基于遗传算法的射影重构

梁 栋 刘春梅 王 年 韦 穗
(安徽大学计算智能与信号处理教育部重点实验室, 合肥 230039)

摘 要 在实现分层重构的过程中, 射影重构是关键的第 1 步. 目前, 大多已有算法对模拟数据是非常有效的, 但对于真实图象效果并不理想. 为了寻求更为鲁棒的算法, 提出了一种基于遗传算法的射影重构算法. 该算法对于射影深度采用十进制编码, 并以测量矩阵的秩为 4 作为约束, 来定义适应度函数, 然后利用遗传算法, 并结合奇异值分解(SVD) 技术来迭代估计射影深度, 进而实现射影重构. 实验表明, 该算法是行之有效的, 且鲁棒性较好.

关键词 射影重构 遗传算法 奇异值分解
中图法分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2002)05-0445-05

The Projective Reconstruction Based on Genetic Algorithms

LIANG Dong, LIU Chun-mei, WANG Nian, WEI Sui
(*Educational Department Key Lab. of IC&SP, Anhui University, Hefei* 230039)

Abstract In computer vision, it is called multi-views 3D reconstruction for recovering both camera and object shapes from multiple images, and it is currently a topic of lively interest. A hierarchical reconstruction method had introduced in 1996. In the course of completing the hierarchical reconstruction, the projective reconstruction is the first key step, having very important effect on the precision of euclidean reconstruction. The existence methods are very efficient for simulation data, but they are not perfect for real image. Namely, they are not robust and the reliable results can only be obtained if images match accurately. In this paper, the projective reconstruction based on genetic algorithms is proposed, the projective depths are coded by using decimal system and the adaptability function is defined by a constraint of the measurement matrix rank 4. The projective depths are iteratively estimated by genetic algorithms and Singular value decomposition (SVD) so that the measurement matrix is made to be as close as possible to rank 4, and then the projective reconstruction is realized. The validity and robusticity of the proposed algorithm is confirmed by experiments.

Keywords Projective reconstruction, Genetic algorithms, Singular value decomposition(SVD)

0 引 言

在计算机视觉领域中, 由多幅图象来恢复摄像机运动参数和空间物体表面的 3D 几何形状问题称为多视点 3D 重构. 1996 年 Frangois 提出了一种分层重构思想[1], 即首先从未标定图象序列进行 3D 射影重构, 然后在给出进一步信息(如摄像机内参数) 的情况下, 从射影重构恢复欧氏重构. 由于在实现分层重构的过程中, 射影重构是关键的第 1 步, 它对欧氏重构结果的精度起着决定性的作用, 因此目前许多研究者都把射影重构作为主要研究对象, 并提出了许多新的方法和技术, 其中, 经理论和数值模拟实验证明矩阵分解方法是非常有效的, 也是近年来研究的热点. 矩阵分解算法[2] 由 Tomasi 和 Kanade 于 1992 年首先提出, 该算法是在正交投影摄像机模型基础上实现的, 后来又被 Poelman 和 Kanade 扩展到弱透视和准透视投影模型[3]. 这些算法的显著特点是可以在仿射模型下, 将由二维图象点坐标构建的测量矩阵分解为表示摄像机运动和空

基金项目: 国家自然科学基金(60075004)
收稿日期: 2001-11-30; 改回日期: 2002-02-27

间物体 3D 几何形状的两个矩阵, 这种分解称为仿射分解(affine factorization). 因为弱透视和准透视投影模型不符合通常意义下的摄像机模型(透视模型), 所以该算法只能在特殊情况下使用. 近年来, 又出现了一些基于透视摄像机模型的矩阵分解算法(称为透视分解 perspective factorization), 透视分解和仿射分解的本质区别在于与每个图象点对应的射影深度未知, 故测量矩阵也未知, 这就意味着不能用传统的仿射分解算法来获取数据. 由此可见, 在透视分解算法中, 最关键的问题是如何估计射影深度. 1996 年 Christy 和 Horaud 提出了一种由透视投影图象来恢复空间物体形状的算法[4], 该算法从准透视摄像机模型开始, 通过测量矩阵的迭代分解来估计射影深度. 1996 年 Sturm 和 Triggs 对于未标定透视投影图象提出了一种非迭代分解算法[5], 该算法是通过基础矩阵(Fundamental matrix) 来计算射影深度. 由于这些算法对图象之间对应点的定位误差相当敏感, 其虽对于模拟数据可以得出很好的结果, 但对于真实图象效果并不理想, 即这些算法不鲁棒, 必须精确匹配才能获得可靠的结果. 因此本文提出一种基于遗传算法的射影重构算法. 该算法对于射影深度是首先采用十进制编码, 并以测量矩阵的秩为 4 作为约束来定义适应度函数; 然后利用遗传算法, 并结合奇异值分解(SVD) 技术来迭代估计射影深度, 进而实现射影重构. 实验结果表明, 该算法是行之有效的, 且与已有算法相比, 鲁棒性较好.

