

ANÁLISIS DE ALGORITMOS EN VISIÓN TRIDIMENSIONAL HUMANA

P. RUBIO DE LEMUS

Universidad Nacional de Educación a Distancia (UNED)

Resumen

La visión en estéreo consiste en calcular la distancia que existe entre los puntos de borde de las imágenes y el plano de las cámaras. En este estudio se va a comparar el rendimiento de dos algoritmos de estereopsis basados en el sistema visual humano. El algoritmo de Kim y Aggarwal (1987) encuentra todos los puntos candidatos de la imagen derecha para cada punto de la izquierda, y asigna una probabilidad inicial a cada candidato en función de la similitud de los patrones de ceros y de la diferencia entre los gradientes de intensidad. Luego aplica un proceso de relajación para determinar los mejores emparejamientos basados en una serie de restricciones de continuidad. El algoritmo PMF (1985) tiene una estructura similar al anterior excepto en que no utiliza probabilidades, sino que se basa en el límite del gradiente de disparidad. Ambos algoritmos se han probado con imágenes geométricas y naturales, utilizando medidas cualitativas y cuantitativas para comparar sus rendimientos.

Palabras clave: Estereopsis, visión tridimensional, visión artificial, búsqueda de correspondencias.

Abstract

Stereo vision consists of calculating the distances from each point in the objects to the plane of the two cameras. In the present study we compare the performance between two stereo algorithms based on aspects of the human visual system. Kim and Aggarwal's algorithm (1987) finds all the candidate points in the right image for each zero-crossing point in the left one and assigns initial weight to every candidate point according to the similarity of zero-crossing patterns, and the difference in intensity gradients. After this, they apply a relaxation process to find the best matches which is based on several continuity constraints. PMF algorithm (1985) has a similar structure to the other model, excepting it does not use probabilities but it is based on the disparity gradient limit. Both algorithms have been tested with geometrical and natural images, using quantitative and qualitative measures to compare their performances.

Key words: Stereopsis, three-dimensional vision, artificial vision, matching.

Introducción

Uno de los objetivos fundamentales de la investigación sobre visión por ordenador es la configuración de un sistema de visión que simule los procesos del sistema visual humano (SVH) y que sea eficiente en su funcionamiento. Un primer análisis del SVH revela que éste, desde que adquiere la imagen hasta que hace uso de una representación o interpretación de la misma, realiza una serie de operaciones encaminadas a ir obteniendo la información más relevante. Dentro de este sistema, la visión estereoscópica formaría parte del procesamiento primario, etapa en la que al final se obtiene una descripción

de las distancias y orientaciones de las superficies y en donde se combina la información obtenida por los diversos módulos de visión estereoscópica, movimiento, color, textura, etc.

La autora de esta investigación lleva varios años estudiando modelos de visión estereoscópica basados en aspectos del sistema visual humano. La visión en estéreo es un método importante para estimar la información sobre la profundidad de una escena, y consiste en calcular la distancia que existe entre los cambios de intensidad luminosa de las imágenes y las cámaras. Esto se consigue partiendo de la información inicial de un par de imágenes captadas por dos cámaras cuyos ejes focales son aproximada-

mente paralelos y utilizando programas de ordenador que simulan la visión. El problema más importante a resolver en estos sistemas es el de encontrar el conjunto de rasgos de una imagen que se correspondan con los de la otra; éste es el denominado problema de la correspondencia. En este estudio se va a comparar el rendimiento de dos algoritmos de estereopsis basados en el sistema visual humano, los cuales utilizan los puntos de cambio de luminancia de las imágenes filtradas o puntos de borde (*zero-crossings*) como rasgos a poner en correspondencia. Se trata del modelo de Kim y Aggarwal de 1987 y del algoritmo PMF de Pollard, Mayhew y Frisby de 1985. Tras analizar ambas teorías y las diferencias más significativas entre ellas, se comparará el rendimiento de ambos algoritmos con imágenes geométricas y naturales, utilizando medidas cualitativas y cuantitativas. Entre las medidas cuantitativas estarán: el tiempo de procesamiento, el número de emparejamientos obtenidos y los puntos que han quedado sin emparejar. Una de las medidas cualitativas será el número de emparejamientos correctos y falsos.

Modelo de Kim y Aggarwal (1987)

Este modelo se implementó siguiendo las descripciones teóricas del mismo (Kim y Aggarwal, 1987). En líneas generales, el modelo se puede encuadrar dentro de los métodos pasivos (que emplean la luz natural para su solución), basados en aspectos fisiológicos y psicológicos, cuyo representante más característico es el modelo de Marr y Poggio de 1979 [modelo también simulado en ordenador por la autora de este artículo (Rubio de Lemus, 1986; Rubio de Lemus y Pérez, 1987)]. El planteamiento del problema por parte de Kim y Aggarwal es similar al método de Marr y Poggio: la obtención de los datos de profundidad de una escena se efectúa a partir de la búsqueda de elementos correspondientes entre las dos imágenes captadas. Al igual que Marr y Poggio, en este modelo se emplean los ceros de la segunda derivada (o puntos de borde) como elementos de la imagen o primitivas sobre las que efectuar el proceso de búsqueda de correspondencias, y además, también se plantean y utilizan restricciones que orientan adecuadamente el funcionamiento de dicho proceso. No obstante, existe una diferencia importante entre ambos modelos: Kim y Aggarwal, cuando comparan dos puntos de borde para establecer su posible correspondencia, no sólo se basan en el gradiente de los mismos, sino que además usan una medida de la similitud entre los alrededores de ambos puntos. La justificación para emplear un mecanismo de este tipo es que está demostrado que nuestra retina transmite las imágenes al cerebro a través de la excitación de grupos o patrones de células visuales corticales. Así, determinando todas las posibles agrupaciones de puntos en torno a uno dado, Kim y Aggarwal etiquetan cada punto de las imágenes con un nuevo valor, con el objeto de facilitar el proceso de búsqueda de correspondencias.

