Algebra proyectiva en un problema de Optimización

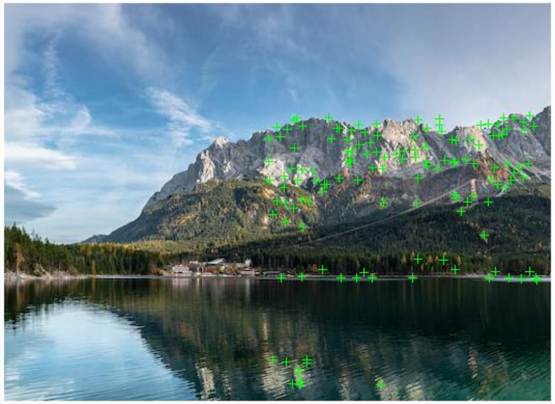
SURF (Speeded-Up Robust Features)

SURF es un algoritmo capaz de obtener una representación visual de una imagen y extraer información detallada y específica del contenido.

Esta información es tratada para realizar operaciones como por ejemplo la localización y reconocimiento de objetos, personas o caras, realización de escenas 3D, seguimiento de objetos y extracción de puntos de interés.

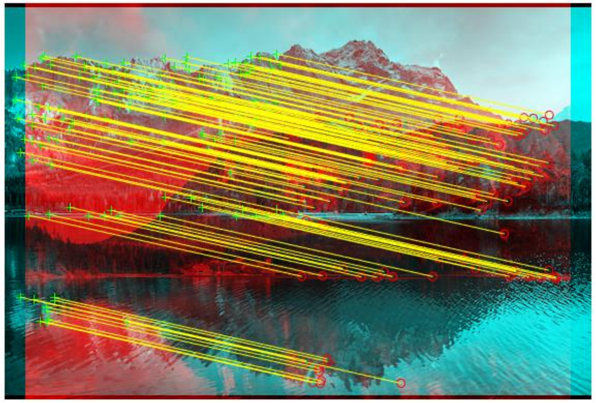
SURF (Speeded-Up Robust Features)

Se suele utilizar SURF para extraer características (puntos de interés) de dos imágenes.



