Cuestionario de Teoría-3
Visión por Computador

Daniel Bolaños Martínez

1. Justificar adecuadamente las respuestas.

1.1. ¿Cuál es la transformación más fuerte de la geometría de una escena que puede introducirse al tomar una foto de ella? Dar algún ejemplo.

La transformación geométrica más fuerte que se puede realizar es la proyección.

Cuando tomamos una foto desde una perspectiva diferente, estamos realizando esta deformación y podemos conseguir que un punto que antes formaba parte de la imagen ahora pase a estar en el infinito (punto de fuga), por lo que estaríamos perdiendo puntos de la imagen original y la pérdida de puntos es la deformación más fuerte que podemos realizar a una imagen.

El ejemplo típico, es el de las vías del tren, donde dependiendo de las perspectiva desde la que tomemos la foto, los dos rieles pueden no cortarse nunca o hacerlo en el infinito.

1.2. Por qué es necesario usar el plano proyectivo para estudiar las transformaciones en las imágenes de fotos de escenas? Dar algún ejemplo.

El plano proyectivo tiene unas propiedades geométricas que lo hacen ideal para trabajar con transformaciones en perspectiva de imágenes y posibilita el uso de aplicaciones que no pueden ser representadas en el plano afín.

El plano proyectivo nos permite operar y representar puntos que no se encontraban en la imagen ya que se encuentran en el infinito. Esto se debe a que todo par de rectas se corta en un único punto, incluso las rectas paralelas (en el punto de fuga del infinito).

Retomando el ejemplo de las vías del tren, surge la necesidad de utilizar el plano proyectivo ya que nos permite tener en cuenta puntos que pueden no existir dependiendo de la perspectiva desde la que se tome la foto.

1.3. Sabemos que en el plano proyectivo un punto no existe en el sentido del plano afín, sino que se define por una clase de equivalencia de vectores definida por $\{k(x, y, 1), k \neq 0\}$. Razone usando las coordenadas proyectivas de los puntos afines de una recta que pase por el $(0, 0)$ del plano afín y verifique que los punto de la recta del infinito del plano proyectivo son necesariamente vectores del tipo $(*, *, 0)$ con $*$ = cualquier número.

Queremos probar que todo punto de la recta del infinito es de la forma $(*, *, 0)$ con $*$ = cualquier número. Sabemos que los puntos de la recta del infinito del plano proyectivo son aquellos que no pertenecen al plano afín, procederemos demostrando el contrarrecíproco, es decir, que un punto es representable en el plano afín sii. tiene tercera coordenada no nula.

Definimos una aplicación que lleve puntos del plano afín al proyectivo $\Pi(x, y) = (x, y, 1)$ y la ecuación de la recta que pasa por el origen del plano afín como $L = \{\lambda(x, y) \ : \ \lambda \in \mathbb{R} \ y \ (x, y) \neq (0, 0)\}$.

Aplicando la proyección a la recta obtenemos $\Pi(L) = (\lambda x, \lambda y, 1)$ cuya tercera coordenada es distinta de 0. Para probar el recíproco, tomamos un punto del proyectivo con tercera coordenada distinta de 0 y veamos que tiene representación en el plano afín. Definimos $\Pi^{-1}(x, y, z) = (\frac{x}{z}, \frac{y}{z})$. Sea (x, y, z) con $z \neq 0$, entonces $\Pi^{-1}(x, y, z) = (\frac{x}{z}, \frac{y}{z})$ que pertenece al plano afín.

1.4. ¿Qué propiedades de la geometría de un plano quedan invariantes cuando se toma una foto de él? Justificar la respuesta.

Cuando tomamos una foto de un plano aplicamos una homografía general que realiza al mismo tiempo una transformación proyectiva y una afín.

La única propiedad que queda invariante con la transformación proyectiva es la colinealidad de los puntos. La colinealidad es la respondable de que la homografía lleve rectas en rectas.

