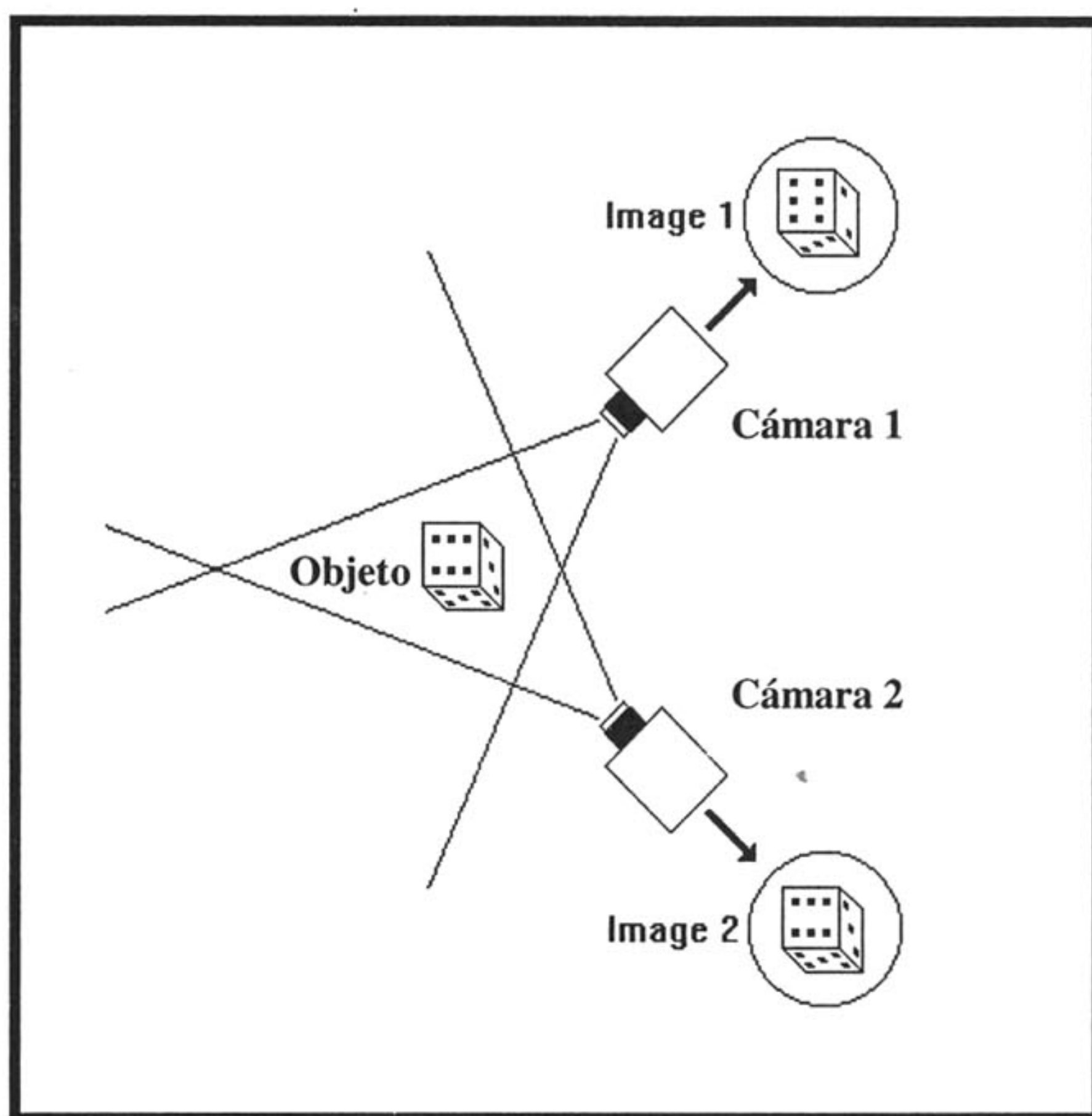


Estereovisión: Algoritmo para establecer correspondencias vía combinación de clasificadores



real. El sensor utiliza dos mini-cámaras dispuestas verticalmente. El procesamiento digital permite detectar y conformar los segmentos de recta presentes en las imágenes. Sobre los segmentos detectados se calcula un número importante de características, que constituyen la base para la etapa de puesta en correspondencia. Para efectuar la correspondencia es necesario responder a la pregunta: ¿«Son los dos segmentos, detectados con las cámaras superior e inferior, la imagen de la misma parte del mundo real del robot?». En este trabajo se da respuesta a dicha pregunta, implementando un novedoso método basado en la combinación de las funciones de clasificación bayesiana y neuronal.

PALABRAS CLAVES

Visión estereoscópica, Puesta en correspondencia, Clasificador Bayesiano, Clasificador neuronal, Robótica móvil.

□ **Humberto Loaiza Correa. Ph.D.**
Profesor Titular.
EIEE. Universidad del Valle.

□ **Sylvie Lelandais. Ph.D.**
Maître de Conférences.
Laboratoire Systèmes
Complexes.
Université d'Evry. France.

Grupo de investigación, PSI ,percepción
y sistemas inteligentes

RESUMEN

Para permitir la localización y navegación de un robot móvil dentro de un ambiente interior se ha diseñado y construido integralmente un sensor estereoscópico de visión. El software del sensor permite efectuar la reconstrucción 3D en tiempo

ABSTRACT

In order to permit the localization and the navigation of a mobile robot within an interior environment we have built a stereoscopic sensor and implemented all the algorithms which allow to obtain 3D coordinates of real objects from data images. Sensor uses two mini cam-

meras on a vertical arrangement. Processing of the images leads us to have segments on the two images. On these segments many features are computed so it is possible to begin a step of matching. To match segments, we have to answer at the question: "Are the two segments, in a pair from top and bottom image segments, two views of the same part of the real world of the robot?". We use a combination of two functions of classification for giving a reply: Bayes and neural functions.

KEY WORDS

Stereoscopic vision – Matching – Bayes classifier – Neural classifier – Mobile robotic

1. INTRODUCCIÓN

En este artículo se describe un sistema de visión estereoscópica vertical para la robótica móvil y los procedimientos originales implementados para resolver el difícil problema de puesta en correspondencia de segmentos de recta presentes en las imágenes superior e inferior del sensor, sin necesidad de recurrir a la restricción epipolar.

En la primera parte de este documento se presentan las características generales de la cabeza estereoscópica y el procesamiento para mejorar las imágenes -las cuales inicialmente se encuentran bastante degradadas- y extraer los segmentos. En la segunda parte, se describen las características calculadas sobre cada segmento y el vector de 16 parámetros obtenidos a partir de estas características. Se presenta el estudio de cuatro clasificadores para asignar cuplas de segmentos a dos clases: *bien apareados* y *mal apareados*. También se describe el procedimiento efectuado para reducir el número de componentes del vector asociado a cada cupla de segmentos. En la tercera parte se describe el algoritmo para resolver el problema de puesta en correspondencia, el cual se basa en los resultados de la *complementaridad* entre dos clasificadores. En la cuarta parte, se muestran y discuten algunos de los resultados obtenidos. Finalmente se presentan las conclusiones del trabajo de investigación y se exponen ideas de trabajos futuros en la misma línea.

1. SENSOR ESTEREOSCÓPICO DE VISIÓN Y PROCESAMIENTO DE IMÁGENES

La cabeza estereoscópica posee dos mini-cámaras, las cuales no presentan ninguna posibilidad de ajuste en sus características ópticas, mecánicas y eléctricas. Las cámaras usadas son monocromáticas, de muy bajo costo, producen imágenes de calidad bastante degradada y no disponen de información sobre sus

características – se desconoce: el tamaño de la matriz CCD, la distancia focal, el tamaño de las celdas fotodetectoras, la distorsión radial y tangencial, entre otros. Las dos mini-cámaras están dispuestas verticalmente, a diferencia de los sistemas estereoscópicos clásicos. En la figura 1 puede apreciarse el prototipo del sensor. La cabeza estereoscópica está montada sobre un motor paso a paso, y el conjunto ha sido embarcado sobre un robot, como se aprecia en el robot móvil *Real Magicol* [1] de la figura 2.

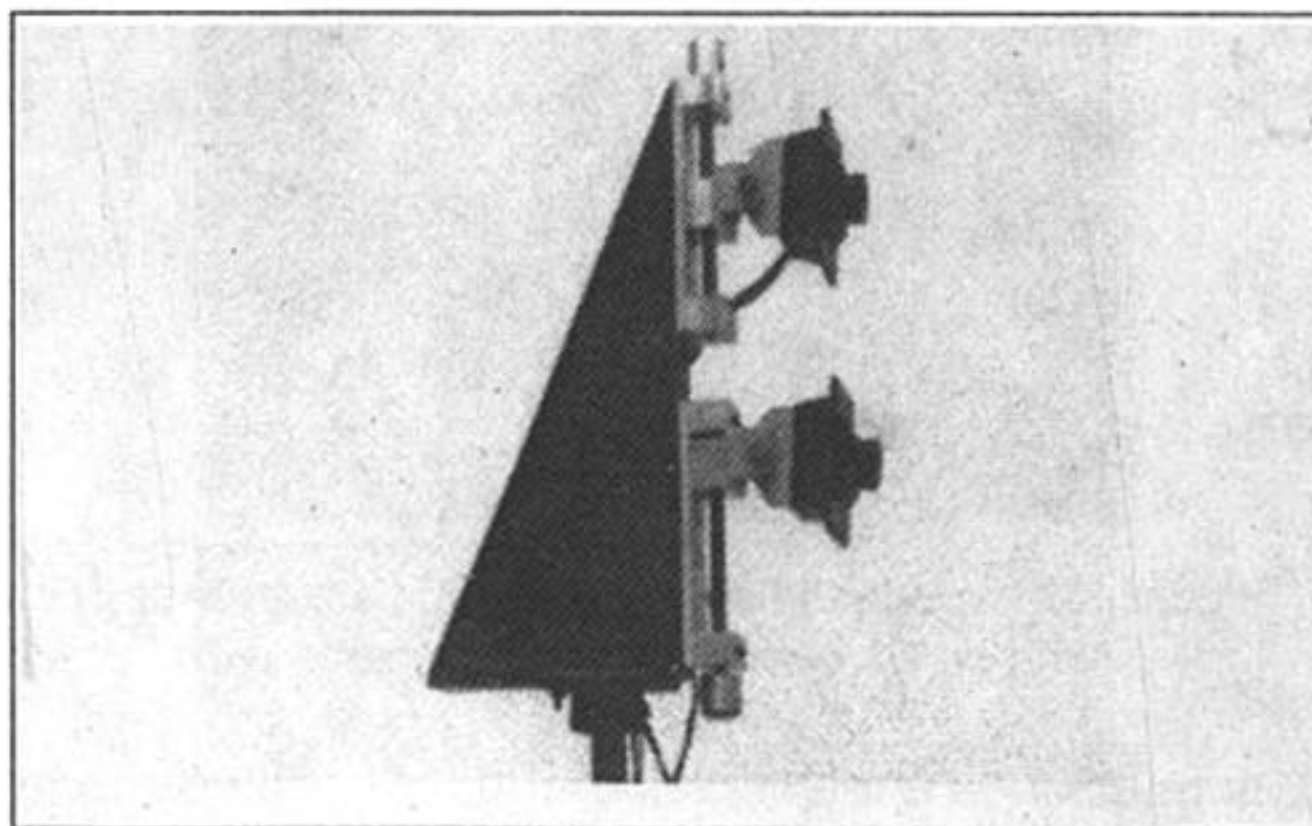


Figura 1. Prototipo del sensor estereoscópico de visión

El campo de visión de la cámara inferior se encuentra orientado ligeramente hacia el suelo, y el de la cámara superior está dirigido hacia la horizontal. Lo anterior permite tener un traslape parcial de los dos campos de visión.