Tanto en el modelo de Marr y Poggio como en el de Kim y Aggarwal, y para obtener finalmente la imagen tridimensional, se han de seguir una serie de procesos previos de pretratamiento de las imágenes, cuya función es la de transformar las imágenes originales hasta que se dispone de la información necesaria para realizar propiamente el proceso de estereopsis. Los pasos más importantes que se siguen en la simulación del modelo de Kim y Aggarwal, y que reflejan procesos que realiza el SVH son los siguientes:

1. El *filtrado de las imágenes*, según Campbell y Robson (1968), es una de las primeras operaciones que el SVH lleva a cabo sobre la escena visual. Este filtrado lo realizan una serie de filtros llamados canales, que son casi lineales y selectivos tanto en frecuencia espacial como en orientación. Aquí, el filtrado de las imágenes izquierda y derecha se ha llevado a cabo en el dominio del espacio mediante la convolución de las imágenes con un filtro gaussiano bidimensional.

2. La *detección de bordes* se realiza sobre las imágenes filtradas. En este caso, se han seguido las teorías de Marr y Hildreth (1980), que propusieron emplear el operador laplaciano para la localización de los puntos de cambio de luminancia. En este programa también se calculan algunas características de estos puntos, las cuales servirán posteriormente para designar los elementos correspondientes u homólogos. Estas características de los gradientes de intensidad son: módulo (o pendiente), signo y dirección. Kim y Aggarwal sólo utilizarán posteriormente el módulo del gradiente.

3. La *búsqueda de correspondencias o emparejamientos (matching)*, consiste en hacer corresponder en las dos imágenes los pares de puntos que sean proyección de la misma entidad física de la escena. Una vez establecidos los elementos de la imagen sobre los que se va a efectuar el proceso y las medidas de comparación para determinar si puede existir o no correspondencia entre dos puntos, Kim y Aggarwal introducen un proceso iterativo en el que, básicamente, la utilización de varias restricciones y un esquema de asignación de probabilidades permite obtener una solución final. Este proceso tiene dos partes:

- a) *Asignación inicial de probabilidades*. A cada punto de una imagen se le buscan todos los posibles correspondientes en la otra, buscando en horizontal dentro de un rango que está en función de las características técnicas de las cámaras, de su separación y de la mínima distancia posible de los objetos a las mismas. A cada punto se le asigna una probabilidad en función de la similitud entre los gradientes y entre los contornos. Este cálculo se realiza a través de una función de peso de ambos factores, pudiendo variar en cada proceso la influencia de cada una en la asignación de la probabilidad ($a = 0,5$ y $b = 0,5$ para el procesamiento de las imágenes de este trabajo, pesos que demostraron ser

los más eficaces en Rubio de Lemus y Pérez, 1990).

b) *Proceso iterativo de cálculo de nuevas probabilidades*. Casi todos los modelos de estereopsis reconocen que, sin establecer algún tipo de restricción, la búsqueda de correspondencias tiene una solución muy difícil. En general, casi todos los investigadores han partido de observaciones sobre los objetos en el mundo real que luego han aplicado sobre los modelos. Kim y Aggarwal proponen un enunciado general que luego se concreta en tres restricciones: un punto de una imagen no puede hacerse corresponder en solitario, hay que considerar cómo se empareja su entorno, lo que implica: 1) continuidad de las disparidades, ya que en general las superficies de los objetos son continuas en pequeños contornos; 2) continuidad en los bordes, lo que obliga a que los ceros estén conectados de forma continua, y 3) continuidad de las probabilidades de correspondencia (esto se deduce a partir de los dos puntos anteriores).

De este modo, en el proceso iterativo, cada probabilidad se vuelve a calcular con una fórmula en la que, básicamente, influyen:

- La probabilidad en la iteración anterior.
- Las probabilidades de los puntos conectados de emparejarse con la misma disparidad.
- Las probabilidades de los puntos conectados de emparejarse con disparidad parecida.

La velocidad de la convergencia del proceso se controla a través de dos constantes, «c» y «d», en la fórmula de iteración (en este trabajo, las figuras se han procesado con $c = 3,0$ y $d = 0,7$, valores que arrojaron el mayor número de emparejamientos en Rubio de Lemus y Pérez, 1990). Una vez recalculadas y normalizadas todas las probabilidades, se eliminan las que no superen un determinado valor y se consideran como correspondencias ciertas las que estén por encima de un umbral concreto. El proceso termina cuando se ha dado un determinado número de iteraciones o cuando se han encontrado suficientes emparejamientos.

4. *La estimación de las distancias* o cálculo de la tercera dimensión se lleva a cabo a partir del mapa de disparidades y se basa en una sencilla argumentación geométrica. Una vez calculadas las distancias de cada uno de los puntos de borde al plano de las cámaras, éstas se pueden representar en forma tridimensional mediante cualquier tipo de utilidad gráfica.

Algoritmo PMF de Pollard, Mayhew y Frisby (1985)

La autora de este trabajo implementó el algoritmo PMF siguiendo las ideas apuntadas por Pollard, Mayhew y Frisby (1985), y por Pollard, Porrill, May-

hew y Frisby (1985). El esquema utilizado por sus creadores es semejante al utilizado por Marr y Poggio en 1979: la solución al problema de la correspondencia consiste en encontrar las restricciones que permitan resolver las ambigüedades que tienen lugar en este proceso. Estas restricciones se derivan principalmente de las propiedades de las superficies de los objetos y/o a través de las características del proceso de formación de las imágenes. La gran diferencia entre los modelos de Pollard, Mayhew y Frisby, y el de Marr y Poggio, es que aquéllos consideran que la estereopsis se produce en una etapa primaria sin necesidad de recurrir a conocimientos o representaciones de alto nivel (experimentos con estereogramas), al contrario que Marr y Poggio. Además introduce una nueva restricción basada en evidencias experimentales de Burt y Julesz (1980) sobre la visión binocular humana. Según muestran estos resultados psicofísicos, el límite del gradiente de disparidad entre emparejamientos correctos será de 1 para la mayor parte de las superficies, mientras que entre emparejamientos incorrectos no será de 1. El gradiente de disparidad entre dos puntos de una imagen es la diferencia de sus disparidades (respecto a sus puntos correspondientes) dividida entre su separación ciclópea. Esta última viene dada por la distancia entre los puntos medios de los dos pares de puntos. Además de la restricción del gradiente de disparidad, el PMF utiliza otras dos restricciones que guían el proceso: la restricción de epipolaridad y la de unicidad. La primera se deriva de la posición geométrica de las cámaras: «siempre que los ejes ópticos de las dos cámaras estén en el mismo plano, todas las primitivas de *matching* se situarán sobre pares izquierdo/derecho de líneas rectas epipolares» (Baker, 1982). La restricción de unicidad impone un límite al número de emparejamientos permitidos para una única entidad de la imagen, es decir, cada primitiva de la imagen participará sólo en una correspondencia. Esta restricción viene del hecho de que los objetos reales raramente proyectan un único punto en uno de los planos de la imagen (izquierdo o derecho), si tiene proyecciones cerca de otro con las dos primitivas correspondientes de la imagen. Sólo en este caso improbable sería legítimo que una primitiva de una imagen se correspondiera con dos primitivas de la otra. Además, esta situación incumpliría la restricción del límite del gradiente de disparidad de 1, la cual se impuso como punto de partida para la mayor parte de los objetos reales. El PMF se divide en tres etapas de procesamiento:

