

Reconocimiento automático de objetos basado en la relación dimensional

Alejandro Israel Barranco Gutiérrez¹ y José de Jesús Medel Juárez²

¹ Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada,
Calzada Legaría 694 Col. Irrigación, 11500, México D.F.
abarrancog0900@ipn.mx

² Centro de Investigación en Computación, IPN, Calle Venus s/n,
Col. Nueva Industrial Vallejo, 07738, México D.F.
jjmedelj@yahoo.com.mx

Resumen. En este trabajo se presenta el reconocimiento de objetos tridimensionales con respecto a las imágenes proporcionadas por el sistema de visión estereoscópico, conformado por cámaras digitales convencionales. La problemática de la flexibilidad requerida por la dinámica en el reconocimiento de los sistemas en la vida real, donde los objetos son captados sin que se requiera una distancia fija a las cámaras, permite reconocerlos a través de sus propiedades dimensionales, al permanecer éstas invariantes. El método propuesto se compone del reconocimiento de forma y su relación dimensional, que en conjunto permite reconocer al objeto de manera confiable sin que se encuentre a una distancia fija a una cámara. Haciendo así que la metodología pueda utilizarse para muchas aplicaciones entre las que se destacan, el seguimiento de objetos por medio de diferentes cámaras al contar con los rasgos dimensionales dentro del sistema de control considerado, permitiendo llegar así a la vigilancia automática.

Palabras clave. Reconocimiento de objetos en 3D, visión estereoscópica, rasgo dimensional.

Automatic Object Recognition Based on Dimensional Relationships

Abstract. This paper presents a methodology for tri-dimensional object recognition with respect to images of the stereoscopic vision system integrated by standard digital cameras. The methodology considered

solving the flexibility problem of a real dynamic recognition system, in which the image object is photographed without a defined fixed distance. The results allow automatic recognition through its dimensional characteristics, such that its invariant form is constant. The recognition technique and the dimensional relationship, without a fixed distance camera were used. The results developed could be considered in many tracking object applications, vision control with dimensional forms, and dynamic classification.

Keywords. Object recognition in 3D, stereoscopic vision, dimensional characteristic.

1 Introducción

En este artículo se presenta una solución al problema del reconocimiento de objetos tridimensionales en imágenes proporcionadas por un sistema de visión estereoscópico. Los trabajos reportados en [0, 5, 7, 15, 17] analizan imágenes de objetos que contienen información tridimensional pero en su mayoría no utilizan cámaras ni un sistema estereoscópico calibrados, por lo tanto no logran identificar diferencias dimensionales entre objetos de acuerdo con [5, 15].

El objetivo dentro del sistema digital es diferenciar a los objetos con la misma forma de acuerdo con sus dimensiones requiriendo para ello del reconocimiento de forma y relación dimensional, considerando que cuentan con invarianza de rasgos [8, 13, 14].

Resumen extendido de tesis doctoral. Graduado: Alejandro Israel Barranco Gutiérrez. Director: José de Jesús Medel Juárez. Fecha de graduación: 19.06.2010.

Extended abstract of PhD thesis. Graduated: Alejandro Israel Barranco Gutiérrez. Advisor: José de Jesús Medel Juárez. Graduation date: 06/19/2010.

2 Componentes del sistema

Los Sistemas de Reconocimiento Automático de Objetos (SRAO) [7, 10, 11], incluyen los componentes: a) *banco de modelos*, b) *módulo de pre-procesamiento*, c) *módulo de segmentación*, d) *módulo de extracción de rasgos*, e) *módulo generador de hipótesis* y e) *módulo verificador de hipótesis*. Y que para una visión estereoscópica se requiere considerar (como se observa en la figura 1): a₁) *adquisición de imagen por las cámaras*, b₁) *conversión a escala de grises y umbralado*, c₁) *detección de puntos de interés*, d₁) *reconocimiento de objetos por medio de sus invariantes*, e₁) *rasgos dimensionales de acuerdo a los puntos de interés de c₁*, f₁) *identificación del objeto con la búsqueda en un ambiente con múltiples objetos con rasgos similares, pero con diferentes dimensiones*. Para que opere la metodología se requiere adicionalmente: a₂) *Dos cámaras digitales* y b₂) *Sistema estereoscópico*: condiciones necesarias para conocer a qué distancia se encuentra el objeto de alguna de las dos cámaras y c₂) *Las dos cámaras deben observar la misma escena para determinar los rasgos dimensionales del objeto a identificar*.

