

RECONOCIMIENTO DE OBJETOS DE FORMA LIBRE A PARTIR DE LOS DATOS DE RANGO DE UNA VISTA PARCIAL USANDO CONO CURVATURAS PONDERADAS¹

Santiago Salamanca*, Antonio Adán**,
Carlos Cerrada***, Miguel Adán****,
Pilar Merchán*, Emiliano Pérez***

* *Escuela de Ingenierías Industriales. UEx, {ssalaman, pmerchan}@unex.es*

** *Escuela de Ingeniería Informática, UCLM, Antonio.Adan@uclm.es*

*** *Escuela Técnica Superior de Ingeniería Informática, UNED, {ccerrada, emiliano}@issi.uned.es*

**** *Escuela Universitaria de Ingenieros Técnicos Agrónomos, UCLM, Miguel.Adan@uclm.es*

Resumen: En este trabajo se muestra un método para el reconocimiento de objetos de forma libre a partir de los datos de rango de una vista parcial. Para ello presentamos una nueva característica que denominamos Cono-Curvatura Ponderada (WCC), que mide el ángulo que sustenta cada uno de los nodos de la malla. Debido a que la información empleada no es completa, es necesario reforzar la búsqueda empleando una técnica basada en componentes principales, así como una basada en el algoritmo ICP. Debido a estas dos últimas fases del algoritmo, además del reconocimiento, se soluciona el problema de localización.
Copyright ©2007 CEA-IFAC

Palabras clave: Reconocimiento, posicionamiento, datos de rango, visión 3D.

1. INTRODUCCIÓN

El reconocimiento de objetos 3D a partir de información visual es una de las tareas centrales en inteligencia artificial. Su resolución tiene una gran utilidad para aplicaciones prácticas como inspección visual, guiado de robots, control de calidad ...

Reconocer un objeto supone identificarlo entre un conjunto de candidatos. Por otro lado, posicionar-

lo o alinearlo conlleva determinar en qué posición se encuentra con respecto a un modelo de referencia almacenado en una base de datos. Aunque conceptualmente son problemas distintos, en la práctica están muy relacionados: si se consigue alinear un objeto incógnita sobre uno de los objetos candidatos de la base de datos, se habrá reconocido. Esto es especialmente útil cuando del objeto a reconocer sólo disponemos una vista.

Para la resolución del problema de reconocimiento 3D se pueden utilizar distintas representaciones de los objetos. Todas ellas se pueden agrupar en alguna de las siguientes categorías: representación

¹ Este trabajo se ha desarrollado gracias a la financiación del Ministerio de Educación y Ciencia, a través de los proyectos de la CICYT DPI2005-03769 y DPI2006-14794-C02.

centrada en vistas o en apariencia y representación centrada en el objeto.

En la primera, el objeto es descrito a partir de distintas imágenes de intensidad de éste, asociada cada una de ellas a una vista. En esta categoría encontramos distintos métodos de representación como los grafos de aspecto, las técnicas basadas en siluetas y, sobre todo, los basados en componentes principales (Campbell and Flynn 1999, Skočaj and Leonardis 2001).

La representación basada en el objeto describe la superficie o el volumen de éste usando características destacadas referidas a un sistema de coordenadas propio. Cuatro grandes grupos se pueden incluir en esta categoría: representaciones basadas en curvas y contornos, representaciones axiales, representaciones volumétricas y representaciones superficiales. Dentro de esta última categoría se encuentra el método presentado en este artículo, por lo que, a continuación, se presentarán con detalle algunas de las representaciones más significativas.

Las firmas de puntos (*Point Signatures*) (Chua and Jarvis 1997) es un procedimiento similar al splash (Stein and Medioni 1992), pero en vez emplear la relación entre las superficies normales de los puntos centrales y sus vecinos, utiliza directamente el conjunto de puntos, eliminando así el cálculo de la primera derivada de las superficies normales. Este método presenta algunos problemas, como es el no poder ser empleado en objetos poliédricos. Por otra parte, en ningún momento se tiene un modelo del objeto completo, por lo que no es posible el reconocimiento de una vista parcial arbitraria.

Las imágenes de revolución (*Spin Images*) (Johnson and Hebert 1999) se definen a través de una colección densa de puntos 3D y sus normales. Cada punto tiene asociado una imagen que describe su entorno cercano. Es, sin duda, la representación más popular de los últimos tiempos. No obstante, existen una serie de aspectos que suponen limitaciones. En primer lugar, se necesita que la resolución de todas las mallas implicadas en el proceso sea la misma. Por otra parte, son muy sensibles al ruido y requieren un espacio de almacenamiento considerable.

Las imágenes de formas armónicas (*Harmonic Shapes Images*) (Zhang and Hebert 1999) representan a un objeto 3D mediante una imagen 2D obtenida mediante el mapeo de cada faceta de la malla que representa al objeto. Para ello utilizan los mapas armónicos. Esta representación no supone un sistema completo de reconocimiento y los autores no han desarrollado un método que fije automáticamente los umbrales en el proceso de acoplamiento.

Las firmas de superficie (*Surface Signature*) (Yamany and Farag 2002) es otra representación en la que se emplean características 2D. Éstas son definidas mediante una imagen en la que en cada píxel se codifica la curvatura, medida como ángulo simplex (Hebert *et al.* 1995), de cada uno de los puntos 3D de la malla. Es un método fácil y robusto, sin embargo presenta algún problema. En primer lugar, es necesario seleccionar unos puntos de interés de antemano, que son aquellos que tengan una curvatura elevada en torno a ellos, para poder generar la firma correspondiente, lo que limita el método a objetos no poliédricos. Además, las escenas que se pueden reconocer son aquellas que están totalmente reconstruidas, es decir, no es posible reconocer en aquellos casos en los que disponemos de una vista parcial.

Recientemente han sido definidas las imágenes de gradiente de profundidad basadas en siluetas (*DGI-BS*) (Merchán *et al.* 2006a). En ellas se sintetizan los gradientes de profundidad de un escena según determinadas direcciones establecidas a través de los contornos de la escena que no son debidos a oclusiones de otros objetos. Resulta ser una técnica eficaz en entornos de escenas complejas, pero necesita un proceso previo de segmentación de la escena, que no siempre es posible.

Como se indicó al comienzo de este apartado, el reconocimiento está íntimamente ligado a la localización o posicionamiento y, así, en los métodos anteriormente descritos, tras la medida de similitud con dichas representaciones, se aplica una fase de localización mediante la resolución del algoritmo *Iterative Closest Point (ICP)* (Besl and McKay 1992). En algunos casos, como en (Johnson and Hebert 1999), de forma previa al ICP, se realizan medidas de consistencia geométrica. Es, por tanto, el resultado del acoplamiento el que confirma cuál es el candidato de la base de objetos buscado.

El presente artículo muestra un método de reconocimiento de objetos de forma libre, incluidos poliédricos, a partir de los datos de rango 3D de una única vista parcial, mediante el empleo de una representación denominada *Cono-Curvatura Ponderada (WCC)*. Ésta se deriva de una representación previa definida en (Adán and Adán 2004) llamada *Cono-Curvatura (CC)* y que ha sido empleada con éxito para el reconocimiento de objetos completos. Con respecto a los procedimientos anteriormente descritos presenta la ventaja de que es una característica que representa de forma muy compacta la información de forma de los objetos a reconocer. Para reforzar la búsqueda, tras el empleo de las WCC, se aplica un método basado en las componentes principales y, por último, se utiliza el algoritmo ICP que es el encargado de confirmar cuál es el objeto reconocido.