Características imagen 1

Características imagen 2

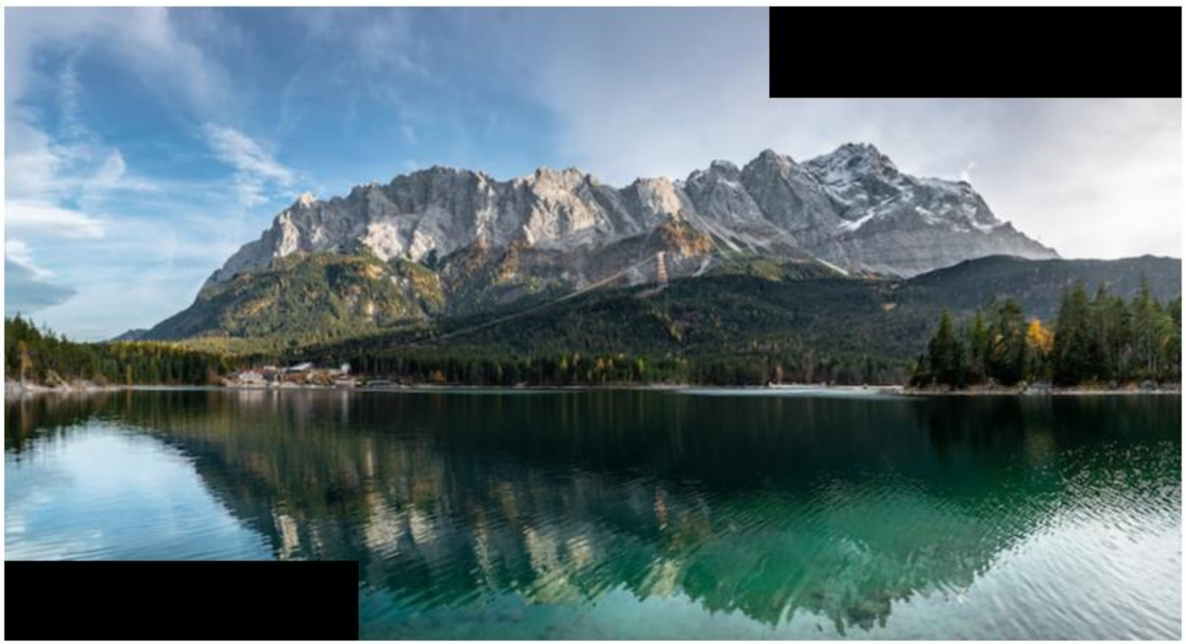


Similitud entre características

Luego, mediante el uso de descriptores de características, se pueden relacionar las características de imagen con la otra.

SURF (Speeded-Up Robust Features)

La relación de las características de imagen con la otra, se puede resolver como un problema de optimización utilizando algebra proyectiva.



Como resultado se puede construir una imagen panorámica.

SURF (Speeded-Up Robust Features)

Entre los algoritmos para extraer características en imágenes mas utilizados tenemos:

- SURF (Speeded Up Robust Features)
- SIFT (Scale Invariant Feature Transform)
- ORB (Oriented FAST and Rotated BRIEF)
- FAST (Maximally Stable Extremal Regions)
- AGAST (Adaptive and Generic Accelerated Segment Test)

Estos algoritmos se pueden encontrar en OpenCV o en paquetes de procesamiento digital de imágenes de Matlab.

Transformaciones del Algebra proyectiva 2D

Entre las transformaciones del algebra proyectiva 2D tenemos:

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

Proyectiva

$$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

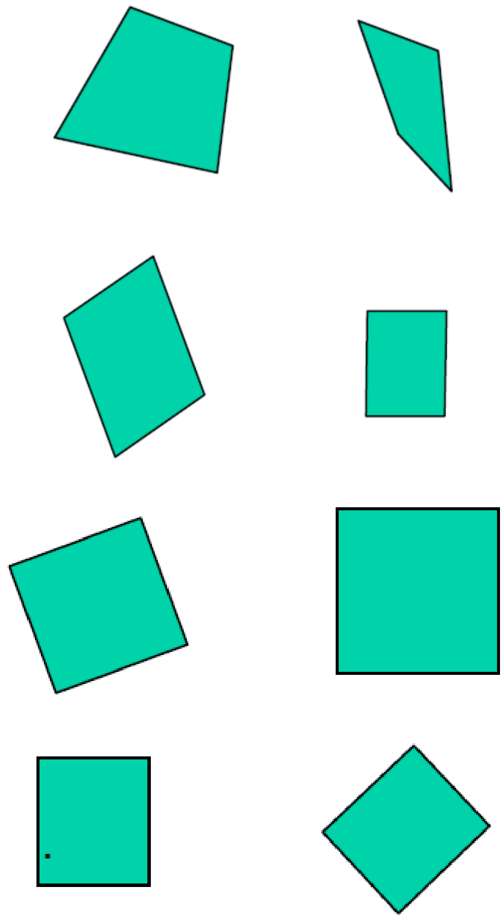
Afín

$$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Similitud

$$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

Euclidiana



- t_x y t_x son translaciones
- s es un escalamiento
- r_{ij} elementos de una matriz de rotación tal que

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}$$

Euclidiana ⊂ Similitud ⊂ Afín ⊂ Proyectiva

Operaciones del Algebra proyectiva 2D

Podemos proyectar una coordenada (x, y) utilizando la siguiente operación

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{T} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

donde $\mathbf{T}$ es alguna transformación proyectiva 2D y (X, Y, Z) son las coordenadas proyectadas.

