

# 基于统计优化的单应矩阵估计算法

杨利峰<sup>1</sup>, 谢世朋<sup>2</sup>

(1. 阜阳师范学院 数学与计算科学学院, 安徽 阜阳 236041;

2. 安徽大学 数学科学学院, 安徽 合肥 230039)

**摘要:** 计算机科学中研究的图像是真实世界(即二维、三维欧式空间)到像平面的射影变换。平面射影变换(单应)估计是特征目标检测、注册、识别、三维重建等方面的关键步骤,但是如何鲁棒、精确地估计单应矩阵是一个没有很好解决的问题。在研究中发现,基于点与直线的直接的单应矩阵估计方法会导致出现较大误差的情况。针对这一情况,文中介绍了一种基于统计优化的单应矩阵估计方法,这种方法是通过单应矩阵的协方差张量的计算和优化来估计单应矩阵的。最后进行了简单的实验,比较了统计优化方法与进行归一化处理后的直接线性方法的估计结果,证明了基于优化统计的估计方法更加有效。

**关键词:** 单应矩阵;估计;统计优化;张量;归一化

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 1673-629X(2012)08-0139-04

## Estimation Method of Homography Matrix Based on Statistical Optimization

YANG Li-feng<sup>1</sup>, XIE Shi-peng<sup>2</sup>

(1. School of Mathematics and Computational Science, Fuyang Teachers College, Fuyang 236041, China;

2. School of Mathematical Science, Anhui University, Hefei 230039, China)

**Abstract:** The study of the images in computational science is projective transformation between the real world (the two-dimensional and three-dimensional Euclidean geometry) and the plane. Plane projective transformation (homography) estimation is a necessary step in many feature-based object detection, registration, recognition, three-dimension reconstruction and some other aspects, however, how to robustly and accurately estimate homography from images is a difficult problem. In this paper found that usual method of homography estimation based on points and lines may result in relatively big errors. Under such configurations, a new estimation method which is based on statistical optimization is proposed in this paper, this method estimates homography matrix through covariance tensor of homography matrix. Have conducted some experiments and compared the estimation results between the method of statistical optimization and the method of normalized DLT, thus proved that the estimation method based on statistical optimization is more effective.

**Key words:** homography matrix; estimation; statistical optimization; tensor analysis; normalization

## 0 引言

图像配准技术在计算机视觉中的应用越来越广泛,例如:目标的识别、注册、检测、跟踪,三维重建,医疗诊断等领域。图像之间的关系是计算机视觉中的基本工作之一,单应矩阵定义了不同图像之间的相互关系,在计算机视觉中有着广泛的应用。

单应矩阵的估计方法有很多种,最基本的方法是线性方法,它具有简单、快速、易实现等优点,所以在实践中得到了非常广泛的应用。但是直接线性方法受测

量数据的影响十分巨大,即使测量数据的误差很小也有可能导致估计的结果不稳定。为了解决直接线性方法的这个缺点, Hartley 于 1997 年提出了归一化的方法<sup>[1]</sup>。

归一化的基本思想是对测量数据做相应的归一化的变换,使得测量数据相较于原始的数据在空间分布上更加均匀,从而降低测量矩阵的条件数,提高算法的鲁棒性。

通过上述的归一化变换后,在进行图像配准的过程中,就会出现更多的点是无法在矫正后的坐标领域中找到另外一个视角中与之相似的点,为了能让其他平面上的点做到更好的配准,文中提出了基于张量分析基础之上的统计优化方法,该方法在单应矩阵的估计中更加有效。

收稿日期:2011-12-07;修回日期:2012-03-14

基金项目:安徽省高等学校青年人才基金项目(2009SQRZ150)