1. *Selección y extracción de primitivas* para realizar el proceso de emparejamiento. Siguiendo las descripciones teóricas de los creadores del PMF y en consistencia con otros modelos implementados por los autores de este trabajo, para las imágenes naturales se eligieron como primitivas los puntos de cambio de luminancia (PCL) que fueran aproximadamente equivalentes a los bordes físicos de los objetos. Se utilizaron los mismos módulos de preprocesamiento que en el caso del modelo de Kim y Aggarwal (véan-

se pasos 1 y 2), y que también se emplearon en la simulación del modelo de Marr y Poggio, es decir, se han extraído los PCL con el laplaciano de la gaussiana, utilizando un sólo tamaño de filtro.

2. *Búsqueda de emparejamientos potenciales.* Basándose en la restricción epipolar y debido a que ambos ejes ópticos son aproximadamente paralelos, la búsqueda de una primitiva situada en el plano izquierdo se realiza en la misma línea horizontal (la correspondiente línea epipolar) del plano derecho, y en un rango cuya amplitud depende de la separación de las cámaras y del tamaño del filtro utilizado en el preproceso. Sólo los puntos del mismo signo y aproximadamente la misma orientación se consideran emparejamientos potenciales. En esta etapa se crea una lista para cada punto de la imagen estéreo de la izquierda con todos sus posibles emparejamientos de la derecha, y se calcula el grado de bondad (corrección) de la correspondencia en función de los dos parámetros antes mencionados (signo y orientación). Además, se hace una lista para cada punto de la imagen derecha con todos sus candidatos de la izquierda y el grado de bondad de la correspondencia.

3. *Selección de los emparejamientos correctos.* Esta etapa tiene dos pasos: en el primero, se calcula lo que los creadores del PMF denominan «la fuerza de la correspondencia» de cada emparejamiento potencial, la cual se estima a partir de la suma de las contribuciones que recibe de todos los emparejamientos potenciales de su alrededor que satisfacen el gradiente de disparidad de 1 con respecto a él. Se utiliza un área definida por un círculo de radio igual a 7 píxeles o elementos de la imagen. En el segundo paso es donde realmente se seleccionan los emparejamientos correctos. Para hacerlo, se emplea un proceso iterativo en el que para cada iteración se eligen como «correctos» aquellos emparejamientos que tengan la fuerza de la correspondencia más elevada para las dos imágenes primitivas que les formaron, y todas las demás correspondencias asociadas con las dos primitivas que componían cada correspondencia elegida se eliminan sin más consideración de acuerdo con la restricción de unicidad.

En este último paso se introdujo la única variación con respecto al modelo original: para que ocurra una correspondencia, la optimización debe provenir de ambos lados, es decir, un punto se emparejará con otro sólo si es su mejor candidato (esto es, si tiene la fuerza de correspondencia más elevada), y si además éste no tiene otro mejor candidato. Con este mecanismo se pretenden eliminar los posibles errores que podrían ocurrir por el orden seguido en el proceso de emparejamiento, es decir, yendo de los puntos que están a la izquierda hacia los puntos que están a la derecha, teniendo en cuenta en cada emparejamiento no sólo los puntos ya procesados, sino también la influencia de los puntos que quedan aún por emparejar.

Finalmente, *el cálculo de las distancias y la representación tridimensional se llevan a cabo de la misma forma que en el modelo de Kim y Aggarwal.*

Descripción de figuras

Se han probado los dos algoritmos con imágenes sencillas compuestas por cilindros, cajas y objetos comunes. Véanse las figuras 1, 2 y 3. En estas figuras, las imágenes superiores (a) son las imágenes originales captadas por las cámaras (izquierda y derecha, respectivamente). El par de imágenes originales de la figura 1 consiste en dos prismas, uno es un prisma cuadrado rectangular dispuesto verticalmente y el otro es un prisma octogonal situado transversalmente en el suelo; en la figura 2 hay un cilindro sobre un prisma cuadrado rectangular situado de forma vertical y un pequeño trozo de papel cuadrado sobre el suelo de la escena; finalmente, la figura 3 es una tetera china. Bajo las imágenes originales (a) se encuentran las imágenes que contienen los puntos de cambio de luminancia de las dos imágenes (izquierda y derecha, respectivamente) (b). En (c), las imágenes representan las distancias de cada punto al plano de las cámaras mediante niveles de gris, donde a menor brillo, mayor es la distancia de los puntos a las cámaras y viceversa. Las imágenes de la izquierda son las imágenes resultantes del modelo de Kim y Aggarwal, y las de la derecha son las resultantes del algoritmo PMF.

Resultados comparativos

El estudio comparativo entre el modelo de Kim y Aggarwal y el PMF se ha llevado a cabo con el propósito de medir de una forma global el rendimiento de cada algoritmo. Para conocer este rendimiento se han tenido en cuenta aspectos tan importantes como los tiempos de procesamiento de cada par de imágenes en estéreo, el número de emparejamientos obtenidos y el análisis de la calidad de estas correspondencias (su corrección). Todos estos aspectos ayudarán a evaluar la implementación real de los algoritmos en sistemas de visión.

