



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



⑪ Número de publicación: **2 152 176**

⑫ Número de solicitud: 009900346

⑬ Int. Cl.⁷: G06T 7/00

⑭

SOLICITUD DE PATENTE

A1

⑮ Fecha de presentación: **19.02.1999**

⑯ Fecha de publicación de la solicitud: **16.01.2001**

⑰ Fecha de publicación del folleto de la solicitud:
16.01.2001

⑱ Solicitante/s: **UNIVERSIDAD DE MÁLAGA**
Plaza de El Ejido s/n
29071 Málaga, ES

⑲ Inventor/es: **Camacho Lozano, Pelegrín**

⑳ Agente: **No consta**

㉑ Título: **Sistema multiatencional adaptativo de estructura multiresolución reconfigurable.**

㉒ Resumen:

Sistema multiatencional adaptativo de estructura multiresolución reconfigurable, cuyo objeto es reducir el número de datos analizables en secuencias de imágenes, para acortar los tiempos de procesamiento.

Consiste en una estructura de datos tridimensionales, representando imágenes con resoluciones decrecientes. Determinando zonas de interés en imágenes con resolución reducida se seleccionan adaptativamente regiones correspondientes de alta resolución y los gradientes de resolución circundantes, obteniéndose estructuras bidimensionales que representan imágenes multifoveales.

Aplicado en sistemas de visión artificial permite procesar las zonas relevantes de imágenes, analizando sus características en tiempo real y reduciendo tiempos dedicados al resto, donde se mantienen capacidades de detección.

Las alternativas para implementarlo partiendo de sensores uniresolución son:

- 1 - técnicas "doble-buffer" con memorias SRAM
- 2 - técnicas "pipeline" con memorias FIFO

aplicadas en dispositivos tipo FPGA, CPLD o ASICs para hacer reconfiguración adaptativa.

ES 2 152 176 A1

DESCRIPCION

Sistema multiatencional adaptativo de estructura multirresolución reconfigurable.

Sector de la técnica

Sistemas de Visión Artificial. Sensores de resolución variable o multirresolución. Automatización de procesos por tratamiento de imágenes con resolución múltiple.

Estado de la técnica

Por la complejidad de los procesos y la cantidad de los recursos computacionales necesarios para extraer información útil a partir de las imágenes suministradas por cámaras, es conocido el interés en emular los sistemas biológicos de visión y, particularmente, la funcionalidad y disposición de fotorreceptores presentes en la retina del ojo humano, donde se observa una fovea de reducidas dimensiones con elevada densidad de conos fotorreceptores y, alrededor de la fovea, con una densidad decreciente radialmente, los bastones fotorreceptores que integran las zonas periféricas de la retina.

Por las diferencias funcionales de conos y bastones, y por su distribución en la retina, se dice que ésta tiene una resolución variable espacialmente. Se han realizado retinas artificiales con estructura log-polar, de las que son ejemplos las desarrolladas en las Universidades de Génova y Lovaina, con una fovea o zona central de resolución uniforme y fotorreceptores de tamaño creciente radialmente, dispuestos en anillos circulares concéntricos cuyas dimensiones crecen a medida que se acercan a la periferia del sensor. También se han realizado retinas artificiales con foveas cuadradas, rodeadas de anillos concéntricos, a modo de marcos, con anchuras verticales y horizontales idénticas. Un ejemplo es el desarrollado entre la Universidad de Nueva York en Buffalo y Anierst Systems. Tanto unos sensores como otros se caracterizan por tener una retinotopología con anillos concéntricos alrededor de una única fovea y, por la complejidad de los procesos que implica su fabricación, ninguno de ellos se comercializa actualmente. Por tener estructura fija no son reconfigurables y, como consecuencia, presentan inconvenientes que complican su uso:

1. Si existen varios objetos de interés en la escena, distanciados entre sí de forma que sus imágenes excedan las dimensiones de la zona de máxima resolución o fovea del sensor, solo uno podrá ser captado en dicha fovea.
2. Si el tamaño de la imagen de un objeto proyectado en la fovea excede el tamaño de ésta, se tienen distorsiones por los cambios de resolución del sensor. Al tener los sensores estructura fija, si la fovea fija se hace mayor para evitar lo anterior, los objetos pequeños se observarán con fovea innecesariamente grande, procesando datos irrelevantes de fondo de escena, lo que aumenta el tiempo y recursos computacionales necesarios.
3. Para captar íntegramente las regiones u objetos de interés en la escena sin los inconvenientes citados en los puntos anteriores,

será preciso mover la cámara en azimut y/o elevación y hacer zoom para encajar la imagen de los objetos de manera óptima dentro de la fovea del sensor, debiendo repetirse el proceso con cada objeto a captar.

4. Los sistemas mecánicos, normalmente precisos, necesarios para controlar la orientación y óptica de la cámara implican compromisos entre coste, velocidad, peso, volumen y otras características del sistema que dificultan su utilización en determinados entornos.

Los sensores que se encuentran en todos los tipos de videocámaras y sensores de imagen comercializados tienen sus elementos fotorreceptores o píxeles en una matriz rectangular con una resolución uniforme, es decir, el tamaño y la densidad de los fotorreceptores son constantes en toda la superficie del sensor. Esto hace que la resolución de las imágenes sea uniforme y determinada por la óptica y las dimensiones físicas de los píxeles del sensor donde se proyecta la imagen de la escena captada. Con estos sensores de resolución uniforme se han desarrollado algoritmos para emular retinotopologías con resolución múltiple, haciendo agrupamientos de píxeles o explorando los fotorreceptores del sensor de forma discontinua, por muestreos. Un ejemplo es la estructura log-polar concéntrica desarrollada por la Universidad de Rochester. Sin embargo, los algoritmos que dan lugar a retinotopologías circulares para imitar la del ojo humano, tienen a cambio una complejidad matemática que dificulta y alarga el tiempo de procesamiento de las imágenes obtenidas con ellos, al no poder aplicar los algoritmos de tratamiento de imágenes utilizados con estructuras cartesianas.

Descripción

El sistema multiatencional adaptativo de estructura multirresolución reconfigurable se obtiene a partir de sensores comerciales de resolución uniforme, aplicando algoritmos que permiten atender con resolución máxima y simultáneamente varias zonas de interés en las imágenes, adaptando el tamaño y posición de zonas a las de objetos o regiones que interesa analizar en escenas. El resto de la imagen se analiza con resoluciones inferiores, obtenidas por reducción o compresión de los píxeles contenidos en la imagen original suministrada por la cámara. El uso de estos sistemas en visión artificial es particularmente eficiente en escenas donde se tienen elementos móviles que deben ser localizados e identificados para actuar en función del reconocimiento realizado, mientras el fondo de escena carece de interés. En esos casos, la reducción selectiva de datos que ofrece el sistema acelera el procesamiento de imágenes, optimizando el tiempo y los recursos dedicados a las tareas de análisis y reconocimiento.

El procedimiento consiste en obtener una serie de imágenes con reducciones progresivas de resolución, conocidas como pirámides de visión, hasta llegar a un nivel mínimo relacionado al contexto de escena y objetos de interés, donde aún se mantienen los detalles que servirán para determinar tamaños y posiciones de los objetos detectados en el campo de visión. Para una mejor comprensión

de lo que sigue, se adjunta el dibujo de la Figura 1. Dado que estas determinaciones se realizan en niveles de baja resolución, con imágenes comprimidas y un número de datos significativamente reducido, el tiempo de procesamiento es mucho menor que si se hiciera a niveles de alta resolución. Con dispositivos electrónicos digitales se obtienen los niveles de la pirámide en tiempo real, entendiéndose como tal las tasas de 25 o 30 imágenes/segundo de las cámaras que dan las imágenes que constituyen la base de la pirámide de visión. Los niveles de la pirámide equivalen a imágenes del mismo campo de visión tomadas con cámaras de resoluciones decrecientes. Las pirámides con gradiente de resolución cuadrático se obtienen a partir de las imágenes base con resolución máxima, denominadas de nivel 0, promediando el brillo o nivel de gris de conjuntos de 4 píxeles adyacentes 2×2 de la imagen base. El promedio es el nivel de gris asignado al superpíxel o máxel correspondiente de nivel 1. Así, por ejemplo, una imagen base de 16×12 píxeles puede considerarse como un conjunto de 8×6 grupos de 2×2 píxeles adyacentes. El promedio de cada grupo 2×2 es el valor del máxel correspondiente de la imagen de 8×6 elementos que corresponde al nivel 1 de la pirámide, que tendrá una resolución lineal $1/2$ y una resolución espacial $1/4$ de la imagen de 16×12 .