作者简介:杨利峰(1981-),男,硕士,讲师,研究方向为模式识别、智能计算。

# 1 直接线性变换

## 1.1 单应变换的基本原理

二维射影平面上的线性变换称为平面射影变换(单应),它可以用非奇异  $3 \times 3$  矩阵  $H$  表示为:

$$\rho \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

其中  $\rho \det H \neq 0$ 。上式可以简洁的用矩阵表示为:  
 $X' = HX$ 。

由于  $H$  是一齐次矩阵, $H$  的九个元素只有八个独立,因此一个射影变换有八个自由度。单应矩阵可以通过 4 对点对应来实现计算,也可以通过 3 个点和一条线的对应或 3 条线和一个点或 4 条线的对应来实现计算,但是不能通过 2 个点和 2 条线来实现计算。在实际情况下,找到的可能的对应点必存在一些偏差,这种偏差可能来自匹配算法的不精确,也可能来自像素本身的不精确性,这就需要更多的来解决这种偏差。当点数超出要求时成了超定问题,就需要通过最小二乘法来求解超定方程<sup>[2]</sup>。

## 1.2 基于点对应的直接线性估计方法

首先介绍一种单应矩阵  $H$  估计的简单的 2D 到 2D 上的点对应计算方法。假设存在点的对应  $X_i \leftrightarrow X'_i$ , 方程  $x' = Hx$  可以表示为向量积的形式:  $X'_i \times (HX_i) = 0$ 。如果把矩阵  $H$  的第  $j$  行记做  $H^j$ , 就有:

$$HX_i = \begin{pmatrix} H^{1T} x_i \\ H^{2T} x_i \\ H^{3T} x_i \end{pmatrix}$$

记  $X'_i = (x'_i, y'_i, z'_i)^T$ , 向量积就可以简明的写成:

$$X'_i \times HX_i = \begin{bmatrix} y'_i H^{3T} X_i - z'_i H^{2T} X_i \\ z'_i H^{1T} X_i - x'_i H^{3T} X_i \\ x'_i H^{2T} X_i - y'_i H^{1T} X_i \end{bmatrix}$$

上述方程  $X'_i \times (HX_i) = 0$ , 可以重新写成:

$$\begin{bmatrix} 0^T & -z'_i X_i^T & y'_i X_i^T \\ z'_i X_i^T & 0^T & -x'_i X_i^T \\ -y'_i X_i^T & x'_i X_i^T & 0^T \end{bmatrix} \begin{bmatrix} H^1 \\ H^2 \\ H^3 \end{bmatrix} = 0$$

这些方程可以简记为  $L_i H = 0$ , 这里  $L_i$  为  $3 \times 9$  的矩阵,  $H = [H^{1T} \ H^{2T} \ H^{3T}]$ 。尽管上式包含 3 组方程,但是只有两组是线性独立的。这样知道一组对应点就能得到关于  $H$  的两组方程:

$$\begin{bmatrix} 0^T & -z'_i X_i^T & y'_i X_i^T \\ z'_i X_i^T & 0^T & -x'_i X_i^T \end{bmatrix} \begin{bmatrix} H^1 \\ H^2 \\ H^3 \end{bmatrix} = 0$$

给出平面上一组对应的点,就能得到一组形如

$L_i H = 0$  这样的方程。

在实际的工作中,在图像中采样的对应点通常不满足方程  $X' = HX$ , 因为图像中采样点含有噪声。假设采样点  $X'_i$  受到了均值为 0 协方差矩阵为  $\Sigma_x$  的高斯噪声的污染。 $H$  的最大似然估计为:

$$J = \sum (X'_i - \hat{X}_i)^T \Sigma_x^{-1} (X'_i - \hat{X}_i)$$

这里  $\hat{X}_i = \frac{1}{H^{3T} X_i} \begin{bmatrix} H^{1T} X_i \\ H^{2T} X_i \end{bmatrix}$ ,  $H^i$  为矩阵  $H$  的第  $i$  行。给定

$n$  组对应点,就可以写成矩阵方程  $LH = 0$ , 这里  $L$  为  $2n \times 9$  的矩阵。上述问题就变成了一个非线性最小二乘估计求使得  $\|X'_i - \hat{X}_i\|$  最小的  $H$  问题<sup>[3-6]</sup>。

