterize the pixel response to very low stimulus frequencies [9], we repeated the bandwidth results of Figure 5.16 to determine the pixel bandwidth. In this case, we increased

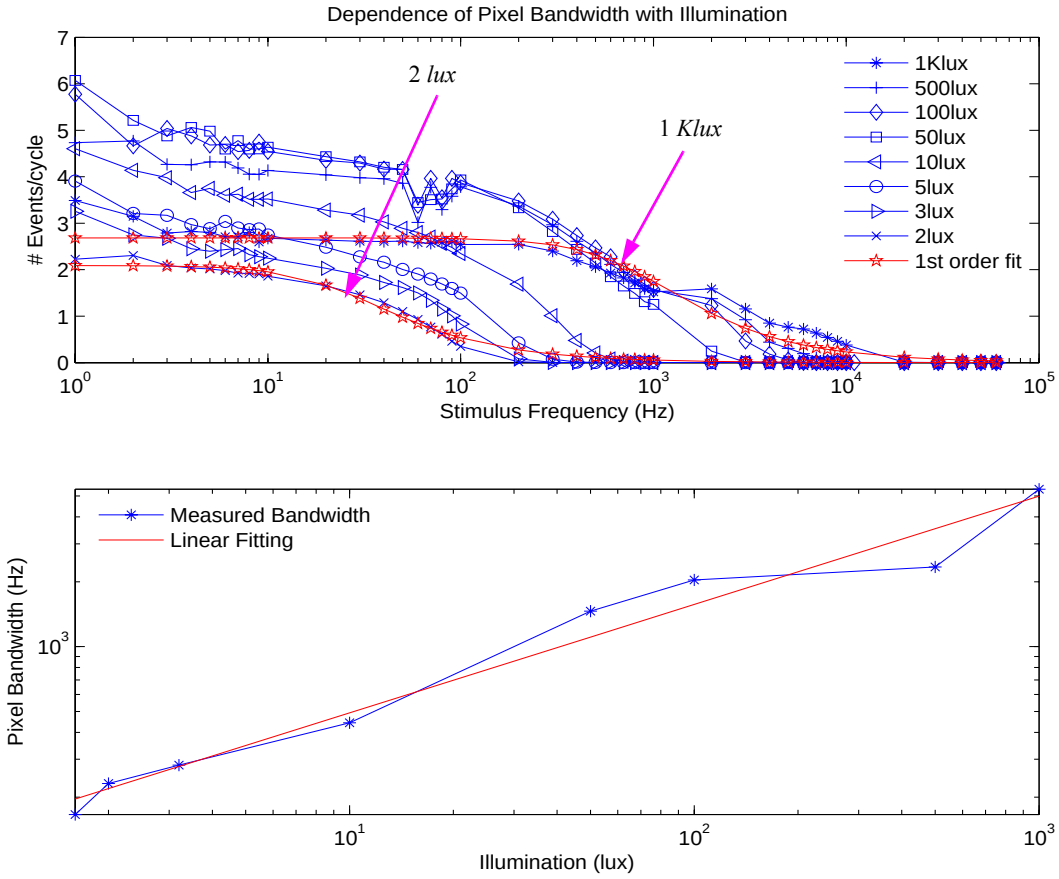


Fig 5.17: Top panel shows several transfer functions of one individual pixel when varying chip illumination. Initially, illumination was about 1*Klux*. We used neutral density filters to decrease it. Legend indicates the chip illuminance in lux. Pixels transfer functions are approximately first-order filters. We have plotted in red the first-order transfer functions that fit the measured data for the highest and the lowest illumination values. Bottom panel displays how pixel bandwidth depends on illumination. A linear fitting of the bandwidth measurements is also displayed.

the scan period to 240s to assure that at least a few events per cycle were generated during the scan period. The leakage current of the reset transistor in the differencing circuit generates positive events with a frequency of 46*mHz*. These events have to been take into account when we study the pixel response to very low frequencies. Figure 5.19 shows the average number of positive and negative events generated by one pixel per stimulus cycle. We have also plotted a third curve showing how the leakage current of the reset transistor affects the measurements of positive events. The effective number of positive events per cycle is equal to the difference between the measured events per cycle and the number of positive events generated due to the leakage current during the scan period. Looking at Figure 5.19, we can state that the retina can detect changes of illumination of 10*mHz*. According to the results of Figure 5.17, the pixel frequency range of operation is approximately $[10^{-2}, 7 \cdot 10^{3}]Hz$.

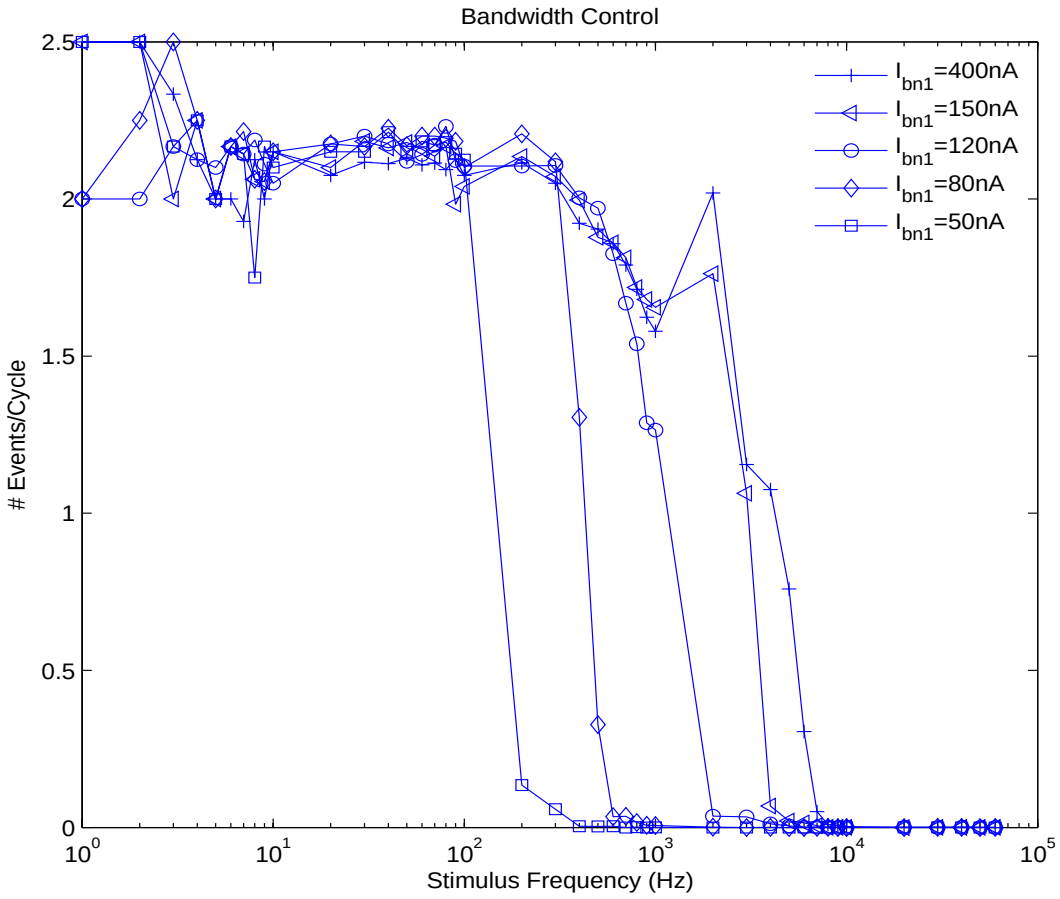


Fig 5.18: Bandwidth control varying the bias current I_{bn_1}.

5.4.5 Latency Characterization

As we explained in section 4.5.6, latency is the amount of time that it takes from the occurrence of a illumination change at the photo diode and the output of an address event. This imager is specially useful for high speed applications an it could be easily used as part of a feedback controller. Latency can be considered as effective measurement of the pixel bandwidth. It also depends on the illumination level. For low values of illumination, it is reciprocal with illuminance. Figure 5.21 displays the latency measurements of one individual pixel under different illumination conditions. Error bars indicate the deviation (jitter) between different measurements. Under standard operation conditions, latency is always below $1ms$. The lowest experimental value of the latency was $3.67\mu s$ at *25Klux*. Latency measurements were taken under nominal bias settings with $I_{bn_1} = 200nA$ and $|\Delta V_{diff}| = 0.2\text{V}$. If we set low thresholds, we will reduce the latency period.

