

Percepción Computacional

Fis. Mat. Pablo Manrique Ramírez
Profesor e Investigador del CINTEC-IPN.
Mat. Ricardo Barrón Fernández
Jefe del Departamento de Matemáticas Aplicadas del CINTEC-IPN.

La percepción computacional (percepción artificial) tiene la finalidad de la reproducción artificial (que engloba subsistemas no informáticos) del sentido de la vista. Se trata de un objetivo excepcionalmente ambicioso y complejo, que se encuentra actualmente en sus inicios, pero que avanza excesivamente rápido. Parece natural pensar que el objetivo de dotar a las máquinas del sentido de la vista supondrá un salto cualitativo en sus capacidades de actuación.

Una manera rudimentaria de definir **percepción artificial** o **percepción computacional**: es la que se basa en un computador. En estos tiempos no se ha podido llegar a dotar a una máquina del sentido de la vista en su totalidad, se han logrado pequeños avances desde un punto de vista sensorial, gracias a la tecnología de los microprocesadores, pero aún existe una gran complejidad de procesamiento.

Haciendo una analogía entre el sistema artificial de procesamiento de imágenes y el sistema visual humano, se tendría el **diagrama 1**.

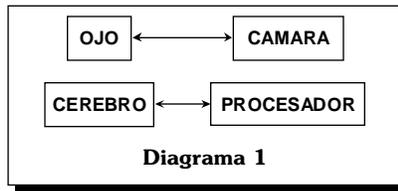


Diagrama 1

Desarrollo

Antes de abordar el tema, hablaremos acerca de una señal eléctrica (un voltaje de tensión) que varía en forma continua en el tiempo denominada señal de video. En ella existen dos tipos de impulsos de sincronía, la vertical y la horizontal. En la práctica se manejan dos formatos de señal de video: el americano y el europeo, cuyas comparaciones entre sí son:

- Difieren en los límites de los tiempos de sincronización (Esto se debe a las diferentes frecuencias de la red eléctrica a la que trabajan; 60Hz para la red americana y 50Hz para la red europea).

- Formatos iguales (Ambos tratan de codificar líneas y cuadros de imagen).

Una vez que se hizo una breve aclaración de la señal de video, para lograr un análisis automático es necesario definir las etapas que se muestran en el **diagrama 2**.

- o El pre-procesamiento es la adquisición de imágenes el cual consiste en digitalizar una señal de video, con la información luminosa de la imagen a interpretar. Naturalmente, la señal continua de video deberá ser:

- muestreada y,
- digitalizada binariamente para su tratamiento en el computador.

El efecto combinado del número de líneas (consustancial a la cámara de video) y el de la frecuencia de muestreo (es decir, el número de muestras de la señal de video que se

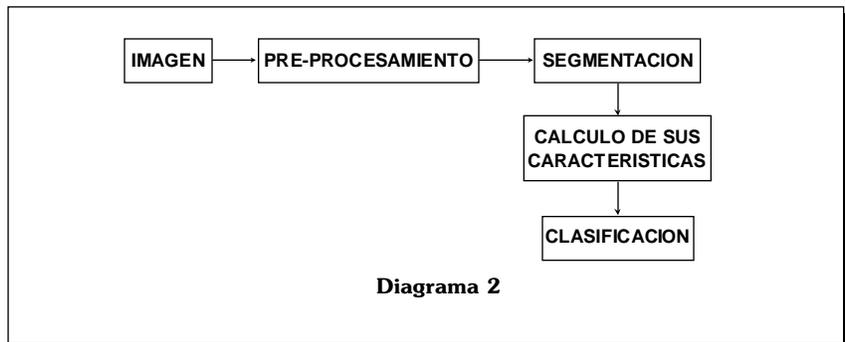


Diagrama 2

toma por cada línea) da lugar a lo que se conoce como **nivel de resolución espacial de la imagen**.