## 1.3 方法存在的问题及改进

对于线性问题而言,测量矩阵的条件数是影响算法稳定的一个重要原因。条件数是方程病态程度的一个重要的指标值,条件数越大说明方程越病态,该方法越不鲁棒,提高鲁棒性的有效途径之一就是降低测量矩阵的条件数<sup>[7,8]</sup>。

如果观测点为  $n$  个  $(X_1, X_2, \dots, X_n)$ , 对于点  $X_i = (x_i, y_i, z_i)^T$ , 令  $k_1 = \sum_{i=1}^n x_i, k_2 = \sum_{i=1}^n y_i, k_3 = \sum_{i=1}^n z_i$ , 构造变换矩阵如下:

$$N_1 = \begin{bmatrix} 1 & 0 & -k_1/k_3 \\ 0 & 1 & -k_2/k_3 \\ 0 & 0 & 1 \end{bmatrix}$$

经过  $X_i^1 = N_1 X_i$  变换后的  $X_i^1 = (x_i^1, y_i^1, z_i^1)^T$  满足  $\sum_{i=1}^n x_i^1 = 0, \sum_{i=1}^n y_i^1 = 0$ , 即变换后的重心落在  $z$  轴上,变换后直线坐标的  $z$  轴分布更加均匀。

$$\text{令 } s = \left[ \frac{\sum_{i=1}^n ((x_i^1)^2 + (y_i^1)^2)}{2 \sum_{i=1}^n (z_i^1)^2} \right]^+, \text{ 构造变换矩阵 } N_2 \text{ 如}$$

下:

$$N_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & s \end{bmatrix}$$

经过  $X_i^2 = N_2 X_i^1$  变换后的直线坐标  $X_i^2 = (x_i^2, y_i^2, z_i^2)^T$  满足:  $\sum_{i=1}^n ((x_i^2)^2 + (y_i^2)^2) = 2 \sum_{i=1}^n (z_i^2)^2$ , 即变换后直线坐标的空间分布更加均匀。

归一化算法的主要步骤(假定对应直线为  $X_i \leftrightarrow \tilde{X}_i$ ):

Step1: 对于直线  $X_i$  根据上述变换的需要构造  $N_1, N_2$  得到  $X_i^2 = N_2 N_1 X_i$ ;

Step2: 对于直线  $\tilde{X}_i$  也根据上述变换的需要构造  $N_1, N_2$  得到  $\tilde{X}_i^2 = N_2 N_1 \tilde{X}_i$ ;

Step3:由直接线性方法求解对应直线  $X_i^2 \leftrightarrow \bar{X}_i^2$  所对应的单应矩阵  $\bar{H}$ ;

Step4:做归一化变换的逆变换,得到对应直线  $X_i \leftrightarrow \bar{X}_i$  的单应矩阵。

通过这样归一化的处理就会使利用直接线性方法估计单应矩阵更加有效。

## 2 统计优化方法

### 2.1 统计优化方法的主要思想

最小二乘法广泛应用于单应矩阵的简易计算,但这样容易出现统计偏差<sup>[9]</sup>,因为在运用该方法时往往忽略微小的物理量之间的度量。图像坐标中的噪声是欧氏空间下的,因此在欧几里得框架下讨论问题变得更加可取。Kanatani 等人提出了一种在欧几里得空间中计算单应矩阵的方法<sup>[10,11]</sup>。不确定的数据点  $(x, y)$  和  $(x', y')$  可以通过它们的协方差矩阵  $\Sigma$  和  $\Sigma'$  表示。向量  $X$  和  $X'$  具有下列单独的协方差矩阵:

$$V[X_\alpha] = \frac{1}{f^2} \begin{bmatrix} \Sigma & 0 \\ 0^T & 0 \end{bmatrix}$$

$$V[X'_\alpha] = \frac{1}{f'^2} \begin{bmatrix} \Sigma' & 0 \\ 0^T & 0 \end{bmatrix}$$

