

# TEMA 1

**Explique qué son y en qué se diferencian los parámetros intrínsecos y extrínsecos de una cámara.**

La transformación entre puntos 3D de la escena y puntos 2D de la imagen es una función no lineal que depende de un vector de varios parámetros. Estos pueden ser extrínsecos o intrínsecos.

Los **intrínsecos** son propiedades inherentes de la cámara y la óptica:  
distancia focal, desplazamiento centro imagen y coeficiente de distorsión

Los **extrínsecos** definen la posición y la orientación de la cámara respecto al sistema de referencia absoluto:  
traslación o rotación de los ángulos

**Con carácter general, ¿Qué pasos seguiría para calibrar una cámara?**

Determinar conjunto de puntos 3D

Determinar sus proyecciones en la imagen

Calcular parámetros que mejor resuelvan las correspondencias

**Explique las fases de muestreo y cuantización dentro del proceso de formación de una imagen digital.**

Muestreo

Digitalización de coordenadas espaciales (Dominio) Conversión que sufren (x,y) como coord. espaciales de la señal analógica generando un pixel

El muestreo es la conversión que sufren las dos dimensiones espaciales de la señal analógica y que genera la noción de pixel.

Cuantización

Digitalización de los valores de la función (Rango) Conversión que sufre el rango de la señal analógica generando el nivel de gris o intensidad

La cuantización es la conversión que sufre el rango de la señal analógica y genera el concepto de nivel de gris o intensidad.

**¿Qué es un espacio de representación del color? Comente las principales características de alguno que conozca.**

Un espacio de representación del color es un método por el que se puede especificar, crear o visualizar cualquier color.

*Es la forma en la que se organizan los colores en una imagen. Son una combinación de colores compuestos de señales de luminancia cromática*

El espacio RGB se basa en la combinación de tres señales de luminancia cromática distinta: el rojo, el verde y el azul.

La manera de conseguir un color concreto es determinar la cantidad de color rojo, verde y azul que se necesita combinar mediante la suma aritmética de las componentes

**En el contexto de una geometría proyectiva 3D, ¿que establece la matriz de transformación en sistemas de coordenadas no coincidentes?**

La transformación de perspectiva se deberá multiplicar por la concatenación de transformaciones : traslación, rotación y traslación

De esta forma los sistemas de cámara y mundo serán coincidentes

## **T2: Extracción Características**

**Dentro del ámbito del procesamiento de imágenes, ¿que entendemos por característica? Ponga algún ejemplo.**

Son atributos extraídos de la imagen que poseen mayor carga semántica que la colección de píxeles que forman la propia imagen.

Ejemplos

-Lineas, regiones, puntos de interes, texturas, color, etc.

Muchas tecnicas de procesamiento, mejora e interpretación emplean dichas características, previamente extraídas.

**Explique en que consiste el proceso de realce de una imagen**

Consiste en la aplicación de técnicas para mejorar la apariencia de una imagen o transformar esta para facilitar posteriores análisis, sin alterar la información presente.

Estos procesos de realce no incrementan la información inherente contenida en los datos, simplemente incrementan el rango dinámico de determinadas características para facilitar su posterior análisis.

Hay tres tipos de transformaciones usadas en el realce de imágenes: lineales, logarítmicas y exponenciales.

**Defina el concepto de correlación espacial de dos señales 2D. ¿Como se implementa? ¿Para que sirve?**

Correlación: Mover la máscara sobre la imagen y calcular la suma de los productos en cada posición

I Los filtros de alisamiento son utilizados para emborronar y reducir ruido en una imagen

**Defina el concepto de convolución espacial de dos señales 2D. ¿Como se implementa? ¿Para que sirve?**

La convolución es el proceso de mover una máscara sobre la imagen e ir calculando la suma de productos en cada posición, rotando previamente el núcleo de convolución inicialmente 180°.

Se implementa multiplicando los valores de vecindad del píxel por los de la máscara de convolución.

Sirve para alisar (emborronar y eliminar ruido).

