

## PROCESAMIENTO DIGITAL DE IMÁGENES

Dr. Johnson Garzón

# SISTEMA DE VISION ARTIFICIAL PARA RECONOCIMIENTO DE OBJETOS APLICADO A UN BRAZO ROBÓTICO



Realizado por :  
Carlos Pillajo

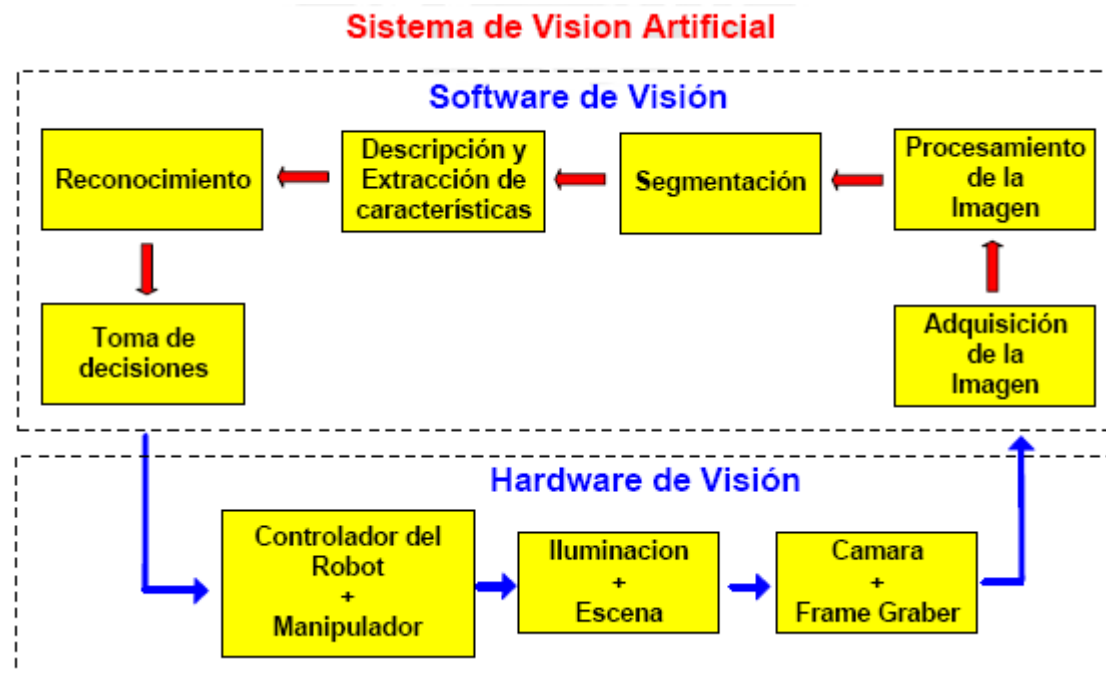
En este proyecto, un brazo robot permitirá seleccionar objetos (Triángulos, Cuadrados, Esferas, etc) que se encuentran en una mesa, independiente de la posición y orientación. El problema se aborda mediante un esquema de Visión Artificial consistente en 6 etapas: obtención de la imagen, pre procesamiento, segmentación, extracción de características, clasificación y manipulación con el brazo robótico tipo SCARA.

Se implementa técnicas de clasificación automáticas para un sistema de visión. La clasificación en si se puede enfrentar de múltiples maneras, generalmente los clasificadores tradicionales, que se basan en métodos estadísticos y/o estructurales y otro esquema como la red neuronal el que presenta algunas ventajas frente a los métodos tradicionales y el cual será realizado.

Una vez reconocida y localizada una pieza determinada, se dará la señal de mando al manipulador robótico para que este lo recoja y lo coloque en una posición determinada previamente por el operador.



