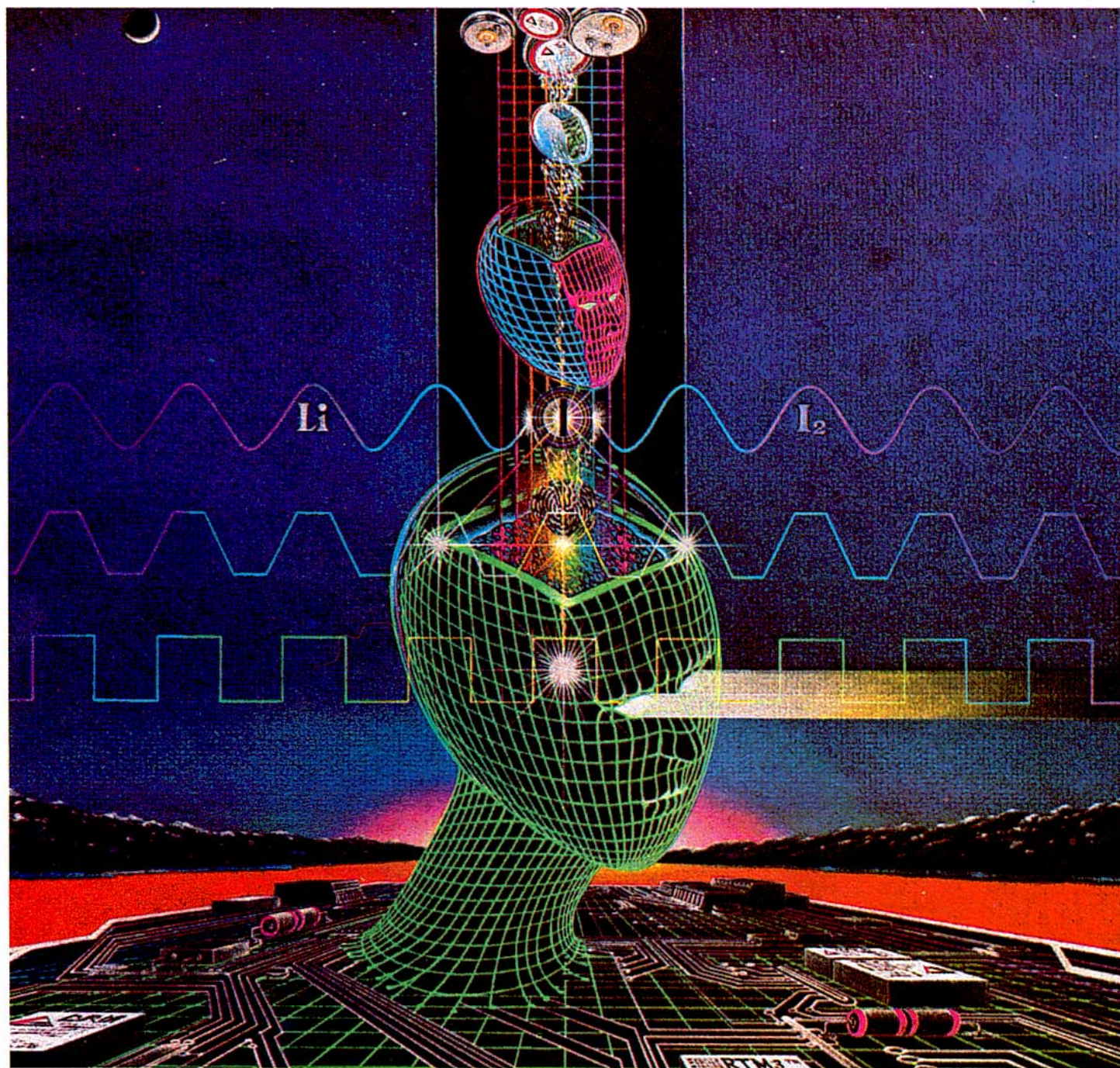




Arquitecturas para visión artificial

El procesamiento digital de imágenes, debido a la gran cantidad de datos que se manejan, es uno de los campos que ha impulsado la investigación de nuevas arquitecturas de ordenadores orientadas a reducir el tiempo de proceso. Sobre este tema escriben MATEO LUQUE SANCHEZ y JOSE P. RUEDA CRUZ.

Las primeras manifestaciones en el campo del tratamiento digital de la imagen pretendían hacer uso de la posibilidad de comunicación entre seres humanos mediante imágenes. Para ello se requería una adecuación de las señales eléctricas transformadas de escenas naturales, a la "calidad" y "cantidad" necesaria para que se efectuase la comunicación.

Nacen las técnicas de procesamiento digital de imágenes como intento de superar este reto: mejoras de calidad mediante el uso de técnicas de filtrado digital, técnicas de compresión, reconstrucción de escenas, ..., que van ampliando sus fronteras con la intervención del ordenador.

La aplicación de las técnicas de Procesamiento de Imágenes dentro del marco de la Inteligencia Artificial, nos conducen al concepto de Visión por Ordenador, o Visión Artificial, con el objeto de dotar a la máquina de una nueva capacidad sensorial como es la percepción visual.