La implementación de ambos modelos se ha hecho sobre un PC-XT trabajando a 10 Mhz y equipado con coprocesador matemático. El lenguaje de programación empleado ha sido el TURBO-PASCAL, herramienta que ha permitido utilizar salidas gráficas a pantalla para controlar y comprobar el procesamiento de los diferentes programas. En cuanto al tipo de programación la norma ha sido el tratar de modularizar al máximo, desarrollando cada función en un solo módulo o rutina. Esto facilita la labor de estudio, ya que el análisis en profundidad de cada módulo conduce a una comprensión completa de todo el modelo. Otro factor importante en relación a la programación es que se ha intentado que la implementación de los dos algoritmos sea lo más parecida posible, utilizando las mismas estructuras de datos y de procesos. Esto permite que la comparación se centre en los aspectos funcionales de los modelos, dejando de lado los aspectos técnicos de la implementación.

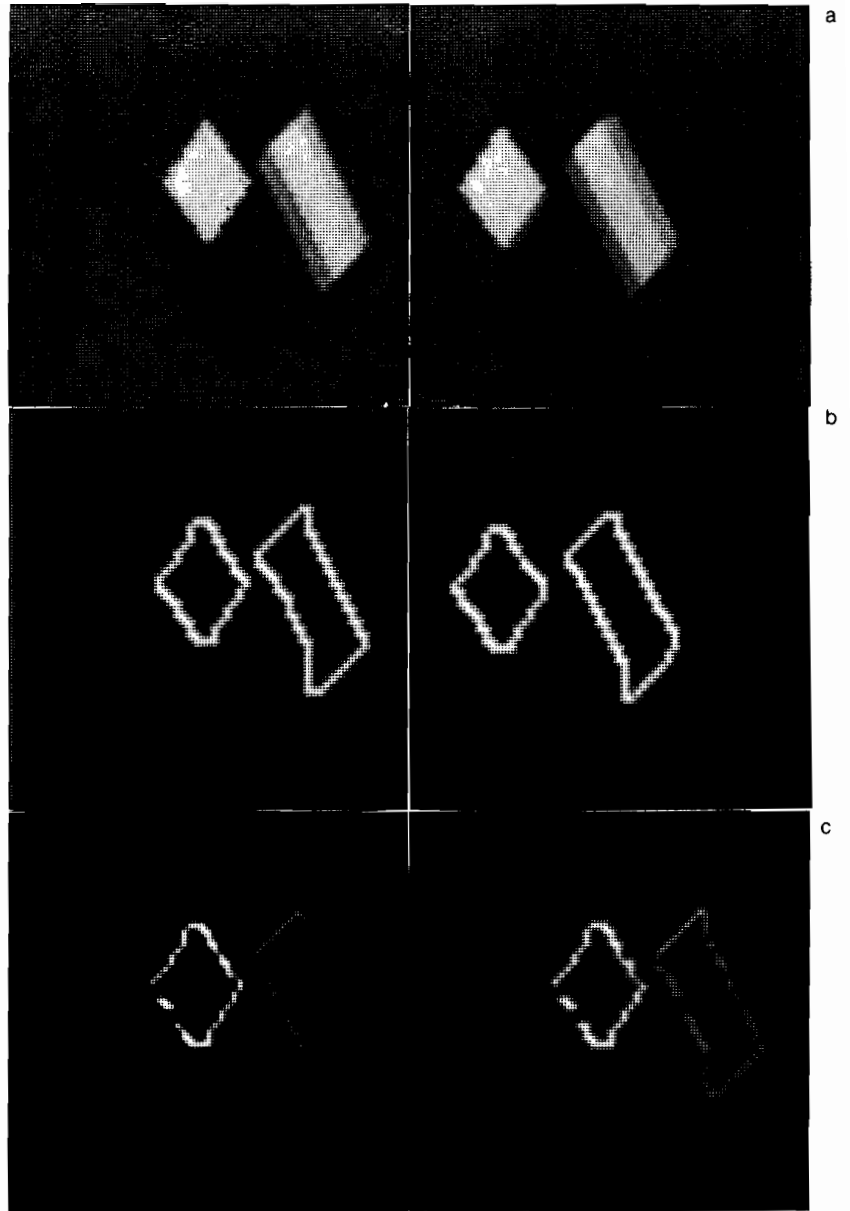


Figura 1. Procesamiento de un par estéreo de imágenes naturales. a) Par estéreo (izquierdo y derecho, respectivamente) tomado de una escena que consistía en dos prismas, uno es un prisma cuadrado rectangular dispuesto verticalmente y el otro es un prisma octogonal situado transversalmente en el suelo. b) Imágenes que contienen los puntos de cambio de luminancia. c) Emparejamiento encontrados con el modelo de Kim y Aggarwal (izquierda) y con el PMF (derecha), utilizando la intensidad de los niveles de gris para representar las distancias de los puntos a las cámaras (oscuro = lejos; claro = cerca).

Tiempo de procesamiento

El primer punto importante para considerar el rendimiento son los tiempos de proceso de cada algoritmo. En este caso, dado el tipo de máquina empleada, los valores obtenidos sólo sirven de referencia para establecer cuál de los dos es más rápido y qué fases se podrían mejorar. Tal como se ve en las figuras 4, 5 y 6, para este análisis se han dividido los algoritmos en dos fases, correspondientes en el caso de Kim y Aggarwal al cálculo de las probabilidades iniciales y al proceso de relajación, y en el

caso del PMF, correspondientes a la creación de las listas de candidatos y al proceso de búsqueda del mejor candidato.

El examen de los resultados confirma dos previsiones iniciales. La primera de ellas era que en el primer paso sería más lento el PMF, dado que la construcción de la lista de candidatos, implementada con asignaciones dinámicas de memoria, requiere más tiempo de proceso que el cálculo inicial de probabilidades soportado por el coprocesador matemático. Y la segunda, y más importante, era que globalmente, el modelo de Kim y Aggarwal sería más lento,