Con el número de muestras que se toman de la señal de video por cada línea se forma una matriz de $N \times M$ elementos (N = número de líneas, M = número de muestras por línea); cada elemento de esta matriz va conformar un pixel que incorpora el nivel de luminosidad del punto correspondiente de la imagen digital. Por lo general se trabaja con matrices cuadradas $M=N$.

Antes de enviar la imagen electrónica al computador, se codifica en binario los niveles de intensidad de cada pixel. El número de bits empleados por pixel se conoce como la **resolución en niveles de gris**; cabe destacar que con los efectos combinados de las resoluciones espacial y de los niveles de gris se pueden apreciar las imágenes.

- La segmentación es un proceso orientado a seccionar la imagen digital en zonas disjuntas con significado propio. Este significado dependerá de la imagen, así como de los objetivos que se persigan con su interpretación. La operación es común en las dos segmentaciones; lo que varía es aquello que se desea segmentar. En un caso, zonas; en el otro, ya dentro de una zona concreta, objetos individuales.

Las técnicas de segmentación están orientadas a objetos o a agrupaciones de objetos, y se pueden dividir en dos grandes grupos:

Aplicación de umbrales de nivel de gris, mediante el empleo de histogramas.

Aplicación de rasgos comunes,

en base al reconocimiento de formas.

- El cálculo de las características consiste en la extracción de rasgos (o bordes); es decir, dado un objeto físico dentro de una imagen se puede obtener, a partir de su contorno (resultado de la segmentación), un conjunto de parámetros característicos como: perímetro, grado de circularidad, momentos de orden, etc. Estos rasgos hacen posible distinguir objetos mediante el uso de un número reducido de propiedades. Los rasgos de un objeto forman un vector $\mathbf{X}=(X_1, X_2, \dots, X_n)^T$, que particularizando en ciertos valores numéricos (los del objeto a clasificar) determina un punto en el espacio n -dimensional de las características.

- La clasificación de los objetos, se realiza al definirlos en la memoria una vez que se segmentaron y se extrajeron ciertas características propias para reconocer un objeto.

cepción Artificial, teniéndose a continuación una breve descripción de cada bloque funcional.

- Sensor óptico. Normalmente entrega una señal de video estandarizada para su posterior digitalización. El conjunto sensor y óptica se integra en una cámara de video.

La etapa sensorial es la adquisición digital de la imagen, iluminación, sensores ópticos, formato de la imagen digital, así como la geometría de la formación de imágenes digitales.

- Convertor analógico/digital; permite digitalizar la señal de video entregada por el subsistema anterior. Aparecen dos parámetros esenciales:

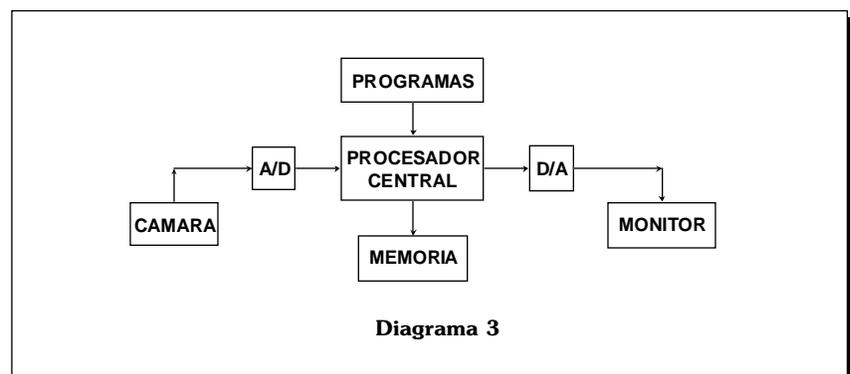
Definición espacial: es decir, número de puntos de digitalización por imagen o cuadro.

Resolución digital: número de bits para codificar los niveles de intensidad luminosa de cada punto de digitalización (que suele denominarse pixel, es decir, elemento de imagen).