A semejanza del Sistema Visual Humano, todo sistema de procesamiento de imágenes requiere la utilización de un sensor óptico capaz de captar la información contenida en una escena natural. El tipo de sensor óptico más utilizado es una cámara de vídeo, que realiza la conversión eléctrica de la imagen reflejada por la escena que se quiere analizar.

Normalmente, un sistema de procesamiento de imágenes conlleva el uso de un ordenador digital capaz de procesar la gran cantidad de información contenida en una imagen. La naturaleza analógica de la señal eléctrica procedente de la cámara de vídeo hace necesario el uso de una etapa de conversión analógico/digital para obtener una representación numérica de la imagen capaz de ser procesada por un ordenador, que llamaremos imagen digital, o simplemente imagen.

La imagen digitalizada es una representación numérica de la escena, organizada en forma matricial, de manera que cada elemento de la matriz, llamado pixel (picture element), se corresponde con un punto de la imagen, y su valor numérico representa el nivel de intensidad luminosa o nivel de gris (para el caso de imágenes monocromáticas) de la misma, realizándose normalmente la codificación de cada pixel con 4, 6 u 8 bits que proporcionan 16, 64 ó 256 niveles de gris, respectivamente. Según se hayan tomado más o menos muestras de la imagen obtendremos una mayor o menor resolución, que se suele expresar en (nº de líneas) x (nº de pixels/línea), siendo resoluciones típicas: 64 x 64, 256 x 256, 512 x 512, 1024 x 1024, ..., según aplicaciones, que como se puede ver son potencias de la base 2 con igual resolución horizontal y vertical, aunque existan sistemas que no se atienen a estos valores.

Los algoritmos de visión determinan la arquitectura

Tradicionalmente, un Sistema de Visión Artificial (S.V.A.) presenta dos etapas

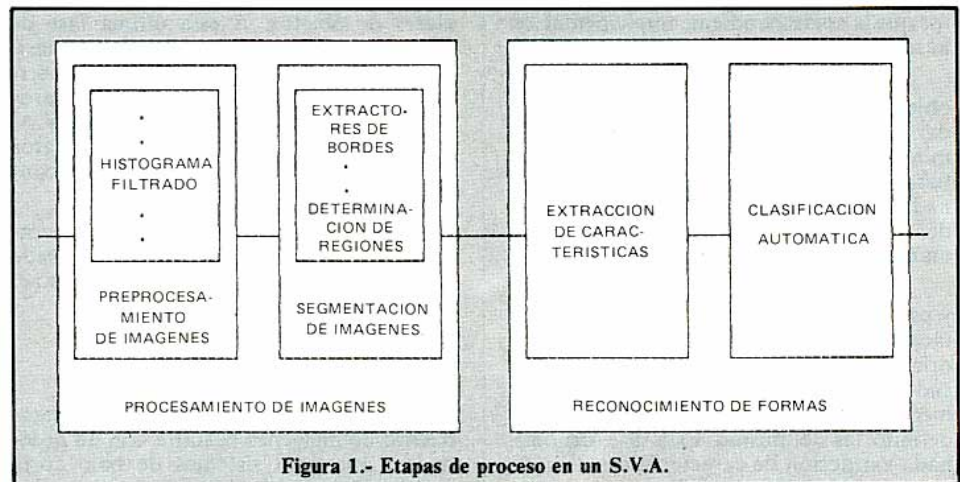


Figura 1.- Etapas de proceso en un S.V.A.

de proceso funcionalmente distintas, que se suelen denominar: Procesamiento de Imágenes y Reconocimiento de Formas.

La primera engloba todas las operaciones destinadas a meras transformaciones de imágenes para lograr una representación individualizada de los objetos presentes en la escena. Los S.V.A. utilizan gran parte de las técnicas de Procesamiento de Imágenes, con una problemática común a todas ellas que se puede resolver con arquitecturas semejantes.

Como se puede observar en la figura 1, la etapa de Procesamiento de Imágenes, a su vez, se puede subdividir en dos grandes bloques de proceso, una primera fase en la que se pretende mejorar la "calidad" de la imagen, llamada procesamiento previo o

preprocesado, y una segunda fase en la que se pretende distinguir objetos o separar regiones presentes en la imagen, que se suele llamar segmentación.

Una característica común a todas estas técnicas de procesamiento es la organización interna de los datos en forma matricial, apropiada para un procesamiento paralelo. Por ello, una arquitectura para un S.V.A. debería estar dotada de la capacidad de procesamiento en paralelo mediante estructuras hardware añadidas al computador central. No obstante, es muy útil disponer de dispositivos hardware, implementados con tecnología VLSI, que realicen algoritmos específicos de procesamiento, como calculadores del histograma, convolucionadores para filtrado espacial, extractores de bordes, etc., en tiempos mucho más cor-