Así, por ejemplo, partiendo de una escena captada en una imagen con resolución uniforme de 640×480 píxeles, se obtendría una pirámide cuyo nivel 1 tendría 320×240 máxeles con resolución $1/4$ de los píxeles del nivel 0, el nivel 2 tendría 160×120 máxeles con resolución $1/4$ de los del nivel 1, el nivel 3 tendría 80×60 máxeles de resolución $1/4$ de los de nivel 2, etc. El mínimo nivel de resolución aprovechable es aquel por debajo del cual se pierden detalles necesarios en el procesamiento de la imagen, ya sean objetos pequeños que se confunden con el fondo o detalles de esos objetos, -texturas, perfiles, brillos, etc.- degradados o no identificables por la caída de resolución en el nivel donde se desearía observarlos. El nivel mínimo se relaciona, por tanto, al contexto de la escena y objetos en ella.

Una estructura multirresolución rectangular reconfigurable con N niveles de resolución, para cuya mejor comprensión se adjuntan los dibujos de Figura 2, se define, para el caso monofoveal o monoatencional, como una disposición de elementos de N resoluciones diferentes en la que se distinguen:

- 1) Una fóvea rectangular integrada por píxeles o elementos de máxima resolución.
- 2) Un conjunto de N-1 anillos que rodean la fóvea a modo de marcos rectangulares de dimensiones reconfigurables, cada uno de ellos compuesto de subanillos de la misma resolución constituidos por elementos cuya resolución decrece a medida que los anillos se alejan de la fóvea. La pérdida de resolución se asocia al tamaño creciente de los ángulos sólidos elementales abarcados por los elementos de los anillos, cuyas dimensiones crecen cuadráticamente con el número del anillo.

La estructura multirresolución rectangular monofoveal de N niveles de resolución viene determinada por la dimensión horizontal, H, y la vertical, V, de la imagen captada, ambas medidas en píxeles y múltiplos de 2^{N-1} y por cuatro parámetros i, d, a, b, que hacen referencia a distancias entre los bordes de la fóvea y de la estructura, siendo las distancias en píxeles múltiplos de sumas de potencias de 2, por la ley de crecimiento cuadrático del tamaño de los elementos que integran los anillos. Dependiendo del valor de los parámetros, se obtienen estructuras más o menos asimétricas, situándose la fóvea excéntricamente en relación a los bordes de la estructura y comprimiendo o dilatando la anchura de los anillos, o lo que es equivalente, reduciendo o aumentando el número de subanillos que componen los anillos.

La excentricidad de la estructura determina la posición relativa de la fóvea y anillos en el marco rectangular de dimensiones HxV, quedando fijadas:

- 1) La posición de la fóvea dentro de la estructura multirresolución rectangular.
- 2) El tamaño o relación de aspecto de la fóvea, ancho/alto, que puede ser mayor, igual o menor que la unidad.
- 3) La anchura, número de subanillos o máxeles, en cada lado de los anillos.

La excentricidad de la fóvea condiciona la anchura o excentricidad de los anillos que la rodean, es decir, el número de subanillos. Sin embargo, con el fin de mantener un perfil de resolución gradual a cada lado de la fóvea, se impone la condición de que el número de subanillos en los fragmentos de anillo a cada lado de la fóvea sea el mismo. Así, por ejemplo, en sentido horizontal, todos los lados derechos de los anillos tienen igual número de subanillos, y todos los lados izquierdos tienen también un número idéntico de subanillos pero, por la excentricidad posible en la estructura el número de subanillos en el lado derecho e izquierdo pueden ser distintos.

A cada elemento de las estructuras multirresolución rectangulares se asocia un dato o byte en el que se contiene la información de la intensidad o brillo de una zona de la escena. El tamaño del elemento de la estructura indica el ángulo sólido con que se capta la zona de la escena cuyo promedio de brillo es el dato asociado a ese elemento. Así pues, la estructura multirresolución rectangular representa las distintas resoluciones con que se analizan las zonas de una escena: cuanto mayor es el tamaño de un elemento, mayor es la zona captada y comprimida en un dato y, por tanto, menor es la resolución asociada a dicha zona. La fóvea contiene los elementos de menor tamaño y, en consecuencia, las zonas captadas con la fóvea son las que se observan con máxima resolución.

Lo que se pretende con el sensor multirresolución es poder observar una escena con un campo de visión amplio y bajo nivel de resolución, pero suficiente para detectar eventos que puedan requerir atención y, en tal caso, dirigir la fóvea a dichas zonas para analizarlas en detalle o con máxima resolución. Dependiendo del tamaño de la

zona en que se detecte el evento, el tamaño de la fovea deberá adaptarse al de la zona, por lo que la fovea debe ser adaptativa en anchura y altura. Por extensión, si se observan simultáneamente varios eventos, interesa disponer de varias foveas; adaptativas para poder atender a la vez a todos los eventos, en lugar de hacer exploraciones secuenciales con una sola fovea. Con este fin se utilizarían las estructuras multirresolución multifo-
veales o multiatencionales de foveas adaptativas. Su interés se justifica, por ejemplo, en una escena de tráfico, donde con poca resolución se pueden detectar objetos identificables como “camión” o “coche”, pero donde será necesaria una resolución mucho mayor para poder reconocer la matrícula. Así, será suficiente una baja resolución para detectar camiones o coches en un campo de visión amplio, como un aparcamiento, por ejemplo, una resolución algo mayor para poder observar las partes traseras o delanteras y buscar las matrículas, y otra resolución aún mayor para poder leer los caracteres de las matrículas que, dependiendo de la distancia a la que estén los móviles, podrá requerir una resolución aún mayor para leer matrículas que estén lejanas. Para realizar estos procesos lo que se hace es, dentro de la misma escena, generar imágenes de distintos tamaños y con distintas resoluciones: Se observa el aparcamiento con un gran campo de visión y baja resolución. Con menores ángulos sólidos y mayor resolución se observan los vehículos, aún menor ángulo y mayor resolución para las matrículas, etc. Es evidente que a medida que la resolución sea mayor, interesa que el ángulo sólido sea menor, para que lo sea la cantidad de datos asociada.

El procedimiento a seguir para reducir y extraer los datos de regiones de interés de las escenas con estructuras multiatencionales o multifo-
veales de N niveles de resolución consta de los siguientes pasos:

1. Obtener una pirámide de visión de N niveles a partir del nivel base o nivel 0.

Al nivel más alto o de mínima resolución se le denomina nivel N-1.

2. Definir las zonas de interés dentro del nivel N-1 y proyectarlas sobre el nivel 0.

Cada zona de interés en el nivel N-1 es un conjunto de superpíxeles, máxeles o elementos de mínima resolución, que contienen de forma comprimida la imagen del objeto o región que se desea examinar con la máxima resolución en el nivel 0. Cada elemento del nivel N-1 contiene comprimidos $2^{N-1} \times 2^{N-1}$ píxeles del nivel 0.

3. Determinar las coordenadas horizontales y verticales, mínimas y máximas, de cada una de las regiones de interés proyectadas en el nivel 0, proporcionales a las definidas en el nivel N-1, siendo 2^{N-1} la constante de proporcionalidad.

Examinando los niveles desde el vértice de la pirámide, se considera origen de coordenadas cartesianas al elemento superior-izquierdo de cada nivel.

De forma genérica se denomina a estas coordenadas $X_{\min}$, $X_{\max}$, $Y_{\min}$, $Y_{\max}$.

4. Determinar los factores de resolución a cada lado de las regiones de interés en el nivel 0. Como a cada lado de las foveas hay un número igual de subanillos en cada uno de los N-1 anillos y, dado el crecimiento cuadrático de los elementos en los anillos, el espacio en píxeles entre bordes de foveas y bordes de estructura serán múltiplos de la suma de las N-1 primeras potencias de 2.

Los factores de resolución o número de subanillos, en lados izquierdos y altos son

$$\text{factor i} = \text{INT} \left[\frac{X_{\min}}{S_0} \right]$$

$$\text{factor a} = \text{INT} \left[\frac{Y_{\min}}{S_0} \right]$$

y, de forma similar, los de los lados derechos y bajos son

$$\text{factor d} = \text{INT} \left[\frac{H - X_{\max}}{S_0} \right]$$

$$\text{factor b} = \text{INT} \left[\frac{V - Y_{\max}}{S_0} \right]$$

INT indica que son números enteros, por truncamiento de la división efectuada, lo que implica una aproximación por exceso para las dimensiones de las zonas de interés o foveas adaptativas delimitadas por los subanillos calculados por defecto.

