

# 基于仿射变换模型的图象特征点集配准方法研究

章权兵 罗斌 韦穗 杨尚骏

(安徽大学计算智能与信号处理实验室, 合肥 230039) (安徽大学数学系, 合肥 230039)

**摘要** 图象配准是计算机视觉中目标识别的一种基本方法,其目的是在待识别图象中寻找与模型图象的最佳匹配。目前,对于图象间的变换为相似变换的情形已有闭合公式,本文则分别运用最小二乘和矩阵伪逆两种方法,对图象间的变换为仿射变换的情形进行了研究,并给出了简单的闭合公式,实验表明这种方法精确、稳定、受噪声影响小。

**关键词** 计算机图象处理(520·6040) 配准 仿射变换 最小二乘 伪逆

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2003)10-1121-05

## Registration for Feature Point Sets Based on Affine Transformation

ZHANG Quan-bing, LUO Bin, WEI Sui

(Key lab of Intelligent Computing and Signal Processing of Anhui University, Hefei 230039)

YANG Shang-jun

(Department of Mathematics of Anhui University, Hefei 230039)

**Abstract** This work investigates the image registration from feature point sets. Image registration is a fundamental object recognition method in computer vision and it aims to find best matches between two or more point sets when there are geometric distortions, point measurement errors and contamination present. Up to now, closed form solution has been developed only when the geometric distortion is similarity transformation. This paper concentrates on image registration from feature point sets when the geometric distortion between the two images is affine transformation and gives closed form solution for the transformation parameters that minimize the root mean-squared residual error of the image points by the linear least-squares techniques and the pseudo-inverse of matrix respectively. In order to give the simple closed form solution, the image points are represented by homogeneous coordinates and the theories of matrix are used. The algorithms are evaluated on both synthetic and real world images and the experiment results show that the methods given in this paper are accurate, stable and are only affected slightly by noise.

**Keywords** Computer image processing, Image registration, Affine transformation, Least-squares, Pseudo-inverse

## 0 引言

数字图象配准是计算机视觉及图象分析中的一个基本问题<sup>[1]</sup>。众所周知,立体视觉中需要配准由多个摄像机从不同角度和位置所获取的图象,运动估计需要配准同一场景不同时间的多幅图象<sup>[2]</sup>。图象配准在医学图象和遥感遥测图象处理和分析方面的应用也非常普遍,它还可用于图象拼接、交通检测和管理视频系统等方面。图象配准技术的广泛应用背景使其

成为近年来发展迅速的图象处理技术之一,图象特征点集配准长期以来一直受到研究人员的重视。

相关是图象配准的经典方法,该方法易于理解和实现,但具有运算量大的缺点。更为严重的是,图象间的畸变使得图象直接相关运算变得毫无意义。由于图象间存在各种畸变,因此要进行这种畸变图象的配准,就需要进行畸变矫正。目前研究人员针对不同的畸变模型已研制出多种不同的算法<sup>[3-8]</sup>,对于相似变换模型已有闭合公式<sup>[9]</sup>,但对于更具有—般性的变换——仿射变换却没有一个简单的闭合公式。

基金项目:国家自然科学基金项目(60143003);安徽省教育厅自然科学基金项目(2003KJ005)

收稿日期:2003-01-13;改回日期:2003-05-27

给定  $n$  维空间中的两个特征点集  $\{x_i\}, \{y_i\}, i=1, 2, \dots, m$ , 假设其对应关系为  $y_i = f(x_i), \forall i=1, 2, \dots, m$ , (注: 对图象配准问题,  $n$  一般取为 2 或 3, 但对于这里所讨论的数学模型而言,  $n$  可取任意正整数, 当然对应的点数  $m$  要足够大). 由于噪声的存在, 上式一般不可能精确成立, 因此特征点集配准就是估计变换  $f$ , 使得两特征点集在某种意义上的差别极小化.

