

Informe Tarea No 1

Tatiana Lopez Guevara
Universidad Tecnológica de Pereira
zepolitat@utp.edu.co

Resumen—El presente documento explica los resultados obtenidos al aplicar y remover una distorsión proyectiva sobre un conjunto de imágenes. El objetivo principal era aplicar las técnicas vistas en la primera parte del curso de Computer Vision para afianzar los conceptos aprendidos.

Index Terms—Computer Vision

I. INTRODUCCIÓN

PARA el desarrollo del proyecto se construyeron los siguientes archivos fuente:

- **zpart1.m**
Archivo principal de la parte 1 de la tarea. Contiene los parámetros (ruta de la imagen y coordenadas de los 8 puntos) así como las invocaciones a las funciones que en su conjunto, remueven la distorsión de la imagen.
- **zimread.m**
Carga una imagen y retorna su representación como un vector 2D de $W * H$ filas y 3 columnas de color. Por la agilidad de su operación, esta representación es la usada principalmente para realizar las operaciones sobre la imagen. Adicionalmente, retorna una matriz tridimensional de H filas, W columnas y 3 canales de color RGB principalmente para el despliegue en pantalla.
- **zfindH.m**
Encuentra una homografía H a partir de las 2 parejas de 4 puntos aplicando (1).
- **transformCorner.m**
Calcula las esquinas de una imagen dado el ancho y el alto y aplica una homografía sobre estas. Esta función invoca a transformX.
- **transformImg.m**
Genera la combinación de puntos (x, y) de una imagen y aplica una homografía invocando a transformX.m.
- **transformX.m**
Aplica una homografía a una matriz de puntos de la forma $[x_1 x_2 \dots x_N; y_1 y_2 \dots y_N]$ y la normaliza.
- **bilineal.m**
Realiza una interpolación bilineal de una imagen en representación vectorial de tamaño $[(W * H), 3]$ y de un conjunto de índices, posiblemente de valores no enteros, que la referencian.
- **zpart2.m**
Archivo principal de la parte 2 de la tarea. Contiene los parámetros (ruta de las imágenes y coordenadas de los 8 puntos) así como las invocaciones a las funciones que en su conjunto, sobrepone una imagen planar sobre otra que tiene alguna distorsión.

Las homografías H usadas, corresponden al grupo de

transformaciones PL(3) de 8 DoF y por lo tanto requiere de 8 puntos para caracterizarla [1]

II. PRIMERA PARTE

El objetivo de la primera parte era remover la distorsión de una imagen tomada con una cámara aplicando interpolación bilineal.

Se tomaron las siguientes convenciones en el código fuente:

- La imagen distorsionada se toma como el conjunto de puntos sobre los que existe una distorsión x' y todo lo relacionado con éstas tiene el prefijo la letra i (inicial).
- La imagen que será creada al remover la distorsión representa el conjunto de puntos x y la información relacionada con éstos tiene como prefijo la letra f (final).

Los parámetros de la imagen y de las coordenadas de los 4 puntos en la imagen inicial y sus 4 puntos equivalentes en la imagen final se ingresan en las primeras líneas del script *zpart1.m*.

```
[iVecImg, iImg, iW, iH] =
    zimread( '../imgs/scan.jpg' );

xp=[32  354 528 138];
yp=[204 138 568 677];
x =[100 460 460 100];
y =[150 150 670 670];

imshow(iImg);
hold on;
scatter(xp,yp,5,'r');
```

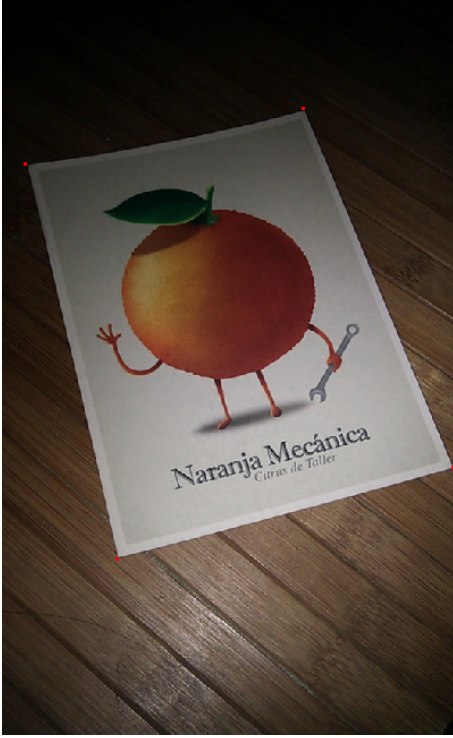
El archivo *scan.jpg* representa la imagen distorsionada de la figura 1.

