

# *Visión por computador*

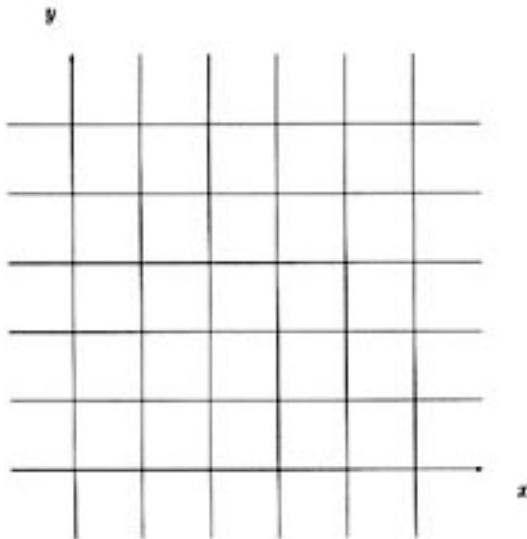
## *Trabajo 2*

### *Cuestionario*

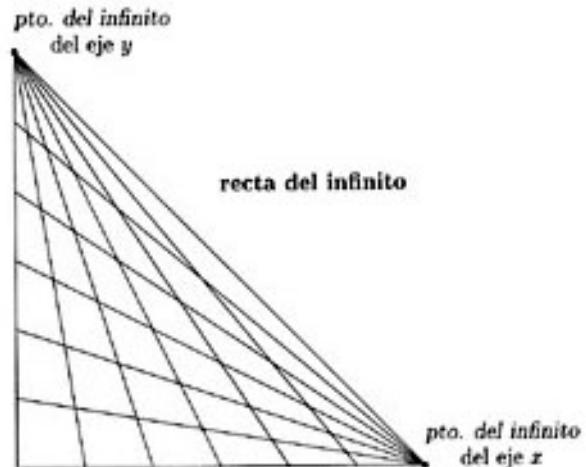


## Cuestionario

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



Plano afín  $A^2$



Completado proyectivo  $P^2$

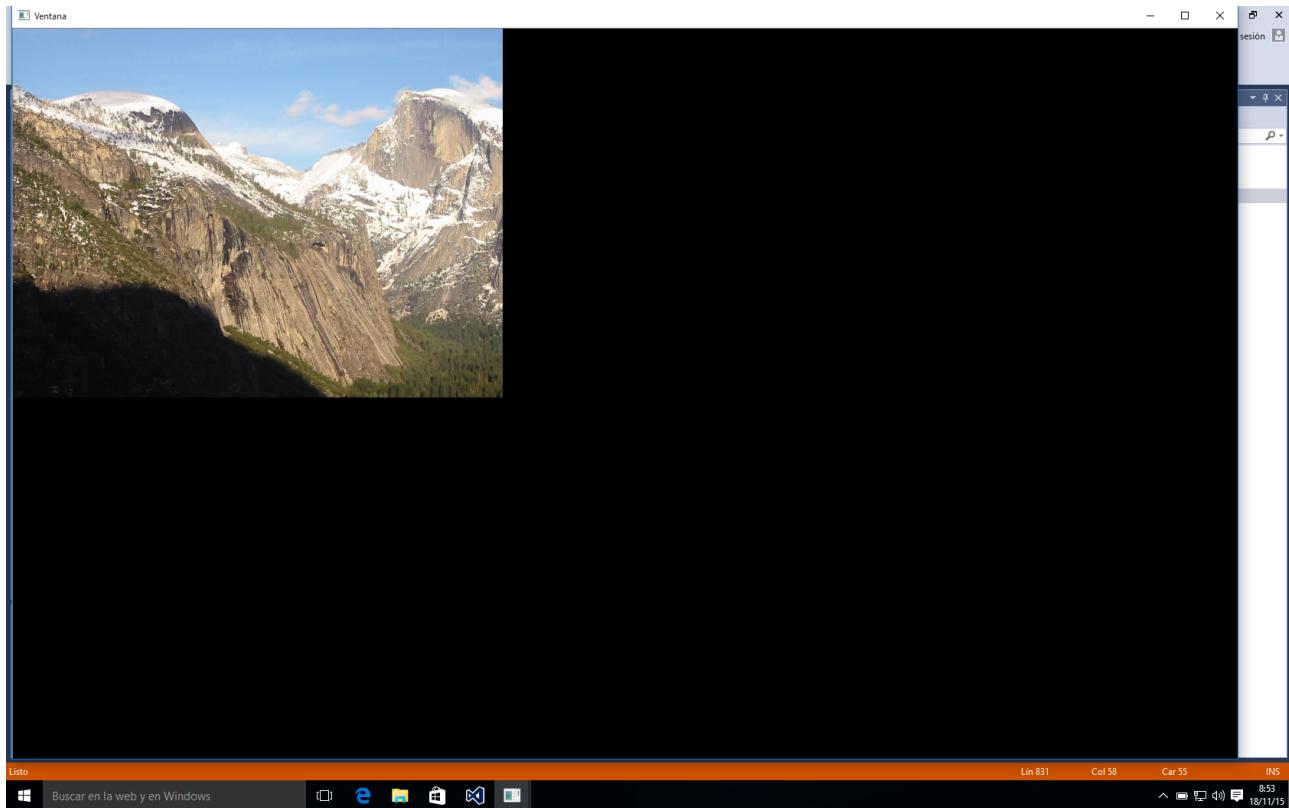
El plano proyectivo es el plano afín más una recta (proyectiva) de puntos de infinito. Las consecuencias serían que en el plano proyectivo se podría apreciar que las rectas se cortan en un punto del infinito, mientras en el plano afín se ve que no se cortarían.

