

Descriptores de Fourier para identificación y posicionamiento de objetos en entornos 3D

Elizabeth González García
Departamento de Informática, Universidad de Matanzas,
Carretera Varadero Km 31/2, CP 40100, Matanzas, Cuba
elizabeth.gonzalez@umcc.cu

Vicente Feliú Batlle
E.T.S.I Industriales, Universidad de Castilla-La Mancha,
Ave. Camilo José Cela s/n, 13071, Ciudad Real, España
vicente.feliu@uclm.es

Antonio Adán Oliver
Escuela Superior de Informática, Universidad de Castilla-La Mancha,
Ronda de Calatrava 5, 13071 Ciudad Real, España
antonio.adan@uclm.es

Luis Sánchez Rodríguez
E.T.S.I Industriales, Universidad de Castilla-La Mancha,
Ave. Camilo José Cela s/n, 13071, Ciudad Real, España
luis.sanchez@uclm.es

Resumen

En este trabajo se presentan los resultados obtenidos en tareas de reconocimiento y posicionamiento espacial utilizando descriptores de Fourier sobre siluetas de objetos. El método propuesto está basado en la creación de modelos de siluetas extraídos de la imagen de modelos 3D sintetizados según un conjunto de puntos de vista prefijados. A través del rápido cálculo de parámetros en el espacio de Fourier, se logra resolver el problema de identificación de punto de vista, escalado y rotación en el eje óptico en tiempos muy reducidos. A su vez, se propone un método discriminación de descriptores que supone una reducción adicional. Este trabajo constituye una de las primeras etapas dentro de un proyecto general sobre estrategias de agarre de objetos guiadas por visión.

Palabras Clave: Visión por Computador
Descriptores de Fourier, Reconocimiento de
Objetos, Métodos Basados en Siluetas.

1. INTRODUCCION

El uso de sensores de visión acoplados a manipuladores robóticos proporciona una amplio

conjunto de posibilidades y aplicaciones en entornos industriales. Sin embargo, lograr aplicaciones flexibles y robustas que trabajen en tiempo real constituye, en estos momentos, el reto de muchos investigadores.

El objetivo de nuestra investigación es el desarrollo de un sistema integrado por una cámara de visión artificial y un robot con pinza de manera que éste sea capaz de agarrar automáticamente objetos de forma libre en movimiento.

Para lograr este objetivo el proyecto se divide en dos partes fundamentales. La primera etapa soluciona el problema de reconocimiento del objeto junto con su posición orientación y velocidad. La segunda etapa trata sobre planificación y generación de trayectorias del extremo del robot de modo que el objeto pueda ser agarrado de una forma eficiente.

En este trabajo presentamos los resultados parciales obtenidos durante el desarrollo del sistema de visión referentes al reconocimiento y posicionamiento del objeto respecto de un modelo almacenado utilizando técnicas basadas en siluetas. Es importante resaltar que, durante esta fase, es de vital importancia que el sistema trabaje lo más cercano posible al tiempo real.

Uno de los métodos de reconocimiento 2D a través de siluetas más populares son los basados en momentos invariantes y descriptores de Fourier [1,2,3]. Los momentos invariantes tienen como dificultad que dos siluetas completamente diferentes pueden tener los mismos momentos de bajo orden. En esos casos los descriptores de Fourier caracterizan mejor al contorno de un objeto pues sólo objetos con siluetas similares tienen descriptores similares.

El artículo está organizado de la siguiente manera: en la próxima sección se comentan brevemente los descriptores de Fourier y el efecto de las transformaciones espaciales en el espacio de Fourier. En la sección 3 se explica el procedimiento de reconocimiento de una silueta así como la determinación de su posición, orientación y escalado. En la sección 4 se aborda la fase experimental y los resultados obtenidos durante la misma. Finalmente en la sección 5 se enuncian las conclusiones y trabajos futuros.

2. DESCRIPTORES DE FOURIER Y TRANSFORMACIONES

Los descriptores de Fourier tienen como característica ser invariantes a transformaciones geométricas y tolerantes ante ruido [4].

