

<h1>GIAR</h1>	Universidad Tecnológica Nacional Facultad Regional Buenos Aires Secretaría de Ciencia y Tecnología	Versión 01-01  12 y 13 /mayo/2010
	<h2>Seminario IA y R</h2>	Página 1 de 6

### 3 "3D Mapping

## Mapeo tridimensional de la topología del entorno median sistema de visión artificial estereoscópica"

Leandro Di Matteo, Ramiro Verrastro, Rodolfo Cignoli,  
Ignacio Bertacchini, Ing. Alexis Peppe Pardini

**Resumen** – En este trabajo se presenta el Desarrollo de Sistema de Visión Artificial de visión estereoscópica. El sistema es capaz de construir un MAPA DE PROFUNDIDAD de una toma de imagen dada en tiempo real, dicha información es sumamente útil para que un robot móvil pueda esquivar obstáculos inmediatos y dinámicos, mediante la navegación reactiva.

La característica más importante de este trabajo es la capacidad de creación de una REPRESENTACIÓN VIRTUAL del mundo estático a través de las imágenes adquiridas. Este mapa puede ser utilizado por algoritmos de navegación planificada, por campos de potencial por ejemplo.

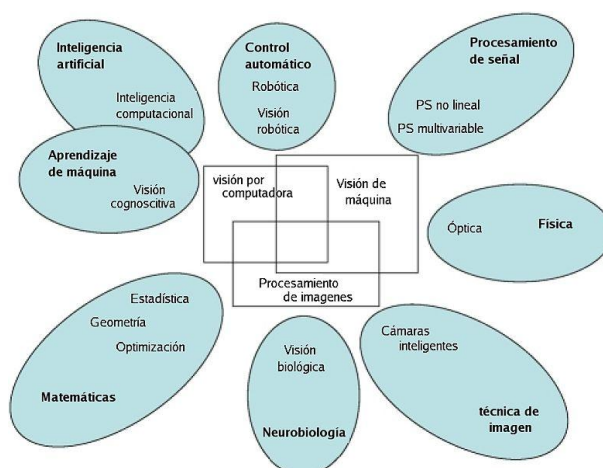
### 3.1 Introducción

La Visión artificial, también conocida como Visión por Computador (del inglés Computer Vision) o Visión técnica, es un subcampo de la inteligencia artificial.

El propósito de la visión artificial es programar un computador para que "entienda" una escena o las características de una imagen

Los objetivos típicos de la visión artificial incluyen:

- \* La detección, segmentación, localización y reconocimiento de ciertos objetos en imágenes (por ejemplo, caras humanas).
- \* Registro de diferentes imágenes de una misma escena u objeto, ej: hacer concordar un mismo objeto en diversas imágenes.
- \* Seguimiento de un objeto en una secuencia de imágenes.
- \* Mapeo de una escena para generar un modelo tridimensional de la escena; tal modelo podría ser usado por un robot para navegar por la escena.
- \* Estimación de las posturas tridimensionales de humanos.
- \* Búsqueda de imágenes digitales por su contenido.



**Figura 1** - Esquema de relaciones entre visión por computadora y otras áreas afines

<b>GIAR</b>	Universidad Tecnológica Nacional Facultad Regional Buenos Aires Secretaría de Ciencia y Tecnología	Versión 01-01
	<b>Seminario IA y R</b>	12 y 13 /mayo/2010  Página 2 de 6

En Óptica, Optometría y Oftalmología, la visión binocular o visión estereoscópica es la capacidad que tiene un ser vivo de integrar las dos imágenes que está viendo en una sola por medio del cerebro (Sistema nervioso central). Éste último es el encargado de percibir las sensaciones que tanto un ojo como otro están viendo y de enviar una respuesta única y en tres dimensiones.

### 3.2 Objetivo

Desarrollo de Sistema de Visión Artificial de visión estereoscópica

El sistema debe ser capaz de construir un MAPA DE PROFUNDIDAD de una toma de imagen dada en tiempo real.