## 1 仿射变换参数的最小二乘估计

如果两个特征点集或两幅图象之间的对应关系是仿射的, 那么

$$y_i = f(x_i) = Ax_i + t, \forall i = 1, 2, \dots, m \quad (1)$$

其中,  $A$  是任意的可逆线性变换 (对应于  $n \times n$  矩阵),  $t$  是  $n$  维平移矢量. 所要解决的问题就是估计极小化配准后的均方根残差

$$\mathcal{E}(A, t) = \left( \frac{1}{m} \sum_{i=1}^m \|y_i - (Ax_i + t)\|^2 \right)^{1/2} \quad (2)$$

的仿射变换参数  $A, t$ . 若记

$$X = [x_1, x_2, \dots, x_m], Y = [y_1, y_2, \dots, y_m]$$

则  $Y = AX + t\mathbf{1}^T$ , 其中  $\mathbf{1}$  是元素全为 1 的  $m$  维列向量, 因此问题等价于极小化误差函数

$$\mathcal{L}(A, t) = m\mathcal{E}^2(A, t) = \text{tr}((Y - (AX + t\mathbf{1}^T))(Y - (AX + t\mathbf{1}^T))^T) \quad (3)$$

其中,  $\text{tr}(A)$  表示方阵  $A$  的迹, 即  $A$  的主对角线元素之和.  $\mathcal{L}(A, t)$  是关于  $A$  和  $t$  的元素的多元函数, 极小化  $\mathcal{L}(A, t)$  的  $A^*, t^*$  应满足下列条件:

$$\left. \frac{\partial \mathcal{L}(A, t)}{\partial t} \right|_{t=t^*} = 0 \quad (4)$$

$$\left. \frac{\partial \mathcal{L}(A, t)}{\partial A} \right|_{A=A^*} = 0 \quad (5)$$

由式(4)可以求得 (详细过程见附录)

$$t^* = \frac{1}{m}(Y - AX)\mathbf{1} \quad (6)$$

令  $J = I - \frac{1}{m}\mathbf{1}\mathbf{1}^T$ , 则  $JJ^T = J^T J = J^T = J$ , 从而将式(6)代入式(3)并作一些化简后可得

$$\begin{aligned} \mathcal{L}(A, t^*) &= \text{tr}((YJ - AXJ)(YJ - AXJ)^T) \\ &= \text{tr}(YJY^T) + \text{tr}(AXJX^T A^T) - 2\text{tr}(YJX^T A^T) \end{aligned}$$

记  $\tilde{X} = XJX^T, \tilde{Y} = YJX^T$ , 得

$$\mathcal{L}(A, t^*) = \text{tr}(YJY^T) + \text{tr}(A\tilde{X}A^T) - 2\text{tr}(\tilde{Y}A^T) \quad (7)$$

可以证明 (详见附录)

$$\frac{\partial \mathcal{L}(A, t^*)}{\partial A} = 2(A\tilde{X} - \tilde{Y}) \quad (8)$$

代入式(5)得  $A^* \tilde{X} = \tilde{Y}$ . 对于一般位置上的点,  $\tilde{X}$  为

非奇异矩阵, 因此有

$$A^* = \tilde{Y}\tilde{X}^{-1}$$

$$\text{进而 } t^* = \frac{1}{m}(Y - A^*X)\mathbf{1} = \frac{1}{m}(Y - \tilde{Y}\tilde{X}^{-1}X)\mathbf{1}.$$

至此, 就求出了极小化配准后的均方根残差  $\mathcal{E}(A, t)$  的仿射变换参数  $A^*, t^*$ .

归纳起来, 最小二乘估计仿射变换参数  $A, t$  的步骤如下:

(1) 构造  $X = [x_1, x_2, \dots, x_m], Y = [y_1, y_2, \dots, y_m], J = I - \frac{1}{m}\mathbf{1}\mathbf{1}^T$ .

(2) 求出  $\tilde{X} = XJX^T, \tilde{Y} = YJX^T$ .

(3) 线性变换  $A = \tilde{Y}\tilde{X}^{-1}$ .

(4) 平移矢量  $t = \frac{1}{m}(Y - \tilde{Y}\tilde{X}^{-1}X)\mathbf{1}$ .

## 2 利用矩阵伪逆求解仿射变换参数

如果利用矩阵的伪逆, 则可以更简单、更直接地求解仿射变换参数. 令  $\tilde{X} = [X^T \ \mathbf{1}]^T, H = [A \ t]^T$ , 则将特征点集对应的关系式  $Y = AX + t\mathbf{1}^T$  转置后可以改写为  $\tilde{X}H = Y^T$ . 因对一般位置上的点,  $\tilde{X}$  为列满秩, 所以  $\tilde{X}$  的伪逆矩阵  $\tilde{X}^+ = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T$ , 故有  $H = \tilde{X}^+ Y^T = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T Y^T$ . 由  $H$  很容易就可求得  $A, t$ .

归纳起来, 利用矩阵的伪逆求解仿射变换参数  $A, t$  的步骤如下:

(1) 构造矩阵  $\tilde{X} = [X^T \ \mathbf{1}]^T, H = [A \ t]^T$ .