El divisor S_0 es, para N=0, el valor del salto o desplazamiento mínimo de la fovea en el nivel 0, obtenido de la expresión

$$S_N = 2^{N+1} + 2^{N+2} + \dots + 2^A = \sum_N^{A-1} 2^{i+1}$$

El salto mínimo hace que el número de subanillos o factor de resolución en el lado correspondiente varíe en una unidad, o que los píxeles que salta la fovea sean la suma de tantas potencias de 2 como anillos, A, tenga la estructura.

5. Determinar, para cada región de interés, las coordenadas de orígenes y vértices opuestos de los rectángulos que definen las foveas y los anillos internos que las rodean en el nivel 0, en base a los respectivos factores i,d,a,b definidos en paso 4. Las coordenadas se obtienen a partir de las expresiones:

$$\text{Origen} \quad O_{xN} = S_N \cdot i \quad O_{yN} = S_N \cdot a$$

$$\text{Vértice opuesto} \quad Z_{xN} = H - S_N \cdot d \quad Z_{yN} = V - S_N \cdot b$$

El rectángulo que contiene el anillo externo es el borde de la estructura, por lo que las coordenadas de su origen y vértice opuesto son siempre (1,1) y (H,V).

6. Repetir el paso 5 para los niveles intermedios, niveles 1 a N-2, y todas las regiones de interés definidas en el paso 2, utilizando para cada nivel intermedio, N_i , el divisor S_i obtenible con la expresión de S_N definida en el paso 4.

El procedimiento para componer la estructura multiresolución multifoveal o multiatencional con los rectángulos cuyas coordenadas se calculan según pasos 5 y 6, para cuya mejor comprensión se adjuntan los dibujos de la Figura 2, se inicia con una estructura de dimensión $H \times V$ píxeles, compuesta de máxeles del nivel N-1 cuyas dimensiones son $2^{N-1} \times 2^{N-1}$ píxeles, sobre la que se posicionan en las coordenadas calculadas en el paso 6 los rectángulos del nivel N-2, compuestos por máxeles de dimensión $2^{N-2} \times 2^{N-2}$ píxeles, que sustituyen parcialmente a los máxeles preexistentes de nivel N-1. El proceso continúa posicionando los rectángulos del nivel N-(i+1), cuyos elementos sustituirán parcialmente a los preexistentes del nivel N-i, repitiéndose hasta llegar al nivel 0 en el que los rectángulos corresponden a las foveas, cuyos píxeles sustituyen parcialmente a los máxeles de nivel 1 preexistentes.

Las estructuras multiatencionales adaptativas obtenidas por superposición de los niveles de las correspondientes estructuras monoatencionales adaptativas, fusionándolas en una única estructura multiresolución, si se considera una estructura jerárquica de niveles piramidales como muestra el dibujo ejemplo de la Figura 1, son el resultado de proyectar el nivel N-1 sobre el plano base de la pirámide, tantas veces como regiones de interés existan, con ángulos de proyección tales que las proyecciones del nivel N-1 sean las foveas adaptativas en el nivel 0. Los anillos que rodean a las foveas se obtienen al proyectar los niveles intermedios sobre el nivel base de la pirámide con los mismos ángulos que se proyectó el nivel N-1. Dependiendo del tamaño de las foveas adaptativas y de la distancia entre ellas, los anillos del mismo nivel pueden llegar a fusionarse o solaparse, es decir, pueden existir elementos comunes a dos o más anillos.

Los datos y tiempos de procesamiento de las regiones de interés se reducen tanto por procesar sólo una pirámide donde todas las regiones están contenidas, como por la selección adaptativa de las regiones a partir del nivel de baja resolución.

Descripción de una realización preferida

Para adaptarse a entornos dinámicos, donde tanto la cámara como los objetos de la escena pueden ser móviles, el sistema debe operar con elevada velocidad. Para ello, la realización preferible será en *hardware*, pues los algoritmos *software* darían, en general, respuestas más lentas. El control de los procesos de lectura-escritura en memorias, ya sean SRAMs -*Static Random Access Memory*- o FIFOs, -*First In First Out*- así como la circuitería utilizada para realizar los promedios que obtienen los máxeles y enviarlos al

Sistema Central que analiza sus características, deben ser rápidos y permitir la reconfigurabilidad del sistema. Este hardware rápido y reconfigurable se tiene, por ejemplo, en los dispositivos conocidos como FPGA's -*Field Programmable Gate Arrays*- y CPLD's -*Complex Programmable Logic Devices*-. Ambos permiten realizar circuitos reconfigurables por tener estructuras compuestas de una matriz de bloques básicos y otra de bloques de entrada-salida, cuyas funcionalidades se combinan a través de otra matriz de interconexiones para obtener la funcionalidad deseada dentro del sistema que los contiene, característica conocida como ISP, acrónimo de *In-System-Programmability*. Esta característica les confiere una gran versatilidad y permite obtener funcionalidades diferentes por reconfiguración interna ejecutable en menos de 100 milisegundos. Otra alternativa es el diseño basado en ASICs -*Application Specific Integrated Circuits*-.

Con FPGAs, CPLDs o con diseños reconfigurables basados en ASICs y utilizando SRAMs y/o FIFOs externas, el sistema puede implementarse, al menos, de dos modos:

1 - Arquitectura de "Doble-Buffer"

Para su mejor comprensión se adjunta la Figura 3. La función de los bloques y señales que se muestran en ella es la que se describe a continuación:

Cámara: Suministra los Datos o píxeles de la imagen de resolución uniforme de una escena explorada de forma progresiva, en un solo campo, para evitar las distorsiones asociadas al vídeo entrelazado, PAL o NTSC. *P,L,C* son las señales de sincronismo suministradas por la cámara, que temporizan respectivamente "Píxeles", intervalos entre "Líneas" e intervalos entre "Cuadros" o imágenes de la secuencia.

Preprocesado: La secuencia de imágenes o Datos que da la cámara se preprocesa según el contexto de la escena, por una parte realizando un recorte o ventana para seleccionar el ángulo sólido que convenga dentro del proporcionado por el conjunto de óptica y sensor de la cámara y por otra parte efectuando convoluciones de tipo variable para realzar bordes, aumentar contraste, fijar umbrales, eliminar ruidos, etc., en función normalmente de la luminosidad de la escena y características de los objetos en ella. Los preprocesados se determinan por ordenes recibidas a través del bus de Instrucciones, desde el Sistema Central de Proceso de Imagen.

Sincronismos: A partir de los sincronismos *P,L,C* y en función de las características de conmutación de la lógica a controlar, en una serie de contadores dentro de este bloque se generan las direcciones A, que permitirán almacenar ordenadamente las líneas de píxeles de cada cuadro o imagen en una de las Memorias Impar y Par. Las señales B y C se activan y desactivan con cada cuadro o imagen, de forma alternada.

Memoria Impar y Memoria Par: Los datos preprocesados se almacenan alternativamente en una de las dos memorias. Mientras en una se almacena la imagen que está entrando, en las direcciones indicadas por A, la imagen anterior de la secuencia contenida en la otra memoria, se procesa según la configuración asignada al sistema y las direcciones D. Los datos a escribir o leer en las memorias fluyen por los buses bidireccionales etiquetados como *Buses Impar y Par*.

MPX1 y MPX2: Multiplexores de doble vía controlados por las señales B y C citadas. Alternadamente llevan a sus salidas los datos preprocesados o los elaborados en el bloque etiquetado como Promedios, para almacenarlos en las *Memorias Impar y Par*.

MPX3 y MPX4: Multiplexores de doble vía controlados por las señales B y C descritas. Alternadamente llevan a sus salidas, conectadas a las entradas de dirección de las memorias, las direcciones A elaboradas en el bloque de Direcciones controlado por los sincronismos de la cámara para cargar la imagen entrante, o las direcciones D generadas en el bloque *Rectángulos que controla al de Promedios*.

MPX5: Multiplexor de doble vía. Lleva a su salida los datos que se leen de la memoria en que se almacenó la imagen anterior a la entrante. Según se conecten a sus entradas los *Buses Par e Impar*, puede ser controlado por B o C.

Promedios: Hace el promedio de cuatro elementos de un nivel para sacar un elemento del nivel inmediato superior o de resolución inmediata inferior. Contiene dos sumadores de nueve bits. El primer sumador recibe en una de sus entradas la salida de un registro REG que almacena los bytes impares que salen de *MPX5* y, en la otra entrada, los bytes directos desde *MPX5*. La suma se realiza en el instante en que la salida de *MPX5* tiene estable el byte par, por lo que la salida del sumador es la suma de dos bytes consecutivos. Para que la salida sea el promedio de ambos, división por 2, se toman los ocho bits más significativos de los nueve, descartándose el menos significativo. Los promedios de dos elementos de filas o líneas impares se almacenan en una FIFO cuya salida se conecta a una de las entradas del segundo sumador, a cuya otra entrada se conecta directamente la salida del primer sumador. Las sumas se realizan sincronamente cuando entran promedios de dos elementos de líneas o filas pares, sumándose con los de las líneas impares almacenados en la FIFO. Para tener el promedio de la suma de los dos pares de elementos, es decir, el promedio de cuatro elementos en disposición 2x2, al igual que a la salida del primer sumador, se toman los ocho bits más significativos. La sincronización del bloque de *Promedios* se hace

desde el bloque *Rectángulos* gobernado por la FIFO del bloque *Coordenadas*.

