

# 一种灵活的从未知视角观测平面的相机标定方法

张正友

## 摘要

我们提出了一种灵活的新技术用来标定摄像机。这种方法只需要相机从很少的位置（至少两个）观测平面，相机和平面都可以以未知速度自由移动，透镜畸变被考虑到该模型中。本方法先给出了封闭解，然后再基于**最大似然估计准则**给出了理想的非线性解。计算机模拟数据和真实数据都用来测试这种方法，并取得了很理想的结果。常规方法需要昂贵的设备，比用两个或者三个正交平面标定的方法，而本方法灵活易用。它推动了3D 计算机视觉从实验室走向现实世界前进了一步。作者的主页上有本方法相应的软件。

**关键词：** 摄像机标定 内参 镜头畸变 灵活的平面标定 运动分析 模型获取

## 1 目的

为了从二维图像中提取几何信息，在三维计算机视觉中，摄像机标定是必要步骤。从摄影测量学[2, 4] 到最近的计算机视觉[9, 8, 20, 7, 23, 21, 15, 6, 19]，前人在这方面已经做过大量工作。这些技术大致可以分为如下几类：

**摄影测量法标定。**通过观测一个三维尺寸精确已知的物体来标定，这种方法[5]非常高效。标定的物体通常包含两个或者三个互相正交的平面。有时，也用一个精确已知变化的平面来标定。这些方法需要昂贵的标定设备和复杂的步骤。

**自标定。**这种方法不需要任何辅助标定的对象。只需要在静态场景里移动摄像机，只通过图像信息来完成标定，这种场景对摄像机的内参有两个约束[15]。因此，如果用相同的相机和固定的内部参数来获取图像，三张图片足够得到内参和外参，以此我们可以重建出相似的三维结构[14, 12]。虽然这种方法很灵活，但是还不成熟，因为有许多参数需要估计，我们无法得到可靠的结果。

**其他方法。**用正交方向的灭点[3, 13]和纯粹的旋转标定方法[10, 18]。

我们当前的研究集中在桌面视觉系统，因为随着摄像机变得越来越便宜和常见，桌面视觉系统具有极好的应用前景。桌面视觉系统主要针对那些不擅长计算机视觉的普通大众，典型的计算机用户只是偶尔用一下视觉功能，所以他们不会愿意为昂贵的设备投资。因此灵活性，鲁棒性，低成本很重要，而本文提出的相机标定方法考虑到了这些事项。

本方法只需要摄像机从不同方位（至少两个）来观测标定板。标定板的图案可以用激光打印机打印出来并贴到一个合适的平面上（比如一个硬的图书封面），摄像机或则标定板可以以未知速度移动。本文提出的方法介于摄影测量法标定和自标定方法之间，因为我们利用了二维几何信息而不是三维或者纯粹的特征信息。计算机模拟的数据和真实数据都被用来测试该方法，并且得到了非常理想的结果。和经典的摄影测量标定法相比，该方法非常灵活。和自标定方法相比，本方法又具有较强的**鲁棒性**。我们相信本方法使计算机视觉从实验室向现实世界的普遍应用迈进了一步。

Bill Trigger[19]最近提出了一种新的自标定方法，这种方法至少需要一个平面图案场景的五个视图。他的技术比我们的更灵活，但是初始化很困难。Liebowitz和Zisserman [13] 描述了一种对标定平面的透视图进行度量矫正的技术。这种方法利用度量信息，比如一个已知的角度或者两个相等

但是未知的角度，和已知的长度比率。虽然没有算法和实验结果，但是他们仍认为基于这些矫正过的平面，相机内参的标定是可能的。

本文结构如下。第二部分描述了观测一个标定平面的基本约束条件。第三部描述了标定步骤。我们以一个封闭解开始，然后得到非线性的最优化，透镜畸变也被考虑到该模型中。第四部分研究了这个方法的不足之处，实践中很容易避免这些地方。第五部分提供了实验结果，计算机模拟出的数据和现实数据都用来验证了该方法。附录中，我们给出了大量详细的说明，包括了怎样估计位图模型和图像单应性关系的内容。

## 2 基本方程

我们通过观测一个待标定的平面来得到相机内参的约束条件。本文以以下注释开始。

### 2.1 注释

二维点用  $\mathbf{m} = [u, v]^T$ ，三维点用  $M = [X, Y, Z]^T$  表示。 $\tilde{x}$  表示增广向量，最后一个元素加1:  $\tilde{m} = [u, v, 1]^T$ ， $\tilde{M} = [X, Y, 1]^T$ 。建立一个常见的小孔成像模型：三维点  $M$  和它图像投影点  $\mathbf{m}$  关系如下

$$s\mathbf{m} = \mathbf{A} [\mathbf{R} \quad \mathbf{t}] \tilde{M} \quad \text{with } \mathbf{A} = \begin{bmatrix} \alpha & c & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