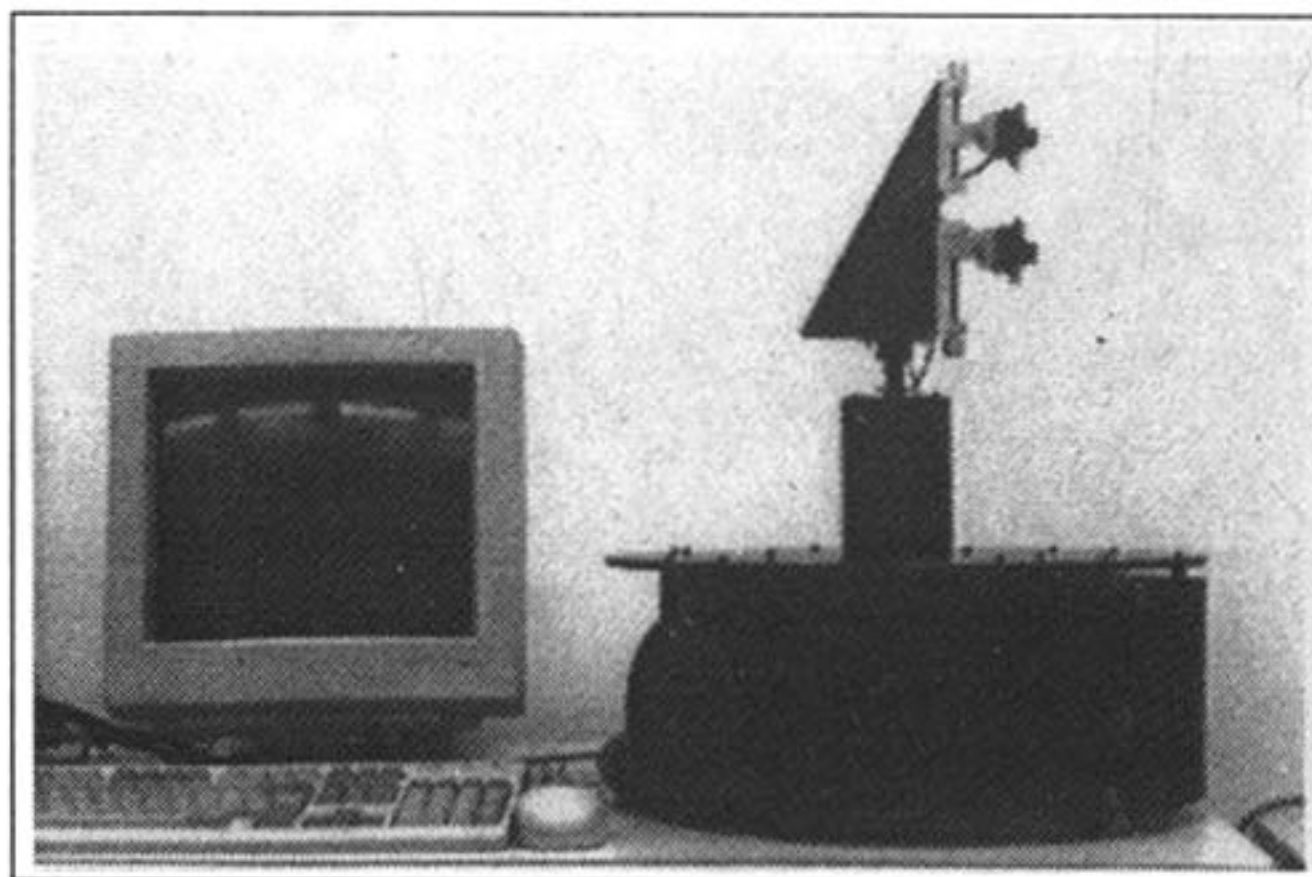


Figura 2. Sistema estereoscópico montado sobre el robot móvil *Real Magicol*.

Con miras a efectuar la reconstrucción 3D, es necesario realizar una etapa de calibración para determinar los parámetros intrínsecos y extrínsecos de cada una de las cámaras. Para tal efecto, se utilizaron dos modelos: el modelo pinhole, que desprecia todas las distorsiones ópticas [2] y un modelo que toma en cuenta la distorsión radial [3]. El proceso de calibración requiere en su etapa inicial el conocimiento de las coordenadas del mundo y del plano imagen de un conjunto de puntos no

coplanares [4]. Para la calibración se implementó un patrón de dos planos ortogonales en los cuales se dispusieron rectángulos negros sobre un fondo blanco (Figura 3). Las coordenadas de los puntos de las aristas de los rectángulos se detectaron y utilizaron para inicializar el algoritmo de calibración.

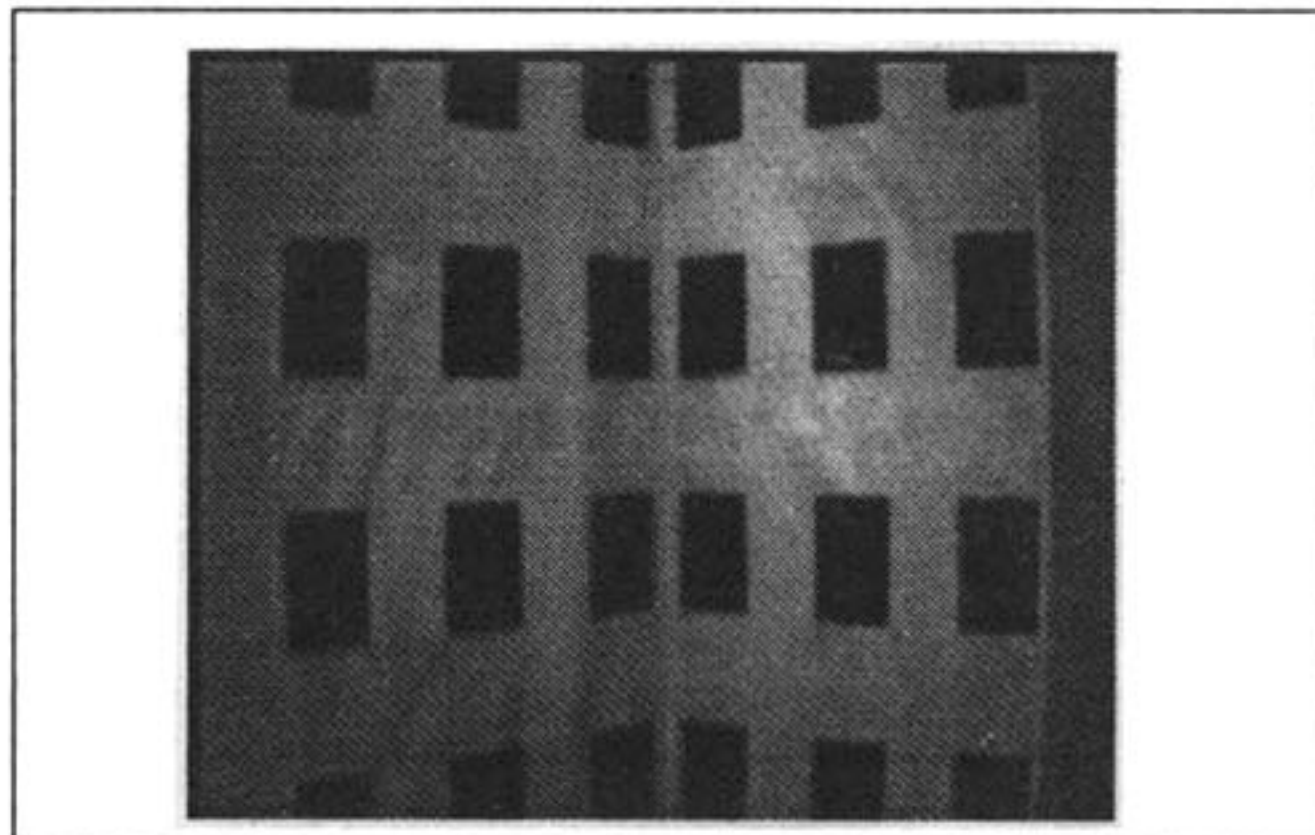


Figura 3. Imagen del patrón de calibración, tomada por una de las cámaras del sensor estéreo.

Las relaciones entre las coordenadas de las cámaras y las coordenadas de la imagen de un punto están establecidas para los modelos pinhole lineal y no lineal por las ecuaciones (1) y (2) respectivamente:

$$u = u_0 + k_u \frac{x_c}{z_c} \quad v = v_0 + k_v \frac{y_c}{z_c} \quad (1)$$

$$u = u_0 + k_u \frac{x_c}{z_c} (1 + k_d r^2) \quad (2a)$$

$$v = v_0 + k_v \frac{y_c}{z_c} (1 + k_d r^2) \quad (2b)$$

donde: $r^2 = \frac{x_c^2 + y_c^2}{z_c^2}$, x_c, y_c, z_c

coordenadas de un punto en el sistema coordenado de la cámara; u, v coordenadas del pixel en el plano retinal; u_0, v_0 coordenadas del centro de la imagen; k_u, k_v componentes horizontal y vertical de la longitud focal; k_d parámetro de distorsión radial.