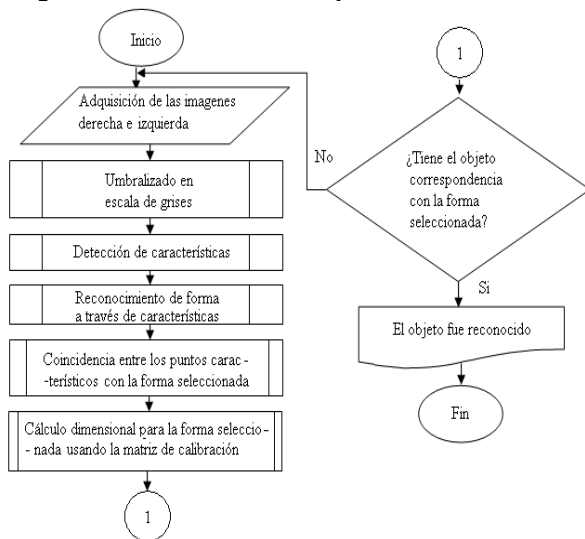


Fig. 1. Diagrama de flujo para el reconocimiento de objetos por rasgos y dimensiones

3 Reconocimiento de objetos

El reconocimiento está basado en la representación matricial de imágenes en escala de grises (tanto derecha (D) como izquierda (I)) donde cada imagen se denota por una matriz de $n \times m$ píxeles.

Una imagen digital $f(x, y)$ es un arreglo bidimensional acotado en Z_+^2 , descrito en escala de grises (desde 1 hasta 256 tonos) con $\{x\}, \{y\} \in Z_+$.

La imagen digital $f(x, y)$ cuenta con una transformación binaria $Ob(x, y)$ [11], formada por un grupo de objetos no vacíos $\{Ob_i\}$ que cumplen con la condición de cerradura para la unión y la intersección $\bigcup_{i=1}^n Ob_i \subseteq Ob$, $\bigcap_{i=1}^n Ob_i \subseteq Ob$ considerando $Ob_i^c = (O \setminus Ob_i)$.

El objeto Ob_i posee n propiedades respecto a un esquema de extracción de características, descritas como $Ob_i = \{ob_{1,i}, ob_{n,i}\}$, $i, n \in Z_+$, como los contornos de la imagen binarizada $b(x, y)$, y con base a los cuales se obtienen sus invariantes [8]. El objeto Ob cuenta con una envolvente o contorno (1).

$$r_i = \overline{\lim}_{i=1}^k \bigcap_{m=1, n} ob_{m,i}, ob_{m,i} \in Ob_i, i, k, m, n \in Z_+. \quad (1)$$

Definición 1. El objeto no vacío $Ob_i, i \in Z_+$ está envuelto en por una región r_i en el sentido de Hausdorff, de acuerdo con (2).

$$r_i := \inf \{ \delta_{i,j} : \delta_{\max} := \overline{\lim} S(Ob_i), i, j \in Z_+ \}. \quad (2)$$

Con $S(Ob_i)$ como superficie del objeto expresada en píxeles.

Definición 2. La región r en el sentido de Hausdorff que acota al objeto Ob contiene a la colección de regiones r_i , donde $r_{\max} := \overline{\lim} \{r_i\}$ en la que $r_{\max} \subseteq r$.

