

Bloque-1.pdf



Daysapro



Fundamentos de Sistemas Inteligentes en Visión



3º Grado en Ingeniería Informática



**Escuela Politécnica Superior de Córdoba
Universidad de Córdoba**

quieres trabajar en Wuolah??

tú puedes ayudarnos a llevar **WUOLAH**
al siguiente nivel (o alguien que conozcas)

**TE
BUSCAMOS**



sin ánimo de lucro, chequea esto:

quieres trabajar
en Wuolah??

TE BUSCAMOS

BLOQUE 1

Las sesiones de las clases resumidas

Después de leer el resumen intentad hacer las preguntas de otros años que algunas están en wuolah

- Daysa -

Tema 1: Introducción a la visión por computador

1.1 Introducción

La visión por computador es un campo de la inteligencia artificial que busca conocer, modelar y emular los procesos cognitivos humanos que nos permiten obtener información a partir de los imágenes y resolver problemas con ellas. Como objetivos principales podemos destacar cómo tratar una imagen a nivel de código, cómo corregirla o mejorarla, analizar características de la misma o incluso comprender e interpretarla.

1.2 ¿Qué es un SIVA?

Cualquier sistema planteado en base a estos objetivos recibe el acrónimo SIVA (Sistema Inteligente basado en Visión Artificial). Las partes de este sistema serían:

- Adquisición
- Procesado
- Extracción de características
- Segmentación y detección
- Procesamiento de alto nivel
- Toma de decisiones

Tema 2: Adquisición

2.1 Introducción

Son los elementos hardware que forma un SIVA. Sus componentes son:

2.2 Iluminación

La escena que queramos observar debe tener una buena iluminación. La luz, como cualquier onda, tiene distintas longitudes de onda que determinan el color.

Las cámaras reciben la luz por un sensor que captura luz durante un tiempo para después digitalizarla y transmitirla al usuario.

A la vez, los materiales respecto a la luz tienen 3 propiedades: absorbentes (determinan el color del material), reflexivas y transmitivos (opacidad, transparencia...).

También existen distintos sistemas de iluminación: direccional (tiene dirección), difusa (llega a todos los ángulos posibles), a contraluz (desde atrás) y estructurada (patrón definido).

2.3 Lente

Dispositivo por donde la luz se redirige para llegar al sensor. Parámetros:

- Distancia focal: cuanto más pequeña es la distancia focal mayor es el ángulo de vista que puede capturar. Con distancias focales pequeñas se deforma la imagen.
- Diafragma: dispositivo que controla cuanto luz va a entrar en la lente.
- Coeficiente de magnificación M: relación entre el tamaño proyectado digitalmente y el tamaño real del objeto.

$$M = P'/P; f = (S_1 M) / (M + 1)$$

Siendo P' tamaño de la imagen proyectada y P tamaño de la original.

- Profundidad de campo: parte de la imagen que sale enfocada.

2.4 Sensor

sin ánimo
de lucro,
chequea esto:



tú puedes
ayudarnos a
llevar
WUOLAH
al siguiente
nivel
(o alguien que
conozcas)

WUOLAH

Digitaliza la imagen para que la pueda visualizar el usuario. Existen 2 tipos: CCD y CMOS.

- CCD: interpreta la imagen completa para su digitalización. Tiene mejores características de sensibilidad, rango dinámico, ruido...
- CMOS: cada píxel es independiente. Es más rápido pero da menos calidad.

2.5 Hardware de adquisición (capturadora)

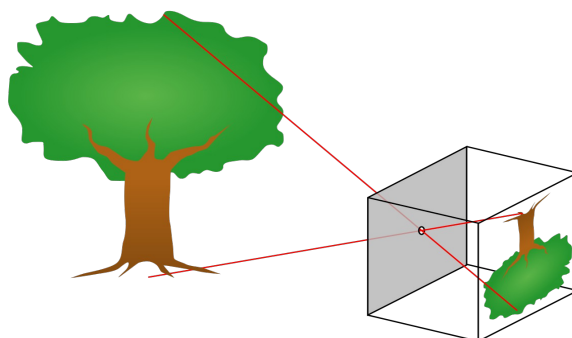
Pasan la información al ordenador. Existen analógicas y digitales.

Tema 3: Formación de la imagen digital

- Modelo fotométrico: lo primero que hay que tener en cuenta es la luz que recibe el sensor. Depende de las propiedades de los materiales en relación a la luz, como vimos anteriormente. La luz se puede expresar en relación a la iluminancia L_i y de la reflectancia f_r , tal que:

$$L_r(\hat{v}_r; \lambda) = \int L_i(\hat{v}_i; \lambda) f_r(\hat{v}_i, \hat{v}_r, \hat{n}; \lambda) \cos^+ \theta_i d\hat{v}_i,$$

- Modelo geométrico: comparamos la lente con un pequeño agujero por el que pasa cada punto a captar de luz en forma de un solo rayo, para captarlo en el sensor.