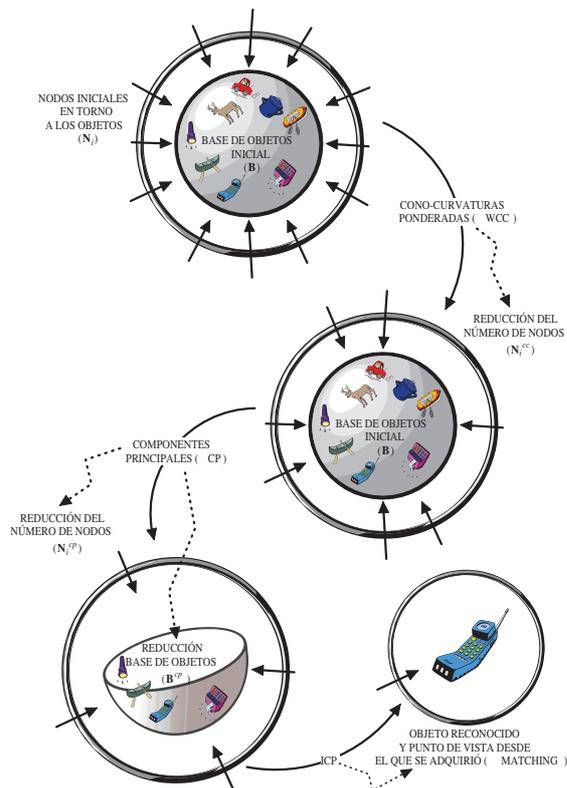


Figura 1. Esquema de los distintos pasos seguidos para el reconocimiento de vistas parciales.

La estructura que seguida a lo largo del trabajo es la siguiente: en la sección 2 se describe de forma general el método propuesto. En la sección 3 se definen las Cono-Curvaturas Ponderadas, empleadas en la primera fase del algoritmo. A continuación, en la sección 4, se proponen los métodos asociados a las dos últimas fases del algoritmo, que buscan el acoplamiento de la vista parcial sobre los objetos de la base de objetos. En la sección 5 se expone, a modo de resumen, la integración de las tres fases del algoritmo. Por último, en la sección 6 se muestran los resultados experimentales y en la sección 7 se exponen las conclusiones y futuras líneas de investigación.

2. MÉTODO PLANTEADO: PRINCIPIO DE FUNCIONAMIENTO

El método de reconocimiento 3D descrito en el presente artículo emplea un conjunto de características determinadas sobre las mallas esféricas y los datos de rango de las vistas parciales de los objetos que se quieren reconocer (figura 1)

En primer lugar se utiliza un nuevo invariante de forma que se denomina *Cono-Curvatura Ponderada (WCC)*, derivado de la *Cono-Curvatura (CC)* (Adán and Adán 2004), y que permitirá determinar una primera aproximación a los posibles ejes de visión desde los que se ha adquirido la vista parcial del objeto incógnita. La discretización

del espacio de visión se obtiene circunscribiendo una malla esférica en torno al modelo del objeto completo (Horn 1984). Cada nodo de esta malla, junto con el origen de coordenadas, definen los ejes de visión iniciales en torno a dicho modelo. Por tanto, determinar los ejes de visión desde los que se ha podido adquirir la vista parcial es equivalente a seleccionar un conjunto de nodos sobre la malla. Es importante tener en cuenta que al realizar esta selección lo que se está haciendo, de forma implícita, es una reducción de las posibles rotaciones que podrían aplicarse sobre la vista parcial para acoplarla sobre el modelo del objeto completo.

Se denotarán a los nodos obtenidos tras este primer paso como $N_i^{cc} \subset N_i$, donde N_i son los nodos iniciales de la malla esférica circunscrita sobre el i -ésimo objeto de la base de datos. Como se deduce de la explicación, en este paso no se reduce el número de modelos de la base de objetos.

Sobre los nodos seleccionados se aplicará otro invariante basado en las componentes principales (autovalores + autovectores) de los datos de rango de la vista parcial y de los modelos completos. Tras esta comparación se creará una lista para cada uno de los objetos de la base de datos con los nodos N_i^{cc} ordenados según el error existente en la comparación de los autovalores. A su vez, esta ordenación permite identificar cuál es el objeto sobre el que existe mayor probabilidad de que la vista parcial pueda ser acoplada. Al finalizar esta segunda fase se obtiene una reducción de los modelos de la base de objetos, así como una reducción de los nodos determinados en la fase anterior. Si se denota a la base de objetos inicial como \mathbf{B} , la obtenida al finalizar la comparación de los autovalores, será denominada $\mathbf{B}^{cp} \subset \mathbf{B}$, y a los nodos por cada objeto $N_i^{cp} \subset N_i^{cc}$.

Los autovectores permitirán realizar una primera aproximación a la rotación existente entre la vista parcial y cada uno de los objetos de \mathbf{B}^{cp} , que será utilizado en la última fase en la que se aplica el algoritmo *Iterative Closest Point (ICP)* (Besl and McKay 1992). Éste permite realizar el acoplamiento entre los datos de rango, y será el error de convergencia del algoritmo el que nos indique cuál es el objeto al que pertenece la vista parcial.

3. FASE DE DETERMINACIÓN ROBUSTA DEL PUNTO DE VISTA

En la primera fase del método de reconocimiento se estima cuáles son los posibles puntos de vista desde los que se ha podido adquirir la vista parcial. Para ello se propone una nueva característica calculada a partir de modelos esféricos parciales

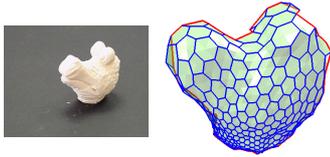


Figura 2. Imagen de intensidad de un objeto y modelo esférico para una vista parcial de éste.

de los datos de rango de dicha vista parcial. La técnica de modelado esférico parcial queda fuera del presente trabajo. Sólo comentar que tras ésta se obtiene una malla aproximada a los datos de rango cuyas teselas son de tipo hexagonal o pentagonal. Cada uno de los nodos de la malla tiene una conectividad tres excepto para aquellos nodos que se encuentran en el contorno que tienen conectividad menor. En la figura 2 se muestra la imagen de intensidad de un objeto y el modelo esférico para una vista parcial de éste. Para una revisión profunda del tema se recomienda la lectura de (Salamanca *et al.* 2000).

Sobre los modelos esféricos se puede crear una estructura denominada *Conjunto de Ondas de Modelado (MWS)* (Adán *et al.* 2000), que se genera a partir de una estructura más sencilla denominada *Onda de Modelado (MW)* definida de la siguiente forma:

Definición 1. Se llama onda de modelado con foco inicial N a una partición $\{F^1, F^2, \dots, F^q\}$ de la esfera teselada, T , que verifica:

$$\begin{aligned} F^1 &= \{N\} & (1) \\ &\vdots \\ F^{j+1} &= V(F^j) - \bigcup_{m=1}^{j-1} F^m \quad \forall j = 1, \dots, q-1 & (2) \end{aligned}$$

Donde V es la relación de vecindad de conectividad 3 impuesta por T .