La región r descrita por un sistema estereoscópico calibrado de acuerdo con [3, 7, 18], en el que se conoce el vector de posición \bar{p} del centro de referencia de la cámara izquierda respecto del de la derecha, permite describir por medio de la matriz de proyección de perspectiva $M_{3,4}$ los puntos del objeto en 3D a una fotografía.

Teorema 1. *El objeto Ob en cualquier posición a partir de los siete invariantes [8, 13] y el rasgo dimensional r_d [14], dada la imagen digital*

$G(x, y, z) \subseteq Z_+^3$ *captada por un sistema estereoscópico de cámaras digitales con $\{x\}, \{y\}, \{z\} \in Z_+$ es posible reconocerlo con respecto a otros objetos con características físicas similares.*

Prueba. *Considérese que existe: a₄) un conjunto de objetos con rasgos similares $\{c_1, c_n : n \in Z_+\} \subseteq C, C \subseteq Z_+^3$, dentro del área del sistema estereoscópico de cámaras., b₄) Existe un vector característico extendido V_C^* respecto del objeto Ob buscado., c₄) Cada objeto Ob tiene un vector característico extendido formado por el rasgo dimensional y los siete invariantes de Hu, $V_{c_i} = (rd_{c_i}, \overline{\phi_{1_{c_i}}}, \overline{\phi_{7_{c_i}}})$, así como el rasgo dimensional (rd) es un invariante del objeto. De acuerdo con las definiciones 1 y a los puntos del a₄) al c₄) se cumple que $\exists! V_C^* \neq 0$ y que $\limsup V_{c_i} \rightarrow V_C^*$.*

4 Desarrollo experimental

Se propone la identificar utilizando a una esfera que tiene como diámetro 5 cm y que se encuentra en una escena con un conjunto de esferas del mismo material y color pero de diferentes diámetros.

El preprocesado requiere del filtro mediana [16], que permite eliminar algunas impurezas en las imágenes. El umbralado de acuerdo con [11] permite obtener una forma binaria, como se muestra en la figura 2. Lo que permite obtener un conjunto de regiones que representan *objetos de*

interés, además de una disminución del universo de búsqueda [6]. Habiendo así *objetos candidatos* para ser analizados y determinar si alguno corresponde con el objeto buscado.

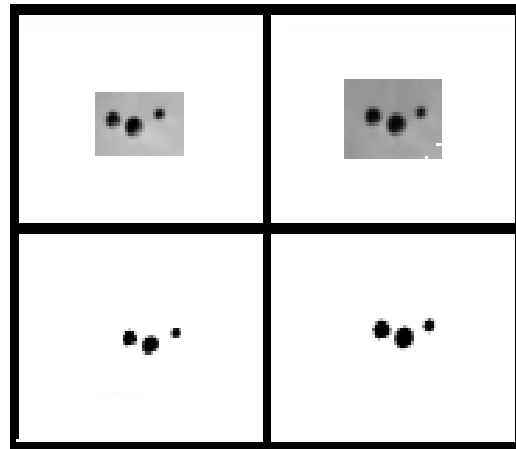


Fig. 2. Imágenes filtradas en escala de grises (superior) y binarizadas (inferior)

De acuerdo con [10, 12] se obtienen los contornos de las imágenes como se muestra en la figura 3, reduciendo la información redundante.

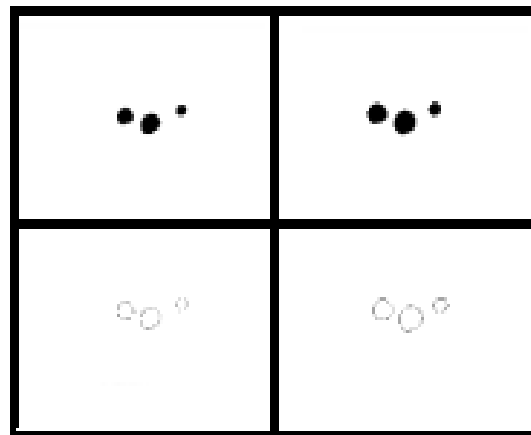


Fig. 3. Umbralizado (superior) y la detección de bordes (inferior) de las imágenes