式中  $s$  是任意的比例因子。 $(\mathbf{R}, \mathbf{t})$  称为外参， $\mathbf{R}$  是旋转矩阵， $\mathbf{t}$  是平移矩阵。 $\mathbf{A}$  是相机内参矩阵， $(u_0, v_0)$  是坐标的主点， $\alpha$  和  $\beta$  是图像在  $u$  和  $v$  轴的比例因子， $c$  是描述两个坐标轴倾斜角的参数。

我们简称  $\mathbf{A}^{-T}$  为  $(\mathbf{A}^{-1})^T$  或者  $(\mathbf{A}^T)^{-1}$ 。

### 2.2 模型平面与其图像间的单应性关系

一般情况下，我们假设模型平面在世界坐标系中的  $Z$  坐标为0。 $\mathbf{R}$  的第  $i$  列旋转矩阵位  $\mathbf{r}_i$ 。从公式(1)中可知：

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

我们仍然用  $M$  表示位于位图中的一点，当  $Z$  为0 后， $M$  可以表示为  $M = [X, Y]^T$ ，同样地  $\tilde{M} = [X, Y, 1]^T$ 。因此点  $M$  和它的在图像上映射点  $\mathbf{m}$  关系用单应矩阵  $\mathbf{H}$  联系

$$s\tilde{\mathbf{m}} = \mathbf{H}\tilde{M} \quad \text{with } \mathbf{H} = \mathbf{A} [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}] \quad (2)$$

很显然， $\mathbf{H}$  是一个  $3 \times 3$  的系数矩阵。

### 2.3 内参的约束条件

对标定平面的一张图片，它的单应矩阵可以估计出来(参见附录A)。令  $\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3]$ ，由(2)式可知

$$[\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3] = \lambda \mathbf{A} [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}]$$

式中 $\lambda$ 是任意的标量。由所学知识可知 $\mathbf{r}_1$ 和 $\mathbf{r}_1$ 是正交的，于是我们有

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0 \quad (3)$$

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 \quad (4)$$

对一个给定的单应性矩阵，对内参有两个基本的约束条件。因为这个矩阵有8个自由度和6个外参（3个是旋转矩阵的和3个是平移向量的），这个条件下，我们只能得到两个内参的约束条件。

### 3 相机标定的解决方法

该节详细描述了怎样有效地解决摄像机的标定问题。我们先给出一个封闭解，然后根据最大似然估计给出非线性的最优化解，最后再考虑透镜的径向畸变，给出解析解和非线性解。

#### 3.1 封闭解

令

$$\mathbf{B} = \mathbf{A}^{-T} \mathbf{A}^{-1} \equiv \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{c}{\alpha^2 \beta} & \frac{cv_0 - u_0 \beta}{\alpha^2 \beta} \\ -\frac{c}{\alpha^2 \beta} & \frac{c^2}{\alpha^2 \beta^2} + \frac{1}{\beta^2} & -\frac{c(cv_0 - u_0 \beta)}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} \\ \frac{cv_0 - u_0 \beta}{\alpha^2 \beta} & -\frac{c(cv_0 - u_0 \beta)}{\alpha^2 \beta^2} - \frac{v_0}{\beta^2} & \frac{(cv_0 - u_0 \beta)^2}{\alpha^2 \beta^2} + \frac{v_0^2}{\beta^2} + 1 \end{bmatrix} \quad (5)$$

注意 $\mathbf{B}$ 是对称的，定义一个六维向量

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \quad (6)$$

假设 $\mathbf{H}$ 的第 $i$ 列向量为 $\mathbf{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$ ，然后我们可以得到

$$\mathbf{h}_i^T = \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b} \quad (7)$$

式中 $\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]$  于是，两个基本约束(3)(4) 式可以写为齐次式

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \quad (8)$$

如果观测了 $n$ 张图片，迭代可以得到

$$\mathbf{V} \mathbf{b} = 0 \quad (9)$$

式中 $\mathbf{V}$ 是一个 $2n \times 6$ 矩阵。若 $n \geq 3$ ，通常会得到一个唯一解 $\mathbf{b}$ 。若 $n = 2$ ，令倾斜约束参数 $c = 0$ ，即 $[0, 1, 0, 0, 0, 0] \mathbf{b} = 0$ ，这作为额外的方程带入(9)式。(9)式的解是被熟知的与最小奇异解相关的特征向量 $\mathbf{V}^T \mathbf{V} \mathbf{p}$ 。

$\mathbf{b}$ 一旦估计出来，我们可以计算内参矩阵 $\mathbf{A}$ 。详情请参见附录B。

求出 $\mathbf{A}$ 后，相应的外部参数也可以被计算出来，从(2)式中，可以得到

$$\mathbf{r}_1 = \lambda \mathbf{A}^{-1} \mathbf{h}_1, \mathbf{r}_2 = \lambda \mathbf{A}^{-1} \mathbf{h}_2, \mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2, \mathbf{t} = \lambda \mathbf{A}^{-1} \mathbf{h}_3$$