**5.- Identificar los movimientos elementales (traslación, giro, escala, cizalla, proyectivo) representados por las homografías H1, H2, H3 y H4:**

Si tenemos que  $H_0$  es la siguiente matriz:

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

El resultado de aplicar esa homografía sería:

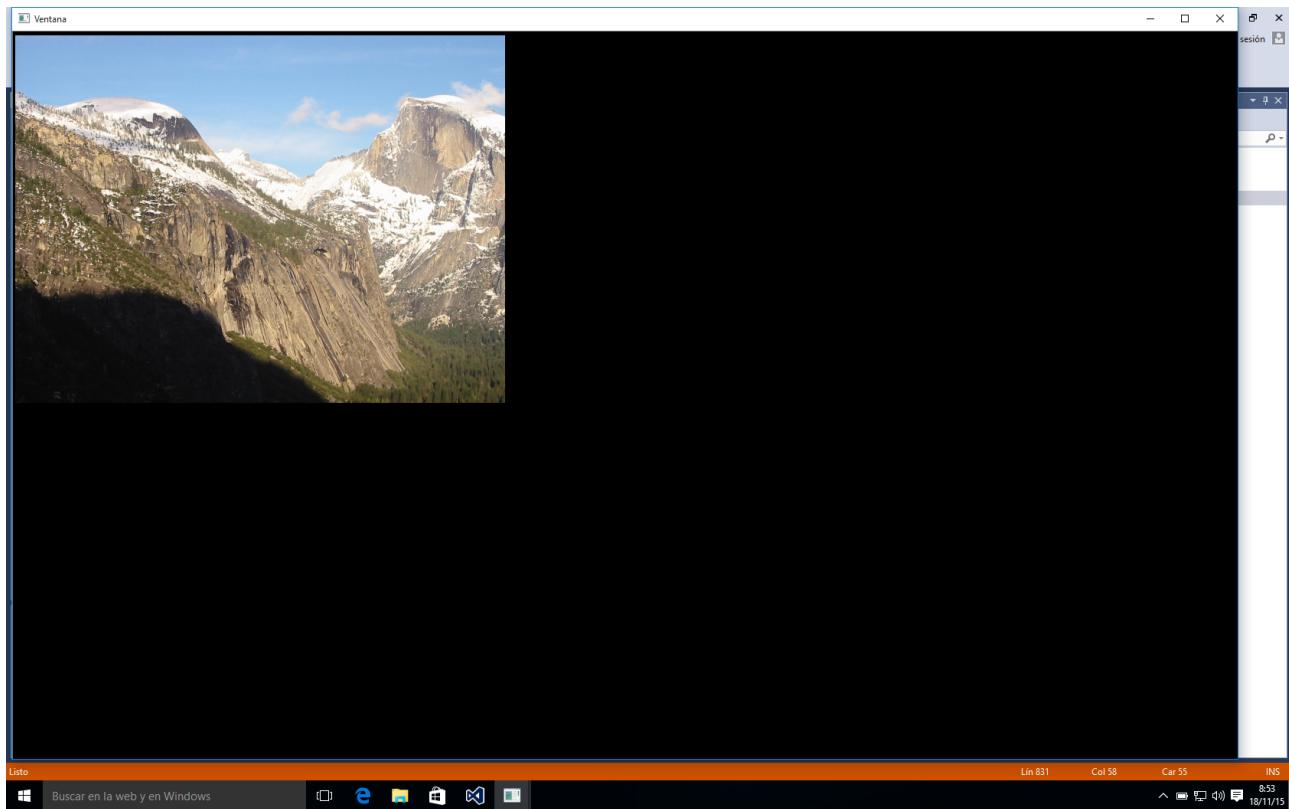


Ahora tenemos H1:

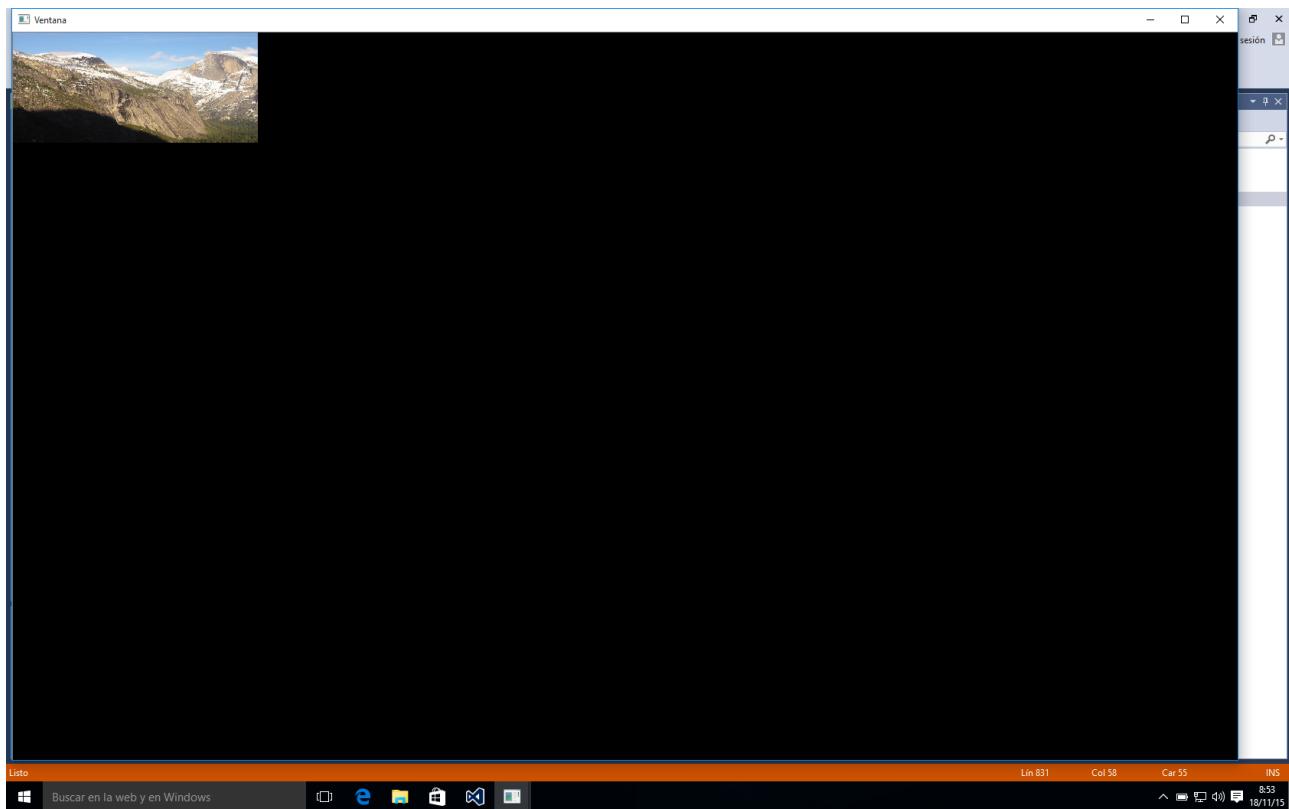
$$H1 = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0.5 & 0 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Voy a hacer las pruebas poco a poco, es decir, voy a llamar H1 primera, a la primera matriz que aparece H1, H2 segunda a la segunda, y H1 tercera a la tercera. De esta forma vemos como afecta cada una de ellas a la imagen, y finalmente a la multiplicación de todas.