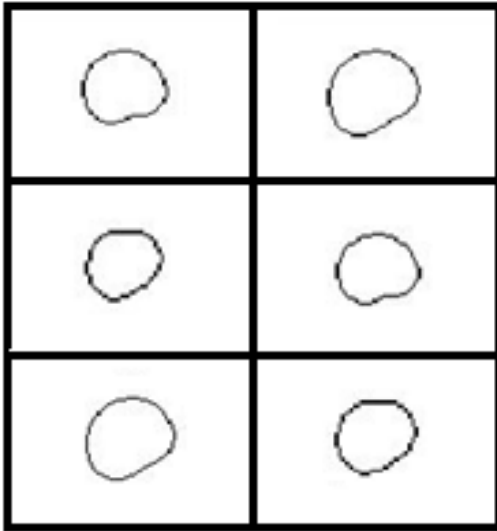


Fig. 4. Fronteras de las tres imágenes binarizadas por cada cámara para ser analizadas de acuerdo a su tamaño

Para hacer un análisis detallado de cada una de las formas obtenidas en la detección de bordes, se procede a separar los objetos aislándolos en una imagen independiente que conserve su posición dentro de la imagen original y analizar los rasgos estadísticos de la forma independiente, comparándolos con los rasgos del banco de modelos almacenados en SRAO, para determinar conforme a una función de pertenencia [9] si dichos parámetros son similares o no a alguno de los almacenados en el banco de modelos, como se muestra en la figura 4.

4.1 Reconocimiento de forma de los objetos

En esta etapa se analiza si la silueta bajo condiciones de iluminación constantes corresponde al objeto buscado mediante el uso de los invariantes de Hu [8], que se derivan de los momentos estadísticos y que mapean a cada silueta formada de una imagen de dos dimensiones binaria a los siete invariantes de Hu. En el caso de las esferas podemos definir rangos en nuestras variables para identificar a la forma

deseada. El proceso desarrollado para establecer dichos límites, se realizó observando los valores típicos de cada coeficiente para treinta esferas, después se realizó un redondeo de los valores. Los límites para este caso se presentan en la tabla 1.

Tabla 1. Límites tomados experimentalmente para la clasificación de las esferas

	F_{i_1}	F_{i_2}	F_{i_3}	F_{i_4}	F_{i_5}	F_{i_6}	F_{i_7}
Límite Superior	3.50	0.60	0.20	0.0060	0.3e-3	0.0100	0.3e-4
Límite Inferior	1.10	0.25	0.01	0.0001	0.06e-3	0.0005	-0.9e-4

4.2 Rasgo dimensionales

Los rasgos dimensionales proporcionan información del objeto a reconocer por medio de algunas de sus dimensiones características; por ejemplo, en este trabajo se muestra un conjunto de esferas del mismo color (blancas), que por tener la misma forma, su análisis estadístico de Hu, muestra valores muy similares; pero si se requiere reconocer en la imagen de la figura 4 a la esfera de diámetro 0.05 m.; será necesario comprobar la dimensión del diámetro de la esfera.

Para realizar la medición automática de alguna dimensión a partir de imágenes, es necesario definir las dimensiones a medir y un sistema estereoscópico de cámaras calibradas, que en conjunto nos permitirán reconocer al objeto buscado, por medio de su forma y algunos rasgos dimensionales, estos últimos son muy importantes para el reconocimiento en tres dimensiones, ya que este rasgo es la característica que nos ayudará a comprobar la identidad del objeto buscado.

La calibración de cámaras digitales de acuerdo con [4, 7, 17, 18], consiste en estimar los parámetros de un modelo que relacione los píxeles de la fotografía con las dimensiones del

objeto capturado en la fotografía, parámetros que dependen de la resolución y el enfoque.