式中 $\lambda = 1/\|\mathbf{A}^{-1} \mathbf{h}_1\| = 1/\|\mathbf{A}^{-1} \mathbf{h}_2\|$ 。当然，由于噪声数据，因此计算的矩阵 $\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3]$  一般不能满足一个旋转矩阵的属性。附录C从一般的 $3 \times 3$ 矩阵介绍了一种方法来估计最佳的旋转矩阵。

### 3.2 最大似然估计

上面的解决方案通过最小化代数距离获得，一般情况下这是没有物理意义的。我们可以通过最大似然估计理论来完善。我们得到模型平面的 $n$ 幅图片，模型平面上有 $m$ 点。假设图像上像素点的噪声服从独立的同一分布。最大似然估计可以从通过求以下函数的最小值得到

$$\sum_{i=1}^n \sum_{j=1}^m \|\mathbf{m}_{ij} - \hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, M_j)\|^2 \quad (10)$$

根据(2)式，上式中 $\hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, M_j)$ 是 $M_j$ 点在第 $i$ 幅图像上的投影。旋转矩阵 $\mathbf{R}$ 用三个参数的向量 $\mathbf{r}$ 表示， $\mathbf{r}$ 平行于旋转轴，而且大小等于旋转角度。 $\mathbf{R}$ 和 $\mathbf{r}$ 关系符合罗德里格斯公式[5]。求(10)式的最小值是一个非线性优化问题，可以通过在[18]中应用的Levenberg-Marquardt 算法解决。它需要矩阵 $\mathbf{A}$ 的一个初始的猜测值， $\{\mathbf{R}_i, \mathbf{t}_i | i = 1..n\}$ 可使用在上一小节中描述的方法得到。

### 3.3 径向畸变的处理

到现在为止，我们没有考虑过摄像机的镜头畸变。然而，桌面相机的镜头通常会有显著的畸变，尤其是径向畸变。在本节中，我们只考虑前两个方面的径向畸变。读者可以参考[17, 2, 4, 23]获取详细信息。根据[2, 20, 22]的报告，镜头的畸变以径向分量为重，尤其是第一项占主导地位。同时文献[20, 22]发现任何更复杂的模型不仅没有帮助，而且会导致数值不稳定。

设 $(u, v)$ 是理想（畸变可忽略）像素的图像坐标， $(\tilde{u}, \tilde{v})$ 对应的实际观测到的图像坐标。同样， $(x, y)$ 和 $(\tilde{x}, \tilde{y})$ 的理想（无畸变）和实时（含有畸变）情况下的归一化图像坐标。我们有[2, 22]

$$\begin{aligned} \tilde{x} &= x + x[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \\ \tilde{y} &= y + y[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \end{aligned}$$

其中， $k_1$ 和 $k_2$ 是径向畸变系数。径向畸变中心的主点是相同的。从 $\tilde{u} = u_0 + \alpha\tilde{x} + c\tilde{y}$ 和 $\tilde{v} = v_0 + \beta\tilde{y}$ 中可以得到

$$\tilde{u} = u + (u - u_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \quad (11)$$

$$\tilde{v} = v + (v - v_0)[k_1(x^2 + y^2) + k_2(x^2 + y^2)^2] \quad (12)$$

**交替径向畸变的估计** 径向畸变一般很小，简单的忽略径向畸变后，可以用3.2中的方法估计另外的五个参数。一种策略是估计其他参数后再估计 $k_1$ 和 $k_2$ 。然后从(11)和(12)中，我们对每幅图的每个点可以得到两个方程

$$\begin{bmatrix} (u - u_0)(x^2 + y^2) & (u - u_0)(x^2 + y^2)^2 \\ (v - v_0)(x^2 + y^2) & (v - v_0)(x^2 + y^2)^2 \end{bmatrix} \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \begin{bmatrix} \tilde{u} - u \\ \tilde{v} - v \end{bmatrix}$$

$n$ 幅图像中共有 $m$ 个点，迭代所有方程得到一个 $2mn$ 的方程组，或者以矩阵形式 $\mathbf{D}\mathbf{k} = \mathbf{d}$ ，其中 $\mathbf{k} = [k_1, k_2]^T$ ，由最小二乘法得

$$\mathbf{k} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{d} \quad (13)$$

$k_1$ 和 $k_2$ 估计出来后，可以通过解(10)式来重新估计其他参数。其中可以用式(11)(12)来代替 $\hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, M_j)$ 交替使用这两个步骤，直到收敛为止。

**完整的最大似然估计** 实验中，我们发现上述收敛过程比较慢。通过最小化下列函数来估计参数的完整集

$$\sum_{i=1}^n \sum_{j=1}^m \|\mathbf{m}_{ij} - \check{\mathbf{m}}(\mathbf{A}, k_1, k_2, \mathbf{R}_i, \mathbf{t}_i, M_j)\|^2 \quad (14)$$

