Single View Metrology

Van-Thinh Vo

June 2025

Mục lục

Nếu có bất kỳ sai sót nào,mọi người có thể phản hồi qua email thinh.vovan@hcmut.edu.vn để hỗ trợ mình với nha :v

1 Giới thiệu

Ta đã phân tích và tìm hiểu cách để hiệu chỉnh camera từ ảnh. Tiếp tục, ta sẽ tìm cách xây dựng lại cấu trúc 3D của vật thể chỉ qua một ảnh, cùng tìm hiểu nhé. Trước hết ta sẽ tìm hiểu về một số phương pháp biến đổi (transformation) trong không gian 2D và 3D:

- **Isometric Transformation**: giữ lại tính chất điểm, đường thẳng, độ dài, góc, song song,... Phép biến đổi này chỉ bao gồm xoay, tịnh tiến và lật hình. **Rigid Transformation** là một dạng đặc biệt của IT vì không có phép lật hình.

$$\mathbf{M} = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix}$$

- **Similarity Transformation**: ngoài xoay và tịnh tiến còn có scale (bằng nhau theo mọi chiều), vì thế sẽ làm thay đổi về kích thước nhưng vẫn giữ về tỉ lệ.

$$\mathbf{M} = \begin{bmatrix} RS & t \\ 0 & 1 \end{bmatrix}, \mathbf{S} = \begin{bmatrix} s & 0 \\ 0 & s \end{bmatrix}$$

- **Affine Transformation**: chỉ giữ các đặc trưng về điểm, đường thẳng, song song.

$$\mathbf{M} = \begin{bmatrix} RS & t \\ 0 & 1 \end{bmatrix}, \mathbf{S} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}$$

- **Projective Transformation**: chỉ còn tính chất về điểm và đường thẳng.

$$\mathbf{M} = \begin{bmatrix} A & t \\ v & b \end{bmatrix}$$

2 Điểm và đường thẳng ở vô cực

Xét trong không gian 2D, ta định nghĩa một đường thẳng $l = \begin{bmatrix} a & b & c \end{bmatrix}^T$ với a, b, c là ba hệ số của đường thẳng. Vậy một điểm bất kì x, ta có $x \in l \Leftrightarrow x^T l = 0$. Gọi đường thẳng $l' = \begin{bmatrix} a' & b' & c' \end{bmatrix}^T$. Ta định nghĩa được giao của hai đường thẳng $x = l \times l'$ (với $\times$ là tích có hướng, x sẽ vuông góc với l và l', suy ra dpcm).

Trước giờ chúng ta cho rằng không tồn tại giao điểm của hai đường thẳng song song, nhưng ở trường hợp này, ta có thể mô hình một điểm là giao của hai đường thẳng song song trong không gian đồng nhất. Giả sử $l \parallel l' \Rightarrow \frac{a}{a'} = \frac{b}{b'}$. Ta có

$$x_\infty = l \times l' \propto \begin{bmatrix} b \\ -a \\ 0 \end{bmatrix}$$

. Theo định nghĩa của hệ tọa độ đồng nhất, nếu chia các tọa độ cho tọa độ cuối (tức là chia cho 0), ta sẽ có một điểm có tọa độ (∞, ∞)! Hợp lý với một điểm ở vô cùng.

Các đường thẳng song song với nhau sẽ cùng giao tạo một điểm ở vô cực x_∞ và tập hợp các điểm vô cực tạo thành một đường thẳng ở vô cực $l_\infty = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$

Phép Affine Transformation sẽ chuyển điểm và đường thẳng vô cực vẫn ở dạng vô cực, trong khi Projective Transformation thì không. Cho ma trận H của phép PT, ta có

$$x' = Hx$$

$$x^T I l = 0 \Rightarrow x^T H^T H^{-T} l = 0$$

Ta biết rằng PT vẫn giữ tính thẳng hàng, do đó $(x')^T l' = 0$, nên $l' = H^{-T} l$.

3 Điểm và đường thẳng tiêu biến (Vanishing)

Xét trong không gian 3D, điểm trên màn ảnh sau khi thực hiện phép chiếu là giao nhau của hai đường thẳng song song trong 3D gọi là *điểm tiêu biến* (**vanishing point**). Với K, R, T là tham số của camera và $d = \begin{bmatrix} a & b & c \end{bmatrix}^T$ là hướng của đường thẳng trong không gian 3D, ta có tọa độ điểm tiêu biến v. Ta nói vanishing point chỉ phụ thuộc vào hướng của đường thẳng trong 3D mà không quan tâm vị trí của nó.

$$v = Kd$$

Ta có thể suy ra từ công thức trên:

$$d = \frac{K^{-1}v}{\|K^{-1}v\|}$$

Một mặt phẳng bất kì sẽ giao với mặt phẳng vô cực $\Pi_\infty = [0, 0, 0, 1]^T$ tại l_∞. l_{horiz} hay *vanishing line* là hình chiếu của l_∞ lên màn ảnh:

$$l_{horiz} = H^{-T} l_\infty$$

Ta còn có thể tính được vector pháp tuyển của mặt phẳng khi biết K và l_{horiz}. Điều đó cho thấy ta có thể xây dựng lại hình dạng 3D với một camera biết đã hiệu chỉnh.

$$n = K^T l_{horiz}$$

Một số khái niệm liên quan:

- *Mặt phẳng ở vô cực* (**Infinity plane**): Π_∞ là mặt phẳng vô cực nếu nó chứa tất cả các đường thẳng ở vô cực. Ta có $\Pi_\infty = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^T$ trong không gian đồng nhất.
- Ta có thể tính góc giữa hai đường thẳng trong không gian 3D thông qua điểm tiêu biến của chúng:

$$\begin{aligned} \cos\theta &= \frac{d_1 \cdot d_2}{\|d_1\|\|d_2\|} \\ &= \frac{v_1^T \omega v_2}{\sqrt{v_1^T \omega v_1}\sqrt{v_2^T \omega v_2}} \end{aligned} \tag{11}$$

  với $\omega = (KK^T)^{-1}$.
- Ta cũng có thể tính góc giữa hai mặt phẳng thông qua vector pháp tuyến của chúng.

$$\begin{aligned} \cos\theta &= \frac{n_1 \cdot n_2}{\|n_1\|\|n_2\|} \\ &= \frac{l_1^T \omega^{-1} v_2}{\sqrt{l_1^T \omega^{-1} l_1}\sqrt{l_2^T \omega^{-1} l_2}} \end{aligned} \tag{12}$$

4 Single view metrology

Ta có thể xây dựng được tham số camera K chỉ với một bức ảnh được hay không? Câu trả lời là có. K có 6 tham số tự do , vì thế, ta có thể chọn 3 điểm tiêu biến nằm trong 3 mặt phẳng vuông góc đôi một, theo [11], ta sẽ có được 3 phương trình. Tuy nhiên như thế là chưa đủ, ta cần giả sử rằng máy ảnh skewness và pixel hình vuông và scale để giảm số tham số của ω còn 3. Sau khi tính được ω, ta có thể tìm được K.

Tài liệu