Coordenadas: FIFO que contiene las coordenadas de origen y vértice opuesto de los rectángulos de interés en cada nivel piramidal y el código del nivel. Las coordenadas, calculadas en el Sistema Central de Proceso de Imagen según los criterios o funciones de la aplicación de visión que seleccionan las áreas de interés, se envían por el bus de *Instrucciones* para almacenarlas ordenadamente en la FIFO.

Rectángulos: La primera función de este bloque es extraer de cada nivel los rectángulos que contienen los datos de interés y la segunda, promediar de 2x2 a 1 el nivel del que se han extraído los rectángulos para obtener el siguiente nivel y almacenarlo para repetir de nuevo las dos operaciones de extracción y promediado en niveles sucesivos. El bloque inicia su operación cuando cambian las señales B y C, indicando que la imagen entrante ha terminado de cargarse en una de las memorias y está lista para proceso. Las direcciones D generadas a partir de las coordenadas que contiene la FIFO de *Coordenadas* controlan, a través de *MPX3* o *MPX4*, la lectura de la memoria que acaba de recibir la última imagen entrante. La memoria vuelca los datos en el *Bus Impar o Par* que, a través de *MPX6*, salen a la FIFO de *Datos* de donde serán extraídos bajo petición del Sistema Central de Proceso de Imagen. Cuando se genera la última dirección del rectángulo a leer de la memoria, se emite una señal hacia la FIFO de *Coordenadas* que, si existe otro rectángulo de interés en ese nivel, entregará las coordenadas de origen y vértice opuesto para generar las direcciones D correspondientes y, si no, entregará las coordenadas del rectángulo del siguiente nivel y el código asociado al nivel que activará el proceso de generación de promedios. Este proceso supone leer todo el nivel anterior, generando direcciones D, inyectar los datos a través de *MPX5* en el bloque *Promedios* que, operando como se ha indicado al describirlo, inyectará el promedio obtenido a través de *MPX1* o *MPX2* en la misma memoria. La dirección D donde se escribe el máxel o promedio será la del elemento del conjunto 2x2 con menor dirección o coordenadas en X e Y. Una vez promediado todo el nivel i se reinicia el proceso de lectura de rectángulos de interés en el nivel i+1 generado con los promedios. El proceso total se repite hasta generar el nivel N-1 de resolución mínima, transmitido íntegramente a la *FIFO de Datos* como rectángulo de interés total, sin realizar ya el proceso de almacenaje en memoria.

MPX6: Multiplexor de doble vía que, controlado por el bloque *Rectángulos*, selecciona el *Bus Impar o Par* en la fase de lectura de rectángulos de interés para su envío a la *FIFO de Datos*.

FIFO de Datos: Almacén temporal de los datos de rectángulos de interés a transmitir, según convenga, al Sistema Central de Proceso de Imagen.

2 - Arquitectura "Pipeline"

Esta arquitectura es aplicable en sistemas de elevada capacidad de proceso de imágenes, necesarios cuando la velocidad de los objetos observados causa una variación temporal y espacial entre imágenes sucesivas incompatible con el retardo de un cuadro introducido por la arquitectura "Doble Buffer", que procesa la imagen previa a la que en un determinado momento se está recibiendo de la cámara. En la arquitectura "Pipeline" se procesa la imagen que está entrando y, con retardo mucho menor a 2,8 milisegundos, intervalo típico entre cuadros sucesivos en la secuencia de imágenes, se tienen disponibles todos los datos que deben ser enviados al Sistema Central antes de iniciarse la generación de los correspondientes a la imagen siguiente.

Para su mejor comprensión se adjunta la Figura 4. La función de bloques y señales que se muestran en ella, donde con carácter ilustrativo y no limitativo se representa un sistema de 4 niveles de resolución, es la que se describe a continuación:

Cámara: Idéntica a la descrita en la arquitectura "Doble Buffer".

Preprocesado: Idéntica a la descrita en la arquitectura "Doble Buffer".

Sincronismos: A partir de las señales P, L, C genera las señales de sincronismo necesarias en los diferentes bloques del sistema "Pipeline". El diseño y tecnología del bloque han de ser acordes a las características de conmutación del resto de la arquitectura.

PN1, PN2, PN3: Bloques de elaboración de promedios de los niveles N1, N2 y N3. Constituyen una estructura "Pipeline" o cadena de procesamiento de datos en la que, a partir de los píxeles a la salida del bloque *Preprocesado*, simultáneamente se generan los promedios o máxeles de los niveles piramidales de resolución inferior. El funcionamiento de los bloques *PN* es idéntico al descrito para el bloque *Promedios* de la arquitectura "Doble Buffer", con la diferencia de que en aquél se procesaban los promedios o máxeles de forma sucesiva, mientras que cada uno de éstos opera a su nivel y todos ellos simultáneamente, permitiendo una mayor velocidad de procesamiento. La cadena es expandible con bloques similares si se desea configurar un sistema con mayor número de niveles de resolución.

REG0, REG1, REG2, REG3: Registros temporales para almacenar datos de cada nivel de resolución. *REG0* contiene píxeles y *REG1, REG2, REG3* máxeles. Si el comparador *CMPi* asociado indica que es dato de interés, el registro añade una cabecera relativa al nivel del dato, y si no, el contenido del registro se repone a todo a ceros.

MPX: Multiplexor de N-1 a 1 vías, siendo N el número de niveles de resolución. En sus entradas recibe las salidas de los registros *REG1* a *REG(N-1)*, una de las cuales es seleccionada para su transmisión a la salida del multiplexor.

N0 Próximo: Contador incrementable con los impulsos P generados por la *Cámara*, partiendo del estado inicial de cuenta "1". Su estado o salida es el número del píxel que estará próximamente en la salida del *Preprocesado*. Pasa al estado inicial "1" cada vez que se genera la señal L de la *Cámara*. Existe otro contador que contiene el número de líneas que se van procesando. Se repone a su estado inicial "1" cada vez que se recibe la señal C de la *Cámara*, al terminar cada imagen.

Ni Próximo: Contadores incrementales similares al anterior, para contar el número de máxeles generados por los bloques *PN*. Existen N-2 contadores además del, contador de Píxeles, ya que el contador de máxeles de nivel N-1 no es necesario, puesto que el nivel N-1 de mínima resolución siempre se transmite íntegro al Sistema Central. Sus estados iniciales son "1" y evolucionan con impulsos generados en los bloques *PN* cada vez que se genera un máxel. El contador de máxels N1, de nivel 1, se reinicia a "1" cada 2 impulsos (2^1) L de la *Cámara* y el contador de nivel 2 lo hace cada 4 impulsos L (2^2). Existen otros contadores conteniendo el número de líneas de máxeles que se van procesando en cada nivel. Se reponen a sus estados iniciales "1" cada vez que se recibe la señal C de la *Cámara*, al terminar cada imagen.

FIFO N0: De direcciones de píxeles, FIFO de cuádruple contenido, consistente en:

- a - Conjuntos de dos direcciones correspondientes a los márgenes horizontales de una zona de píxeles de interés, pudiendo existir varias zonas en una misma línea.
- b - Conjuntos de dos direcciones correspondientes a los márgenes verticales de una zona de píxeles de interés. Cada conjunto "b" puede tener varios conjuntos "a".

Las direcciones marcadas por un código cabecera se cargan en la FIFO, vía el bus de *Instrucciones*, desde el Sistema Central que define zonas de interés según criterios de selección utilizados en la aplicación de visión artificial que se procesa. La extracción de datos FIFO se hace en función del margen vertical en que está la línea procesada, con la información que suministran a la FIFO las señales del bloque *N0 Próximo*.

FIFO Ni: De direcciones de máxeles, similares en función y contenidos a los descritos en

el párrafo anterior para píxeles, pero en relación a los máxeles correspondientes del nivel N_i definidos como máxeles de interés en el Sistema Central y transmitidos para almacenar en las FIFOs conforme a los códigos de cabecera que los preceden.

