

利用几何代数进行线段模型匹配和运动估计

黄良明 彭立中 程民德

(北京大学数学科学学院, 北京 100871)

摘要 首先探讨了 Clifford 代数(几何代数)在计算机视觉中的应用,并得到了 2D 与 3D 旋转的统一表达公式,进而探讨了该公式在直线模型匹配和运动估计中的应用.在改进 2D 多角弧匹配算法的基础上,提出了一个同时进行线段模型的匹配和运动估计的算法.该算法通过最小化模型线段与被检测线段间的距离(距离函数定义为对应点间欧氏距离的积分)而求得的最佳运动估计中的旋转,可由一个矩阵的奇异值分解来表示,从而为首次同时解决这两个问题,进行了初步尝试.且该算法不受维数限制.最后的模拟实验结果表明,该算法效果良好.

关键词 线段 运动估计 Clifford 代数 几何代数 匹配 多向量

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2001)03-0270-05

Matching and Estimating Motion of Line Model Using Geometric Algebra

HUANG Liang-ming, PENG Li-zhong, CHENG Min-de

(Department of Mathematics, Peking University, Beijing 100871)

Abstract Matching and estimating motion are basic problem of computer vision. Classical methods are first to find the matching point (or line etc.) and then estimating motion. This paper discussed the application of Clifford algebra (Geometric algebra) in the area of computer vision, presented the uniform formula of 2D and 3D rotation and their application in matching and estimation motion of the line segments model. Based on improving the algorithm of matching 2D polygonal arcs in reference [4], this paper provides an algorithm solve both of matching and estimating motion simultaneously using Clifford algebra. Via minimizing the distance between the model and the detected characteristic (the distance measure is defined as the integral of the Euclidean distance between corresponding points), The algorithm results with that the rotation of the best estimation can be represented by the SVD of a matrix. To our knowledge, this paper is the first investigation to solve both of them. And the algorithm is free from the dimension of the line segment model. Synthetic data has been used to test the algorithm, and excellent result has been obtained.

Keywords Line segment, Estimating motion, Clifford algebra, Geometric algebra, Match, Multivector

0 引言

匹配问题是计算机视觉中的一个基本而重要的问题,许多学者对此做了广泛而深入的研究,但这些研究大部分是基于点与点的匹配.直线段作为视觉信息中最显著而重要的特征之一,在计算机视觉的各个领域均受到重视. Zhang Zhengyou 第一次成功地将直线段用于运动与结构估计^[1]和传统方法相比,

传统方法需要 3 幅图象,而该方法只需要 2 幅就给出了, Ross 则给出了二维线段模型与图象线段匹配的局部搜索算法^[2],该算法在机器人导航及场景理解等方面已得到应用,说明直线段这一基本结构在视觉中的重要作用.另外,还有不少学者对二次曲线乃至高次曲线的运用也作了探索.如封举富利用二次型的可分解理论,研究了平面二次曲线和二次曲面的整体恢复方法^[3],并进一步给出了一种基于二次曲线的整体立体视觉的统一恢复方法和求解算法.

模型匹配作为形状匹配的一部分也是计算机处理的一个重要而基本的课题. 本文提出了一个直线模型并以 Clifford 代数(或几何代数)为工具, 把匹配和运动估计问题转化为求矩阵的奇异值分解问题. 算法中通过最小化模型与被检测特征的距离(其中距离利用分析的方法, 即对应点间欧氏距离的积分)来求得模型与待匹配直线集之间的运动变换, 从而同时解决了匹配与运动估计的问题. 这是作者为同时解决这两个问题所做的首次尝试, 并且该算法不受维数的限制.

Douglas 等也利用同样一种距离函数, 将四元数用于二维多角弧的匹配, 并得到最小距离函数的闭形式解, 其最小距离为一个 2×2 矩阵的特征值, 同时还对三维多角弧的匹配作了研究^[4]. 这里四元数是 Clifford 代数的一个典型特例, 文献[4]中方法实际上也是本文的一个特殊情形.