Dentro de los procedimientos efectuados para especificar las cámaras y definir la calidad de las imágenes generadas por el sensor, se realizaron comparaciones entre una cámara de la cabeza estereoscópica y una cámara estándar [5]. El análisis y comparación de las imágenes de una misma escena tomadas bajo idénticas condiciones con las dos cámaras estudiadas,

pusieron en evidencia una fuerte distorsión en «barril», una débil definición de los niveles de gris y un rango dinámico reducido presentes en la cámara del sensor estéreo.

Teniendo en cuenta lo anterior, se diseñó una compleja etapa de pre-procesamiento de imágenes para corregir la distorsión y mejorar las imágenes para permitir una mejor detección y extracción de segmentos de recta. La decisión para retener los segmentos de recta como primitivas fue tomada con base en el hecho de que éstos pueden ser detectados más fácilmente en los ambientes estructurados por donde evoluciona el robot móvil. Además, esta clase de primitivas permite calcular una gran variedad de descriptores y derivar en una función que es muy sensible a las variaciones de estos descriptores [6].

Los diferentes pasos de pre y procesamiento de imágenes realizados, se resumen a continuación:

- *Corrección de la distorsión y Mejoramiento de imagen.* En esta parte se corrige la distorsión radial optimizando el algoritmo a través de la utilización de una LuT (Look up Table) y se aumenta el contraste con la técnica de especificación –gausiana– del histograma.
- *Filtraje óptimo.* Se utiliza un filtro de alisamiento y derivativo para resaltar los puntos de los bordes de los objetos. El cálculo de los componentes del gradiente se efectúa separadamente con ayuda del filtro óptimo de Deriche [7].
- *Máximos locales.* Se detectan los puntos que presenten una magnitud del gradiente superior a sus vecinos en la dirección del gradiente. Lo anterior permite obtener bordes muy finos de espesor 1 pixel.
- *Cálculo automático de umbrales.* Los umbrales se obtienen utilizando el algoritmo original ApEMG, el cual permite calcularlos dinámicamente sobre cada nueva imagen y en tiempo real [8].
- *Segmentación por histéresis.* Procedimiento que emplea la *umbralización por histéresis* [9] para detectar, extraer y seguir puntos de bordes. El método también permite simultáneamente construir y fusionar los segmentos de recta presentes en las imágenes. El resultado es una lista de segmentos definidos por las coordenadas imagen de sus puntos extremos [8].

La figura 4a muestra una imagen de un ambiente tipo oficina tomada con una de las cámaras del sensor estéreo; la figura 4b muestra, la imagen luego de corrección de la distorsión y extracción de máximos locales de la magnitud del gradiente; la figura 4c presenta los pixeles de contornos detectados; y la figura 4d, presenta una imagen reconstruida a partir de la lista de segmentos detectados con una longitud superior a 10 pixeles. Puede apreciarse que la can-

tividad de segmentos retenidos es suficiente para calcular muchas coordenadas 3D de los objetos presentes en la escena.

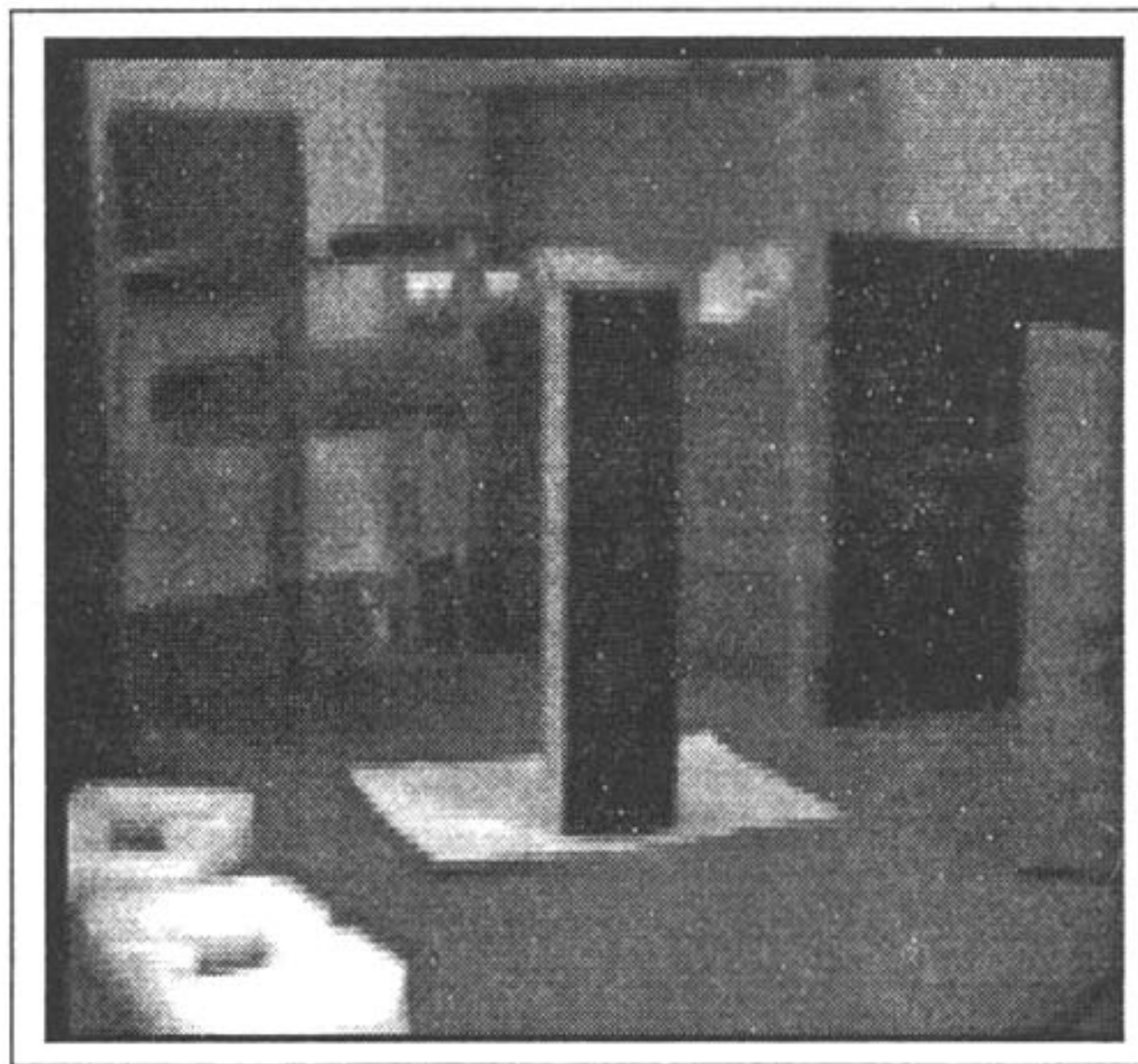
2. CLASIFICACIÓN DE DATOS

2.1. Características de segmentos

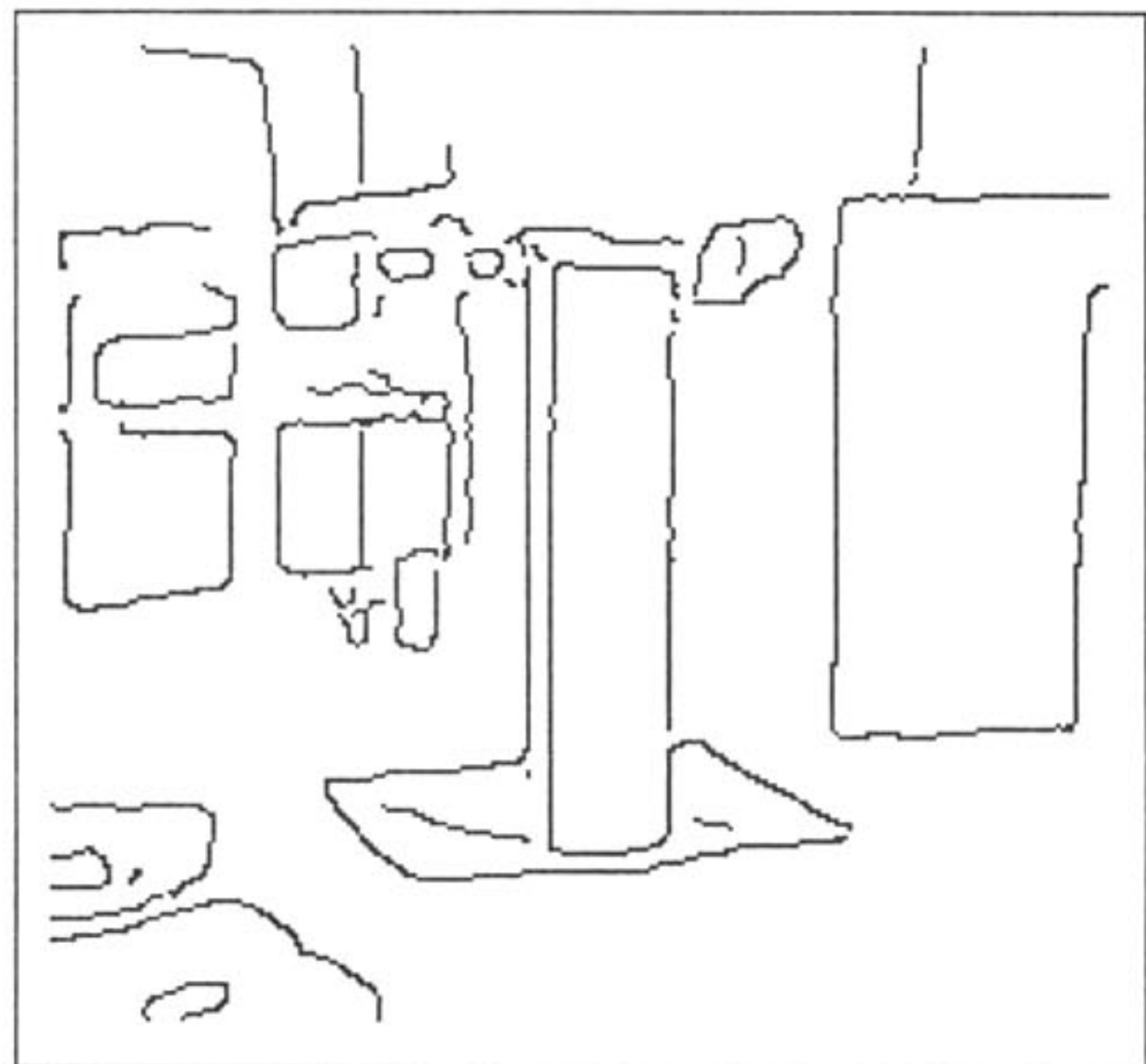
Luego de la etapa de generación de segmentos se calculan los valores de los descriptores. Se han escogido cuatro clases de descriptores que involucran características geométricas, de nivel de gris, de textura y de vecindad. Durante la etapa de creación de segmentos se calculan los parámetros geométricos y el valor medio del gradiente del segmento. Para calcular las otras características que consideran el nivel de gris

y aspectos de textura, se definieron dos regiones de píxeles, una al lado izquierdo y otra al lado derecho de cada segmento, como se ilustra en la figura 5.