Información útil para que un robot móvil pueda esquivar obstáculos inmediatos y dinámicos.  
Navegación reactiva.

Creación de una REPRESENTACIÓN VIRTUAL del mundo estático a través de las imágenes adquiridas.

Utilizada por algoritmos de navegación planificada, por campos de potencial por ejemplo.

El SISTEMA debe ser ABIERTO para constituir un subsistema dentro de un sistema de visión artificial con más funciones.

Ej: interacción con subsistema de detección de movimiento, de identificación de objetos, de identificación de seres animados (personas, animales, móviles en general), de identificación de landmarks, de OCR para detección de indicaciones, etc.)

El SISTEMA debe ser FLEXIBLE y ESTRUCTURADO, para poder adicionarle más atributos y poder mejorar sus capacidades

Cada etapa debe estar construida como una caja negra con parámetros de entrada y salida (homologados) para poder intercambiar cualquier bloque por otro mejorado sin alterar el funcionamiento del sistema.

### 3.3 Desarrollo

El proyecto consta de las siguientes etapas:

#### 1. PREPROCESAMIENTO

- 1.1 CALIBRACIÓN DE VIDEO CAMARAS
- 1.2 PRE PROCESAMIENTO DE IMAGENES

#### 2. CORRELACION DE PUNTOS CARACTERISTICOS

- 2.1 BUSQUEDA DE PUNTOS CARACTERISTICOS EN UNA IMAGEN
- 2.2 CORRELACION DE PUNTOS CARACTERISTICOS

#### 3. UBICACIÓN DE PUNTOS EN ESPACIO 3D

- 3.1 UBICACIÓN DE PUNTOS EN EL ESPACIO
- 3.2 ROBUSTECIMIENTO DE PUNTOS ENCONTRADOS
- 3.3 SEGMENTACION DEL ESPACIO POR REGIONES

#### 4. VISUALIZACION VIRTUAL DEL "MUNDO"

- 4.1 MAPA BIDIMENSIONAL DE PROFUNDIDAD

<h1>GIAR</h1>	Universidad Tecnológica Nacional Facultad Regional Buenos Aires Secretaría de Ciencia y Tecnología	Versión 01-01  12 y 13 /mayo/2010
	<h2>Seminario IA y R</h2>	

### 4.2 MAPA TRIDIMENSIONAL “MUNDO VIRTUAL”

## 5. INFORMACIÓN PARA NAVEGACION DE ROBOT MOVIL AUTÓNOMO

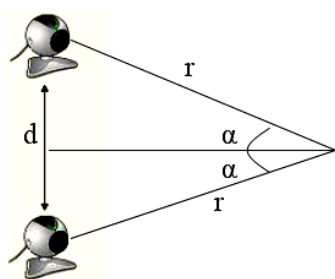
### 5.1 NAVEGACIÓN REACTIVA - Obstáculos inmediatos

### 5.2 NAVEGACIÓN PLANIFICADA - Mapa de ocupación

Se describirán brevemente cada una de dichas etapas:

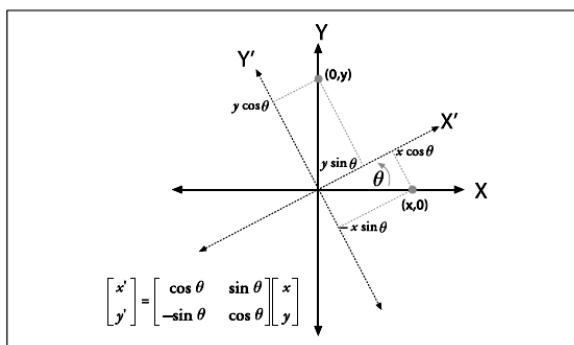
#### 1. PREPROCESAMIENTO

Se realiza la calibración de parámetros extrínsecos de las cámaras y se preprocesa la imagen para que ambas tomas queden normalizadas.