4.3 Medición automática de dimensiones

Después de la etapa de pre-reconocimiento, donde se verificó si alguna de las formas encontradas en la imagen correspondía con la forma buscada (en este caso la silueta de una esfera), se trabaja con la etapa de medición automática de dimensiones de las siluetas. Para esto se escoge una dimensión característica del objeto, que llamaremos *rasgo dimensional*. En este caso medimos el radio de las esferas encontradas en la imagen, con un sistema de visión estereoscópico previamente calibrado.

El rasgo dimensional seleccionado fue el diámetro de la esfera medido desde el extremo superior al inferior de esta, así como se muestra en la figura 5. Se escogen estos puntos debido a que en las dos imágenes del sistema estereoscópico estos puntos corresponden al mismo punto en el sistema de referencia espacial 3D, ya que asumimos que el sistema está calibrado y colocado sobre una línea horizontal.

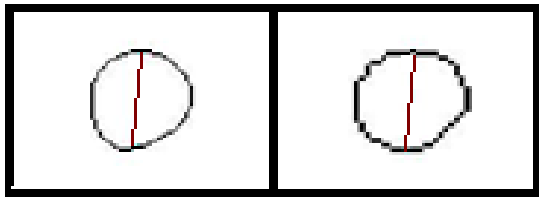


Fig. 5. Medida del diámetro de la imagen a través de sus fronteras superior e inferior

Al encontrar los puntos (x, y) de los extremos, superior e inferior de cada una de las esferas en las imágenes derecha e izquierda, se procede a calcular el diámetro de dicha esfera obteniendo directamente las coordenadas (x, y, z) del extremo superior e inferior de cada una de las imágenes (derecha e izquierda) de la esfera. Para lograr esta tarea utilizamos el modelo de cámaras *PINHOLE* estereoscópico calibrado, donde sus entradas son los puntos: (x_I, y_I) y (x_D, y_D) de los extremos de las esferas en las

cámaras izquierda y derecha respectivamente y la salida es un punto en el espacio (x, y, z) . Este proceso se ilustra en las Tablas 2, y Tabla 3., Donde se muestran los resultados para los puntos superiores e inferiores respectivamente para el caso de las imágenes presentadas en las figuras 2, 3 y 4.

Tabla 2. Coordenadas de los puntos superiores de la esfera captada por las cámaras

Coordenada	x_izquierdo_l (píxeles)	x_derecho_l (píxeles)	Xc_l_izquierdo (mm)
X	93	220	-70.60
Y	102	71	41.654
Z			46.028

Tabla 3. Coordenadas de los puntos superiores de la esfera captada por las cámaras

Coordenada	x_izquierdo_l (píxeles)	x_derecho_l (píxeles)	Xc_l_izquierdo (mm)
X	97	230	69.7538
Y	185	159	51.9970
Z			365.9866

El último paso es verificar si las dimensiones de la esfera corresponden con el diámetro buscado. En este trabajo se eligió reconocer a la esfera de diámetro 0.05 m. Y con ayuda de los dos puntos en el espacio (superior e inferior) de la esfera se calcula el diámetro de la misma por medio de (3).

$$d_{L_2} = \left(\sum_{i=1}^3 (x_i - x_0)^2 \right)^{0.5} \quad (3)$$

El resultado en el ejemplo que hemos venido manejando fue: 0.05002476 m. Lo cual verifica que la esfera mediana que se encuentra en la imagen de las tres esferas, es la esfera que buscábamos y esto fue calculado de manera automática.

5 Conclusiones

Se presentó la metodología para el reconocimiento automático de objetos basado en las dimensiones y se probó que el sistema de reconocimiento es flexible respecto a diferentes formas de los objetos de interés. Describiendo cómo seleccionar un objeto sin contar con las cámaras calibradas, de acuerdo a las características establecidas en el sistema de cómputo de entre objetos de forma y colores similares pero de diferente dimensión.