根据方程(2)(11)(12)，其中 $\check{\mathbf{m}}(\mathbf{A}, k_1, k_2, \mathbf{R}_i, \mathbf{t}_i, M_j)$ 是点 $M_j$ 在图象 $i$ 中的投影。这是一个非线性最小化问题，可以通过在[16]中应用的Levenberg-Marquardt算法解决。旋转矩阵仍然还可用一个3.2节中所示的3 向量 $\mathbf{r}$  来表示。 $\mathbf{A}$ 初始值的选取和 $\{\mathbf{R}_i, \mathbf{t}_i | i = 1..n\}$ 可以采用3.1 或者3.2中描述的技术。 $k_1$ 和 $k_2$ 初始值的选取可采用上一段所述的方法，或着直接将它们设置为0。

### 3.4 小结

建议校准程序如下：

- 打印图案，并将它贴到一个平面；
- 移动摄像机或者平面，从不同角度拍下几张图片；
- 在图像中检测特征点；
- 利用3.1中所述方法，估计五个内参和全部的外参；
- 通过最小二乘法(13)估计径向畸变系数；
- 改进全部参数通过最小化(14)式

## 4 退化情形

本节我们研究不满足相机内参约束的退化情形。因为(3)(4)式是从旋转矩阵的性质推导出来的，如果 $\mathbf{R}_2$ 和 $\mathbf{R}_1$ 不相互独立，那么第二幅图对于摄像机标定来说不起作用。下面我们将考虑更复杂的情况。

**命题1** 如果在第二个位置的模型平面平行于第一个位置的模型平面，那么第二幅图片不提供额外的约束条件。

由于篇幅原因，证明过程就省略了，具体细节可以参见技术文档[24]。实际过程中，非常容易避免退化：只需要改变平面与相机拍照点的相对位置即可。

## 5 实验结果

本文提出的算法测试了计算机模拟数据和真实数据。封闭解是来分解一个 $2n \times 6$ 矩阵从而求得奇异解，其中 $n$ 是图片数。用Levenberg-Marquardt算法求非线性解需要3到5次迭代才收敛。

### 5.1 计算机模拟

模拟的摄像机具有如下性质： $\alpha = 1250, \beta = 900, c = 1.09083$ （相当于 $89.95^\circ$ ）， $u_0 = 255, v_0 = 255$ 。图像分辨率为 $512 \times 512$ 像素。标定板包含 $10 \times 14 = 140$ 个点，尺寸是 $18cm \times 25cm$ 。平面的旋转矩阵用三维向量 $\mathbf{r}$  来表示，平移矩阵用一个三维向量 $\mathbf{t}$ （米制单位）表示。

**关于噪声水平的性能** 实验中，我们用三个平面

$$\begin{aligned} \mathbf{r}_1 &= [20^\circ, 0, 0], \mathbf{t}_1 = [-9, -12.5, 500]^T, \\ \mathbf{r}_2 &= [0, 20^\circ, 0]^T, \mathbf{t}_2 = [-9, -12.5, 510]^T, \\ \mathbf{r}_3 &= \frac{1}{\sqrt{5}}[-30^\circ, -30^\circ, -15^\circ]^T, \mathbf{t}_3 = [-10.5, -12.5, 525]^T \end{aligned}$$

期望为0，标准差为 $\sigma$ 的高斯噪声加入投影图像。估计出的摄像机参数与真值相对比。我们测试了 $\alpha, \beta$ 的相对误差， $u_0, v_0$ 的绝对误差。加入的噪声水平分布在0.1 像素到1.5 像素之间，对每个噪声等级，做了100次独立试验，并得到平均结果。从Figure 1可以看到，误差随噪声水平呈线性增长。当 $\sigma = 0.5$ （大于实际噪声的标准差）时， $\alpha, \beta$ 的误差小于0.3%， $u_0, v_0$ 误差在1 像素附近。 $u_0$ 误差大于 $v_0$ 误差，主要原因是 $u$ 轴方向上的数据比 $v$ 方向上数据多。

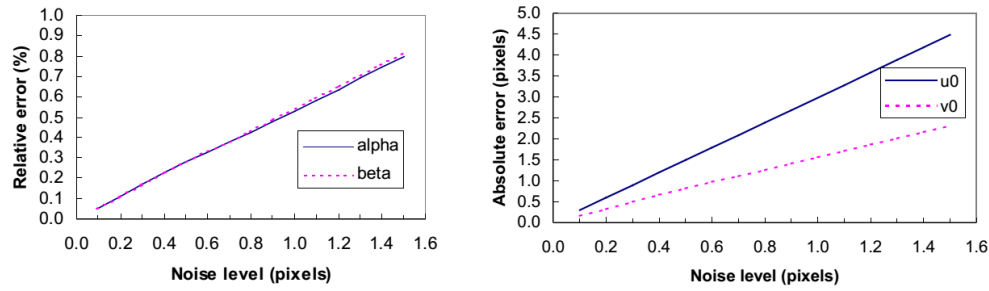


Figure 1: 不同噪声水平标定的误差对比

**不同平面数的性能测试** 该实验探索了不同模型平面数与性能的关系。前三幅图像的方向和位置和上一段的相同。第四幅图是在均匀的球体中任选一个旋转轴，然后再旋转 $30^\circ$ 得到。图片数目从2到16不等，对每个数目，做了100次独立实验，实验中噪声相互独立，并服从期望为0，标准差为0.5像素的高斯分布。Figure2显示了平均结果，图示可以看出，平面数从2到3，误差水平显著减小。

