# Modelo Óptico de Estéreo Visión usando Visión por Computadora

Marco Antonio Moreno Armendáriz, Julián Aguilar Luna, Eduardo Gómez Ramírez & Xavier Vilasis Cardona<sup>1</sup> Laboratorio de Investigación y Desarrollo de Tecnología Avanzada (LIDETEA) UNIVERSIDAD LA SALLE <u>E-mail: eqr@ci.ulsa.mx</u>

<sup>1</sup>U. Ramon Llull, Departament d'Electrónica, Enginyeria i Arquitectura La Salle, Pg. Bonanova 8, 08022 Barcelona, España, E-mail: xvilasis@salleURL.edu Recibido: Julio de 2003. Aceptado: Julio de 2003.

#### RESUMEN

A través de los tiempos, el hombre ha observado y estudiado la naturaleza para entender y replicar algunos de sus comportamientos por medio de elementos artificiales. Los resultados obtenidos lo han ayudado a simplificar sus tareas en algunas áreas de trabajo.

Gracias a los estudios tecnológicos hoy día se cuenta con herramientas de apoyo que tienen funciones similares a las funciones humanas, tales como las memorias permanentes que simulan una de las tantas actividades cerebrales, brazos mecánicos que simulan una de las dos extremidades humanas, etc. La visión no es la excepción de estudio, en el presente trabajo se da a conocer una propuesta de desarrollo y aplicación de la visión artificial.

Palabras claves: Visión por computadora, estéreo visión, cinemática.

#### ABSTRACT

The research of robotics and computer vision aims to realize some aspects of human functions into mechanical systems. Trying to reproduce these functions allow researchers to reach very simple and functional algorithms to be used within the industry.

Currently there are appropriate technologies based on this idea, like a special memory with the capacity to simulate some kind of brain activities or robot manipulators used to reproduce simple repetitive tasks. The computer vision aims to produce new schemes to bring autonomy and more flexibility to these systems. In this paper a new computer vision model is developed with the ability to locate automatically the position in the space of an unknown point.

Keyword: Computer vision, stereovision, kinematics.

# **INTRODUCCIÓN**

En tiempos recientes se ha llevado a cabo la automatización de diferentes tipos de procesos, como los industriales o los administrativos. Esto trae como beneficio la reducción del tiempo. gracias al avance en los sistemas de cómputo y al surgimiento de nuevas tecnologías, como

son: los sistemas de visión por computadora, la lógica difusa y las redes neuronales.

Estos nuevos desarrollos han facilitado el trabaio del ser humano haciéndolo más eficiente. Existen diversos tipos de sistemas de visión por

# Artículo

computadora, catalogados por el número y tipo de sensores. Se componen principalmente de cámaras de video especializadas, encargadas de realizar la adquisición de las imágenes y de un conjunto de algoritmos cuyo objetivo es extraer de éstas la información necesaria para el proceso a realizar. Actualmente, la mayoría de los sistemas de visión por computadora que se emplean se encuentran instalados en las plantas manufactureras para efectuar tareas de control de la calidad de productos, esto puede llevarse a cabo mediante algoritmos de conteo y detección de patrones, los cuales se seleccionan por un supervisor experto, de acuerdo con las necesidades de cada proceso. Para este tipo de actividades es suficiente utilizar una sola cámara o sensor, ya que no es necesario conocer la ubicación de un objeto en el espacio.

Por otra parte, existen tareas como en una línea de producción automatizada donde un conjunto de robots se encargan de elaborar un producto en etapas previamente programados con rutinas especificas, que ejecutan repetidamente. Sin embargo, si se deseara implementar una celda de manufactura flexible, donde los robots tuvieran la capacidad de realizar diferentes actividades, sin la necesidad de tener que reprogramarlos. Una opción sería implementar algunas de las tecnologías anteriormente mencionadas con la finalidad de brindar mayor flexibilidad al robot. Un ejemplo de esta tecno- logías son los sistemas de visión por computadora. Una de las tareas más comunes de estos sistemas es obtener la posición del órgano terminal del robot en el espacio de trabajo mediante la información visual. El valor obtenido de la posición se incorpora en el lazo de control del robot para que éste lleve a cabo la tarea asignada.