Como se deduce de la definición 1 cualquier nodo de T puede ser foco y, por tanto, puede generarse a partir de él una MW. Por tanto, una esfera teselada con un número de nodos $n = \text{ord}(T)$ contendrá n ondas de modelado. Definimos por tanto al conjunto de ondas de modelado de la siguiente forma:

Definición 2. El conjunto de ondas de modelado se define como $MWS = \{MW^1, MW^2, \dots, MW^n\}$, donde MW^i es la onda de modelado generada a partir de un foco que se corresponde con el i -ésimo nodo de T .

Como se ha comentado en el apartado anterior, cada nodo de la malla esférica, junto con el ori-

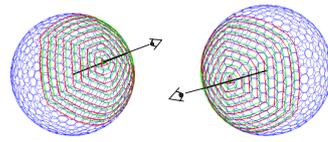


Figura 3. Representación de dos frentes de onda, asociados a dos nodos sobre la esfera teselada que definen dos ejes de visión.

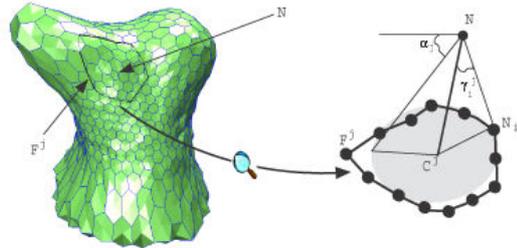


Figura 4. Definición de la j -ésima Cono Curvatura (CC).

gen de coordenadas, forma un eje de visión. Por tanto, para cada eje de visión existe una onda de modelado asociada. En la figura 3 se muestran dos ejemplos.

3.1 Cono-Curvatura

Debido a que la WCC es determinada a partir de una representación previa llamada *Cono-Curvatura (CC)*, pasaremos a continuación a describirla y analizar brevemente sus propiedades.

Es una representación que se calcula para cada nodo de la malla a partir del MWS, y cuya definición es la siguiente:

Definición 3. Sea N el foco inicial de una onda de modelado sobre el modelo esférico parcial, T' . Se llama Cono Curvatura j -ésima de N , al ángulo del cono con vértice N cuya superficie está aproximada al j -ésimo frente de ondas, F^j , de la onda de modelado asociada a N . Formalmente:

$$\alpha^j = \text{sign}(F^j) \cdot \left| \frac{\pi}{2} - \frac{1}{t_j} \sum_{i=1}^{t_j} \gamma_i^j \right| \quad (3)$$

donde $\gamma_i^j = \angle C^j N N_i \in F^j$, t_j es el número de nodos de F^j y C^j es el baricentro de F^j . El rango de valores de la CC es $[-\pi/2, \pi/2]$, donde el signo toma en cuenta la localización relativa de O , C^j y N , siendo O el origen de coordenadas del sistema de referencias fijo a T' . Valores negativos indican la existencia de zonas cóncavas, valores cercanos a cero se corresponden con superficies planas, y valores positivos con zonas convexas.

En la figura 4 se ilustra la definición de la CC.

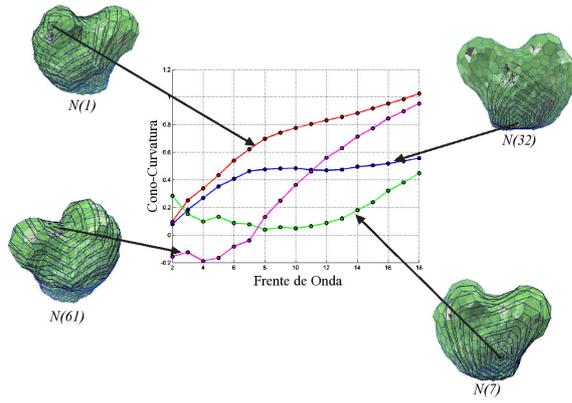


Figura 5. Representación de las cono-curvaturas para los primeros dieciocho frentes de los nodos $N(1)$, $N(7)$, $N(32)$ y $N(61)$.

Dado un foco N , existe un conjunto de frentes de onda, q , que definen las CC's para dicho foco $\{\alpha^1, \alpha^2, \dots, \alpha^q\}$. Éstas proporcionan una información completa de la curvatura del objeto existente desde el punto de vista que determina N . Decir por último que q , que denominamos *orden del frente*, puede tomar valores desde 2 (el caso $q = 1$ no tiene sentido) hasta el máximo número de frente que disponga el objeto. Al trabajar con modelos parciales el orden del frente será, como máximo, igual al número de frentes completos que disponga el modelo parcial.

A continuación se realizará un análisis de las CC's que permita sacar conclusiones con respecto a algunas propiedades fundamentales en cualquier representación empleada para tareas de reconocimiento: unicidad de la representación, invariancia frente a transformaciones afines (traslación, rotación y escalado) y robustez frente a la aparición de ruido.

En primer lugar representamos las CC's (eje de ordenadas) para cuatro focos, con respecto al frente de onda (eje de abscisas) para un objeto de forma libre (figura 5). Esta figura pone de manifiesto una de las características que deben verificar las CC's, que es que los valores de éstas se encuentren lo suficientemente alejados como para poder ser utilizados en tareas de reconocimiento.

En la figura 6 se ha coloreado la malla para representar los valores de la CC para distintos órdenes (2, 8, 14 y 16) sobre ésta. De esta representación se pueden extraer las siguientes conclusiones: la consistencia existente en la definición de las CC's, ya que, como se puede observar, existe una continuidad en los colores de la malla tanto para un mismo orden, es decir, entre nodos de una misma malla, como para distintos órdenes, es decir, entre los nodos de las cuatro mallas; y que la variación de color, equivalentemente de CC, es mucho menor en los frentes de orden bajo, de lo que se

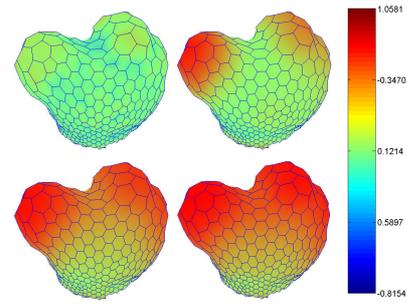


Figura 6. Representación en color de las CC de cada nodo de la malla para ordenes 2, 8, 14 y 16 (izquierda a derecha y de arriba hacia abajo). Se muestra también una barra en la que se puede visualizar los colores para las máximas, medias y mínimas CC del objeto.

puede deducir que éstos ofrecen una información más pobre que los de orden superior.

La demostración analítica de la invariancia de las CC's cuando el cuerpo sufre transformaciones afines es trivial. Según la ecuación (3) ésta se define como el valor medio del ángulo γ_i^j mostrado en la figura 4. Por el teorema del coseno,

$$\gamma_i^j = \arccos \left(\frac{|\overrightarrow{NN_i}|^2 + |\overrightarrow{NC^j}|^2 - |\overrightarrow{C^jN_i}|^2}{2|\overrightarrow{NN_i}||\overrightarrow{NC^j}|} \right) \quad (4)$$

en donde $|\overrightarrow{NN_i}|^2$ es el módulo al cuadrado del vector $\overrightarrow{NN_i}$ (equivalentemente para el resto de términos de la ecuación). Este término puede escribirse de forma matricial como:

$$|\overrightarrow{NN_i}|^2 = [\mathbf{N} - \mathbf{N}_i]^T \cdot [\mathbf{N} - \mathbf{N}_i] \quad (5)$$

donde $[\cdot]^T$ indica la trasposición de la matriz.