Para proyectar la coordenada (X, Y, Z) en el plano donde $z = 1$, utilizamos

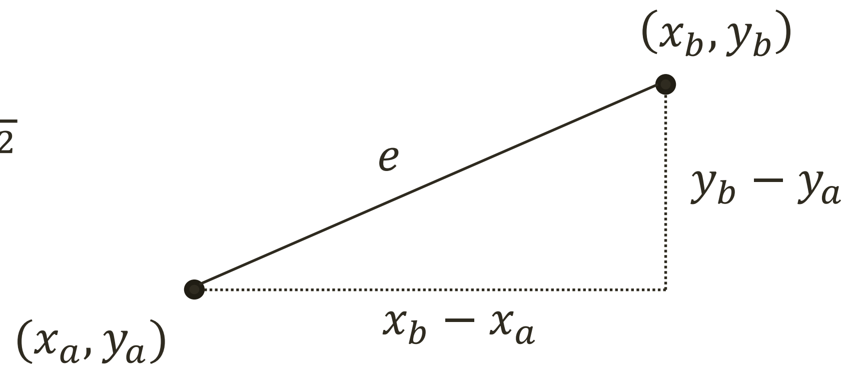
$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} X/_{Z} \\ Y/_{Z} \end{bmatrix}$$

donde (x', y') son las coordenadas transformadas.

Distancia Euclidiana

La distancia Euclidiana expresa la distancia entre una coordenada (x_a, y_a) y otra coordenada (x_b, y_b) utilizando:

$$e = \sqrt{(x_b - x_a)^2 + (y_b - y_a)^2}$$



donde e es una medida de error que indica la distancia entre los dos puntos.

Nota: Si una coordenada es igual al otra, entonces $e = 0$.

Función objetivo

Podemos proponer una función objetivo f con base en errores $(e_1, e_2, \cdots, e_n)$ como sigue:

Error Absoluto Medio

$$f = \frac{1}{n}\sum_{i=1}^{n}|e_i|$$

Raíz del Error Cuadrático Medio

$$f = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i)^2}$$

Error Cuadrático Medio

$$f = \frac{1}{n}\sum_{i=1}^{n}(e_i)^2$$

Planteamiento del problema

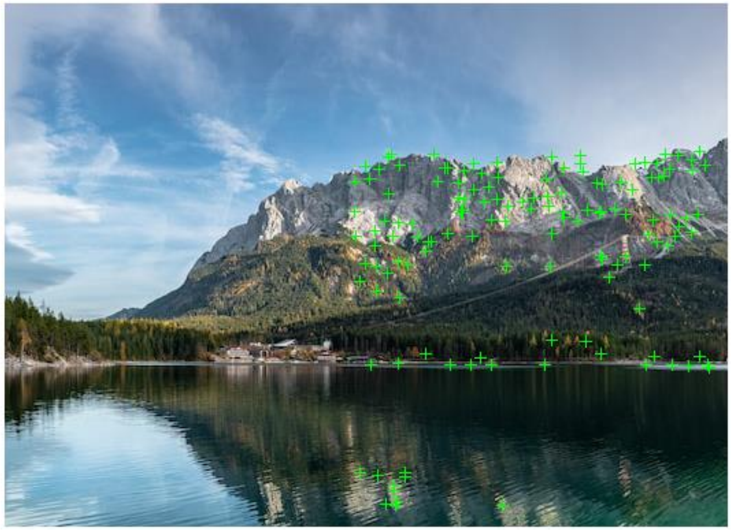
Objetivo: Crear alguna imagen panorámica utilizando el algebra proyectiva y la técnica SURF para la comparación de imágenes.

Metodología:

- Seleccionar dos imágenes (imagen A, imagen B) para crear una imagen panorámica.
- Utilizar la técnica SURF para comparar las imágenes y extraer los puntos de interés (pixeles) en común.
- Optimizar la transformación proyectiva que relaciona los puntos de interés de la imagen B con los puntos de interés de la imagen A.
- Crear la imagen panorámica utilizando la transformación proyectiva óptima.

Planteamiento del problema (continuación)

Después de seleccionar dos imágenes (A y B), aplicamos la técnica SURF para identificar los puntos de interés.