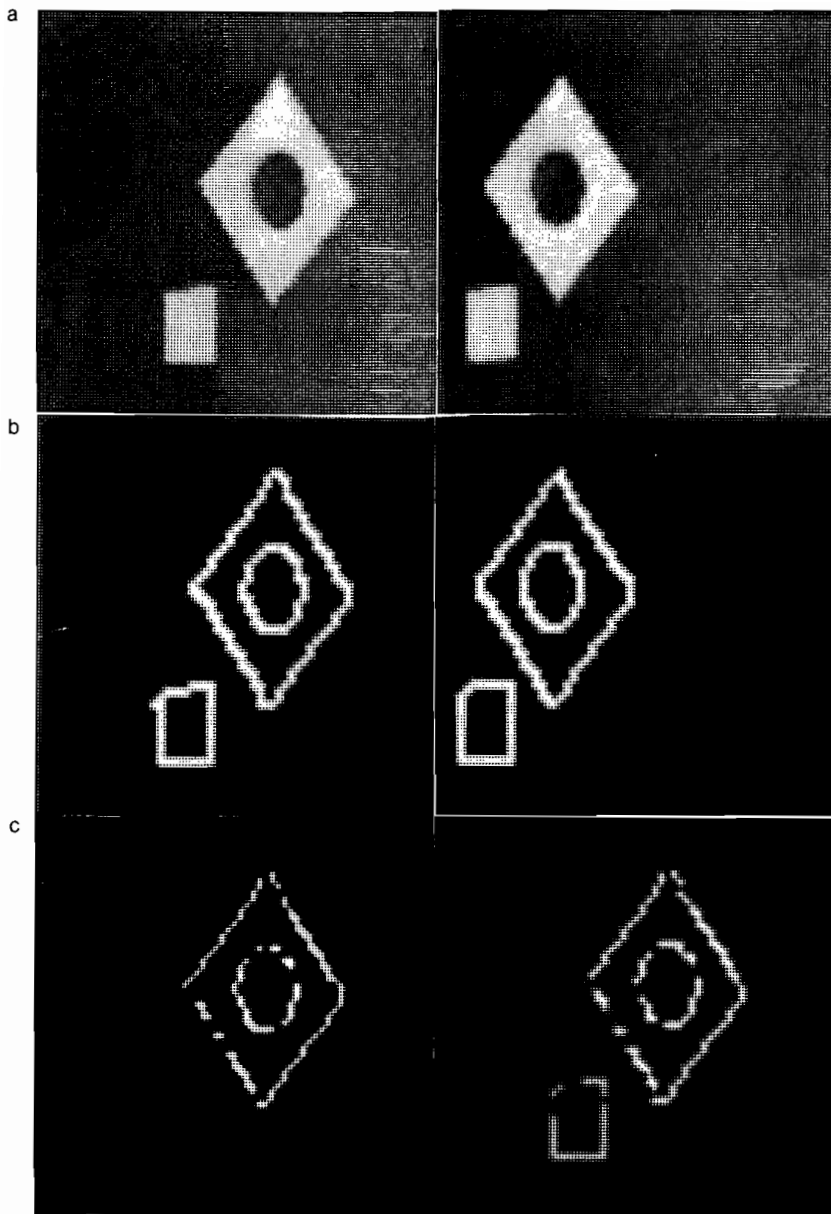


Figura 2. Otro ejemplo del procesamiento de un par estéreo de imágenes naturales. En esta ocasión, la escena consiste en un cilindro sobre un prisma cuadrado rectangular situado de forma vertical y un pequeño trozo de papel cuadrado sobre el suelo de la escena. (Véanse detalles en la figura 1.)

dada la complejidad del proceso de relajación y el número de iteraciones necesarias para encontrar la solución. Este punto es quizá el más reseñable, dado que los tiempos se obtuvieron realizando 10 iteraciones, límite impuesto ya que hasta el momento es el que garantiza completar mejor el proceso de correspondencias.

Número y bondad de los emparejamientos

Quizá el factor más importante a la hora de considerar el rendimiento de estos modelos es el número y la

bondad de los emparejamientos que encuentran. Se ha añadido una nueva medida sobre el comportamiento de los algoritmos: los emparejamientos que no puede efectuar. ¿Por qué esta medida? Porque al considerar la experiencia en la captación y el procesamiento de imágenes se ha encontrado que en la mayor parte de los casos existen discrepancias entre los dos miembros del par estéreo, bien debidos a defectos de las cámaras bien debidos a fallos en los preprocesos, o incluso por la propia configuración de los objetos de las escenas captadas. En cualquier caso, siempre hay un conjunto de puntos que no de-