Si aplicamos una traslación, \mathbf{t} , rotación, \mathbf{R} , y escalado, α , sobre todos los nodos de la malla que modelan al objeto, tendremos que el módulo al cuadrado del vector $\overrightarrow{NN_i}$, que denotaremos como $|\overrightarrow{NN_i}|^2$, pasará a ser:

$$\begin{aligned} \overline{|\overrightarrow{NN_i}|^2} &= [\alpha \cdot (\mathbf{R} \cdot \mathbf{N} + \mathbf{t}) - \alpha \cdot (\mathbf{R} \cdot \mathbf{N}_i + \mathbf{t})]^T \cdot \\ & \quad [\alpha \cdot (\mathbf{R} \cdot \mathbf{N} + \mathbf{t}) - \alpha \cdot (\mathbf{R} \cdot \mathbf{N}_i + \mathbf{t})] \quad (6) \\ &= \alpha^2 \cdot [\mathbf{N} - \mathbf{N}_i]^T [\mathbf{N} - \mathbf{N}_i] \\ &= \alpha^2 \cdot |\overrightarrow{NN_i}|^2 \end{aligned}$$

De forma equivalente se modificarían el resto de parámetros de la ecuación (6), y ésta pasaría a ser:

$$\begin{aligned} \overline{\gamma_i^j} &= \arccos \left(\frac{\alpha^2 |\overrightarrow{NN_i}|^2 + \alpha^2 |\overrightarrow{NC^j}|^2 - \alpha^2 |\overrightarrow{C^jN_i}|^2}{2\alpha |\overrightarrow{NN_i}| \alpha |\overrightarrow{NC^j}|} \right) \\ &= \gamma_i^j \quad (7) \end{aligned}$$

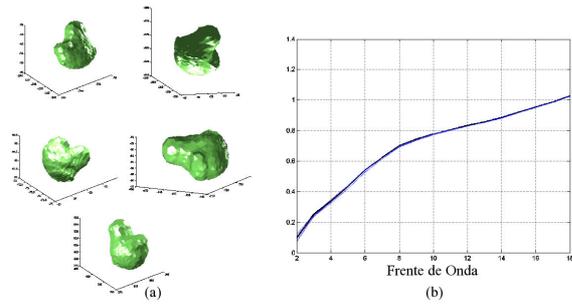


Figura 7. CC para cinco modelos de forma libre obtenidos mediante la rotación, traslación y escalado de los datos de rango originales. Se observa la invariancia existente en esta medida.

por lo que queda demostrada la invariancia de las CC's frente a transformaciones afines. En la figura 7 se muestra experimentalmente este mismo hecho. En (a) se encuentra representada la malla de un objeto sobre el que se han aplicado distintas transformaciones afines. En (b) se ha representado la CC con respecto al frente de onda para cada uno de los cinco casos. Se observa que en todos los casos éstas coinciden.

La última propiedad que vamos a analizar es la robustez de las CC's. En este caso se medirá cómo varían las CC's cuando sobre cada dato de rango se aplica un ruido de tipo aleatorio. Éste lo vamos a medir como un porcentaje con respecto al valor medio de la longitud de la aristas que forman la malla, d_m . De esta forma si, por ejemplo, se ha aplicado un ruido del 10 %, querrá decir que las coordenadas de cada nodo se habrán movido dentro de una esfera con origen en el nodo y radio $0,1d_m$. En la figura 8(a) se muestran, de izquierda a derecha y de arriba hacia abajo, las mallas sobre las que se han aplicado distintos niveles de ruido (0 %, 5 %, 10 %, 25 %, 50 % y 100 %), correspondiéndose la primera (0 %) con la original. El color en el que se dibujan las mallas se corresponde con el color de las gráficas en las que se muestran las CC's. En todas estas gráficas se puede apreciar que la robustez del invariante es mayor cuanto mayor es el orden del frente.

3.2 Cono-Curvatura Ponderada

Las CC se han aplicado con éxito al reconocimiento de objetos completos (Adán and Adán 2004), sin embargo su utilización en reconocimiento a partir de una única vista parcial es compleja debido a las siguientes causas:

1. El número de frentes de onda completos en una vista parcial es variable, lo que impide a priori realizar una comparación frente a frente de las CC de la vista parcial y del modelo completo.

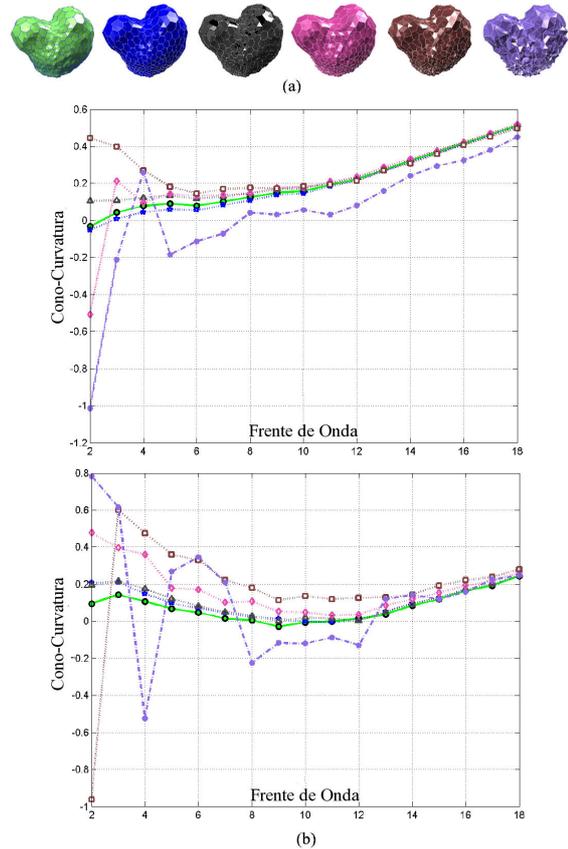


Figura 8. Representación de CC en objetos de forma libre para dos focos distintos ($N(1)$ y $N(32)$) y distintos niveles de ruidos en la malla (0 %, 5 %, 10 %, 25 %, 50 % y 100 %). El color empleado en cada una de las mallas es el mismo que en el de la representación de la CC.

2. La superficie representada por un determinado número de frentes de onda puede variar entre un modelo parcial y su modelo completo, por lo que frentes del mismo orden en ambos modelos podrán ser similares pero nunca iguales.
3. La longitud media de las aristas entre nodos es distinta para el modelo parcial y el modelo completo, lo que introduce errores en los valores de las CC.

Es por esto por lo que se propone una nueva característica, la *Cono-Curvatura Ponderada* (*WCC*), determinada a partir de la CC. En el apartado anterior se vio que la información suministrada por las CC's es similar entre frentes cercanos. Esto posibilitará utilizar una información más compacta derivada de las CC que reduzca su dimensionalidad, sustituyendo a cada vector $\{\alpha^1, \alpha^2, \dots, \alpha^q\}$ calculado desde $N' \in T'$, por un escalar que caracterizará al objeto desde ese nodo. Es a esta nueva característica a la que llamaremos *Cono-Curvatura Ponderada* (*WCC*).