To take latency results, we use a LED (as is shown in Figure 5.15). In this case, the diode was initially OFF. Then, we illuminated it using a step signal. The latency was the delay between the step signal and the first RQST of the pixel under test in response to the luminance change. To

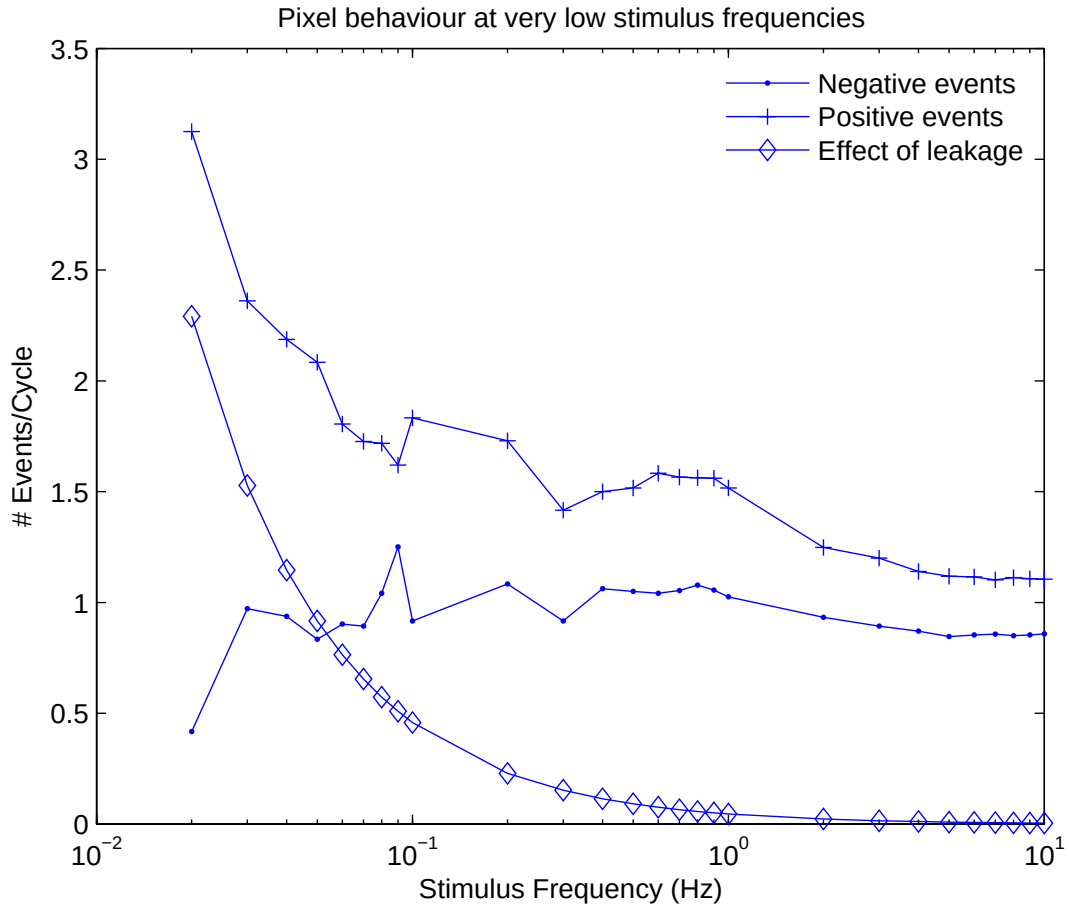


Fig 5.19: Pixel response at very low frequencies and influence of leakage current on measured positive events for each frequency.

avoid confusions with the request signals sent out by other pixels, a Mapper device (connected as is indicated in Figure 5.6) was programed to remove all the Request signals from other pixels. The delay introduced by the Mapper was known (~270ns). Latency was measured with an oscilloscope working on single mode. Figure 5.21 shows the two recorded signals with the oscilloscope (step stimulus and Request signal) and the delay (latency) between them. The experiment was repeated 30 times to determine the pixel deviation (jitter).

5.4.6 Dynamic Range

We define the dynamic range of the sensor as the ratio of the lowest to the highest level of illumination at which the imager can detect contrast stimuli. The logarithmic sensor compresses the dynamic range and allows to work under very diverse illumination conditions. We tested the imager successfully with very illumination ratios under 1*lux*. It was also able to work under very bright lighting conditions up to 50*Klux*. The measured dynamic range was higher than 100dB. Photo diode dark current limits the lower end of the range. The precision of the photometer used to measure the level of illumination was 1 lux. To know the exact value of the chip illumination, we measure overall level of illumination of the room where the sensor was placed (~ 400lux). To

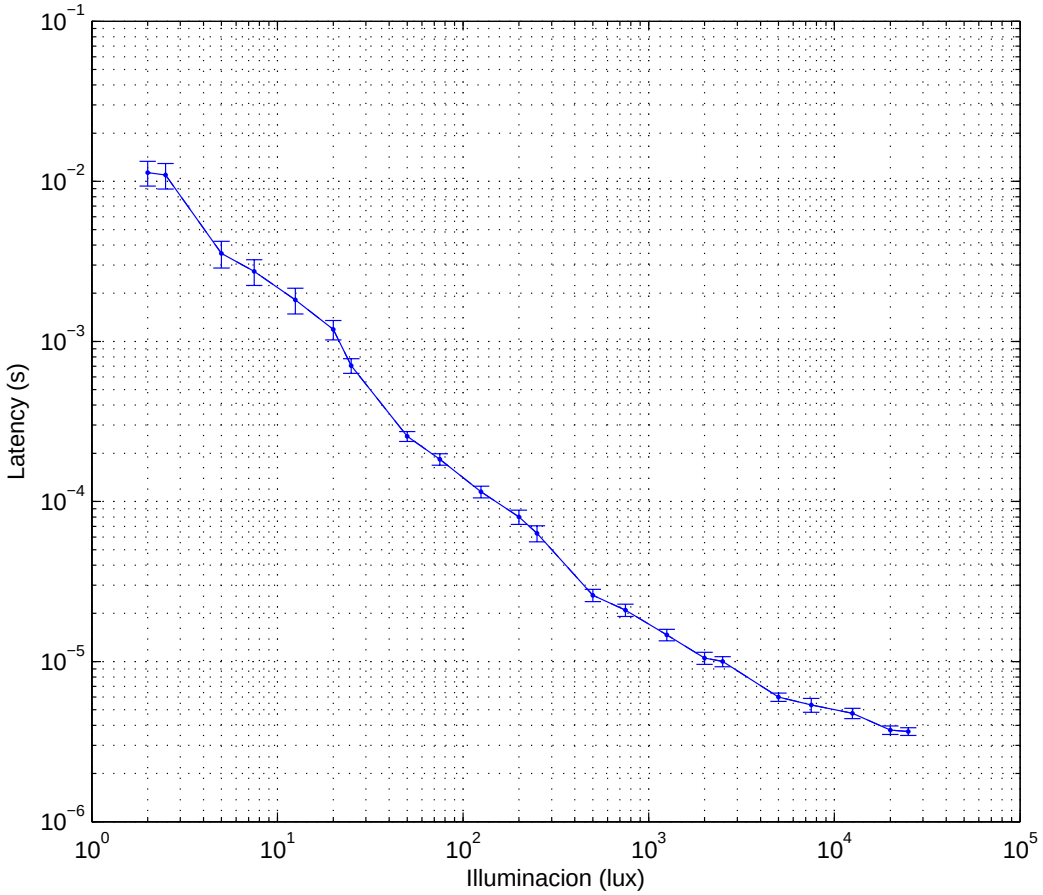


Fig 5.20: Latency and latency deviation (error bars) versus illumination in response to a step increase of single pixel illumination. The experiment was repeated 30 times to determine latency deviation. The circuit was fed with nominal bias settings described in section 5.4.

illuminate the chip with lower and controlled levels of illumination, we used neutral density filters.