CMP0, CMP1, CMP2: Comparadores de las direcciones presentes en las salidas de las FIFOs y en las de bloques *N0 Próximo* y *Ni Próximo*. Cuando la dirección presente en la salida de los bloques *N0 Próximo* o *Ni Próximo* está entre el margen del conjunto “a” presente en la salida de los bloques *FIFO N0* o *FIFO Ni*, se habilita la carga del registro *REGi* correspondiente que, con el sincronismo adecuado, almacenará el píxel que salga del bloque *Preprocesado* o el máxel que suministre el bloque *PNi* correspondiente. Cuando en una determinada línea de datos, de píxeles o máxeles, después de detectar direcciones entre márgenes se encuentra una dirección de dato próximo que ya no está dentro del margen horizontal, el comparador *CMPi* emite una señal hacia la FIFO a él conectada para extraer el nuevo margen horizontal que pudiese existir en esa línea, según lo descrito en relación a los bloques *FIFO N0* y *FIFO Ni*. Los bloques *N0 Próximo* y *Ni Próximo*, al incluir contadores de línea entrante, cada vez que cambia la línea solicitan a los bloques *FIFO N0* y *FIFO Ni* un nuevo conjunto “b”, si la línea nueva no está dentro del margen vertical presente en las salidas de las FIFOs.

FIFO de Salida: Su función es aislar los procesos de generación interna de datos y su transmisión al Sistema Central de Proceso de Imagen. El proceso de generación es ininterrumpible, realizándose a una velocidad determinada por el “Pipeline” y los sincronismos asociados, mientras que el proceso de transmisión es fraccionable.

Dado que en la arquitectura “Pipeline” se tiene generación simultánea de píxeles y máxeles, es frecuente la situación en que simultáneamente se generan elementos de los cuatro niveles que, por pertenecer todos ellos a rectángulos de interés, deben ser transmitidos al Sistema Central. Para evitar rebosamientos de datos en la FIFO de salida, el ancho de ésta es, al menos, el doble del ancho de los datos a transmitir, cargándose los píxeles directamente desde *REG0* en la parte izquierda de la FIFO, mientras en la parte derecha se cargan los máxeles que, desde sus respectivos *REGi* y a través de *MPX*, se seleccionan con el sincronismo apropiado. El aumento del ancho de los datos a transmitir disminuye el tiempo de su transmisión, siendo posible por la anchura del bus de datos del Sistema Central, el *Peripheral Components Interconnect* o PCI-bus de 32 bits y en evolución a 64. Se adjunta la Figura 5 para una mejor comprensión de la temporización de píxeles y máxeles que determina la composición de

los datos de doble ancho, o palabras, que se cargan en la FIFO de salida.

Descripción de los dibujos

Para complementar la descripción, con objeto de ayudar a una mejor comprensión de las características del invento y como parte integrante de esta memoria, se incluye un conjunto de dibujos donde, con carácter ilustrativo y no limitativo, se ha representado lo siguiente:

Figura número 1: Pirámide cuya base es el campo de visión o imagen suministrada por la cámara, identificado como nivel 0 o de máxima resolución. Los sucesivos niveles, en el caso de gradiente de resolución cuadrático, tienen $1/4$ del área del nivel inmediato inferior. Todos los niveles abarcan el mismo campo de visión, pero con resoluciones decrecientes a medida que se acercan al vértice. Sobre el nivel 0 se han dibujado, con gris oscuro, dos posibles regiones de interés. Los rectángulos que delimitan esas regiones forman troncos de pirámide o túneles de visión con superficie superior común coincidente con el último nivel de resolución de la pirámide construida que, en la figura, es el nivel 3. Los niveles intermedios 1 y 2, al intersectar a los troncos de pirámide, fijan las zonas de interés en cada nivel de resolución. Las proyecciones de los niveles de resolución sobre el campo de visión, realizadas paralelamente a las caras laterales de los troncos de pirámide citados, marcan sobre el campo de visión las zonas o anillos de resolución de lo que se conoce como imagen multifoveal, por su imitación a la resolución gradual del ojo humano. Para mayor claridad, solo se muestran las líneas de proyección de las esquinas izquierdainferior y derecha-superior, terminadas en punta de flecha.

Figura número 2: Está relacionada al ejemplo de la Figura número 1 y muestra separadas las estructuras foveales que corresponden a las regiones de interés o fóveas señaladas en la Figura 1 en tono gris oscuro. La determinación de posiciones y tamaños de las zonas de resolución en las Figuras 2a y 2b, se hace por los parámetros i, d, a, b que indican saltos de resolución en lado izquierdo, derecho, alto y bajo. La Figura 2c muestra la fusión de las figuras 2a y 2b, o imagen bifoveal que se obtiene por las proyecciones descritas para la Figura 1. En el dibujo 2c se observa la zona fusionada de máxeles del nivel 2 de la pirámide, correspondiente al gris claro en la base de la pirámide de Figura 1 y las zonas de resolución de nivel 1, o primeros anillos, con tono gris medio en Figura 1. Los dibujos 2d, 2e, 2f y 2g se acompañan para mostrar como se pueden obtener imágenes multifoveales con más de dos regiones de interés.

Figura número 3: Presenta una arquitectura “Doble Buffer” adaptada a la obtención de

los datos de interés o rectángulos de los niveles piramidales que se muestran en Figura 1. La arquitectura “Doble Buffer” contiene dos estructuras gemelas de memorias y circuitos de direccionamiento y lectura-escritura, conectados a bloques externos relacionados a la entrada de datos -*Cámara*, *Sincronismos*, *Preprocesado*-, a la selección y procesado de datos -*Promedios*, *MPX5*, *Rectángulos y Coordinadas*- y a la extracción o salida de datos -*MPX6* y *FIFO de Datos*. El diagrama muestra los buses de direcciones con flechas anchas de fondo blanco y los buses de datos e instrucciones con flechas anchas de fondo negro. Las señales de control se muestran con flechas finas.

Figura número 4: Presenta una arquitectura “Pipeline” adaptada a la obtención de los datos de interés o rectángulos de los niveles piramidales que se muestran en la Figura 1.

A partir de los datos de la cámara, preprocesados o acondicionados para resaltar sus características y con los sincronismos de la cámara que permiten situar píxeles y máxeles en relación a coordenadas de los niveles piramidales, la arquitectura “Pipeline” contiene una cadena de procesado de datos -*PN1*, *PN2*, *PN3*- que genera simultáneamente los máxeles que constituyen los niveles piramidales. Con cada máxel se emite un impulso que, contabilizado en los bloques *N1 Próximo* y *N2 Próximo*, relaciona el máxel a sus coordenadas en el nivel piramidal correspondiente. De forma similar se hace con los píxeles en el bloque *N0 Próximo*. Para determinar si los píxeles y máxeles corresponden a los situados dentro de rectángulos piramidales de interés, se disponen unos comparadores -*CMP0*, *CMP1*, *CMP2*- que indican si la dirección o número de píxel o máxel está dentro del rango de interés contenido en las *FIFOs N0*, *N1* y *N2*, donde el Sistema Central ha cargado los rangos de interés. Si la comparación realizada con antelación a la llegada del píxel o la generación del máxel fuese positiva, los píxeles y máxeles se cargan en registros -*REG0*, *REG1*, *REG2*, *REG3*-, donde se les añade una cabecera para relacionar elementos a niveles. Los comparadores y contadores generan señales que permiten ir extrayendo de las *FIFOs* los sucesivos rangos de direcciones para comparación progresiva de píxeles y máxeles. El último nivel piramidal se transmite integro y, por tanto, no necesita comparación, pero

sí registro de los máxeles.

Para evitar rebosamiento de datos a transmitir, se dispone una *FIFO de salida* de doble ancho, en la que se cargan los datos con la temporización que detalla la Figura número 5.