- Una señal de video estandarizada, una vez digitalizada, se transforma en una matriz cuadrada (que contiene la información visual digi-

Configuración de un sistema mínimo de Percepción Artificial

El siguiente diagrama (**Diagrama 3**) muestra la organización básica de un sistema mínimo de Per-



talizada) y se almacena en la memoria central de un procesador, para que posteriormente se procese y manipule a través de programas.

- o Monitor de video; permite visualizar tanto las imágenes captadas como los resultados del procesamiento de imágenes. En realidad, se requiere un conversor D/A y, finalmente un monitor de video, ya que este último trabaja con una señal analógica.

Hasta ahora se ha hablado a grosso modo de la etapa sensorial de la percepción artificial. A partir de aquí se habla del procesamiento y de los verdaderos problemas, pues involucra los algoritmos, el tipo de lenguaje de programación y el tipo de procesador que se este utilizando. Uno de los microprocesadores que se utiliza para la adquisición y procesamiento de imágenes es el DSP (Digital Signal Processor).

El análisis automático (informático) propiamente dicho de una imagen digital es aplicado a ciertas operaciones orientadas a mejorar la calidad de la misma; ésto nos lleva a aplicar técnicas y algoritmos de procesamiento, como por ejemplo:

- 1) Conversión de los niveles de gris. (Transformación adecuada de los niveles de gris para mejorar la sensibilidad de la imagen).
- 2) Transformaciones geométricas:
 - Homogéneas
 - Traslación (desplazamiento)
 - Escalamiento
 - Rotación (Se utilizan en la corrección de la perspectiva, y reconstrucción tridimensional de los objetos de una imagen).
- 3) Transformación del histograma. (Es una función donde su dominio son los niveles de gris y su

contradominio el número de pixeles de cada nivel. La función es un diagrama de barras de la propia imagen, que frecuentemente se normaliza entre 0 y 1; por tanto, se emplea la frecuencia relativa de cada nivel de gris).

- 4) Filtrado espacial. (Actúa sobre los niveles de gris de los pixeles de la imagen).
- 5) Filtrado frecuencial. (Es costoso en tiempo y en memoria y se lleva a cabo utilizando la transformada rápida de Fourier)

Para aplicar estas técnicas y algoritmos es necesario definir formalmente una imagen digital. Una imagen digital es una función discreta bidimensional: $f(x,y)$ donde $0 \leq x, y \leq N$ que se denota como sigue:

$$f(x,y) = \begin{matrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,N-1) \end{matrix} \quad (1)$$

Una imagen digital es una matriz cuadrada $N \times N$ binaria, salvo en casos excepcionales, conteniendo cada elemento de la matriz un valor discreto que cuantifica el nivel de intensidad luminosa (también denominado nivel de escalas de grises) del correspondiente elemento de imagen o pixel. Esta matriz binaria es accesible y manejable por un programa ordenador, con lo que se puede en principio realizar cualquier operación sobre ella.

La coordenada (x,y) denota una coordenada espacial de una imagen digital, y el conjunto de puntos en el plano $x-y$ forman el espacio de posición. La resolución espacial $N \times N$ viene determinada por la calidad que se persigue en las imágenes

digitales. A mayor N , mayor calidad, aunque esto obliga a una carga computacional mayor. Por necesidades del formato binario, N es siempre múltiplo de 2, es decir, $N=2^m$ donde m es un entero positivo. En cuanto a la resolución en niveles de gris también es un múltiplo de 2, ya que la cuantificación de los niveles de intensidad luminosa se realizan en código binario.

Por lo tanto, si q es el número de bits empleados para codificar los niveles de gris, entonces:

$$0 \leq f(x,y) \leq 2^q \quad (2)$$

El convenio habitual es codificar los niveles oscuros o negros con valores binarios más bajos, mientras que los niveles claros o blancos con los valores binarios más altos.

Se observa que con la definición anterior de una imagen digital ya se puede realizar un tratamiento matemático discreto, que nos facilitará el tratamiento del procesamiento de las imágenes por computadora.

