Spring 2015,CIS II Project #4

Paper Seminar Presentation

Han Xiao
hxiao9@jhu.edu

Mentors
Dr. Nassir Navab (nassir.navab@jhu.edu)
Bernhard Fuerst (be.fuerst@jhu.edu)
Javad Fotouhi (fotouhi@jhu.edu)

1. Project Overview

Our goal is to integrate a depth sensor (Kinect sensor) into the CamC (Camera Augmented Mobile C-arm) system.

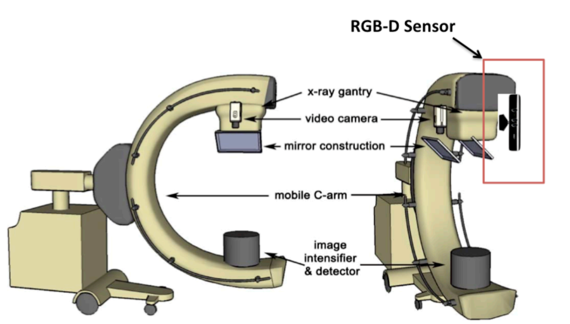


Figure 2. Illustration for Kinect mounting. (Navab, Nassir, IEEE Transactions 2010)

- **Hands and tools segmentation**
- **Spatial relationships determination**
- **Enhanced X-ray overlay without blocking**

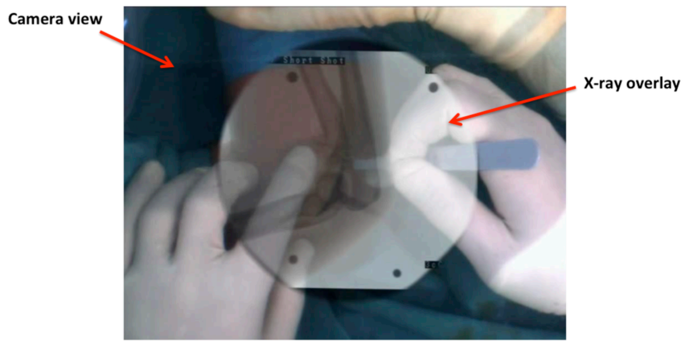


Figure 1. Overlay view of CamC. (Navab et al. IEEE TMI 2010)

2. Paper Selection

- Paper

Zhang, Zhengyou. "**A flexible new technique for camera calibration.**" Pattern Analysis and Machine Intelligence, IEEE Transactions on 22.11 (2000): 1330-1334.

- Why it is important

Multiple cameras calibration and image registration are keys for our project. The paper provide a flexible, robust and low cost camera calibration technique.

3. Problems and Key Results

- **Photogrammetric calibration**

Camera calibration is performed by observing a calibration object whose geometry in 3-D space is known with very good precision.
Problem: These approaches require an expensive calibration apparatus, and an elaborate setup.

- **Self-calibration**

By moving a camera in a static scene, without a calibration object. Using image information only.
Problem: This is not yet mature. There are many parameters to estimate, cannot always obtain reliable results.

- **A flexible new calibration (By zhang)**

Results: Provide a flexible, robust and low cost camera calibration technique.

4. Background knowledge

A 2D point $\mathbf{m} = [u, v]^T$

A 3D point $\mathtt{M} = [X, Y, Z]^T$

For convenience $\widetilde{\mathbf{m}} = [u, v, 1]^T$ $\widetilde{\mathtt{M}} = [X, Y, Z, 1]^T$.

the relationship between a 3D point M and its image projection m is given by

$$s\widetilde{\mathbf{m}} = \mathbf{A}\begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}\widetilde{\mathtt{M}}$$

$$\mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Assume the model plane is on Z = 0 of the world coordinate system.
Let's denote the i-th column of the rotation matrix R by ri:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} .$$

Therefore, a model point M and its image m is related by a homography H (Up to a scalar):

$$s\widetilde{\mathbf{m}} = \mathbf{H}\widetilde{\mathtt{M}} \quad \text{with} \quad \mathbf{H} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$$

campar.in.tum.de + camp.lcsr.jhu.edu

Set the homography $\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix}$

Then $\begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix} = \lambda \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix}$

r1 and r2 are orthonormal for a rotation matrix

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0$$
$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2$$

Homography: 8 degrees of freedom
extrinsic parameters: 6 DOF (3 for rotation and 3 for translation)
We get 2 constraints on the intrinsic parameters.