Descartamos la invarianza de rotaciones o escalas, ya que no todas las transformaciones afines respetan estas propiedades y sabemos que las transformaciones afines son homografías. Por tanto, la colinealidad será la única propiedad que se cumpla al aplicar una homografía general.

1.5. 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 $x^T l = 0$, es decir $(x_1, x_2, x_3)(a \ \ b \ \ c)^T = 0$. Considere una homografía H que transforma vectores de puntos, $x' = Hx$. Dado que una homografía transforma vectores de tres coordenadas también existen homografías G para transformar vectores de rectas $l' = Gl$. Suponga una recta l y un punto x que verifican $x^T l = 0$ en el plano proyectivo y suponga que conoce una homografía H que transforma vectores de puntos. En estas condiciones ¿cuál es la homografía G que transforma los vectores de las rectas? Deducirla matemáticamente.

Se nos pide encontrar una homografía G que cumpla que $l' = Gl$. Sea x' un punto de la recta l', por contener a ese punto cumplirá $(x')^T l' = 0$.

Consideramos una homografía que cumple $x' = Hx$, por tanto tendremos que $(Hx)l' = 0$. Sabemos que G es una homografía que transforma vectores de rectas $l' = Gl$. Sustituyendo en la ecuación anterior tenemos que $x'l' = (Hx)(Gl) = HxGl = x^T H^T Gl = 0$ obteniendo el resultado pedido cuando $H^T G = I$.

Multiplicando a la izquierda por $(H^T)^{-1}$ en ambos lados de la igualdad $H^T G = I$ (podemos hacerlo puesto que H es invertible por definición de homografía), tenemos que $(H^T)^{-1} H^T G = (H^T)^{-1} I$ por lo que la homografía buscada es $G = (H^T)^{-1}$.

1.6. ¿Cuál es el mínimo número de escalares necesarios para fijar una homografía general? ¿Y si la homografía es afín? Justificar la respuesta

Una homografía general es una transformación lineal invertible de coordenadas homogéneas en el espacio proyectivo. Es una generalización de la transformación afín y su representación matricial viene dada por $\mathbf{y} = H\mathbf{x}$:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\mathbf{y} = H_P\mathbf{x} = \begin{pmatrix} A & \mathbf{t} \\ \mathbf{v}^T & v \end{pmatrix} x, \ \textit{con A invertible}, \ \mathbf{v}^T = (h_{31}\ h_{32}) \ y\ \mathbf{t} = (h_{13}\ h_{23})^T.$$

La matriz H está definida en coordenadas homogéneas por lo que solo se define hasta una escala, y dos matrices H en coordenadas homogéneas serán equivalentes si difieren solo por esta escala. Por tanto, aunque la matriz H tiene 9 parámetros, podemos definirla salvo el factor de escala $v = h_{33}$, por lo que solo necesitaremos **8 escalares** correspondientes a los 4 elementos para la transformación lineal de coordenadas homogéneas, 2 para la traslación y 2 para el vector $\mathbf{v}$.

Por otro lado, una homografía afín es una transformación lineal invertible seguida de una traslación cuya representación matricial viene dada por $\mathbf{y} = A\mathbf{x} + \mathbf{t}$:

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \equiv \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Que extendemos a coordenadas homogéneas añadiendo:

$$\begin{pmatrix} y_1 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix}$$

$$\mathbf{y} = H_A\mathbf{x} = \begin{pmatrix} A & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix} x, \ \textit{con A matriz invertible.}$$

Por tanto, una homografía afín podrá ser representada con **6 escalares** correspondientes a los 4 elementos que definen la transformación lineal y 2 elementos que definen la traslación.

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

Tenemos que definir una homografía H como una matriz 3×3 que transforme el punto $\mathbf{x} = (3,0,2)$ de un plano a otro. Como queremos que $\mathbf{x}$ se tranforme en un punto de la recta del infinito, como ya vimos en la pregunta 3, la última coordenada debe ser igual a 0. Por lo que el punto buscado será $\mathbf{y} = H\mathbf{x}$ donde $\mathbf{y} = (y_1, y_2, 0)$

Sea

$$H = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

Para que se cumpla $\mathbf{y} = H\mathbf{x}$ se debe cumplir $3g + 2i = 0$. Es decir, $i = \frac{-3g}{2}$. Las homografías que son válidas para transformar el punto son:

$$H = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & \frac{-3g}{2} \end{pmatrix}$$

Donde H debe ser invertible ($det(H) \neq 0$) por ser una homografía.