如果将协方差矩阵可以表示成  $V[X_\alpha] = \varepsilon^2 V_0$   $[X_\alpha]$  和  $V[X'_\alpha] = \varepsilon'^2 V_0 [X'_\alpha]$ , 这里未知量  $\varepsilon$  可以认为是噪声水平。归一化的协方差矩阵  $V_0 [X_\alpha]$  和  $V_0 [X'_\alpha]$  可以表示噪声发生的位置和方向。如果没有其他的信息可以利用,可以假设它们为各向同性的和齐次的。这里  $\hat{H}$  表示单应矩阵的估计值,  $H$  表示单应矩阵的真实值,未确定的估计值  $\hat{H}$  可以通过其协方差张量来度量:

$$V[\hat{H}] = E[P((\hat{H}-H) \otimes (\hat{H}-H))P^T]$$

这里  $E[\cdot]$  表示期望,运算符  $\otimes$  表示张量积。张量  $P$  的第  $ijkl$  元素通过  $P_{ijkl} = \delta_{ik}\delta_{jl} - H_{ij}H_{kl}$  得到。

理论上的精确解可以通过下式利用最小化 Mahalanobis 距离的平方得到:

$$W_\alpha = X'_\alpha \times H V_0 [X] H^T \times X'_\alpha + H X_\alpha \times V_0 [X_\alpha] \times (H X_\alpha)$$

利用 Lagrange 乘子和一阶近似可以得到:

$$J = \sum (X'_\alpha \times H X, W_\alpha (X'_\alpha \times H X_\alpha))$$

协方差张量  $V[\hat{H}]$  的特征矩阵  $U$  对应的最大特征根  $\lambda$  表明了 9 维空间中误差最可能发生的方向。 $H$  的结果受到两个方向  $H^+ = N[H + \sqrt{\lambda}U]$ ,  $H^- = N[H - \sqrt{\lambda}U]$  的干扰,这里  $N[\cdot]$  表示归一化计算<sup>[12-15]</sup>。

### 2.2 统计优化方法的主要步骤

利用上述方法进行单应矩阵估计主要步骤如下:

Step1:给定初值  $c=0, W_\alpha=I, \alpha=1, 2, \dots, N$ ;

Step2:定义张量  $M$ :

$$M = \frac{1}{N} \sum_{\alpha=1}^N \sum_{k,l=1}^3 W_\alpha^{(kl)} (e^{(k)} \times X'_\alpha) \otimes X_\alpha \otimes (e^{(l)} \times X'_\alpha) \otimes X_\alpha;$$

$X_\alpha$ ;

Step3:计算张量  $N = (N_{ijkl})$ :

$$N_{ijkl} = \frac{1}{N} \sum_{\alpha=1}^n \sum_{m,n,p,q=1}^3 \in_{imp} \in_{knq} W_\alpha^{mn} (V_0 [X_\alpha]_{jl} X'_{\alpha(p)} X'_{\alpha(q)} + V_0 [X'_\alpha]_{jl} X_{\alpha(l)} X_{\alpha(l)});$$

Step4:计算张量  $M = M - cN$  的特征值并按从大到小的顺序排好  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_9$ , 如果  $\lambda_9 \approx 0$ , 停止。否则按照下面的表达式更新  $c$  和  $W_\alpha$  的值:

$$c \leftarrow c + \frac{\lambda_9}{(H_9, NH_9)}$$

$$W_\alpha \leftarrow X'_\alpha \times H_9 V_0 [X_\alpha] H_9^T \times X'_\alpha + (H_9 X_\alpha) \times V_0 [X'_\alpha] \times V_0 [X'_\alpha] \times (H_9 X_\alpha)$$

将更新后的值带入第二步继续进行计算<sup>[16]</sup>。

## 3 实验结果

为了测试文中提出的单应矩阵统计优化估计方法的效果,准备了同一场景的两幅图像,并且标明理想状态下一些对应的点。对于每一幅图像,噪声都可以根据实验的需要而进行添加。用第一幅图像经过变换后的点的坐标与第二幅图像中对应的点进行比较。计算第二幅图像中的对应点与变换点之间的欧氏距离,假设第一幅图像中的点为  $X$ , 在第二幅图像中对应的点为  $X'$ , 点  $HX$  为其匹配点,用  $d(X, Y)$  来表示非齐次的两点间的欧氏距离,则对应的点集间的变换误差<sup>[17]</sup>为:

$$E = \sum_i d(X_i, HX_i)^2$$

当然还有其他的一些误差衡量的指标,在这里就不过多的计算了。

通过实验比较不同噪声水平下直接线性方法(DLT)、归一化的直接线性方法(N-DLT)和基于统计优化的方法(S-O)进行单应估计的结果如表1:

表1 不同噪声水平下三种估计方法的结果

噪声水平( $\sigma$ )	DLT	N-DLT	S-O
0	0.0000	0.0000	0.0000
0.5	0.9738	0.7125	0.6127
1	1.8629	1.2640	0.9046

根据表1很容易看出,在噪声水平越高的情况下, S-O 算法的优势越明显,所以在噪声水平比较高的图像匹配中更应该选择统计优化的估计方法。

根据表2的结果很容易看出对应的点越多统计优化算法相较于其他两种方法误差下降的越快,所以在具有较多的测量数据点时更应该选择统计优化算法。

表 2 不同匹配点数目下的三种估计方法结果

选择的点数	DLT	N-DLT	S-O
20	0.7056	0.7148	0.6926
40	0.6835	0.6610	0.6242
60	0.6672	0.6097	0.5532
80	0.6493	0.5931	0.5207

#### 4 结束语

文中提出了一种基于张量分析基础之上的统计优化算法,用该算法来估计不同场景间对应的单应矩阵具有其独特的优势。通过上面的实验可以看到当图像噪声较大的时候,选择统计优化算法可以有效地降低误差;在观测数据比较多的时候选择统计优化算法也能得到更优的结果。总之,用统计优化的方法来进行单应矩阵的估计是十分有效的。

#### 参考文献:

- [1] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision[M]. Cambridge: Cambridge Press, 2003: 281-285.
- [2] Lowe D G. Object recognition from local scale-invariant features[C]//Proc. of the International Conference on Computer Vision. [s. l.]: [s. n.], 1999: 1148-1160.
- [3] 胡茂林. 空间和变换[M]. 北京: 科学出版社, 2007: 160-181.
- [4] Azuma R, Baillot Y, Behringer R, et al. Recent Advances in Augmented Reality[J]. IEEE Computer Graphics and Applications, 2001, 21(6): 34-47.
- [5] Saito S, Hiyama A, Tanikawa T, et al. Indoor Marker-based Localization Using Coded Seamless Pattern for Interior Decoration[C]//IEEE Virtual Reality Conference, 2007. Charlotte, North Carolina, USA: [s. n.], 2007: 67-74.
- [6] Kim J B, Kim H J. Efficient region-based motion segmentation for video monitoring system[J]. Pattern Recognition Letters, 2003, 24(1-3): 113-128.
- [7] Kanatani K, Ohta N, Kanazawa Y. Optimal homography computation with a reliability measure [C]//Proceedings of the IAPR Workshop on Machine Vision Applications. [s. l.]: [s. n.], 1998.
- [8] Kumar M P, Goyal S, Kuthirummal S, et al. Discrete contours in multiple views: approximation and recognition [J]. Image and Vision Computing, 2004, 22(14): 1229-1239.
- [9] Kumar M P, Jawahar C V, Narayanan P J. Geometric structure computation from conics[C]// Proc. of Indian Conference on Computer Vision, Graphics and Image Processing. [s. l.]: [s. n.], 2004.
- [10] Kumar M P, Kuthirummal S, Jawahar C V, et al. Planar homography from fourier domain representation[C]//Proc. of International Conference on Signal Processing and Communications (SPCOM). [s. l.]: [s. n.], 2004.
- [11] 薄 华, 马缚龙, 焦李成. 图像纹理的灰度共生矩阵计算问题的分析[J]. 电子学报, 2006(1): 187-192.
- [12] Ilonen J, Kamarainen J K, Paalanen P, et al. Image feature localization by multiple hypothesis testing of Gabor features[J]. IEEE Transactions on Image Processing, 2008, 17(3): 311-326.
- [13] 张春燕, 汤 进, 赵海峰, 等. 基于 MDS 的统计形状聚类[J]. 计算机技术与发展, 2007, 17(3): 58-62.
- [14] 何传江, 田巧玉. 几何活动轮廓模型中停止速度函数的尺度变换[J]. 计算机工程与应用, 2008(8): 73-78.
- [15] 刘秀平, 常先堂, 李治隆. 一种基于边缘和区域信息的变分水平集图像分割方法[J]. 大连理工大学学报, 2008(5): 162-166.
- [16] 闫卫杰, 杨 丹, 王洪星, 等. 轮廓结构张量 B-样条多尺度表示的角点检测[J]. 计算机技术与发展, 2010, 20(4): 80-83.
- [17] 陈 亮, 郭 雷, 王雅萍, 等. 一种基于结构张量的 MAS 边缘检测算法[J]. 计算机科学, 2009(1): 115-123.