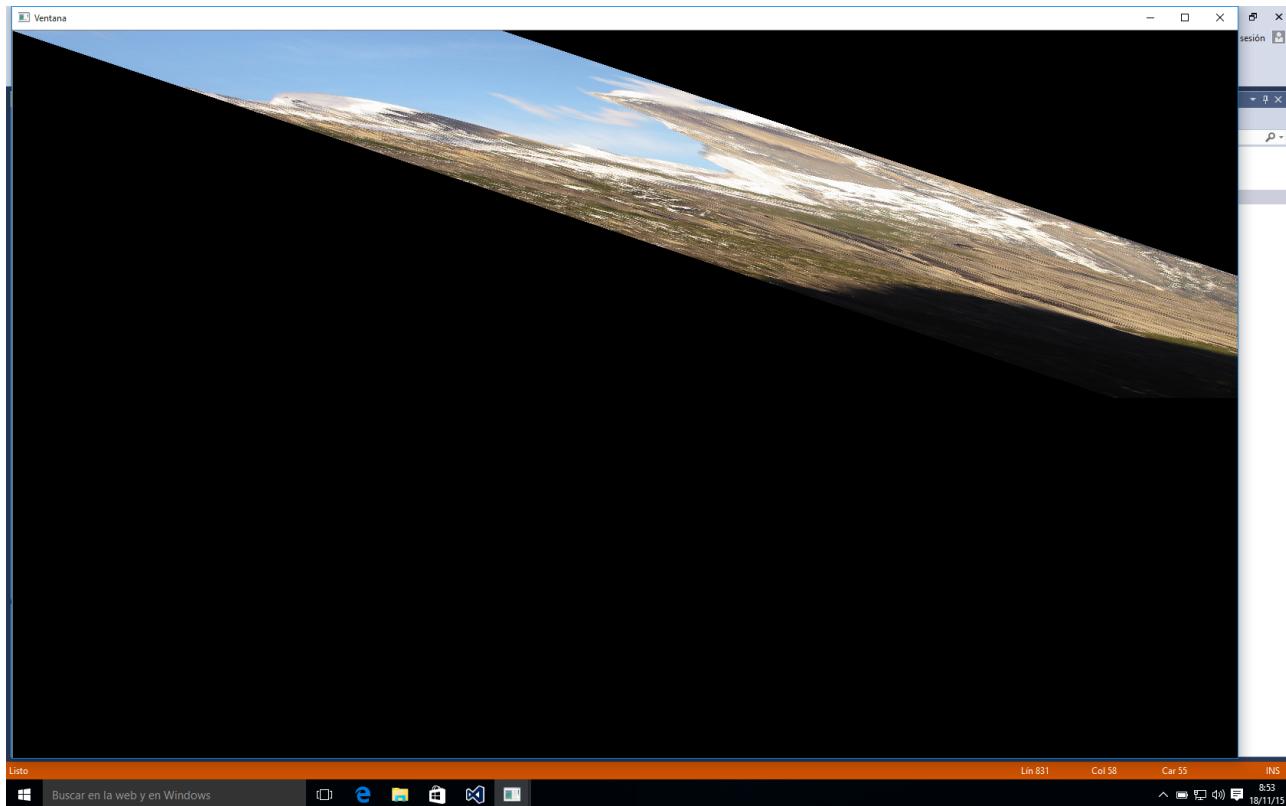
H1 primera, podemos ver por su estructura que es una translación:



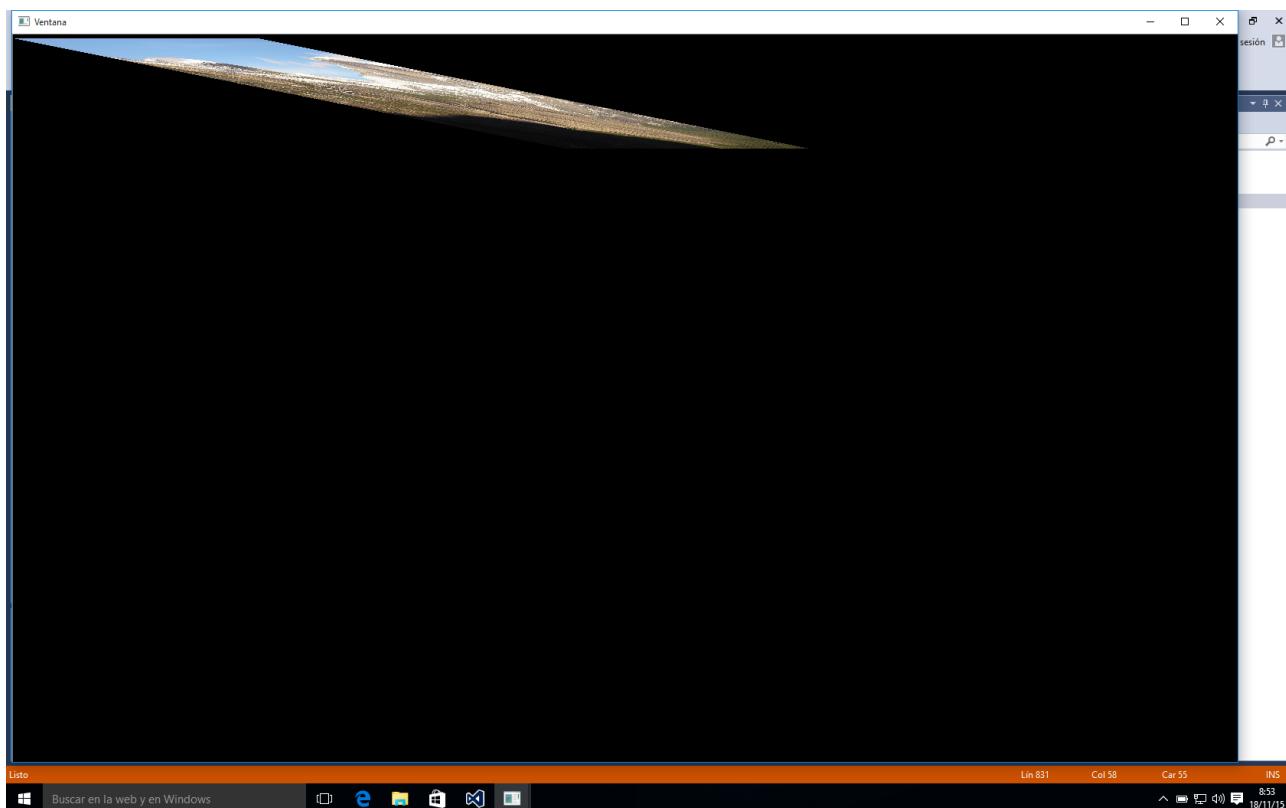
H1 segunda sería una escalada:



H1 tercera sería un shear:



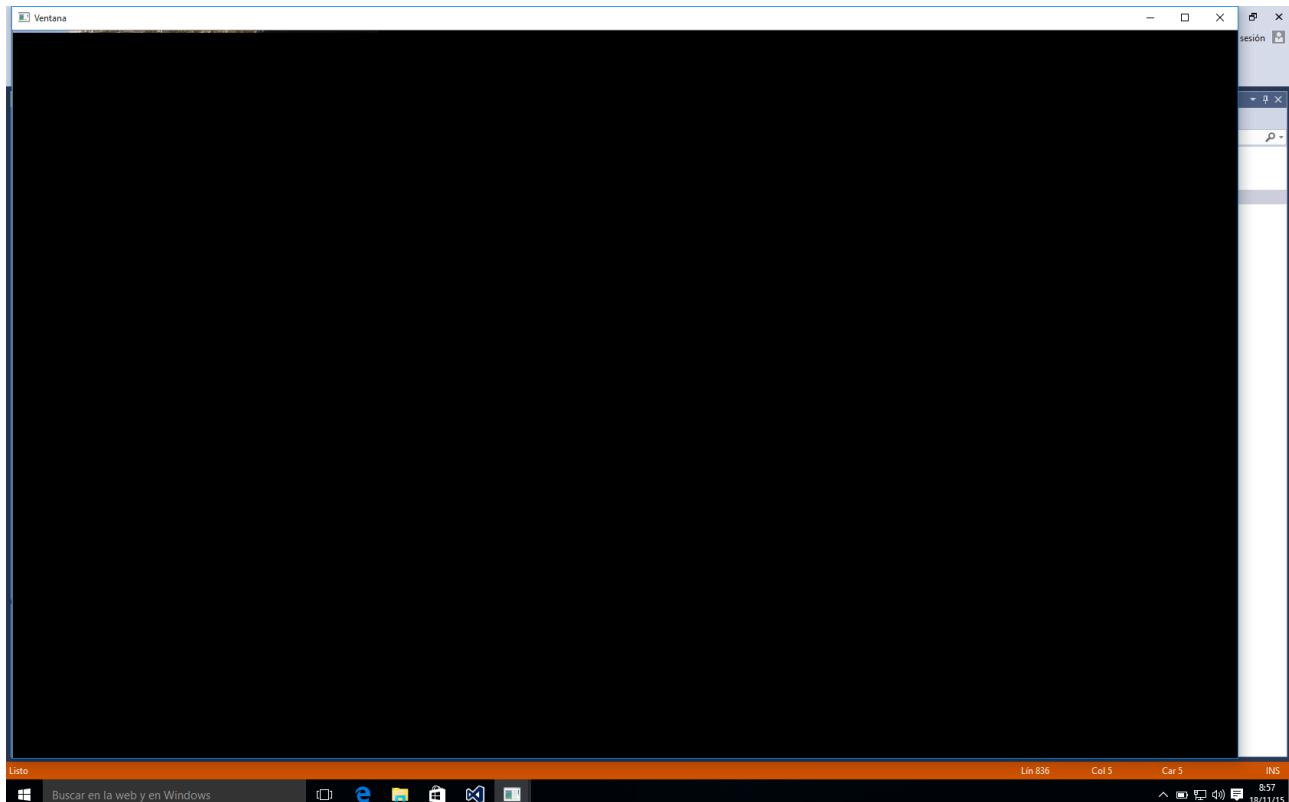
Finalmente la multiplicación de todas las matrices H1 quedaría:



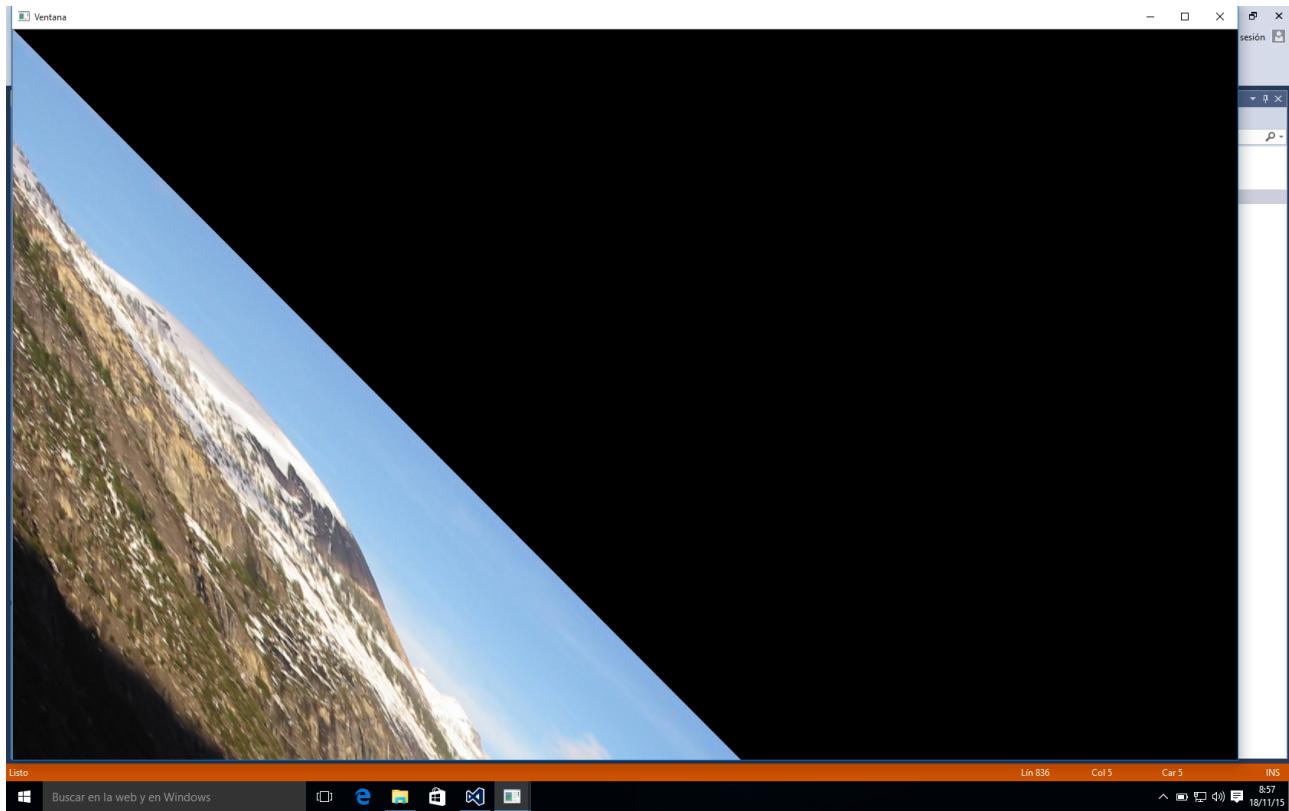
Ahora voy a realizar lo mismo que en H1 con H2:

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

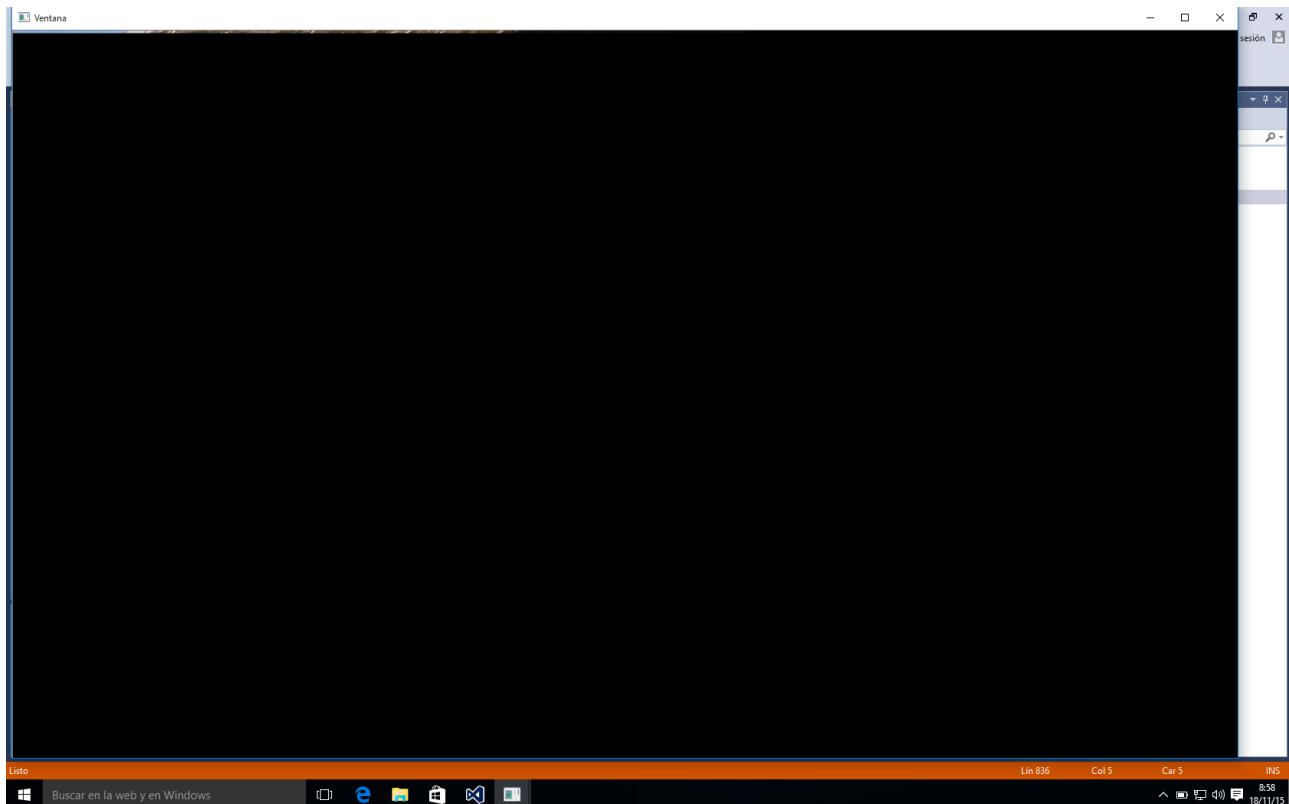
H2 primera se aprecia muy poco, pero la imagen aparece arriba del todo en un tamaño muy pequeño:



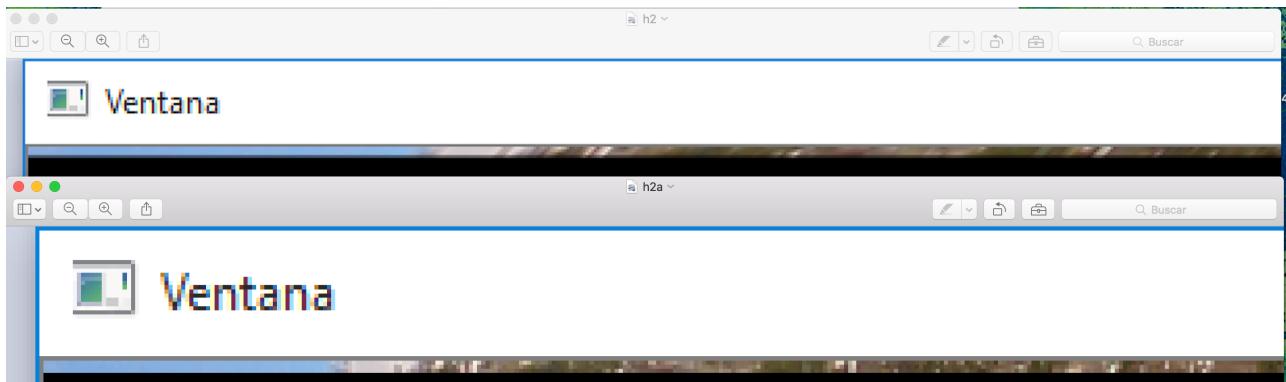
H2 segunda correspondería a una rotación:



H2 quedaría:



Si ampliamos este resultado, por intentar apreciar algo, quedaría:

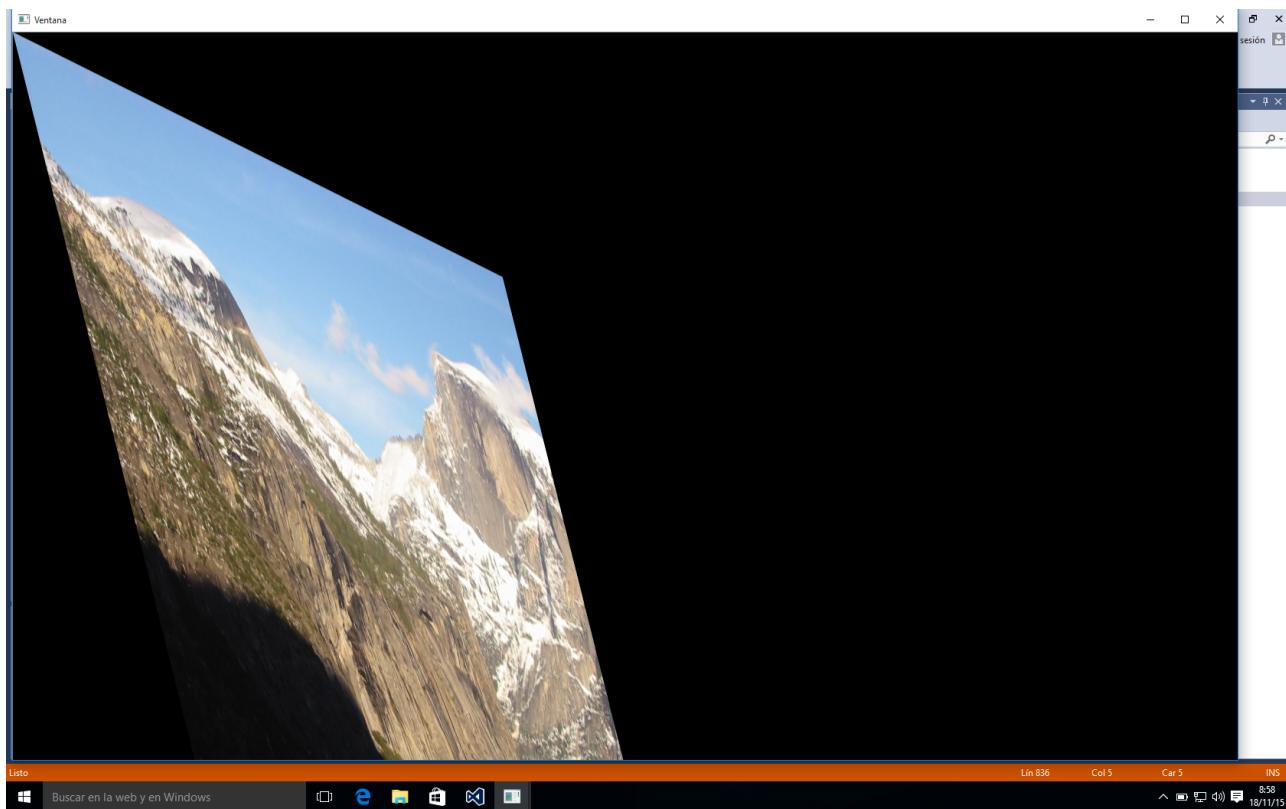


Siendo H2 la imagen superior, y H2 primera la imagen inferior.

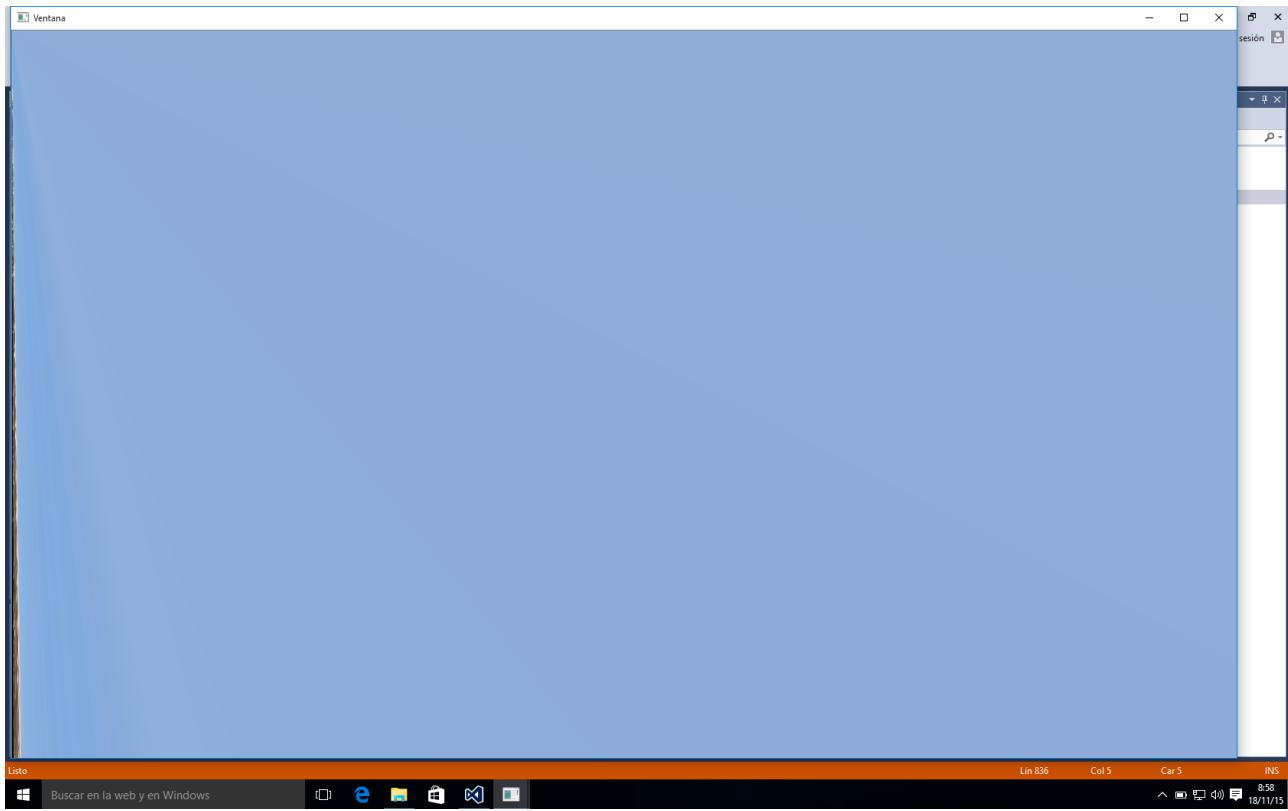
Si hacemos lo propio con H3:

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

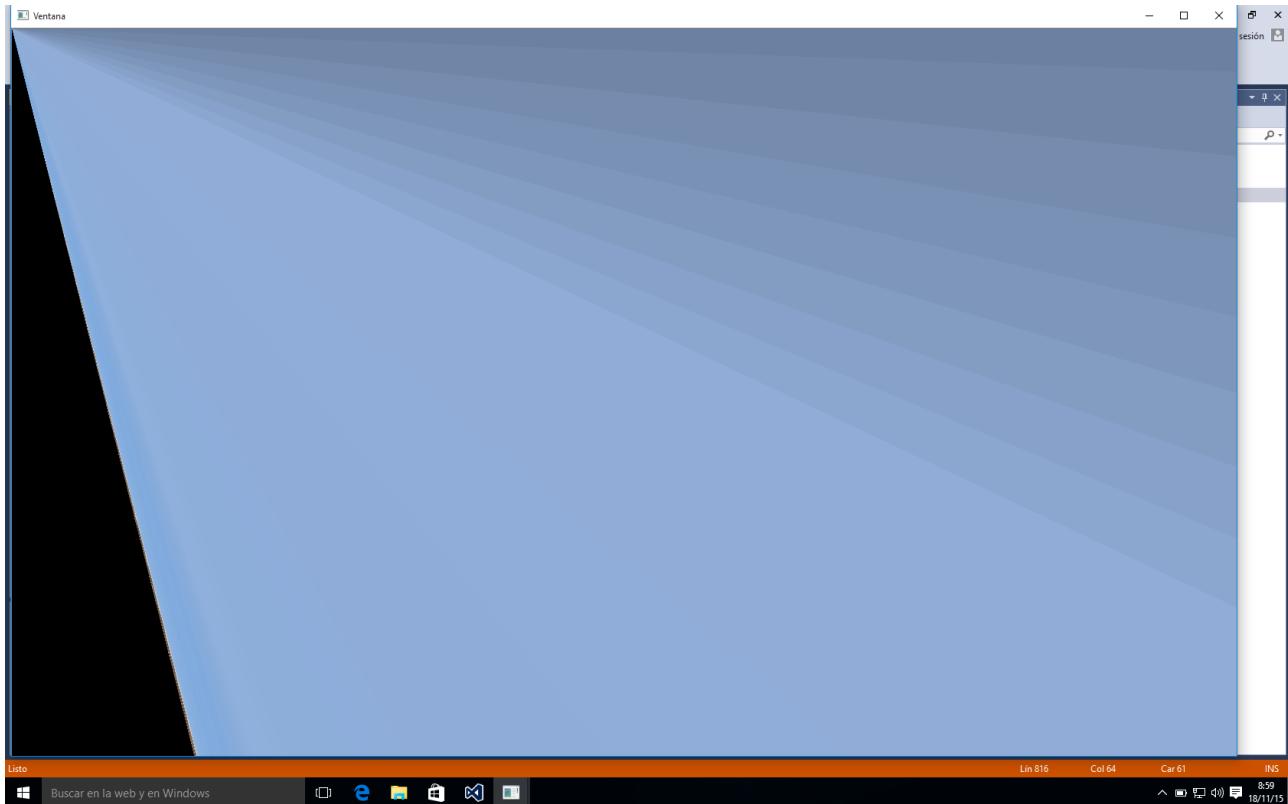
Quedando H3 primera sería una rotación/shear:



H3 segunda sería:



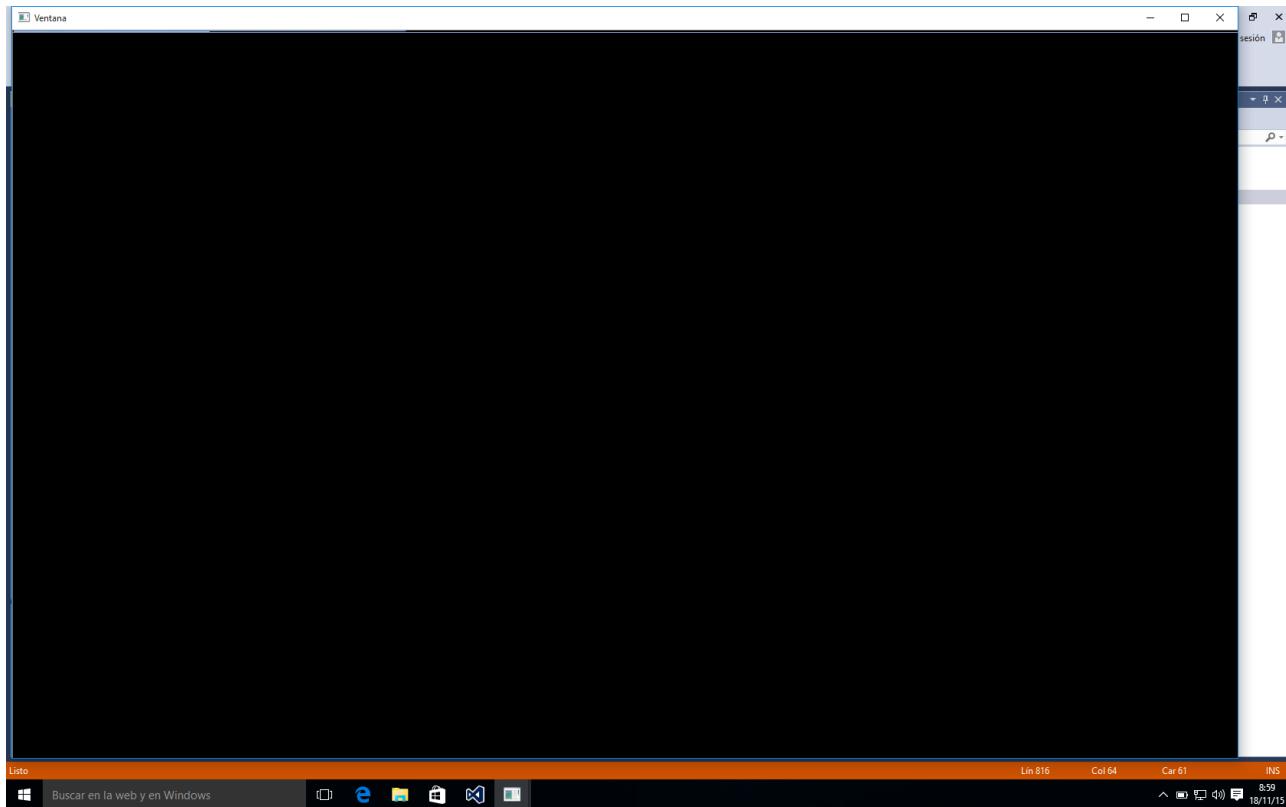
Y por último, H3 quedaría:



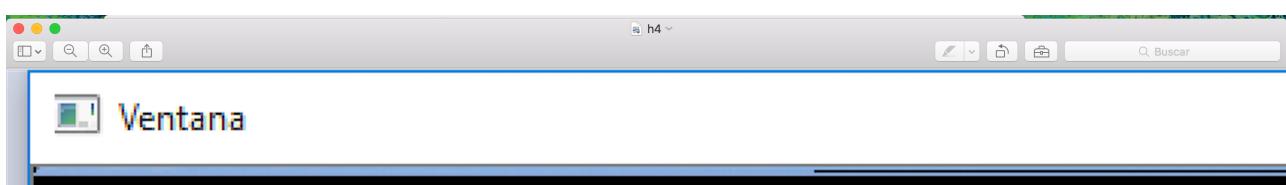
Por último, H4 es:

$$H4 = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

Quedando:



Que aumentado sería:



## 6.- ¿Cuáles son las propiedades necesarias y suficientes para que una matriz defina una homografía entre planos?

En esta imagen vemos la correspondencia lineal:  $x' = H x$ .