1 几何代数简介

19 世纪后期产生了许多代数系统. 其中有 Clifford^[1878] 和 Grassmann^[1877] 代数. 虽然当时引起了人们很大的兴趣, 但不久就被物理学家们抛弃了. 因为在他们寻找解时空实问题的代数系统时发现, Gibbs 的向量代数更有用(如在电磁学领域); 后来物理学家们又发现了许多新问题, 由此而产生了旋量 (spinor) 代数, 张量代数等. Clifford 代数是一个结合的非交换代数, 即 $n = 1$ 时为复数; $n = 2$ 时是四元数; 当 $n \geq 3$ 时则是非可除代数. 从 19 世纪后期开始, Clifford 代数被多次引用到各个领域并得到发展和完善^[3]. 本文所用的方法是由 David Hestenes 在本世纪 60 年代建立的, 也称之为几何代数 (Geometric algebra).

1.1 几何积和多向量

几何积和多向量是几何代数中最基本的概念, 两个向量 a 与 b 的几何积定义为

$$ab = a \cdot b + a \wedge b \quad (1)$$

其中, $a \cdot b$ 为 a 与 b 的内积, 其值为实数 $|a| |b| \cos \theta$; $a \wedge b$ 称为 a 与 b 的外积, 具有模 $|a| |b| \sin \theta$, 它是 a 与 b 所在平面的一个有向面积, 也就是将向量 a 沿向量 b 平移所扫过的面积, 称为二向量 (bivector). 即两个向量的几何积为一个纯量和一个二向量之和, 这种形式的量称为多向量.

几何积具有以下性质:

$$a(bc) = (a)b c \quad (\text{结合律})$$

$$a(b+c) = ab + ac \quad (\text{加法分配律})$$

$$(b+c)a = ba + ca \quad (\text{加法分配律})$$

$$a\lambda = \lambda a \quad (\text{数乘})$$

$$a^2 = \pm |a|^2$$

对 n 维空间可以有 0 阶 (纯量), 1 阶 (向量), 2 阶 (二向量)... 直到 n 阶的齐次多向量, 还可以证明两个齐次多向量 A_r 与 B_s 的几何积可以写成

$$A_r B_s = AB_{r+s} + AB_{r+s-2} + \dots + AB_{|r-s|} \quad (2)$$

其中, 齐次多向量 Q_t 表示多向量 Q 的 t 阶成分^[5].

1.2 三维空间的几何代数、基及互反框架

三维空间的几何代数有 $2^3 = 8$ 个基, 其中正交基写成

$$1, \{\sigma_1, \sigma_2, \sigma_3\}, \{\sigma_1\sigma_2, \sigma_2\sigma_3, \sigma_3\sigma_1\}, \sigma_1\sigma_2\sigma_3 = i \quad (3)$$

这里, 三向量 i 称为伪纯量 (trivector), 因为 $i^2 = -1$, 且可与空间中任一多向量交换.

考虑向量 a 对法向量为 n 的平面反射, 可得到向量 $a_1 = -nan$; 再对法向量为 m 的平面反射, 得 $a_2 = mnana = RaR$, 其中, R 是一个二向量 (不一定齐次), 且满足 $RR = 1$. 若考虑单位向量 n_1 经旋转 R , 即可得到单位向量 n_2 , 则由 $n_2 R = R n_1$ 可解得

$$R = \frac{1 + n_2 n_1}{\sqrt{2(1 + n_2 \cdot n_1)}} \quad (4)$$

由式 (4) 可以看出, R 是只含有纯量和二向量的多向量. 由所有的只含偶数阶向量的多向量组成的一个子代数, 称为偶子代数.

若给定一组线性无关向量集 $\{e_i\}$ (不一定正交), 即可建其互反框架 (reciprocal frame) $\{e^i\}$ ^[5], 并使得

$$e^i \cdot e_j = \delta_j^i \quad (5)$$

1.3 对多向量的微分

设 X 为一混合阶数的多向量, 且 $X = \sum_r X_r$, 其中 X_r 为 X 的第 r 阶齐次分量, $F(X)$ 为 X 的多量值函数. 则 $F(X)$ 对 A 方向 '的微分 (A 与 X 同阶), 记作 $A * \partial_X F(X)$ (这里, $*$ 表示两个多向量之积的数量部分), 定义为

$$A * \partial_X F(X) \equiv \lim_{\tau \rightarrow 0} \frac{F(X + \tau A) - F(X)}{\tau} \quad (6)$$