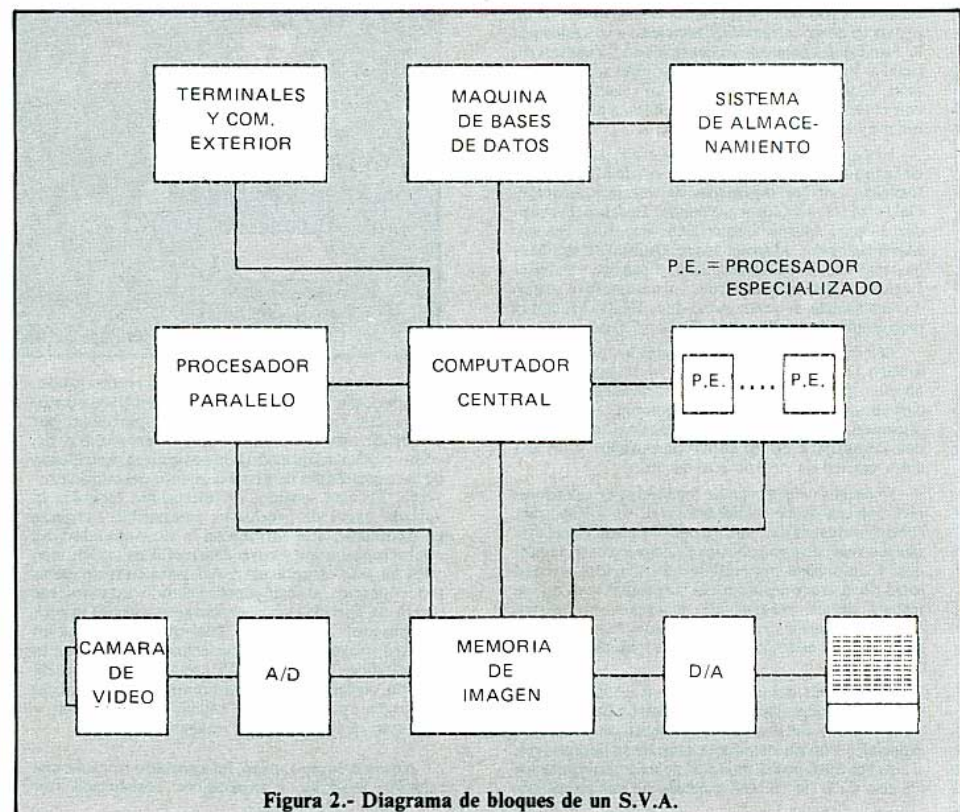


Figura 2.- Diagrama de bloques de un S.V.A.

tos que la correspondiente implementación hardware en otras partes del sistema.

Una vez aislados y diferenciados los objetos que componen la escena, se puede dar por concluida la etapa que se ha denominado de Procesamiento de Imágenes, para pasar a una segunda etapa denominada Reconocimiento de Formas, que puede encuadrarse propiamente dentro del marco de la Inteligencia Artificial.

En una primera fase de reconocimiento, y partiendo normalmente de una representación simbólica de los objetos, como podría ser la codificación de sus bordes en cadena, se pretende representar el objeto mediante un conjunto de características definitorias del mismo. Esta fase, denominada extracción de características, tiene la particularidad de extraer de un mismo conjunto de datos, resultados que globalmente forman la representación vectorial del objeto llamada vector de características. Se podría pensar en utilizar para esta fase una arquitectura paralela capaz de ejecutar simultáneamente distintos algoritmos sobre un mismo conjunto de datos.

El objetivo final del Reconocimiento de Formas es identificar el objeto, representado por su vector de características, como perteneciente a una clase de objetos determinada, cuyos parámetros característicos podrían estar almacenados en una base de datos, que integrara los conocimientos adquiridos sobre un universo de objetos y

clases de objetos. A esta última fase de reconocimiento se le suele denominar clasificación automática, según se puede observar en la figura 1, y su optimización hardware podría obtenerse dotando al S.V.A. de la facilidad de gestión de bases de datos mediante una máquina especializada controlada por el ordenador central.

Tras la exposición de los rasgos generales que debiera reunir un S.V.A., se puede proponer un diagrama de bloques como el representado en la figura 2.

Procesamiento en tiempo real

En principio, para realizar procesamiento de imágenes bastaría con un ordenador, que para sistemas de bajo coste podría ser un ordenador de los llamados personales, un sistema de captación compuesto de una cámara de vídeo u otro tipo de sensor y un digitalizador, una memoria donde almacenar la información obtenida de la etapa anterior que puede ser parte de la propia memoria central del ordenador, o mejor, en una memoria independiente, en cuyo caso se le denomina "frame buffer" o "memoria de imagen", y por último una etapa de presentación donde se visualicen, a través de una pantalla gráfica, los resultados del procesamiento realizado.

Por supuesto, a un sistema mínimo como el descrito en el párrafo anterior no se le pueden pedir grandes prestaciones, ni

hablar de ejecución en tiempo real en la mayoría de las aplicaciones.

El principal problema que se presenta en el procesamiento de imágenes y visión artificial es el enorme coste en términos de tiempo de proceso, debido a la complejidad de los algoritmos de visión y a la cantidad de datos a manejar, a modo de ejemplo, el satélite de observación terrestre LANDSAT envía a la Tierra alrededor de 200 imágenes diarias con una resolución de 2.500 x 3.300 pixels, teniendo en cuenta que algunos algoritmos de segmentación en regiones para estas imágenes tardan en ejecutarse del orden de 7 horas en un computador VAX 11/780, y que tan sólo se trata de una parte del procesamiento total necesario en un S.V.A., nos podemos hacer una idea del número de horas necesarias para procesar toda esta información en un ordenador convencional.