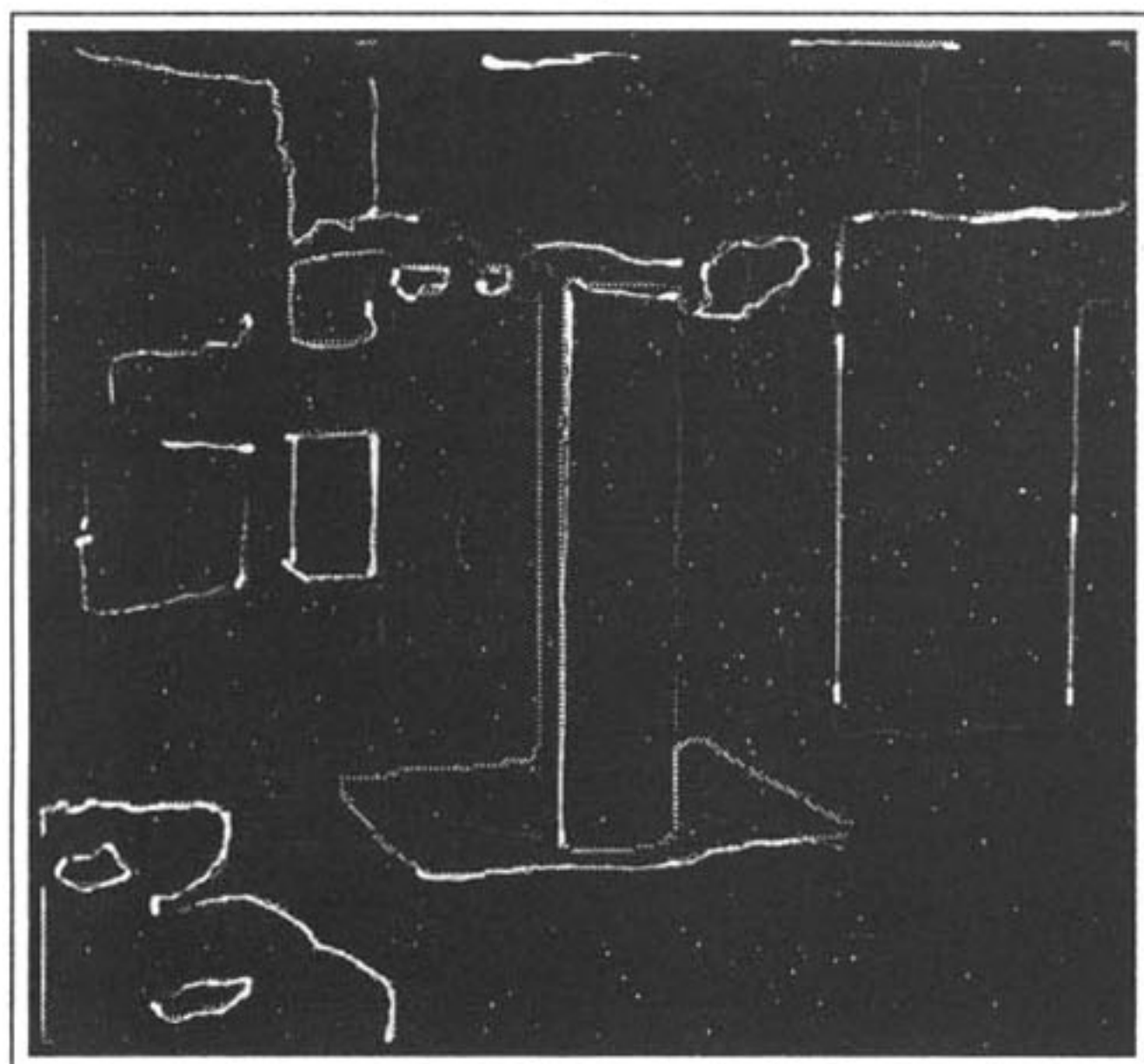
Sobre estas regiones se calcula el valor medio de los niveles de gris, el contraste interno y las diferencias de niveles de gris en cuatro direcciones (0° , 45° , 90° y 135°) que representan la variación local. Estos últimos cuatro parámetros texturales se calcularon como simplificaciones de la matriz de concurrencia [10] para reducir el tiempo de cómputo. Para el cálculo de las características de vecindad, se utiliza el método de «buckets» [11] que particiona la imagen en cuadrículas «deslizantes» para optimizar el número de vecinos de cada segmento. Hasta esta parte del procedimiento, se tienen 16 parámetros para describir un segmen-



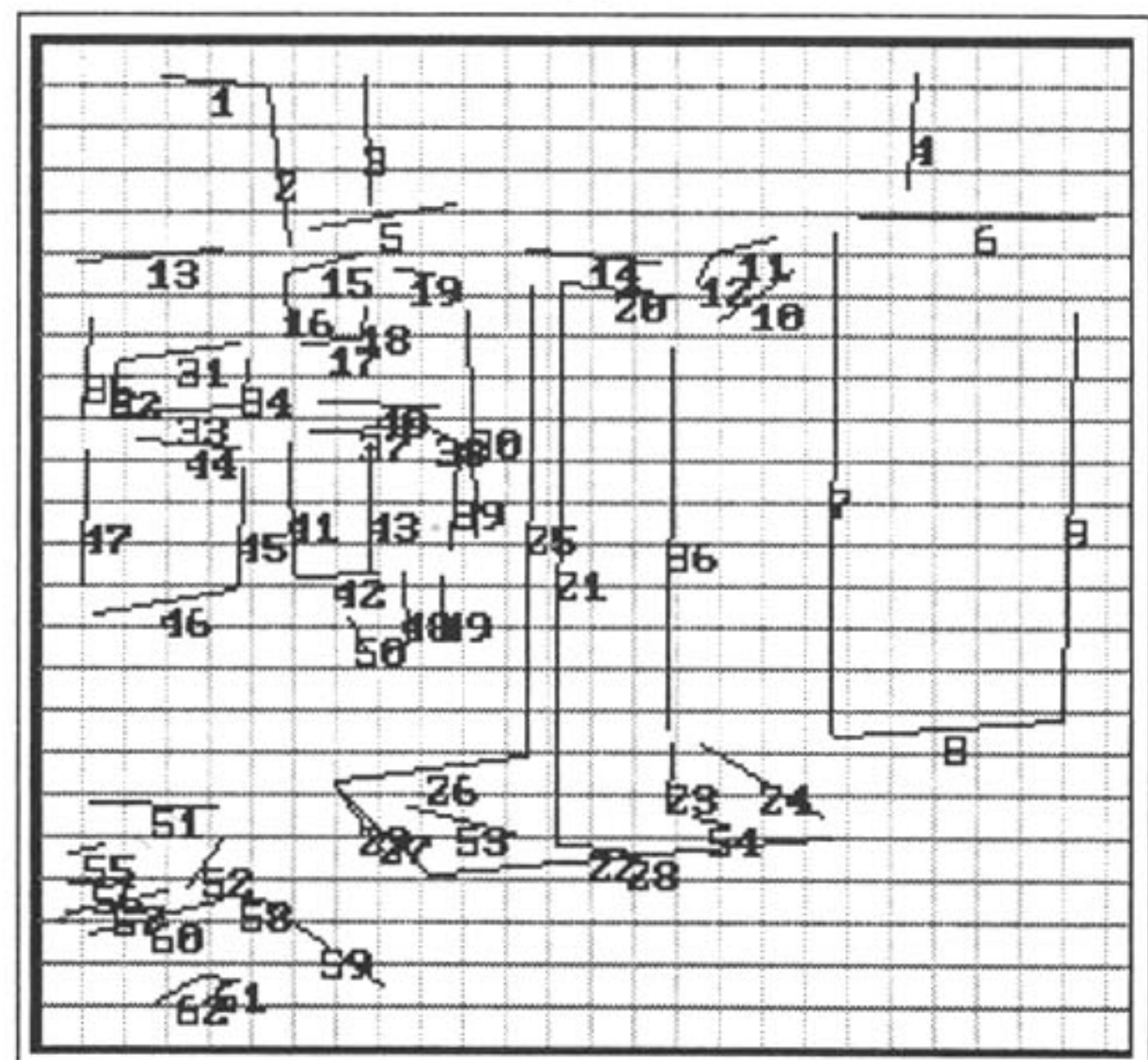
(a)



(c)



(b)



(d)

Figura 4. Secuencia de imágenes correspondiente a diferentes etapas de extracción de primitivas.