1 遗传算法原理

遗传算法是模拟生物进化过程的计算模型. 遗传算法作为一种新的全局优化搜索算法, 以其简单通用、鲁棒性强、适于并行处理以及应用范围广等显著特点, 奠定了它作为 21 世纪关键智能计算技术之一的地位.

在问题解决的过程中, 遗传算法作为一种迭代算法, 它可从一组变量中搜寻一个最优解. 对于一个具体的问题, 用遗传算法来搜寻最优解的过程有如下 4 个主要步骤:

(1) 编码　在候选解的范围内, 可随机选择一个候选解作为初始解, 且每一个候选解都可用一个序列表示, 它们均可以用二进制或者十进制编码, 例如, 对于二进制编码, 8 被编码成 00001000, 而被编码后的候选解被称为个体.

(2) 选择　根据每个个体的适应度函数的值, 并基于一个确定的概率来选择一些个体作为父代, 一个个体的适应度函数值越大, 被选择的概率就越大.

(3) 交叉　从当前交配池中, 基于一个确定的概率随机地选择一些个体作为下一代的父代, 每两个父代个体再按一定的概率通过交叉, 产生两个新的个体, 这两个新的个体既保留了它们父代的部分基因(它们的一部分编码未变), 也引进了一些新的基因.

(4) 变异　基于一个确定的概率, 随机地选择一些个体的部分基因, 并且把它们变成新的基因. 例如, 对于二进制编码, 将一个初始位元为 1 的位变成 0, 而将初始位元为 0 的位变成 1; 对于十进制编码, 即在设定的范围内, 随机地选择一个新的数值来替代原来的基因. 变异在遗传算法中非常重要, 虽然变异的概率很小, 但在解决问题的过程中, 却能够防止因早熟收敛所引起的局部最优解.

与传统优化算法相比较, 遗传算法具有如下特征:

(1) 最优解搜寻的过程不影响变量, 但是变量的编码对寻优过程会产生影响;

(2) 搜寻最优解的过程是从一组解到另外一组解的过程, 而不是从一个解到另外一个解的过程, 由于这个过程能够防止局部收敛, 因此它的全局收敛能力较强;

(3) 对搜寻空间没有特别的要求, 由于它仅仅利用适应值, 而不需要辅助的信息, 因此它的应用范围很广泛.

2 基于遗传算法的射影重构

假设 P 个三维空间物体点通过 F 个 3× 4 射影矩阵 $\boldsymbol{Q}_i(i=1,2,\cdots,F)$ 投射到 F 幅透视图象上, 其中, $\boldsymbol{x}_j(j=1,2,\cdots,P)$ 为第 j 个三维空间物体点, 它在第 $i(i=1,2,\cdots,F)$ 幅图象上的象点为 $\boldsymbol{m}_{ij}$. 从 P 个三维空间物体点中任取一点 $\boldsymbol{x}$ 作为参考点, 令它在第 i 幅图象上的象点为 $\boldsymbol{m}_i$. 将世界坐标系原点平移到 $\boldsymbol{x}$, 并将每一个图象坐标系原点平移到 $\boldsymbol{m}_i$, 记: $\Delta\boldsymbol{m}_{ij}=\boldsymbol{m}_{ij}-\boldsymbol{m}_i$, $\Delta\boldsymbol{x}_j=\boldsymbol{x}_j-\boldsymbol{x}$, . 引入齐次坐标向量 $\underline{\boldsymbol{x}}_j=\left|\Delta\boldsymbol{x}_j^{\mathrm{T}}\quad 1\right|^{\mathrm{T}}$ 和 $\underline{\boldsymbol{m}}_{ij}=\left|\Delta\boldsymbol{m}_{ij}^{\mathrm{T}}\quad 1\right|^{\mathrm{T}}$, 分别表示空间物体点和其象点的齐次坐标. 这样就可以得到射影方程

$$\mu_{ij}\,\underline{\boldsymbol{m}}_{ij}=\boldsymbol{Q}_i\underline{\boldsymbol{x}}_j \tag{1}$$

其中, μ_{ij} 表示射影深度. 那么, 对于 F 幅图象和 P 个空间物体点, 则有

$$\underbrace{\begin{bmatrix} \mu_{11}\boldsymbol{m}_{11} & \cdots & \mu_{1P}\boldsymbol{m}_{1P} \\ \vdots & \ddots & \vdots \\ \mu_{F1}\boldsymbol{m}_{F1} & \cdots & \mu_{FP}\boldsymbol{m}_{FP}\end{bmatrix}}_{W} = \underbrace{\begin{bmatrix}\boldsymbol{Q}_1\\ \vdots \\ \boldsymbol{Q}_F\end{bmatrix}}_{Q}\underbrace{\begin{bmatrix}\boldsymbol{x}_1 & \cdots & \boldsymbol{x}_P\end{bmatrix}}_{X} \tag{2}$$