Para evaluar cuantitativamente si es posible la reducción propuesta, se ha calculado la correlación

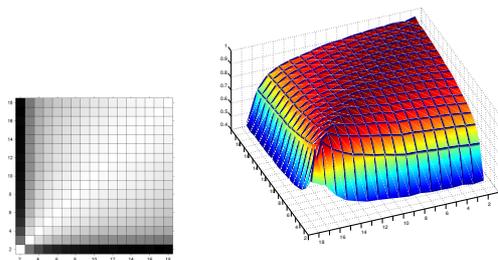


Figura 9. Ilustraciones de la correlación existente entre las CC de los frentes de onda de orden 2 a 18 de la base de datos.

existente en las CC's, que se ha calculado como el valor medio de todos los coeficientes de correlación de las mallas pertenecientes a la base de datos. Ésta se muestra a la izquierda de la figura 9 en niveles de gris (blanco para valores altos y negro para valores bajos), y a la derecha como una superficie 3D en el que el eje X e Y se corresponde con los frentes de onda y el eje Z con la correlación. Del análisis de las figuras, y como habíamos previsto, se puede concluir que existe alta correlación entre frentes de órdenes cercanos, la cual aumenta a medida que aumenta dicho orden.

Si denotamos a la WCC, para cada $N' \in T'$, como c^w , ésta se define mediante la siguiente combinación lineal:

$$c^w = \sum_{j=1}^q v^j \alpha^j \quad (8)$$

donde v^j son las coordenadas del autovector asociado al mayor autovalor de la matriz de covarianza para las q variables iniciales.

Este autovector ha sido determinado empíricamente mediante la evaluación de las componentes principales sobre las CC de todos los nodos de la malla. Con respecto a los órdenes considerados se han estudiado tres posibilidades:

1. Frentes de onda desde $q = 2$ hasta $q = 18$.
2. Frentes de onda desde $q = 4$ hasta $q = 18$.
3. Frentes de onda desde $q = 4$ hasta $q = 9$.

El primer caso se corresponde con los frentes de onda que se han estado considerando hasta este momento en los análisis de las CC's. En el segundo caso se han eliminado las CC asociadas a los dos primeros frentes de onda debido a su alta sensibilidad al ruido. Por último, en el tercer caso, se han eliminado, al igual que en el caso anterior, los dos primeros frentes de onda y se ha considerado un orden menor de frentes de onda para su mejor adecuación al reconocimiento de vistas parciales. En todos los casos, su cálculo se ha realizado de igual forma: determinación de los autovectores para todas las mallas de la base de

Tabla 1. Coordenadas de los autovectores calculadas para los casos (1), (2) y (3).

	v^1	v^2	v^3	v^4	v^5	v^6
(1)	0	1,50	2,63	3,50	4,41	5,15
(2)	0	0	0	3,62	4,58	5,35
(3)	0	0	0	11,79	14,52	16,51
	v^7	v^8	v^9	v^{10}	v^{11}	v^{12}
(1)	5,80	6,29	6,70	6,98	7,18	7,29
(2)	6,03	6,56	6,99	7,29	7,49	7,61
(3)	18,13	19,15	19,89	0	0	0
	v^{13}	v^{14}	v^{15}	v^{16}	v^{17}	v^{18}
(1)	7,34	7,31	7,23	7,08	6,90	6,67
(2)	7,67	7,64	7,55	7,40	7,21	6,98
(3)	0	0	0	0	0	0

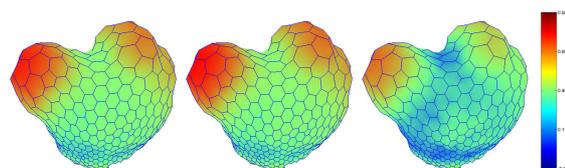


Figura 10. Representación en color de las WCC de cada nodo de la malla para los casos (1), (2) y (3) (izquierda a derecha). Se muestra una barra en la que se puede visualizar los colores para las máximas, medias y mínimas WCC del objeto.

datos, selección de aquel que tenga un autovalor mayor y cálculo del valor medio de todos los autovectores elegidos. En la tabla 1 se muestran las componentes de éstos tras la realización de estos pasos.

Para evaluar la bondad de la reducción de la dimensionalidad de las WCC es necesario realizar una análisis del porcentaje de la varianza total que se corresponde con las nuevas variables. Si existe un alto porcentaje las nuevas variables representarán satisfactoriamente a las variables iniciales y, por tanto, la reducción habrá tenido sentido. En nuestro caso se obtienen unos porcentajes en torno al 90% (89.70%, 92.42% y 95.34% para los casos (1), (2) y (3) respectivamente), lo que reafirma la utilidad de la reducción realizada.

En la figura 10 se ha representado sobre la propia malla y mediante código de colores, al igual que en la figura 6, las WCC. Podemos ver que las correspondientes al primer y segundo caso son muy similares.

3.3 CC's y WCC's aplicadas al reconocimiento de vistas parciales

La utilización de las WCC's en el proceso de reconocimiento implica una comparación de las almacenadas en la base de datos con las calculadas en la vista parcial a reconocer. Esta comparación implica una medida de error que a continuación definimos:

Definición 4. Sea $N' \in T'$ el nodo de la malla parcial T' que es el más cercano al eje de visión. Se define el error o distancia de comparación de WCC para cada $N \in T_i$, donde T_i es el i -ésimo modelo total de la base de objetos, como:

$$e_j(N) = \left| c^{w'}(N') - c^w(N_j) \right| \quad (9)$$

donde el subíndice j se extiende desde 1 hasta el máximo número de nodos existente en T_i y $|\cdot|$ representa el valor absoluto.

Condicionar la reducción de los nodos en torno a T_i a una única medida de error puede acarrear errores significativos en dicha reducción, por lo que, para reforzar la reducción, para cada $N \in T_i$ se medirán dos errores. El primero tendrá en cuenta las WCC hasta los frentes más alejados generados a partir de N' . A este error lo denominaremos *error en profundidad* y lo denotaremos como $e_j^p(N)$. El segundo tendrá en cuenta los frentes más cercanos generados a partir de N' . En este caso lo denominamos *error superficial* y lo denotamos como $e_j^s(N)$.

En ambos casos se obtiene un conjunto de errores igual al número de nodos existente en T_i , y a partir de ellos se determinarán los nodos, \mathbf{N}_i^{cc} , de la malla que serán pasados a la siguiente fase del algoritmo. Si denotamos como $\mathbf{N}^p = \{N_1^p, N_2^p, \dots, N_j^p\} \subset \mathbf{N}$ al conjunto de nodos de T_i con menor e_i^p y como $\mathbf{N}^s = \{N_1^s, N_2^s, \dots, N_j^s\} \subset \mathbf{N}$ al conjunto de nodos de T_i con menor e_i^s , \mathbf{N}_i^{cc} será $\mathbf{N}_i^{cc} = \mathbf{N}^p \cup \mathbf{N}^s$.

4. FASES DE COMPONENTES PRINCIPALES Y DE ICP

En este apartado se comentarán las dos últimas fases del algoritmo de reconocimiento, que buscan el acoplamiento de vistas parciales sobre los modelos completos (Salamanca *et al.* 2004). Éstas son las fases de componentes principales y de ICP.