Con la idea general de reconstrucción de imágenes 3D a partir de 2D se hace necesario el tener al menos dos patrones lógicos que ayuden a analizar el comportamiento de las imágenes, bajo esta estructura muchos experimentos publicados se han llevado a cabo, como la estructura a partir del movimiento[1,2], estéreo correspondencia [3-7] y forma [8].



Figura 1. Partes del robot A465.

En el presente trabajo se diseñará un sistema de visión por computadora para obtener la información necesaria para el modelo óptico de estéreo visión propuesto. Finalmente, se implementa este modelo para obtener visualmente algunos parámetros del modelo cinemático del robot A465. (Fig. 1)

# I. PLATAFORMA DE VISIÓN.

#### I.1 Sistema de visión artificial.

El sistema de estéreo visión está formado por los siguientes elementos:

- 1. Dos cámaras análogas Pulnix TM72-EX
- 2. Dos lentes Cosmicar C815B
- 3. Una tarjeta de video NI PCI-1408
- 4. Una computadora Dell Optiplex GX240.



Figura 2. Sistema de visión por computadora.

Con el fin de reducir la cantidad de información a procesar, la escena y el robot se cubren en color negro (ver Figura 3). Después, se colocan tres marcas visuales (círculos blancos) sobre las partes indicadas en la Figura 1. A partir de estas marcas se obtendrá toda la información visual necesaria para ubicar la posición en el espacio del órgano terminal del <u>robot A465.</u>



Figura 3. Imagen capturada con el sistema de visión por computadora

#### I.2 Distorsión radial

Las lentes utilizadas para las cámaras son las Cosmicar C815b, que tienen un área visual de 1194x1728 mm, a 2 m de distancia, de un objeto enfocado. La convexidad de las lentes hace que las imágenes que pasan por ellas tengan una distorsión radial, como puede observarse en la Figura 4:



Figura 4. Imagen con distorsión radial.

Dado que el modelo óptico de estéreo visión se basa en la información extraída de imágenes, es necesario tener una imagen sin distorsión. Para eliminar la distorsión de las lentes utilizamos el algoritmo de corrección de imágenes presentado por Chavarría, Soria y Wiederhold [9], que usa la siguiente expresión matemática:

$$r = \frac{r^4}{1 + d_2 r^2 + d_4 r^4}$$

(Ec.1)

Donde *r* es el radio de corrección y *r*' es el radio de distorsión. Experimentalmente, se obtienen los valores de corrección de la imagen que son:  $d_2 = -6.166e^{-7}$  y  $d_4 = 2.083e^{-12}$ . Así, las coordenadas sin distorsión se obtienen mediante:

$$\begin{split} x &= X + \frac{x' - X}{1 + d_2 r'^2 + d_4 r'^4} \\ y &= Y + \frac{y' - Y}{1 + d_2 r'^2 + d_4 r'^4} \end{split}$$

(Ec.2)

Donde (x',y') son las coordenadas de los centroides de las marcas con distorsión, (X, Y) es la coordenada del centro de la imagen que, para este caso, (320,240) y (x,y) son las nuevas coordenadas sin distorsión. En la Figura 5 se muestra la imagen sin distorsión.



Figura 5. Imagen con distorsión radial eliminada.

# II. ALGORITMOS DE VISIÓN POR COMPUTADORA

El sistema de visión por computadora es el encargado de calcular el centroide de las tres marcas visuales que aparecen en las imágenes capturadas con las cámaras analógicas (ver Figura 6).



Figura 6. Diagrama de proceso de los algoritmos de visión por computadora.