La implementación hardware de ciertos algoritmos de procesamiento de imágenes dotarían al sistema de una mayor velocidad de cálculo, que permitiría la ejecución en tiempo real para algunas aplicaciones muy concretas. No obstante, un sistema dotado exclusivamente de estos procesadores específicos sería demasiado rígido habida cuenta que la mayoría de los algoritmos de visión varían según el tipo de imágenes que hayan de procesar, por ejemplo el tipo de filtrado espacial a realizar depende de las condiciones de captación, ruido, contraste,

Un sistema de visión artificial en la UPM

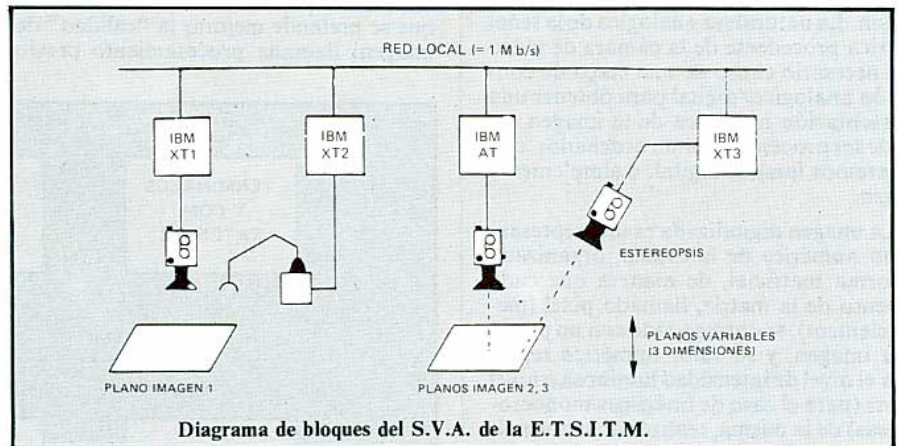
Los autores de este artículo pertenecen a un equipo interdisciplinario de ingenieros de telecomunicación y psicólogos, trabajando en el marco de un proyecto de investigación financiado a tres años por la Fundación Ramón Areces titulado "Visión Artificial y Visión Humana. Aplicaciones a la Percepción Visual en Robótica". Por su relación directa con el tema del artículo aprovechamos para describir muy brevemente este proyecto.

El equipo de ingenieros, que trabaja en la ETS de Ingenieros de Telecomunicación de Madrid, está formado por los ingenieros de telecomunicación: Darío Maravall Gómez-Allende, Eusebio Fernández López, Mateo Luque Sánchez, José Ignacio Martí Sempere, Manuel Mazo Quintas, Félix Monasterio-Huelin y Maciá, Juan Francisco Paredes Lozano y José Rueda Cruz, y un numeroso grupo de estudiantes de sexto curso de la ETSITM, con la realización de su Proyecto Fin de Carrera.

El equipo de psicólogos, dirigidos por el catedrático D. Luis Jánhez Escalada, se ha propuesto el objetivo de llegar a algoritmos inspirados en el funcionamiento del sistema visual humano, para lo que disponen de una potentísima instalación informática compuesta de un supercomputador y un sistema gráfico de tiempo real de vídeo.

El objetivo del grupo de ingenieros de telecomunicación consiste en el desarrollo de un S.V.A. para robots industriales, trabajando en condiciones de tiempo real, que en robótica son unos pocos segundos. Como base material se ha escogido una red local de microcomputadores (ver figura), ya que se trabaja en un entorno de procesamiento distribuido, con varias cámaras enfocadas hacia diferentes escenas y, además, con un manipulador mecánico.

Aunque en principio se pensó en emplear un superminiordenador de 32 bits, con un bus estandarizado (a fin de incorporar tarjetas de propósito especial) y con un sistema operativo de tiempo real, al final se optó por la red local de microcomputadores con CPU de 16 bits, también de bus estandari-



zado y con tarjetas de procesamiento especializado. Esta decisión se realizó no sólo por motivos económicos, con ser importantes, sino en particular por constituir una solución modular y expandible, así como por su adaptación a la naturaleza distribuida de las necesidades de procesamiento del citado proyecto. En este sentido, es interesante recordar la variada gama de productos compatibles, tarjetas en particular, que enriquecen la funcionalidad del microcomputador y cuyo desarrollo en los últimos años ha sido espectacular. Así, para citar un ejemplo concreto, se está desarrollando actualmente (fuera de España) un coprocesador para el microcomputador IBM PC-AT (del que se dispone en nuestro proyecto, con arquitectura SIMD y con la finalidad de dotar al PC-AT con una capacidad de cálculo cercana a un supercomputador actual, al menos en aplicaciones de cálculo intensivo matricial, cual es el proceso de imágenes).

