

# Cuestionario de teoría

## 2

Alejandro García Montoro<sup>\*</sup>  
agarciamontoro@correo.ugr.es

18 de noviembre de 2015

### 1. Cuestiones

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

**Solución.** La diferencia esencial desde un punto de vista axiomático entre la geometría afín y la proyectiva es que en la primera, dado un punto que no está en una recta, existe una única paralela a la recta que pasa por el punto, mientras que en la segunda, dadas dos rectas cualesquiera, éstas se cortan siempre en un punto.

Evidentemente, de la diferencia axiomática derivan todas las diferencias en las geometrías. Así, una de las más importantes y directamente relacionadas con el axioma citado es que las transformaciones entre planos afines conservan el paralelismo y las transformaciones entre planos proyectivos en general no lo hacen.

Esta *libertad* permite dotar de una estructura matemática a unas transformaciones que no tendrían sin la geometría proyectiva un modelo en el que reflejarse y con el que poder trabajar.

**Ejercicio 2.** Verificar que en coordenadas homogéneas el vector de la recta definida por dos puntos puede calcularse como el producto vectorial de los vectores de los puntos ( $l = x \times x'$ ). De igual modo el punto intersección de dos rectas  $l$  y  $l'$  está dado por  $x = l \times l'$ .

**Solución.** En coordenadas cartesianas, la ecuación genérica de una recta es:

$$ax_1 + bx_2 + c = 0$$

donde  $(-b, a) \in \mathbb{R}^2$  es el vector director de dicha recta.

---

<sup>\*</sup>He recibido ayuda de Antonio Álvarez Caballero.

Por tanto, los puntos  $(x_1, x_2, x_3)$  en coordenadas homogéneas que están en esa recta son los que cumplen la ecuación

$$ax_1 + bx_2 + cx_3 = 0$$

Nuestro problema es ahora, dados dos puntos,  $x = (x_1, x_2, x_3)$  y  $x' = (x'_1, x'_2, x'_3)$ , encontrar la recta que los une; es decir, encontrar los  $a, b, c$  tales que se cumplen estas dos ecuaciones:

$$\begin{cases} ax_1 + bx_2 + cx_3 = 0 \\ ax'_1 + bx'_2 + cx'_3 = 0 \end{cases}$$

Pero resolver ese sistema se reduce a calcular el producto vectorial de los dos puntos:

$$x \times x' = \begin{vmatrix} i & j & k \\ x_1 & x_2 & x_3 \\ x'_1 & x'_2 & x'_3 \end{vmatrix} = (x_2x'_3 - x_3x'_2, x_3x'_1 - x_1x'_3, x_1x'_2 - x_2x'_1)$$

Es inmediato comprobar que el vector  $l = (x_2x'_3 - x_3x'_2, x_3x'_1 - x_1x'_3, x_1x'_2 - x_2x'_1)$  cumple las dos ecuaciones:

$$\begin{cases} (x_2x'_3 - x_3x'_2)x_1 + (x_3x'_1 - x_1x'_3)x_2 + (x_1x'_2 - x_2x'_1)x_3 = 0 \\ (x_2x'_3 - x_3x'_2)x'_1 + (x_3x'_1 - x_1x'_3)x'_2 + (x_1x'_2 - x_2x'_1)x'_3 = 0 \end{cases}$$

y es, por tanto, el vector de la recta que buscamos.

De forma simétrica, el punto intersección de las rectas  $l = (l_1, l_2, l_3)$  y  $l' = (l'_1, l'_2, l'_3)$  es el punto  $x = (x_1, x_2, x_3)$  que cumple a la vez las dos ecuaciones que determinan las rectas:

$$\begin{cases} l_1x_1 + l_2x_2 + l_3x_3 = 0 \\ l'_1x_1 + l'_2x_2 + l'_3x_3 = 0 \end{cases}$$

