

Visión por Computador

Cuestionario – Teoría 2

Curso 2017/2018

Francisco Javier Caracuel Beltrán

`caracuel@correo.ugr.es`

Índice

1. Identificar la/s diferencia/s esencial/es entre el plano afín y el plano proyectivo. ¿Cuáles son sus consecuencias? Justificar la contestación.	3
2. En coordenadas homogéneas los puntos y rectas del plano se representan por vectores de tres coordenadas (notados x y 1 respectivamente), de manera que si una recta contiene a un punto se verifica la ecuación $x^T l = 0$. Verificar que en coordenadas homogéneas.....	4
3. Defina una homografía que haga que el punto $(2, 0, 3)$ del plano proyectivo finito en el plano transformado esté en el infinito. Si recta del infinito l se define como aquella que contiene a todos los puntos del infinito y por tanto debe verificarse que $x^T l = 0$ para todos los puntos del infinito. ¿Cuál son las coordenadas del vector l ? Justificar la respuesta.	4
4. Descomponer en composición de movimientos elementales (traslación, giro, escala, cizalla, proyectivo) cada una de las matrices de las siguientes homografías H_1 , H_2 y H_3 :.....	4
5. ¿Cuáles son las propiedades necesarias y suficientes para que una matriz defina una homografía entre planos? Justificar la respuesta.	5
6. ¿Qué propiedades de la geometría de un plano quedan invariantes si se aplica una homografía general sobre él? Justificar la respuesta.	5
7. ¿Cuál es la deformación geométrica más fuerte que se puede producir sobre la imagen de un plan por el punto de vista de la cámara? Justificar la respuesta.....	6
8. ¿Qué información de la imagen usa el detector de Harris para seleccionar puntos? ¿El detector de Harris detecta patrones geométricos o fotométricos? Justificar la contestación. .	6
9. ¿Sería adecuado usar como descriptor de un punto Harris los valores de los píxeles de su región de soporte? En caso positivo identificar cuando y justificar la respuesta.	6
10. ¿Qué información de la imagen se codifica en el descriptor de SIFT? Justificar la contestación.	7
11. Describa un par de criterios que sirvan para establecer correspondencias (matching) entre descriptores de regiones extraídos de dos imágenes. Justificar.....	7
12. ¿Cuál es el objetivo principal en el uso de la técnica RANSAC? Justificar	8
13. Si tengo 4 imágenes de una escena de manera que se solapan la 1-2, 2-3 y 3-4. ¿Cuál es el número mínimo de puntos en correspondencias necesarios para montar un mosaico? Justificar la respuesta.....	8
14. En la confección de un mosaico con proyección rectangular es esperable que aparezcan deformaciones de la realidad. ¿Cuáles y por qué? ¿Bajo qué condiciones esas deformaciones podrían desaparecer? Justificar la respuesta.	9
15. Referencias:	9

1. Identificar la/s diferencia/s esencial/es entre el plano afín y el plano proyectivo. ¿Cuáles son sus consecuencias? Justificar la contestación.

Un plano afín es un sistema de puntos y líneas que tienen las siguientes propiedades:

- Dos puntos distintos se encuentran en una única línea.
- Cada línea tiene al menos dos puntos.
- Dada cualquier línea y un punto no perteneciente a la línea, existe una única línea que contiene al punto no se corta con la primera línea.

Estas propiedades indican que en un plano afín se pueden tener líneas delimitadas por dos puntos cada una, el primero de ambas líneas situado en cualquier posición y el segundo situado en el infinito y dichas líneas pueden no cortarse nunca.