Una curva cerrada puede ser representada mediante series de Fourier con una parametrización adecuada. Consideremos un contorno $x(n)$ constituido por N puntos en el plano XY :

$$x(n) = [x_n, y_n], n = 1..N \quad (1)$$

donde cada punto se puede considerar como un número complejo:

$$x(n) = x_n + jy_n \quad (2)$$

Empezando en un punto arbitrario del contorno y siguiéndolo alguna dirección de avance (por ej., a favor de las manecillas del reloj), se tiene una secuencia de números complejos. La transformada discreta de Fourier $X = F(x)$ de la secuencia se obtiene mediante la siguiente expresión:

$$X(k) = \sum_{n=0}^{N-1} x(n) \exp(-j2\pi kn/N) \quad 0 \leq k \leq N-1 \quad (3)$$

De esta forma los descriptores de Fourier pueden representar un contorno cerrado arbitrario. Esta representación tiene la ventaja de que,

generalmente, se logra una buena descripción en base a pocos términos.

Sea la base de datos de $nbase$ siluetas:

$$S_R(i), 1 \leq i \leq n_R \quad \text{con} \quad 1 \leq R \leq nbase$$

Cada silueta se regulariza a N puntos (N es potencia de 2):

$$S_R(i) \longrightarrow Z_R(n) \quad 1 \leq n \leq N$$

Supóngase que se tiene una silueta escena \hat{S} obtenida de una imagen y regularizada a N puntos:

$$\hat{S} \longrightarrow \hat{Z}(n) \quad 1 \leq n \leq N$$

El problema consiste en identificar \hat{Z} con alguna silueta Z_R bajo la hipótesis de que puede haber sufrido:

- 1) Desplazamiento en el origen de referencia del vector de la silueta: δ
- 2) Rotación respecto al centro de masas del objeto: θ
- 3) Escalado en la imagen: λ
- 4) Desplazamiento del centro de masas del objeto c_x, c_y en la coordenadas x - y respectivamente.

A continuación se analiza cuál es la transformación en la transformada de Fourier de una silueta bajo cada una de estas operaciones

1) Desplazamiento δ

Sea la secuencia:

$$x(n) \longrightarrow X(k)$$

$$= \sum_{n=0}^{N-1} x(n) \exp(-j2\pi kn/N) \quad 0 \leq k \leq N-1 \quad (4)$$

Se define la nueva secuencia

$$\hat{x}(n) = \begin{cases} \hat{x}(n) = x(n+\delta) & ; 0 \leq n \leq N-1-\delta \\ \hat{x}(n) = x(n-N+\delta) & ; N-\delta \leq n \leq N-1 \end{cases} \quad (5)$$

Sustituyendo en (4) se tiene

$$\hat{X}(k) = \sum_{n=0}^{N-1} \hat{x}(n) \exp(-j2\pi kn/N) \quad (6)$$

$$= \sum_{n=0}^{N-1-\delta} x(n+\delta) \exp(-j2\pi kn/N) + \sum_{n=N-\delta}^{N-1} x(n-N+\delta) \exp(-j2\pi kn/N) \quad (7)$$

Intercambiando coeficientes y reagrupando se obtiene que

$$\hat{X}(k) = X(k) \exp(j2\pi k\delta/N) \quad (8)$$

2) Rotación θ

Sea la secuencia $x(n)$, la nueva secuencia es

$$\hat{x}(n) = \exp(j\theta)x(n) \quad (9)$$

La relación entre transformadas de Fourier es:

$$\hat{X}(k) = \exp(j\theta)X(k) \quad (10)$$

3) Escalado λ

Sea la secuencia $x(n)$, la nueva secuencia es

$$\hat{x}(n) = \lambda x(n) \quad (11)$$

La relación entre transformadas es:

$$\hat{X}(k) = \lambda X(k) \quad (12)$$

4) Desplazamiento $c = c_x + jc_y$ (13)