通过引入任意框架 $\{e_j\}$ 并把它扩展成一组基, 而由这组基生成的代数基记作 $\{e_J\}$, 则 ∂_X 可以定义为

$$\partial_X \equiv \sum e^J e_J * \partial_X \quad (7)$$

Lasenby 等人证明,对 ∂_X 有下列等式成立^[6]

$$(1) \partial_X \overline{XB} = P_X(\overline{B}), \quad (8)$$

$$(2) \partial_X \overline{XB} = P_X(\overline{B}), \quad (9)$$

$$(3) \partial_X \overline{XB} = P_X(\overline{B}) = P_X(\overline{B}), \quad (10)$$

$$(4) \partial_\psi M\psi^{-1} = -\psi^{-1}P_\psi(M)\psi^{-1}, \quad (11)$$

$$(5) \partial_a = e^i \frac{\partial}{\partial a^i}, \text{ 如果 } a = a^i e_i \quad (12)$$

多向量微分的应用可看下面例子.

求旋转 R 使得

$$\varphi = \sum_{i=1}^n (v_i - Ru_i \overline{R})^2$$

取最小,由

$$\varphi = \sum_{i=1}^n (v_i - Ru_i \overline{R})^2 = \sum_{i=1}^n \{ (v_i^2 + u_i^2) - 2 v_i Ru_i \overline{R} \}$$

可得

$$\partial_R \varphi(R) = 4 \overline{R} \sum_{i=1}^n v_i \wedge (Ru_i \overline{R})$$

从而 R 必须满足

$$\sum_{i=1}^n v_i \wedge (Ru_i \overline{R}) = 0 \quad (13)$$

为解式(13)本文引入了代数上的线性函数.

1.4 线性函数

设 $f(a)$ 是向量空间到向量空间的线性函数,定义 f 的扩张 $\tilde{f}$ 为

$$\tilde{f}(a_1 \wedge a_2 \wedge \dots \wedge a_r) = \tilde{f}(a_1) \wedge \tilde{f}(a_2) \wedge \dots \wedge \tilde{f}(a_r) \quad (14)$$

其伴随 $\tilde{f}(a) = e^i \tilde{f}(e_i) a$. 如果 $f = \tilde{f}$ 则称 f 是对称的. Hestenes 证明了 f 对称 $\Leftrightarrow \partial_a \wedge \tilde{f}(a) = 0$

由此可将式(13)写成

$$\partial_a \wedge \left[\sum_{i=1}^n R(a \cdot v_i) u_i \overline{R} \right] = 0 \quad (15)$$

定义 $\tilde{f}(a) = \sum_{i=1}^n (a \cdot v_i) u_i$, 则有 $\partial_a \wedge R \tilde{f}(a) \overline{R} = 0$.

定义 $\underline{Ra} = Ra \overline{R}$, 则有 $\partial_a \wedge R \tilde{f}(a) = 0$, 从而 $\underline{G} = \underline{R} \tilde{f}$ 对称. 由于 $R^{-1} = \overline{R}$, 可得 $\underline{R}^{-1} = \overline{R}$, 从而可得 $\tilde{f}(a) = \overline{R} \underline{G}(a)$.

若将 $\tilde{f}$ 作奇异值分解 $\tilde{f} = \underline{U} \underline{S} \underline{V}$, 则有 $\tilde{f} = \underline{U} \underline{V} (\overline{V} \underline{S} \underline{V}) = \overline{R} \underline{G}$, 所以 $\underline{R} = \overline{V} \underline{U}$.

$\tilde{f}$ 的矩阵形式为: $F = (F_{\alpha\beta})$ 其中

$$F_{\alpha\beta} = \sigma_\alpha \cdot \tilde{f}(\sigma_\beta) = \sum_{i=1}^n (\sigma_\alpha \cdot u_i) \tilde{f}(\sigma_\beta \cdot v_i) \quad (16)$$

2 问题的提出与解决

将上述理论用于直线段模型的匹配与运动估计. 考虑直线段模型, 其中每条直线段用它的中点向量 c 、长度 u 、方向向量 v 组成的三元组 (c, u, v) 表示, 则模型 A 可用以下三元组集合表示

$$A = \{ (c_1, u_1, v_1), (c_2, u_2, v_2), \dots, (c_n, u_n, v_n) \} \quad (17)$$