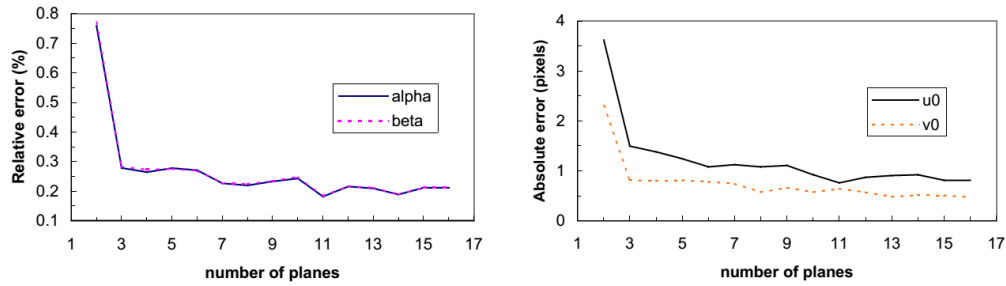


Figure 2: 不同数目平面标定误差对比

**模型平面方向的性能测试** 实验研究了位面方向对于标定结果的影响。实验用了三幅图片。平面方位选取如下：开始模型平面与图像平面平行，旋转轴从均匀球体中随机选取，然后绕轴旋转 $\theta$ 角，图像中加入期望为0，标准差为0.5 像素的高斯噪声。重复测试100次，计算平均误差。 $\theta$ 变化范



围为 $5^\circ$ 到 $75^\circ$ ，结果如Figure3所示。当 $\theta$ 为 $0.5^\circ$ 时，40%的实验因为平面之间几乎是相互平行而失败，实验结果中已经包含了这种情况。角度为 $45^\circ$ 左右时，效果似乎最佳。值得注意的是，实践中当角度增加时，透视缩减现象让角点检测变得困难，但是实验中并未考虑这点。

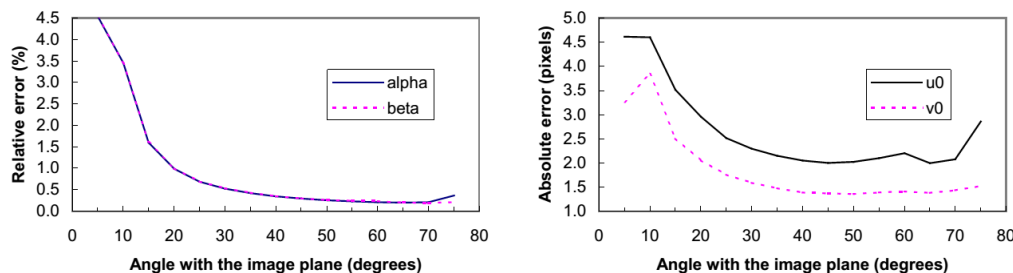


Figure 3: 不同角度标定的误差对比

## 5.2 真实数据测试

本文提出的方法已在我们的视觉组和图像组得到普遍应用。在这里，我们以一实际例子来说明标定过程。

标定PULNiX牌镜头为 $6mm$ 的CCD摄像机，图像分辨率为 $640 \times 480$ 。标定板含 $8 \times 8$ 的正方形，因此有256个角点。标定板尺寸为 $17cm \times 17cm$ ，得到五幅不同方位的图像，如Figure4所示，图中可以看到明显的畸变。

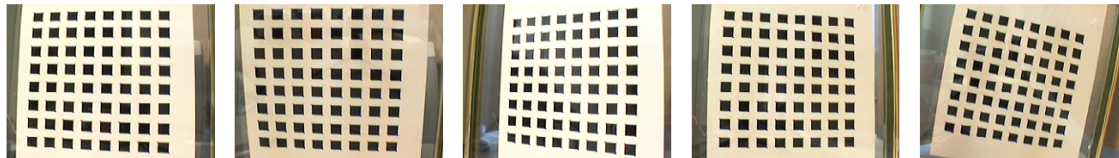


Figure 4: 标定平面的五幅图片以及提取的角点

实验中，分别采用不同数目即2,3,4,5幅图像来标定。结果如Table1所示，对每种情形，第一列是对封闭解的估计，第二列是最大似然估计结果，第三列是标准差的估计，以此表示结果的不确定性。