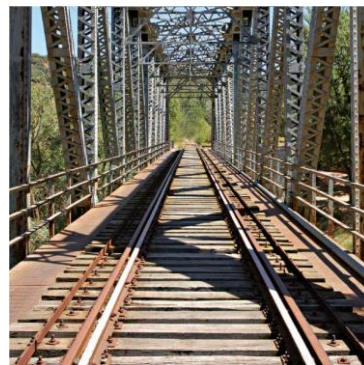
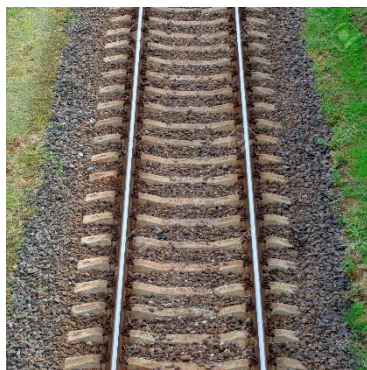
Un plano proyectivo es un sistema de puntos y líneas que tienen las siguientes propiedades:

- Dos puntos distintos se encuentran en una única línea.
- La intersección de dos rectas distintas cualesquiera contienen un único punto.
- Existe un conjunto de cuatro puntos, sin haber tres puntos que pertenezcan a la misma recta.

Estas propiedades indican que en un plano proyectivo no existen líneas paralelas, sino que siempre existe un punto situado en el infinito en el que se cortan dos líneas.

La diferencia fundamental entre el plano afín y el plano proyectivo es la existencia de líneas paralelas en el primero y la ausencia de éstas en el segundo.

Las consecuencias son que, dependiendo del plano en el que se encuentre la imagen, los puntos que tienden al infinito se distorsionan. Como se vio en clase, dependiendo del plano en el que se visualicen las vías del tren, pueden no cortarse nunca o hacerlo en el infinito.



2. En coordenadas homogéneas los puntos y rectas del plano se representan por vectores de tres coordenadas (notados x y l respectivamente), de manera que si una recta contiene a un punto se verifica la ecuación $x^T l = 0$. Verificar que en coordenadas homogéneas el vector de la recta definida por dos puntos afines puede calcularse como el producto vectorial de los vectores de ambos puntos ($l = x \times x'$). De igual modo el punto intersección de dos rectas l y l' está dado por $x = l \times l'$. ¿Qué conclusiones extrae de las anteriores propiedades de cara a construir un algoritmo que calcule la intersección de dos rectas del plano? Justificar la contestación.

3. Defina una homografía que haga que el punto $(2, 0, 3)$ del plano proyectivo finito en el plano transformado esté en el infinito. Si recta del infinito l se define como aquella que contiene a todos los puntos del infinito y por tanto debe verificarse que $x^T l = 0$ para todos los puntos del infinito. ¿Cuál son las coordenadas del vector l ? Justificar la respuesta.

4. Descomponer en composición de movimientos elementales (traslación, giro, escala, cizalla, proyectivo) cada una de las matrices de las siguientes homografías $H1$, $H2$ y $H3$:

$$H1 = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.2 & 0 & 0 \\ 0 & 0.4 & 0 \\ 0 & 0 & 0.2 \end{pmatrix} \begin{pmatrix} 1.5 & 0 & 3 \\ 0 & 0.8 & 5 \\ 0 & 0 & 2 \end{pmatrix}$$

$$H2 = \begin{pmatrix} 0 & 0.5 & -3 \\ -0.5 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$






$$H3 = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0.5 & 0 \\ 0.5 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

5. ¿Cuáles son las propiedades necesarias y suficientes para que una matriz defina una homografía entre planos? Justificar la respuesta.

Una homografía es un isomorfismo de espacios proyectivos inducidos por un isomorfismo de los espacios vectoriales, desde los cuales el espacio proyectivo deriva. Esto quiere decir que, en toda homografía, existen correspondencias entre los puntos y rectas de ambos espacios vectoriales. Para que una matriz defina una homografía entre planos, debe cumplir unas propiedades:

- La matriz debe ser 3×3 .
- La matriz debe ser invertible, es decir, su determinante debe ser distinto de 0.

6. ¿Qué propiedades de la geometría de un plano quedan invariantes si se aplica una homografía general sobre él? Justificar la respuesta.

Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} I & t \end{bmatrix}_{2 \times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} R & t \end{bmatrix}_{2 \times 3}$	3	lengths	
similarity	$\begin{bmatrix} sR & t \end{bmatrix}_{2 \times 3}$	4	angles	
affine	$\begin{bmatrix} A \end{bmatrix}_{2 \times 3}$	6	parallelism	
projective	$\begin{bmatrix} H \end{bmatrix}_{3 \times 3}$	8	straight lines	

Teniendo en cuenta la tabla 2.1 (pag.38) del libro de Richard Szeliski, dependiendo del tipo de transformación se tendría una homografía distinta. Las propiedades que quedan invariantes dependen de la transformación aplicada. De este modo, las propiedades que se preservan son las que se encuentran en la misma fila que la transformación que se aplica y todas las que aparecen en las filas inferiores a dicha transformación.

Por ejemplo, si se aplica una homografía que realiza una traslación sobre la imagen, las propiedades que se preservan son la orientación, la longitud, ángulos, paralelismo y puntos sobre una recta.

7. ¿Cuál es la deformación geométrica más fuerte que se puede producir sobre la imagen de un plan por el punto de vista de la cámara? Justificar la respuesta.

Utilizando la tabla del ejercicio anterior (6), se puede apreciar como la transformación más fuerte que se puede realizar es la proyección, ya que es la que menos propiedades preserva en la imagen.

8. ¿Qué información de la imagen usa el detector de Harris para seleccionar puntos? ¿El detector de Harris detecta patrones geométricos o fotométricos? Justificar la contestación.

La información que utiliza el detector de Harris para seleccionar puntos es el cambio de intensidad en varias direcciones existente entre ellos.

Puede detectar patrones geométricos, ya que es capaz de encontrar las esquinas de objetos geométricos y patrones fotométricos, ya que propiamente utiliza los cambios de la intensidad, aunque no sean zonas con objetos geométricos.

9. ¿Sería adecuado usar como descriptor de un punto Harris los valores de los píxeles de su región de soporte? En caso positivo identificar cuando y justificar la respuesta.

Para usar como descriptor de un punto Harris los valores de los píxeles de su región de soporte se tiene que comprobar que la relación entre los píxeles de esa región no varía al aplicar las transformaciones.

Si se aplica una traslación, los puntos al final mantienen la misma estructura que antes de realizar la transformación. Aplicando una rotación, escalado, cizalla o proyección, los puntos ya no mantienen la misma relación entre sí, por lo que ya no es adecuado usarlos como descriptor de un punto.

**10. ¿Qué información de la imagen se codifica en el descriptor de SIFT?
Justificar la contestación.**

Para calcular el descriptor SIFT se detectan los keypoints con diferentes escalas, se localizan los keypoints, se calcula la orientación y se obtiene el descriptor de cada punto.

Se toma una matriz de 16x16 alrededor del punto, dividida en 16 sub-bloques de tamaño 4x4. Para cada sub-bloque se crea un histograma con 8 direcciones del gradiente. Finalmente, se puede decir que la información que se codifica en el descriptor de SIFT son los gradientes que se encuentran alrededor de un punto.

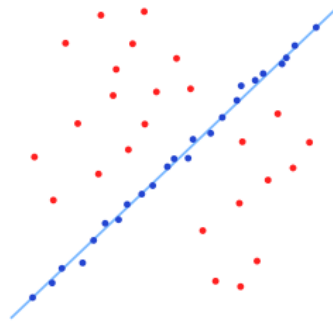
**11. Describa un par de criterios que sirvan para establecer correspondencias (matching) entre descriptores de regiones extraídos de dos imágenes.
Justificar la idoneidad de los mismos.**

Se pueden describir los métodos utilizados en el Trabajo 2:

- BruteForce+crossCheck: este método consiste en comprobar punto a punto los descriptores de la primera imagen, compararlos con los de la segunda y escoger la mejor relación entre ambos puntos. Se repite el proceso al contrario y finalmente solo se eligen los puntos que coincidan en ambas operaciones.
- K vecinos más cercanos: se elige cada par de puntos de las imágenes y se calcula la distancia que existe entre sus descriptores. A continuación, se comprueba los k vecinos más cercanos y si éstos coinciden, se marcan como válidos.