La entrada a un sistema de VA es una imagen obtenida por un elemento de adquisición, como en este caso una cámara asociada a una tarjeta FPGA's , mientras que su salida es una descripción de la escena, la cual ha sido obtenida a partir de la imagen. Por un lado, esta descripción debe estar relacionada de algún modo con aquella realidad que produce la imagen y, por el otro, debe contener toda la información requerida para la tarea de interacción con el medio ambiente que se desea llevar a cabo, por ejemplo mediante un brazo robótico tipo SCARA



## Hardware de Visión

El hardware que se utiliza, que tiene como finalidad capturar imágenes digitales de los objetos a reconocer, son:

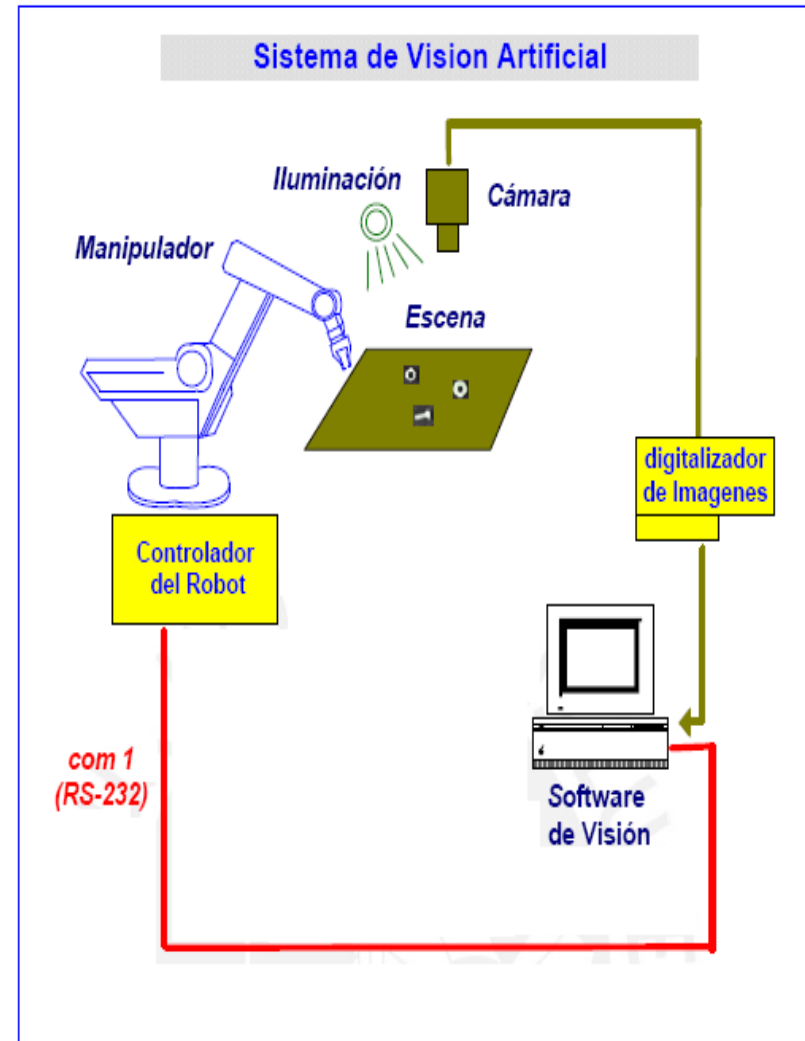
**a. Sistema de iluminación:**

**b. Cámara de Video:** corresponde a una cámara de video CCD con formato NTSC (National Television Standard Committee), la cual se utiliza en la obtención de imágenes de los objetos.

**c. Tarjeta FPGA's (Field Programmable Gate Array):** es una tarjeta de adquisición de datos, imágenes Analógico/Digital

**d. Computadora:** para el desarrollo y ejecución de los programas

**e. Brazo manipulador tipo SCARA** y su controlador: es el elemento que seleccionará las piezas deseadas.



El Sistema de VA debe realizar las siguientes funciones:

- Presprocesamiento de la Imagen, conversión a escala de grises, filtrado
  - Segmentación ,
  - Operaciones morfológicas,
  - Correspondencias.
-

La segmentación es el proceso mediante el cual una imagen se descompone en regiones o elementos que pueden corresponder a objetos o parte de objetos. **El proceso de segmentación se encarga de evaluar si cada pixel de la imagen pertenece o no al objeto de interés.** Esta técnica de procesamiento de imágenes idealmente genera una imagen binaria, donde los píxeles que pertenecen al objeto se representan con un 1, mientras que los que no pertenecen al mismo se representan con un 0. Este tipo de particionamiento está basado en el análisis de alguna característica de la imagen, tal como los niveles de gris o la textura.