其中, $3F\times P$ 矩阵 $\boldsymbol{W}$ 称为测量矩阵. 式(2)表明, $\boldsymbol{W}$ 可以分解为一个表示摄像机运动的 $3F\times P$ 矩阵 $\boldsymbol{Q}$ 和一个表示空间物体形状的 $4\times P$ 矩阵 $\boldsymbol{X}$. 在一般情况下, 由于 μ_{ij} 未知, 从而 $\boldsymbol{W}$ 也未知, 其分解式(2)更不得而知.

如果由某种方法给出了正确的射影深度 μ_{ij}, 则可将 $\boldsymbol{W}$ 作奇异值分解(SVD), 得

$$\boldsymbol{W} = \boldsymbol{V}\,\mathrm{diag}(\sigma_1, \cdots, \sigma_N)\,\boldsymbol{U}^{\mathrm{T}} \tag{3}$$

式中, $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_N\geq 0$ 为 $\boldsymbol{W}$ 的奇异值, $N=\min(3F, P)$, $\boldsymbol{V}$、$\boldsymbol{U}$ 分别为 $3F\times N$、$P\times N$ 的次正交矩阵. 如果图象坐标无噪声(即 F 幅图象精确匹配), 则 $\boldsymbol{W}$ 的秩为 4, 即 $\sigma_5=\sigma_6=\cdots=\sigma_n=0$, 从而 $\boldsymbol{V}_1\mathrm{diag}(\sigma_1, \cdots, \sigma_4)\boldsymbol{U}_1^{\mathrm{T}}$ 为 $\boldsymbol{W}$ 的一个满秩分解, $\boldsymbol{V}_1$ 和 $\boldsymbol{U}_1$ 分别为 $\boldsymbol{V}$ 和 $\boldsymbol{U}$ 的前 4 列. 结合式(2)可得

$$\boldsymbol{Q} = \boldsymbol{V}_1, \quad \boldsymbol{X} = \mathrm{diag}(\sigma_1, \cdots, \sigma_4)\boldsymbol{U}_1^{\mathrm{T}} \tag{4}$$

但此满秩分解不是唯一的, 因为式(2)在相差任意非零常数因子的情况下, 可以写为

$$\boldsymbol{W} = \boldsymbol{Q}\boldsymbol{H}\boldsymbol{H}^{-1}\boldsymbol{X} \tag{5}$$

其中, $\boldsymbol{H}$ 为任意非奇异 4×4 矩阵, 所以 $\boldsymbol{Q}'=\boldsymbol{Q}\boldsymbol{H}$, $\boldsymbol{X}'=\boldsymbol{H}^{-1}\boldsymbol{X}$ 均为 $\boldsymbol{W}$ 的分解. 这就表明, 可以在一个未知射影变换下来恢复摄像机运动和空间物体形状, 并可以任意选择一个 3D 射影坐标系来描述这种运动和形状, 这种恢复就称为射影重构.

由此可见, 实现射影重构的关键是获得正确的射影深度. 下面利用遗传算法, 并结合奇异值分解技术来估计射影深度.

(1) 解的编码与适应度函数

对于参数 $\mu_{ij}(i=1, \cdots, F,\ j=1, \cdots, P)$ 采用十进制编码. 令 $\{\mu_{ij}\}$ 是射影深度的一个解, 以 $\tilde{\mu}=\{\mu_{11}, \cdots, \mu_{1P}, \cdots, \mu_{F1}, \cdots, \mu_{FP}\}$ 作为一个个体进行编码. 由于透视投影模型可以利用仿射投影模型来迭代逼近[6], 且仿射投影模型下的仿射深度是一个常量(通常取 1), 因此 $\{\mu_{ij}\}$ 可以在(0, 2]区间内随机产生.

在图象坐标无噪声, 且射影深度为正确值的理想情况下, 测量矩阵 $\boldsymbol{W}$ 的秩 $\mathrm{rank}\boldsymbol{W}=4$, 当有噪声, 但不大时, $\mathrm{rank}\boldsymbol{W}\approx 4$. 据此可以定义遗传算法的适应度函数为

$$f(\tilde{\mu}) = \frac{1}{1+\sigma_5} \tag{6}$$

$f(\tilde{\mu})$ 是依赖于 $\tilde{\mu}$ 的非负实数, 为使 $\mathrm{rank}\boldsymbol{W}=4$, 必有 $\sigma_5=0$, 于是, 求使 $f(\tilde{\mu})$ 取最大值的 $\tilde{\mu}$, 即是所需要的正确射影深度. 由于 $f(\tilde{\mu})$ 不是线性函数, 故最大化 $f(\tilde{\mu})$ 是一个非线性优化问题.