Sea la secuencia $x(n)$, la nueva secuencia es

$$\hat{x}(n) = x(n) + c \quad (14)$$

La relación entre transformadas es:

$$\hat{X}(k) = X(k) + cN \quad (15)$$

3. RECONOCIMIENTO Y DETERMINACIÓN DE PARÁMETROS DE POSICIONAMIENTO

Supongamos una imagen del objeto que se quiere reconocer. Su silueta puede aparecer, con respecto a la almacenada en la base de datos z_l bajo transformaciones de: desplazamiento, rotación,

escalado y traslación de su c.d.m. Formalmente en el dominio espacial:

$$\hat{z}(k) = D(z_l, \delta)\lambda \exp(j\theta) + c \quad (16)$$

donde $D(z_l, \delta)$ es un desplazamiento δ de la silueta almacenada z_l en su origen de referencia, entonces la transformada de Fourier quedaría como:

$$\hat{Z}(k) = \lambda \exp(j\theta) \exp(j2\pi k\delta/N) Z_l + cN \quad (17)$$

donde:

$$\hat{Z} = F(z); Z_l = F(z_l)$$

- Cálculo del desplazamiento del c.d.m

Se supone que todas las siluetas de la base de datos tienen su centro geométrico en el origen. En este caso $Z_R(0) = 0$ y la expresión (17) se reduce a:

$$\hat{Z}(0) = cU(0) \Rightarrow c = \frac{\hat{Z}(0)}{N}$$

- Identificación de la silueta más próxima en la base de datos.

Se define $\hat{\hat{Z}}(k) = \hat{Z}(k) - cN$. Entonces

$$\hat{\hat{Z}}(k) = \lambda \exp(j\theta) \exp(j2\pi k\delta/N) Z_l(k)$$

Tomando módulos

$$\left| \hat{\hat{Z}}(k) \right| = \lambda |Z_l(k)| \quad (19)$$

A continuación, se optimiza λ utilizando la función de costo

$$f = \left(\left| \hat{\hat{Z}} \right| - \lambda |Z_l| \right)^t \left(\left| \hat{\hat{Z}} \right| - \lambda |Z_l| \right) \quad (20)$$

donde $\left| \hat{\hat{Z}} \right|$ es el vector columna que contiene las coordenadas de la silueta a identificar y $|Z_l|$ es otro vector columna que contiene las coordenadas de la silueta de la base de datos.

Haciendo:

$$\frac{\partial f}{\partial \lambda} = 0 = -2 \left(\widehat{Z} - \lambda |Z_l| \right)^t |Z_l|$$

$$\Rightarrow \left| \widehat{Z} \right|^t |Z_l| = |Z_l|^t |Z_l| \lambda$$

con lo que se obtiene:

$$\lambda_{opt} = \frac{\left| \widehat{Z} \right|^t |Z_l|}{|Z_l|^t |Z_l|} \quad (21)$$

$$f_{opt}(l) = \left| \widehat{Z} \right|^t \left| \widehat{Z} \right| - \frac{\left(\left| \widehat{Z} \right|^t |Z_l| \right)^2}{|Z_l|^t |Z_l|} \quad (22)$$

Una vez seleccionada la silueta optima l_{opt} (con transformada de Fourier Z_{opt}) que minimiza $f_{opt}(l)$ se pasa a calcular los parámetros λ, θ, δ

- Cálculo de los parámetros λ, θ, δ

Si denominamos $\lambda \exp(j\theta) = q$, el objetivo es minimizar la función de costo:

$$f(\delta) = (\widehat{Z} - q\Lambda(\delta)Z_{opt})^* (\widehat{Z} - q\Lambda(\delta)Z_{opt}) \quad (23)$$

donde:

$$\Lambda(\delta) = \text{diag} (1, \exp(j2\pi\delta / N), \exp(j2\pi 2\delta / N), \dots, \exp(j2\pi(N-1)\delta / N),)$$

Si se optimiza respecto a $q = q_2 + jq_1$ se llega a:

$$q_{opt} = \frac{\widehat{Z}^t \Lambda^*(\delta) Z_{opt}^*}{|Z_{opt}|^t |Z_{opt}|} \quad (24)$$

y finalmente:

$$f_{opt}(\delta) = \left| \widehat{Z} \right|^t \left| \widehat{Z} \right| - \frac{\left| Z_{opt}^{t*} \Lambda^*(\delta) \widehat{Z} \right|^2}{|Z_{opt}|^t |Z_{opt}|} \quad (25)$$

donde * es el complejo conjugado.

