Cuestionario 2 - VC

Óscar David López Arcos.[redacted]

---

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

Dos diferencias esenciales:

- En el plano proyectivo el paralelismo no se conserva al realizar una transformación.
- El plano afín no incluye la recta proyectiva de puntos de infinito.

La consecuencia más relevante que provocan estas diferencias es la posibilidad que ofrecen los planos proyectivos, permitiendo manejar, trabajar y representar puntos que antes no estaban en la imagen (ya que se encontraban en el infinito).

2. Demuestre que los puntos de la recta del infinito del plano proyectivo son vectores del tipo (*,*,0) con *=cualquier número.

Por definición, cualquier coordenada del plano proyectivo (x, y, z) equivale a una coordenada bidimensional de la forma $(x/z, y/z)$.

Los puntos de la recta del infinito del plano proyectivo son aquellos que en el plano bidimensional se encuentran en el infinito. Un punto se encuentra en el infinito si y solo si el valor de z es igual a 0. Los valores de x e y podrían ser cualquier número, verificando así la demostración.

3. 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 $x^T l=0$. Puede 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 × x'). De igual modo el punto intersección de dos rectas l y l' está dado por x = l × l' ¿Considera de interés las anteriores propiedades de cara a construir un algoritmo que calcule la intersección de dos rectas cualesquiera del plano? Justificar la contestación.

Un punto $X = (x, y)$ pertenece a la recta l definida por la ecuación $ax + by + c$ si se verifica $ax + by + c = 0$. Si expresamos este punto como coordenadas homogéneas $(x, y, 1)$, la ecuación de la recta puede representarse de la forma $(a, b, c)\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = 0$.

Que el producto de dos vectores sea igual a 0 significa que su posición en el espacio es perpendicular $\left(\overrightarrow{a} \cdot \overrightarrow{c} = |a| \cdot |c| \cdot \cos(90)\right)$. De esta forma, acabamos de demostrar que que X pertenezca a l es equivalente a decir que $\overrightarrow{X}$ es perpendicular a $\vec{l}$ en el espacio proyectivo.

Si tanto X como X' pertenecen a l, en el espacio proyectivo l será perpendicular a ambos puntos, verificándose $l = X \times X'$.

Del mismo modo, si X pertenece a dos rectas l y l' simultáneamente, $\overrightarrow{X}$ será perpendicular a ambas, verificándose también $x = l \times l'$.

De cara la construcción de un algoritmo que calcule la intersección de dos rectas en el plano, es preferible computacionalmente calcular el producto escalar de dos rectas que no se cruzan a resolver el sistema.

4. Defina una homografía entre planos proyectivos que haga que el punto (2,0,3) del plano proyectivo-1 se transforme en un punto de la recta del infinito del plano proyectivo-2.? Justificar la respuesta.

Definimos una Homografía 3x3 que transforme el punto de un plano a otro. Para que se transforme en un punto de la recta del infinito, la última coordenada debe ser igual a 0 (como vimos en el ejercicio 2).

A continuación, desarrollamos los cálculos y concluimos que nos vale cualquier homografía, siempre y cuando la multiplicación del punto por ésta dé como resultado 0 en la última coordenada del nuevo punto:

$$(2,0,3) \times \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = (x, y, 0) \Longrightarrow 2c + 3i = 0 \Longrightarrow i = \frac{-2c}{3} \Longrightarrow H = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & -2c/3 \end{pmatrix} \forall [a - h]$$

5. Descomponer en composición de movimientos elementales (traslación, giro, escala, cizalla, proyectivo) cada una de las matrices de las siguientes homografías H1, H2y H3:

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

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

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

Las secciones 2.1.2 y 2.1.3 del libro de Szeliski muestran las distintas transformaciones que provocan las diferentes homografías. Siguiendo lo aquí descrito, tenemos:

H1 = cizalla * escalado * (escalado + traslación)
H2 = (rotación + translación) * (escalado + cizalla)
H3 = proyección * (rotación + escalado)

Como norma básica, podemos fijarnos en la matriz $\begin{pmatrix} s_x & g_x & t_x \\ g_y & s_y & t_y \\ 0 & 0 & 1 \end{pmatrix}$ donde s controla el escalado, g las rotaciones y t los desplazamientos. La última fila (eje z) controla el cambio de proyección

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

La matriz de homografía H debe ser de dimensión 3x3, para definir una relación entre dos planos proyectivos y transformar un punto (x, y, z) en otro (x', y', z').

Además, el $|H| \neq 0$ para que la matriz H que define la transformación tenga inversa y pueda realizarse la transformación inversa, valga la redundancia.

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

Al aplicar una homografía general sobre un plano estamos aplicando simultáneamente tanto una transformación afín como una proyectiva. La única propiedad que se mantiene en el caso de la proyectividad es la colinealidad de los puntos (y las propiedades que dependen de ésta: concurrencia, ratio cruzado, etc.). Al mantenerse también en la transformación afín, será la única propiedad que se cumpla al aplicar una homografía general.