Figura número 5: Presenta a modo de ejemplo, con carácter ilustrativo y no limitativo, un fragmento de 16 columnas y 8 filas de la zona superior-izquierda, o inicial, de una posible imagen de H columnas x V filas de píxeles, siendo ambas dimensiones H y V múltiplos de 2^{N-1} píxeles, donde N es el número de niveles de resolución. A su lado derecho, otro dibujo muestra el tamaño relativo de píxeles y máxeles en el caso o ejemplo de ser $N=4$. Se supone, como caso peor, que el fragmento de imagen corresponde a una configuración multirresolución en la que todos los elementos que se generan al ser explorada esa zona están en rectángulos de interés, por lo que deben ser transmitidos a la *FIFO de salida*. Sobre el fragmento de imagen están marcados, con los símbolos que figuran al pie, los píxeles que generan en su tiempo los tipos de máxeles que corresponden al caso de $N=4$. Para ilustrar la temporización con que se envían a la *FIFO de salida* los datos que se generan en la arquitectura “Pipeline”, en relación con el fragmento de imagen, se presenta una Tabla en cuya primera columna y en sentido descendente, se indican los tiempos de los píxeles que componen la fila 8 del fragmento de imagen, recorrida de izquierda a derecha. En la segunda columna de la Tabla se indican los elementos, píxeles y máxeles, generados con cada píxel de dicha fila. En la tercera columna se indica cómo se transfieren los elementos de la columna segunda, por conjuntos de dos, para ser cargados en instantes sucesivos en las dos partes de que consta la *FIFO*, lo que permite observar la periodicidad con que salen los elementos de los diferentes niveles, cuyo ciclo completo es 16 tiempos de píxel, en los que aún existen tres donde la *FIFO* no se completa, pudiendo utilizarse la misma estructura de *FIFO* para sistemas de más de 4 niveles de resolución. Se han incluido en la Tabla los tiempos de píxeles 17 a 20 que, aunque no se muestran en el fragmento de imagen del dibujo, se entiende que existen en la imagen íntegra. En esos píxeles se repite la secuencia de salida que se presentaba en los tiempos de píxeles 9 a 12.

REIVINDICACIONES

1. Sistema multiatencional adaptativo de estructura multirresolución reconfigurable, **caracterizado** por su capacidad de determinar simultáneamente varias zonas de interés en imágenes de baja resolución, obtenidas por N-1 compresiones sucesivas de 2x2 elementos a 1, siendo N un número entero y procediendo los elementos de imágenes suministradas con resolución uniforme y máxima, cuyas regiones de interés se analizan a partir del tamaño y posición de las zonas determinadas en las imágenes de baja resolución del nivel N-1 y analizando el resto de la imagen original con resoluciones periféricas decrecientes, permitiendo tiempos de procesamiento menores que los necesarios si las determinaciones se hicieran secuencialmente o en alta resolución.

2. Sistema multiatencional adaptativo de estructura multirresolución reconfigurable, según reivindicación 1, **caracterizado** por la estructura compuesta por superposición lineal en sentido creciente, de los niveles de resolución de varias estructuras rectangulares monofoveales con N niveles de resolución, cada una de ellas **caracterizada** por una disposición de elementos de N resoluciones diferentes, en la que se distinguen:

- 1) Una fóvea rectangular integrada por píxeles o elementos de máxima resolución.
- 3) Un conjunto de N-1 anillos que rodean la fóvea a modo de marcos rectangulares de dimensiones reconfigurables, cada uno de ellos compuesto de subanillos de la misma resolución, compuestos a su vez de elementos o máxeles cuya resolución decrece a medida que los anillos se alejan de la fóvea, correspondiendo a los elementos ángulos sólidos elementales de dimensión creciente al captar zonas periféricas de la escena.

3. Sistema multiatencional adaptativo de estructura multirresolución reconfigurable, según reivindicación 2, con estructuras multifresolución rectangulares **caracterizadas** por dimensiones externas H y V, que pueden ser iguales, y cuatro parámetros definidos como:

- 1) Dimensión horizontal, H, y dimensión vertical, V, especificadas en píxeles.
- 2) Parámetros i, d, a, b, números enteros que definen la excentricidad de la estructura, fijando la posición y relación de aspecto de las zonas de diferente resolución, o sea:
 - a) Posición de la fóvea dentro de la estructura multirresolución rectangular.
 - b) Relación de aspecto de la fóvea, ancho/alto, mayor, igual o menor que la unidad.
 - c) Anchura, número de subanillos o máxeles, de cada uno de los lados de los anillos.

4. Sistema multiatencional adaptativo de estructura multirresolución reconfigurable, según reivindicaciones 1 a 3, cuya estructura multirresolución rectangular de N niveles de resolución

para varios objetos o regiones de interés se **caracteriza** por una disposición de elementos de N resoluciones diferentes, resultante de superponer en una estructura única las estructuras monoa-tencionales de todas las regiones de interés que existan en la escena que, al fusionar sus respectivos niveles, originan una estructura única de igual tamaño que las estructuras componentes, que conduce a una reducción de los datos o elementos que resultarían si se considerasen aisladamente las estructuras multirresolución integra-tes.

5. Sistema multiatencional adaptativo de estructura multirresolución reconfigurable, según reivindicaciones 1 a 4, en que el procedimiento a seguir para obtener las estructuras multiatencio-nales con N niveles de resolución se **caracteriza** por los siguientes pasos:

1. Obtener una pirámide de visión de N niveles, cuyo nivel 0 o nivel de partida es la base de máxima resolución y el nivel N-1 el más alto o de mínima resolución.
2. Definir todas las zonas de interés dentro del nivel N-1.
3. Determinar las coordenadas horizontales y verticales, mínimas y máximas, de cada región de interés en el nivel N-1, denominadas $X_{\min}$, $X_{\max}$, $Y_{\min}$, $Y_{\max}$.
4. Determinar los factores de resolución lateral para todas y cada una de las regiones de interés en el nivel 0, por medio de las fórmulas:

$$\text{factor i} = \text{INT} \left[\frac{X_{\min}}{S_0} \right]$$

$$\text{factor a} = \text{INT} \left[\frac{Y_{\min}}{S_0} \right]$$

$$\text{factor d} = \text{INT} \left[\frac{H - X_{\max}}{S_0} \right]$$

$$\text{factor b} = \text{INT} \left[\frac{V - Y_{\max}}{S_0} \right]$$

donde INT indica truncamiento al valor entero de la división efectuada, siendo el divisor S_0 el valor del salto o desplazamiento mínimo de la fóvea en el nivel 0, obtenido haciendo N=0 en la expresión:

$$S_N = 2^{N+1} + 2^{N+2} + \dots + 2^A = \sum_{N=0}^{A-1} 2^{i+1}$$

5. Determinar, para cada región de interés, las coordenadas del origen y vértice opuesto de los rectángulos que definen la fóvea y anillos internos en el nivel 0, en base a los respectivos factores i,d,a,b definidos en el paso 4, y utilizando para la determinación las expresiones:

- Origen $O_{xN}=S_N.i$ $O_{yN}=S_N.a$
 Vértice
 opuesto $Z_{xN}=H-S_N.d$ $Z_{yN}=V-S_N.b$
6. Repetir el paso 5 para los niveles intermedios, niveles 1 a N-2, y todas las regiones de interés definidas en el paso 2, utilizando con cada nivel intermedio N_i , el divisor S_i definido en el paso 4.
 7. Componer la estructura multifoveal o multiatencional por superposición de las estructuras multirresolución de N niveles correspondientes a cada uno de los objetos o regiones de interés, obtenidas siguiendo los pasos 2 a 6.
6. Sistema electrónico con arquitectura “Doble Buffer”, para extracción ordenada y, secuencial por niveles de los datos correspondientes a los rectángulos de interés situados, en los niveles piramidales, según las reivindicaciones 1 a 5, realizable sobre FPGAs, CPLDs o ASICs y **caracterizado** por tener:
- a) Dos conjuntos o “Buffers” gemelos, compuestos de memorias SRAM, multiplexores, y buses asociados al direccionamiento y escritura-lectura de datos en las memorias.
 - b) Una circuitería común a los conjuntos gemelos, que conmuta su funcionalidad de uno a otro para, alternadamente, escribir en una de las memorias los datos preprocesados de las imágenes de resolución uniforme entregadas por una cámara de exploración progresiva, controlando la escritura y su alternancia con sincronismos y controles derivados de los sincronismos generados en la cámara.
 - c) Una FIFO común a los conjuntos gemelos, donde se escriben las coordenadas de los rectángulos de interés de cada nivel piramidal, enviadas por un bus de Instrucciones conectado al Sistema Central que selecciona y actualiza las regiones de interés en los niveles piramidales de las imágenes, calculando las coordenadas correspondientes.
 - d) Una circuitería común a los conjuntos gemelos, que conmuta su funcionalidad de uno a otro para, alternadamente y en función de las coordenadas que contiene la FIFO citada en c), leer los datos de los rectángulos de interés del nivel base en la memoria donde se cargó la última imagen, promediar el nivel reduciendo sus datos de 2x2 a 1, escribir los promedios resultantes en la misma memoria donde estaba el nivel origen para tener en ella el siguiente nivel de resolución, y repitiendo todo el proceso hasta obtener el nivel de mínima resolución que, sin reescribir en la memoria, se envía a una FIFO que actúa de “buffer” común a la salida de datos del “Doble Buffer”.