Como se nos pide definir una homografía en particular, tomamos $g = 2$, $a = 1$, $e = 1$ y el resto de componentes 0. Obteniendo:

$$H = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & -3 \end{pmatrix},\ det(H) = -3 \neq 0$$

Sea H la homografía definida y $\mathbf{x} = (3,0,2)$, obtenemos $\mathbf{y} = (3,0,0)$ que pertenece a la recta del infinito del plano proyectivo.

1.8. Una homografía $H = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix}$, $det(H) \neq 0$ admite una descomposición única en movimientos elementales de la siguiente forma $H = H_S H_A H_P$, donde H_S representa la homografía de una similaridad (escala, giro y traslación), H_A la homografía de un movimiento afín puro y H_P una transformación proyectiva pura. Es decir,

$$H_S = \begin{pmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{pmatrix} \equiv \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix}, s > 0$$

$$H_A = \begin{pmatrix} a & c & 0 \\ 0 & b & 0 \\ 0 & 0 & 1 \end{pmatrix} \equiv \begin{bmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}, det(K) = 1$$

$$H_P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ v_1 & v_2 & v \end{pmatrix} \equiv \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{v}^T & v \end{bmatrix}, v \neq 0$$

Describir un algoritmo que permite encontrar las matrices de la descomposición de una matriz H dada. Aplicarlo para encontrar la descomposición de

$$H = \begin{pmatrix} 1{,}707 & 0{,}586 & 1{,}0 \\ 2{,}707 & 8{,}242 & 2{,}0 \\ 1{,}0 & 2{,}0 & 1{,}0 \end{pmatrix}$$

Por el enunciado sabemos que:

$$H = H_S H_A H_P = \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{K} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{v}^T & v \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix}$$

Multiplicando las tres matrices, obtenemos que $\mathbf{A} = s\mathbf{R}\mathbf{K} + \mathbf{t}\mathbf{v}^T$. Sustituyendo cada matriz por su valor con incógnitas tenemos:

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} = s \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_1 & a_2 \\ 0 & a_3 \end{pmatrix} + \begin{pmatrix} c \\ f \end{pmatrix} \begin{pmatrix} g & h \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} = s \begin{pmatrix} a_1 \cos\theta & (a_2 \cos\theta - a_3 \sin\theta) \\ a_1 \sin\theta & (a_2 \sin\theta + a_3 \cos\theta) \end{pmatrix} + \begin{pmatrix} cg & ch \\ fg & fh \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ d & e \end{pmatrix} = \begin{pmatrix} sa_1 \cos\theta + cg & s(a_2 \cos\theta - a_3 \sin\theta) + ch \\ sa_1 \sin\theta + fg & s(a_2 \sin\theta + a_3 \cos\theta) + fh \end{pmatrix}$$

Donde conocemos a, b, d, e, c, h, f, g puesto que son los valores de la matriz H conocida. Por lo tanto tenemos un sistema de 4 ecuaciones con 5 incógnitas. Obtenemos la 5ª ecuación de la condición de que $det(K) = 1$, por lo que $a_1 a_3 = 1$. Por tanto, para obtener la descomposición pedida, debemos resolver un sistema de 5 ecuaciones con 5 incógnitas.

$$\begin{cases} sa_1 \cos\theta + cg = a \\ s(a_2 \cos\theta - a_3 \sin\theta) + ch = b \\ sa_1 \sin\theta + fg = d \\ s(a_2 \sin\theta + a_3 \cos\theta) + fh = e \\ a_1 a_3 = 1, \;\; a_1 \; y \; a_3 \neq 0 \end{cases}$$