Referencias

1. **Amit, Y. (2002).** *2D Object Detection and Recognition: Models Algorithms, and Networks.* Cambridge, Mass.: MIT Press.
2. **Barranco, A.I. & Medel, J.J. (2008).** Identificación de formas y color en imágenes a partir de la media y la varianza de los niveles de color RGB. *1er Simposio de Tecnología avanzada, México, D. F., 79.*
3. **Barranco, A.I., & Medel, J.J. (2008).** Proceso de calibración de cámaras digitales basado en el modelo pin-hole. *2do Simposio de Tecnología Avanzada, México, D.F., 66.*
4. **Barranco, A.I. & Medel, J.J. (2008).** *Visión estereoscópica por computadora* (Informe técnico, serie azul, no. 235). México, DF: Instituto Politécnico Nacional, Centro de Investigación en Computación.
5. **Barranco, A.I. & Medel, J.J. (2009).** Digital Camera Calibration Analysis Using Perspective Projection Matrix. *8th WSEAS International Conference on Signal Processing, Robotics and Automation, 321–325.*
6. **Gonzalez, R.C. & Woods, R.E. (2008).** *Digital Image Processing (3rd ed.).* Upper Saddle River, N.J.: Prentice Hall.
7. **Hartley, R. & Zisserman, A. (2003).** *Multiple view geometry in computer vision (2nd ed.).* Cambridge, UK: Cambridge University Press.
8. **Hu, M.K. (1962).** Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory, 8(2), 179–187.*
9. **Medel, J.J., García, J.C. & Guevara, P. (2009).** Real-time neuro-fuzzy digital filtering: a technical scheme. *Automatic Control and Computer Sciences, 43 (1), 22–30.*
10. **Lam, L., Lee, S.W. & Suen, C.Y. (1992).** Thinning Methodologies – A comprehensive Survey. *IEEE Transactions on pattern analysis and machine intelligence, 14(9), 869–885.*
11. **Otsu, N. (1979).** A threshold selection method from gray level histograms. *IEEE transactions on systems, man and cybernetics, 9(1), 62–66.*
12. **Pajares, G. & de la Cruz, J.M. (2002).** *Visión por computador, imágenes digitales y aplicaciones.* Madrid: Alfa Omega.
13. **Peebles, P.Z. (2006).** *Principios de probabilidad, variables aleatorias y señales aleatorias (4^a ed.).* Madrid: McGrawHill.
14. **Rudin, W. (1980).** *Principios de análisis matemático (3ra ed.).* México: Mc Graw Hill.
15. **Sobel, I. & Feldman, G. (1968).** *A 3x3 Isotropic Gradient Operator for Image Processing,* presented at a talk at the Stanford Artificial Project.
16. **Sossa, J.H. (2006).** *Rasgos descriptores para el reconocimiento de Objetos.* México, D.F.: Instituto Politécnico Nacional.
17. **Voss, K., Marroquin, J.L., Gutiérrez S.J., & Suesse, H. (2006).** *Análisis de imágenes de Objetos Tridimensionales.* México, D.F.: Instituto Politécnico Nacional.
18. **Zhang, Z. (2000).** A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11), 1330–1334.*



Alejandro Israel Barranco Gutiérrez es egresado del Doctorado en Tecnologías Avanzadas. Actualmente el profesor a tiempo completo del TEC unidad Tlahuac y del grupo ICEL. Es coautor de una patente y de artículos internacionales.



José de Jesús Medel Juárez es profesor a tiempo completo del Centro de Investigación en Computación, egresado del doctorado en Ciencias en Control Automático del CINVSTAV. Es reconocido como miembro del SNI, así como de la AMC, IBC y Whos' who in Sciences and Eng. Ha graduado tanto en doctorado como en maestría a más de 25 profesionistas. Cuenta con más de 6 libros publicados por editoriales nacionales como internacionales. Es coautor de diversos artículos indizados en JCR, EI, CONACyT, entre otros. Ha desarrollado y sometido diversas patentes. El Dr. Medel fue creador y fundador del posgrado en Tecnología Avanzada que se imparte en el CIITEC.

Article received on 22/08/2010; accepted 10/05/2011.