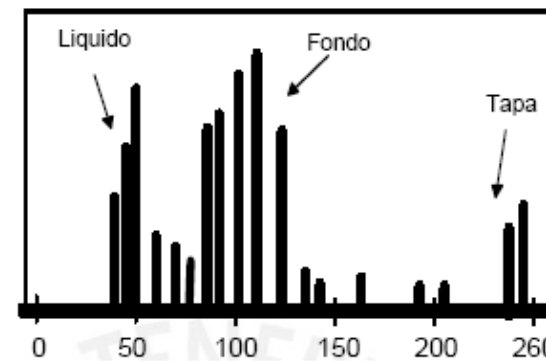
## **Segmentación basado en Píxeles**

Este método de segmentación toma en cuenta sólo el valor de gris de un píxel, para decidir si el mismo pertenece o no al objeto de interés. Para ello, se debe encontrar el rango de valores de gris que caracterizan dicho objeto, **lo que requiere entonces la búsqueda y el análisis del histograma de la imagen.**

---

El objetivo de este método, es el de encontrar de una manera óptima los valores característicos de la imagen que establecen la separación del objeto de interés, con respecto a las regiones que no pertenecen al mismo; debido a esta característica y si los valores de gris del objeto y del resto de la imagen difieren claramente, entonces el histograma mostrará una distribución bimodal, con dos máximos distintos, lo que debiera generar, la existencia de una zona del histograma ubicada entre los dos máximos, que no presenten los valores característicos, y que idealmente fuera igual a cero, con lo cual se logrará una separación perfecta entre el objeto y la región de la imagen que lo circunda, **al establecer un valor umbral ubicado en esta región del histograma**. Por lo tanto cada pixel de la imagen, es asignado a una de dos categorías, dependiendo si el valor umbral es excedido o no.

Segmentación de la imagen de botella



Histograma de la imagen de botella

**son métodos para procesar imágenes binarias basado sobre formas.** Estas operaciones toman una imagen binaria como entrada y dan como resultado una imagen binaria como salida. El valor de cada pixel en la imagen de salida esta basado sobre el correspondiente pixel de entrada y sus vecinos. Dentro de las operaciones morfológicas tenemos la dilatación y erosión.

**La dilatación adiciona pixeles a los limites del objeto (es decir los cambia de off a on), y**

La erosión remueve pixeles sobre los limites del objeto (los cambia de on a off).

---



Sea  $A$  y  $B$  conjuntos de  $Z^2$  y  $\emptyset$  representando al conjunto vacío, la dilatación de  $A$  por  $B$ , se representa  $A \oplus B$ , en la ecuación (1):

$$A \oplus B = \{x / (B^{\wedge} )x \cap A \neq \Phi\} \quad (1)$$

Por tanto, el proceso de dilatación consiste en obtener la reflexión de  $B$  sobre su origen, después, cambiar esta reflexión por  $x$ . La dilatación de  $A$  por  $B$  es entonces, el conjunto de todos los desplazamientos  $x$ , tales que  $B$  y  $A$  se solapen en al menos un elemento distinto de cero. Basándose en esta interpretación, la ecuación (1) se puede volver a representar mediante la ecuación (2):

$$A \oplus B = \{x / [(B^{\wedge} )x \cap A] \subseteq A\} \quad (2)$$

Al conjunto  $B$ , se le conoce normalmente como el *elemento de estructura de la dilatación*. Como vemos, se toma el elemento de estructura  $B$  como una máscara de convolución. Aunque la dilatación se basa en operaciones de conjunto, mientras que la convolución se basa en operaciones aritméticas, el proceso básico de mover a  $B$  respecto a su origen, y desplazarlo después sucesivamente de tal forma que se deslice sobre el conjunto (imagen)  $A$ , es análogo al proceso de convolución