$$d = 2r \cdot \text{sen}(\alpha)$$

**Figura 2** – Sistema de cámaras estereoscópicas



$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

**Figura 3** – Modelo de roto traslación de las imágenes

#### 2. CORRELACION DE PUNTOS CARACTERISTICOS

Se buscan puntos característicos en ambas cámaras, empleándose diferentes técnicas como por ejemplo el filtro Canny, el cual deriva la imagen para resaltar sus bordes, y funciones para resaltar las esquinas notables dentro de cada imagen cuya función característica se representa en la matriz H(p) indicada.

$$H(p) = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix}_p$$

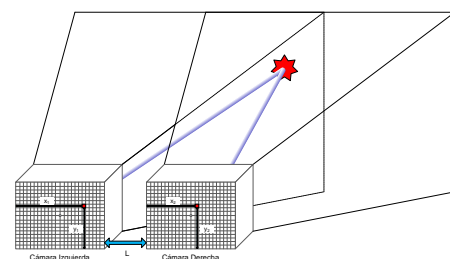
**Figura 4** – Matriz utilizada para la detección de esquinas

#### 3. UBICACIÓN DE PUNTOS EN ESPACIO 3D

Una vez encontrados los puntos característicos en ambas imágenes se trata de encontrar puntos homólogos en ambas cámaras utilizando diferentes técnicas, entre ellas la correlación cruzada normalizada, indicada en la expresión de la derecha.

$$R_{\text{coeff\_normed}}(x,y) = \frac{R_{\text{coeff}}(x,y)}{Z(x,y)}$$

Una vez hallados los puntos característicos homólogos, se aplican las ecuaciones fotogramétricas de estereoscopía para determinar la coordenda (x,y,z) del punto correspondiente. Esto se ilustra en la figura de la derecha.

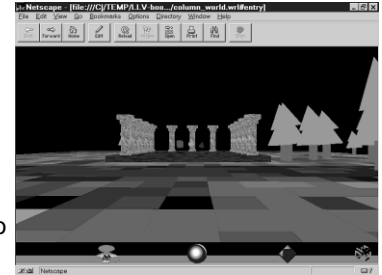


**Figura 5** – Determinación de un punto en el espacio tridimensional objeto mediante las cámaras estereoscópicas

<h1>GIAR</h1>	Universidad Tecnológica Nacional Facultad Regional Buenos Aires Secretaría de Ciencia y Tecnología	Versión 01-01
	<h2>Seminario IA y R</h2>	12 y 13 /mayo/2010  Página 4 de 6

#### 4. VISUALIZACION VIRTUAL DEL "MUNDO"

Cada punto encontrado en el espacio es mapeado dentro de un entorno virtual tridimensional, y punto a punto se va reconstruyendo la topología del entorno, conformándose nubes de puntos que agrupadas corresponden a objetos en el espacio.



**Figura 6** – Ejemplo de mundo virtual

#### 5. INFORMACIÓN PARA NAVEGACION DE ROBOT MOVIL AUTÓNOMO

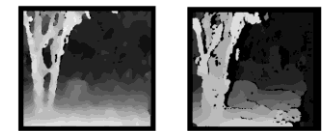
La información del mundo virtual anterior es utilizada como mapa para la navegación de robots móviles autónomos.

En NAVEGACIÓN REACTIVA se utilizan mapas de profundidad para esquivar obstáculos inmediatos

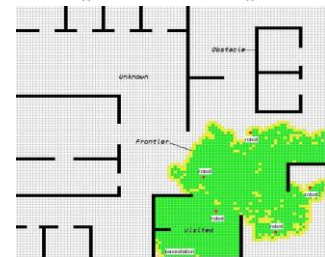


**Figura 7** – Mapa de profundidades

En NAVEGACIÓN PLANIFICADA se utiliza el Mapa de ocupación para planificar una trayectoria valida a seguir.