Nuestro equipo, pues, ha apostado por una instalación modular y expandible, constituida por

varios microcomputadores con CPUs de 16 bits conectados en red local, cada uno de ellos encargado de una tarea específica: Captación y reconocimiento de imágenes digitalizadas (provenientes de varias escenas) y control de un robot manipulador, aparte de su función específica como estación de desarrollo software. En este sentido se ha desarrollado un completo sistema de software propio que incluye las siguientes funciones: captación y presentación de imágenes provenientes de señal de vídeo, almacenado y lectura de disco de imágenes digitales, creación y gestión de una base de datos visual, y toda una gama de procesamiento de imágenes digitales: filtrado, extracción de contornos, realce, codificación, transmisión, restauración, reconocimiento automático de objetos, etc. Se trata, pues, de la posibilidad de compaginar la universalidad de un sistema informático estándar y compatible (en donde la principal labor a desarrollar es de programación) y la eficacia de un hardware específico (en este caso constituido por los productos compatibles, tarjetas y chips de procesamiento especial).

etc., de la imagen y de la aplicación a que esté destinado.

Queda claro que para dotar al sistema de flexibilidad y versatilidad necesaria, sería imprescindible disponer de procesadores de muy alta velocidad de cálculo que implementasen algoritmos software. Sobre dos factores se puede incidir para mejorar la velocidad de procesamiento; un método es aumentar la frecuencia del reloj del sistema mediante el uso de tecnologías avanzadas en circuitos digitales. En la actualidad existen ordenadores como el X-MP de Cray Research con un ciclo de reloj aproximado de 10 nseg, o lo que es lo mismo una frecuencia de 100 MHz, y se espera llegar en el futuro a un ciclo de reloj de 1 nseg que presenta un límite tecnológico difícilmente superable. Otra alternativa es la ejecución simultánea de operaciones mediante un solapamiento temporal, paralelismo temporal, o bien mediante réplica de las unidades de procesamiento, paralelismo espacial.

Hoy día coexisten dos tendencias arquitectónicas en diseño de ordenadores de alto rendimiento. Una adoptada por los diseñadores de supercomputadores, normalmente

los grandes fabricantes, que junto a un incremento de la frecuencia de reloj incorporan un alto grado de paralelismo temporal, estructuras pipeline, e incluso un cierto grado de paralelismo espacial en algunos de sus elementos de proceso. La otra tendencia, nacida en Universidades y Centros de Investigación y que cada vez más empieza a ser tenida en cuenta por las casas comerciales, se basa en la incorporación de un alto grado de paralelismo espacial mediante la utilización masiva de elementos de proceso.

Paralelismo en visión artificial

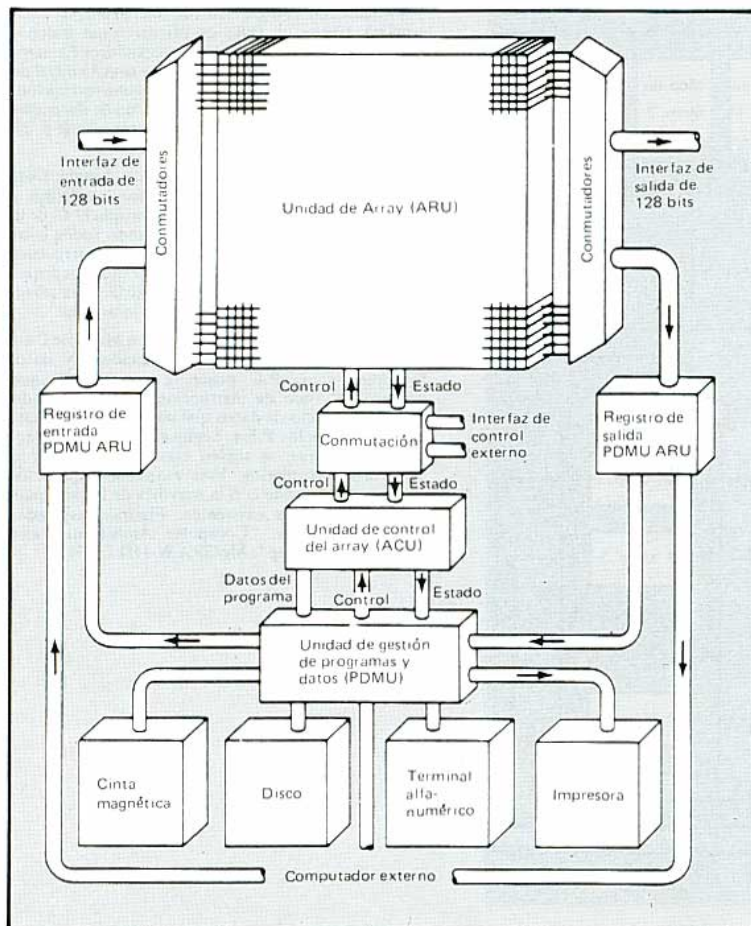
Aunque hoy día no parece probable que las estructuras altamente paralelas lleguen a trabajar con ciclos de tiempo del orden de magnitud de los supercomputadores, éstas son muy apropiadas en S.V.A., debido a la estructura de los datos que se manejan. Téngase en cuenta que una imagen digital está organizada en forma matricial y es susceptible de ser dividida en partes más elementales, incluso a nivel de pixels, para ser procesadas en paralelo.