设在图象中检测到的线段集 Γ 为

$$\Gamma = \{ (c'_1, u'_1, v'_1), (c'_2, u'_2, v'_2), \dots, (c'_n, u'_n, v'_n) \} \quad (18)$$

利用对应点间距离的积分作为测量函数, 则匹配问题转化为在 A 中寻找 $n/2 < m \leq l$ 条线段 (考虑半数以上的线段匹配), 使得下式距离误差最小

$$M(A, \Gamma) = \sum_{i=1}^m \left\{ l(c_i - Rc'_i \overline{R} - t)^2 + \frac{l_i^3}{12} (v_i - Rv'_i \overline{R})^2 \right\} \quad (19)$$

其中, R 为转子 (rotor), t 为平移向量, $l_i = \min(u_i, u'_i)$.

因为

$$\frac{\partial M}{\partial t} = \frac{\sum_{i=1}^m l(c_i - Rc'_i \overline{R})}{\sum_{i=1}^m l_i}$$

并且, 要使 M 取得最小的 R 必须满足

$$\sum_{i=1}^m \left[l(c_i - t) \wedge Rc'_i \overline{R} + \frac{l_i^3}{12} v_i \wedge Rv'_i \overline{R} \right] = 0 \quad (20)$$

设

$$\tilde{f}(a) = \sum_{i=1}^m \left[l(a \cdot (c_i - t)) c'_i + \frac{l_i^3}{12} (a \cdot v_i) v'_i \right]$$

$\underline{U}, \underline{V}$ 为 $\tilde{f}$ 的奇异值分解, 即 $\tilde{f} = \underline{U} \underline{S} \underline{V}$, 其中 $\underline{U}, \underline{V}$ 是正交阵, $\underline{S}$ 是对角的, 则旋转矩阵 $R = \overline{V} \underline{U}$. 由此可得如下匹配算法.

算法 1

输入: 直线段模型 A (n 条线段) 检测到的直线段集 A' (m 条线段).

输出: 判断 A' 是否有 A 的一部分, 若有, 则同时输出它们之间的运动变换.

第 1 步: 令 $l_{\min} = \min\{m, n\}$, $l = l_{\min}$.

第 2 步: 选定 A' 中 l 条线段;

第 3 步: 置初值 $t_0 = \frac{\sum_{i=1}^l u_i (c_i - c'_i)}{\sum_{i=1}^l u_i}$, $t = t_0$;

第 4 步: 将 t 代入 $\tilde{f}$ 求得 $R = \overline{V} \underline{U}$;

第 5 步: 将求得的 R 代入 $\frac{\partial M}{\partial t} = 0$ 求得 t ;

第 6 步:判断收敛准则;若满足,则计算 M ,若 M 足够小,则输出匹配,否则转到第 7 步;若不满足,则回到第 4 步;

第 7 步:选定新的 l 条线段,回到第 3 步;若没有新的线段可选,则转到第 8 步;

第 8 步: $l=l-1$,若 $l>l_{\min}/2$,则回到第 2 步,否则输出不匹配。

3 模拟实验结果

为了测试上面的算法是否有效,还作了一些有代表性的模拟实验。

首先,进行简单的二维直线段模型匹配实验。直线段模型为一个等腰直角三角形(图 1 中的模型 I)。待匹配线段集 J 中有一个三角形与 I 匹配,其余的线段组成的两个三角形则与 I 不匹配。为论述方便起见,把模型和待匹配图形在同一图中画出。其中,模型数据与待匹配数据如下:

$I = \{ \langle (4.5, 4), 1 \rangle, \langle (1, 0), 1 \rangle, \langle (5, 4.5), 1 \rangle, \langle (0, 1), 1 \rangle, \langle (4.5, 4.5), 1.414 \rangle, \langle (0.707, 0.707), 1.414 \rangle \}$