Es clara la simetría entre este problema y el anterior —situación que no es casual, ya que responde al principio de dualidad del plano proyectivo; es decir: el espacio dual del plano proyectivo es isomorfo al plano proyectivo—, así que es evidente que la solución a este sistema es el producto vectorial de las dos líneas:

$$x = l \times l' = \begin{vmatrix} i & j & k \\ l_1 & l_2 & l_3 \\ l'_1 & l'_2 & l'_3 \end{vmatrix} = (l_2l'_3 - l_3l'_2, l_3l'_1 - l_1l'_3, l_1l'_2 - l_2l'_1)$$

**Ejercicio 3.** Sean  $x$  y  $l$  un punto y una recta respectivamente en un plano proyectivo  $P_1$  y suponemos que la recta  $l$  pasa por el punto  $x$ , es decir  $l^T x = 0$ . Sean  $x'$  y  $l'$  un punto y una recta del plano proyectivo  $P'$  donde al

igual que antes  $l'^T x' = 0$ . Supongamos que existe una homografía de puntos  $H$  entre ambos planos proyectivos, es decir  $x' = Hx$ . Deducir de las ecuaciones anteriores la expresión para la homografía  $G$  que relaciona los vectores de las rectas, es decir  $G$  tal que  $l' = Gl$ . Justificar la respuesta.

**Solución.** Por hipótesis, tenemos que

$$l'^T x' = 0 \quad (1)$$

Por la homografía  $H$ , podemos escribir 1 como

$$l'^T Hx = 0$$

Pero por hipótesis sabemos que  $l'^T x = 0$ , luego podemos igualar y obtener

$$l'^T Hx = l'^T x$$

de donde deducimos directamente que

$$l'^T H = l'^T$$

Ahora podemos deshacernos de la  $H$  multiplicando por su inversa a la derecha —sabemos que existe porque las homografías, por definición, tienen inversa—, y llegamos a una expresión que se parece mucho a lo que queríamos:

$$l'^T = l'^T H^{-1} \quad (2)$$

Ahora sólo falta tomar transpuestas en 2. Teniendo en cuenta que la transpuesta de un producto es el producto de las transpuestas con el orden invertido —es decir,  $(AB)^T = B^T A^T$ —, llegamos a la expresión que buscábamos:

$$l' = (H^{-1})^T l$$

Concluimos que la homografía  $G$  que relaciona los vectores de las rectas  $l$  y  $l'$  es

$$G = (H^{-1})^T$$

**Ejercicio 4.** Suponga la imagen de un plano en donde el vector  $l = (l_1, l_2, l_3)$  representa la proyección de la recta del infinito del plano en la imagen. Sabemos que si conseguimos aplicar a nuestra imagen una homografía  $G$  tal que  $l' = Gl$ , siendo  $l' = (0, 0, 1)$ , entonces habremos rectificado nuestra imagen llevándola de nuevo al plano afín. Suponiendo que la recta definida por  $l$  no pasa por el punto  $(0, 0)$  del plano imagen, encontrar la homografía  $G$ . Justificar la respuesta.

**Solución.** Podemos tantear una solución. Como queremos llevar un vector con tres coordenadas a otro mediante una matriz, vamos a tener que resolver

tres ecuaciones. De todas las matrices posibles, podemos tomar la identidad con la última fila arbitraria e imponer la relación que queremos que satisfaga.

Así, sea  $G$  la matriz

$$\begin{pmatrix} 1 & 0 & a \\ 0 & 1 & b \\ 0 & 0 & c \end{pmatrix}$$

Entonces, imponiendo  $l' = Gl$ , tenemos las siguientes ecuaciones:

$$\begin{cases} 0 &= l_1 + cl_3 \\ 0 &= l_2 + bl_3 \\ 1 &= cl_3 \end{cases}$$

De aquí deducimos que

$$c = \frac{1}{l_3} \quad b = \frac{-l_2}{l_3} \quad a = \frac{-l_1}{l_3}$$

que están bien definidos porque, al no intersectar la línea  $l$  con el origen del plano, es claro que  $l_3 \neq 0$ .