**Explique que es el ruido de una imagen. ¿Como lo eliminaría? Proponga al menos dos alternativas.**

Información de la imagen que es añadida de forma artificial por el dispositivo de captura debido a la imprecisión del mismo o interferencias del entorno. Se puede eliminar difuminando la imagen con filtros espaciales (media, mediana) o frecuenciales (paso bajo o Butterworth)

**Explique que ventajas tiene trabajar en el espacio transformado de Fourier.**

Permite analizar desde el punto de vista espectral (por frecuencias)

- La conversión inversa (de frecuencia a imagen) no tiene pérdidas.
- Se pueden expresar todo tipo de funciones sin importar su complejidad mediante sumas de senos y cosenos.
- Detectar patrones que se repiten en el espacio.
- Algunos algoritmos son más rápidos de realizar en la transformada

**Comente algunas técnicas de realce frecuencial que conozca. Ventajas e inconvenientes.**

#### FILTRADO FRECUENCIAL

Los saltos bruscos tienen componentes en todas las frecuencias  
Butterworth y gaussiano es gradual

Paso alto elimina bajas frecuencias, quedando los bordes

**Explique las diferencias entre los procesos de filtrado espacial y frecuencial. ¿en qué situaciones funciona mejor cada uno?**

El filtrado espacial consiste en filtrar en el plano de la imagen, modificando así los valores de sus píxeles, mientras que el filtrado frecuencial consiste en hacer dicha modificación pero filtrando a través de un plano proyectado de la imagen en la que la característica frecuencial esté presente

**Teorema de Convolución. ¿Cuál es su aplicación práctica?**

Análisis de señales de audio, eléctricas y eliminar ruido en estas

La convolución bidimensional se usa en procesamiento de imágenes para aplicar filtros y efectos visuales. En este caso la función de entrada es una matriz de números, cada uno representando el brillo de un píxel. La función kernel es otra matriz diseñada para lograr algún efecto, y la salida es la imagen filtrada, donde cada píxel es la convolución del kernel con la imagen, calculada por filas y columnas.

**Describe los principales pasos del algoritmo de detección de fronteras de Canny.**

Alisamiento de la imagen con una función Gaussiana circular 2D.

Cálculo del gradiente de la imagen resultante en el paso anterior.

Usar la magnitud del gradiente y su dirección para estimar la fuerza del salto y dirección en cada punto.

Canny, maximos locales son fronteras, intermedios dudosos, si al lado tiene frontera y gradiente parecido se marca

**¿En qué consiste el proceso de cálculo de un operador de gradiente sobre una imagen? ¿Para qué se utiliza?**

La obtención del gradiente de una imagen requiere calcular las derivadas parciales  $\partial f / \partial x$  y  $\partial f / \partial y$  en cada pixel de la imagen.

Píxeles de frontera conjuntos conectados

Componente x e y, diferencia de dos píxeles para la derivada

Valor negativo grande en el centro y alrededor pequeños que lo igualen (sea 0)

Vector gradiente es perpendicular

Tener en cuenta un entorno mayor y eliminar altas frecuencias

## T3: Segmentación

**¿Que entiende por segmentar una imagen? ¿Cuales son las entradas y salidas de este proceso?**

La segmentación subdivide la imagen en una serie de regiones, que comparten alguna característica común

La entrada la imagen y la salida las regiones

**Explique los dos paradigmas básicos de segmentación de una imagen.**

Uno de los paradigmas de segmentación es la discontinuidad del nivel de gris. Consisten en segmentar la imagen a partir de los cambios grandes en los niveles de gris entre los píxeles. Las técnicas que utilizan las discontinuidades como base son la detección de líneas, fronteras, de bordes, de puntos aislados, etc.

El otro es la similitud de niveles de gris. Es lo contrario al método anterior, las divisiones de la imagen se hacen agrupando los píxeles que tienen unas características similares en regiones. Algunas técnicas que usan esto son la umbralización, el crecimiento de regiones entre otras

**Explique brevemente dos métodos para unir trozos de frontera, e indique las diferencias entre ellos.**

### **Procesamiento local**

Este método utiliza información del vecindario más próximo a cada punto etiquetado como frontera en el detector. Todos los puntos considerados similares según un criterio previo son enlazados. Normalmente el criterio se construye a partir de la magnitud y la orientación del vector gradiente.