$J = \{ \langle (1.5, 2), 1 \rangle, \langle (1, 0), 1 \rangle, \langle (2, 1.5), 1 \rangle, \langle (0, -1), 1 \rangle, \langle (1.5, 1.5), 1.414 \rangle, \langle (-0.707, 0.707), 1.414 \rangle; \langle (3.5, 3), 1 \rangle, \langle (1, 0), 1 \rangle, \langle (4, 4), 2 \rangle, \langle (0, 1), 1 \rangle, \langle (3.5, 4), 2.236 \rangle, \langle (-0.4472, -0.8944), 2.236 \rangle; \langle (7.5, 1.5), 1.414 \rangle, \langle (0.707, 0.707), 1.414 \rangle, \langle (7.75, 2.5), 1.118 \rangle, \langle (0.8945, -0.4472), 1.118 \rangle, \langle (7.25, 2), 2.0616 \rangle, \langle (-0.9701, -0.2425), 2.0616 \rangle \}$

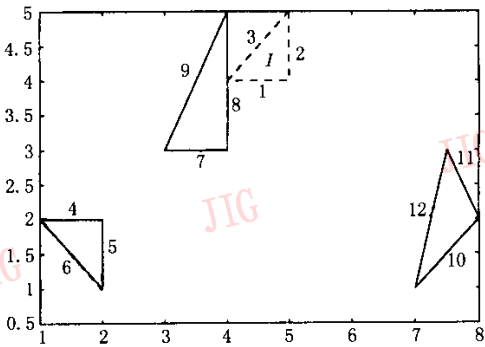


图 1
图中虚线三角形为 I , 实线为待匹配线段集 J ,
左下角小三角形为与 I 匹配的三角形及经变换后的 I

实验中,所取的收敛准则是相邻两次迭代所得的 R_i 与 R_{i+1} 之差的 1-模小于 10^{-4} ,且 t_i 与 t_{i+1} 的 1-模小于 10^{-6} ,而所取的判别准则是经变换后模型与待匹配线段集之间的距离误差 $M \leq 10^{-3}$ 。

实验中,待匹配的 9 条线段集组成 3 个三角形,其中街匹配线段集 J 中,由线段 4、5、6 组成的实线

三角形与 I (图 1 右上方的由线段 1、2、3 组成的虚线三角形)相匹配,而由线段 7、8、9 组成的大三角形中,由于线段 7 和线段 1 相等,并且线段 7 与 8 的夹角也和线段 1 与线段 2 的夹角相同,因而匹配过程中也会形成一个极值点,但此时得到的 M 值较大,不满足判断准则。

经算法 1 计算后得到

$$R = \begin{bmatrix} 0.0222 & 0.9998 \\ -0.9998 & 0.0222 \end{bmatrix}, \quad t = [5.8953 \quad -3.0951]$$

两者匹配后结果如图 1 所示,从图中可以看出 I 经过旋转和平移后,基本与由线段 4、5、6 组成的三角形重合。

由线段 1、2、3 组成的三角形到由线段 4、5、6 组成的三角形的变换,其真实值为

$$R = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad t = [6 \quad -3]$$

由此可以看出,实验结果与正确结果的误差较大。这是因为一般软件中,实矩阵的特征分解算法用的是迭代法(目前也没有更好的算法),而迭代法本身的计算误差较大,再加上乘法带来的累积误差,因此使得计算得到的结果出入较大。

当平面形物体从一个位置运动到另外一个位置时,本模拟实验可以判断从两幅图象中检测到的物体是否为同一物体。若是同一物体,则可以同时求出它们之间的运动关系。

对三维线段模型,我们利用文献 [6] 的数据(如图 2)。为了验证上面的算法的效果,先将模型绕 z 轴旋转 30° ,然后再将平移后所得到的数据作为待匹配数据。其中,旋转矩阵和平移向量的真实值为

$$R = \begin{bmatrix} 0.866 & 0.5 & 0 \\ -0.5 & 0.866 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad t = (1, 1, 1)$$