En resumen, el proceso de cálculo de los parámetros quedaría como:

- Se calcula $f_{opt}(\delta)$ para todo δ , $0 \leq \delta \leq N-1$
- Se busca el valor δ que minimiza la función de costo $f_{opt}(\delta): \delta_{opt}$

- Se calcula q_{opt} según (24) y se obtienen los parámetros λ y θ .

$$\lambda_{opt} = |q_{opt}| \quad (25)$$

$$\theta_{opt} = \angle q_{opt} \quad (26)$$

4. EXPERIMENTACIÓN

La base de datos usada durante el proceso de experimentación cuenta con 28 objetos de forma libre. Para cada objeto se ha obtenido previamente un modelo tridimensional de alta resolución y precisión a través de un láser scanner. A continuación, se extraen las siluetas 2D de las imágenes de cada modelo visto desde un conjunto homogéneo de puntos de vista. La implantación de estos puntos de vistas se realiza a través de los vértices de una esfera teselada semiregular con origen en el c.d.m del objeto. La figura 2 muestra dos modelos dentro de una esfera de 80 vértices mientras que en la figura 3 se observa la apariencia de los modelos desde determinados punto de vista. Para cada una de estas vistas se extrae la silueta correspondiente (Figura 4 y 5) y se almacena en la base de datos sus descriptores de Fourier.

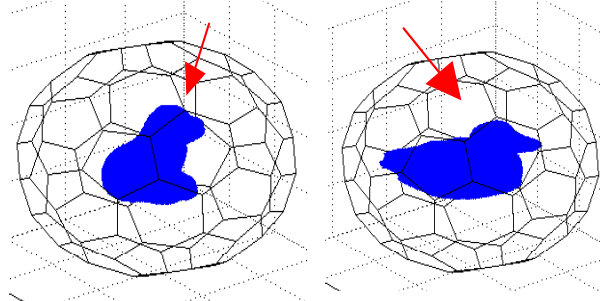


Figura 2. Cada nodo de la esfera teselada constituye un punto de vista diferente a partir del cual se extrae las siluetas del objeto

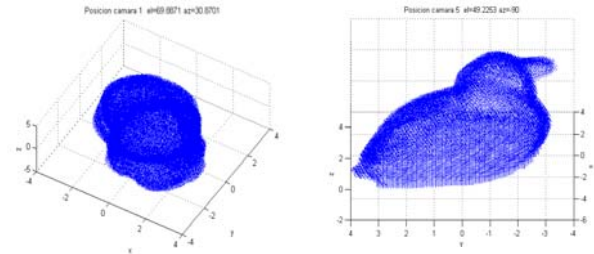


Figura 3. Vista parcial desde uno de los nodos de la esfera.

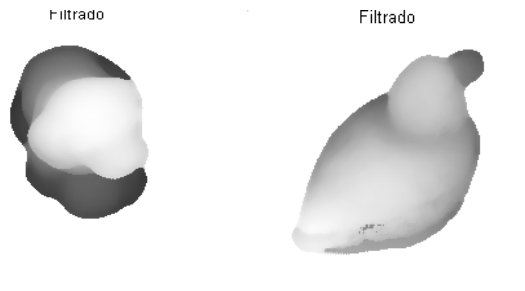


Figura 4. Imagen de profundidad sobre la que se realiza un proceso de filtrado adicional.