Por tanto, la homografía  $G$  que buscamos es:

$$G = \begin{pmatrix} 1 & 0 & -l_1/l_3 \\ 0 & 1 & -l_2/l_3 \\ 0 & 0 & 1/l_3 \end{pmatrix}$$

**Ejercicio 5.** Identificar los movimientos elementales (traslación, giro, escala, cizalla, proyectivo) representados por las homografías  $H_1$ ,  $H_2$ ,  $H_3$  y  $H_4$ :

$$H_1 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0,5 & 0 & 0 \\ 0 & 0,3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$H_2 = \begin{pmatrix} 0 & 1 & -3 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$H_3 = \begin{pmatrix} 1 & 0,5 & 0 \\ 0,5 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$H_4 = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

**Solución.** La matriz  $H_1$  es trivial: es una transformación afín —la última fila es 0 0 1— formada por la composición de una cizalla, una escala y una traslación, aplicadas en ese orden.

La matriz  $H_2$  es también una transformación afín, aunque ahora las matrices que la forman no son de los tipos que conocemos y tenemos que aplicar una descomposición sencilla.

Para el primer factor del producto, vemos que parece que hay una rotación de ángulo  $3\frac{\pi}{2}$  y una traslación. Llamando  $A$  a la parte de la rotación:

$$A = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

podemos comprobar que lo que queda es efectivamente una traslación. Para eso, resolvemos la ecuación matricial

$$H_{2_1} = A \cdot X$$

que tiene solución por ser  $A$  una homografía —en particular tiene inversa—:

$$X = A^{-1} \cdot H_{2_1}$$

Usamos la misma técnica para el segundo factor, donde observamos que hay una escala y una cizalla. Así, podemos escribir ambos factores como sigue:

$$\begin{aligned} H_{2_1} &= \begin{pmatrix} 0 & 1 & -3 \\ -1 & 0 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \\ H_{2_2} &= \begin{pmatrix} 2 & 0 & 0 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

Y, por tanto, podemos escribir  $H_2$  en función de movimientos que conocemos:

$$H_2 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

es decir,  $H_2$  es la composición de una cizalla, una escala, una traslación y una rotación de ángulo  $3\frac{\pi}{2}$ , aplicadas en ese orden.

La matriz  $H_3$  ya no es una transformación afín, pues la última fila no es  $0 \ 0 \ 1$ . El primer factor sí que lo es, pero la segunda matriz es un movimiento proyectivo. Para identificar el primer factor, lo descomponemos con la misma técnica que hemos usado antes:

$$H_{3_1} = \begin{pmatrix} 1 & 0,5 & 0 \\ 0,5 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0,5 & 0 \\ 0,25 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

y vemos que no es más que la composición de una cizalla y una escala. Podemos escribir entonces  $H_3$  de forma más clara:

$$H_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0,5 & 0 \\ 0,25 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

y concluir que  $H_3$  es la composición de un movimiento proyectivo, una cizalla y una escala, aplicadas en ese orden.

Identificar qué movimiento representa la matriz  $H_4$  requiere un poco más de trabajo, pero no es difícil tampoco. En primer lugar, como sabemos que dos homografías que sólo difieren en un factor no nulo representan el mismo movimiento, podemos trabajar con  $H'_4 = \frac{1}{2}H_4$ , que representa el mismo movimiento y nos permite trabajar más cómodos con la última fila para la descomposición que queremos hacer —pues queremos tener una última fila de 0 0 1 en todos los movimientos afines y queremos que el proyectivo sea la identidad con la última fila arbitraria—.

Por tanto, podemos escribir  $H'_4$  como sigue:

$$H_4 \cong H'_4 = \begin{pmatrix} 1 & 0 & \frac{3}{2} \\ 0 & 1 & \frac{-1}{2} \\ 0 & \frac{1}{2} & 1 \end{pmatrix} = A \cdot P$$

donde  $P$  es la parte de  $H_4$  correspondiente al movimiento proyectivo

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$$

y  $A$  la parte correspondiente a la transformación afín.