### **Enlazado de fronteras**

Este método crea una imagen binaria a partir de un umbral T, un ángulo A y un umbral de ángulo Ta. Se recorre la imagen por filas y se rellenan los huecos

más cortos que el umbral. Para otras orientaciones, se gira la imagen y se repite el proceso. La diferencia es que el entrelazo de fronteras es más eficiente

**Cuando una imagen tiene mucho ruido, ¿conoce algún filtro que permita mejorar su segmentación mediante umbralización? ¿en que se basa?**

Para reducir el ruido se utilizan filtros de alisamiento no lineal o lineal, aunque los no lineales funcionan mejor. Se puede utilizar por ejemplo un filtro de paso bajo o un filtro espacial con una gaussiana o en general cualquier función de densidad. Estos filtros se basan en la idea de reducir valores dispares entre pixeles vecinos.

**¿Desde el punto de vista de la segmentación de imágenes, ¿cuál es la principal utilidad de la transformada de Hough?**

A veces tenemos una imagen de fronteras sin informacion acerca de d'onde pueden estar las regiones de interes.

Los pixeles frontera solo pueden ser enlazados a partir de propiedades globales.

Usaremos una aproximación que permite enlazar pixeles pertenecientes a curvas con una forma determinada.

Ejemplo:

Dados n puntos de una imagen, determinar conjuntos de los mismos que pertenezcan a lineas rectas.

Una aproximación puede consistir en tomar pares de puntos, calcular la recta que los une, y buscar todos los demás puntos cercanos a esa recta.

Esta aproximacion es computacionalmente muy costosa.

Hough (1962) propuso una aproximacion alternativa, mas eficiente

**Explique la diferencia entre un sistema de umbralización global frente a otro local. ¿Qué ventajas presenta uno con respecto al otro?**

Segmentación por Umbralización

En función de un umbral se asignan los píxeles de una a imagen a una clase u otra

Problema: Al aumentar el ruido, las modas del histograma se mezclan

Umbralización Global (Iterativo)

Seleccionar umbral inicial T

Segmentar la imagen con ese umbral.  $G1 < T$  y  $G2$  el resto

Calcular la media de  $G1$  y  $G2$  para actualizar el umbral:  $T = (m1+m2)/2$

Repetir pasos 2-3 hasta que se estabilice

Funciona siempre que el umbral inicial esté dentro de las modas

Umbralización Otsu (Directo)

Se basa en minimizar el error medio que se comete al segmentar la imagen · Partiendo de una imagen generada por dos distribuciones gaussianas (Tonos de gris) · Busca el umbral donde, al superponer las distribuciones, la suma de los errores de cada una es mínima

Permite evaluar la separabilidad de las modas del histograma

Problemas Umbralización Otsu

Imagen con ruido -> Histograma más homogéneo Solución: Suavizar (filtro paso bajo)

Objetos pequeños -> Descompensan el histograma Solución: Crear máscara binaria a partir de la magnitud de gradiente de la imagen Así solo se cogen los píxeles a la frontera (donde la magnitud del gradiente aumenta)

Umbralización Variable

Útil si la iluminación es variable

Particionando en losetas la imagen

Calcular umbral de cada punto en base a su entorno

**Explique las diferencias entre el algoritmo de segmentación basado en crecimiento de regiones y el de inundación.**

### **Crecimiento de Regiones**

Se asigna una etiqueta a un pixel si cumple determinada condición y tiene algún vecino que también la cumpla o ya forme parte de esa región:

Comienza con varias semillas (puntos iniciales de cada región)

Se van añadiendo a cada región los vecinos que cumplan las condiciones

Se trata de poner las semillas donde sepamos que hay una región de interés o tomando los centroides de cada posible región

### **Algoritmo**

Partiendo del array de semillas  $S(x,y)$ , se erosiona hasta que cada componente conexa sea de 1 pixel

Formar una imagen  $F$  que valga 1 en los puntos que cumplan el predicado

Formar una imagen  $G$  añadiendo cada punto que esté 8-conectado con algún punto semilla o que haya sido marcado y cumpla el predicado

Etiquetar cada componente conexa en  $G$

### **Inundación**

Genera resultados estables y fronteras cerradas

Proporciona un marco para añadir restricciones basadas en conocimiento previo de la imagen

Fundamentos

Tratar la imagen como una superficie 3D, donde la altura de un pixel es la magnitud de su gradiente. Así, los puntos más altos serán posibles fronteras