- Muestreo y “aliasing”: es la cantidad de píxeles que queremos mostrar en la imagen. Cuanto más sean, más afín será la imagen a la realidad. Si se usa un muestreo grande aparecerán distorsiones.
- Cuantificación de la energía: los fotones recibidos en el sensor deben convertirse a voltaje y digitalizarse a valores discretos (de 0 a 255) para ya obtener la señal análoga.
- Codificación: con los valores discretos cada píxel se descompone en RGB (red, green, blue), siendo también cada color un valor discreto de 0 a 255. Todo esto se almacena a nivel digital con una matriz de datos.

Tema 4: Procesamiento de la imagen digital

Vamos a trabajar el procesamiento de la imagen en el dominio espacial, trabajando con la matriz de datos y la relación de los mismos en el espacio y el plano. Trabajamos con 2 tipos métodos:

4.1 Métodos puntuales con una imagen

Con una imagen de entrada obtenemos una de salida, teniendo en cuenta solo el valor de cada píxel de ella. En estos métodos veremos técnicas de cambio del rango dinámico y también técnicas para el espacio de color.

- Normalización del rango del rango dinámico: se cambia el rango en el que están almacenados los valores. Por ejemplo, solemos tener un rango de 0 a 255 y podemos querer que el rango sea de 0 a 1 tratándolo en valores flotantes, con el objetivo de, por ejemplo, realizar raíces cuadradas sobre la imagen. Esta normalización no tiene obligatoriamente que cambiar la apariencia de la imagen. Hay un procedimiento lineal y uno no lineal para obtener este resultado.

Lineal: $R = \min V + \frac{[\min V, \max V] - \min V}{[\max V - \min V]} \cdot [\min V - \max V]$

No lineal (normalización de la media y de la varianza): $I(\mu, \sigma^2) \rightarrow R(\mu=0, \sigma^2=1) \quad R = [I - \mu] / \sigma$



#ESTASREADYCOLACAO

ColaCao®

- Ajuste de gamma, contraste y brillo: un ajuste se realiza según la siguiente fórmula.

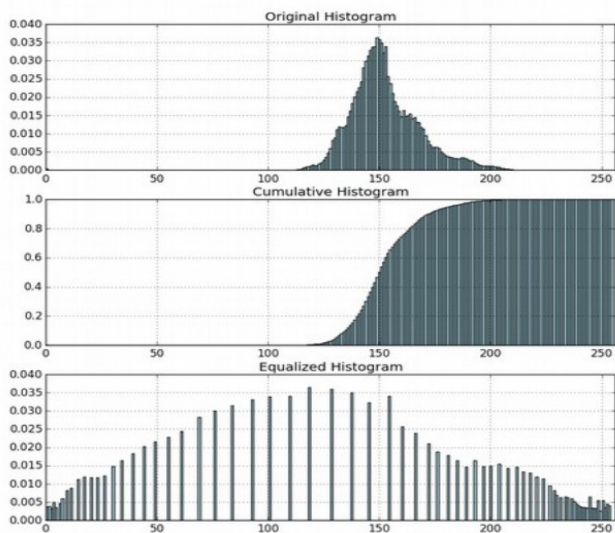
$$R(x, y) = c \cdot I(x, y)^\gamma + b$$

$c \rightarrow$ contraste

$\gamma \rightarrow$ gamma o no linealidad de la respuesta del sensor

$b \rightarrow$ brillo

- Procesamiento de histograma, ecualización: usa el histograma de una imagen viendo el número de píxeles que tiene cada color para normalizarlo haciendo que sus valores sean acumulativos a todos los tonos anteriores (como si de una función de densidad se tratara). Con este histograma normalizado acumulativo buscamos en él cada valor de la imagen original y vemos que nuevo valor obtendría. El resultado final sería una imagen mucho más contrastada y píxeles con todos los valores del rango.



- Look-up tables: son tablas que almacenan los cambios de color para obtener ciertos resultados. Por ejemplo, hacerle el negativo a una imagen. Sobre ellas solo buscamos el valor a cambiar y ponemos el nuevo valor. Esto codificado es mucho más eficiente que hacer la operación píxel a píxel.

- Cambios del espacio de color, HSV/HSI (hue, saturation, value/intensity): a veces nos interesa editar la imagen en vez de tomar la referencia RGB, tomar el tono, saturación e intensidad. Para eso tenemos HSV. Cabe señalar 2 ejemplos de equilibrado de color:

- White patch: busca el punto más luminoso de la escena y se pone blanco. Tras esto se calcula el cambio (factor de corrección) que se ha dado a ese píxel y se hace el mismo a toda la imagen.