Figure 5.22 shows two different snapshots. The first one is an image taken under very low illumination conditions (<1 *lux*). The second one is a snapshot of an image with two regions with very different levels of illumination (intra-scene variations). The illumination ratio of the bright and the dark region was 500. The proposed amplifiers AGC mechanism works well under very different illumination conditions. We achieved good results in very different natural environments. Its main goal is just to keep the DC gain constant, independently of the global luminance level.

5.4.7 Example Data and High Speed Applications

Figure 5.23 displays some scenes taken with the temporal contrast sensor. All of them and the data of Figure 5.24 and Figure 5.26 were captured with indoor light and nominal bias settings with $I_{bn_1} = 200nA$.

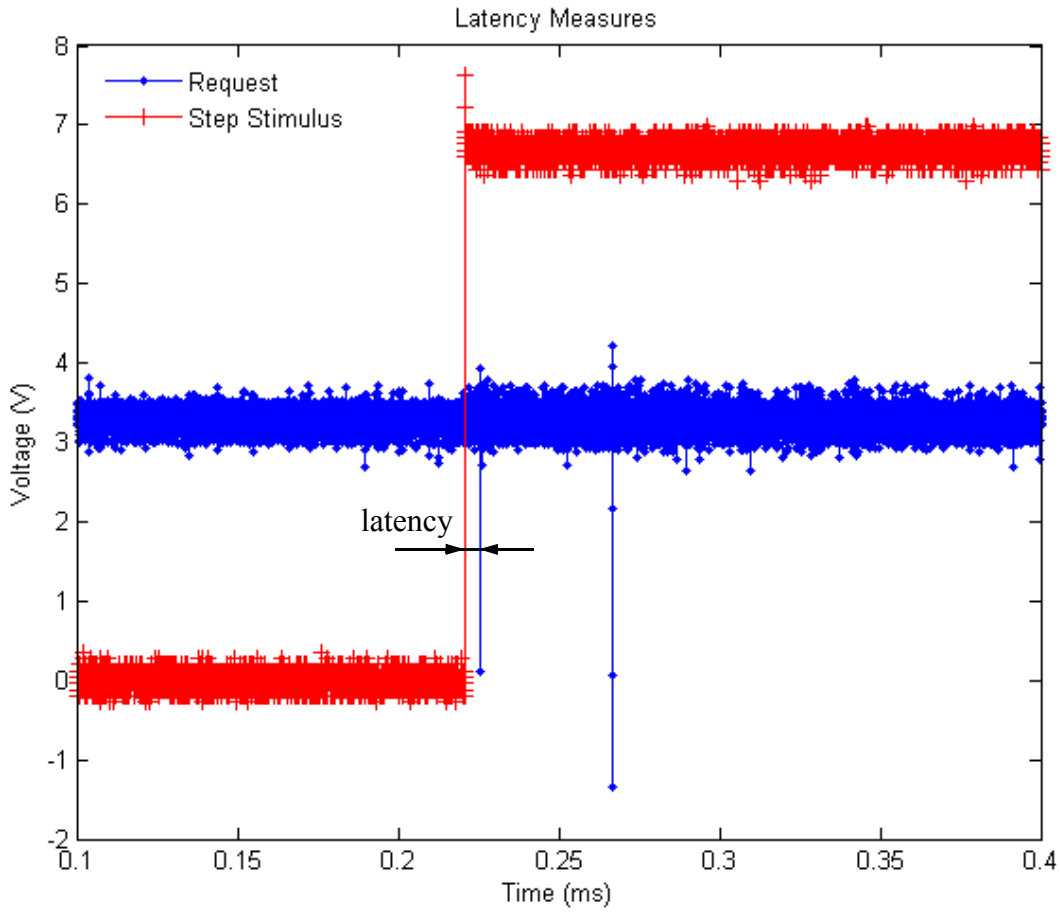


Fig 5.21: Recorded signals used to measure the pixel latency. The delay between the step stimulus and the first pixel request is the latency. Delay was measured with an oscilloscope.

The sensor is specially suitable for high speed applications. According with the experimental results of Figure 5.16, the imager can detect temporal changes in illumination with a frequency higher than 10*KHz*. It is difficult to generate such fast stimuli with mechanical devices. Figure 5.24 shows the events generated during a period of time by a rotating white dot at 4*00Hz* (24,000*rpm*). It was painted on the black blade of an electrical fan. The retina easily detected it. Unfortunately, it was not possible for us to work at faster speed with mechanical devices. In order to generate faster and controlled stimuli, we used an analog oscilloscope working on *XY* mode. Its inputs were two signals with a controlled delay between them. Thus, a bright point rotating at controlled speed could be generated and seen in the phosphor oscilloscope screen.

Firstly, we stimulated the imager with bright points moving through an elliptical trajectory. The oscilloscope inputs were two sinosoidal signals with a delay between them. However, we found out that the phosphor time response could be a limitation to generate high speed stimuli. Emitted light by the phosphor screen did not go out immediately. There was a time response and it could occur that one point would be always illuminated if the stimulus was very fast. Using spiral shape paths, the effect of phosphor persistence was very attenuated because the trajectory of the stimulus was longer.

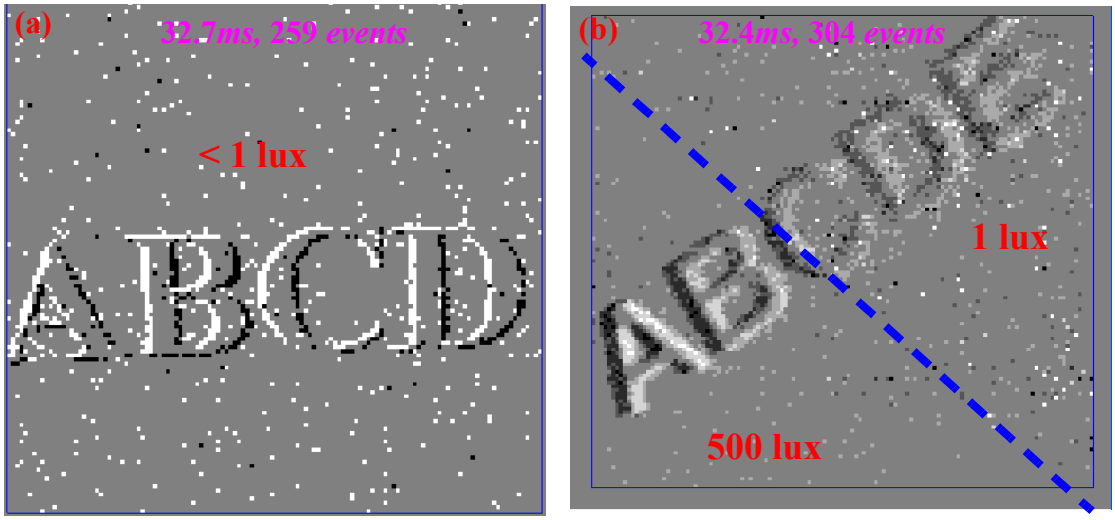
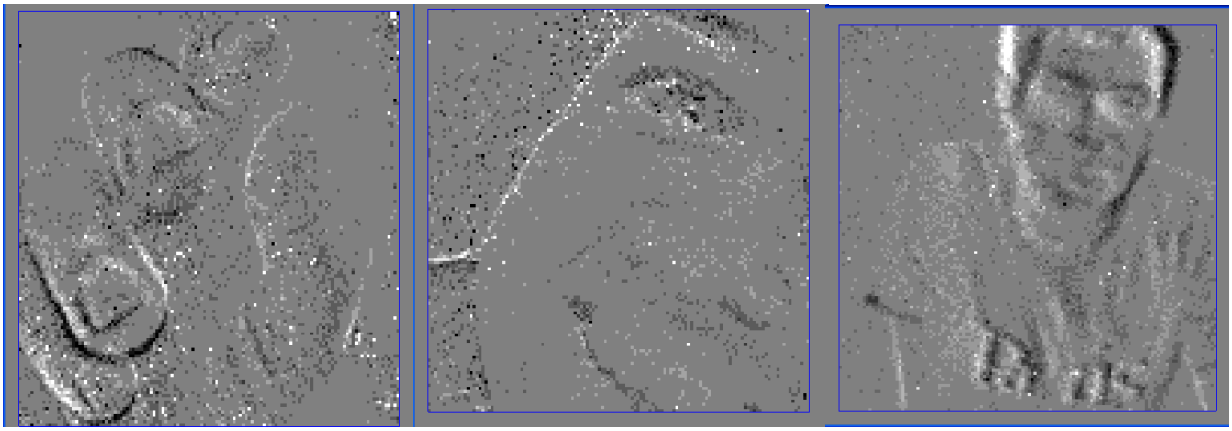


Fig 5.22: (a) Captured snapshot under very low illumination conditions (< 1 lux). (b) Captured snapshot in a scene with two very different levels of illumination. Dotted line is the boundary between the two regions.