Para nuestro ejemplo concreto, tenemos que resolver:

$$\begin{cases} sa_1 \cos\theta + 1{,}0 = 1{,}707 \\ s(a_2 \cos\theta - a_3 \sin\theta) + 2{,}0 = 0{,}586 \\ sa_1 \sin\theta + 2{,}0 = 2{,}707 \\ s(a_2 \sin\theta + a_3 \cos\theta) + 4{,}0 = 8{,}242 \\ a_1 a_3 = 1 \end{cases}$$

$$\begin{cases} sa_1 \cos\theta = 0{,}707 \\ s(a_2 \cos\theta - a_3 \sin\theta) = -1{,}414 \\ sa_1 \sin\theta = 0{,}707 \\ s(a_2 \sin\theta + a_3 \cos\theta) = 4{,}242 \\ a_1 = \frac{1}{a_3} \end{cases}$$

$$\begin{cases} sa_1 \cos\theta = sa_1 \sin\theta \\ s(a_2 \cos\theta - a_3 \sin\theta) = -1{,}414 \\ s(a_2 \sin\theta + a_3 \cos\theta) = 4{,}242 \\ a_1 = \frac{1}{a_3} \end{cases}$$

De la primera ecuación, obtenemos que $\cos\theta = \sin\theta$ por tanto $\theta = 45^{\circ}$ y obtenemos $s = \frac{0{,}707}{a_1 \cos 45}$ por lo que $s = a_3 \frac{0{,}707}{\cos 45} = a_3 \sqrt{2}\; 0{,}707$

$$\begin{cases} a_3 \sqrt{2}\; 0{,}707(a_2 \cos 45 - a_3 \sin 45) = -1{,}414 \\ a_3 \sqrt{2}\; 0{,}707(a_2 \sin 45 + a_3 \cos 45) = 4{,}242 \end{cases}$$

Tenemos que solucionar un sistema sencillo con dos ecuaciones y dos incógnitas a_2, a_3. Cuyas soluciones son $a_3 = 2$ y $a_2 = 1$. Por lo que $a_1 = 0{,}5$, $\theta = 45^o$, $s \approx 2$, $\mathbf{t} = (1\ 2)^T$ y $\mathbf{v}^T = (1\ 2)$. Por lo que podemos expresar la homografía H como descomposición de las siguientes matrices:

$$H = \begin{pmatrix} 2\cos 45^o & -2\sin 45^o & 1 \\ 2\sin 45^o & 2\cos 45^o & 2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0{,}5 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

1.9. ¿Cuáles son las propiedades necesarias y suficientes para que una matriz defina un movimiento geométrico no degenerado entre planos? Justificar la respuesta

Los movimientos geométricos no degenerados entre planos son las homografías que realizan trasnformaciones del tipo: traslaciones, rotación+traslación, semejanzas, afinidades o proyecciones:

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

Las cuatro primeras transformaciones, están definidas como matrices 2×3, por lo que necesitan ampliar una fila añadiendo $(0\ 0\ 1)$ o una proporcional a esta para que la última coordenada sea distinta de cero y no lleve puntos a la recta del infinito del plano proyectivo.

- La **traslación** se define como una matriz arbitraria de dimensión 3×3 tal que:

$$\mathbf{y} = \begin{bmatrix} \mathbf{I} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{x}$$

Donde $\mathbf{I}$ es la matriz identidad de orden 2.

- La **rotación+traslación** se define como una matriz de dimensión 3×3 tal que:

$$\mathbf{y} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{x}$$

  Donde $\mathbf{R}$ es una matriz ortonormal que cumple:

$$\mathbf{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \mathbf{R}\mathbf{R}^T = \mathbf{I} \ y \ |\mathbf{R}| = 1$$

- La **semejanza** se define como una matriz de dimensión 3×3 tal que:

$$\mathbf{y} = \begin{bmatrix} s\mathbf{R} & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{x} = \begin{pmatrix} a & -b & t_x \\ b & a & t_y \\ 0 & 0 & 1 \end{pmatrix} \mathbf{x}$$

  Con $\mathbf{R}$ una matriz de rotación que debe cumplir que $a^2 + b^2 = 1$.

