

**ĐẠI HỌC QUỐC GIA HÀ NỘI**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**



**Object-based camera calibration via direct linear transformation**

**Sinh viên: Nguyễn Công Anh Tuấn – 21020566**  
**Khóa: QH-2021-I/CQ-R**

**HỌC PHẦN: XỬ LÝ ẢNH VÀ THỊ GIÁC ROBOT**  
**Mã học phần: RBE3015**  
**Ngành: Kỹ thuật Robot**

Hà Nội, 2023

## Contents

1. Giới thiệu bài toán. ....	3
2. Một số nghiên cứu liên quan. ....	3
3. Mô phỏng hệ thống và thuật toán. ....	3
a) Mô phỏng hệ thống:.....	3
b) Thuật toán:.....	4
4. Kết quả và đánh giá. ....	5
5. Kết luận.....	5
6. Tham khảo .....	6

## 1. Giới thiệu bài toán.

Camera calibration [1] là bài toán ước lượng tham số của một mô hình máy ảnh lỗ kim (pinhole camera model). Từ những tham số này, ta có thể xây dựng mối quan hệ giữa toán học giữa tọa độ một điểm ở không gian ba chiều và hình chiếu của nó lên mặt phẳng ảnh.

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Trong đó:

$u, v$  là tọa độ trong ảnh.

$x, y, z$  tọa độ thực.

$\mathbf{M}$  là ma trận hình chiếu, hay còn gọi là ma trận camera.

$\mathbf{M}$  có thể được tách ra làm tích của ba ma trận:

$$\mathbf{M} = \mathbf{K} * [\mathbf{R} \ \mathbf{T}]$$

Trong đó:

$$\mathbf{K} \text{ là ma trận intrinsic, } \mathbf{K} = \begin{bmatrix} \alpha_x & \gamma & u_0 & 0 \\ 0 & \alpha_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{3 \times 4}$$

$$[\mathbf{R} \ \mathbf{T}] \text{ là ma trận extrinsic, } [\mathbf{R} \ \mathbf{T}] = \begin{bmatrix} \mathbf{R}_{3 \times 3} & \mathbf{T}_{3 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix}_{4 \times 4}$$

$\mathbf{R}$  và  $\mathbf{T}$  là ma trận xoay và tịnh tiến để chuyển từ hệ tọa độ thế giới sang hệ tọa độ theo camera.

## 2. Một số nghiên cứu liên quan.

Camera calibration là một quá trình quan trọng trong thị giác máy tính 3D, cho phép trích xuất các thông tin về độ đo từ ảnh 2D. Một số phương pháp đã được đề xuất để giải bài toán này như phương pháp Zhang [2], hay phương pháp plane based [3], calibration from pure rotation [4],... Hầu như các phương pháp đều yêu cầu ảnh chụp mặt phẳng caro ở các góc độ khác nhau hoặc thay đổi góc chụp của camera để thực hiện quá trình calibrate.

Phương pháp trong bài viết này chỉ yêu cầu thay đổi vị trí vật mẫu và ghi chép tọa độ thực, từ đó sử dụng phương pháp DLT [1] để đưa ra ma trận camera.

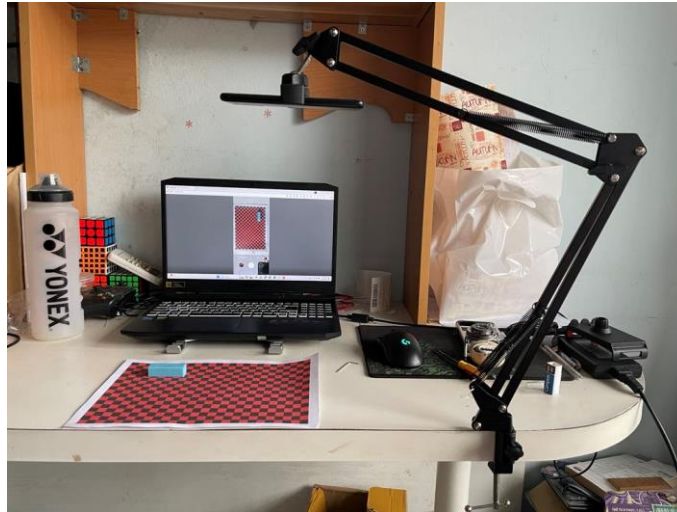
## 3. Mô phỏng hệ thống và thuật toán.

a) Mô phỏng hệ thống:

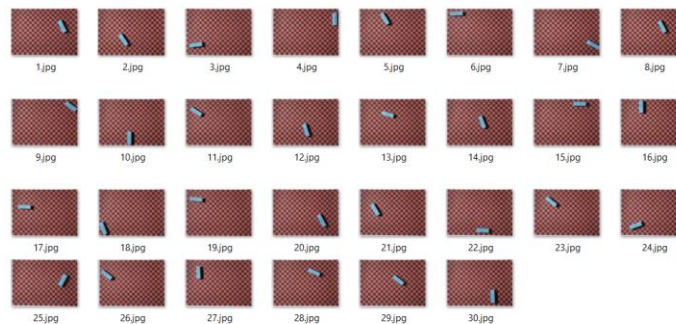
Hệ thống gồm:

- Một mặt phẳng được chia thành các ô vuông, mỗi ô có kích thước 1.05x1.05(cm)(góc tọa độ là góc trái trên).

- Một camera cách mặt phẳng 50cm lên phía trên và chiếu thẳng xuống mặt phẳng.
- Vật mẫu: khối hộp hình chữ nhật đặt nằm ngang, màu sắc có độ tương phản cao với mặt phẳng.
- Ảnh chụp ra có kích thước là: 2052x1431.
- Hệ thống cho ra dataset với 30 ảnh ở cùng một góc chụp với vật mẫu ở các vị trí khác nhau.