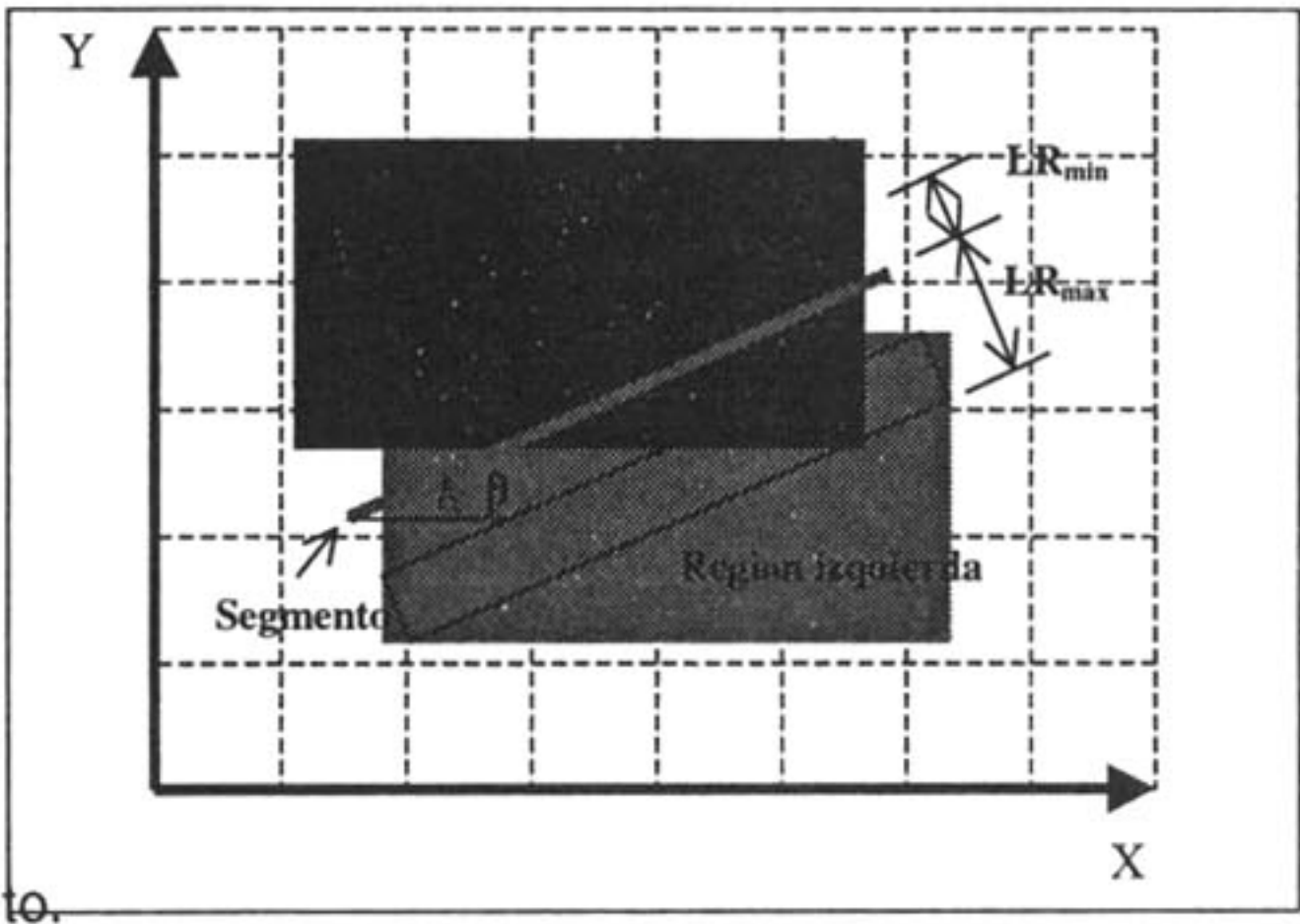


Figura 5. Definición de regiones a los lados del segmento

La meta de la puesta en correspondencia es comparar un par de segmentos, uno de la imagen superior y otro de la imagen inferior, y decidir si ellos representan la misma parte de la escena del mundo. Para lograrlo, se crearon dos clases: Clase 1 → pares de segmentos bien apareados, Clase 2 → pares de segmentos mal apareados. Un vector de 16 componentes se asocia a cada miembro de la clase. Estas 16 variables (x^i) se obtienen como la diferencia entre los parámetros de cada segmento del par a clasificar, como lo indica la ecuación (3):

$$\begin{pmatrix} x^1 \\ x^2 \\ \dots \\ x^{16} \end{pmatrix} = \begin{pmatrix} x_{upp}^1 \\ x_{upp}^2 \\ \dots \\ x_{upp}^{16} \end{pmatrix} - \begin{pmatrix} x_{low}^1 \\ x_{low}^2 \\ \dots \\ x_{low}^{16} \end{pmatrix} \tag{3}$$

donde X_{upp}^i , X_{low}^i , indican características de los segmentos de la imagen superior e inferior respectivamente. La tabla 1 indica el significado de cada variable.

Para analizar y probar el poder descriptivo de cada variable se construyeron cuatro bases de datos, cada una con 660 miembros. Los miembros se obtuvieron luego de procesar una importante cantidad de pares de imágenes. La base 1 y 2 están conformadas por segmentos extraídos de imágenes altamente texturadas. Las bases 3 y 4 se obtuvieron a partir de imágenes con segmentos poco texturados. Los miembros de las bases 1 y 3 son de clase 1 (bien apareados), y los miembros de las bases 2 y 4 son de clase 2 (mal apareados).

2.2. El problema de reconocimiento de patrones
Es claro que el problema ahora consiste en determinar

Variables	Calculada como la diferencia entre :
x^1	Coordenadas vertical del centro del segmento
x^2	Coordenadas horizontal del centro del segmento
x^3	Magnitud del gradiente a lo largo del segmento
x^4	Orientación del segmento (grados)
x^5	Contraste interno en la región izquierda
x^6	Contraste interno en la región derecha
x^7	Promedio de nivel de gris en la región derecha
x^8	Promedio de nivel de gris en la región izquierda
x^9	Variaciones de gris a 0° en la región derecha
x^{10}	Variaciones de gris a 45° en la región derecha
x^{11}	Variaciones de gris a 90° en la región derecha
x^{12}	Variaciones de gris a 135° en la región derecha
x^{13}	Variaciones de gris a 0° en la región izquierda
x^{14}	Variaciones de gris a 45° en la región izquierda
x^{15}	Variaciones de gris a 90° en la región izquierda
x^{16}	Variaciones de gris a 135° en la región izquierda

Tabla 1. Definición de los 16 componentes del vector para un par de segmentos.

el mejor clasificador. La literatura propone varios métodos para resolver el problema de clasificación automática. La selección depende de las restricciones de la aplicación y del conocimiento a priori de los datos de entrada y del fenómeno físico involucrado. Generalmente el mejor clasificador para una tarea dada puede encontrarse al evaluar y comparar diferentes criterios como el error de clasificación, la complejidad computacional y la eficiencia de la implementación en hardware [12]. Para este trabajo se seleccionaron cuatro métodos de clasificación supervisada bien conocidos: Método Bayesiano (BM), Análisis Discriminante Lineal (LDA), Análisis discriminante Cuadrático (QDA), y el Perceptrón Multi-Capa (MLP) [13].

2.2.1. Aproximación Estadística

Sea X un vector de p dimensiones, definido como $K=[X_1, X_2, \dots, X_p]^T \in R^p$ y sea K la clase a la que pertenece el vector X , donde $K \in \{1, \dots, K\}$. Un clasificador puede considerarse una función $c: R^p \rightarrow \{1, \dots, K\}$ que asigna un vector de características dado a una de las clases K [14]. La regla optimal de Bayes $c(x)$ para dos clases que minimiza la tasa de error esperado es:

$$c(x) = \begin{cases} \text{Clase 1 si } h(x) > 1 \\ \text{Clase 2 si } h(x) < 1 \end{cases} \tag{4}$$



con,
$$h(x) = \frac{C_{1/2} p_1 f_1(x)}{C_{2/1} p_2 f_2(x)} \tag{5}$$

Donde, p_k denota la probabilidad a priori de la clase k , $f_k(x)$ es la densidad de probabilidad de la clase k , $C_{k1/k2}$ representa el costo de clasificar un vector de la clase k_1 como de clase k_2 , y $h(x) = 1$ constituye la superficie de discriminación.

Para la siguiente discusión, se considera que los vectores de los patrones de la clase k presentan una distribución normal con vector medio u_k y matriz de covarianza Σ_k . Si la función de costo se considera igual para todas las clases, la regla anterior deriva en la regla QDA. Si se asume que las clases presentan distribución normal, diferentes vectores promedio, pero con igual matriz de covarianza Σ , la regla de decisión deriva en la regla LDA. El método de verificación de la calidad se define por el factor D_k que indica el porcentaje de éxito para la clase k .

2.2.2. Aproximación Bayesiana

La red neuronal que se utilizó en este trabajo es un perceptrón multicapa entrenada con el algoritmo «back-propagation». La red está compuesta de una capa oculta con funciones de activación no lineales sigmoidales, 16 entradas y 2 salidas con funciones de activación lineales. La mejor arquitectura se escogió después de comparar el error de generalización para diferentes inicializaciones y cantidad de neuronas en la capa oculta. Para maximizar las capacidades de la generalización se utilizó un algoritmo de validación limitada [15]. El conjunto de datos se dividió en tres partes iguales: datos de aprendizaje, datos de validación, y datos de generalización [16]. Para evaluar la probabilidad a posteriori de la clasificación neuronal se utilizó una función softmax [17].

En la siguiente sección se presentan los resultados de generalización alcanzado con todos los métodos. La configuración final de la red neuronal, se seleccionó considerando el más alto porcentaje de generalización.

2.3. Resultados de clasificación

Los cuatro métodos de clasificación fueron probados con todos los patrones de las bases de datos. El rendimiento de la clasificación para los métodos Bayesiano, LDA, QDA y MLP se presentan en la tabla 2. Los resultados muestran que los clasificadores Bayesiano y Neuronal conducen a los más altos índices

Método	D ₁ %	D ₂ %
LDA	63,4	54,5
QDA	97,3	82,7
MLP	91,0	93,6
Bayes	97,3	90,0

Tabla 2. Resultados de clasificación. D_k = éxito para la clase k . Clase 1= Bien apareado; Clase 2= Mal apareado.

de buena clasificación para las dos clases, e igualmente indican que los descriptores seleccionados para la conformación de los vectores es acertada.

2.4. Reducción del número de variables

Siendo consciente de que el cálculo de los segmentos y sus atributos debe realizarse dinámicamente para cada nuevo par de imágenes estereoscópicas, que la aplicación requiere implementarse en tiempo real, es natural cuestionarse sobre si el vector de 16 variables para caracterizar cada patrón es demasiado grande y si algunas de las variables están correlacionadas. Por tales razones, se decidió analizar la posibilidad de reducir el espacio de variables. Se probaron tres métodos: Análisis en Componentes Principales (ACP), Análisis Factorial Discriminante (AFD) y un método Combinatorio que utiliza el criterio de la Traza (CT) [5]. El resultado arrojado por los tres métodos indica que la cantidad óptima de variables es 8, pero las ocho características seleccionadas fueron diferentes para cada método!. Para determinar el mejor sub-espacio de variables, se procedió a repetir el procedimiento de clasificación de patrones pero sólo con los clasificadores Bayesiano y neuronal. Se encuentra que el sub-espacio obtenido con el AFD conduce a los mejores resultados. Los componentes retenidos son: $x^1, x^2, x^4, x^5, x^6, x^8, x^9, x^{14}$ (ver tabla 1). Es importante comprobar, que la composición de este sub-espacio, incluye variables de tres tipos de características: geométricas, nivel de gris y textura.