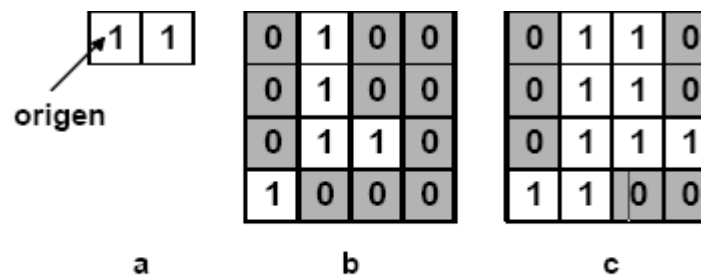


Figura Ejemplo de dilatación. (a) Elemento estructural  $B$  (b) Imagen  $A$ . (c) Resultado de la dilatación  $A \oplus B$

Para los conjuntos A y B de  $Z^2$ , la erosión de A por B representada por  $A \ominus B$ , se define como la ecuación (3):

$$A \ominus B = \{x / (B)x \subseteq A\} \quad (3)$$

que dice, que la erosión de A por B es el conjunto de todos los puntos x tal que B, trasladado por x, está contenido en A. Como en el caso de la dilatación, la ecuación (3) no es la única definición de la erosión.

Sin embargo, esta ecuación, normalmente, es más adecuada en implementaciones prácticas de la morfología. Si se toma como ejemplo el elemento estructurador que se muestra en la figura (a), y la imagen de entrada de la figura (b), se obtiene la imagen de salida que se muestra en la figura (c), en la cual el proceso de erosión ha sido aplicado.

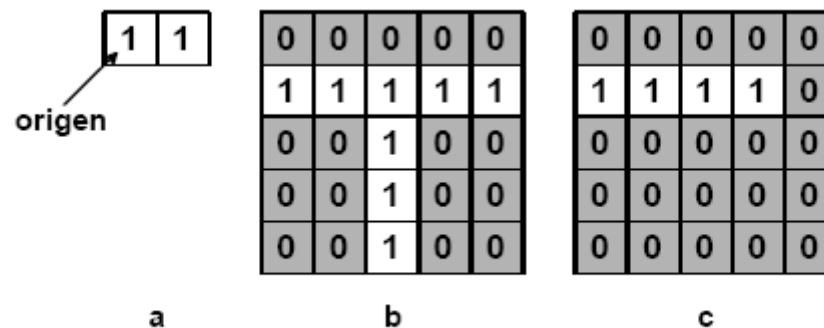


Figura (a) Elemento estructurador (b) Imagen original, (c) Imagen erosionada.

La puesta en correspondencia de elementos homólogos pertenecientes a diferentes imágenes **es una de las tareas actuales en curso de continuo desarrollo** por parte de disciplinas como el Análisis de Imagen, la Visión Computacional y la Fotogrametría. No obstante, el grado de dificultad varía en cada una de estas disciplinas. Por tanto, **el problema de la correspondencia de imágenes ó 'matching' es una línea de investigación incipiente y en continua evolución.**

Método de correspondencia	Índice de correspondencia	Entidades de correspondencia
Por áreas	Correlación, mínimos cuadrados	Niveles de gris
Basada en características	Función de costo	Bordes, regiones
Simbólica	Función de costo	Descripción simbólica

En la correspondencia por áreas las entidades son los niveles de gris. *La idea de este método es comparar la distribución de niveles de gris de una pequeña subimagen, denominada *tesela imagen*, con su correspondiente en la otra imagen.*

**La correspondencia entre dos píxeles se establece a partir de la correlación (grado de semejanza)** de los niveles de gris correspondientes a los píxeles pertenecientes a sendas máscaras (matrices) centradas sobre los píxeles candidatos.

A partir de un píxel que actúa como referencia (dato) en una imagen (cuya máscara no se mueve) se busca en la otra imagen (o imágenes) el píxel (o píxeles) cuya máscara ofrece el mayor nivel de correlación. Esta máscara de correlación debe moverse por el espacio de búsqueda, con lo que se hace crucial poder acotar este espacio.

La idea es medir el grado de semejanza del patrón con la ventana de correspondencia calculando el factor de correlación. (Kreiling, 1976) fue de los primeros que utilizó la correlación digital

**Factor de correlación cruzado.** - El coeficiente de correlación  $\rho$  se define como:

$$\rho = \frac{\sigma_{LR}}{\sigma_L \sigma_R}$$

Si  $\rho$  está normalizado se verifica  $-1 \leq \rho \leq 1$ . En la ecuación anterior los términos se definen como:

$\sigma_{LR}$ , covarianza de las teselas de imagen L (izquierda) y R (derecha).