(2) 求出  $H = \tilde{X}^+ Y^T = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T Y^T$ .

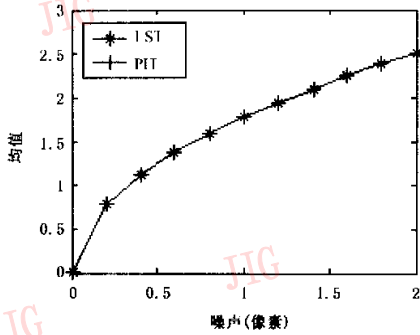
(3) 线性变换  $A$  即为  $H$  的前  $n$  行的转置, 平移矢量  $t$  即为  $H$  的第  $i+1$  行的转置.

## 3 实验

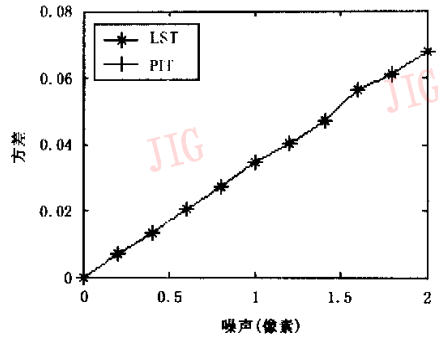
为了验证本文算法理论的正确性, 分别用模拟与真实场景进行了实验.

### 3.1 模拟实验

首先利用随机数发生器产生一个点列  $X, X$  为含有 30 个二维点的坐标阵列, 并通过对  $X$  施加仿射变换得到含有 30 个二维点的坐标阵列  $Y$ ; 然后对  $X$  和  $Y$  加上已知方差的零均值的高斯噪声, 并分别运用最小二乘 (LST) 和矩阵伪逆求解 (PIT) 算法求出仿射变换参数. 对不同水平的高斯噪声 (单位: 像素) 各运算 1 000 次, 所得结果的均方根残差  $\mathcal{E}(A, t)$  的均值和方差如图 1 所示.



(a)  $\varepsilon(A, t)$  的均值



(b)  $\varepsilon(A, t)$  的方差

图 1 对施加不同水平的高斯噪声各运算 1 000 次,残差  $\varepsilon(A, t)$  的均值和方差

### 3.2 真实图象的配准

图 2 为实际拍摄的 6 幅图象(CMU/VASC), 实验中,首先对每幅图象检测出对应角点(图中用“+”或“×”标出),然后分别运用最小二乘和矩阵伪逆对所有角点和部分角点(空间中同一平面上的点的像,图中用“+”标出)估计第 1 幅图象与其他各幅图象之间的变换参数,所得结果的误差如表 1 所示.由于最小二乘和矩阵伪逆求解的结果非常接近(相

差  $10^{-8}$  左右),因此表 1 中给出的误差既是最小二乘的结果,又是利用矩阵伪逆直接求解的结果.

表 1 对角点计算第 1 幅图象与其他各幅图象之间的变换参数所得结果的均方根残差

	对应图象				
	a & b	a & c	a & d	a & e	a & f
全部角点	2.803 1	11.521 5	12.698 6	16.114 1	17.950 1
部分角点	1.025 3	1.407 1	1.415 0	1.431 3	1.746 2