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x \\ y \\ w \end{bmatrix}$$

Por lo tanto, una de las propiedades de una homografía  $H$  es que tiene correspondencia lineal entre vectores expresados en coordenadas homogéneas de 3 dimensiones. Además el determinante de la matriz  $H$  tiene que ser distinto de 0.

Como  $H$  es una matriz no singular, se podrán invertir las transformaciones proyectivas y la inversa de  $H$  será también una proyectividad.

La matriz  $H$  se podrá multiplicar por un valor que no sea cero, sin que la transformación proyectiva varíe.

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

Las homografías son transformaciones proyectivas que establecen correspondencia entre elementos iguales. Por lo tanto, la homografía conserva la naturaleza de los elementos transformados teniendo las unas propiedades invariantes como que un punto se transforma en otro punto, una recta se transforma en otra recta, y un plano se transforma en otro plano. Por lo tanto, la translación, giro y afinidad serán homografías.

Cuando se aplica una homografía, se conservará la linealidad de la transformación y si dos líneas se cortasen, se conservaría la congruencia/concurrencia.

## 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?

En la proyectiva se conservará de la imagen inicial, solo los vértices y lo que es recto. Mientras tanto, en la proyectiva no sólo no hay noción de distancia, sino que no hay noción de paralelismo. Esto sería como en la euclídea y en la afín.

## **9 .-¿Qué información de la imagen usa el detector de Harris para seleccionar puntos?**

De una imagen podremos extraer información que resume el comportamiento de esa imagen. Hay técnicas para ello como son extraer regiones homogéneas, extraer esquinas...

Esta información puede estar en las transformaciones, en los cambios de intensidad o en la estructura. Para comparar, se intenta llevar el mismo punto de representación, por lo que el detector Harris usa el gradiente (representa las variaciones de textura y grises). Como se umbraliza la información del gradiente, esta es relevante, ya que se reduce el número de píxeles interesantes, quedando así los más altos.

La información importante del gradiente se suele encontrar en las esquinas, ya que si se desplaza un pixel, todo cambia, con lo que tienen el valor máximo del gradiente, por lo que se extraerá el máximo local. Como una esquina tiene dos direcciones, el gradiente es alto.

Los autovalores representan la matriz H de los gradientes (cambios de dirección), por lo que si el resultado de ambos es alto, los cambios de dirección son fuertes.

En definitiva, la información importante se concentra en la esquina, de donde se extraen los puntos que serán máximos locales, y donde con el detector Harris se extraerá un alto valor del gradiente al tener dos direcciones y representar las variaciones.

## **¿El detector de Harris detecta patrones geométricos o fotométricos?**

Los patrones geométricos son los que detectan formas geométricas, mientras que los patrones fotométricos son descripciones distintivas en función del valor del pixel. Por lo tanto, el detector Harris detectaría patrones de tipo geométrico ya que detecta esquinas, o lo que es lo mismo, formas geométricas.

## **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**

El detector Harris no es un descriptor, y esto se dividiría en detección y descripción, independientemente. Por lo tanto el detector Harris calcularía el gradiente en una región, seleccionando los máximos locales, contando con las esquinas.

Dicho esto, sería adecuado usar como descriptor de un punto Harris los valores de los píxeles de su región de soporte cuando las imágenes puedan tener figuras geométricas o alguna transformación.

## **11.- ¿Qué información de la imagen se codifica en el descriptor de SIFT?**

Una vez encontrados los KeyPoint, vamos a calcular los descriptores para cada punto de forma que sea fácil identificarlo y parcialmente invariante a la iluminación, punto de vista...

Los descriptores SIFT son invariantes a cambios afines menores y para probar el carácter distintivo de estos, se iguala la precisión que se mide contra un número variable de puntos clave en la imagen, demostrando que la adecuación de la precisión disminuye los tamaños grandes de imágenes.

En definitiva, los descriptores SIFT contienen puntos de interés de las imágenes que serán usados para encontrar coincidencias con otro conjunto de descriptores de otra imagen.

## 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

El matching tiene como finalidad calcular un valor, el cual representa un grado de similitud entre dos imágenes. A partir de ese grado de similitud vamos a establecer diferentes conclusiones.

Para calcular este valor, primero, estableceremos las correspondencias entre los puntos clave, y segundo, se realizará una aplicación métrica o cálculo de distancia entre las imágenes participantes. La correspondencia de los puntos clave será la distancia euclídea entre los vectores de diferentes puntos de interés de la imagen. Este cálculo, además generará un valor que nos permite determinar que punto de la imagen corresponde con su homólogo en la otra imagen.

## 13.- ¿Cuál es el objetivo principal en el uso de la técnica RANSAC?

RANSAC es un consenso de muestra aleatoria, el cual tiene 5 pasos a seguir:

1. Elegir aleatoriamente  $s$  muestras (siendo  $s$  el tamaño mínimo que permite encajar un modelo)
2. Colocar un modelo (por ejemplo, una línea) para las muestras
3. Contar el número de inliers que caben en el modelo aproximadamente
4. Repetir  $N$  veces
5. Elegir el modelo que tiene el mayor conjunto de inliers.

Por lo tanto, esta técnica tiene distintos Pros:

- Es simple y general.
- Es aplicable a muchos problemas diferentes
- A menudo funciona bien en la práctica
- Existen implementaciones para tiempo real

Por otra parte, las contras serían:

- Los parámetros para sintonizar
- A veces se requieren muchas iteraciones
- Puede fallar con inliers ratios extremadamente pequeños

En definitiva, el objetivo de RANSAC sería calcular, mediante un método iterativo, los parámetros de un modelo matemático de un conjunto de datos observados que contiene valores atípicos.

**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 correspondencias necesarios para montar un mosaico?**

Si tenemos 4 imágenes de una escena de forma que se solapan de 1-2, 2-3 y 3-4 el número mínimo de puntos en correspondencias que necesitaríamos serán 4, igual que si fuese para dos solas imágenes. En este caso hay una restricción para ello, que los 4 puntos estén y se encuentren como correspondencias en las 4 imágenes.

En la siguiente imagen de ejemplo se puede apreciar como podrían ser los 4 puntos necesarios para montar el mosaico:



Por lo tanto, el número mínimo de puntos en correspondencias necesarios para montar el mosaico sería 4.

**15.- En la confección de un mosaico con proyección rectangular es esperable que aparezcan deformaciones de la realidad. ¿Cuáles y porqué?**

Si es esperable que aparezcan deformaciones de la realidad como que las líneas paralelas dejen de serlo.

**¿Bajo qué condiciones esas deformaciones podrían desaparecer?**

En el caso en el que las transformaciones solo sean translaciones.