$\sigma_L$ , desviación típica o estándar de la tesela de imagen L (patrón).

$\sigma_R$ , desviación típica de la tesela de imagen R (ventana de correspondencia).

Si se introducen las funciones imagen  $g_L(x,y)$   $g_R(x,y)$  para las teselas de imagen izquierda y derecha (en este caso el patrón y la ventana de correspondencia) y se notan en la forma  $g_L$ ,  $g_R$  se obtienen las siguientes ecuaciones de definición:

$$\bar{g}_L = \frac{\sum_{i=1}^n \sum_{j=1}^m g_L(x_i, y_j)}{n \cdot m}$$

$$\bar{g}_R = \frac{\sum_{i=1}^n \sum_{j=1}^m g_R(x_i, y_j)}{n \cdot m}$$

$$\sigma_L = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (g_L(x_i, y_j) - \bar{g}_L)^2}{n \cdot m - 1}}$$

$$\sigma_R = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (g_R(x_i, y_j) - \bar{g}_R)^2}{n \cdot m - 1}}$$

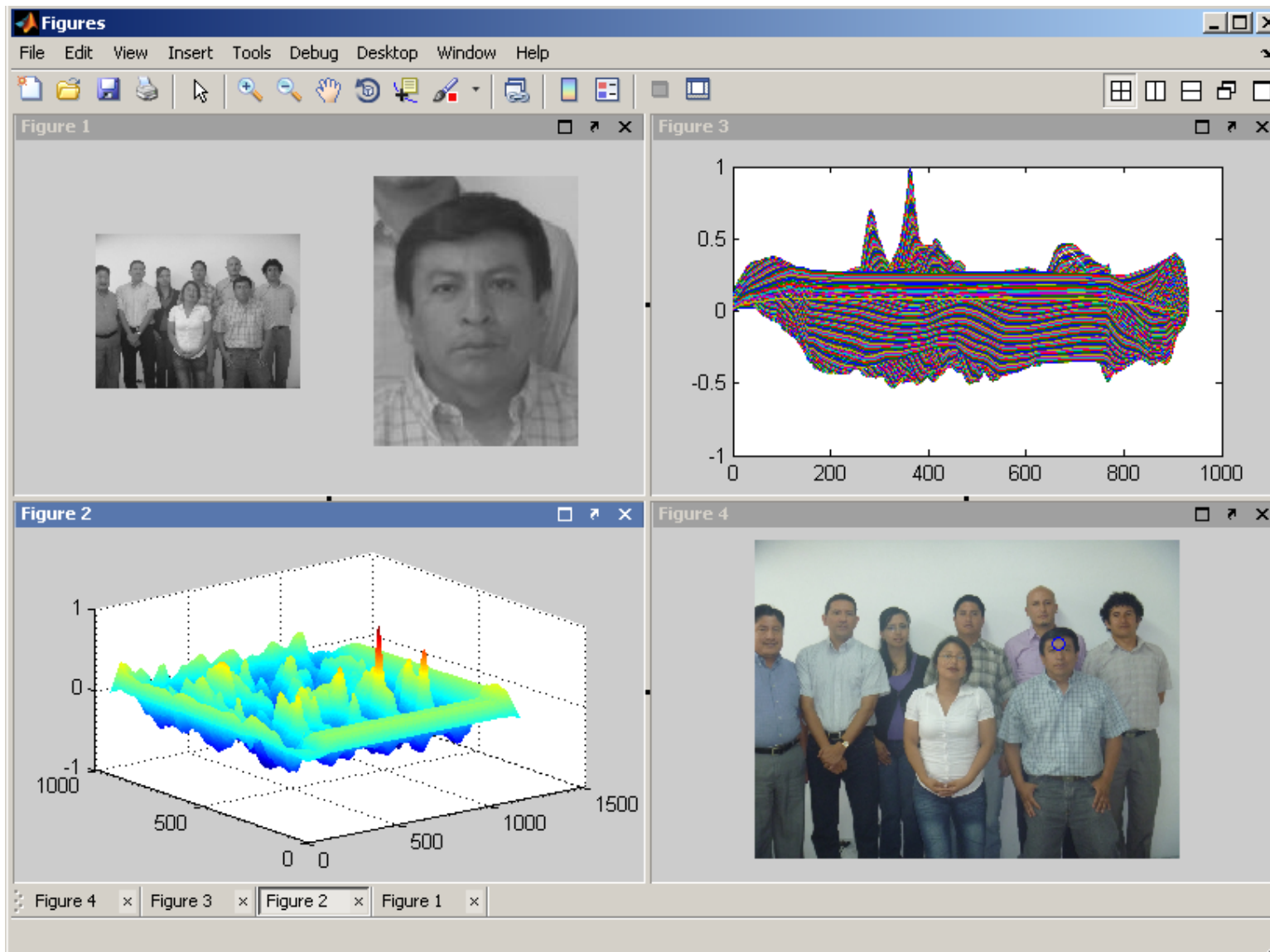
$$\sigma_{LR} = \frac{\sum_{i=1}^n \sum_{j=1}^m ((g_L(x_i, y_j) - \bar{g}_L) (g_R(x_i, y_j) - \bar{g}_R))}{n \cdot m - 1}$$