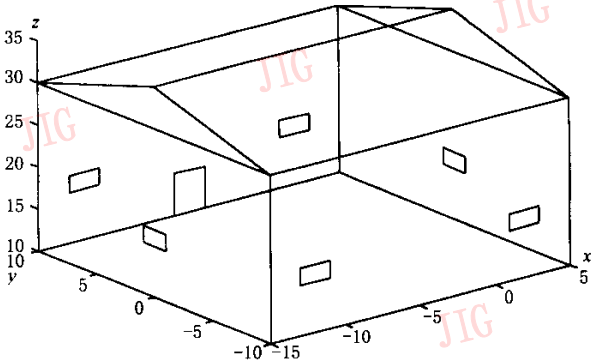


图 2 由 44 条线段组成的房屋框架线段模型

由于实验数据由模型数据计算得到,理应全部匹配.表 1 显示了算法过程中判断准则距离误差 M 、迭代次数 n 、求得的平移向量 t 之间的关系.从实验可以看到,随着迭代次数的增加,误差 M 逐渐减小,但在误差减小到 10^{-4} 以后, M 即很难再得到改进,其原因在于该模型的线段较多,且由一点匹配误差所带来的所有线段的对应点间的距离积分之和也不小.

表 1 误差 M 、迭代次数 n 、平移向量 t 之间的关系

M	n	t
10	21	(0.938 3 ,1.267 4 0.883 7)
0.1	236	(0.970 1 ,1.128 6 0.937 9)
0.001	1 147	(0.987 9 ,1.053 4 0.989 9)
0.000 1	4 360	(0.993 8 ,1.017 3 0.995 3)

本文也做了文献 6 的实验,相比之下,同样是迭代算法,本文的算法会慢一些,这一方面是因为文献 6 利用的是顶点坐标信息,而本文用的是整条线段的信息;另一方面,文献 6 是在已知匹配的情况下求运动估计,而本文的方法是同时判断匹配与否,并求运动估计.随着计算机硬件速度的提高,对运算量的限制越来越低,从而,原来认为海量计算的不能解决的问题变得容易解决.

4 结 论

本文初步探讨了 Clifford 代数在计算机视觉中的初步应用,并提出了由几何代数所表示的二维和三维旋转的统一求解公式及其在直线模型匹配中的应用,第一次把匹配问题与运动估计问题结合起来同时解决,而且利用几何代数作为工具的另一优点是,可以用同一种方法解决不同维数的问题.本文用此方法解决了二维和三维直线模型的匹配问题,并同时估计了它们之间的相对运动.由于任何曲线都可以用直线来近似,所以,这方法有希望用来作形状匹配,但因为用直线近似曲线时误差较大,因此在匹配时还需处理许多细节,这也是下一步将展开的工作.

致 谢 特别感谢北京大学科瑞奇图象处理实验室对本文研究的资助;特别感谢北京大学科瑞奇图象处理实验室主任张绪定教授对本文的帮助.

参 考 文 献

1 Zhengyou Zhang. Estimating motion and structure from correspondences of line segments between two perspective images. IEEE PAMI, 1995, 17(12):11291139.

2 Ross Beveridge J, Edward M Riseman. How easy is matching 2D line models using local search? IEEE PAMI 1997, 19(6) 564579.

3 封举富. 计算机视觉中若干问题的研究 博士研究生学位论文]. 北京 北京大学, 1997.

4 Douglas R. Heisterkamp and prabir bhattacharya. Matching 2D Polygonal Arcs by Using a Subgroup of the Unit Quaternions. Computer Vision and Image Understanding, 1998, 69(2) 246249.

5 David Hestenes, Garret Sobczyk. Clifford algebra to geometric calculus. D.Reidel Publishing Company, 1985.

6 Lasenby J, Lasenby A N, Doran C J L *et al.* New geometric methods for computer vision: An application to structure and motion estimation. International journal of computer vision, 1998, 26(3):191215.

黄良明 1972 年生, 1995 年获北京大学数学系学士学位, 现为北京大学数学科学学院博士研究生. 主要研究方向为计算机视觉、模式识别、Clifford 代数在计算机视觉中的应用等.

彭立中 1944 年生, 教授, 博士生导师, 曾获北京大学数学系学士学位、瑞典斯得哥尔摩大学数学系博士学位, 现任北京大学数学科学学院基础数学系主任. 主要研究方向为调和分析、小波分析等.

程民德 1917 年生, 中国科学院院士, 1940 年和 1942 年先后在浙江大学数学系获学士学位和硕士学位, 1949 年在美国普林斯顿大学数学系获博士学位, 曾在《数学年刊》(Annals of Mathematics)上发表论文 3 篇, 多次获得各种国家科技进步奖. 主要研究方向为调和分析、逼近理论、计算机视觉、模式识别、计算语言学、数学机械化证明等.