Puntos de interés – Imagen A

Puntos de interés – Imagen B

$$p_i^A = \begin{bmatrix} x_i^A \\ y_i^A \end{bmatrix}$$

Definimos m como el total de puntos.

$$i = 1,2,3,\cdots,M$$

$$p_i^B = \begin{bmatrix} x_i^B \\ y_i^B \end{bmatrix}$$

Planteamiento del problema (continuación)

Utilizamos una transformación proyectiva para relacionar los puntos de interés de la imagen B con los puntos de interés de la imagen A.

Proyectamos cada punto de interés $\left(x_i^B, y_i^B\right)$ utilizando

$$\begin{bmatrix} X_i^B \\ Y_i^B \\ Z_i^B \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x_i^B \\ y_i^B \\ 1 \end{bmatrix}$$

$$\begin{aligned} X_i^B &= x_i^B h_{11} + y_i^B h_{12} + h_{13} \\ Y_i^B &= x_i^B h_{21} + y_i^B h_{22} + h_{23} \\ Z_i^B &= x_i^B h_{31} + y_i^B h_{32} + h_{33} \end{aligned}$$

O bien,

$$\begin{bmatrix} X_i^B & Y_i^B & Z_i^B \end{bmatrix} = \begin{bmatrix} x_i^B & y_i^B & 1 \end{bmatrix} \begin{bmatrix} h_{11} & h_{21} & h_{31} \\ h_{12} & h_{22} & h_{32} \\ h_{13} & h_{23} & h_{33} \end{bmatrix}$$

Planteamiento del problema (continuación)

Si contamos con todos los puntos de interés de la imagen B en un arreglo matricial, entonces una operación mas sencilla seria

$$\begin{bmatrix} X_1^B & Y_1^B & Z_1^B \\ X_2^B & Y_2^B & Z_2^B \\ & \vdots & \\ X_M^B & Y_M^B & Z_M^B \end{bmatrix} = \begin{bmatrix} x_1^B & y_1^B & 1 \\ x_2^B & y_2^B & 1 \\ & \vdots & \\ x_M^B & y_M^B & 1 \end{bmatrix} \begin{bmatrix} h_{11} & h_{21} & h_{31} \\ h_{12} & h_{22} & h_{32} \\ h_{13} & h_{23} & h_{33} \end{bmatrix}$$

Después, transformamos cada coordenada $\left(X_i^B, Y_i^B, Z_i^B\right)$ utilizando

$$\begin{bmatrix} x_i^{B'} \\ y_i^{B'} \end{bmatrix} = \begin{bmatrix} X_i^B / Z_i^B \\ Y_i^B / Z_i^B \end{bmatrix}$$

Planteamiento del problema (continuación)

Para relacionar cada punto de interés $\left(x_i^B, y_i^B\right)$ de la imagen B con el punto de interés $\left(x_i^A, y_i^A\right)$ de la imagen A, definimos el siguiente error

$$e_i = \sqrt{\left(x_i^A - x_i^{B'}\right)^2 + \left(y_i^A - y_i^{B'}\right)^2}$$

Finalmente, utilizamos la siguiente función objetivo para optimizar la transformación proyectiva

$$f = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(e_i)^2}$$

Planteamiento del problema (continuación)

Debido a que se quiere optimizar la transformación proyectiva

$$\mathbf{T} = \begin{bmatrix} h_{11} & h_{21} & h_{31} \\ h_{12} & h_{22} & h_{32} \\ h_{13} & h_{23} & h_{33} \end{bmatrix}$$

La dimensión del problema de optimización es $D = 9$, entonces los parámetros de ajuste se pueden considerar como

$$\mathbf{x}^T = [h_{11} \quad h_{12} \quad h_{13} \quad h_{21} \quad h_{22} \quad h_{23} \quad h_{31} \quad h_{32} \quad h_{33}]$$

Planteamiento del problema (continuación)

Además, podemos agregar un termino de regularización a la función objetivo, entonces tenemos

$$f = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(e_i)^2 + \lambda\left(\frac{1}{D}\right)\sum_{j=1}^{D}\left(x_j\right)^2}$$

donde λ escala la contribución de la regularización. La regularización evita sobre ajustar los parámetros $\mathbf{x}$.