可以清楚的看到，封闭解比较合理，最终的最大似然估计值也具有很好的一致性。同时可以看出随着图片数目的增多，估计结果的不一致性降低了。Table1的最后一行是检测出的角点和投影点的均方根距离，单位是像素，可知最大似然估计方法使结果提高了很多。

细心的读者可能发现了封闭解和最大似然估计解中的 $k_1$ 和 $k_2$ 的一致性不好。原因在于解封闭解时，摄像机内参估计的前提是假设没有畸变，预测点更靠近图像中心而不是检测值，而随后对畸变的估计试着延伸外部点，并且增大系数以此减小距离，使畸变形状和真实畸变不一致。最大似然估计最终覆盖了畸变形态。估计出的畸变参数让我们校正原图像，Figure5展示了校正后的第1,2幅图像，可以看出方框之间的线条变直了

**标定结果的变化。**Table1中，我们展示了用2到5张图片标定的结果，而且我们发现结果的一致性非常好。为了进一步探究该算法的稳定性，我们从中任取四张图片，再运用该方法标定。结

Table 1: 用2幅图像到5幅的标定数据

nb	2 images			3 images			4 images			5 images		
	initial	final	$\sigma$	initial	final	$\sigma$	initial	final	$\sigma$	initial	final	$\sigma$
$\alpha$	825.59	830.47	4.74	917.65	830.80	2.06	876.62	831.81	1.56	877.16	832.50	1.41
$\beta$	825.26	830.24	4.85	920.53	830.69	2.10	876.22	831.82	1.55	876.80	832.53	1.38
$c$	0	0	0	2.2956	0.1676	0.109	0.0658	0.2867	0.095	0.1752	0.2045	0.078
$u_0$	295.79	307.03	1.37	277.09	305.77	1.45	301.31	304.53	0.86	301.04	303.96	0.71
$v_0$	217.69	206.55	0.93	223.36	206.42	1.00	220.06	206.79	0.78	220.41	206.56	0.66
$k_1$	0.161	-0.227	0.006	0.128	-0.229	0.006	0.145	-0.229	0.005	0.136	-0.228	0.003
$k_2$	-1.955	0.194	0.032	-1.986	0.196	0.034	-2.089	0.195	0.028	-2.042	0.190	0.025
<b>RMS</b>	0.761	0.295		0.987	0.393		0.927	0.361		0.881	0.335	

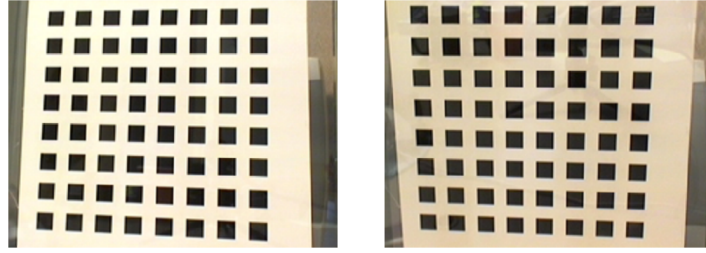


Figure 5: 标定畸变后的第一幅和第二幅图片

果如Table2 所示。例如第三列，实现了用第1,2,3,5幅图片标定的结果。最后两列代表均值和标准差。代表倾斜度的参数 $c$ 在0附近，偏差系数 $0.086/0.1401 = 0.6$ ，是一个相对较大的数。 $c = 0.1401$ ， $\alpha = 832.85$ 对应 $89.99^\circ$ ，非常接近 $90^\circ$ ，我们同样计算了比率 $\alpha/\beta$ ，平均值为0.99995，偏差为0.00012，非常接近1，即标定后图案是正方形的。

Table 2: 用四幅不同图片标定结果对比

quadruple	(1234)	(1235)	(1245)	(1345)	(2345)	mean	deviation
$\alpha$	831.81	832.09	837.53	829.69	833.14	832.85	2.90
$\beta$	831.82	832.10	837.53	829.91	833.11	832.90	2.84
$c$	0.2867	0.1069	0.0611	0.1363	0.1096	0.1401	0.086
$u_0$	304.53	304.32	304.57	303.95	303.53	304.18	0.44
$v_0$	206.79	206.23	207.30	207.16	206.33	206.76	0.48
$k_1$	-0.229	-0.228	-0.230	-0.227	0.229	0.229	0.001
$k_2$	0.195	0.191	0.193	0.179	0.190	0.190	0.006
<b>RMS</b>	0.361	0.357	0.262	0.358	0.334	0.334	0.04

**基于图像的建模。** 用上文中提到的摄像机拍摄一个茶叶罐的两幅图片Figure6所示。茶叶罐的两边是可见的，我们在每边上手动挑选8个点，然后用我们之前开发的运动建模软件得到该茶叶罐的部分模型。该重构模型是用虚拟现实标记语言实现的，三维渲染细节如Figure7所示，每边重构出的点确实共面，计算重构的两个平面的夹角为 $94.7^\circ$ 。虽然我们没有确切值，但是茶叶罐的两面确实几乎正交。