El proceso de obtención del centroide se divide en cinco etapas (ver Figura 7):

- 1. Binarización de la imagen.
- 2. Redimensionamiento de la imagen binarizada.
- 3. Segmentación de las marcas visuales.
- 4. Obtención de los centroides de las marcas visuales.
- 5. Etiquetado de las marcas visuales.

#### II. 1 Binarización

La imagen capturada por las cámaras analógicas se envía a la tarjeta PCI-1408 que se encarga de digitalizarla y binarizarla.

# II. 2 Redimensionamiento

Completado el proceso de binarización se cuenta con una imagen de 640x480 píxeles, esta cantidad podría incrementar mucho el tiempo de procesamiento si se maneja durante el proceso. Dado que el objetivo es obtener solamente el centroide de las marcas visuales, la imagen se reduce (redimensiona) de tal forma que los píxeles resultantes proporcionen esta información. Para redimensionar la imagen utilizamos la siguiente ecuación:

$$l = \frac{2 \times r \times \sin(45)}{2.5}$$

(Ec.3)

donde *r* es radio de los círculos en píxeles y *l* el número de píxeles.

Para esta aplicación r = 13 píxeles, por lo que l = 7 píxeles, dando como resultado que una imagen redimensionada es de 91x68 píxeles.

#### II. 3 Segmentación

En esta parte del proceso las marcas visuales se separan para después calcular su centroide y posteriormente etiquetarlas. Este proceso se lleva a cabo mediante una búsqueda de la 8*conexidad*,[10,11] que se encarga de agrupar a los puntos en tres objetos separados. Esto es posible dado que las marcas visuales se encuentran en las extremidades del robot A465, pudiendo garantizar que éstas nunca se encimarán.



Figura 7. Etapas del proceso de los algoritmos de visión por computadora.

# II. 4 Cálculo del centroide

Para la obtención del centroide de las marcas visuales se localizan los puntos máximos y mínimos horizontales y verticales, (ver Figura 8) siendo el centroide la intersección de ellas.



Figura 8. Cálculo de los valores máximos y mínimos

El valor de los centroides se obtiene con:

$$c_x = \frac{\max abajo + \max arriba}{2}$$
$$c_y = \frac{\max der + \max izq}{2}$$

(Ec.4)

En la Tabla 1, se presenta el valor de los centroides para el ejemplo de la Figura 7.

Objetos	Centroides	
	u	V
Objeto O1	267.5	213
Objeto O2	269	283
Objeto O3	411	207.5

Tabla 1. Listado de los objetos y el valor de sus centroides.

#### II.5 Etiquetado de las marcas

Por último, los centroides se etiquetan dependiendo el lugar en el que se encuentren colocados sobre el robot A465 (ver Figura 1).



Figura 9.Segundo criterio usado para etiquetar las marcas visuales.

Para etiquetar los objetos, se toma el valor de la coordenada (u,v) del primer centroide de la Tabla 1, con estos valores se posiciona sobre la imagen binarizada, y si el bit en esa posición es de color negro, se etiqueta como órgano terminal, de no ser así se toma el siguiente valor hasta encontrar el centroide con estas características. Para los dos puntos restantes se utiliza el criterio de posición, suponemos que uno de ellos siempre estará por arriba del otro (ver Figura 9). Bajo este criterio se toman los valores de la coordenada v de la Tabla 1, nombrando al de mayor valor base y el restante codo como se muestra en la Tabla 2.

Etiqueta	Centroide	
	u	V
Codo	267.5	213
Base	269	283
Órgano terminal	411	207.5

#### Tabla 2. Etiquetado de las marcas visuales

Con este proceso la computadora tiene la capacidad de correlacionar cada una de las marcas visuales con su correspondiente parte en el robot.

Hasta este punto se han descrito las diferentes herramientas encargadas del procesamiento de las imágenes para obtener la información visual que se usará en el modelo óptico de estéreo visión.