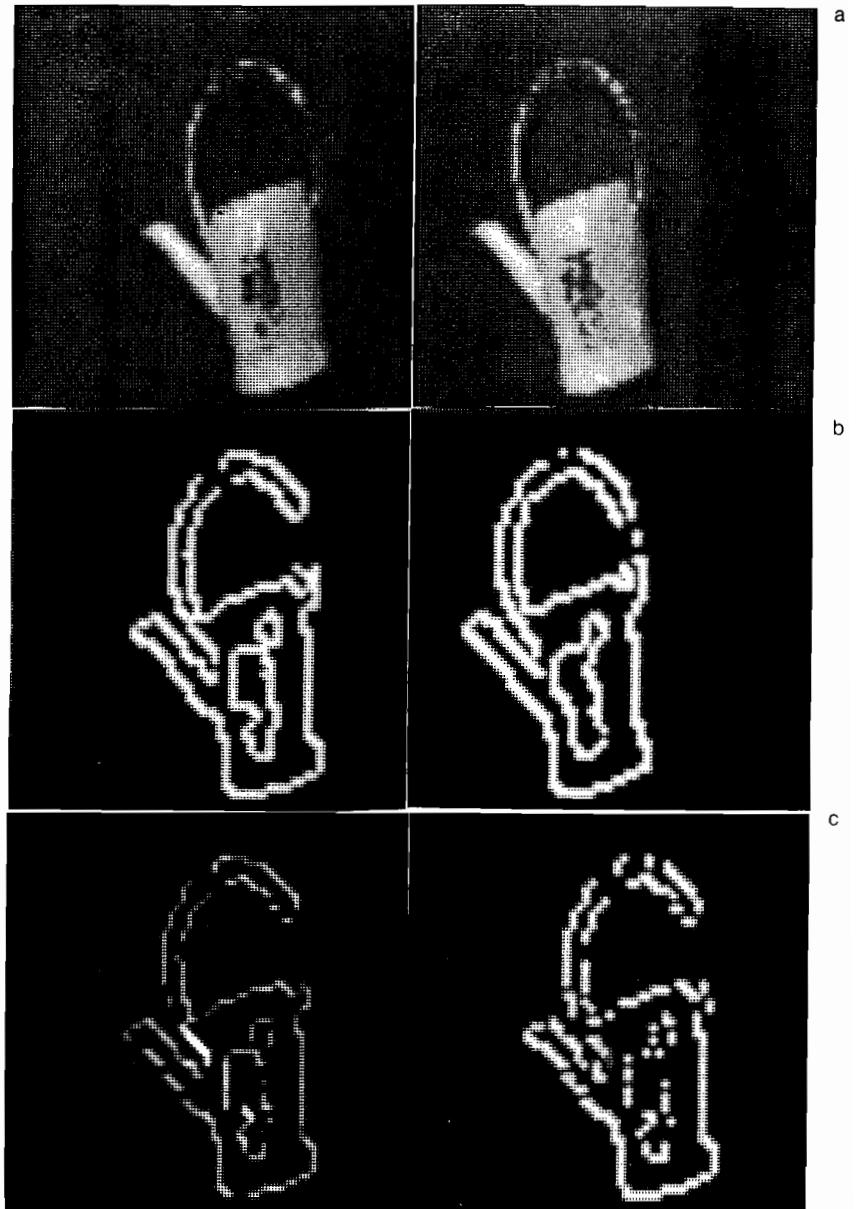


Figura 3. Otro ejemplo del procesamiento de un par estereó de imágenes naturales. En esta ocasión, la escena consiste en una tetera china. (Véanse detalles en la figura 1.)

ben encontrar correspondiente y que si el algoritmo empareja, deben considerarse como errores. De esta forma, se han considerado los siguientes tipos de emparejamientos:

- Puntos emparejados:
 - Puntos emparejados correctamente (EMPAR. CORRECTOS).
 - Puntos emparejados incorrectamente (EMPAR. FALSOS).
- Puntos no emparejados:

- Puntos no emparejados correctamente (NO EMPAR. CORRECTOS).
- Puntos no emparejados incorrectamente (NO EMPAR. FALSOS).

En las figuras 7, 8 y 9 se muestra el análisis de los emparejamientos, comparando ambos algoritmos en empar. correctos, empar. falsos, no empar. correctos y no empar. falsos, con los tres pares de imágenes naturales en estereó.

El análisis de las figuras 7, 8 y 9 muestra varios puntos importantes, siendo quizá el más destacable

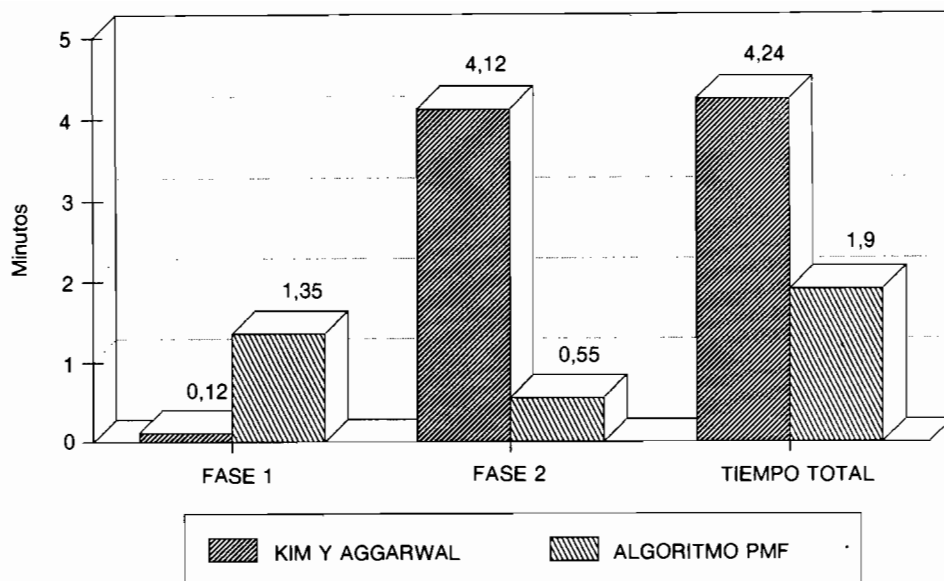


IMAGEN 1 (133 puntos)

Figura 4. Rendimiento comparativo en tiempos de procesamiento de los dos algoritmos con el par estéreo de la figura 1.

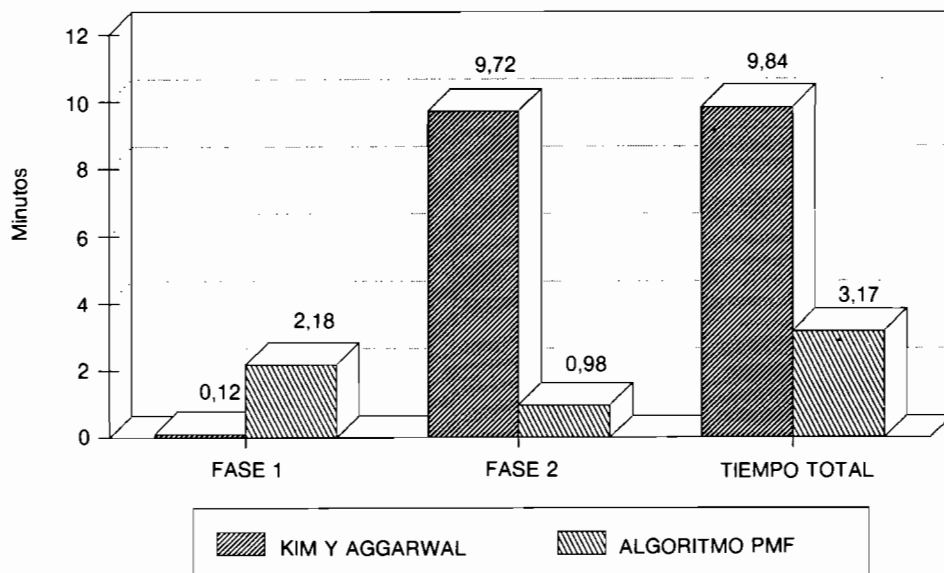


IMAGEN 2 (173 puntos)

Figura 5. Rendimiento comparativo en tiempos de procesamiento de los dos algoritmos con el par estéreo de la figura 2.