4.1 Componentes principales

En la fase de componentes principales el método que se propone se basa en el cálculo de las componentes principales sobre los datos de rango que hayamos empleado en la obtención del modelo T' usado en el paso anterior. Las componentes principales se definen como los autovalores y autovectores $\{(\lambda_i, \bar{\mathbf{e}}_i) | i = 1, \dots, m\}$ de la matriz de covarianza de dichos datos de rango.

Los autovalores son invariantes frente a rotaciones y desplazamientos y los autovectores frente a desplazamientos, correspondiéndose estos últimos

con un sistema de referencia ligado a los propios datos. Ésto hace que se puedan emplear los primeros para evaluar qué parte de los datos de rango del modelo completo se corresponden con la escena, y los segundos para calcular una primera aproximación a la transformación que deben sufrir los datos parciales para ser acoplados sobre los totales. Ésta sólo reflejará la submatriz de rotación de la transformación total ya que los orígenes de los dos triedros coincidirán.

Para la aplicación de la técnica comentada es necesario evaluar, de forma previa al proceso de reconocimiento, las posibles vistas parciales existentes sobre los datos de rango del objeto completo. Para ello se ha discretizado el espacio de los posibles puntos de vista existentes en torno al objeto, y se ha desarrollado un método para la generación de *vistas parciales virtuales (VPV)* basado en el algoritmo *z-buffer* (Salamanca *et al.* 2004). De cada una de estas VPV se calcularán sus componentes principales que se emplearán en la fase de acoplamiento según lo expuesto anteriormente.

La comparación de los autovalores da información de cuáles son las posibles zonas en donde la vista parcial puede acoplarse, pero dicha información es de carácter global y, como hemos dicho, sólo da información de la rotación. De esta forma será necesario realizar una última fase de cálculo que refine el acoplamiento y, por tanto, nos permita calcular cuál es la transformación definitiva. Para ello emplearemos el algoritmo ICP sobre un número de posibles candidatos marcados en la fase de comparación de autovalores. El ICP debe partir de una aproximación a la transformación inicial, que en nuestro caso se corresponde con la dada en el acoplamiento de los autovectores. La zona donde se acoplará la vista, y la transformación definitiva será medida mediante el error de fin en el algoritmo ICP.

En la figura 11 se muestra un esquema del proceso de cálculo de componentes principales sobre los datos de rango del objeto completo y del proceso de acoplamiento.

La comparación de los autovalores de la vista parcial y de las vistas parciales virtuales se realiza mediante la medida de un índice de error que viene dado por la siguiente expresión:

$$e_i^{cp}(N) = \|\Lambda_i^v(N) - \Lambda^r\| \quad (10)$$

donde $N \in \mathbf{N}_i^{cc}$ es el nodo desde el que hemos generado la VPV, $\Lambda_i^v(N)$ es el vector formado por los autovalores de la VPV generada a partir del nodo N del i -ésimo objeto de la base de objetos ($i = 1, \dots, K$), $\Lambda^r = \{\lambda_1^r, \lambda_2^r, \lambda_3^r\}$ es el vector formado por los autovalores de la vista parcial real y $\|\cdot\|$ es la distancia euclídea.

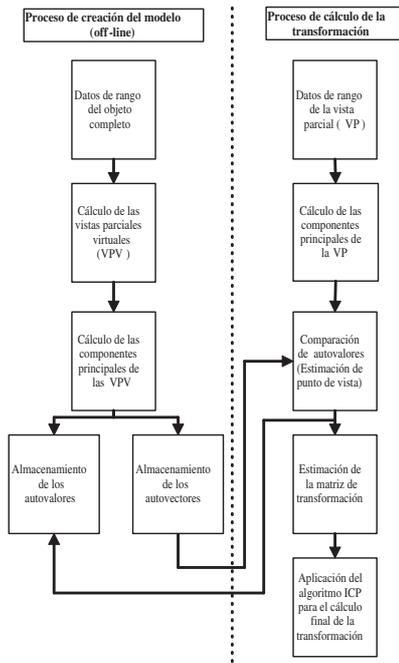


Figura 11. Diagrama de bloques del proceso propuesto para el acoplamiento de una vista parcial sobre los datos de rango del objeto completo.

Una vez calculado el error para todos los nodos \mathbf{N}_i^{cc} y todos los objetos de la base de objetos, obtenemos una lista de estos errores, e_i^{cp} ($i = 1, \dots, K$), ordenados de menor a mayor. Si comparamos el primer error (menor error para un determinado objeto) en todas las listas, se obtendrá, a su vez, una ordenación de los distintos objetos de la base de objetos. De esta forma, podremos aplicar en la última fase el algoritmo ICP sobre un subconjunto de la base de objetos, que denotaremos como \mathbf{B}^{cp} y para cada uno de ellos, utilizando las transformaciones asociadas al subconjunto de nodos \mathbf{N}^{cp} que han producido los errores comentados.

Para la resolución del ICP es necesario determinar una aproximación a la matriz de transformación entre la vista parcial y el objeto de la base de objetos \mathbf{R}_1 . Esta se calcula teniendo en cuenta que los autovectores son ortonormales, y por tanto:

$$\mathbf{R}_1 = \mathbf{E}^r (\mathbf{E}_i^v(N))^{-1} = \mathbf{E}^r (\mathbf{E}_i^v(N))^T \quad (11)$$

donde $\mathbf{E}_i^v(n)$ son los autovectores de la VPV generada a partir de $N \in \mathbf{N}^{cp}$ del i -ésimo modelo de la base de objetos \mathbf{B}^{cp} , y \mathbf{E}^r son los autovectores de la vista parcial.

4.2 ICP

El algoritmo ICP minimiza el error cuadrático medio:

$$e(k)^{icp} = \frac{1}{n} \sum \|\mathbf{Y} - \mathbf{Y}'\|^2 \quad (12)$$

donde \mathbf{Y}' son los datos de rango del objeto incógnita, denominados también *datos de la escena*, normalizados y transformados mediante la matriz homogénea \mathbf{T} , e \mathbf{Y} es el subconjunto de los datos de rango del objeto almacenado en la base de objetos, denominados *datos del modelo*, asociados con \mathbf{Y}' .

Inicialmente, la matriz \mathbf{T} viene dada por:

$$\mathbf{T} = \begin{bmatrix} \mathbf{R}_1 & \mathbf{t}_1 \\ \mathbf{0} & 1 \end{bmatrix} \quad (13)$$

donde \mathbf{R}_1 es la matriz de rotación calculada mediante la ecuación (11) y \mathbf{t}_1 es el vector de traslación definido por el centroide de los datos de rango del objeto completo.

Para el resto de pasos de iteración, denotamos a la matriz de transformación como \mathbf{T}_2 , en donde la submatriz de rotación, \mathbf{R}_2 , se determinará aplicando el método de Horn (Horn 1988):

$$\mathbf{R}_2 = \text{eigv} [\mathbf{Y} (\mathbf{Y}')^T] \quad (14)$$

donde $\text{eigv}(\cdot)$ son los eigenvalores de la matriz.

Por otro lado, el vector de traslación de \mathbf{T} se calculará de igual manera que para \mathbf{T}_1 , es decir, como la diferencia de centroides.