Al insertar estos 4 puntos x y sus equivalentes x' en 1, se hallaron los parámetros de la matriz de proyección H .

$$\begin{pmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{pmatrix} = \begin{pmatrix} x_1 & y_1 & 0 & 0 & 0 & -x'_1 x_1 & -x'_1 y_1 \\ 0 & 0 & 0 & x_1 & y_1 & -y'_1 x_1 & -y'_1 y_1 \\ & & & & \vdots & & \\ x_4 & y_4 & 0 & 0 & 0 & -x'_4 x_4 & -x'_4 y_4 \\ 0 & 0 & 0 & x_4 & y_4 & -y'_4 x_4 & -y'_4 y_4 \end{pmatrix}^{-1} \begin{pmatrix} x'_1 \\ y'_1 \\ \vdots \\ x'_4 \\ y'_4 \end{pmatrix} \quad (1)$$

Lo primero que se hizo fue hallar el tamaño de la imagen final. Para esto se aplicó la homografía inversa sobre las esquinas de la imagen cargada empleando (2).

Figura 1: Imágen Inicial



$$x = H^{-1}x' \quad (2)$$

En este punto, existían 2 opciones para remover la distorsión:

1. Partir de la imagen inicial, calculando para cada uno de sus puntos (x', y') el correspondiente valor de (x, y) usando (2).
2. Partir de la imagen final, calculando para cada uno de sus puntos (x, y) el correspondiente valor de (x', y') mediante (3).

La primera opción, aunque parece más natural, tiene el inconveniente de que no necesariamente se asigna un valor para todos los píxeles de la imagen final, dejando espacios vacíos (de color negro) que se debe rellenar en un paso posterior. A pesar de que se ensayaron las 2 opciones, sólo se dejó la segunda en el código y es la que se continuará explicando.

Teniendo entonces el tamaño de la imagen, se procedió a encontrar para cada punto (x, y) de la imagen final, su equivalente (x', y') con el fin de conocer el color RGB del píxel correspondiente a (x, y) . Cabe notar que a pesar de que (3) hace referencia a la transformación de un solo punto, en el desarrollo se usó la representación matricial en donde x es el conjunto de puntos de la forma $[x_1 y_1 1; x_2 y_2 1; \dots x_N y_N 1]'$ y por lo tanto x' contiene todos los puntos transformados de x ;

$$x' = Hx \quad (3)$$

Sin embargo, los puntos (x', y') que se obtienen de (3) pueden no ser enteros y por lo tanto se aplicó la técnica de interpolación bilineal (4) [2].

Figura 2: Imágen Final



$$f(x, y) \approx \frac{1}{(x_2 - x_1)(y_2 - y_1)} (f(Q_{11})(x_2 - x)(y_2 - y) + f(Q_{21})(x - x_1)(y_2 - y) + f(Q_{12})(x_2 - x)(y - y_1) + f(Q_{22})(x - x_1)(y - y_1)) \quad (4)$$

Donde para este caso la distancia entre las coordenadas de las cuales se obtendrá el promedio ponderado de un punto desconocido es 1, con lo que el término del denominador se vuelve 1.

Finalmente, al aplicar la homografía y la interpolación bilineal sobre la figura 1 se obtuvo la figura 2.

III. SEGUNDA PARTE

La segunda parte del proyecto consistió en utilizar la misma técnica para aumentar una escena, sobreponiendo una imagen planar sobre una imagen base, con la perspectiva correcta.

La imagen planar usada en el proyecto corresponde a la mostrada en la figura 3. El objetivo entonces era aplicar una distorsión proyectiva sobre ésta con el fin de sobreponerla en los puntos marcados con rojo de la figura 4.

Figura 3: Imágen Planar



Figura 4: Imágen con distorsión proyectiva



De forma similar a la parte 1, se realizó el cálculo de la matrix de transformación H a partir de los 8 puntos dados. Luego, ésta se usó para aplicar la distorsión proyectiva sobre cada punto de la imágen planar mediante (3). El resultado obtenido se puede ver en la figura 5.

Figura 5: Imágen planar sobrepuesta en imágen con distorsión proyectiva



IV. CONCLUSIONES

- Al no tener limitaciones de las posibles transformaciones que tenía la imágen origen, se debía asumir que éstas podían tener elementos de rotación, traslación y escalamiento. Por lo tanto, las homografías H con las

que se debía trabajar pertenecen al grupo $PL(3)$ de 8 DoF requiriendo 8 puntos para encontrarla.

- Existen dos posibilidades de remover la distorsión proyectiva de una imágen. La primera es partir de la imágen inicial, calculando para cada uno de sus puntos el valor equivalente en la imágen destino. La segunda es el proceso inverso partiendo de cada pixel de la imágen destino hacia la inicial.
- La interpolación bilineal permite realizar una aproximación suave de un pixel no entero a partir de sus pixeles vecinos, mientras que aproximar directamente a un entero, hace que la imágen se vea más pixelada ya que es una aproximación "dura".

REFERENCIAS

- [1] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*, volume 2. Cambridge Univ Press, 2000.
- [2] Wikipedia. Bilinear interpolation. http://en.wikipedia.org/wiki/Bilinear_interpolation. [Online; accessed Mar-2013].