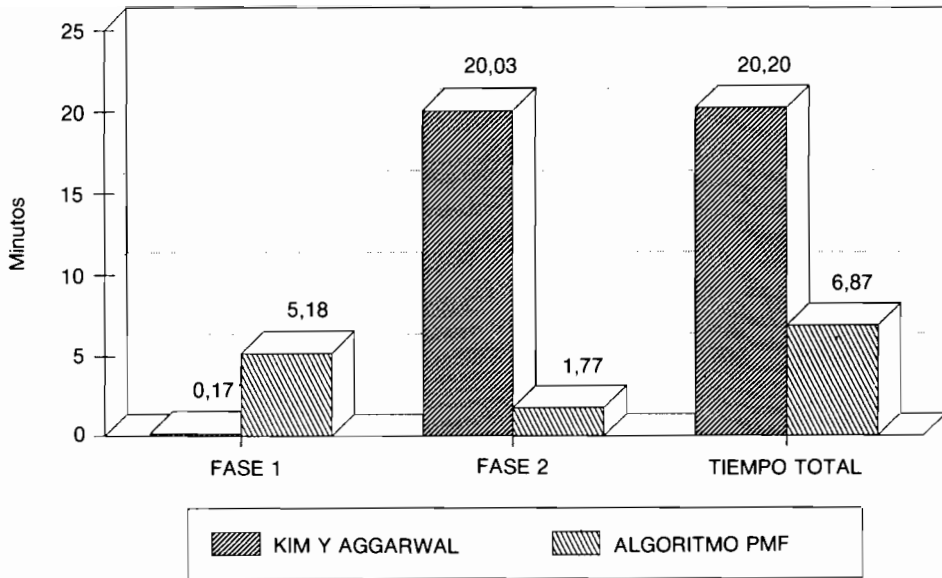


IMAGEN 3 (285 puntos)

Figura 6. Rendimiento comparativo en tiempos de procesamiento de los dos algoritmos con el par estéreo de la figura 3.

que el modelo de Kim y Aggarwal tiende a producir más emparejamientos que el PMF, lo cual a su vez provoca que los resultados muestren más correspondencias correctas y falsas, sobre todo en el procesamiento de las figuras más complejas. Estos comportamientos son fácilmente comprensibles si atendemos a la estructura de los procesos de relaja-

ción de ambos algoritmos. En el PMF, el límite del gradiente de disparidad, sobre el que se basa el modelo, restringe el número de emparejamientos posibles. Además, la propia mecánica evita que los errores se propaguen. En cambio, en el algoritmo de Kim y Aggarwal, el proceso se inicia con un número de posibles emparejamientos muy alto, lo que en

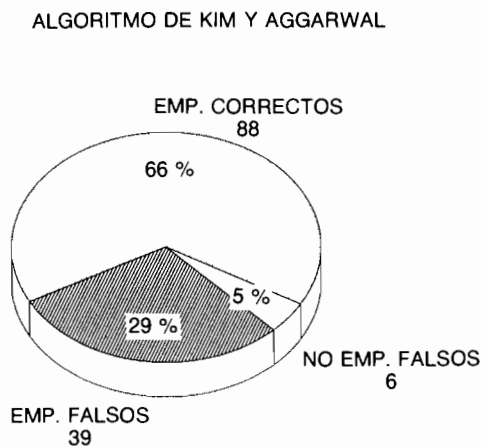


IMAGEN 1 (133 puntos)

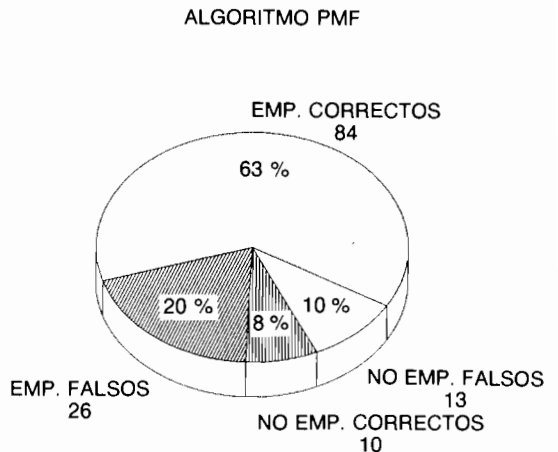


IMAGEN 1 (133 puntos)

Figura 7. Análisis de los emparejamientos comparando ambos algoritmos en emparejamientos correctos, emparejamientos falsos, no emparejamientos correctos y no emparejamientos falsos sobre el par estéreo de la figura 1.

ALGORITMO DE KIM Y AGGARWAL

ALGORITMO PMF

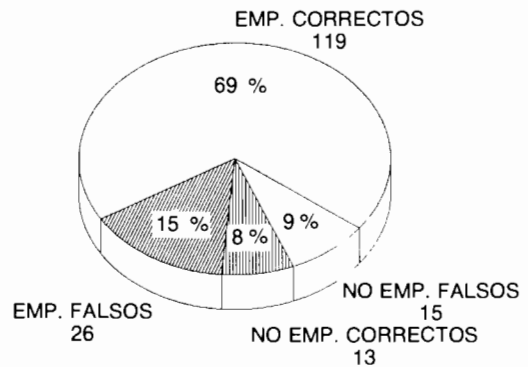
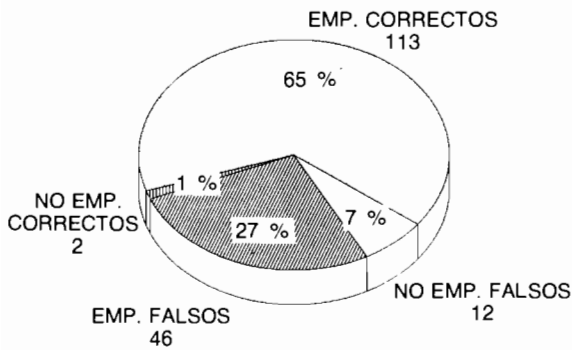


IMAGEN 2 (173 puntos)

IMAGEN 2 (173 puntos)

Figura 8. Análisis de los emparejamientos comparando ambos algoritmos en emparejamientos correctos, emparejamientos falsos, no emparejamientos correctos y no emparejamientos falsos sobre el par estéreo de la figura 2.

determinadas zonas de las imágenes permite el emparejamiento de puntos que en el PMF no entran ni en las consideraciones iniciales. Luego el proceso de relajación, basado fundamentalmente en la consideración de las probabilidades de emparejamiento de las proximidades de un punto, provoca en ocasiones la propagación de los errores o el emparejamiento de puntos con probabilidades iniciales muy bajas. No obstante, esta característica del modelo de Kim y Aggarwal hace que su rendimiento sea en

general más alto, sobre todo en figuras más complicadas. En estos casos, el PMF saca poco partido a las disparidades calculadas, de forma que en puntos con muchos candidatos la dificultad de selección se mantiene durante todo el proceso. En cambio, en el proceso de relajación de Kim y Aggarwal, los emparejamientos calculados ayudan activamente durante todas las iteraciones al resto de correspondencias por calcular, diferencia que se nota mucho más en figuras con muchos puntos y candidatos.