- La **afinidad** se define como una matriz de dimensión 3×3 tal que:

$$\mathbf{y} = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{0}^T & 1 \end{bmatrix} \mathbf{x}$$

  Donde A es una matriz invertible.

- La **proyección** se define como una matriz de dimensión 3×3 tal que:

$$\mathbf{y} = \begin{bmatrix} A & \mathbf{t} \\ \mathbf{v}^T & v \end{bmatrix} \mathbf{x}$$

  La matriz A debe tener determinante distinto de 0, es decir, la matriz debe ser invertible. Esto se debe cumplir para que ningún punto tenga coordenadas homogéneas $(0, 0, 0)$, ya que este punto no existe en el proyectivo.

1.10. ¿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 usa la información que ofrece el gradiente de las derivadas en cada punto con el objetivo de detectar esquinas en la imagen. A partir del gradiente, obtenemos una matriz H (simétrica y diagonalizable) a partir de la cual podremos estimar cambios en la curvatura de la imagen. Los valores propios de la matriz corresponderán con las curvaturas principales de la imagen en cada punto y se utilizarán para obtener cada punto Harris como:

$$f_p(H) = \frac{det_p(H)}{tr_p(H)} = \frac{\lambda_{1p} \cdot \lambda_{2p}}{\lambda_{1p} + \lambda_{2p}}$$

Un valor alto de la función indicará una gran diferencia entre los valores de los gradientes en cada dirección, por lo que el entorno del píxel será una zona con curvatura elevada. Posteriormente, nos quedaremos con los píxeles que superen cierto umbral de respuesta que son los que nos ofrecerán mayor información sobre las esquinas de la imagen.

El detector Harris detecta patrones de ambos tipos: geométricos y fotométricos. Geométricos puesto que usa el gradiente de la imagen para detectar esquinas mediante el proceso anterior y fotométricos porque es invariante a cambios de intensidad de iluminación a partir de cierto umbral.

1.11. ¿Sería adecuado usar como descriptor de un punto Harris los valores de los píxeles de su región de soporte? Identifique ventajas, inconvenientes y mecanismos de superación de estos últimos.

Un descriptor de un punto Harris basado en los valores de los píxeles de la región de soporte no es invariante a transformaciones geométricas tales como escalados, rotaciones, afinidades o cambios en la intensidad de la imagen a partir de cierto umbral, por lo que en general es bastante malo si queremos usarlo posteriormente para establecer correspondencias entre imágenes.

Si tenemos la certeza de que todas las imágenes cuentan solo con transformaciones de traslación y tienen misma escala, orientación y no presentan variaciones de intensidad notables, podemos usar el descriptor definido, ya que la región de soporte es invariante a traslaciones. Si no estamos seguros de esto, lo mejor es usar un descriptor como SIFT que funciona bien en la mayoría de los casos y es invariante entre otras propiedades, a escalados. Otra ventaja es, como hemos visto en la práctica 3, que el detector de puntos Harris es relativamente sencillo de implementar por lo que extraer la información de la región de soporte de los puntos también lo será.

1.12. Describa un par de criterios que sirvan para seleccionar parejas de puntos en correspondencias (”matching”) a partir de descriptores de regiones extraídos de dos imágenes. ¿Por qué no es posible garantizar que todas las parejas son correctas?

Describiremos los métodos utilizados en la práctica 3 para el cálculo de parejas de correspondencias reales entre los descriptores de dos imágenes.

- **Fuerza Bruta + crossCheck**: este método selecciona para cada descriptor de la primera imagen un descriptor de la segunda que minimice la distancia entre ambos. A continuación, se realiza el mismo proceso para cada descriptor de la segunda imagen con la primera. La correspondencia será válida si la pareja de descriptores es la misma en ambos sentidos.