2.5. Complementaridad Bayes/Neuronal

La tabla 3a presenta los resultados de clasificación obtenidos con la representación de 8 variables por pareja de segmentos. El método neuronal arroja unos porcentajes ligeramente superiores al método Bayesiano. Interesados en determinar cuál era el conjunto de parejas mal clasificadas, se efectuó un estudio que permitió demostrar que estos conjuntos no eran iguales para los dos métodos. Para verificar lo anterior, se definió el *porcentaje de complementaridad* (PC) entre las dos aproximaciones. El PC del método A sobre el método B para la clase i , se define como:

$$PC_{A \rightarrow B}^i = 1 - \frac{card(F_A^i \cap F_B^i)}{card(E^i)} \tag{6}$$

donde, F_M^i es el conjunto de patrones de la clase i mal clasificados por el método M y E^i es el conjunto total de patrones de la clase i .

En la tabla 3b se consignan los resultados obtenidos para los métodos Bayesiano y Neuronal. La interpretación de estos resultados indica que es posible aumentar potencialmente la capacidad de clasificación

de la aplicación si se combinan los dos métodos. Esta idea original en estereo-visión se desarrollará en las dos secciones siguientes.

Método	Bayesiano %	Neuronal %
Class 1	93,75	96,43
Class 2	92,73	92,73
Mean	93,24	94,58

Tabla 3a. Resultados de clasificación con ocho variable.

Complement	Bayes/Neural %	Neural/Bayes %
$CP_{A \rightarrow B}^1$	4,54	2,67
$CP_{A \rightarrow B}^2$	3,63	0,00

Tabla 3b. Porcentaje de complementaridad entre métodos.

2.6. Combinación de dos clasificadores

Dadas las características del problema, para este trabajo se decidió utilizar la regla de la suma [18] como estrategia de combinación de clasificadores. Bajo la suposición que las representaciones usadas son independientes y estadísticamente condicionadas, se establece la regla:

Asignar: $Z_k \rightarrow w_j$ Si

$$P_B(w_j | \mathbf{x}_k) + P_N(w_j | \mathbf{x}_k) = \max_{i=1}^{n_{\text{candida}}} [P_B(w_j | \mathbf{x}_i) + P_N(w_j | \mathbf{x}_i)] \quad (7)$$

donde $P_B(W_j | X_k)$ y $P_N(W_j | X_k)$ son las probabilidades a posteriori de que la pareja Z_k pertenezca a la clase W_j . Estas probabilidades a posteriori son suministradas por los clasificadores bayesiano y neuronal.

Los resultados de aplicar esta estrategia de combinación a las bases de datos se presentan en la tabla 4. Se puede apreciar que los resultados en promedio para las dos clases son mejores que los arrojados por un solo clasificador. Otra ventaja que se obtiene, bajo este esquema es la posibilidad de confirmar la respuesta de un método con la respuesta del otro.

Combinación	Clase 1 %	Clase 2 %
Sum(Bayes, Neural)	93.75	96.36

Tabla 4. Resultado promedio de clasificación con combinación de clasificadores.

3. ALGORITMO DE PUESTA EN CORRESPONDENCIA

Los algoritmos clásicos de puesta en correspondencia utilizan la restricción epipolar para encontrar candidatos y construir parejas de primitivas [6]. El algoritmo propuesto por los autores es construido a partir de dos hechos simples pero robustos. De una parte, se utiliza una función de decisión proveniente de la fusión de dos métodos de discriminación eficiente. Esta función permite determinar si un par de segmentos constituye un buen apareamiento. De otra parte la información de vecindad utilizada por el algoritmo reduce significativamente el campo de búsqueda para posibles apareamientos. Esto se apoya en el hecho de que los vecinos de una pareja de segmentos bien apareados conforman un espacio reducido con alta probabilidad de encontrar nuevos apareamientos correctos.

Puesto que el problema de inicialización del algoritmo no es trivial, se explotan dos aspectos relevantes de la aplicación: la configuración vertical del sensor estereoscópico y el espacio de evolución de los robots. De lo anterior se desprende que los segmentos verticales constituyen elementos confiables para la inicialización del algoritmo. Los segmentos verticales pueden detectarse y poner en correspondencia con una alta confiabilidad debido a que sus coordenadas horizontales son similares tanto en la imagen superior como en la inferior. Además, estos segmentos constituyen un conjunto reducido de los segmentos detectados en cada imagen de la escena. Lo anterior confirma que los segmentos verticales constituyen un apropiado sub-conjunto para la etapa de inicialización del algoritmo de establecimiento de correspondencia. Este sub-conjunto se genera inmediatamente después de que el procedimiento de extracción y caracterización suministra la lista de segmentos de ambas imágenes. El sub-conjunto se representa en dos listas ordenadas en relación con la coordenada horizontal del punto medio de los segmentos, v_m .

El algoritmo de puesta en correspondencia empieza aplicando la regla de la ecuación (7) a aquellos segmentos verticales que satisfacen la condición $|v_{ms} - v_{mi}| < \Delta v$ pixels. Para reducir las listas de candidatos potenciales se definió una nueva regla (ecuación 8). Esta regla es necesaria puesto que un segmento de la imagen superior debe confrontarse con varios segmentos de la imagen inferior. Por consiguiente, esta regla permite reducir significativamente el número de parejas que deberán evaluarse según la regla (7).

Adicionar Z_k como elemento potencial de la clase w_j si

$$P_B(w_j | \mathbf{x}_k) \geq 0.5 \quad \text{ó} \quad P_N(w_j | \mathbf{x}_k) \geq 0.5 \quad (8)$$

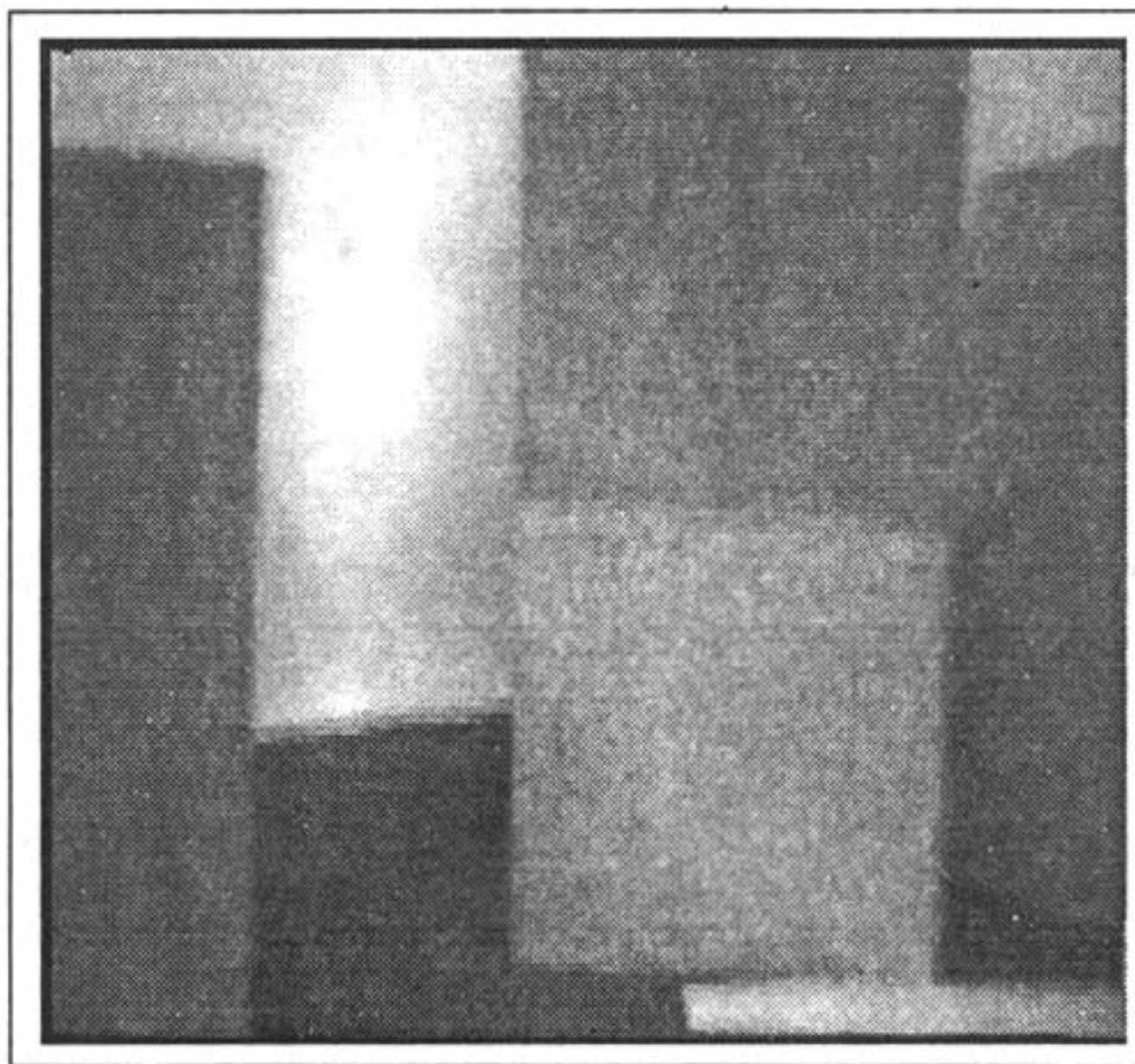
El hecho de imponer que una de las dos probabilidades sea superior o igual a 0.5 implica simplemente que al menos uno de los clasificadores considera este apareamiento como correcto. De otra parte, la relación «o» permite tener en cuenta posteriormente la *complementariedad* de los métodos de clasificación.