Este algoritmo, como hemos dicho en la sección previa, se aplicará sobre los objetos de \mathbf{B}^{cp} , y en cada caso tomando como matriz inicial de rotación la calculada para cada nodo de \mathbf{N}^{cp} . Tras esto se podrán comparar los errores de convergencia del ICP y aquel objeto en el que se verifique que dicho error es el menor, se corresponderá con el objeto reconocido por el algoritmo.

5. INTEGRACIÓN DE LOS MÉTODOS

Para terminar la descripción del método propuesto en este trabajo vamos a realizar una integración, a modo de resumen, de las distintas etapas por las que está formado. En la figura 12 se muestra un esquema de todo el proceso.

Como se puede observar en esta figura, los datos de entrada al algoritmo son, por un lado, el modelo y los datos de rango de la vista parcial y, por otro, la base de datos de objetos que almacenan los modelos de cada uno de ellos para cada punto de vista con los que se ha discretizado el espacio de visión en torno a ellos.

En la **fase 1** se toma la malla esférica parcial, T' , y se comparan las WCC según se explicó en la sección 3. Esta comparación determina de forma robusta los puntos de vista que tienen más probabilidad de ser los utilizados en la adquisición de la

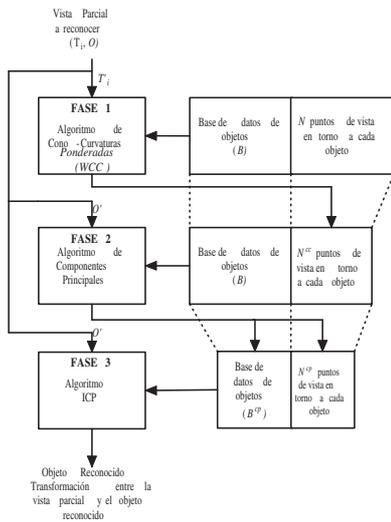


Figura 12. Esquema general de reconocimiento de vista parciales.

vista parcial, N_i^{cc} , aunque no habremos reducido los objetos que forman la base de datos.

En la **fase 2** tomaremos los datos de rango de la vista parcial y haremos una comparación de los autovalores obtenidos para las Vistas Parciales Virtuales generadas a partir de los nodos N_i^{cc} . Logramos de esta forma una reducción de candidatos tanto de la base de objetos, que denotamos como B^{cp} , como de los nodos que marcan las puntos de vista para cada uno de estos objetos, N_i^{cp} . Para cada uno de estos nodos candidatos se ha calculado una primera aproximación a la matriz de transformación basada en los autovectores de las componentes principales que, aplicada sobre los datos de rango de la vista parcial la acoplaría sobre los datos de rango del objeto completo.

En la **fase 3** aplicamos el algoritmo ICP sobre B^{cp} y tomando N_i^{cp} . Éste permite refinar el cálculo de las matrices de transformación estimadas en la fase anterior, y medir una distancia media, o error de proximidad, entre los datos de rango de la vista parcial y de la zona del objeto completo sobre la que iría acoplada que nos permite la verificación de reconocimiento definitiva.

6. RESULTADOS EXPERIMENTALES

El algoritmo propuesto en este trabajo se ha probado sobre una base de datos de 20 objetos (número habitual en este tipo de algoritmos (Merchán *et al.* 2006b)) adquiridos mediante el sensor de rango GRF-2 (Sato and Inokuchi 87). Estos objetos son de forma libre y poliédricos. La única restricción es que los objetos deben ser de genus 0, ya que no se pueden crear modelos esféricos de objetos de distinto genus.

El reconocimiento se realizó para tres vistas parciales por cada objeto, excepto en uno de ellos en el que tras su modelado parcial se vio que no disponía de frentes de onda suficientes como para poder calcular su WCC. Por tanto el número total de pruebas para la verificación del algoritmo ha sido 59.

Como hemos explicado, en primer lugar se comparan las CC del modelo parcial para un nodo que disponga de un máximo de número de frentes de onda. A partir de esta comparación se ha determinado N_i^{cc} , en donde el número de nodos que forman los conjuntos N_p (búsqueda en profundidad) y N_s (búsqueda superficial) es 32. Debido a que la malla utilizada para obtener el modelo completo, T_i ha sido de 1280, la reducción mínima que se puede obtener es del 95 %. Sobre la base de datos empleada, el valor medio de la reducción ha sido del 95,377 %. El tiempo medio de esta fase es de 7,95 s.

Partiendo de estos nodos se han comparado los autovalores, lo que nos permite hacer una reducción de la base de datos, que como ya hemos dicho, denotamos como B^{cp} , y de los nodos obtenidos en la fase anterior, que denotamos como N_i^{cp} . Se ha determinado experimentalmente que si B^{cp} está constituida por el 35 % aproximadamente de los objetos de B y N_i^{cp} por el 8 % aproximadamente de los nodos de N^{cc} por objeto, los resultados de reconocimiento son satisfactorios. El tiempo medio de ejecución de esta fase del algoritmo es de 0,86 s.

En la tabla 2 se muestran las tasas de aciertos en el reconocimiento tras la realización de las dos primeras fases del algoritmo. Ésta se ha creado realizando una lista ordenada de los objetos en función de los errores que se hayan obtenido. De esta forma, por ejemplo, en la primera fila se indica que tras las dos primeras fases, de las 59 pruebas realizadas, en 27 casos (segunda columna), que se corresponde con el 45,763 % (tercera columna) de los casos, el objeto que se encuentra en la primera posición de la lista, es el objeto que se quiere reconocer. Por último, en la cuarta columna de la tabla se indica el porcentaje acumulado. De estos datos se deduce que, tras esta fase del algoritmo, en un porcentaje relativamente alto (76,272 %) la coincidencia entre el objeto que se está intentando reconocer y el objeto de la base de datos se encuentra entre los tres mejores candidatos. Por otro lado, sólo existe un caso en el que el objeto buscado no se encuentra en B^{cp} .

Por último el algoritmo ICP se ha aplicado sobre los siete objetos (35 %) de la base de datos B^{cp} y tres nodos (8 %) por cada uno de ellos N_i^{cp} . La tasa de aciertos en el reconocimiento es del 90 %, lo que demuestra la validez del método. El grave

Tabla 2. Tasas de aciertos tras las dos primeras fases del algoritmo.

Posición	Casos	Porcentaje	Porcentaje acumulado
1	27	45,763 %	45,763 %
2	15	25,424 %	71,187 %
3	3	5,085 %	76,272 %
4	5	8,474 %	84,746 %
5	5	8,474 %	93,222 %
6	1	1,695 %	94,915 %
7	2	3,390 %	98,305 %

problema de esta fase es el tiempo de ejecución, ya que éste es, en término medio, de 9,73 m.

En la figura 13 se muestran algunos ejemplos de reconocimiento antes y después de la aplicación del algoritmo.

7. CONCLUSIONES

En este trabajo se ha presentado un método para el reconocimiento de objetos a partir de los datos de rango de sus vistas parciales. Éste se encuentra dividido en tres fases:

1. Fase de comparación de Cono-Curvaturas Ponderadas (WCC).
2. Fase de comparación de autovalores.
3. Fase de ICP.

En la primera se emplea una representación medida sobre el modelo esférico parcial de la vista, que hemos llamado Cono Curvatura Ponderada. Como hemos visto éstas presentan propiedades adecuadas para ser empleadas en tareas de reconocimiento. Su aplicación en la primera fase del algoritmo supone una reducción de los posibles ejes de visión desde los que se ha podido adquirir la vista parcial. La reducción que se obtiene es, como mínimo, del 95 %, que supone en la práctica una reducción muy significativa.