**Figura 8** – Mapa virtual subdividido en celdas



### 3.4 Experiencia

Se logra una muy buena detección de puntos característicos, de todos modos tendremos que continuar trabajando con nuevas técnicas para obtener más cantidad de dichos puntos



**Figura 9** – Muestra de hallazgo de puntos característicos sobre la puerta del laboratorio del grupo GIAR

Se obtiene la correlación y se hallan aproximadamente 50 puntos homólogos por cuadro. Es un número suficiente para nuestro algoritmo de reconstrucción. De todos modos seguiremos investigando algoritmos que introduzcan mejoras en técnica de correlación de puntos

<h1>GIAR</h1>	Universidad Tecnológica Nacional Facultad Regional Buenos Aires Secretaría de Ciencia y Tecnología	Versión 01-01  12 y 13 /mayo/2010
	<h2>Seminario IA y R</h2>	Página 5 de 6



**Figura 10** – Hallazgo de puntos característicos homólogos mediante técnica de correlación cruzada

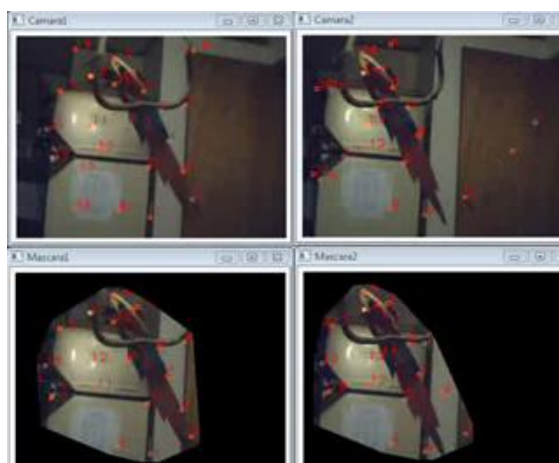
De acuerdo a los ensayos, toda nueva operación que se incorpora al procesamiento consume recursos adicionales de procesador. A medida que el nivel de complejidad de los cálculos aumenta, también aumenta la necesidad de tener un procesamiento a mayor velocidad, para ello comenzaremos a trabajar con procesamiento a mayor velocidad con el sistema CUDA.

### 3.5 Resultados

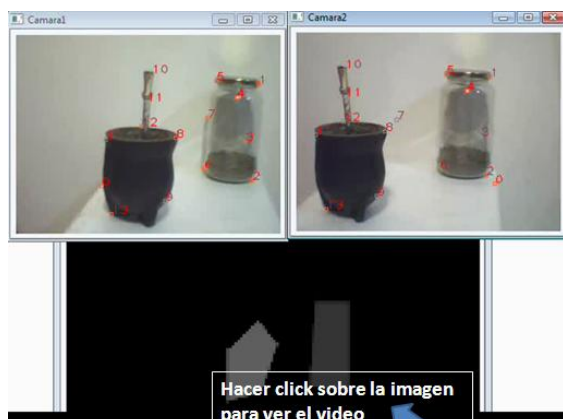
Se segmentaron objetos de acuerdo a la distancia relativa a la que se encuentra del sistema de cámaras estereoscópicas.

Se Logró conformar el mapa de profundidades, observado en la fgura de la derecha, donde para una mejor representación visual, se asigna un nivel de gris a cada objeto de acuerdo a la distancia que lo separa de las cámaras.

A través de la herramienta XNA se construyó el mapa tridimensional dentro del mundo virtual. Los resultados logrados se han obtenido con caras poligonales conformadas por los puntos detectados. Como próximo paso se trabajará en construir el mapa en base a cúmulos de puntos aislados, utilizando técnicas de segmentación de puntos en el espacio.