Cada punto de la imagen pertenece a una clase:

### **Mínimo Local**

Punto donde al dejar caer una gota -> 1 Mínimo Local concreto (Cuenca)

Punto donde al dejar caer una gota -> +1 Mínimo Local (Línea Divisoria)

Algoritmo

La superficie 3D se inunda de manera uniforme

Cuando se van a unir dos cuencas, se coloca un muro (presa) para evitarlo

Al inundar toda la imagen, solo se ven las presas, correspondientes con las Fronteras

Las "islas" intermedias no corresponden a una frontera entre dos regiones, se descartan

**Explique brevemente el esquema general de reconocimiento de patrones en una imagen.**

**¿Cual cree que es su principal diferencia de los algoritmos de segmentación tradicionales?**

1. Extraer las características de la imagen

2. Representar en el espacio de clasificación
3. Clasificar la imagen en función de su posición.

La diferencia está en qué usa propiedades globales de la imagen (?)

**Explique en que consiste el reconocimiento de objetos basado en apariencia.**

Usan una representación completa del objeto (modelo). Este modelo contiene parámetros que generan las posibles apariencias del objeto. Se intenta alinear el modelo con la imagen (matching).

Hay que tener en cuenta cambios de perspectiva, color, tamaño...

**Explique en que consiste el reconocimiento de objetos basado en características.**

En base a Ejemplos se determinan las Características del objeto en diversas condiciones

Las características se buscan en puntos interesantes: Esquinas, Blobs...

Se construyen Descriptores Locales en base a esas características

Se considera que el objeto está si hay suficientes características en posiciones coherentes

**Explique brevemente que es un descriptor SIFT y que es lo que caracteriza.**

Detector de Blobs (pegotes de píxeles)

Basado en Histogramas Locales de Orientación de Gradiente: Suma los gradientes en las distintas direcciones y los agrupa

Útil para Alinear imágenes (Panorámicas) y encontrar Puntos de Referencia (Landmarks)

Invariante a Traslación - Escalado – Rotación

El detector SIFT proporciona un algoritmo para detectar Blobs de diferentes escalas en una imagen a través de histogramas locales de orientación gradiente. No solo permite detectar objetos, sino que también puede hacer registros y detectar “landmarks” o puntos de reconocimiento, haciéndolo útil para detectar caras. Es invariable a las operaciones de escala, rotación y traslación, parcialmente invariable a cambios de luminosidad y robusto

## T4: Análisis del Movimiento

**¿Que es una secuencia de imágenes entrelazada? ¿Cual es su utilidad?**

En una secuencia entrelazada, la matriz de puntos que compone la imagen se divide en dos campos: uno contiene las líneas pares, y otro las impares.

El muestreo de cada campo se hace en diferentes instantes de tiempo.

Su utilidad es que permite mantener una tasa de refresco alta con la mitad del ancho de banda

**Explique la diferencia entre desentrelazado espacial y temporal. ¿En que situaciones funciona mejor cada uno?**

Desentrelazado espacial (“Bob”): Explora la correlación espacial entre muestras vecinas verticales en un campo. Se calcula promediando las líneas superior e inferior.

Desentrelazado temporal (“Weave”): Explora la correlación temporal entre muestras. Se calcula insertando (o repitiendo) el campo anterior en el cuadro actual.

El desentrelazado espacial funciona mejor con secuencias dinámicas. Por otro lado, el desentrelazado temporal funciona mejor con secuencias estáticas.

**Explique el problema de la apertura en el cálculo del flujo óptico.**

Calcular el Flujo

Necesario añadir Restricciones Adicionales al no disponer de suficiente información: \* Problema de la Apertura = 1 Ecc. con 2 incógnitas Ya que la Ecc. de Constancia de Brillo solo sirve con diferencias pequeñas, partimos de una aprox. del flujo (u,v):

**¿Cuándo son aplicables y qué ventaja tienen los métodos de cálculo de flujo paramétrico sobre los métodos locales?**

**Explique la diferencia entre reconocimiento (matching) y seguimiento (tracking).**

**Explique cada una de las tres fases del proceso de fusión temporal para seguimiento.**

**¿Existe alguna situación en la que el filtro de Kalman sea óptimo? ¿Esta condición se cumple habitualmente cuando se hace seguimiento? Justifique sus respuestas**

**¿Que es el muestreo factorizado? ¿Cuales son las fases que lo componen? ¿Que distribuciones de probabilidad permite representar?**

Es un algoritmo de filtrado de partículas que implementa seguimiento (tracking). Sirve para aproximar la distribución a posteriori.