Durante la siguiente etapa se evalúan los apareamientos entre cada vecino del segmento vertical apareado de la imagen superior y todos los vecinos del segmento apareado de la imagen inferior. Los apareamientos que satisfacen la regla (8) conformarán la nueva lista de candidatos probables al apareamiento entre un segmento de la imagen superior y varios de la imagen inferior. El apareamiento retenido es aquel que satis-

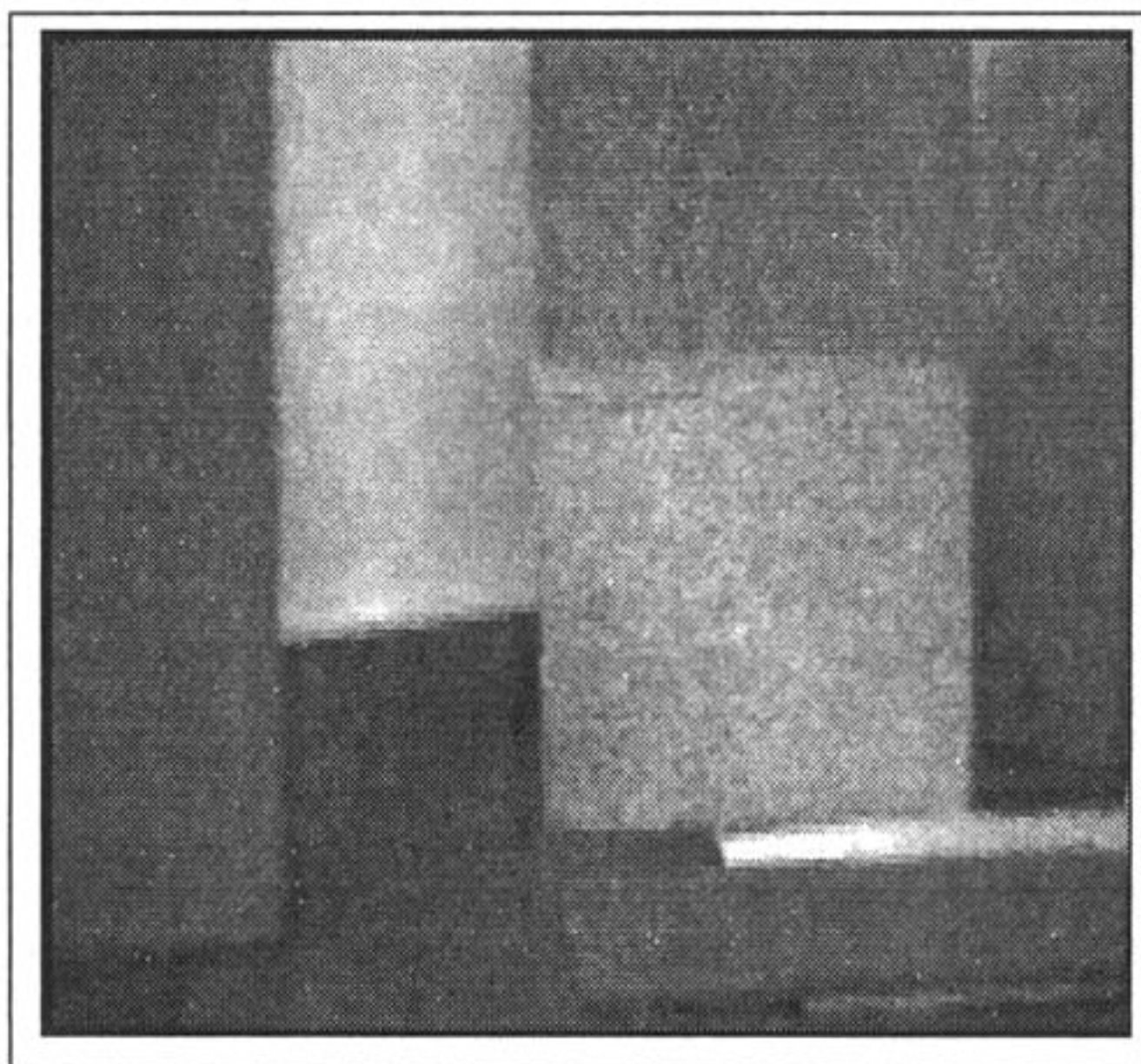
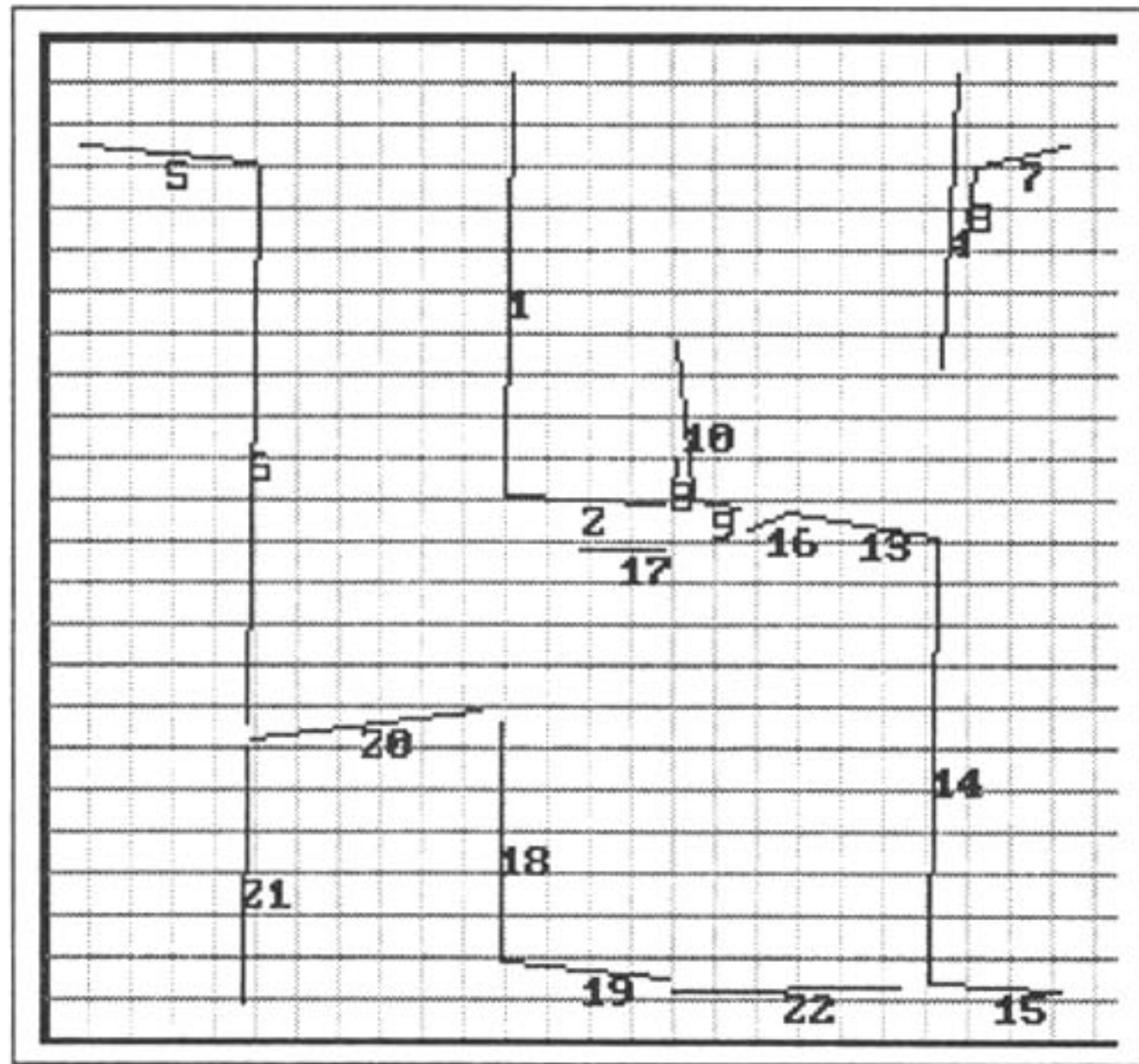
face la regla (7). El procedimiento anterior se repite luego de establecer una nueva correspondencia. El resultado final del algoritmo es una lista de segmentos con sus respectivas probabilidades a posteriori según los criterios neuronal y bayesiano.

En el caso de no detectarse segmentos verticales, es posible inicializar el algoritmo con segmentos horizontales. En el peor de los casos, cuando no se detecten ni segmentos verticales ni horizontales, puede usarse la restricción epipolar sólo para reducir el área de búsqueda de candidatos potenciales.

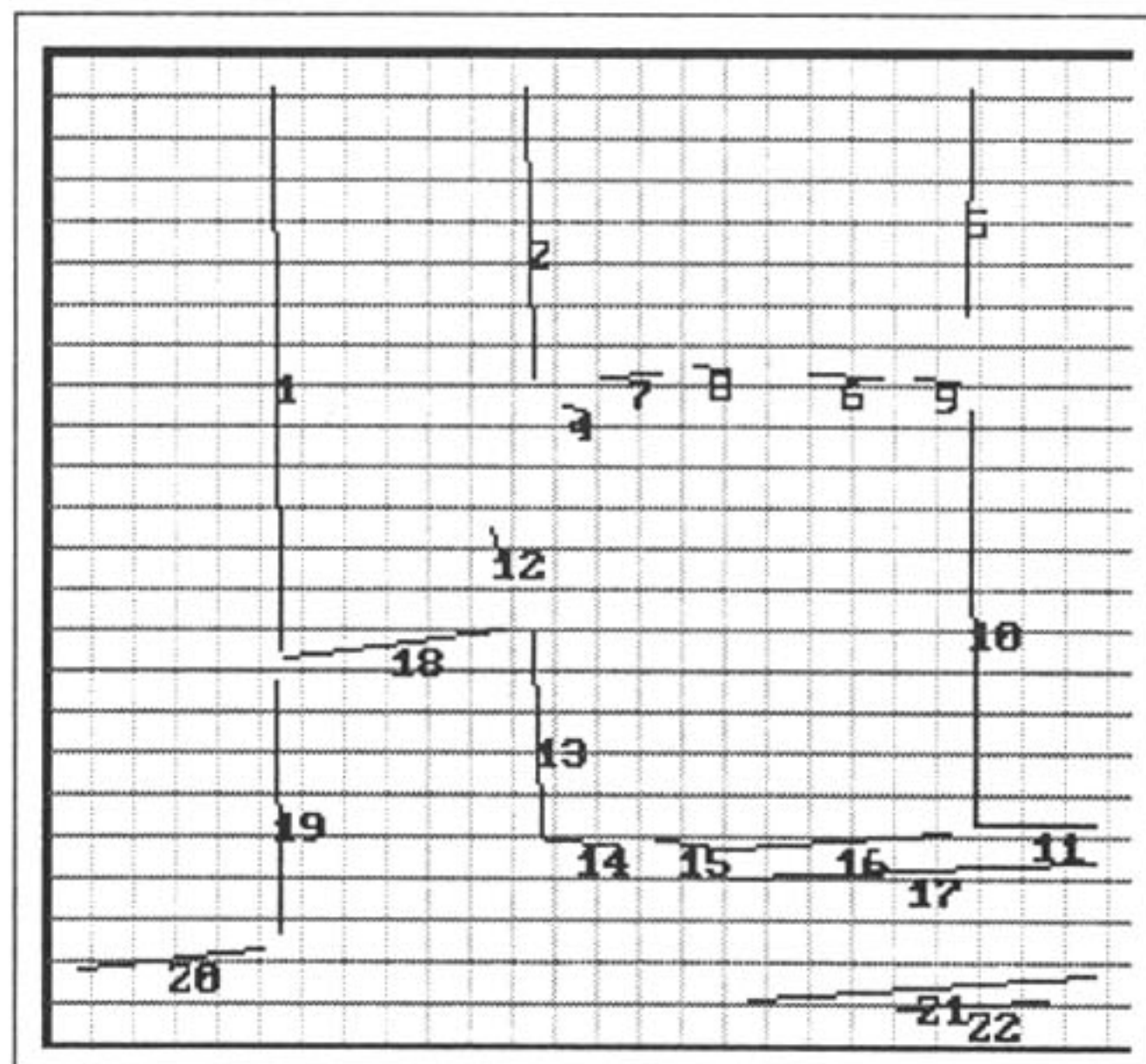
4. RESULTADOS



(a)



(b)



(d)

Figura 6.a,b). Pareja de imágenes tomadas por el sensor estéreo; c,d) Imágenes de síntesis.