12. ¿Cuál es el objetivo principal en el uso de la técnica RANSAC? Justificar la respuesta.

El objetivo principal de la técnica RANSAC es encontrar una homografía entre dos imágenes de manera que el error que se produce entre los puntos de las dos imágenes sea mínimo.



La técnica consiste en realizar un proceso parecido a la regresión lineal, de manera que se trazan líneas entre los puntos de ambas imágenes y se elige la que finalmente mejor resultado ofrece de entre todas.

13. Si tengo 4 imágenes de una escena de manera que se solapan la 1-2, 2-3 y 3-4. ¿Cuál es el número mínimo de puntos en correspondencias necesarios para montar un mosaico? Justificar la respuesta.

Como se ha visto en clase y como se ha realizado en el Trabajo 2, los puntos mínimos necesarios para realizar la homografía entre dos imágenes son 4. Si se quiere realizar un mosaico entre las 4 imágenes, solapándose la 1-2, 2-3 y 3-4, es necesario realizar 3 homografías. Teniendo en cuenta los 4 puntos mínimos por homografía y las 3 homografías, se puede decir que el número mínimo de puntos en correspondencias necesarios son 12.

No es necesario encontrar ningún punto en correspondencia para llevar las imágenes a su lugar correspondiente del mosaico, ya que solo es necesaria una homografía que se encargue de realizar la traslación.

14. En la confección de un mosaico con proyección rectangular es esperable que aparezcan deformaciones de la realidad. ¿Cuáles y por qué? ¿Bajo qué condiciones esas deformaciones podrían desaparecer? Justificar la respuesta.

Cuando se realiza un mosaico se toma la imagen central como la principal y en base a ella se crea, concatenando las respectivas imágenes del mosaico hacia los extremos.

Una deformación que se produce son cambios en la tonalidad de los colores debido al cambio de iluminación que se produce al mover la cámara. Estos cambios bruscos en la tonalidad de una imagen a otra se pueden reducir aplicando una máscara gaussiana en las zonas donde se produce esa diferencia.

Otra deformación se produce por el cambio de perspectiva. Al avanzar en la realización del mosaico, el movimiento es circular hacia los extremos, por lo que, al hacer la homografía de las imágenes, tienden a deformarse para ajustarse entre ellas. La solución a este efecto sería proyectar el mosaico en una esfera.

La última deformación viene dada principalmente por el tiempo que transcurre entre la toma de las distintas imágenes en lugares dinámicos. Siempre que aquello sobre lo que se quiera realizar el mosaico no cambie con el paso del tiempo en el que se toman todas las imágenes, no existe ningún problema. Si durante la toma de cada imagen los elementos que salen en ella cambian, se debe tomar alguna medida para su corrección.

15. Referencias:

- Apuntes de Teoría – Visión por Computador – Curso 2017/2018
- Computer Vision: Algorithms and Applications – Richard Szeliski
- https://es.wikipedia.org/wiki/Plano_af%C3%ADn
- https://es.wikipedia.org/wiki/Geometr%C3%ADa_finita#Proyectivo
- [https://es.wikipedia.org/wiki/Homograf%C3%ADa_\(geometr%C3%ADa\)](https://es.wikipedia.org/wiki/Homograf%C3%ADa_(geometr%C3%ADa))
- https://www.gti.ssr.upm.es/~jir/comp_vis/gp.pdf
- https://es.wikipedia.org/wiki/Scale-invariant_feature_transform
- http://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_feature2d/py_sift_intro/py_sift_intro.html
- <https://es.wikipedia.org/wiki/RANSAC>