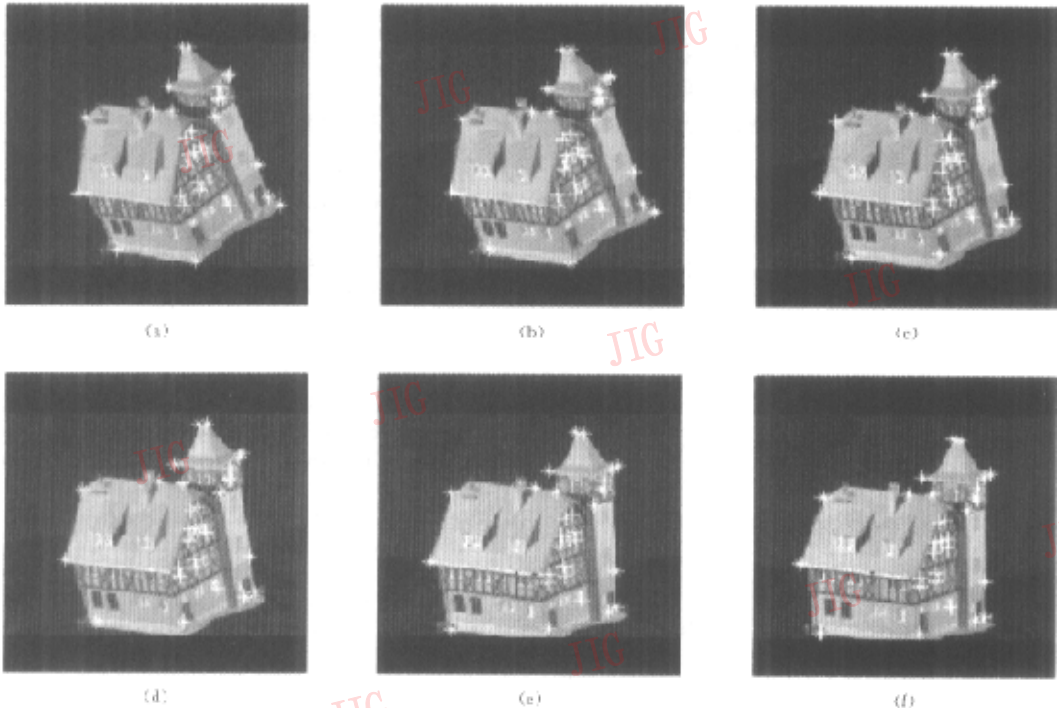


图 2 实际拍摄的 6 幅图象(CMU/VASC 图象库)

## 4 分析与结论

因为利用矩阵的伪逆求解线性方程组的公式是由最小二乘原理推导得到的,所以本文中两种算法本质上是相同的,从而它们的结果也就相差很小,模拟数据和真实场景实验结果都说明了这一点。

由于到目前为止,还没有发现文献中有其他相应的闭合公式,因此无法进行比较分析。但从图1(a)可以看出,当噪声在2个像素以内时,配准的误差随噪声的增加而增加,但最高不超过2.5个像素,因此,文中的算法是比较精确的,并且受噪声的影响比较小。另外,从图1(b)可以看出,误差的方差受噪声的影响非常小,这说明该算法具有很好的稳定性。

真实场景实验中,由部分角点的实验结果可以看出,该算法具有很高的精度,并且受噪声的影响比较小。对于空间同一平面上的点而言,它们在两幅图象上的像点之间的对应关系近似于仿射变换,图象中物体的方位变化越大,它们与仿射变换的差别就越大,从而结果的误差就越大。对于所有角点(空间中处于不同平面上的点的像)来说,当图象中物体的方位变化很小时(如图2中(a)和(b)两幅图象),它们之间的对应关系近似于仿射变换,因此结果的误差也不大。但当图象中物体的方位变化稍大时(如图2中(a)与其他图象),它们与仿射变换的差别就很大,从而误差也就很大。

基于以上分析可知,如果能找到由空间同一平面上的点投影得到的像点,然后由这些像点进行配准,结果会有所改善。因此,进一步将研究平面的自动检测方法以及拟采用更一般的变换,如射影变换等来进行图象配准,以期获得更好的配准结果。

## 附 录

文中式(6)的证明

首先,利用矩阵迹的性质将误差函数(文中式(3))展开为

$$\mathcal{L}(A, t) = \text{tr}(YY^T) - 2\text{tr}(AXY^T) - 2\text{tr}((Y-AX)\mathbf{1}t^T) + \text{tr}(\mathbf{1}t^T\mathbf{1}^T) \quad (1)$$

因为对任意  $n$  维列向量  $S$  有  $\text{tr}(St^T) = \sum_{j=1}^n S_j t_j$ , 所以  $\frac{\partial \text{tr}(St^T)}{\partial t_j} = S_j$ , 从而  $\frac{\partial \text{tr}(St^T)}{\partial S} = S$ 。因为  $\text{tr}(\mathbf{1}t^T\mathbf{1}^T) =$

$m \sum_{i=1}^n t_i^2$ , 所以  $\frac{\partial \text{tr}(\mathbf{1}t^T\mathbf{1}^T)}{\partial t_i} = 2mt_i$ , 从而  $\frac{\partial \text{tr}(\mathbf{1}t^T\mathbf{1}^T)}{\partial \mathbf{t}} = 2m\mathbf{t}$ 。将式(1)对  $t$  求偏导得  $\frac{\partial \mathcal{L}(A, t)}{\partial \mathbf{t}} = \mathbf{0} - \mathbf{0} - 2(Y-AX)\mathbf{1} + 2m\mathbf{t}$ , 代入文中式(4)得  $2m\mathbf{t}^* - 2(Y-AX)\mathbf{1} = \mathbf{0}$ , 因此  $\mathbf{t}^* = \frac{1}{m}(Y-AX)\mathbf{1}$ , 证毕。