7.8*Keps* in 32.8*ms* 20*Keps* in 30*ms* 1.8*Keps* in 32.8*ms*

Fig 5.23: Some snapshots of natural scenes. From left to right: hand, woman face and man.

Figure 5.25 shows the input signals to each oscilloscope channel. Two sine wave signals were modulated with an exponential signal. The resulting *XY* signal had a spiral shape. By this way, the number of times that one individual screen point was turned ON was reduced approximately 5 times. Stimulus speed could be controlled varying the sine wave frequencies. The two signals of Figure 5.25 were repeated periodically with a signal generator.

Figure 5.26 displays some recorded data with the *Datalogger* when we vary the speed of the rotating dot. Sine wave frequencies are indicated on each graph. The average number of events is also highlighted. The imager can detect stimuli frequencies up to 7*KHz*. For stimuli frequencies higher than 1*KHz,* negative events (red points) are not easily detected. The reason is that negative temporal contrast is lower than positive contrast. This can be understood with the help of Figure 5.8. In one region of the retina, we counted the number of positive and negative events generated

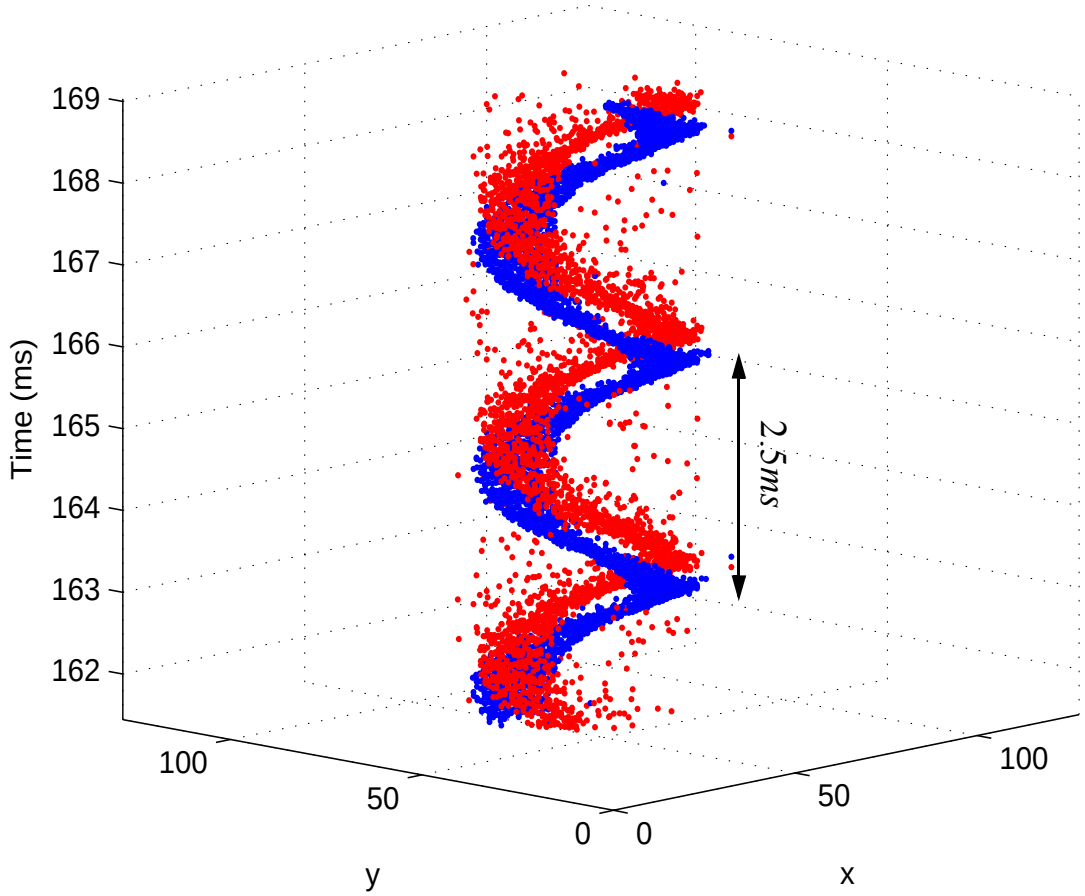


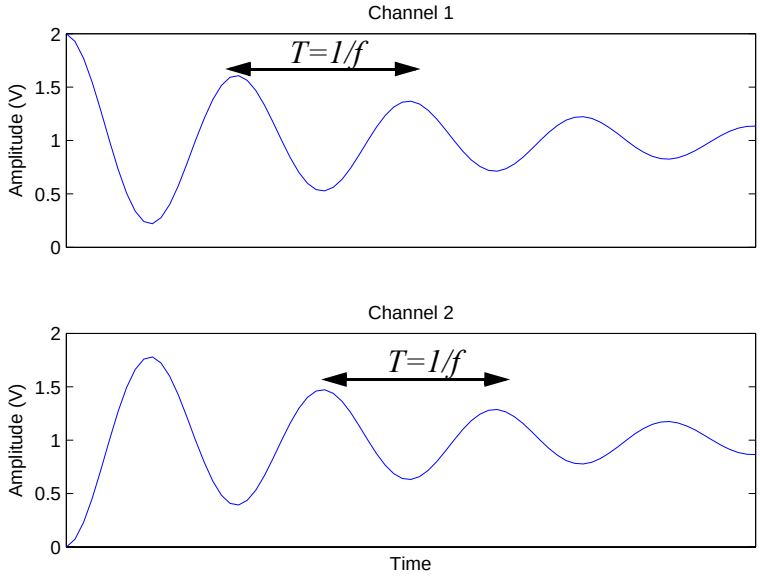
Fig 5.24: Events generated by a rotating dot at 4*00Hz*. Blue points are positive events. Red points are negative ones. The dot stimulus was white and its background black.

by the pixels of such region during time steps of 1*ms*. We can notice that positive events (transition from dark to bright illumination) have a different response time than negative ones (transition from bright to dark illumination). In fact, pixels are generating negative events all the time if the stimulus frequency is higher than 500*Hz*. This suggests, that phosphor light is never completely OFF if the stimulus is very fast. In Figure 5.27, frequency was 500*Hz*. Obviously, when the frequency is increased, it is more difficult to detect negative events because the decay time response is relatively low. Positive temporal contrast is higher and positive events are generated firstly.

We can see that our sensor was able to detect rotating dots at frequencies of 10*KHz* (600,000*rpm*). That is something that we could expect from the results of Figure 5.16. All the pixels do not have the same bandwidth. For this reason, some pixels stop spiking for lower frequencies than another ones.

If positive and negative thresholds are very low, output event rate can be very high and the arbiter system could not send out all the events. It is desirable to work with events rates under 1 *Meps* to avoid information loss. If we look at the events rates of Figure 5.26, we can notice that when stimulus frequency is lower than pixels bandwidth, the output rate is proportional to stimu-

(a)



(b)

Fig 5.25: (a) Input signals to each oscilloscope channel. Both of them were repeated periodically. The resulting *XY* signal, (b), was a bright point moving through a spiral trajectory. We varied frequency values to find out how fast was able the sensor to detect temporal contrast.

lus frequency. For stimuli frequencies higher than the cut off frequency, event rate starts to decay. For frequencies higher than 15*KHz*, only the background ON events due to reset transistor leakage were measured.