- Gray world: se calcula si los tonos medios de RGB son cercanos a 128 para saber si está equilibrado el color. Se hace la transformación:

$$S = I(128/R_m, 128/G_m, 128/B_m), \text{ siendo } R_m, G_m \text{ y } B_m \text{ los tonos medios RGB.}$$

4.2 Métodos puntuales con 2 imágenes

- Operaciones aritméticas (+ - * /):

- Suma: a veces una imagen tiene ruido, y sumar numerosas imágenes iguales con ruido hace que gane nitidez.

- Diferencia: para detectar movimientos en planos estáticos. Si hacemos 2 fotos a la misma escena y algo se ha movido la diferencia sería todo color 0 (negro), menos la silueta del elemento en movimiento.

- Multiplicación/división: pueden servir para reducir iluminaciones no deseadas dividiendo imágenes mal iluminadas entre una imagen tipo que represente los tonos de luz que no queremos

- Operaciones lógicas (AND, OR, XOR, NOT): nos sirve para la composición de imágenes.

4.3 Métodos basados en la vecindad

Las transformaciones no solo dependen de un píxel, si no que también dependen de los píxeles de alrededor de él de la imagen de entrada. La selección de un píxel tomando en cuenta un número de

TE BUSCAMOS


$$g(x, y) = \sum_{s=-a}^a \sum_{t=-b}^b w(s, t) \cdot f(x + s, y + t)$$

Figure 1 illustrates the convolution operation. The diagram shows the original image $f(x, y)$ (a 4x4 grid of 0s and 1s), the padded image $f(x, y)$ (a 6x6 grid with a border of 0s), the initial position for w (a 3x3 grid of 1s, 2s, and 3s), the full correlation result (a 6x6 grid of values), the cropped correlation result (a 4x4 grid of values), the rotated w (a 3x3 grid of 1s, 2s, and 3s), the full convolution result (a 6x6 grid of values), and the cropped convolution result (a 4x4 grid of values).

- Realce de imagen usando “unsharp mask”: sirve obtener una clara diferenciación de las estructuras

tú puedes
ayudarnos a
llevar
WUOLAH
al siguiente
nivel
(o alguien que
conozcas)

Reservados todos los derechos.
No se permite la explotación económica ni la transformación de esta obra. Queda permitida la impresión en su totalidad.

de la imagen. Depende de una ganancia g , y a partir de $g > 1$ se denomina “high-boost filtering”. El proceso es:

$$G = (1 + g)I - gI_L$$

Siendo I_L es la imagen I con un filtro de bajas frecuencias.

- Filtros paso alta, derivadas de la imagen: hacer la derivada de la imagen nos da unos resultados que en algunas situaciones nos puede interesar. La primera derivada sigue el filtro $[-1, 1]$ y la segunda $[1, -2, 1]$.

- Filtros paso alta, filtro de Sobel: hace las derivadas parciales.

$$dx: [[-1, 0, 1], [-2, 0, 2], [-1, 0, 1]]$$

$$dy: [[-1, -2, -1], [0, 0, 0], [1, 2, 1]]$$

- Filtros paso alta, operador laplaciano: hace las segundas derivadas.

$$d^2x + d^2y: [[0, 1, 0], [1, -4, 1], [0, 1, 0]]$$

$$d^2x + d^2y: [[1, 1, 1], [1, -8, 1], [1, 1, 1]]$$

El primero solo toma en cuenta las aristas al píxel y el segundo también las diagonales (3x3 completo).

- Realce de imagen “sharpening”: la imagen original se suma esa imagen con el operador laplaciano.

- Filtrado no lineal:

- Filtro mediana: se coge la mediana de los valores de la ventana y es el valor que se pone. Se obtiene un resultado de suavizado. Se produce una reducción de ruido.

- Filtro bilateral: otro filtro de suavizado. Cuando se usa un gaussiano se suelen perder muchos tonos claros. En el bilateral no. Se basa en hacer un filtro gaussiano y un filtro de similitud a la vez, por lo que se le da mucho protagonismo a los puntos parecidos cercanos.

- Morfología: se basa en 2 operaciones básicas: dilatación y erosión. En la dilatación las zonas claras se expanden y en erosión al contrario.

Image I



Erosion $I \ominus B$



Dialation $I \oplus B$



Opening $I \ominus B = (I \ominus B) \oplus B$



Closing $I \bullet B = (I \oplus B) \ominus B$



Grad(I) = $(I \oplus B) - (I \ominus B)$



TopHat(I) = $I - (I \ominus B)$



BlackHat(I) = $(I \bullet B) - I$