En las dos últimas fases del algoritmo se busca el candidato definitivo buscando el mejor acoplamiento de la vista parcial sobre los candidatos de la fase previa. Para ello, en primer lugar, se comparan los autovalores de las componentes principales de la vista parcial con los de vistas extraídas de los modelos completos, llamadas Vistas Parciales Virtuales, según los ejes de visión determinados en la fase de WCC's. Mediante esta fase reducimos nuevamente los nodos que definen los ejes de visión y la base de datos original, y calculamos, a partir de los autovectores de las componentes principales, una primera aproximación a la matriz de transformación entre la vista parcial y el objeto completo. Por último, en la fase final de algoritmo, se aplica el algoritmo ICP y, en función del error de convergencia de éste, se determina cuál es el objeto de la base de datos con el que se corresponde la vista parcial, así como la

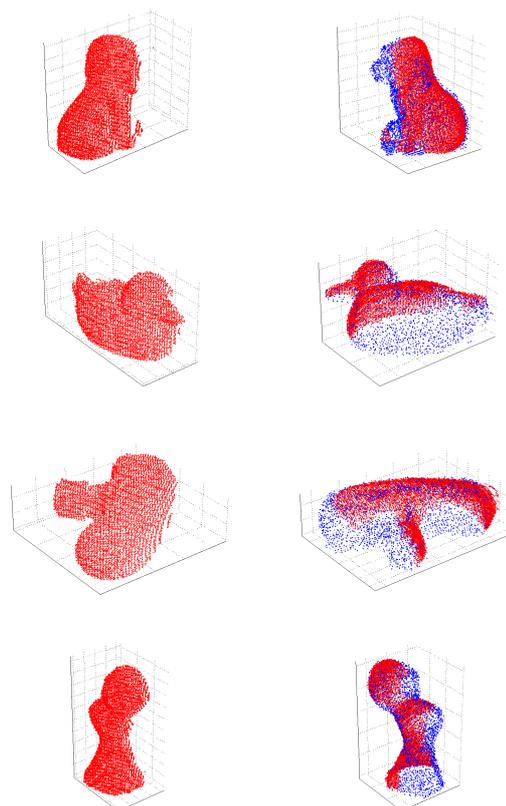


Figura 13. Resultados de reconocimiento sobre objetos de forma libre. En cada fila se muestra, a la izquierda, la vista parcial que se desea reconocer y, a la derecha, el resultado obtenido tras la aplicación del algoritmo.

transformación que resuelve la localización de la vista parcial.

Para demostrar la validez del método hemos trabajado con una base de datos de 20 objetos, que incluyen objetos de forma libre y poliédricos. La tasa de aciertos obtenida es del 90 %.

Como limitaciones del algoritmo podemos citar dos. En primer lugar, emplear modelos esféricos restringe la aplicabilidad del método ya que no permite el empleo de objetos de genus distinto a 0 y, sobre todo, el modelado en sí es bastante complejo. Actualmente estamos trabajando en la determinación de las CC's y WCC's sobre los modelos triangulares generados por los sensores de rango (Adán and Adán 2006). Por otro lado, el tiempo de ejecución del algoritmo es alto, lo que supone que no es posible su aplicación en sistemas de tiempo real. Hay que decir con respecto a este hecho que ha sido programado en matlab[®], lo que ralentiza significativamente su tiempo de ejecución. Por otro lado, todos los algoritmos que emplean representaciones como la que presentamos en este artículo, citados en la introducción, presentan esta limitación.

REFERENCIAS

- Adán, Antonio and Miguel Adán (2004). A flexible similarity measure for 3D shapes recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(11), 1507–1520.
- Adán, Antonio, Carlos Cerrada and Vicente Feliu (2000). Modeling wave set: Definition and application of a new topological organization of 3D object modeling. *Computer Vision and Image Understanding* **79**(2), 281–307.
- Adán, Miguel and Antonio Adán (2006). Cone curvature in triangular meshes: an adaptable curvature measure for 3d shapes. In: *International Congress of Mathematicians*.
- Besl, Paul.J. and N.D. McKay (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(2), 239–256.
- Campbell, Richard J. and Patrick J. Flynn (1999). Eigenshapes for 3D object recognition in range data. In: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. Vol. 2. Fort Collins, Colorado. pp. 2505–2510.
- Chua, Ching S. and Ray Jarvis (1997). Point signatures: A new representation for 3D object recognition. *International Journal of Computer Vision* **25**(1), 63–85.
- Hebert, M., K. Ikeuchi and H. Delingette (1995). A spherical representation for recognition of free-form surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**(7), 681–690.
- Horn, Berthold K.P. (1984). Extended gaussian images. *Proceeding of the IEEE* **72**(12), 1656–1678.
- Horn, Berthold K.P. (1988). Closed form solutions of absolute orientation using orthonormal matrices. *Journal of the Optical Society A* **5**(7), 1127–1135.
- Johnson, Andrew E. and Martial Hebert (1999). Using spin images for efficient object recognition in cluttered 3d scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(5), 433–449.
- Merchán, Pilar, Antonio Adán and Santiago Salamanca (2006a). Recognition of free-form object in complex scenes using DGI-BS models. In: *Proc. of the International Symposium on 3D Data, Processing, Visualization and Transmission*.
- Merchán, Pilar, Antonio Adán and Santiago Salamanca (2006b). Revisión y análisis comparativo de técnicas de reconocimiento en entornos tridimensionales. In: *Actas de las XXVII Jornadas de Automática*.
- Salamanca, Santiago, Carlos Cerrada and Antonio Adán (2000). HWM: a new spherical representation structure for modeling partial views of an object. In: *Proc. of the International Conference on Pattern Recognition (ICPR'2000)*. Vol. 3. Barcelona, Spain. pp. 770–773.
- Salamanca, Santiago, Carlos Cerrada, Antonio Adán and Miguel Adán (2004). Partial views matching using a method based on principal components. In: *Proc. of the First International Conference on Informatics in Control, Automation and Robotics (ICINCO 2004)* (Helder Araújo, Alves Vieira, José Braz, Bruno Encarnação and Marina Carvalho, Eds.). Vol. 2. Setubal (Portugal). pp. 3–10.
- Sato, K. and S. Inokuchi (87). Range-imaging system utilizing nematic liquid crystal mask. In: *International Conference on Computer Vision*. pp. 657–661.
- Skočaj, Danijel and Aleš Leonardis (2001). Robust recognition and pose determination of 3-D objects using range image in eigenspace approach. In: *Proc. of 3DIM'01*. pp. 171–178.
- Stein, Fridtjof and Gérard Medioni (1992). Structural indexing: Efficient 3-D object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14**(2), 125–145.
- Yamany, Sameh and Aly Farag (2002). Surfacing signatures: An orientation independent free-form surface representation scheme for the purpose of objects registration and matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(8), 1105–1120.
- Zhang, Dongmei and Martial Hebert (1999). Harmonic maps and their applications in surface matching. In: *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. II: 524–530.