Podemos entonces multiplicar por la inversa de  $P$  a la izquierda y obtener la expresión de la parte afín:  $H'_4 \cdot P^{-1} = A$ , que se escribe como sigue:

$$A = \begin{pmatrix} 1 & \frac{-3}{4} & \frac{3}{2} \\ 0 & \frac{5}{4} & \frac{-1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

Esta expresión nos permite hacer una descomposición como antes de una manera mucho más sencilla —vemos que hay una cizalla, una escala y una traslación—.

Por tanto, escribimos  $H'_4$  como el siguiente producto de matrices:

$$H'_4 = \begin{pmatrix} 1 & 0 & \frac{3}{2} \\ 0 & 1 & \frac{-1}{2} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{5}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \frac{-3}{4} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{pmatrix}$$

y concluimos que  $H_4$  es la composición de un movimiento proyectivo, una cizalla, una escala y una traslación, aplicadas en ese orden.

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

**Solución.** Por definición, una homografía entre dos planos proyectivos es un isomorfismo de dichos planos. Las matrices que definen aplicaciones entre planos proyectivos obedecen a las estructuras sobre las que residen los planos, que son espacios vectoriales.

Por tanto, una matriz definirá una homografía entre planos proyectivos cuando defina una isomorfía entre los espacios vectoriales subyacentes, y esto ocurre si y sólo si la matriz es invertible.

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

**Solución.** La única propiedad invariante por una homografía general es la naturaleza de los dos objetos matemáticos esenciales sobre los que actúa; es decir, los puntos se transforman en puntos y las rectas se transforman en rectas.

La distancia, los ángulos y el paralelismo, que son las propiedades geométricas esenciales restantes, son todas dependientes de la homografía particular, y no tienen por qué conservarse.

Por ejemplo, una homografía de *escala* distorsiona distancias, una del tipo *cizalla* distorsiona distancias y ángulos y una homografía del tipo *proyectivo* distorsiona distancias, ángulos y paralelismo —podemos pensar en una fotografía de unas vías de tren; lo que hacemos es una homografía del plano donde yacen las vías al plano que recoge la cámara y, mientras en el suelo las vías son paralelas, en la fotografía intersecan en el horizonte—.

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

**Solución.** Lo que hacemos, geoméricamente, al tomar una fotografía, es aplicar homografías entre planos del espacio y el plano de la cámara. Por tanto, la deformación geométrica más fuerte es la homografía que menos propiedades geométricas conserva.

Como acabamos de ver, las homografías proyectivas son las que menos propiedades conservan, pues sólo mantienen la naturaleza de puntos y rectas.

**Ejercicio 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.

**Solución.** El detector de Harris se basa en una idea central: en una esquina, la intensidad lumínica cambia drásticamente en muchas direcciones. Por tanto, usa la información obtenida de las derivadas en ambos ejes de la imagen para detectar potenciales puntos clave.

Así, lo que detecta son patrones geométricos —bordes, esquinas, cambios drásticos de forma— basándose en la información fotométrica de la imagen en esos lugares.

**Ejercicio 10.** ¿Sería adecuado usar como descriptor de un punto Harris los valores de los píxeles de su región de soporte? En caso positivo identificar cuándo y justificar la respuesta.

**Solución.** Para que un descriptor de un punto clave sea bueno, tiene que ser invariante ante el mayor conjunto de transformaciones posible.

Pero los valores de los píxeles cercanos a un punto clave se ven afectados por cualquier tipo de transformación que no sea una traslación: las escalas pueden expandir un píxel a varios o condensar varios en uno solo, las cizallas pueden hacer lo anterior y cambiar los ángulos con otros píxeles vecinos, una simple rotación los modifica de posición completamente, y un movimiento proyectivo no deja invariante ninguna propiedad de un píxel concreto.

Por tanto, esta elección es muy pobre en un caso general. Si las transformaciones que se esperan son única y exclusivamente traslaciones, esta elección podría ser buena por su sencillez: si el punto clave lo trasladamos con una distancia concreta en una dirección dada, todos sus píxeles se trasladan con la misma distancia y en la misma dirección, así que su valor queda invariante ante la transformación.

