

# *Dos métodos de emparejamiento automático de imágenes en visión estereoscópica\**

Por D. MARAVALL GÓMEZ-ALLENDE\*, E. FERNÁNDEZ LÓPEZ\*\*

\* *Catedrático de la Facultad de Informática de Madrid*

\*\* *Profesor Titular de la E.U.I.T. de Telecomunicación de Madrid*

## **Resumen**

Los autores presentan algunos resultados obtenidos en visión artificial. En primer lugar, se describe un procedimiento de emparejamiento de pares estereoscópicos de imágenes digitales, en el que se busca ante todo la rapidez de cálculo. A continuación, se presentan las líneas generales, incluyendo escenas con objetos solapados. Los autores trabajan con el Grupo de Control Avanzado y Visión Artificial de la Universidad Politécnica de Madrid.

## **Abstract**

The authors present some results in computer vision. In the first place, a matching algorithm for stereo digital images in which speed of computation is the main objective is presented. Afterwards, the general characteristics of a hierarchical algorithm for matching of stereo images operating in very general conditions (including scenes with overlapped objects) is described. Both authors belong to Advanced Control and Artificial Vision Group (GCAVA in Spanish) at the U.P.M.

## **INTRODUCCION. CALCULO DEL RANGO MEDIANTE EMPAREJAMIENTO DE LOS CENTROIDES**

La reconstrucción tridimensional (3D) de las escenas digitalizadas y procesadas en ordenador constituye uno de los campos de mayor interés en las actuales investigaciones y desarrollos en el campo de la denominada visión artificial (VA) basada en computador. Una problemática específica de la reconstrucción 3D de escenas es la determinación del rango de los objetos presentes (es decir, su distancia respecto a un origen de coordenadas arbitrario y conocido). El cálculo del rango es de enorme interés práctico en robótica industrial, al permitir un salto cualitativo en la automatización de tareas gracias a la independencia respecto a las distancias y a las posiciones de los objetos (piénsese en la flexibilidad adicional de un sistema de visión 2D con determinación de la distancia de los objetos presentes).

La visión estereoscópica constituye el método general más empleado en la reconstrucción 3D de escenas digitalizadas. La idea consiste en emparejar los puntos comunes de dos imágenes de la misma escena, pero tomadas desde perspectivas diferentes (par estereoscópico). El problema con este método, sin embargo, es la tremenda carga computacional, donde el tiempo de procesamiento depende de la arquitectura utilizada.

\* Presentada en la Sesión Científica del día 10 de mayo de 1989.

En esta primera parte de la comunicación se propone un procedimiento original para la determinación del rango (concretamente, la distancia de los objetos al centro focal de la cámara), cuya característica más notable es la velocidad de cálculo: de hecho, no supone más que una mínima carga computacional adicional al proceso de extracción de características y reconocimiento de los objetos presentes en la escena.

La idea básica, que consideramos extraordinariamente simple y clara, es realizar una triangulación con los dos centroides de los objetos comunes de un par estereoscópico, según se describe en la figura 1 adjunta.

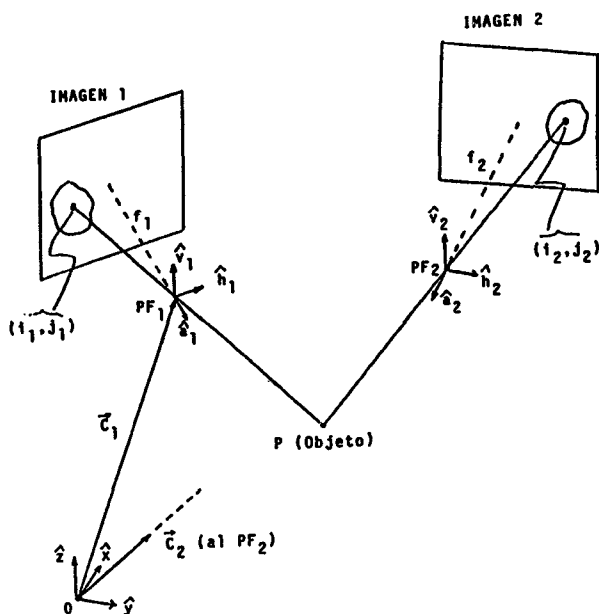


Fig. 1. Fundamentos de la visión estereoscópica.

En [1] se demuestra, siempre en relación a la figura 1 adjunta, que la determinación del punto P en tres dimensiones viene dada por la ecuación vectorial:

$$\vec{C}_1 + r_1 \hat{d}_1 = \vec{C}_2 + r_2 \hat{d}_2$$

En donde  $\vec{C}_1$  y  $\vec{C}_2$  se calculan mediante la calibración de las cámaras. En cuanto a la determinación de los vectores  $\hat{d}_1$  y  $\hat{d}_2$ , puede consultarse la referencia [1], páginas 17 a 20.

Este procedimiento es muy atractivo para su aplicación en robótica industrial por la sencillez del cálculo que implica y que se resume en las siguientes operaciones:

1. Segmentación de la escena, con el objetivo de obtener los contornos cerrados de los objetos presentes; para cada imagen del par estereoscópico.
2. Extracción de rasgos y clasificación automática de los objetos (para cada par estereoscópico).

3. Determinación, en cada par, de los centroides de los objetos.
4. Cálculo del rango, según las expresiones matemáticas obtenidas en [1].

Se describe a continuación otro procedimiento de cálculo del emparejamiento, aplicado a escenas complejas (incluyen objetos solapados), donde los procesos que intervienen son más complicados; por tanto, los tiempos de ejecución son también mayores.