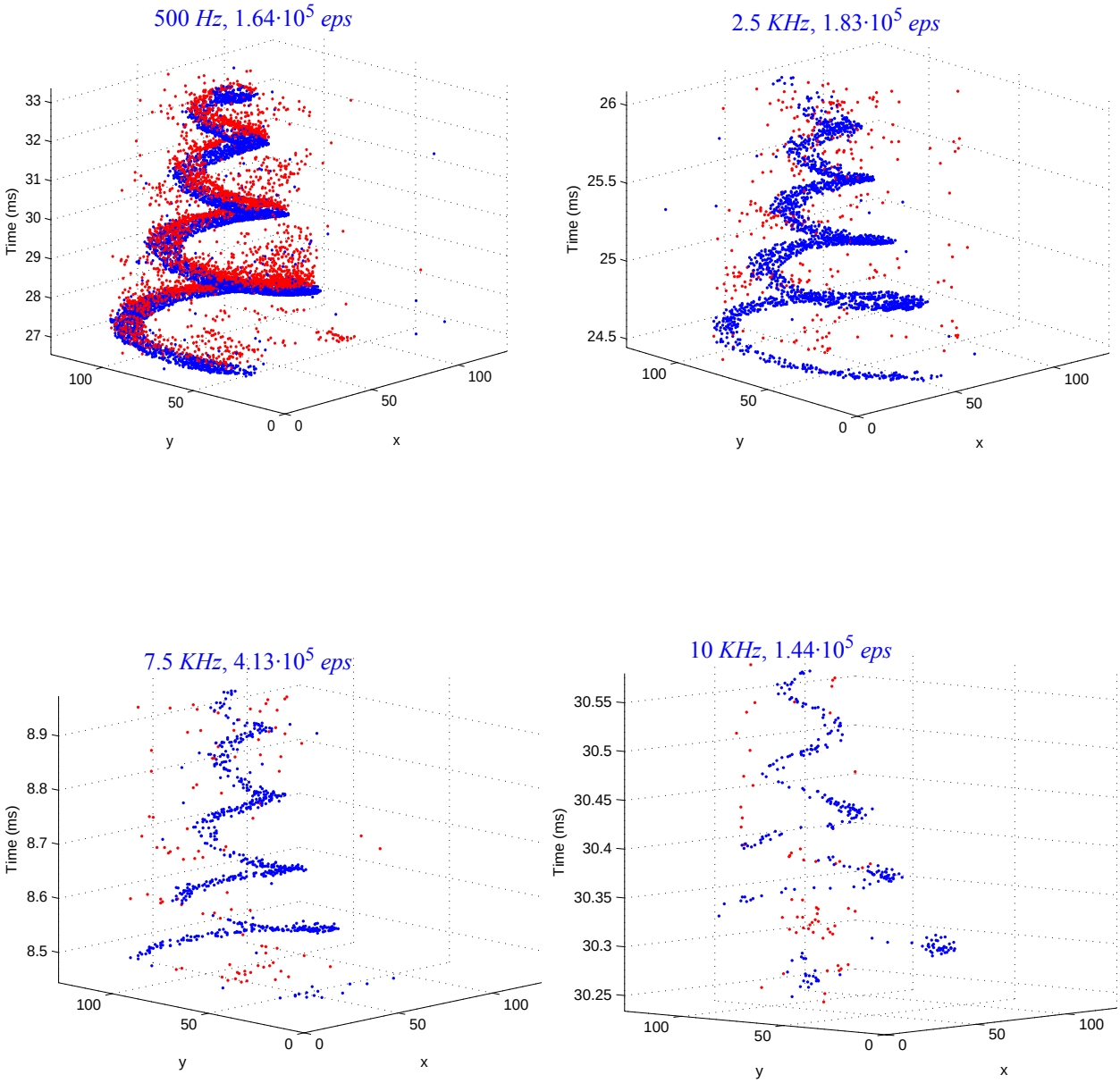


Fig 5.26: Recording events during a *XY* signal period. The stimulus was a dot moving through a trajectory with spiral shape as is shown in Figure 5.25. Results are shown for different frequencies of the input signal. Event rate is also highlighted.

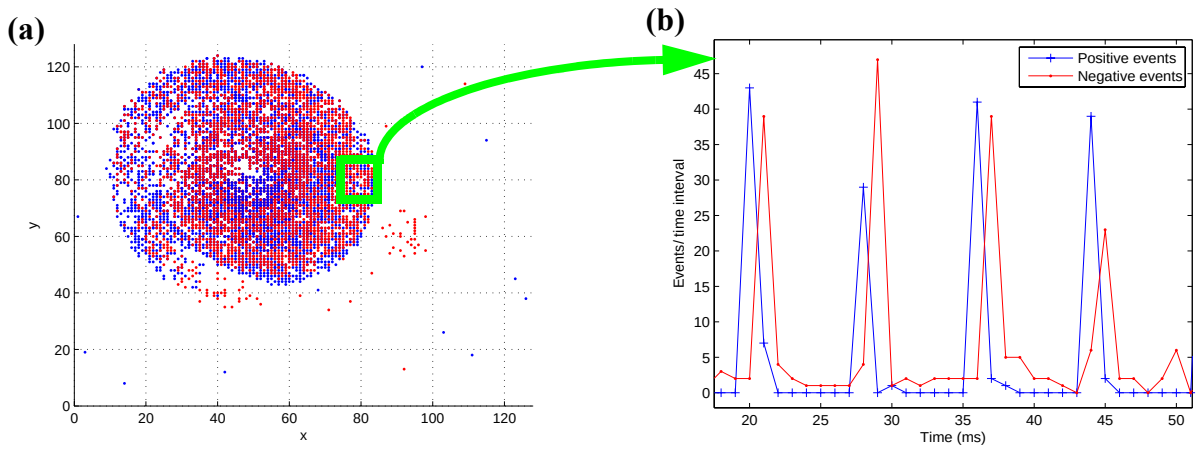


Fig 5.27: (b) Number of events generated by the pixels of the region the highlighted region in (a). Time step was 1*ms*. Phosphor time response is slower for negative events (bright to dark transitions) and faster positive events (dark to bright transitions). Stimulus frequency was 500*Hz*

5.4.8 Power Consumption

Chip power consumption has been characterized. Supply voltage was 3.3*V*. Power consumption depends on static bias conditions and output rate. For this imager, it is dominated by the static bias conditions. The main drawback of using amplifiers is an increase of the standby consumption. In further designs, it could be reduced by using amplifiers operating in subthreshold region. Measured standby chip consumption was approximately 108*mW*. This dissipated power includes the pixel reception stage, amplifiers, AGC block and *I-Pots* consumption. There is also a dynamic power consumption that depends on the output rate. This consumption includes the arbiter and the communication interface.

Typical bias settings were $V_{ref} = 1.5\text{V}$, $V_g = 600mV$, $V_{cas} = 1.8\text{V}$, $V_{bpn_1} = 2V$, $I_{bn} = 100nA$, $I_{bldc} = 100nA$, $I_{brf} = 10nA$ and $I_{bn_1} = 200nA$. Figure 5.28 shows the chip current consumption versus the output event rate. It can vary between 132-185*mW*. As can be seen, for a standard output event rate of 1*Keps,* total power consumption is 142*mW*.

5.5 Discussion

In this chapter, a new AER temporal contrast retina has been presented. Part of the work of this thesis was to test it. The new sensor is specially useful for high speed applications and shows some advantages over prior designs [9]. It uses amplification after the photoreceptor stage. By this way, we achieve more bandwidth, lower latency and better fill factor. The only drawback of amplification is more power consumption. In the future, consumption could be reduced using amplifiers in subthreshold region. Table 5.2 summarizes the imager main specifications. We could think that FPN would be worse in this design due to amplifiers gain mismatch. However, FPN experimental values are similar to the values of the sensor presented by Lichsteiner et al. in [9].