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

La proyección. Tras esta deformación puede ocurrir que, al realizar un cambio de perspectiva, un punto que antes formaba parte de la imagen ahora pase a estar en el infinito. Perder puntos de una imagen sería la deformación más fuerte que podría producir.

9. ¿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.

El detector Harris establece una ventana que desplazará para calcular la variación de los píxeles en los ejes x e y, usando para ello el gradiente de las derivadas y la diferencia de cuadrados. Posteriormente, seleccionará los píxeles con un alto valor de respuesta que superen cierto umbral de calidad (esquinas).

Detecta patrones fotométricos. El detector de Harris no es invariante a la escala, por lo que una misma figura geométrica de distintas dimensiones podría no ser detectada igual, dependiendo del tamaño de ventana que se use.

10. ¿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.

Uno de los principales objetivos de los descriptores es poder utilizarlo posteriormente para establecer correspondencias entre imágenes. Un descriptor basado en los valores de los píxeles de la región de soporte daría problemas a la hora de emparejar imágenes que estén a distinta escala, ya que éste es uno de los principales problemas no resueltos en Harris.

Si tuviésemos la certeza de que todas las imágenes a comparar están en la misma escala, sí podría utilizarse, ya que funciona correctamente ante las rotaciones y cambios de intensidad.

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

Por cada Keypoint detectado, establecemos una ventana de 16x16 en la que calculamos el gradiente de cada píxel, aplicando posteriormente una Gaussiana. Después, generamos un histograma de orientación del gradiente por cada subventana de tamaño 4x4. Finalmente obtenemos 16 celdas con 8 orientaciones cada una, lo que queda codificado como 128 características del punto referentes a su contexto.

12. Describa un par de criterios que sirvan para seleccionar parejas de correspondencias ("matching") entre descriptores de regiones extraídos de dos imágenes. Justificar la idoneidad de los mismos.

- **Lowe-Average-2NN:** Este algoritmo solo calcula las diferencias de los descriptores de una de las imágenes con todos los de la otra. Para elegir las correspondencias, aplica la siguiente regla: Escoger, para cada Keypoint i de la primera imagen, el Keypoint j de la segunda imagen que menor valor de distancia produzca (mínimo $d(i, j)$) siempre y cuando se verifique

$\frac{d(i,j)}{d(i,k)} < 0{,}7$ siendo $d(i,k)$ el segundo valor de d más pequeño para el punto i. Este criterio propone elegir únicamente los puntos cuya correspondencia esté más clara, ya que existe una diferencia notable entre la pareja elegida y el valor d a la siguiente posible.

- **Fuerza bruta + validación cruzada:** Este algoritmo genera la matriz de diferencias de cada descriptor de la imagen 1 con todos los de la imagen 2. Posteriormente realiza el mismo proceso, pero esta vez de la imagen 2 con la imagen 1, generando una segunda matriz. Por último, elegimos las correspondencias cuya pareja mínima (diferencia más baja) coincida en ambas matrices. Este criterio intenta asegurarse por medio de la validación cruzada de que la correspondencia elegida sea correcta (los dos coinciden), sin embargo, si los puntos son muy parecidos podría equivocarse al haber valores de d muy similares.

13. Cual es el objetivo principal en el uso de la técnica RANSAC en el cálculo de una homografía. Justificar la respuesta.

Esta técnica es utilizada con el objetivo de reducir el ruido a la hora de establecer las correspondencias entre dos imágenes definidas por una homografía.

Los algoritmos que seleccionan las parejas de correspondencias no son perfectos y pueden generar parejas erróneas. Por ello, RANSAC tiene en cuenta esta posibilidad y busca la tendencia general que siguen la mayoría de parejas para dar de lado aquellas que considera “outliers”, de forma parecida a como hacían los algoritmos de regresión linear vistos en Machine Learning.

14. ¿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 parejas de puntos en correspondencias necesarios para montar un mosaico? Justificar la respuesta

El mínimo número de correspondencias de puntos para establecer una homografía entre dos imágenes es 4. Si tenemos que establecer 3 homografías distintas (1-2, 2-3, 3-4), necesitaríamos:

$$3 \text{ homografías} \times \frac{4 \text{ correspondencias}}{\text{homografía}} \times \frac{2 \text{ puntos}}{\text{correspondencia}} = 24 \text{ puntos}$$

Por último, para trasladarlo a un mosaico donde quepan todas las imágenes necesitamos definir una homografía de traslación, pero para esta homografía no sería necesario obtener más puntos.

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

El primer error que puede aparecer es el ocasionado por acumulación de pequeños errores a la hora de establecer las correspondencias entre puntos, aunque este es de difícil solución.

Por otro lado, la deformación más común viene ocasionada por el cambio de proyección al rotar la cámara en el eje Y en lugar de desplazarla horizontalmente por X para tomar las distintas imágenes. Al componer el mosaico, las imágenes tienden a deformarse para ajustarse unas a otras. Esta deformación podría desaparecer si en lugar de una proyección rectangular, proyectásemos el resultado sobre una esfera.