文中式(8)的证明:

因为  $\text{tr}(\tilde{Y}A^T) = \sum_{i=1}^n \sum_{j=1}^n \tilde{y}_{ij} a_{ij}$ , 所以  $\frac{\partial \text{tr}(\tilde{Y}A^T)}{\partial a_{ij}} = \tilde{y}_{ij}$ , 从而  $\frac{\partial \text{tr}(\tilde{Y}A^T)}{\partial A} = (\tilde{y}_{ij}) = \tilde{Y}$ 。因为  $\text{tr}(A\tilde{X}A^T) = \sum_{k=1}^n \sum_{i=1}^n \sum_{j=1}^n a_{ik} \tilde{x}_{ij} a_{jk}$ , 所以利用  $\tilde{X}$  的对称性 ( $\tilde{X}^T = XJ^T X^T = XJX^T = \tilde{X}$ ) 可得  $\frac{\partial \text{tr}(A\tilde{X}A^T)}{\partial a_{ij}} = \sum_{k=1}^n \tilde{x}_{jk} a_{ik} + \sum_{k=1}^n a_{ik} \tilde{x}_{ij} = 2 \sum_{k=1}^n a_{ik} \tilde{x}_{kj} = 2(A\tilde{X})_{ij}$ , 从而  $\frac{\partial \text{tr}(A\tilde{X}A^T)}{\partial A} = 2A\tilde{X}$ 。将文中式(7)对  $A$  求导得

$$\frac{\partial \mathcal{L}(A, t^*)}{\partial A} = \mathbf{0} + 2A\tilde{X} - 2\tilde{Y} = 2(A\tilde{X} - \tilde{Y})$$

证毕。

## 参 考 文 献

- 1 Lisa Gottesfeld Brown. A survey of image registration techniques[J]. ACM Computing Surveys, 1992, 24(4): 325~376.
- 2 马颂德, 张正友. 计算机视觉——计算理论与算法基础[M]. 北京: 科学出版社, 1998.
- 3 Feldmar J, Ayache N J. Rigid, Affine and locally affine registration of free-form surface[J]. International Journal of Computer Vision, 1996, 18(2): 99~119.
- 4 Moshfeghi M, Ranganath S, Nawyn K. Three-dimensional elastic matching of volumes[J]. IEEE Transactions on Image Processing, 1994, 3(2): 128~138.
- 5 Jain A K, Zhong Y, Lakshmanan S. Object matching using deformable templates[J]. IEEE Transactions on Pattern Recognition and Machine Intelligence, 1996, 18(3): 267~278.
- 6 Bookstain F L. Shape and the information in medical images: a decade of the morphometric synthesis[J]. Computer Vision and Image Understanding, 1997, 66(2): 97~118.
- 7 Luo B, Hancock E R. Iterative procrustes alignment with the EM algorithm[J]. Image and Vision Computing, 2002, 20: 377~396.
- 8 罗纲, 罗斌. 图特征点集配准的加权相关迭代算法[J]. 中国图象图形学报, 2000, 5A(9): 755~758.
- 9 Shinji Umeyama. Least-squares estimation of transformation parameters between two point patterns[J]. IEEE Transactions on Pattern Recognition and Machine Intelligence, 1991, 13(3): 376~380.



**章权兵** 1977 年生,1998 年获安徽大学数学系理学学士学位,2001 年获安徽大学数学系理学硕士学位,现为安徽大学计算机科学与信息工程学院博士研究生.主要研究计算机视觉、图象处理等.



**韦 穗** 1946 年生,1969 年毕业于南京工学院电子工程系,现为安徽大学教授、博士生导师.长期从事计算机视觉、图象处理方面的研究,多次主持过国家 75、85 攻关,“863”计划和国家自然科学基金项目.曾获中国科学院科技进步二等奖.



**罗 斌** 1963 年生,1984 年获安徽大学无线电系工学学士学位,1991 年获安徽大学电子系工学硕士学位,2002 年获英国 York 大学计算机系博士学位,现为安徽大学教授、博士生导师,中国图象图形学会理事.主要研究领域为图象处理与识别、计算机视觉与模式识别.发表论文 60 余篇.



**杨尚骏** 1937 年生,1959 年毕业于四川大学数学系,现为安徽大学教授.长期从事应用数学研究,近年转向计算机视觉方面的研究.曾获中华人民共和国教育部科技进步三等奖(第一完成人).