# III. MODELO ÓPTICO DE ESTÉREO VISIÓN

En esta sección presentaremos nuestra principal contribución. El modelo óptico de visión estéreo a cargo de localizar un punto desconocido en el espacio, con base en la información visual. Para ello se requiere conocer el centroide (u, v) de las marcas visuales ubicadas en el codo y órgano terminal del robot, además de los ángulos de rotación  $\theta_{1a}$ , $\theta_{1b}$  de las cámaras (ver Figuras 10 y 13).



Figura 10. Proceso realizado por el modelo óptico de estéreo visión.

#### III. 1 Modelo óptico de la cámara

Con base en el modelo del telescopio de Galileo [12], en la Figura 11 se plantea un modelo geométrico que nos ayudará a obtener las distancias entre el robot A465 y las cámaras.



Figura 11. Diagrama óptico de la cámara Pulnix con lentes Cosmicar C815B.

Con base en el modelo óptico de la Figura 12 obtenemos las siguientes ecuaciones:



(Ec.5)



Figura 12. Variables definidas para el modelo óptico

#### III.1.1 Obtención de los parámetros f, d'

De la ecuación 4 se obtienen las siguientes expresiones:

$$f = \frac{dh}{h + h'd^2} \quad ; \quad d' = \frac{h'd}{h}$$
(Ec.6)

Experimentalmente se obtienen los valores para d,h,h,h', posteriormente se sustituyen en la **ec. 6**, obteniendo f,d', teniendo a h en centímetros y las unidades de h' en píxeles. Por ello hacemos uso de un factor de conversión de unidades de píxel a centímetro, del cual resultó experimentalmente que:

$$f_{pix-cm} = 0.002708333.$$

Con este factor se obtuvieron los siguientes valores:

 $f = 0.585091 \ cm \ d' = 1.704799 \ cm$ 

#### III.2 Distancia entre cámara y escena

En este paso resulta el valor de las distancias  $(e_1, e_2)$  y  $(f_1, f_2)$ , que son las entre marcas visuales colocadas en las articulaciones del objeto observado y las cámaras (ver Figura 13). Donde, *c* es la distancia del codo al órgano terminal.

Hay que resaltar que este punto es uno de los más importantes, ya que las distancias calculadas se obtienen cuando el robot se encuentra en posición de inicio, de la que podemos obtener una coordenada de referencia.



Figura 13. Distancia entre las cámaras y la escena.

Con base en el diagrama geométrico del modelo óptico, (ver Figura 13) se obtienen las ecuaciones para el cálculo de las distancias:

$$\begin{aligned} e_1 &= c \frac{\sin(\alpha_2)}{\sin(\beta_2)} \quad ; \quad e_2 &= c \frac{\sin(\beta_3 + \alpha_1)}{\sin(\beta_2)} \\ f_1 &= c \frac{\sin(\alpha_1)}{\sin(\gamma_2)} \quad ; \quad f_2 &= c \frac{\sin(\gamma_3 + \alpha_2)}{\sin(\gamma_2)} \end{aligned}$$
(Ec.7)

Donde, los ángulos  $\alpha_1, \alpha_2, \beta_2, \beta_5, \gamma_2, \gamma_5$  se definen por los triángulos determinados por las líneas  $c, e_1, e_2, f_1, f_2$  (ver Figura 14) Estos valores se usarán en el siguiente paso.



Figura 14. Diagrama geométrico del modelo óptico

Obtención de la posición de un punto en el espacio.

Una vez que se han calculado las distancias  $(e_1, e_2)$  y  $(f_1, f_2)$ , se elige un punto conocido o de referencia, que será el centroide de la marca visual, encontrándose en el codo del robot A465 (ver Figura 15). Posteriormente, éste se mueve, de tal forma que ahora la posición del órgano terminal es desconocida (ver Figura 16).



Figura 15. Cuadro cartesiano a partir del codo

De ahí que, con esta nueva referencia, la posición desconocida del órgano terminal es  $(C_{x}, C_{y}, 0)$ .