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

**Solución.** El descriptor de SIFT de un punto de la imagen es, básicamente, un histograma de la dirección del gradiente —discretizado en 8 direcciones— que tienen los píxeles de una ventana de tamaño  $16 \times 16$  alrededor del punto.

Por tanto, SIFT codifica en qué direcciones hay más cambios bruscos de luminosidad si nos movemos cerca del punto.

**Ejercicio 12.** 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.

**Solución.** El primer criterio que se nos ocurre es el más sencillo: para cada descriptor de la primera imagen, calcular las distancias —como la norma de la resta— entre él y cada uno de los descriptores de la segunda imagen y tomar el que más *cerca* se encuentre.

Esta primera aproximación es muy pobre, ya que es muy probable que dos o más descriptores de la primera imagen tomen un mismo descriptor de la segunda como su mejor *match*.

Otra aproximación podría ser la siguiente: si  $m = \min(n_1, n_2)$  con  $n_i$  el número de descriptores de la imagen  $i$ , podemos considerar todos los conjuntos ordenados de  $m$  descriptores en cada imagen y, para cada par



de ellos, calcular la distancia entre cada par de descriptores en la misma *posición*. Entonces, tomar el par de conjuntos que minimice la suma de distancias y hacer corresponder los elementos en la misma posición.

Este criterio puede ser extremadamente lento si hay muchos descriptores, pero asegura que, en promedio, la distancia entre descriptores en correspondencia es mínima. Sin embargo, no asegura que cada par de descriptores en correspondencia tenga distancia mínima.

**Ejercicio 13.** Cual es el objetivo principal en el uso de la técnica RANSAC. Justificar la respuesta.

**Solución.** El objetivo principal es minimizar el impacto negativo que tienen unos pocos valores atípicos —*outliers*, en inglés— en la obtención de la función que mejor aproxima un conjunto de valores dado.

Este algoritmo surge del problema que presenta la técnica de los mínimos cuadrados cuando existen valores atípicos muy discordantes con el resto de valores buenos, ya que esta técnica termina centrándose en ajustar la función para minimizar el error con los valores atípicos en vez de buscar una función concordante con la mayoría de valores buenos.

RANSAC, a su vez, elimina de una forma cercana a la óptima este problema basándose en la siguiente idea: para cada posible modelo que aproxime el conjunto de datos, el mejor será aquel que tenga menos valores atípicos; o, dicho de otro modo, aquel que tenga más valores concordantes con el modelo.

**Ejercicio 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 puntos en correspondencia necesarios para montar un mosaico? Justificar la respuesta.

**Solución.** Para encontrar una homografía entre dos imágenes necesitamos al menos hacer corresponder 4 puntos de cada imagen. La justificación de esto es simple: las homografías tienen 8 grados de libertad, luego necesitamos al menos 8 puntos para determinar plenamente la matriz que la representa; es decir, necesitamos 4 pares de puntos —cada uno de cada par en una imagen— en correspondencia.

Si tenemos que determinar 3 homografías —una para cada par de imágenes—, es necesario encontrar 12 pares de puntos en correspondencia.

**Ejercicio 15.** 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.

**Solución.** No se puede realizar una proyección rectangular de varias imágenes que abarquen un campo de visión moderadamente grande sin estirar o apretar demasiado los píxeles de los bordes. Esto es una consecuencia directa de trabajar con planos proyectivos: al tener que conservar rectas las

líneas rectas, en ocasiones las deformaciones necesarias son tan grandes que transforman la imagen más de lo deseado.

Estas deformaciones podrían desaparecer si las homografías que hay que aplicar para pasar del sistema de coordenadas de una imagen al sistema de coordenadas de otra nos permiten trabajar sin *estirar* la imagen artificialmente; es decir, si las transformaciones necesarias para pasar de un sistema a otro son únicamente movimientos rígidos: composiciones de traslaciones y rotaciones.