Se puede citar el MPP (Massively Parallel Processor), construido por Goodyear

Aerospace Co. para la NASA, como ejemplo de ordenador altamente paralelo destinado a procesamiento de imágenes. Este ordenador que posee 16.384 procesadores elementales distribuidos en una matriz de 128 x 128 (cuadro 1), puede alcanzar velocidades de 5 a 10 veces mayores que las de los supercomputadores existentes, en las aplicaciones de tratamiento de imágenes, aunque su frecuencia de reloj es 10 veces más lenta. El diseño del MPP permite aumentar el número de sus procesadores en varios órdenes de magnitud, lo cual lo dotaría de una velocidad enormemente más rápida que la de los supercomputadores convencionales. No obstante quede claro que estos rendimientos se obtienen en aplicaciones muy específicas como el procesamiento de imágenes.

El uso de estructuras paralelas conduce a un replanteo de la estructura de programación convencional. El programador debe analizar el problema bajo una óptica distinta, que le conduzca al diseño de algoritmos que se adaptan óptimamente al nuevo concepto de máquina. Este problema es generalizable a todo el soporte software que gestiona la máquina. Aunque hoy día

CUADRO 1
Un Array Processor para tratamiento de imágenes, el M.P.P.



El M.P.P. (Massively Parallel Processor) fue desarrollado para la NASA Goddard Space Flight

El M.P.P. (Massively Parallel Processor) fue desarrollado para la NASA Goddard Space Flight Center en 1983 por Goodyear Aerospace Co. Con sus 16.384 procesadores representa el primer paso hacia el paralelismo en gran escala necesario en los futuros ordenadores.

Es un procesador de alta velocidad diseñado para resolver problemas de procesamiento de datos bidimensionales como los presentes en procesamiento de imágenes.

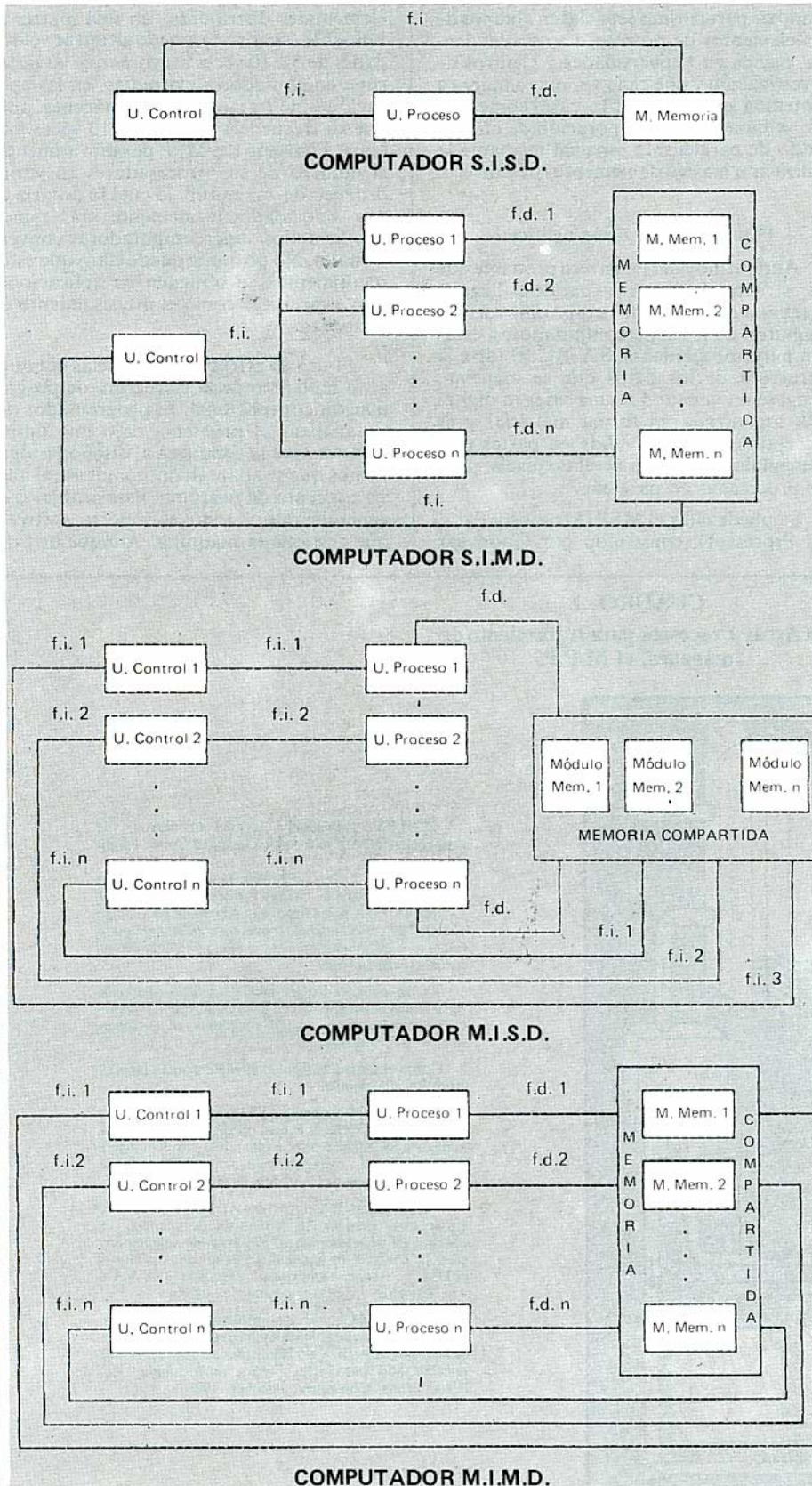
Como muestra la figura, el MPP consta de tres módulos principales:

- La Unidad de Array (ARU) organizada como matriz de 128 x 128 P.E.s de un bit, con un Kbit de memoria local cada uno, conectados con topología de red en malla, capaz de realizar transferencias de datos con el exterior en bloques de 128 bits.

- La Unidad de Control del Array (ACU) que se encarga de realizar las funciones de control del sistema de procesamiento, es capaz de seleccionar bien la Unidad de Manejo de Programas y Datos (PDMU), bien un computador externo (DEC VAX 11/780) para el control global del sistema.

- La P.D.M.U. es un minicomputador PDP 11/34A que se encarga de inicializar, realizar pruebas de diagnóstico, y las funciones de control si está seleccionado para ello. Figura reproducida del IEEE Trans. Computer, Batcher, 1980.

CUADRO 2
Una clasificación de los computadores (Flynn, 1966)



Desde el punto de vista del flujo de control existente en una arquitectura, y más particularmente atendiendo a la multiplicidad de los flujos de instrucciones y datos, Michael J. Flynn en 1966 propuso la clasificación de las arquitecturas de los computadores que se muestra en la figura.

Flujo de instrucciones (f.i.) es una secuencia de instrucciones tal y como se ejecuta en la máquina. Flujo de datos (f.d.) es una secuencia de datos que incluyen los datos de entrada y los resultados parciales o temporales que utiliza el flujo de instrucciones.

Flynn presenta cuatro tipos de máquinas:

- **S.I.S.D.** (Single Instruction Single Data stream), tienen flujos únicos de instrucciones y datos. Este tipo de máquina representa a la mayoría de los computadores secuenciales existentes, pudiendo incorporar algún tipo de paralelismo temporal, en cuyo caso se les denomina Computadores Pipeline.

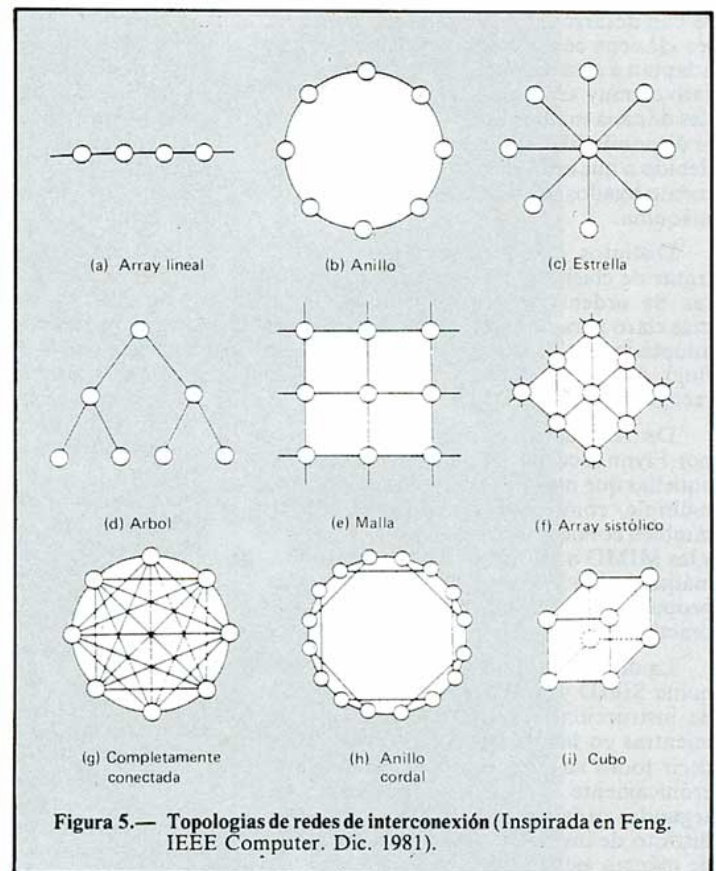
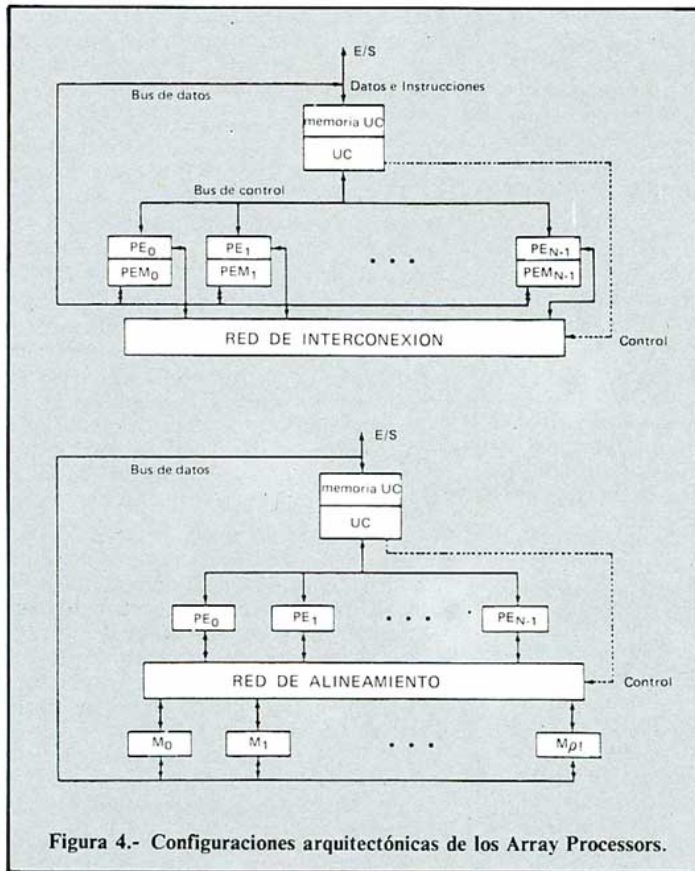
- **S.I.M.D.** (Single Instruction Multiple Data stream), tienen un flujo de instrucciones único y múltiple de datos. Todos los Procesadores Elementales (P.E.s) están gobernados por una Unidad de Control (U.C.) que les difunde la misma instrucción, operando cada P.E. sobre un conjunto de datos distinto. Este tipo de arquitectura corresponde a los Array Processors.