Las figuras 6a y 6b presentan un par de imágenes tomadas por el sensor estéreo luego de la etapa de mejoramiento. Los segmentos detectados en estas imágenes se presentan en las imágenes de síntesis de las figuras 6c y 6d. La lista de segmentos apareados obtenida por el algoritmo se muestra en la tabla 5. Sobre este sencillo ejemplo, se alcanza más del 90% de buenos apareamientos. Sólo la pareja (22,17) corresponde a un mal apareamiento. Las imágenes muestran que el segmento 22 puede aparearse con el segmento 16. La tabla 6 muestra las probabilidades arrojadas por los dos clasificadores. Aquí puede apreciarse que la decisión tomada – pareja (22,17) – no difiere significativamente de la decisión correcta - pareja (22,16)-.

Sobre las imágenes del ejemplo existen otros segmentos que pueden ser apareados, pero estos no se tuvieron en cuenta debido a su corta longitud.

La tabla 7 presenta el conglomerado de los resultados obtenidos sobre otras 10 imágenes estereoscópicas. Se aprecia que la tasa de buenos apareamientos supera el 95 %. Los autores consideran que es posible alcanzar mejores resultados si se realiza un estudio exhaustivo sobre las parejas mal apareadas, y se analiza para cada caso la respuesta individual de los clasificadores. Lo anterior podría conducir, por ejemplo, a la introducción de coeficientes de ponderación en la fusión de los clasificadores. Igualmente se considera que un estudio detallado podría mostrar que los clasificadores responden mejor con diferentes conjuntos de variables. Como podrá recordarse, en este trabajo se utiliza el mismo vector para los dos métodos.

5. CONCLUSIONES Y PERSPECTIVAS

Con el desarrollo del presente trabajo se ha demostrado que los algoritmos implementados son suficientemente robustos y rápidos para hacer posible la implementación de un sensor estereoscópico a partir de cámaras de visión con características fuertemente degradadas. Estos algoritmos cubren todas las etapas desde el pre-procesamiento de imágenes, la detección y caracterización de segmentos, hasta el establecimiento de correspondencia. En este artículo sólo se hace hincapié en el problema del apareamiento. Inicialmente se calcula un número importante de características sobre cada segmento, con las cuales se construyeron vectores de 16 variables que fueron usados en la primera etapa de clasificación automática. Luego se realizó un estudio exhaustivo sobre el poder discriminante de cada característica y sobre la eficiencia de diferentes métodos de clasificación. Este estudio arrojó como resultado la posibilidad de combinar los métodos Bayesiano y Neuronal para lograr índices de eficiencia de clasificación altamente satisfactorios. La estrategia desarrollada se basa en la

Mal apareado	Segmentos apareados		Probabilidad a posteriori	
	Imagen Sup.	Imagen Inf.	Bayes	Neuronal
	1	2	0,999	0,945
	4	5	0,970	0,992
	6	1	0,812	0,982
	20	18	0,884	0,948
	21	19	0,801	0,965
	14	10	0,772	0,965
	13	9	0,934	0,970
	15	11	0,872	0,966
X	22	17	0,999	0,963
	19	14	0,989	0,965
	18	13	0,515	0,992

Tabla 5. Resultados de apareamiento sobre las imágenes 6c y 6d.

Segmentos	P. Bayes	P. Neuronal	Sum
22-16	0,915	0,982	1,897
22-17	0,999	0,963	1,962

Tabla 6. Parejas candidatas al apareamiento

Test	Número de apareamientos			Clasificación %	
	Potencial	Correcto	Incorrecto	Correcto	Incorrecto
1	22	18	1	94,74	5,26
2	26	25	0	100,00	0,00
3	30	26	2	92,86	7,14
4	45	38	3	92,68	7,32
5	29	26	0	100,00	0,00
6	24	22	1	95,65	4,35
7	31	27	1	96,43	3,57
8	24	23	0	100,00	0,00
9	37	34	2	94,44	5,56
10	43	36	4	90,00	10,00

Tabla 7. Resultados de aparear 10 juegos de imágenes estereoscópicas.

complementaridad de los espacios de los patrones sobre los cuales actúa eficientemente cada clasificador. Las pruebas realizadas sobre ambientes reales alcanzan valores superiores al 90% de buena clasificación con un índice muy bajo en las malas clasificaciones. La cantidad importante de segmentos apareados correctamente permite efectuar la reconstrucción 3D de los objetos del ambiente, lo que permite a un robot móvil efectuar las tareas de localización y navegación sin ambigüedad.

La continuación de este trabajo se enfocará en la reducción del espacio de candidatos para el establecimiento de correspondencia, para ello se puede explorar aún más la geometría particular del sensor estéreo. Igualmente, se considera que puede aumentarse el porcentaje de buena clasificación sobre el total de segmentos existentes en cada par de imágenes y reducir a cero el porcentaje de mala clasificación. Esto podría lograrse con la implementación de otras reglas

de combinación, y/o la utilización de vectores de parámetros diferentes para cada clasificador, y/o calcular nuevas características de los segmentos como por ejemplo parámetros derivados de la información en color.

6. AGRADECIMIENTOS

Este trabajo ha sido elaborado gracias al apoyo de Colciencias, la Universidad del Valle y la Université d'Evry.

7. BIBLIOGRAFÍA

- [1.] H. LOAIZA, A. SUAREZ, E. GONZALEZ, S. LELANDAIS, C. MORENO. Real MagiCol: Complex Behavior through simpler Behavior Oriented Commands. Proceedings of the second Robocup Workshop. Robocup-98. Pg.475-482. Paris, July 1998.
- [2.] N. AYACHE, O. D. FAUGERAS, Maintening Representation of the Environnement of a Mobile Robot. Proceédings IEEE of the International Symposium and Exposition on Robots, Esprit Projet, 1989.
- [3.] R. K. LENZ, R. Y. TSAI. Techniques for calibration of the scale factor and image center for high accuracy 3-D machine vision metrology. IEEE Transactions on PAMI, vol. 10, No. 5. September, 1988.
- [4.] W. I. GROSKEY, L. A. TAMBURINO. A unified approach to the linear camera calibration problem. IEEE Trans. on PAMI. vol.12, no. 7, Pg.663-671. July, 1990.
- [5.] H. LOAIZA, Capteur de vision stéréoscopique multiconfiguration pour la localisation d'un robot mobile. Thesis on Robotic, University of Evry, France, November 1999.
- [6.] O. FAUGERAS. Three-Dimensional Computer Vision. MIT Press, Massachusetts, 1996.
- [7.] R. DERICHE. Using Canny's criteria to derive an optimal edge detector recursively implemented. The International Journal of Computer Vision, Pg.167-187, 1987.
- [8.] H. LOAIZA, J. TRIBOULET, S. LELANDAIS, et al. A Multi-configuration Stereoscopic Vision System for Domestic Mobile Robot Localization. Robot Motion and Control , pg. 207-212. Poland, Juin 1999.
- [9.] J. F. CANNY., A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, Pg.679-698, 1986.
- [10] R. M. HARALICK, K. SHANMUGAM, I. DINSTEN. Textural features for image classification. IEEE Trans. on Systems, Man, and Cybernetics. vol. smc-3, no.6, Pg.610-621. November. 1973.

- [11] T. ASANO, M. EDAHIRO, H. IMAI, K. MUROTA, Practical use of bucketing techniques in computational geometry. Computational Geometry. G. Toussaint Editor. North-Holland, Pg.153-195, 1985.
- [12] L. HOLMSTRÖM, P. KOISTINEN, J. LAAKSONEN, E. OJA, Neural and Statistical Classifiers – Taxonomy and Two Case Studies. IEEE Transaction on Neural Networks, Volume 8, Number 1, January 1997.
- [13] C. BARAT, H. LOAIZA, E. COLLE, S. LELANDAIS, Neural and statistical classifiers. Can these approaches be complementary ? IEEE Instrumentation and Measurement Technology Conference, IMCT 2000, Baltimore, Pg. 1480-1486. Maryland, USA, May, 2000.
- [14] B.D. RIPPLEY, Pattern Recognition and Neural Networks. Cambridge University Press, 1996.
- [15] I.E. DROR, M. ZAGAESKI, C.F. Moss, Three-dimensional target recognition via sonar: a neural network model. Neural Networks, Vol. 8, No. 1, Pg.149-160, 1995.
- [16] C. BISHOP, Neural networks for pattern recognition Clarendon Press.Oxford, 1995.
- [17] J.S. BRIDLE, Probabilistic interpretation of feedforward classification network outputs, with relationships to statistical pattern recognition. Neuro-computing: Algorithms, Architectures and Applications, Berlin: Springer, Pg.227-236, 1990.
- [18] J. KITTLER, M. HATEF, R. DUIN, J. MATAS. On combining classifiers. IEEE Transactions on PAMI, vol.20,no.3,Pg.226-238.1998. ■

AUTORES



Humberto Loaiza Correa. Ph.D.

Ingeniero Electricista y Magíster en Automática de la Universidad del Valle. Ph.D. en Robótica y Visión Artificial de L'Université d'Evry, France. Profesor Titular del Area de Informática Industrial de la Escuela de Ingeniería Eléctrica y Electrónica de la Universidad del Valle. Dentro del grupo de investigación Percepción y Sistemas Inteligentes trabaja en las líneas de procesamiento de imágenes y señales, percepción inteligente y robótica móvil.

c.e: hloaiza@eiee.univalle.edu.co

Sylvie Lelandais. Ph.D.

Maître de Conférences.
Laboratoire Systèmes Complexes.
Université d'Evry. France.