ALGORITMO DE KIM Y AGGARWAL

ALGORITMO PMF

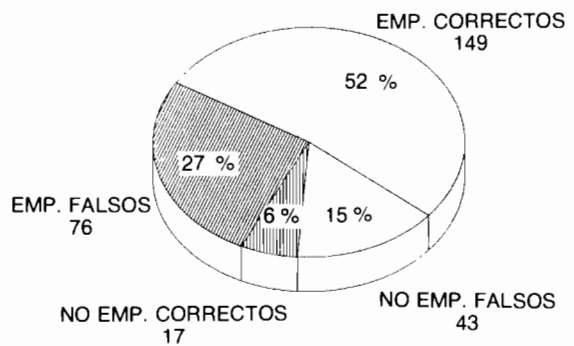
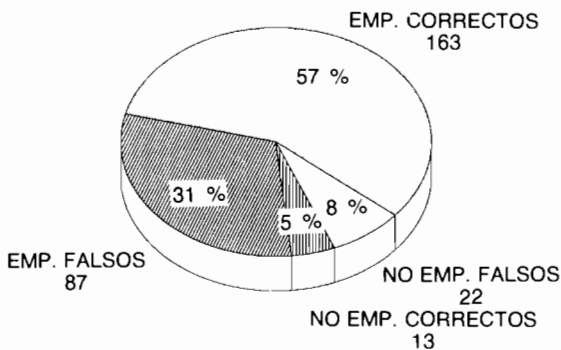


IMAGEN 3 (285 puntos)

IMAGEN 3 (285 puntos)

Figura 9. Análisis de los emparejamientos comparando ambos algoritmos en emparejamientos correctos, emparejamientos falsos, no emparejamientos correctos y no emparejamientos falsos sobre el par estéreo de la figura 3.

Conclusiones

En este trabajo de investigación, el objetivo primordial es el estudio de la eficiencia de los algoritmos de búsqueda de correspondencias más significativos, a través de implementaciones propias que reflejan lo mejor posible las líneas teóricas más importantes de los creadores. En el caso del PMF y el modelo de Kim y Aggarwal, el estudio comparativo que se ha presentado ha permitido conocer a fondo el funcionamiento de ambos y obtener conclusiones acerca de sus puntos fuertes y de posibles cambios que mejorarán sus respectivos rendimientos.

En cuanto a los puntos fuertes de ambos modelos, parece claro que la utilización de los patrones como medida de similitud y el proceso de relajación del modelo de Kim y Aggarwal, representan mecanismos que favorecen una correcta solución del proceso de búsqueda de correspondencias. No obstante, habría que mejorar el trabajo con los patrones, puesto que, por ejemplo, en el caso de puntos no conectados (por problemas de iluminación, captación de imágenes o preproceso) que pueden emparejarse con facilidad, la mecánica del proceso les impide hacerlo, dificultando de alguna forma y haciendo más lento todo el tratamiento en la zona próxima al punto. Otro punto que habría que trabajar es en la mejora de la eficiencia. La complejidad del proceso de relajación y el número de candidatos tratados hace pensar en alguna restricción inicial más y en alguna simplificación de los procesos de cálculo. En este sentido también se trabaja en el estudio del número de iteraciones necesarias para cada figura y con qué factores está ligado, aspecto éste crítico en cuanto a la calidad de la solución que se obtiene.

En relación al modelo PMF, es obvio que su punto más importante es la restricción del límite del gradiente de disparidad. La utilización de este concepto permite que el tratamiento sea más eficiente, al reducir el número de emparejamientos posibles. Sin embargo, se piensa que el algoritmo debería introducir en el proceso de selección las disparidades que ya se consideran seguras, y de esta forma, se

mejorará el proceso de selección en los emparejamientos más complicados y, a su vez, aumentará el rendimiento del proceso.

Referencias

- Baker, H. H. (1982). Depth from edge and intensity based stereo. AI Memo 347, Stanford, CA, Stanford University, USA.
- Burt, P. y Julesz, B. (1980). Modifications of the classical notion of Panum's fusional area. *Perception*, 9, 671-682.
- Campbell, F. W. y Robson, J. (1968). Application of Fourier Analysis to the visibility of gratings. *Journal of Physiology*, 197, 551-566.
- Kim, Y. C. y Aggarwal, J. K. (1987). Positioning three-dimensional objects using stereo images. *IEEE Journal of Robotics and Automation*, RA-3, 4, 361-373.
- Marr, D. y Hildreth, E. (1979). Theory of edge detection. MIT AI Lab. Memo 518.
- Marr, D. y Poggio, T. (1979). A computational theory of human stereo vision. *Proc. R. Soc. Lond.*, B204, 301-328.
- Pollard, S. B., Mayhew, J. E. W. y Frisby, J. P. (1985). PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception*, 14, 449-470.
- Pollard, S. B., Porrill, J., Mayhew, J. E. W. y Frisby, J. P. (1985). Disparity gradient, Lipschitz continuity, and computing binocular correspondences. *Proceedings of the Third International Symposium of Robotics Research*, Gouviex, France (Cambridge, MA: MIT Press).
- Rubio de Lemus, P. (1986). *Simulación de la visión estereoscópica humana*. Memoria de Licenciatura. Laboratorio de Visión. Facultad de Psicología de la UCM.
- Rubio de Lemus, P. y Pérez, R. (1987). Simulación de la visión estereoscópica humana. *Revista de Informática y Automática*, XX, 3, 19-26.
- Rubio de Lemus, P. y Pérez, R. (1990). Simulación de dos sistemas de visión estereoscópica humana: un estudio comparativo. *Resúmenes del VIII Congreso Nacional de Psicología*, Barcelona.
- Rubio de Lemus, P., Ruiz, T. y Pérez Muñoz, R. (1990). A comparative study between Marr & Poggio's model and the PMF stereo algorithm. *Abstracts de la 13th European Conference on Visual Perception*, Paris, 362.