7. Sistema electrónico con arquitectura “Pipeline” que, simultáneamente a la entrada de imágenes, extrae los datos correspondientes a los rectángulos de interés, según las reivindicaciones 1 a 5, realizable sobre FPGAs, CPLDs o ASICs y **caracterizado** por tener:

- a) Una cadena de N-1 promedidores de 2x2 a 1, siendo N el número de niveles de resolución, que utilizando los píxeles generados por una etapa de preprocesado de las imágenes y con sincronismos generados a partir de los suministrados por la cámara de exploración progresiva, generan los máxeles de los N-1 niveles de resoluciones piramidales inferiores de forma continua y en orden progresivo de coordenadas.
- b) Un conjunto de N-1 contadores cíclicos asociados a los niveles piramidales 0 a N-2, siendo sus máximos estados de cuenta respectivos las longitudes H_N de las líneas de elementos que componen el nivel piramidal correspondiente y “1” el estado inicial de todos ellos, al que se reponen tras recibir múltiplos de 2^N impulsos “L”, de intervalos de líneas de la cámara progresiva, siendo N el nivel asociado al contador.
- c) Un conjunto de N-1 FIFOs asociadas a los niveles 0 a N-2, conteniendo los rangos horizontales y verticales de coordenadas correspondientes a los rectángulos de interés en cada nivel piramidal.
- d) Un conjunto de N-1 comparadores, asociados a los conjuntos de contadores y FIFOs citados en b) y c), para detectar si el próximo píxel a entrar o los máxeles a generarse en los promedidores de a), pertenecen a los rangos contenidos en las FIFOs.
- e) Un conjunto de N registros, los primeros N-1 asociados a los niveles piramidales 0 a N-2, para almacenar temporalmente, uno a uno, los píxeles entrantes o los máxeles generados cuyas coordenadas pertenecen al rango de interés comparado, y el último registro asociado al nivel más alto N-1, que almacena temporalmente, uno a uno, todos los elementos generados por el promediador asociado al nivel N-1, añadiendo cada registro una cabecera indicadora del nivel del elemento que contiene.
- f) Un multiplexor o selector de los datos contenidos en los N-1 registros asociados a los niveles 1 a N-1 citados en e).
- g) Una FIFO cuya anchura mínima es el doble de aquella de los datos contenidos en los registros citados en e), una de cuyas mitades carga los datos existentes en el registro asociado al nivel 0, mientras la otra carga el registro seleccionado por el multiplexor.