- **M.I.S.D.** (Multiple Instruction Single Data stream), tiene un flujo de instrucciones múltiple y único de datos. Cada P.E. tiene su propia U.C. de la que recibe las instrucciones, operando todos ellos sobre un mismo conjunto de datos. Esta estructura se ha calificado de impracticable por muchos arquitectos de computadores, lo cual podría ser la razón por la que no existe implementación alguna.

- **M.I.M.D.** (Multiple Instruction Multiple Data stream), tienen flujos de instrucciones y datos múltiples. Cada P.E. posee su propia U.C. que gobierna el flujo de instrucciones, interactuando sobre un espacio de datos que puede estar compartido por todos los P.E.s. Según el grado de interacción que posean se suelen clasificar en fuerte o débilmente acoplados. Este esquema arquitectónico se corresponde con la mayoría de los sistemas multiprocesadores existentes. Figuras inspiradas en: K. HWANG, "Computer Architecture and Parallel Processing", McGRAW-HILL 1985.

Por último, es necesario señalar que un Array Processor se conecta a un ordenador central que se encarga de realizar la carga de programas sobre la memoria de la U.C. y de los datos, bien en esta misma, o en las memorias locales de los P.E.s. Un Array Processor no es pues un conjunto autónomo; se utiliza como un coprocesador para la ejecución más óptima posible de aquellos algoritmos que se puedan adaptar a su estructura.





La tabla 1 resume algunas características de varios Array Processors utilizados para procesamiento de imágenes.

memoria local, y una Red de Interconexión encargada de establecer las comunicaciones necesarias entre los P.E.s.

BIBLIOGRAFIA

- A.G. BATCHELOR, D.A. HILL and D.C. HODGSON. "Automated Visual Inspection". I.F.S. Ltd. 1985.
- M.J.B. DUFF and S. LEVIALDI. "Languages and Architectures for Image Processing". Academic-Press. 1981.
- RAFAEL C. GONZALEZ and PAUL WINTZ. "Digital Image Processing". Addison-Wesley Publishing Co. 1977.
- ERNEST L. HALL. "Computer Image Processing and Recognition". Academic Press 1979.
- KAI HWANG and FAYE A. BRIGGS. "Computer Architecture and Parallel Processing". Mc. Graw-Hill. 1985.
- J.K. ILIFFE. "Advanced Computer Design". Prentice-Hall. 1982.
- JOSEF KITTLER and MICHAEL J. DUFF. "Image Processing Systems Architectures". Research Studies Press. 1985.
- J.L. POTTER. "The Massively Parallel Processor". The M.I.T. Press.



Mateo Luque Sánchez, Ingeniero de Telecomunicación, es responsable del mantenimiento de la instalación informática y forma parte del equipo interdisciplinario de investigación en el proyecto "VISION ARTIFICIAL Y VISION HUMANA. APLICACIONES A LA PERCEPCION VISUAL EN ROBOTICA", financiado por la Fundación Ramón Areces, en colaboración con la U.P.M. y la U.C.M.



José P. Rueda Cruz, Ingeniero de Telecomunicación, ha trabajado varios años en la empresa privada, actualmente es profesor de la U.P.M. y participa en el mismo proyecto.