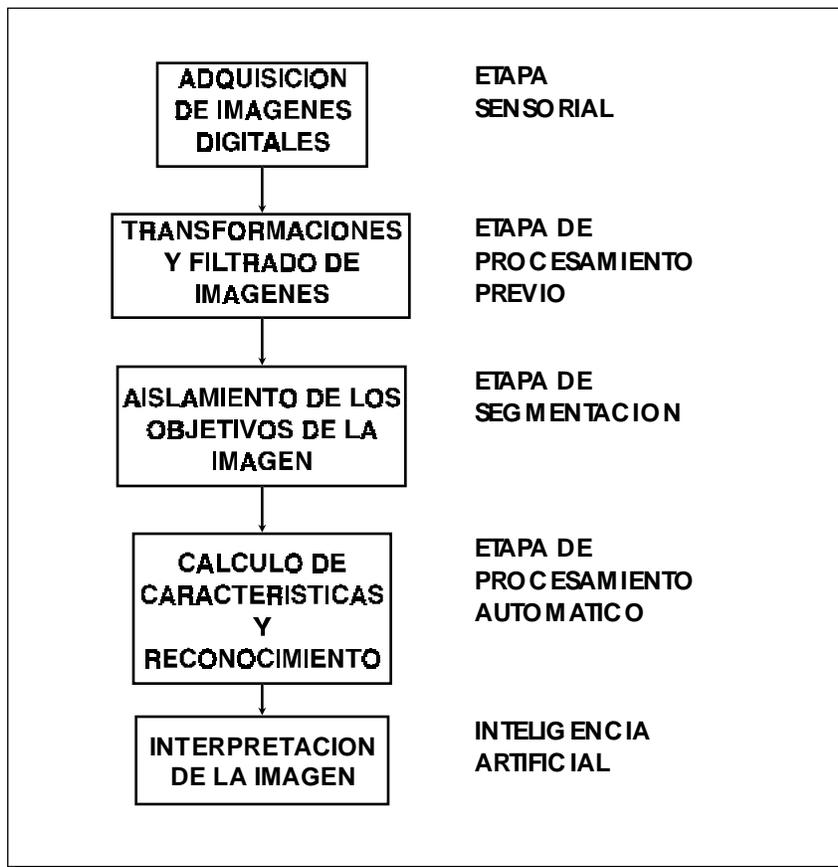
El método en el dominio del espacio, maneja procesamientos definidos por medio de una transformación T , cuyo dominio es el conjunto de pixeles $f(x,y)$ que generan la imagen, es decir:

$$t(x,y) = T(f(x,y)) \quad (3)$$

donde, esta transformación es lineal, homogénea e invariante en la posición.

Un ejemplo de estas transformaciones es: Traslación del procesamiento de las imágenes, al dominio de las frecuencias espacial con la ayuda de las Transformada de Fourier.

$$F\{f(x,y)\} = F(u,v)$$



$$= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) e^{-j2\pi(ux+vy)} dx dy \quad (4)$$

El siguiente diagrama muestra las etapas en que puede organizarse un sistema visual artificial.

Conclusiones

De lo precedente se desprende que el desarrollo inicial de un sistema de procesamiento de imágenes resulta viable, si se conocen de antemano sus características y necesidades, para lograr un ahorro en espacio de memoria y tiempo de ejecución; para esto se eligen adecuadamente las técnicas y algoritmos que se deben de implementar para satisfacer los requerimientos de dicho sistema. La digitalización de la señal de video permite que su información visual pueda ser intro-

ducida y tratada mediante un ordenador. No obstante, el formato de las imágenes digitales es distinto al de la señal de video, entre otros motivos porque se trata de una información discreta, tanto en los valores que puede tomar la propia imagen digital, como en sus argumentos o variables independientes, que son igualmente discretos.

Bibliografía

- o Computer Techniques in Image Processing. H. C. Adresws Academic Press, 1970.
- o Digital Image Processing. W. Pratt. John Willey & Sons, 1978.
- o Fundamental of Digital Image Processing. A. K. Jain. Prentice-Hall, 1989.