Fase 1: Se genera un conjunto a partir de  $P(X)$ ,  $\{x_1, \dots, x_n\}$

Fase 2: Se calcula un peso para cada elemento del conjunto

**¿Que método emplea el algoritmo CONDENSATION para representar las distribuciones de probabilidad involucradas en el proceso de seguimiento? ¿Que ventaja tiene esta representación con respecto a otras?**

Primero, se toman muestras en  $C_{t-1}$ , luego se seleccionan algunas muestras, se emplea la dinámica para generar las muestras nuevas, y finalmente se representan mediante  $C_t$ .

Hacerlo así nos faculta observar el movimiento del objeto, trazar el flujo óptico...

**Explique brevemente como funciona en general el algoritmo Mean Shift, y como puede aplicarse a tareas de seguimiento**

Obteniendo el histograma del objeto, podemos realizar un modelo probabilístico para calcular si un píxel pertenece o no a él. El algoritmo crea un gradiente local de la distribución de probabilidad. Para esta área, el algoritmo comprueba la diferencia con el histograma del objeto y del modelo y realiza un movimiento en la dirección que mejore este ajuste. La distribución de probabilidad genera puntos aleatorios donde es más probable que se encuentre el objeto. Estos puntos tienen diferente peso, y se agrupan la mayoría en el centro y van perdiendo densidad conforme se sale de este. Si el objeto no se encuentra en ninguno de estos puntos, se agranda esta “nube” de puntos



hasta encontrar el objeto. Si el objeto se encuentra en la “nube” de puntos se puede hacer más pequeña.

## OTROS AÑOS

### **19 - Explica y compara dos técnicas de realce espacial que conozcas. ventajas e inconvenientes. (falta por completar)**

- Filtro de alisamiento lineal: se calcula la media de los píxeles contenidos en el entorno de vecindad de la máscara del filtro.
- Filtro de alisamiento no lineal: como por ejemplo el filtrado de mediana. Al no hacer ninguna media como en el caso anterior (hacemos el mismo proceso que antes, pero con la mediana), en este tipo de realce se mantienen las fronteras intactas.

### **20 - ¿En qué consiste el procedimiento de igualación de un histograma? ¿cómo se implementa? (falta por completar)**

Es una transformación que pretende obtener para una imagen un histograma con una distribución uniforme (los píxeles de la imagen se reparten de igual forma en los distintos niveles del rango dinámico de la escena).

### **21 - Define el concepto de convolución espacial entre dos señales 2D. ¿Cómo se implementa? ¿Para qué sirve?**

La convolución es el proceso de mover una máscara sobre la imagen e ir calculando la suma de productos en cada posición, rotando previamente el núcleo de convolución inicialmente 180°.

Se implementa multiplicando los valores de vecindad del píxel por los de la máscara de convolución.

Sirve para alisar (emborronar y eliminar ruido).

### **22 - ¿Qué es un filtro de paso bajo? ¿Qué efectos tiene su aplicación?**

Un filtro de paso bajo es un filtro que deja pasar las frecuencias bajas.

Al utilizarlo podemos eliminar las frecuencias altas, eliminando ruido y produciendo un efecto suavizado de la imagen, aunque también podemos destrozar las fronteras al aplicarlo.

### **23 - Cuando una imagen tiene mucho ruido, ¿conoces algún filtro que permita mejorar su segmentación mediante umbralización? ¿En qué se basa?**

Cuando aumenta el nivel del ruido, las modas del histograma se mezclan y puede resultar muy difícil separarlas. Se puede utilizar el suavizado para evitar este problema. El suavizado elimina mínimos locales poco significativos.

### **24 - Explica el proceso de clasificación bayesiana y la hipótesis MAP**

El proceso de clasificación bayesiano permite incorporar conocimiento previo al proceso de clasificación. Además este proceso se basa en la inferencia probabilística y utiliza la hipótesis MAP, que consiste en quedarse con la hipótesis más probable (máximo a posteriori).