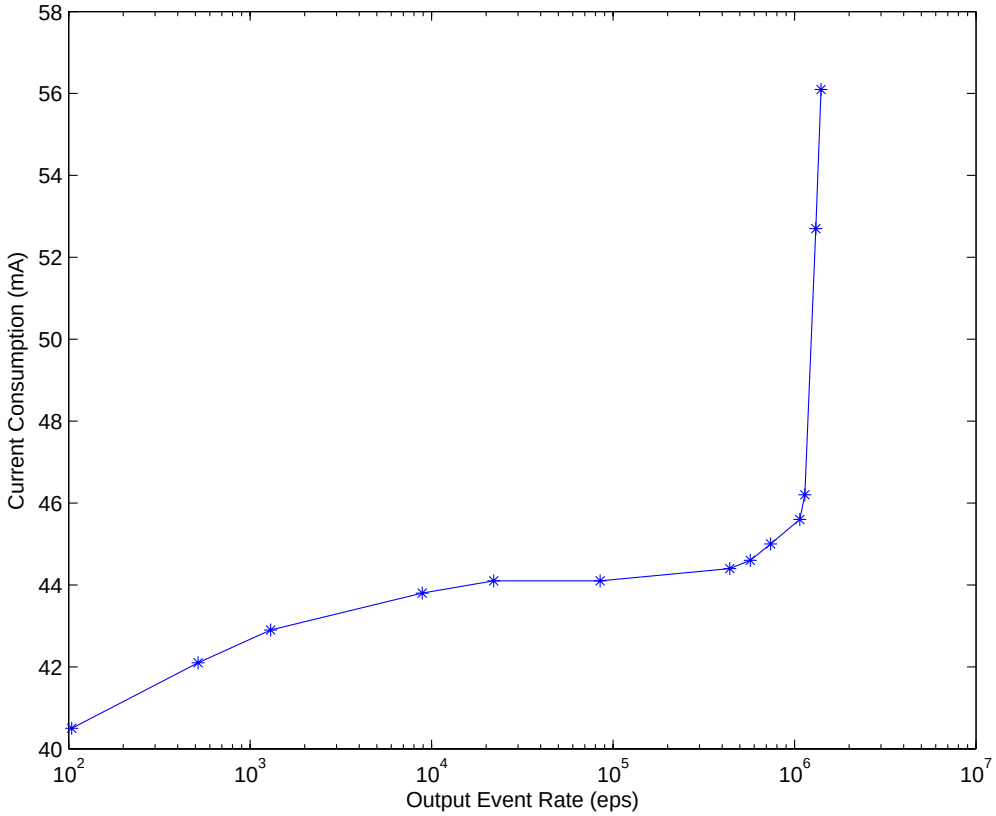


Fig 5.28: Chip total current consumption as a function of output event rate. Standby chip consumption was 108*mW*.

Table 5.2: Sensor Specifications.

Functionality	Light to time restriction	Latency/ Bandwidth	Dynamic Range	FPN	Fill Factor	Pixel Size (μm^2)	Fabrication Process	Power
Temporal Contrast to Number of Events	NO	$3.6\mu s$ / $8KHz$	>100dB	2.8%	8%	35 x 35	$0.35\mu m$ 4M 2P	132-185mW

The ratio between pixel deviation and gain is similar (see comparative between different sensors in Table 1.1).

Finally, some examples of tracking of very fast stimulus (Figure 5.24 and Figure 5.26) have been shown.

CHAPTER 6

Conclusions

6.1 Previous Work

The aim of this thesis was to propose new AER imagers to detect either spatial and temporal contrast. Our first goal was to design a functional spatial contrast retina. Prior attempts to design such sensors were plagued with mismatch and had several drawbacks like bandwidth consumption, when there was no spatial contrast. We started our work, studying the sensor presented by Boahen [21]-[22]. After it, we decided to implement an improved version. To overcome the mismatch problem, we decided to add a new calibration system. The idea of calibration in AER systems was successfully implemented in prior designs [14]-[27]. For our particular necessities, we proposed a new calibration method with improved features.

The second main target was to test a new AER temporal contrast sensor designed in our group. Its main features were high speed and very low latency response. This sensor exploits some ideas previously presented by Lichsteiner et al. in [10] and it is based on their sensor.

6.2 Achievements

In this thesis we have proposed new concepts that could be implemented in further AER designs. The first one is calibration. We decided to design a new calibration system specially suitable for large neuromorphic arrays. Its main advantage over prior calibration systems was the fact that calibration degrades gracefully when we scale bias currents or illumination level changes. The calibration system is specially useful because it only needs to be calibrated once. Calibration accuracy does not depend on illuminance strongly. Only digital registers (one for each bit of resolution) and 3 transistors have to be added to each pixel to implement the calibration procedure.

The second main achievement is the AER spatial contrast retina. It uses the calibration method mentioned above. The sensor has also a thresholding mechanism that allows to remove the residual mismatch after calibration. This mechanism is also useful to remove noise or inhibit the activity of positive or negative events. There are two independent channels to process positive or negative events. For some applications, it could be desirable to work with unsigned events. In these cases, the thresholding mechanism is an efficient method to inhibit completely the activity

of positive or negative events. Moreover, the imager has an optional TFS (time-to-first spike) mode. It is implemented with a global reset mechanism. By adjusting the voltage of the refractory period, the number of spikes that one individual pixel generates can be controlled. If we set the refractory period to allow each pixel just spike once after the global reset signal becomes active, pixels with more activity (more relevant information) will spike first. Thus, if the reset period is chosen suitably, only relevant information will be transmitted. We can achieve a frame-based processing with some advantages of AER systems. Another important features of the sensor are high sensitivity, low fixed pattern noise and low latency.

The second proposed sensor was a temporal contrast retina. Its main features were high speed, low latency and high dynamic range. This sensor has a similar topology to the one proposed by Lichsteiner et al. [10]. Its main advantages are the use of amplifiers after the photoreception stage. By this way, temporal contrast sensitivity and speed response can be improved. At the same time, fill factor was reduced because the gain at the capacitance divider of the differencing circuit could be decreased.

Extensively experimental results showing the validity of the new calibration systems and the main features of the two contrast retinae are provided. In chapter 2 a new circuit to generate biasing currents was presented. In chapter 3, the calibration circuitry is explained, in chapter 4 the spatial contrast retina is presented, and in chapter 5, the temporal contrast retina is showed.

6.3 Future Work

After studying the new circuits thoroughly, some improvements could be proposed for each one. Let us start with the calibration system.

As we explained in chapter 3, the calibration procedure is presently off-line, but it is perfectly possible to implement it on-chip with an extra controller (designed and synthesized via VHDL, for example), since it only requires to expose the chip to uniform illumination (one can simply remove the optics), compare the pixel frequencies (we would only need to count the number of times that each pixel spikes during a time interval) and compute an optimum set of calibration weights (find the calibration words that minimize the number of spikes of each pixel).

If we speak about the spatial contrast retina, the feasibility of the design has been shown. However, the retina resolution is rather low (32 x 32). It would be easy to design a new prototype with more resolution to take images with more quality. One of its drawbacks is the low fill factor. This could be easily improved in the future. The FPN of the circuit is very low (0.6%), so we could reduce the number of bits (3 or 4 bits registers would be enough to calibrate the chip) of the calibration system keeping a reasonable FPN value. Combining calibration and thresholding, better results can be achieved. Maybe, a more precise thresholding mechanism could be designed. It would only be necessary to increase the transistor size of the two transistors of the current mirror used to implement the thresholding method. Moreover, an extra controller to monitor the time-to-first spike mode could be added. It could control the duration of the frames, T_{frame}, or the number of spikes received before resetting the sensor.

With regard to the AER temporal contrast retina, we improved the sensor speed and response time with an amplification stage. We also increased fill factor at the same time. However, amplifiers operate in strong inversion and their power consumption is higher in comparison to the rest of blocks that make up the pixel. In further designs, amplifiers could be designed to work in weak inversion and reduce its power consumption. Pixel size is relatively low ($35x35\mu m^2$). It could be reasonable to implement a temporal contrast retina with more resolu-

tion. An optional time-to-first spike mode could be added to this sensor. Only 3 extra transistors per pixel and one output pin would be necessary to have available the optional operation mode.