**Figura 11** – Segmentación de objeto a una distancia determinada de las cámaras de video



**Figura 12** – Ejemplo de mapa de profundidades



**Figura 13** – Ejemplo de reconstrucción 3D

<b>GIAR</b>	Universidad Tecnológica Nacional Facultad Regional Buenos Aires Secretaría de Ciencia y Tecnología	Versión 01-01
	<b>Seminario IA y R</b>	12 y 13 /mayo/2010  Página 6 de 6

### 3.6 Conclusiones

El sistema ha sido capaz de construir un MAPA DE PROFUNDIDAD de una toma de imagen dada en tiempo real, información útil para que un robot móvil pueda esquivar obstáculos inmediatos y dinámicos. Navegación reactiva.

Se ha logrado crear una REPRESENTACIÓN VIRTUAL del mundo estático a través de las imágenes adquiridas.

Continuaremos ampliando el alcance del proyecto, en tópicos como: adición de Tags sobre el mundo virtual para identificación de objetos dentro del mapa, solapamiento de imágenes del mundo real al texturado del mundo virtual para una mejor interpretación del mapa por parte de operadores humanos, aplicación de métodos tomográficos de mapeo para determinar con mayor precisión objetos dentro del mapa y aplicación de técnica de SLAM para generación dinámica de mapa y auto posicionamiento dentro del mismo.

Este trabajo en curso es el punto de partida para amplios trabajos de investigación y desarrollo con aplicación a distintos campos científicos y tecnológicos.

### 3.7 Referencias

1. Pedram Azad, Tilo Gockel, R. Dillmann (2008). Computer Vision-Principles and Practice. Elektor International Media BV. ISBN 0905705718 <http://ivt.sourceforge.net/book.html>
2. Dana H. Ballard and Christopher M. Brown (1982). Computer Vision. Prentice Hall. ISBN 0131653164. <http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/bandb.htm>.
3. Wilhelm Burger and Mark J. Burge (2007). Digital Image Processing: An Algorithmic Approach Using Java. Springer. ISBN 1846283795 and ISBN 3540309403. <http://www.imagingbook.com/>.
4. James L. Crowley and Henrik I. Christensen (Eds.) (1995). Vision as Process. Springer-Verlag. ISBN 3-540-58143-X and ISBN 0-387-58143-X.
5. E. Roy Davies (2005). Machine Vision : Theory, Algorithms, Practicalities. Morgan Kaufmann. ISBN 0-12-206093-8.
6. Olivier Faugeras (1993). Three-Dimensional Computer Vision, A Geometric Viewpoint. MIT Press. ISBN 0-262-06158-9.
7. R. Fisher, K Dawson-Howe, A. Fitzgibbon, C. Robertson, E. Trucco (2005). Dictionary of Computer Vision and Image Processing. John Wiley. ISBN 0-470-01526-8.
8. David A. Forsyth and Jean Ponce (2003). Computer Vision, A Modern Approach. Prentice Hall. ISBN 0-12-379777-2.
9. Gösta H. Granlund and Hans Knutsson (1995). Signal Processing for Computer Vision. Kluwer Academic Publisher. ISBN 0-7923-9530-1.
10. Richard Hartley and Andrew Zisserman (2003). Multiple View Geometry in Computer Vision. Cambridge University Press. ISBN 0-521-54051-8.
11. Gérard Medioni and Sing Bing Kang (2004). Emerging Topics in Computer Vision. Prentice Hall. ISBN 0-13-101366-1.
12. Tim Morris (2004). Computer Vision and Image Processing. Palgrave Macmillan. ISBN 0-333-99451-5.
13. Nikos Paragios and Yunmei Chen and Olivier Faugeras (2005). Handbook of Mathematical Models in Computer Vision. Springer. ISBN 0-387-26371-3.
14. Azriel Rosenfeld and Avinash Kak (1982). Digital Picture Processing. Academic Press. ISBN 0-12-597301-2.
15. Linda G. Shapiro and George C. Stockman (2001). Computer Vision. Prentice Hall. ISBN 0-13-030796-3.