### **25 - ¿Qué ocurre si entrenamos una red neuronal demasiado compleja para clasificar un conjunto de muestras? ¿Y si es demasiado simple?**

Cuanto más complejo es un sistema de aprendizaje, más probable es que sobreaprenda

(memorize cada uno de los ejemplos de entrenamiento de forma individual), y que sea incapaz de generalizar.

Un sistema demasiado simple no puede representar distribuciones complejas de los datos, pero generaliza mucho más.

Hay que encontrar un equilibrio entre la complejidad del modelo y su capacidad para reconocer los datos.

**26 - ¿Qué papel desempeñan los pesos en un conjunto de muestras dentro de un algoritmo de filtrado de partículas?**

En las zonas donde los pesos son mayores, los valores de densidad también son más altos.

Preguntas del examen del 4 de junio de 2021

**27 - Teorema de Convolución. ¿Cuál es su principal aplicación práctica?**

Al trabajar con la transformada de Fourier de una convolución, esta puede ser igual al producto punto a punto de las transformadas, es decir, que la convolución en un dominio sería igual al producto punto a punto en el otro dominio. (?)...

**28 - Explica al menos dos formas de estimar el gradiente de una imagen, y comenta las ventajas e inconvenientes de cada una.**

**30 - ¿Frente a qué transformaciones resulta invariante el descriptor SIFT?**

**31 - ¿Qué es una pirámide Laplaciana? ¿Cómo se calcula?**

**32 - ¿Qué dos fuentes de información se integran habitualmente en un proceso de seguimiento? ¿Cómo se hace esta integración?**

**33 - ¿Qué tipo de tratamiento daría a la siguiente imagen para mejorar su aspecto?**

**34 - Explica qué es un filtro de mediana, qué propiedades tiene y en qué situaciones su uso resulta apropiado.**

Es un filtro usado para emborronar fronteras, cuyo funcionamiento se basa en reemplazar el valor de la imagen en un punto por la mediana de los niveles de intensidad de sus vecinos.

**35 - ¿Por qué crees que el modelo deep face para identificar caras obtiene unas tasas de acierto tan superiores a los métodos que le precedieron?**

Deep Face se basa en redes neuronales y éstas requieren de entrenamiento por su naturaleza basada en el sistema nervioso. Cuanto más se entrena, mejor funciona, y ya ha sido entrenada con millones de caras.

Además, Deep Face se ayuda de modelos 3D de las caras y de técnicas de alineado de puntos en las caras.

**36 - ¿Cuál es la estructura y componentes de una neurona artificial? ¿Por qué es tan importante en el ámbito de la visión artificial y la Inteligencia Artificial?**

Los componentes principales de una neurona artificial son:

- Un conjunto de entradas  $x_i$ , con pesos  $w_i$  asociados
- Una función de activación  $f()$
- Una salida:  $f(\sum x_i * w_i + b)$

En el ámbito del procesamiento de información visual, el deep learning (basado en redes neuronales) es útil en temas como el Deep Face (reconocimiento de caras). En general, las redes neuronales son claves en el desarrollo de la IA ya que se pueden resolver problemas

complejos con ellas de forma más poderosa que con otro tipo de algoritmos (que alguien concrete esto más, por favor ). 😊

**37 - Explica el concepto de coherencia temporal de una secuencia de vídeo, y su utilidad de cara al análisis de información visual.**

La coherencia temporal ocurre cuando entre un cuadro y otro no hay muchas diferencias en cuanto a posicionamiento de los objetos representados, propiedades lumínicas, etc.

Esto es especialmente interesante si tenemos localizado un objeto en un cuadro y queremos buscarlo en el siguiente, ya que, al estar cerca, nuestro ámbito de búsqueda (y por tanto el coste computacional) se reduce significativamente, y además nos permite no confundir el objeto con otros parecidos de la imagen.

**38 - La ecuación de constancia de brillo no funciona para valores de flujo grandes. ¿Cómo solucionarías este problema?**

Se emplea una pirámide gaussiana de imágenes, y se calcula el flujo para cada imagen de cada nivel de tal pirámide, y se usa como punto de partida para la siguiente imagen. (Intentar poner ejemplo)