  Este método intenta asegurar la validez de las correspondencias por medio de la validación cruzada, sin embargo si los puntos de la imagen son muy parecidos, podríamos obtener correspondencias incorrectas al comparar entre descriptores similares.

- **Lowe-Average-2NN**: este método calcula para cada descriptor d de la primera imagen los dos descriptores de la segunda imagen que estén a menor distancia, definamoslos como d_1 y d_2. La correspondencia de (d, d_1), será válida si el cociente de las distancias:

$$\frac{\mathbf{d}(d, d_1)}{\mathbf{d}(d, d_2)} < k, \; con \; k \; una \; constante \; prefijada.$$

  Este método tiene en cuenta a partir del valor de k, posibles ambigüedades entre correspondencias de descriptores similares. A menor valor de k, obtenemos mayor fiabilidad en las correspondencias pero podemos estar descartando correspondencias válidas. Conforme aumentamos k, aparece el problema del método de Fuerza Bruta. Deberemos elegir un valor k que mantenga cierto equilibrio entre cantidad de correspondencias y validez de las mismas.

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

RANSAC es un algoritmo de estimación de homografías entre dos imágenes que tiene como objetivo minimizar el número de correspondencias con posibles outliers.

Este método toma un conjunto de 4 correspondencias escogidas de forma aleatoria y estima una homografía con ellas. Con esta homografía, se valoran los resultados obtenidos calculando cuantos puntos dentro de todas las correspondencias escogidas son compatibles, descartando aquellas correspondencias con un error mayor a uno prefijado. Si la cantidad de puntos compatibles son suficientes, se utilizan estos puntos para estimar mejor la homografía que habíamos aproximado inicialmente.

1.14. 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 número mínimo de parejas de puntos necesarios para montar un mosaico con 4 imágenes es 12.

Necesitamos 4 parejas de puntos para poder calcular la matriz de la homografía entre dos imágenes. Se nos pide montar un mosaico con 4 imágenes. Por tanto, necesitaremos calcular tantas homografías como número de solapamientos haya en el mosaico. En nuestro caso, necesitamos 3 homografías (H_{12}, H_{23}, H_{34}) y una homografía que lleve la primera imagen al centro del canvas donde se representará el mosaico (H_0). Como para H_0 no es necesario obtener correspondencias a través de una homografía, en total necesitaremos:

$$4\ \frac{parejas\ de\ puntos}{homografia} \cdot 3\ homografias = 12\ parejas\ de\ puntos$$

.

Finalmente, calcularemos las traslaciones de las imágenes al mosaico por medio de la composición de homografías H_{12}, H_{23}, H_{34} y H_0.

1.15. ¿En la confección de un mosaico con proyección rectangular es esperable que aparezcan deformaciones geométricas de la escena real? ¿Cuáles y por qué? ¿Bajo qué condiciones esas deformaciones podrían no estar presentes? Justificar la respuesta.

Cuando estamos realizando la confección de un mosaico sobre un canvas rectangular podemos encontrar varios errores que provocan deformaciones sobre la escena real que estamos construyendo, a saber:

La acumulación de errores al estimar las homografías para cada par de imágenes, lo que podemos solucionar, empezando el mosaico a partir de una imagen situada en el centro del canvas.

Otro error que suele aparecer en la confección de un mosaico, es el cambio de proyección o cambio de perspectiva. Conforme creamos el mosaico, obtenemos deformaciones geométricas hacia los extremos ya que las imágenes tienden a deformarse para ajustarse unas a otras acumulando un error de perspectiva generado al tomar la imagen. Esto ocurre porque los KeyPoints de las imágenes contiguas no tienen por qué estar alineados. Para arreglar este problema, antes de tomar las fotografías, podríamos obtener las imágenes para nuestro mosaico desplazando la cámara horizontalmente y evitar rotarla en la medida de lo posible minimizando así las deformaciones hacia los extremos.

Referencias

[1] Diapositivas de clase.

[2] R.Hartley, A.Zisserman. Multiple View Geometry in computer vision.

[3] R.Szeliski. Computer Vision: Algorithms and Applications.