The paper also discuss a geometric interpretation of these two constrain using abstract concepts like the image of the absolute conic and circular points.

campar.in.tum.de + camp.lcsr.jhu.edu

5. Solving Calibration (Closed-Form)

$$\mathbf{B} = \mathbf{A}^{-T}\mathbf{A}^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2\beta} & \frac{v_0\gamma - u_0\beta}{\alpha^2\beta} \\ -\frac{\gamma}{\alpha^2\beta} & \frac{\gamma^2}{\alpha^2\beta^2}+\frac{1}{\beta^2} & -\frac{\gamma(v_0\gamma-u_0\beta)}{\alpha^2\beta^2}-\frac{v_0}{\beta^2} \\ \frac{v_0\gamma-u_0\beta}{\alpha^2\beta} & -\frac{\gamma(v_0\gamma-u_0\beta)}{\alpha^2\beta^2}-\frac{v_0}{\beta^2} & \frac{(v_0\gamma-u_0\beta)^2}{\alpha^2\beta^2}+\frac{v_0^2}{\beta^2}+1 \end{bmatrix}$$

Note that B is symmetric, defined by a 6D vector

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T$$

campar.in.tum.de + camp.lcsr.jhu.edu

We set $\mathbf{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$

Relate homography and B $\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b}$

Where

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T$$

Recall

$$\mathbf{h}_1^T \mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{h}_2 = 0$$
$$\mathbf{h}_1^T \mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T}\mathbf{A}^{-1}\mathbf{h}_2$$

We get the system

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0}$$

campar.in.tum.de + camp.lcsr.jhu.edu

If n images of the model plane are observed, by stacking n such equations shown on the previous slide, we have:

$$\mathbf{Vb} = \mathbf{0}$$

where V is a 2n × 6 matrix. If n ≥ 3, we will have in general a unique solution b defined up to a scale factor.

The solution to this euqation is well known as the eigenvector of VT V associated with the smallest eigenvalue (equivalently, the right singular vector of V associated with the smallest singular value).

Once b is estimated, we can compute the intrinsic and extrinsic parameter of the camera.

6. Maximum likelihood estimation

The closed form solution is obtained through minimizing an algebraic distance which is not physically meaningful. We can refine it through maximum likelihood estimation.

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\|\mathbf{m}_{ij} - \hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathtt{M}_j)\|^2$$

nonlinear minimization problem, which can be solved with the Levenberg-Marquardt Algorithm

7. Other Concerns

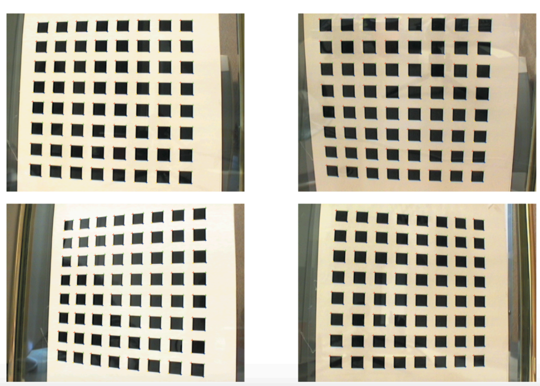
- The paper also discusses about radial distortion correction
- Propose a complete maximum likelihood estimation with radial distortion

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\|\mathbf{m}_{ij} - \breve{\mathbf{m}}(\mathbf{A}, k_1, k_2, \mathbf{R}_i, \mathbf{t}_i, \mathtt{M}_j)\|^2$$

- Special case where parallel model planes occur, and case with pure translation of model plane occurs, and a solution.

8. Experimental Results

- Tested with computer simulated camera
- Tested with real data
- Compared the closed form solution with the maximum likelihood estimation



- Application to image-based modeling

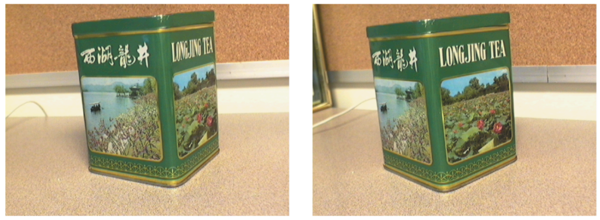
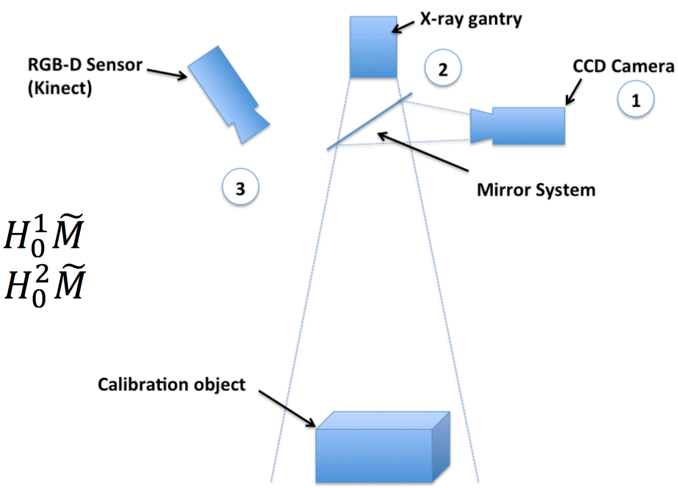


Figure 6: Two images of a tea tin

9. Importance to Our Project



$$\widetilde{m_1} = \lambda_1 A_1 H_0^1 \widetilde{M}$$
$$\widetilde{m_2} = \lambda_2 A_2 H_0^2 \widetilde{M}$$

Thanks for your attention!

campar.in.tum.de + camp.lcsr.jhu.edu