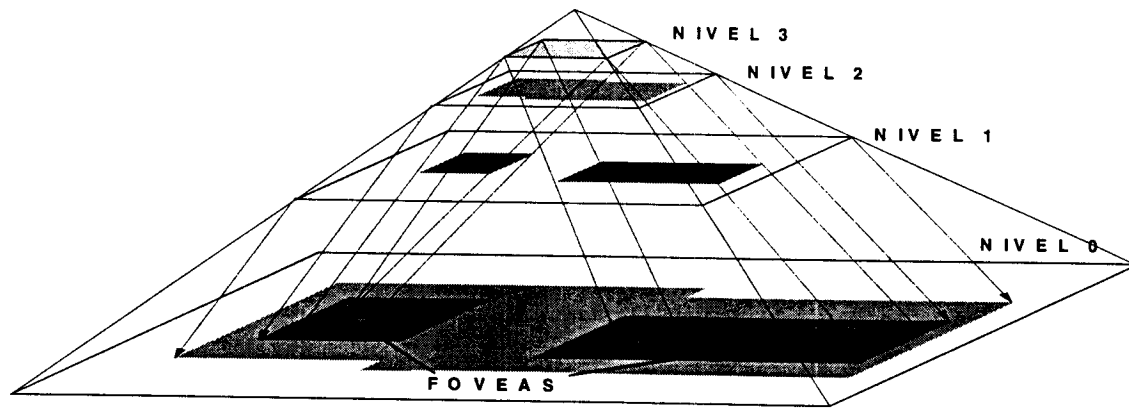


Figura 1

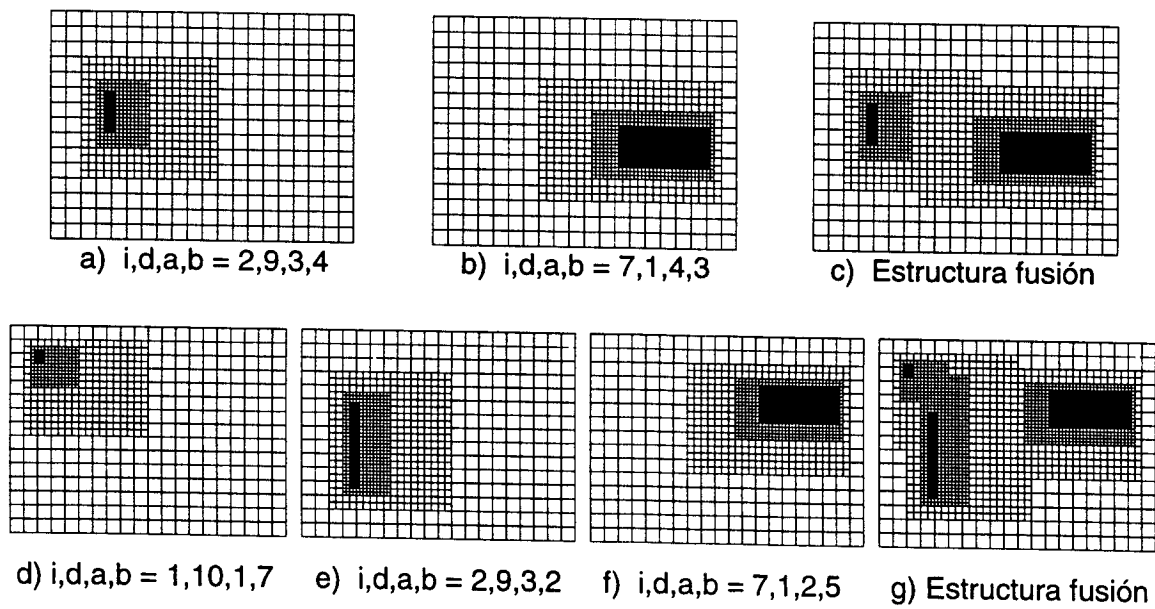


Figura 2

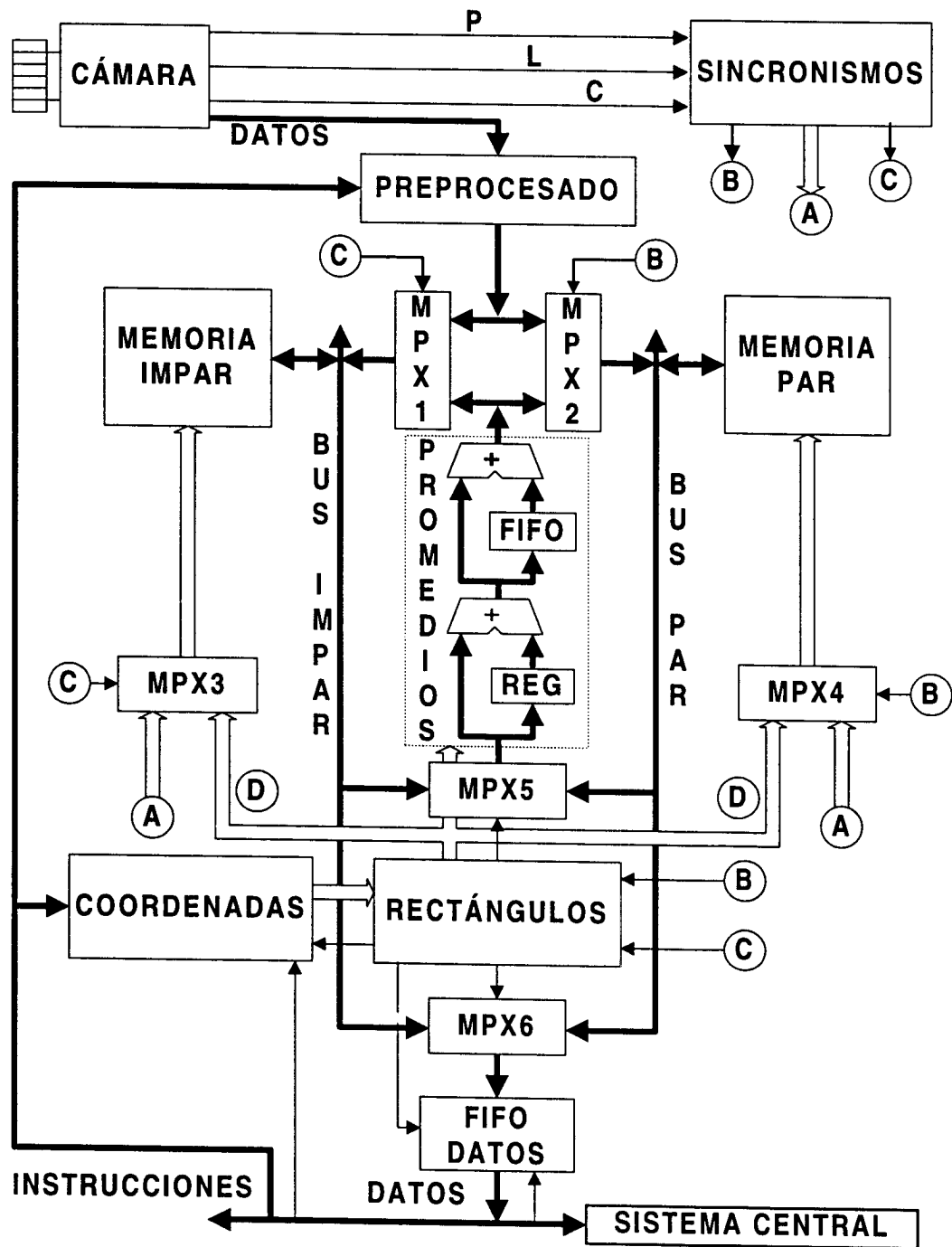


Figura 3

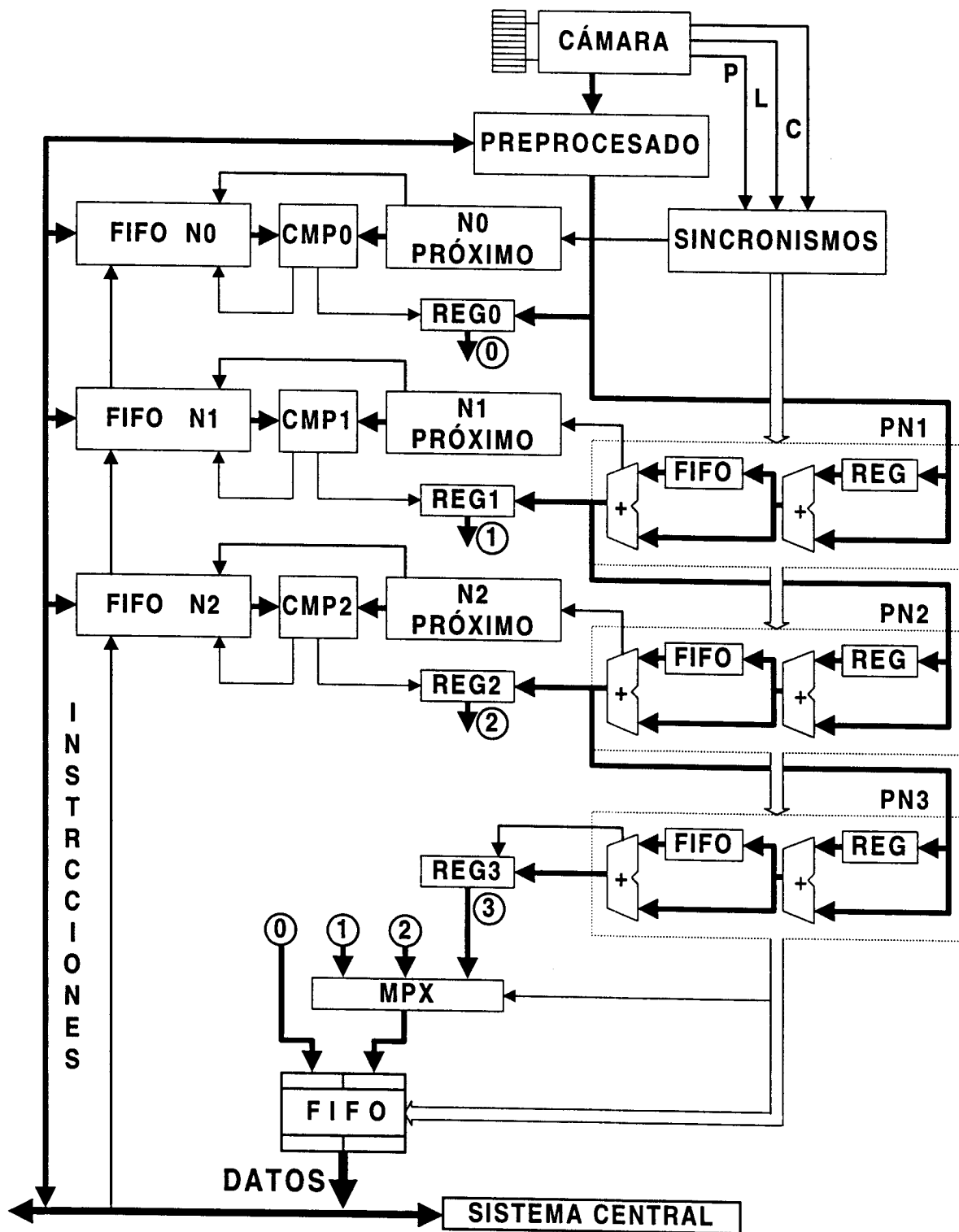
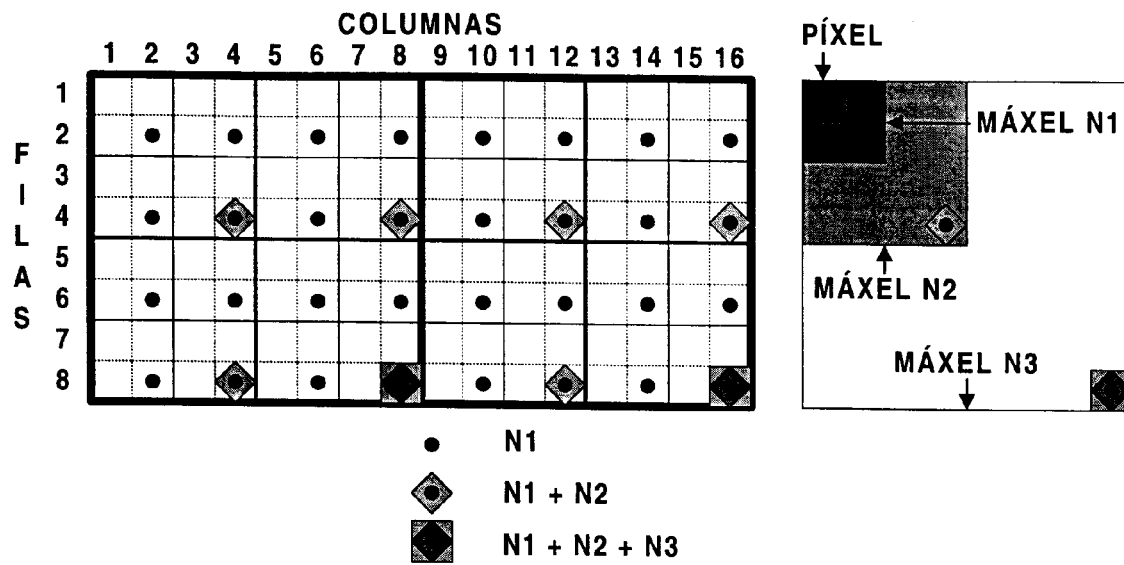


Figura 4



TIEMPO	ELEMENTOS	TRANSFERENCIAS
1	P	-
2	P, N1	P
3	P	P, N1
4	P, N1, N2	P
5	P	P, N1
6	P, N1	P, N2
7	P	P, N1
8	P, N1, N2, N3	P
9	P	P, N1
10	P, N1	P, N2
11	P	P, N1
12	P, N1, N2	P, N3
13	P	P, N1
14	P, N1	P, N2
15	P	P, N1
16	P, N1, N2, N3	P
17	P	P, N1
18	P, N1	P, N2
19	P	P, N1
20	P, N1, N2	P, N3

Figura 5



OFICINA ESPAÑOLA
DE PATENTES Y MARCAS
ESPAÑA

- ⑪ ES 2 152 176
⑫ N.º solicitud: 009900346
⑬ Fecha de presentación de la solicitud: 19.02.1999
⑭ Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑮ Int. Cl.⁷: G06T 7/00

DOCUMENTOS RELEVANTES

Categoría	Documentos citados	Reivindicaciones afectadas
X A	CAMACHO, P. et al. "Multiresolution sensors with adaptive structure". En: Proceedings of the 24th Annual Conference of the IEEE Industrial Electronics Society. 1998, New York. ISBN 0-7803-4503-7. Volumen 2, páginas 1230-1235.	1-6 7

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

El presente informe ha sido realizado

☒ para todas las reivindicaciones

☐ para las reivindicaciones n.º:

Fecha de realización del informe 29.11.2000	Examinador M. Alvarez Moreno	Página 1/1
--	---------------------------------	---------------