Hình 1: Hệ thống.



Hình 2: Dataset.

Quá trình lấy tọa độ thực của tâm vật:

- Để tăng độ chính xác, ta sử dụng một cây kim xuyên qua tâm của vật mẫu, sau đó đặt đầu cây kim trùng với tọa độ mong muốn (sử dụng phân phối đều để sinh ra các cặp tọa độ) rồi rút cây kim ra khỏi vật mẫu rồi tiến hành quá trình chụp ảnh và ghi chép tọa độ đã đặt.

b) Thuật toán:

Xác định tọa độ tâm của vật trong ảnh:

- Sử dụng histogram vào ảnh mẫu vật để xác định vector màu của vật theo hệ RGB.
- Dùng hệ màu hsv, lọc ra vật mẫu trong ảnh để tạo ra mask.
- Sau khi có mask, ta dùng thuật toán pca hoặc dùng moment lên ảnh mask để xác định tâm của vật mẫu trong ảnh (hai phương pháp đều cho ra kết quả tương tự nhau).

Sau khi có được các ground truth, ta tiến hành giải phương trình  $\mathbf{U} = \mathbf{M} * \mathbf{X}$ :

$$\mathbf{U} = [\vec{u}_1 \quad \vec{u}_2 \quad \dots \quad \vec{u}_n]$$

$$\mathbf{X} = [\vec{x}_1 \quad \vec{x}_2 \quad \dots \quad \vec{x}_n]$$

Với  $\vec{u}_i = [u_i \ v_i \ 1]^T$ ,  $\vec{x}_i = [x_i \ y_i \ z_i \ 1]^T$  lần lượt là tọa độ pixel tâm của vật và tọa độ thực của tâm vật trong ảnh thứ  $i$  và  $n$  là số điểm lấy mẫu để giải phương trình. Ở đây, coi  $z_i = 1.5$  do tâm của vật nằm cách mặt phẳng 1.5cm theo hướng lên trên. Do phương trình trên là phi tuyến, ta sử dụng phương pháp direct linear transformation để giải phương trình này.

$$\mathbf{U} * \mathbf{X}^T = \mathbf{M} * (\mathbf{X} * \mathbf{X}^T)$$

$$\Leftrightarrow \mathbf{M} = \mathbf{U} * \mathbf{X}^T * (\mathbf{X} * \mathbf{X}^T)^{-1}$$

Để thuận tiện cho việc xác định tọa độ thực của tâm vật khi biết tọa độ của nó trong ảnh, ta giải phương trình  $\mathbf{X} = \mathbf{M}' * \mathbf{U}$ . Khi đó,  $\mathbf{M}' = \mathbf{X} * \mathbf{U}^T * (\mathbf{U} * \mathbf{U}^T)^{-1}$ . Việc lấy mẫu điểm phù hợp được thực hiện bằng cách duyệt tuần tự không gian khả năng bao gồm  $\binom{\text{total\_sample}}{n}$  khả năng. Thực hiện duyệt với một số vòng lặp nhất định, cho đến khi nhận được ma trận camera cho ra sai số bé nhất. Số vòng lặp tương đối nhỏ vì đa số khả năng đều cho sai số với độ chênh lệch không quá 0.02.

## 4. Kết quả và đánh giá.

Sai số (cm) được tính trên toàn bộ unseen-datas với độ đo là sai số bình phương trung bình gốc(RMSE):

n	2	3	4	6	7	9
Sai số	9.2	1.09	0.4	0.3	0.3	0.3

Bảng 1: Sai số trên tập unseen-data.

Sai số có thể được giảm thiểu qua việc chuẩn hóa ma trận  $\mathbf{X}$  và  $\mathbf{U}$ :

n	2	3	4	6	7	9
Sai số	8.1	0.4	0.2	0.2	0.2	0.2

Bảng 2: Sai số trên tập unseen-data sau khi đã chuẩn hóa  $\mathbf{U}$  và  $\mathbf{X}$ .

Sai số còn phụ thuộc vào những điểm ta lấy mẫu, việc chuẩn hóa làm cho không gian khả năng này bé đi, do đó, cần nhiều thời gian hơn khi tìm kiếm những điểm phù hợp để cho ra được sai số tốt nhất.

## 5. Kết luận.

Sử dụng phương pháp direct linear transformation, ta đã thành công giải được bài toán camera callibration. Việc ước lượng tọa độ thực tâm vật khi đã biết tọa độ tâm vật trong ảnh cũng như ma trận camera cho sai số rất bé. Với phương pháp này, ta không cần ảnh của mặt phẳng ở các góc độ khác nhau, chỉ cần 3 – 4 ảnh mà vật mẫu ở các vị trí khác nhau. Trong tương lai sẽ đưa ra một phương pháp để tìm kiếm không gian nghiệm để đưa ra được ma trận camera cho ta thông tin 3D chính xác nhất.

## 6. Tham khảo

- [1] Wikipedia, "Camera resection," 10 11 2023. [Online]. Available:  
[https://en.wikipedia.org/wiki/Camera\\_resectioning](https://en.wikipedia.org/wiki/Camera_resectioning).
- [2] Z. Zhang, "A Flexible New Technique for Camera Calibration," p. 3, December 2, 1998.
- [3] P. F. Sturm and S. J. Maybank, "On Plane-Based Camera Calibration:," 1999.IEEE.
- [4] R. Hartley, " Self-calibration from multiple views with a rotating camera," p. 471–478, May 1994. Springer-Verlag..