APPENDIX A

Abbreviations and Mathematical Conventions

Throughout this thesis, we used the following abbreviations,

CCD	*Charge-Coupled Device*
VLSI	*Very Large Scale Integration*
Matlab	*Mathematical Program of Mathworks, Inc*
APS	*Active Pixel Sensor*
FPN	*Fixed Pattern Noise*
I_{ph}	*Photo current*
V	*Voltage*
AER	*Address Event Representation*
VHDL	*Very High Speed Integrated Circuit Hardware Description Language*

The following mathematical symbols are also used,

f, w	Temporal frequency
T	Temporal period
$\lvert\ \rvert$	Absolute value

$\boldsymbol{w}_{cal}$	Optimum calibration word
θ	Temporal contrast
$\Delta V_{diff\_\mathrm{ev}}$	Voltage threshold to generate one event (either positive or negative)
ΔV_{diff}	Voltage variation in response to temporal contrast stimulus
θ_{ev}	Contrast threshold to generate one event (either positive o negative)
$\sigma_{\theta_{ev}}$	Deviation of the contrast threshold
θ_{min}	Minimum detectable contrast.
G_i	Gain of the amplification stage, i
τ	System response time

Bibliography

[1] M. Sivilotti, "Wiring considerations in analog VLSI systems with application to field-programmable networks," *Ph.D. dissertation, Calif. Inst.Technol.*, Pasadena, 1991.

[2] M. Mahowald, "VLSI analogs of neural visual processing: a synthesis of form and function" Ph.D. dissertation, California Institute of Technology, Pasadena, 1991.

[3] T. Delbrück, Library essentials, Analog VLSI and Neural Systems by Carver Mead, Addison Wesley, 1986. *The Neuromorphic Engineer,* 1(1):11, 2004. http://ine-web.org/research/newsletters/index.html.

[4] Kwabena A. Boahen, "Retinomorphic Vision Systems: Reverse Engineering the Vertebrate Retina", Ph. D. dissertation, California Institute of Technology, Pasadena, 1996.

[5] G. M. Shepherd, *The Synaptic Organization of the Brain,* 3rd ed. Oxford, U.K.: Oxford University Press, 1990.

[6] M. Mahowald. *An Analog VLSI System for Stereoscopic Vision,* Kluwer, Boston, MA, 1994.

[7] U. Mallik, et al., *"Temporal Change Threshold Detection Imager",* in *ISCCC Dig. of Tech. Papers, San Francisco, 2005, pp362-363.*

[8] Y. Chi et al., *"CMOS Camera with In-pixel Temporal Change Detection and ADC",* IEEE Journal of Solid State Circuits, vol. 42, pp, 2187-2196, OCT 2007.

[9] Patrick Lichsteiner et al, *"A 128x128 120dB 15μs Latency Asynchronous Temporal Contrast Vision Sensor"*, IEEE Journal of Solid State Circuits, vol. 43, pp. 566-576, 2008.

[10] P. Lichsteiner, et al., "*A 128x128 120dB 30mW Asynchronous Vision Sensor that Responds to Relative Intensity Change*", in ISSCC Dig. of Tech. Papers, San Francisco, 2006, pp. 508-509 (27.9)

[11] S. Thorpe, D. Fize, C. Marlot, *"Speed of Processing in the Human Visual System",* Nature 381:520-2, 1996.

[12] X. G. Qi, X.; Harris J., *"A Time-to-first-spike CMOS imager",* in Proc. of the 2004 IEEE International Symposium on Circuits and Systems (ISCAS04), Vancouver, Canada, 2004, pp. 824-827.

[13] J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, *"A Calibration Technique for Very Low Current and Compact Tunable Neuromorphic Cells. Application to*

5-bit 20nA DACs", IEEE Transactions on Circuits and Systems, Part-II: Brief Papers, vol.55, No. 6, pp. 522-526, June 2008.

[14] J. Costas-Santos, T. Serrano-Gotarredona, R. Serrano-Gotarredona and B. Linares-Barranco, *"A Spatial Contrast Retina with On-chip Calibration for Neuromorphic Spike-Based AER Vision Systems",* IEEE Trans. Circuits and Systems, Part-I: Regular Papers, vol. 54, No. 7, pp. 1444-1458, July 2007.

[15] M. Barbaro, P.-Y. Burgi, A. Mortara, P. Nussbaum, and F. Heitger. "*A 100x100 Pixel Silicon Retina for Gradient Extraction with Steering Filter Capabilities and Temporal Output Coding".* IEEE Journal of Solid-State Circuits, 37(2):160-172, Feb. 2002.

[16] P. F. Ruedi, et al., *"A 128x128 Pixel 120-dB Dynamic-range Vision-Sensor Chip for Image Contrast and Orientation Extraction",* IEEE Journal of Solid-State Circuits, 38:2325-2333, Dec. 2003.

[17] P. F. Ruedi, et al., *"An SoC Combining a 132dB QVGA Pixel Array and a 32b DSP/MCU Processor for Vision Applications",* in IEEE ISSCC Dig. of Tech. Papers, 2009, pp.46-47, 47a.

[18] Honghao Ji and Pamela A. Abshire, "*Fundamentals of Silicon-Based Phototransduction"* in *"CMOS Imagers from Phototransduction to Image Processing"*, Kluwer Academic Publishers.

[19] E. Culurciello, R. Etienne-Cummings, and K. A. Boahen, *"A Biomorphic Digital Image Sensor",* IEEE J. Solid-State Circuits, vol.38, pp.281-294, 2003.

[20] S. Chen, and A. Bermak, *"Arbitrated Time-to-First Spike CMOS Image Sensor with On-Chip Histogram Equalization",* IEEE Transactions VLSI Systems, vol. 15, No. 3, pp 346-357, March 2007.

[21] K. A. Zaghloul and K. Boahen, *"Optic Nerve Signals in a Neuromorphic Chip: Part 1",* IEEE Transactions on Biomedical Engineering, vol 51, pp. 657-666, 2004.

[22] K. A. Zaghloul and K. Boahen, *"Optic Nerve Signals in a Neuromorphic Chip: Part 2",* IEEE Transactions on Biomedical Engineering, vol 51, pp. 667-675, 2004.

[23] J. A. Leñero-Bardallo, T. Serrano-Gotarredona, and B. Linares-Barranco, *"A 5-Decade Dynamic Range Ambient-Light-Independent Calibrated Signed-Spatial-Contrast AER Retina with 0.1ms Latency and Optional Time-to-First-Spike Mode",* Transactions on Circuits and Systems, Part I, Under press.

[24] C. Posch, et al., *"High DR, Low Date-rate Imaging Based on an Asynchronous, Self-triggered Address-event PWM Array with Pixel-level Temporal Redundancy Suppression"*, in 2010, in ISCAS 2010.

[25] R. R. Harrison, J.A. Bragg, P. Hasler, B.A. Minch, and S.P. DeWeerth, *"A CMOS programmable analog memory-cell array using floating-gate circuits," IEEE Trans. on*

Circuits and Systems, Part II, vol. 48, No. 1, pp. 4-11, Jan. 2001.

[26] Y. L. Wong, M. H. Cohen, and P. A. Abshire, *"128x128 floating gate imager with self-adapting fixed pattern noise reduction," Proc. of the IEEE 2005 Int. Symp. on Circuits and Systems* (ISCAS'05), vol. 5, pp. 5314-5317, 2005.

[27] R. Serrano-Gotarredona, T. Serrano-Gotarredona, A. Acosta-Jimenez, and B. Linares-Barranco, *"A Neuromorphic Cortical-Layer Microchip for Spike-Based Event Processing Vision Systems," IEEE Trans. on Circuits and Systems I: Regular Papers*, vol. 53, No. 12, pp. 2548-2566, Dec. 2006.

[28] B. Linares-Barranco, T. Serrano-Gotarredona, and R. Serrano-Gotarredona, *"Compact Low-Power Calibration Mini-DACs for Neural Massive Arrays with Programmable Weights," IEEE Trans. on Neural Networks*, vol. 14, No. 5, pp. 1207-1216, September 2003.

