

Quien quiera puede escribir las preguntas que falten, añadir contenido a otras o indicar por qué están mal.

1 - Explique en qué consiste el 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.

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

2 - Explique brevemente cómo funciona en general el algoritmo Mean Shift, y cómo 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.

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

El enlazado de fronteras marca como frontera píxeles que se encuentran al lado de una siempre y cuando satisfaga los criterios de magnitud y ángulo.

El procesamiento a nivel de región, a veces se sabe a qué región pertenecen determinados puntos frontera y pueden enlazarse mediante una aproximación poligonal.

DIFERENCIA(?), son distintos totalmente xd

OTRA SOLUCIÓN

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 T_a . 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.

4 - Defina el concepto de convolución espacial de 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).

5 - Explique qué 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

6 - 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 no es más que una función no lineal que depende de un vector de 11 parámetros, los parámetros del modelo. Estos pueden ser extrínsecos o intrínsecos.

Parámetros intrínsecos: Son aquellos que caracterizan las propiedades inherentes de la cámara y de la óptica, es decir, los parámetros involucrados en la transformación de puntos 3D en el sistema de referencia de la cámara a puntos 2D del plano imagen. Se suelen considerar 5 parámetros.

Es decir, son aquellos que describen el funcionamiento de una cámara. Por ejemplo, la distancia focal, el punto principal y el centro óptico..

Parámetros extrínsecos: Son los seis parámetros que definen la posición y orientación de la cámara con respecto al sistema de referencia absoluto.

Es decir, son aquellos que definen la posición y la orientación del cuadro de referencia de la cámara con respecto al mundo real, es decir, dan la orientación externa de la cámara.

7 - Explique la diferencia entre desentrelazado espacial y temporal. ¿En qué 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.

8 - Explique brevemente el esquema general de reconocimiento de patrones en una imagen. ¿Cuál cree que es su principal diferencia con 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 que usa propiedades globales de la imagen (?)

9 - Explique que es el ruido de una imagen. ¿Cómo 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).

10 - Explique las diferencias entre los procesos de filtrado espacial y frecuencial.

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.

11 - Describa 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.

12 - Describa brevemente qué es un detector SIFT y qué es lo que le caracteriza.

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

frente a las distorsiones geométricas locales. Propiedades similares a las de las neuronas de la corteza temporal inferior de los primates.

13 - Explique en qué consiste el reconocimiento de objetos basado en apariencias.

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...

14 - ¿Qué es una secuencia de imágenes entrelazada? ¿Cuál 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.

15 - ¿Qué es el muestreo factorizado? ¿Cuáles son las fases que lo componen?

¿Qué 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

16 - 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.

17 - ¿Qué método emplea el algoritmo CONDENSATION para representar las distribuciones de probabilidad involucradas en el proceso de seguimiento? ¿Qué 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 facilita observar el movimiento del objeto, trazar el flujo óptico...

18 - Cuando una imagen tiene mucho ruido, ¿conoce algún filtro que permita mejorar su segmentación mediante umbralización? ¿en qué 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 píxeles vecinos.

Preguntas del examen de incidencias del 1 de junio de 2021

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 (memorice 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?



Preguntas del examen de incidencias del 1 de julio de 2021

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)