Figura 5. Siluetas 2-D

Se ha impuesto una normalización de siluetas de 512 puntos. Si se realiza el gráfico de los módulos de los descriptores de Fourier para el conjunto de siluetas del objeto de la figura 2 izquierda, se observa que sólo los primeros y los últimos términos son los representativos (Figura 6). Con vistas a optimizar los tiempos de ejecución manteniendo la información básica de cada silueta, se pretende realizar un método de discriminación de descriptores. Comenzando por el primer descriptor $X(k)$ y su simétrico $X(N-k)$ la idea es ir añadiendo descriptores hasta que la silueta quede representada con un error menor que un valor de umbral prefijado. El número de descriptores de Fourier calculado para cada silueta es también guardado en la base de datos. Notar que este es un proceso *off line* que no afecta a la ejecución de nuestro algoritmo.

La determinación de la discriminación de descriptores y la influencia de ruido de imagen son cuestiones que están todavía en fase investigación. No obstante, las primeras pruebas realizadas demuestran que la cantidad de descriptores se reducen en más de un 60% para un error de un píxel. La Figura 7 muestra los resultados obtenidos después de una discriminación superior al 60% de descriptores.

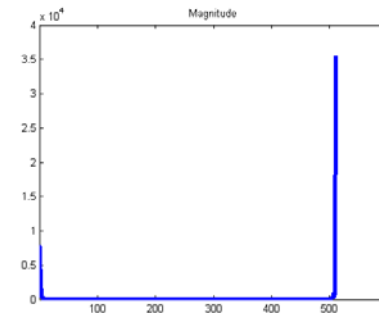
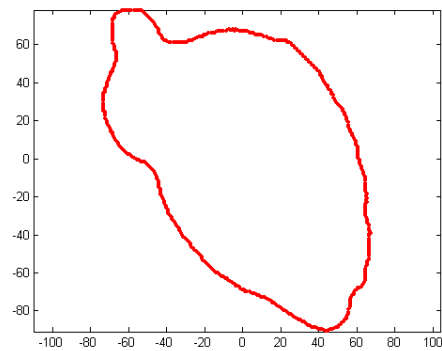
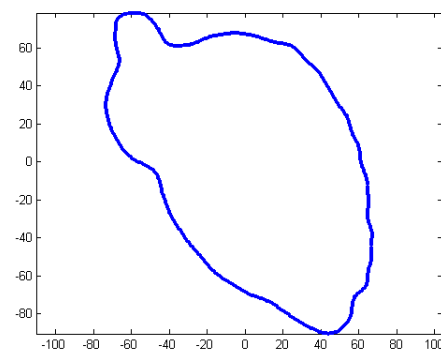


Figura 6. Gráfica de los módulos de los descriptores de Fourier de un objeto: 512 términos

Para estos ejemplos, la disminución de los tiempo de ejecución fue significativa. Los test hechos en un PIII a 700 MHz mostraron que la rutina de reconocimiento para la base de datos de 28 objetos disminuyó de 0.7049s a 0.3923s con el método de reducción de descriptores.



(a)



(b)

Figura 7. (a) Representación de la silueta con 512 descriptores. (b) Representación de la silueta con 134 descriptores obtenidos después de aplicar el algoritmo de discriminación.

A las siluetas de prueba se le aplicaron niveles de ruido aleatorio del orden de hasta 2 píxeles de distorsión ante las cuales el sistema actuó con robustez. La existencia de ruido sólo provoca que el tiempo de ejecución al hacer el reconocimiento

aumente ya que las mejores siluetas candidatas pueden no ser las verdaderas con lo cual aumenta la cantidad de siluetas a evaluar.

A la hora de realizar el proceso de identificación se comprobó que los descriptores de Fourier para siluetas similares así como para siluetas que cumplen la propiedad de reflexión (Figura 8) son muy parecidos. Esta característica implica tomar varias siluetas candidatas. En la pruebas realizadas con siluetas de escena sintéticas, a las que se les aplicó los niveles de ruido indicados anteriormente, se ha trabajado con un margen de 3 siluetas candidatas obteniendo los resultados esperados.