El factor de correlación cruzado se determina dentro del espacio de búsqueda para cada posición  $f,c$  de la ventana de correspondencia. El problema siguiente es determinar qué posición  $u, v$  proporciona el máximo factor de correlación.

El factor de correlación máximo cruzado normalizado  $\rho$  toma valores en el intervalo  $[-1, 1]$ . Se obtiene un valor igual a la unidad si el patrón y la ventana de correlación son idénticos. Si no existe correlación entre las dos teselas de imagen, es decir si no hay coincidencia en absoluto, entonces  $\rho=0$ .

```
%%% Ejercicio básico para identificar patrones
close all , clear all, clc      % limpio variables y pantalla
%%% LEO LOS IMAGENES
A = imread('FotoGrp.JPG'); % Imagen total en donde se desea identificar la imagen tipo
B = imread('FotoCP.JPG'); % Imagen tipo
% convierto a escala de grises las imágenes
Agris = rgb2gray(A);
Bgris = rgb2gray(B);
figure(1)
subplot(1,2,1), imshow(Agris);
subplot(1,2,2) , imshow(Bgris);
%%% Se obtiene la Correlación cruzada normalizada bidimensional
C = normxcorr2(Bgris,Agris);
figure(2), mesh(C)
figure(3), plot (C)
[vx ix] = max(max(C));
[vy iy] = max(max(C'));
figure(4); imshow(A); hold on;
plot((ix-(length(Bgris(:,1)))/2)+20),(iy-(length(Bgris(1,:)))/2)-50,'O');
```

---





## BIBLIOGRAFIA.

- [1] Diego Gonzales Aguilera, “Correspondencia de Imágenes”, Universidad de Salamanca, 2008.
- [2] Moreira Quiroz José ,Chavez. B. Patricia., «Implementación de un algoritmo para la detección y conteo de células en imágenes microscópicas», 2009.
- [3] Eddie Sobrado, “Sistema de Visión Artificial para el reconocimiento y manipulación de objetos utilizando un brazo robot”, Pontificia Universidad Católica del Perú, 2003



MÉTODO	PROPOSITO	CARACTERÍSTICA
técnicas de shape-signature	cuyo objetivo es encontrar la sub-curva común más larga de dos curvas 2D. Estas curvas se convierten primero en cadenas llamadas de shape-signature y se aplican técnicas de comparación de cadenas para encontrar sub-cadenas coincidentes,	Este método resuelve el problema de las curvas en general sin depender de un conjunto de puntos especiales en las curvas
Otro método conocido utiliza modelos ocultos de Markov (HMM).	Las características de una imagen son convertidas en código y representan la entrada de estos modelos. Dichos modelos, con un enfoque probabilístico, tratará de reconocer la propia imagen	Los HMM son modelos matemáticos de procesos estocásticos
metodología de la curvatura del Espacio Escala (CSS)	Una especie de medida local que representa los límites físicos de los objetos	Se utiliza una caja de luz para obtener imágenes de siluetas que se clasifican como convexas o cóncavas