#### III.3 Coordenada z

Para obtener el valor de la coordenada z, suponemos que ambas cámaras están montadas a la misma altura, con lo que se establece la siguiente relación:

$$z = k_r \mathbf{u}_L = k_r \mathbf{u}_R$$

(Ec.8)

• *K<sub>r</sub>*: es el factor de conversión de pixel a cm.

•  $u_L y u_R$  son las coordenadas en ambas imágenes con respecto al eje z, para el centroide de la marca visual que se encuentra en el órgano terminal.

#### III.4 Coordenadas x & y

Para obtener las coordenadas (x,y) se hace uso de un modelo geométrico análogo al mostrado en la Figura 14 para el calculo de las distancias c' y el ángulo  $\alpha'$  fórmula como se muestra en la Figura 16:



Figura 16. Valores requeridos para obtener las coordenadas (x,y).

Así:

 $x=c'cos(\alpha')$   $y=c'cos(\alpha')$ 

(Ec.9)

#### IV. RESULTADOS

En esta sección se implementan nuestros modelos en el robot CRS A465 (ver Figura 17).

# IV. 1 Calculo de parámetros cinemáticos usan do información visual

Cualquier robot puede detallarse cinemáticamente mediante la medición de cuatro variables que describen completamente cada articulación. Dos describen su propia articulación, y dos describen la conexión con las articulaciones. Estas variables se definen por la notación de Denavit-Hartenberg [13]. Aplicando esta convención obtenemos el modelo cinemático para el robot A465 (ver Figura 17). El modelo se probó experimentalmente.



Figura 17. Diagrama del modelo cinemático del robot A465.

El modelo cinemático es:



El experimento consiste en calcular visualmente los ángulos  $\theta_2$  (ángulo entre la base y la primera articulación del robot)  $\theta_3$  (ángulo entre la primera y la segunda articulación) que forman parte del modelo cinemático del robot.

Para obtener estos valores usamos las tres marcas visuales y el siguiente proceso:

- 1. Obtenemos el centroide  $(u_L, v_L) y (u_R, v_R)$  de las tres marcas visuales.
- Usando nuestro modelo óptico de visión estéreo obtenemos la posición correspondiente (*x*,*y*,*z*), coordinada por los tres puntos visuales. Estos valores se marcan como (*P*<sub>1</sub>,*P*<sub>2</sub>,*P*<sub>3</sub>).
- 3. Con esta información calculamos los valores de los parámetros  $\theta_2$ ,  $\theta_3$  como sigue:

$$\begin{split} P_1 &= (x_1, y_1, z_1), P_2 = (x_2, y_2, z_2), P_3 = (x_3, y_3, z_3) \\ V_1 &= (x_2 - x_1)i + (y_2 - y_1)j + (z_2 - z_1)k \\ V_2 &= (x_3 - x_2)i + (y_3 - y_2)j + (z_3 - z_2)k \\ V_3 &= (x_3 - x_1)i + (y_2 - y_1)j \\ V_1 &= V_2 = (x_2 - x_1) \cdot (x_3 - z_2) + (y_2 - y_1) \cdot (y_3 - y_2) \\ &+ (x_2 - z_1) \cdot (x_3 - z_2) \\ V_1 &= V_3 = (x_2 - x_1) \cdot (x_2 - x_1) + (y_2 - y_1) \cdot (y_2 - y_1) \\ V_1 &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} \\ V_2 &= \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2 + (x_3 - z_2)^2} \\ V_3 &= \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \end{split}$$

(Ec.11)

Los ángulos se obtienen mediante:

$$\begin{split} \theta_2 &= \cos^{-1} \frac{V_1 \cdot V_3}{|V_1| \cdot |V_3|} \\ \theta_3 &= \cos^{-1} \frac{V_1 \cdot V_2}{|V_1| \cdot |V_2|} \end{split}$$

(Ec.12)

El experimento consiste en mover la segunda articulación del robot A465 (ver Figura 18) desde -90° a 90°, y obtener usando la información visual el ángulo  $\theta_3$  del robot A465.