## 科技论文摘要四大要素的内容

1. 目的: 研究、研制、实验等课题所涉及的范围和所要解决的问题。
  2. 方法: 所采用原理、理论、思想、技术、条件、材料、工艺、结构等, 如何创建的新理论、新技术、新方法、新材料、新工艺、新结构等。
  3. 结果: 研究的结果、所得数据、被确定的关系、得到的效果和性能等。
  4. 结论: 对结果通过分析、比较、升华所得到的具有普遍意义的规律和适用范围。
- 这四大要素是简明扼要地、全面准确地表述论文关键内容的必要条件, 缺一不可。

作者: 葛新, 董朝阳  
作者单位: 葛新(西安未来国际信息股份有限公司, 陕西西安710075), 董朝阳(西安建筑科技大学机电工程学院, 陕西西安710055)  
刊名: 计算机技术与发展  
英文刊名: Computer Technology and Development  
年, 卷(期): 2012(8)

参考文献(17条)

1. Hartley R, Zisserman A Multiple View Geometry in Computer Vision 2003
2. Lowe D G Object recognition from local scale-invariant features 1999
3. 胡茂林 空间和变换 2007
4. Azuma R, Bailiot Y, Behringer R Recent Advances in Augmented Reality 2001(06)
5. Saito S, Hiwama A, Tanikawa T Indoor Marker-based Localization Using Coded Seamless Pattern for Interior Decoration 2007
6. Kim J B, Kim H J Efficient region-based motion segmentation for video monitoring system 2003(1-3)
7. Kanatani K, Ohita N, Kanazawa Y Optimal homography computation with a reliability measure 1998
8. Kumar M P, Goyal S, Kuthirummal S Discrete contours in multiple views: approximation and recognition[外文期刊] 2004(14)
9. Kumar M P, Jawahar C V, Narayanan P J Geometric structure computation from conics 2004
10. Kumar M P, Kuthirummal S, Jawahar C V Planar homography from fourier domain representation 2004
11. 魏华, 马博龙, 焦李成 图像纹理的灰度共生矩阵计算问题的分析[期刊论文]-电子学报 2006(01)
12. Ilonen J, Kamarainen J K, Paalanen P Image feature localization by multiple hypothesis testing of Gabor features 2008(03)
13. 张存燕, 汤进, 赵海峰 基于MDS的统计形状聚类[期刊论文]-计算机技术与发展 2007(03)
14. 何传江, 田巧玉 几何活动轮廓模型中停止速度函数的尺度变换[期刊论文]-计算机工程与应用 2008(08)
15. 刘秀平, 常光堂, 李治隆 一种基于边缘和区域信息的变分水平集图像分割方法[期刊论文]-大连理工大学学报 2008(05)
16. 闫卫杰, 杨丹, 王洪星 轮廓结构张量B-样条多尺度表示的角点检测[期刊论文]-计算机技术与发展 2010(04)
17. 陈亮, 郭雷, 王雅萍 一种基于结构张量的MS边缘检测算法[期刊论文]-计算机科学 2009(01)

本文链接: [http://d.g.wanfangdata.com.cn/Periodical\\_wjfx201208036.aspx](http://d.g.wanfangdata.com.cn/Periodical_wjfx201208036.aspx)