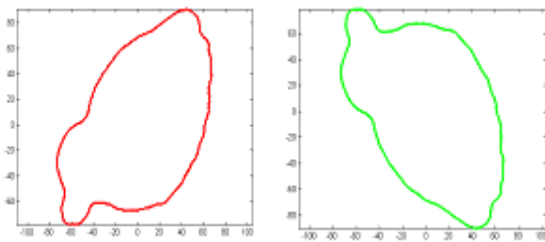
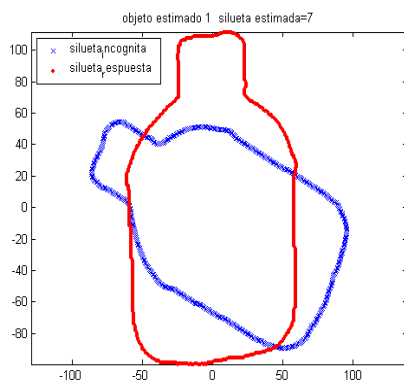


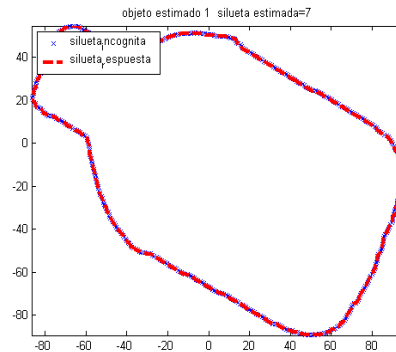
Figura 8. La silueta de la izquierda es la reflexión de la silueta de la derecha

Como resultado más relevante se tiene que la efectividad de los métodos empleados, sobre la base de datos de 28 objetos, en la fase de reconocimiento es de un 100%, siendo el error cuadrático medio que comete el sistema al calcular los parámetros de posición (rotación y traslación) y escalado de 0.0005.

La siguiente figura muestra la identificación correcta y *matching* realizado a una silueta de prueba.



(a)



(b)

Figura 9. (a) en azul la silueta incógnita a identificar y en rojo la silueta correspondiente en la base de datos. (b) se muestra el *matching* después de calcular los parámetros de posición y pose.

5. CONCLUSIONES

El uso de descriptores de Fourier para llevar a cabo el reconocimiento de objetos y la determinación de parámetros de posicionamiento y escalado, ha proporcionado resultados satisfactorios en tiempos cercanos al tiempo real. Por otra parte, la estrategia de reducir el número de descriptores de Fourier sin pérdida significativa de información, permitirá optimizar los tiempos de ejecución.

En trabajos futuros relacionados con el sistema de visión, se estudiará un procedimiento óptimo de selección de bandas de descriptores considerando el factor de ruido en la imagen. Otro problema a resolver es el referente a la discretización de puntos de vista en la base de datos y su influencia en los resultados.

Agradecimientos

Queremos agradecer la colaboración de las siguientes entidades:

- Programa de Cooperación Internacional de la UCLM, España.
- Ministerio de Educación y Cultura por la financiación de los proyectos: DPI2002-03999-C02.
- Laboratorio de ISA de la Universidad de Extremadura por su colaboración en la síntesis de los modelos tridimensionales.

Referencias

- [1] Diaz de Leon, R. , Sucar, E.(2000) Human Silhouette Recognition with Fourier Descriptors, *Proceedings of the International*

- [2] Chaker, F. , Tarak, M. Ghorbel, F. (2003). A complete and stable set of Affine-Invariant Fourier Descriptors, *Proceedings of the 12th International Conference on Image Analysis and Processing (ICIAP'03)*, 0-7695-1948-2/03
- [3] Persoon, E., King-Sun, F. (1977), Shape Discrimination using Fourier descriptors. *IEEE Trans on Systems, Man and Cybernetics*, vol. Smc-7, No. 3 march 1977.
- [4] Proakis, J., Manolakis, D. (2001), "Tratamiento digital de señales". Prentice Hall pp 401-441.
- [5] Smith, C. , Papanikolopoulos, N.(1995) Grasping of static and moving objects using vision-based control approach. *Proceedings of the International Conference on Intelligent Robots and Systems (IROS'95)*, 0-8186-7108-4/03