Figura 18. Experimento variando el ángulo q3 del robot A465

En la Tabla 3 se muestran los resultados, comparando la primera columna con la última, observamos que el error entre ambas es de tan sólo  $\pm 2^{\circ}$ . Este rango es aceptable para tareas que no requieran de alta precisión.

$\theta_3$ real	$\theta_3$
-90	-92.7840
-75	-76.7100
-60	-60.7780
-45	-45.6620
-30	-30.6810
-15	-15.7780
0	0.9057
15	13.8890
30	28.6460
45	43.4890
60	58.2490
75	72.7460
90	88.1120

Tabla 3. Resultados obtenidos con el modelo óptico de estéreo visión para diferentes ángulos de  $\theta_3$ .

# CONCLUSIONES

La principal contribución consistió en el desarrollar e implementar un nuevo modelo óptico de estéreo visión, analizando los componentes ópticos del sistema de visión por computadora. Este modelo se probó exitosamente en el robot A465, calculando los ángulos de su modelo cinemático. Nuestro modelo puede usarse en diferentes aplicaciones de visión por computadora como por ejemplo, en control visual, en robots móviles para la reconstrucción en tres dimensiones, etc.

# AGRADECIMIENTOS

Agradecemos el apoyo brindado por el <u>Labora-</u> torio de Robótica de los Talleres y Laboratorios de la Escuela de Ingeniería de la Universidad <u>La Salle, México.</u>

#### REFERENCIAS

- S.M. Seitz & C.R Dyer, Complete Scene Structure from Four Point Correspondences, Proc. 5th Int. Conf on Computer Vision, Cambridge MA, EUA, pp. 330-337, 1995.
- C.J Taylor & D.J. Kriegman, Structure and Motion from Line Segments in Multiple Images, *IEEE Trans. on Pattern Analysis.* Machine Intelligence, núm. 17, vol. 11, 1995.
- J.E.W. Mayhew & J.P Frisby, 3D Model Recognition from Stereoscopic Cues, MIT Press, 1991.
- R. M. Haralick and L.G.Shapiro, *Computer* and Robot Vision II (Addison-Wesley Pub. Co., 1993).

- E. Grosso, G.Metta, A.Oddera & G.Sandini, Robust Visual Servoing in 3-D Reaching Tasks, *IEEE Trans. on Robotics and Automation*, vol. 12, núm. 5, pp. 732-741, 1996.
- S. B. Pollard, J. E. W. Mayhew and J. P. Frisby, A stereo correspondence algorithm using a disparity gradient limit, *Perception*, 14:449-470, 1985.
- E. Trucco & A. Verri, *Introductory Techniques for 3-Computer Vision*, Prentice Hall, pp. 26-28, 1998.
- R. Szeliski, From Image to Model (and Beyond): A Personal Restrospective, *Vision Interface*, núm 97, Kelowna, B.C, Mayo 21, 1997.
- M. Chavarría, A. Soria & P. Wiederhold, Percepción monocular para el control de robots, Congreso Latinoamericano de Control Automático, Guadalajara – México, D.F., 2002.
- R. C. González y R. E. Woods, *Tratamiento Digital de Imágenes*, Addison Wesley / Díaz de Santos, 1996.
- J. H. Sossa, Características Invariantes para el Reconocimiento de Formas en Visión por Computadora: una Panorámica. Instituto Politécnico Nacional, *Centro de Investigación en Computación*, núm. 33, serie Azul, México, 1999.
- 12. E. Hecht y A. Zajac, *Óptica*, Colección Fondo Educativo Intermericano, 1977.
- J. Denavit & R.S. Hartenberg, A Kinematic Notation for Lower-Pair Mechanisms Based on Matrices, *Journal of Applied Mechanics*, pp. 215-221, 1955.