[29] C. A. Laber, C. F. Rahim, S. F. Dreyer, G. T. Uehara, P. T. Kwok, and P. R. Gray, *"Design Considerations for a high-Performance 3mm CMOS Analog Standard-Cell Library," IEEE J. Solid-State Circuits*, vol. SC-22, No. 2, pp. 181-189, April 1987.

[30] T. Delbrück and A. Van Shaik, *"Bias Current Generators with Wide Dynamic Range" Int. Journal of Analog Integrated Circuits and Signal Processing*, No. 43, pp. 247-268, June 2005.

[31] T. Delbrück and A. Van Shaik, *"Bias Current Generators with Wide Dynamic Range"* , 2004 International Symposium on Circuits and Systems (ISCAS 2004), Vancouver, Canada, May 23-25 2004 pp. I-337-340.

[32] T. Delbruck, P. Lichtsteiner, R. Berner, C. Dualibe, *"32-bit Configurable bias current generator with sub-off-current capability", ISCAS 2010, (in press)*

[33] R. R. Harrison, J.A. Bragg, P. Hasler, B.A. Minch, and S.P. DeWeerth, *"A CMOS programmable analog memory-cell array using floating-gate circuits," IEEE Trans. on Circuits and Systems, Part II*, vol. 48, No. 1, pp. 4-11, Jan. 2001.

[34] K. Bult and G.J.G.M. Geelen, *"An inherently linear and compact MOST-only current division technique," IEEE J. Solid-State Circuits,* vol. 27, No. 12, pp. 1730-1735, Dec. 1992.

[35] T. Serrano-Gotarredona and B. Linares-Barranco, *"CMOS Mismatch Model valid from Weak to Strong Inversion"*, *Proc. of the 2003 European Solid State Circuits Conference*, (ESSCIRC'03), pp. 627-630, September 2003.

[36] P.R. Gray, P.J. Hurst, S.H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits,* 4th Edition, John Wiley, 2001.

[37] Rafael Serrano-Gotarredona, Luis Camuñas-Mesa, Teresa Serrano-Gotarredona, Juan A. Leñero-Bardallo and Bernabé Linares-Barranco, *"The Stochastic I-Pot: A Circuit Block for Programming Bias Currents", IEEE Transaction on Circuits and Systems-II: Brief Papers,*

vol 19, No. 7, pp. 1196-1219, July 2008.

[38] B. Linares-Barranco and T. Serrano-Gotarredona, *"On the Design and Characterization of Femtoampere Current-Mode Circuits", IEEE Journal of Solid-State Circuits,* vol. 38, No. 8, pp. 1353-1363, August 2003.

[39] M. Azadmehr, J. Abrahamsen, and P. Häfliger, *"A Foveated AER Imager Chip",* in *Proc. of the IEEE International Symposium on Circuits and Systems (ISCAS2005),* pp. 2751-2754, Kobe, Japan, 2005.

[40] R.J. Vogelstein, U. Mallik, E. Culurciello, R. Etienne-Cummings, and G. Cauwenberghs, "Spatial acuity modulation of an address-event imager," *Proc. of the 2004 11th IEEE International Conference on Electronics, Circuits and Systems (*ICECS 2004), pp. 207-210, Dec. 2004.

[41] J. Kramer, *"An Integrated Optical Transient Sensor", IEEE Transactions on Circuits and Systems, Part II: Analog and Digital Signal Processing,* v. 49, No. 9, pp. 612-628, Sep. 2002.

[42] M. Arias-Estrada, D. Poussart, and M. Tremblay, *"Motion Vision Sensor Architecture with Asynchronous Self-Signaling Pixels", Proceedings of the 7th International Workshop on Computer Architecture for Machine Perception (CAMP07),* pp. 75-83, 1997.

[43] K. Boahen, *"Retinomorphic Chips that see Quadruple Images", Proceedings of International Conference of Microelectronics for Neural, Fuzzy and Bio-inspired Systems* (Microneuro99), pp. 12-20, Granada, Spain, 1999.

[44] S. Thorpe, et al., *"SpikeNet: Real-time Visual Processing in the Human Visual System", Neurocomputing 58-60: 857-64, 2004.*

[45] K. Boahen and A. Andreou, *"A Contrast-Sensitive Retina with Reciprocal Synapses",* in J. E. Moddy (Ed.), *Advances in neural information processing,* vol. 4, pp- 764-772, San Mateo CA, 1992. Morgan Kaufman.

[46] A. G. Andreou and K. Boahen, *"Translinear Circuits in Subthreshold CMOS", Analog Integrated Circuits and Signal Processing,* Kluwer, no. 9 , pp. 141-166, Apr. 1996.

[47] K. Boahen, *"Point-to-Point Connectivity Between Neuromorphic Chips Using Address Events", IEEE Transactions on Circuits and Systems Part-II,* vol. 47, No. 5, pp. 416-434, May 2000.

[48] R. Berner, T. Delbruck, A. Civit-Balcells and A. Linares-Barranco,"A 5 Meps $100 USB2.0 Address-Event Monitor-Sequencer Interface", *IEEE International Symposium on Circuits and Systems,2007, ISCAS 2007,* pp. 2451-2454.

[49] jAER Open Source Project 2007 [Online]. Available: http://jaer.wiki.sourceforge.net.

[50] V. Gruev and Etienne-Cummings, R. A. pipelined temporal difference imager. *IEEE J. Solid State Circuits,* 39:538-543, 2004.

[51] M. A. Mahowald, *"VLSI Analogs of Neuronal Visual Processing: A Synthesis of Form and Function",* PhD, Computation and Neural Systems, Caltech, Pasadena, California, 1992.

[52] D. A. Baylor, T. D. Lamb, and K.W. Yau, *"Responses of Retinal Rods to Single Photons",* J. Physiology., 288:613-634, 1979.

[53] D. A. Baylor, B. J. Nunn, and J. L. Schnapf. *"The Photo Current, Noise, and Spectral Sensitivity of Rods of the Monkey Macaca Fascicularis".* J. Physiology, 357:575-607, 1984.

[54] Patrick Lichtsteiner, *"An AER Temporal Contrast Vision Sensor",* Ph.D. Dissertation, ETH Zürich, 2006.

[55] Foveon , Inc., Website, www.foveon.com.

[56] H.R. Blackwell. Contrast threshold of the human eye, *"J. Opt. Soc. Am., 36:624-643, 1946.*

[57] P. Sterling, E. Cohen, M. Freed, and R.G. Smith. *"Microcircuitry of the On-beta Ganglion Cell in Daylight, Twilight, and Starlight",* Neuroscience. Res. (Suppl.), 6:269-285,1987.

[58] D.C. Burr. *"Motion Smear",* Nature, 284:164-165, 1990.

[59] Tobi Delbrück, Bernabe Linares-Barranco, Eugenio Culurciello, Christoph Posch, *"Activity-Driven, Event-Based Vision Sensors", IEEE International Symposium on Circuits and Systems, 2010, ISCAS 2010, Paris.*

[60] J. Zihl, D. Von Cramon, and N. Mai. *"Selective Disturbance of Movement Vision After Bilateral Brain Damage".* Brain, 106(2):313-340.

[61] J. Kramer, *"An ON/OFF Transient Imager with Event-driven, Asynchronous Readout",* In Procc. *IEEE International Symposium on Circuits and Systems, May 2002.*

[62] J. Kramer, *"An Integrated Optical Transient Sensor", IEEE Transactions of Circuits and Systems II,* 49(9):612-628, Sep 2002.

[63] T. Delbrück and C.A. Mead. *"Adaptive Photoreceptor with Wide Dynamic Range",* In 1994 *IEEE International Symposium On Circuits and Systems",* volume 4, pages 339-342, London, May 1994. ISCAS'94: London, England, 30 May-2 June.

[64] J. Kramer, R. Sarpeshkar, and C. Koch. *"Pulse-based Analog VLSI Velocity Sensors", IEEE Transactions of Circuits and Systems II,* 44(2):86-101, February 1997.