真实的数据、结果和软件可以从下列网页获得  
<http://research.microsoft.com/p~zhang/Calib/>

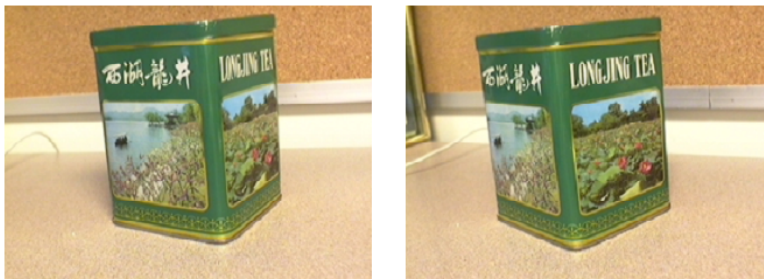


Figure 6: 茶叶罐的两幅图片



Figure 7: 茶叶罐重构后模拟的三面

### 5.3 结论

本文提出了一种新的灵活标定摄像机的方法。本方法只需要摄像机从不同方向（至少两个）观测标定板，可以任意移动摄像机或者标定板，镜头的畸变也已被模型考虑。本方法先给出封闭解，然后根据最大似然估计准则给出非线性优化结果，计算机模拟数据和真实数据都用来测试了该方法，并取得了极好的结果，和经典方法相比，该方法相当灵活。

**致谢:**感谢Brian Guenter的关于角点提取的软件,以及Bill Triggs具有洞察力的评论。感谢Andrew Zisserman的CVPR98成果[13],它使用相同的约束但不同的形式。要感谢Bill Triggs 和Gideon Stein给实验提出的宝贵意见[24]。感谢Anandan 和Charles Loop校对了本文英语。

## A 标定平面与其图像的单应性估计

有多种方法来估计标定平面与其图像的单应性关系，本文我们用最大似然估计来处理。令  $M_i$  和  $\mathbf{m}_i$  分别是标定平面和图像上的点，理想情形下它们符合(2)式，但实际上，因为图像上存在噪声，所以不符合上述假设。假设  $\mathbf{m}_i$  上加均值为0，协方差矩阵为  $\Lambda_{\mathbf{m}_i}$  的高斯噪声。通过求下列最小值，得到  $\mathbf{H}$

$$\sum_i (\mathbf{m}_i - \tilde{\mathbf{m}}_i)^T \Lambda_{\mathbf{m}_i}^T (\mathbf{m}_i - \tilde{\mathbf{m}}_i)$$

$$\text{式中 } \tilde{\mathbf{m}}_i = \frac{1}{\mathbf{h}_3^T M_i} \begin{bmatrix} \bar{\mathbf{h}}_1^T M_i \\ \bar{\mathbf{h}}_2^T M_i \end{bmatrix}, \bar{\mathbf{h}}_i \text{ 是 } \mathbf{H} \text{ 的第 } i \text{ 行}$$

在工程实际中，我们简单的假设对于所有的  $i$  有  $\Lambda_{\mathbf{m}_i} = \sigma^2 \mathbf{I}$ 。如果图像的点在同一过程中被独立的提取出来，这种假设就不合理了。在这种情况下，以上的问题便转化成一个非线性最小二乘法的问题，例如  $\min_H \sum_i \|\mathbf{m}_i - \tilde{\mathbf{m}}_i\|^2$  这种非线性最小化可以用 Levenberg-Marquardt 算法[16]来实现的。这需要一个初始猜测值，它可由如下得到：

令  $\mathbf{x} = [\bar{\mathbf{h}}_1^T, \bar{\mathbf{h}}_2^T, \bar{\mathbf{h}}_3^T]$ 。(2) 式可以写为

$$\begin{bmatrix} \tilde{M}^T & 0^T & -u\tilde{M}^T \\ 0^T & \tilde{M}^T & -v\tilde{M}^T \end{bmatrix} \mathbf{x} = \mathbf{0}$$

当给已知  $n$  个点，就可得到  $n$  个上述方程，它们可以写成矩阵方程如： $\mathbf{L}\mathbf{x} = \mathbf{0}$ ，这里  $\mathbf{L}$  是一个  $2n \times 9$  的矩阵，定义  $\mathbf{x}$  为一个比例因子，这种方法被称作伴有最小奇异值的  $\mathbf{L}$  矩阵的右奇异向量（等价于矩阵  $\mathbf{L}$  最小特征值对应的特征向量）。在  $\mathbf{L}$  中的元素一部分是常数1，一部分在图像坐标系，一部分在世界坐标系里，一部分是两者的相乘，这就使得  $\mathbf{L}$  矩阵没有充分的数字化。可以在运行以上程序之前，将数据归一化可以得到很好的结果，例如在[11] 中提出的方法。