## CODIFICACION DE SEGMENTOS DE BORDE

La primera operación que se realiza con cada una de las imágenes del par estereoscópico es la extracción de bordes. El tipo de detector utilizado en este trabajo se debe a Marr y Hildreth [2]; se emplea un operador gradiente a la función de intensidades de los puntos de la imagen, aunque se somete previamente la imagen a un suavizamiento, mediante un filtrado paso bajo de tipo Gaussiano. Una vez aplicado dicho operador habrá que detectar qué puntos son cruces por cero, además del valor de la pendiente en esos puntos.

Se somete la imagen de bordes a un adelgazamiento con el fin de poder realizar un seguimiento de bordes. El método de seguimiento empleado es el código cadena; intervienen dos fases, una de búsqueda de puntos borde y otra de seguimiento para extracción del código cadena. Se emplea un código cadena de ocho direcciones en el seguimiento del borde, en caso de que se detecte un punto de ramificación se emplea automáticamente un código de cuatro direcciones, con el fin de que no haya puntos que se tomen más de una vez.

Se emplea una codificación de segmentos de borde, basada en una estructura de grafos [4], como paso previo a la extracción de características. Se entiende aquí por segmento de borde cualquier borde abierto, cerrado, o que sale de un punto de ramificación. Cada nodo del grafo representa un segmento de borde, mientras que los enlaces entre dos nodos se refieren al caso de que esos segmentos salgan o entren, según el sentido del seguimiento, del punto de ramificación considerado.

Una vez codificados los segmentos de borde, por la flexibilidad que permite la propia estructura de datos, se eliminan aquellos bordes de pequeña longitud. Esto se hace con la finalidad de ahorro computacional, y porque en la mayoría de los casos es un número pequeño de características las que se extraen; supone entonces una reducción de la ambigüedad la eliminación de estos segmentos.

## EXTRACCION DE CARACTERISTICAS

Extraído el código cadena de cada segmento, podemos realizar una aproximación lineal de cada segmento. El procedimiento empleado consiste en reconstruir las coordenadas cartesianas en el plano digital de los puntos borde hasta encontrar intersección con línea epipolar, se realiza una aproximación del borde por un segmento de línea que pase por dicho punto de intersección.

El método de aproximación empleado aquí es el de línea autovector [3]. Si la distancia media de los puntos borde a la línea de aproximación es mayor que un umbral constante, se hace la aproximación para un conjunto distinto de puntos, hasta conseguir que sea menor que el umbral. El conjunto de puntos se desplaza a lo

largo del segmento, incluyendo siempre el punto borde de intersección. Se toma al principio un número grande de puntos, reduciendo dicho número hasta encontrar el segmento de línea buscado.

Las características que se extraen a partir de los segmentos lineales son de orientación, longitud y pendiente media de los puntos borde en los que se ha realizado la aproximación. Estas características se codifican en listas encadenadas, con el fin de facilitar los cálculos en la fase de emparejamiento de imágenes.

## EMPAREJAMIENTO DE IMAGENES

Se emplea un algoritmo jerárquico [4] para el emparejamiento de imágenes; en el primer nivel se obtiene el emparejamiento entre puntos borde de cada segmento de una imagen con cada uno de los segmentos de la otra imagen, siempre que al menos exista una línea epipolar común que intersecte a los dos segmentos emparejados. Los resultados del etiquetado en el primer nivel van a influir en los valores de las probabilidades iniciales y de las restricciones en el segundo nivel.

Se emplea un etiquetado relajado basado en autómatas de aprendizaje [5] para el cálculo de la correspondencia en cada nivel, entendiendo por objetos puntos o segmentos de una imagen, y etiquetas esos mismos elementos en la otra imagen. Los valores de las probabilidades iniciales en los dos niveles dependen de la comparación de características de puntos borde de ambas imágenes. Los puntos que se comparan pertenecen a la misma línea epipolar o líneas adyacentes.

Las restricciones del etiquetado se basan en la similitud de la relación de tipo espacial entre puntos o segmentos de borde en cada uno de los planos de imagen. Este tipo de relación varía de una imagen estereoscópica a otra por el desplazamiento de las cámaras; sin embargo, se tolera cierta componente de ruido en los valores de las restricciones en estos procesos, debido a la aleatoriedad de los mismos.

El problema principal del emparejamiento en estereopsis es el de resolver la ambigüedad, es decir, cuando hay que discernir entre varios candidatos con características semejantes. Se ha experimentado con escenas industriales que incluyen objetos de formas idénticas, para comprobar el funcionamiento de los algoritmos ante problemas de ambigüedad. Se han obtenido resultados satisfactorios, los pocos errores aparecidos se deben a rupturas de borde en la fase de extracción de bordes, que hacen variar las características de una imagen a otra.

## BIBLIOGRAFIA

- [1] MARAVALL, D.: *Una contribución al Cálculo del Rango en Visión Artificial*, Trabajo de Investigación, Cátedra de Percepción Computacional de la Facultad de Informática de Madrid, Madrid 1988.
- [2] MARR, D. y HILDRETH, E. C.: *Theory of edge detection*, Proc. R. Soc. London B., vol. 207, pp. 187-217, 1980.
- [3] DUDA, D., y HART, P. E.: *Pattern classification and scene analysis*, Wiley, N. Y., 1973.
- [4] FERNÁNDEZ, E.: *Resolución de la oclusión y de la correspondencia en estereopsis, mediante un etiquetado de imágenes*, Tesis Doctoral, E.T.S.I. de Telecomunicación de Madrid, 1989.
- [5] THATHACHAR, M. A. L., y SASTRY, P. S.: *Relaxation labeling with learning automata*, IEEE Trans. on Pattern Anal. Mach. Intell., vol. PAMI-8, pp. 256-268, Mar. 1986.