## B 从B矩阵提取内部参数

在3.1节中描述的矩阵  $\mathbf{B}$  被估计为一个比例因子，即  $\mathbf{B} = \lambda \mathbf{A}^{-T} \mathbf{A}$ ， $\lambda$  为一个任意的因子。我们可以很容易的从矩阵  $\mathbf{B}$  中提取出内部参数。

$$\begin{aligned} v_0 &= (B_{12}B_{13} - B_{11}B_{23}) / (B_{11}B_{22} - B_{12}^2) \\ \lambda &= B_{33} - [B_{13}^2 + v_0(B_{12}B_{13} - B_{11}B_{23})] / B_{11} \\ \alpha &= \sqrt{\lambda / B_{11}} \\ \beta &= \sqrt{\lambda B_{11} / (B_{11}B_{22} - B_{12}^2)} \\ c &= -B_{12}\alpha^2\beta / \lambda \\ u_0 &= cv_0 / \alpha - B_{13}\alpha^2 / \lambda \end{aligned}$$

## C 用旋转矩阵近似 $3 \times 3$ 矩阵

本节是为了构造一个最佳的旋转矩阵来近似一个  $3 \times 3$  的矩阵  $\mathbf{Q}$ 。这里的最佳是基于  $\mathbf{R} - \mathbf{Q}$  差值的最小F范数。解决方法参见我们给出的技术报告[24]。

## References

- [1] S. Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In Proc. 6th International Conference on Computer Vision, pages 790 – 796, Jan. 1998.
- [2] D. C. Brown. Close-range camera calibration. Photogrammetric Engineering, 37(8):855 – 866, 1971.
- [3] B. Caprile and V. Torre. Using Vanishing Points for Camera Calibration. The International Journal of Computer Vision, 4(2):127 – 140, Mar. 1990.
- [4] W. Faig. Calibration of close-range photogrammetry systems: Mathematical formulation. Photogrammetric Engineering and Remote Sensing, 41(12):1479 – 1486, 1975.
- [5] O. Faugeras. Three-Dimensional Computer Vision: a Geometric Viewpoint. MIT Press, 1993.
- [6] O. Faugeras, T. Luong, and S. Maybank. Camera selfcalibration: theory and experiments. In Proc. 2nd ECCV, pages 321 – 334, May 1992.
- [7] O. Faugeras and G. Toscani. The calibration problem for stereo. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, pages 15 – 20, Miami Beach, FL, June 1986.
- [8] S. Ganapathy. Decomposition of transformation matrices for robot vision. Pattern Recognition Letters, 2:401 – 412, Dec. 1984.
- [9] D. Gennery. Stereo-camera calibration. In Proc. 10th Image Understanding Workshop, pages 101 – 108, 1979.
- [10] R. Hartley. Self-calibration from multiple views with a rotating camera. In Proc. 3rd European Conference on Computer Vision, pages 471 – 478, Stockholm, Sweden, May 1994.
- [11] R. Hartley. In defence of the 8-point algorithm. In Proc. 5th International Conference on Computer Vision, pages 1064 – 1070, Boston, MA, June 1995.
- [12] R. I. Hartley. An algorithm for self calibration from several views. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, pages 908 – 912, Seattle, WA, June 1994.
- [13] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, pages 482 – 488, Santa Barbara, California, June 1998.
- [14] Q.-T. Luong and O. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. The International Journal of Computer Vision, 22(3):261 – 289, 1997.
- [15] S. J. Maybank and O. D. Faugeras. A theory of selfcalibration of a moving camera. The International Journal of Computer Vision, 8(2):123 – 152, Aug. 1992.
- [16] J. More. The levenberg-marquardt algorithm, implementation and theory. In G. A. Watson, editor, Numerical Analysis, Lecture Notes in Mathematics 630. Springer-Verlag, 1977.

- [17] C. C. Slama, editor. Manual of Photogrammetry. American Society of Photogrammetry, 4th ed., 1980.
- [18] G. Stein. Accurate internal camera calibration using rotation, with analysis of sources of error. In Proc. 5th International Conference on Computer Vision, pages 230 - 236, Cambridge, Massachusetts, June 1995.
- [19] B. Triggs. Autocalibration from planar scenes. In Proc. 5th European Conference on Computer Vision, pages 89 - 105, Freiburg, Germany, June 1998.
- [20] R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. IEEE Journal of Robotics and Automation, 3(4):323 - 344, Aug. 1987.
- [21] G. Wei and S. Ma. A complete two-plane camera calibration method and experimental comparisons. In Proc. Fourth International Conference on Computer Vision, pages 439 - 446, Berlin, May 1993.
- [22] G. Wei and S. Ma. Implicit and explicit camera calibration: Theory and experiments. IEEE Transactions on Pattern Analysis and Machine Intelligence, 16(5):469 - 480, 1994.
- [23] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(10):965 - 980, Oct. 1992.
- [24] Z. Zhang. A Flexible New Technique for Camera Calibration. Technical Report MSRTR-98-71, Microsoft Research, December 1998. Available together with the software at <http://research.microsoft.com/~zhang/Calib/>