(2) 基于遗传算法的射影深度估计

射影深度估计算法流程如下:

① 从 F 幅图象中选择 P 个特征对应点, 并提取它们的图象坐标 $\boldsymbol{m}_{ij}(i=1, \cdots, F, j=1, \cdots, P)$, 然后, 在其中任取一点 $\boldsymbol{m}_i$ 作为图象坐标原点, 并计算 $\Delta\boldsymbol{m}_{ij}=\boldsymbol{m}_{ij}-\boldsymbol{m}_i$;

②对图象进行归一化, 即每一幅图象中的 $\Delta\boldsymbol{m}_{ij}$ 乘上一公共因子 s_i, $s_i = \sqrt{2}\Big/\frac{1}{P}\sum_{j=1}^{P}\|\Delta\boldsymbol{m}_{ij}\|_2$ $(i=1, \cdots, F)$, 使它们的平均范数为 $\sqrt{2}$;

③在(0, 2]区间内, 随机选取个体 $\tilde{\mu}$, 取交配池的容量为 20;

④对与每一个个体 $\tilde{\mu}^{(k)}$ 对应的测量矩阵 $\boldsymbol{W}(\tilde{\mu}^{(k)})$ 进行奇异值分解(SVD), 并由适应度函数公式来计算每个个体的适应度 $f(\tilde{\mu}^{(k)})$, $k=1, \cdots, 20$;

⑤每次进行遗传操作, 均以概率 $f(\tilde{\mu}^{(k)})\Big/\sum f(\tilde{\mu}^{(k)})$ 复制;

⑥以概率 p_c 用赌轮法随机选取两个个体, 并对其进行交叉操作;

⑦以概率 p_m 进行变异操作;

⑧把经过遗传操作后得到的个体都放在交配池中. 当个体的个数超过交配池容量时, 就将适应度小的个体从交配池中删去;

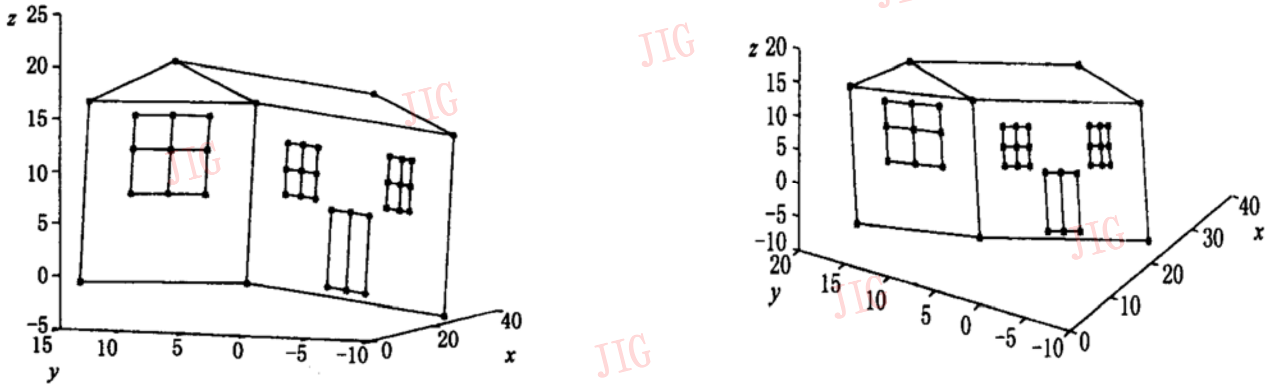
⑨进行上述遗传操作至第 s 代后(s 是预先给定的常数), 再以预先给定的常数 $a\%$ 来对交配池中适应值小的个体进行删除, 代之以新的个体;

⑩进行上述遗传操作至第 t 代后(t 是预先给定的常数). 在第 t 代的交配池中取适应度最大的个体, 即为所求的个体.

3　实验结果

本实验用一架数码相机从 3 个不同的视点拍摄 3 幅图象(如图 1 所示), 实验时, 首先从 3 幅图象中提取 41 个特征对应点, 并利用本文算法来实现射影重构; 然后在摄像机内参数已知的情况下, 求解一个满足欧氏重构条件的 4×4 非奇异矩阵, 最后通过此矩阵将射影重构变换为欧氏重构. 重构结果如图 2 所示. 实验用

图 1 从 3 个不同的视点拍摄的 3 幅原始图象



(a) AZ: − 68°, EL: 6° (b) AZ: − 60°, EL: 38°

图 2 重构结果(AZ 为俯视角, EL 为方向角)

参数 $p_c = 0.7$, $p_m = 0.02$, $s = 100$, $a = 20$, $t = 1\,500$, 摄像机内标定参数为:

$$K = [(1\,369.7, 0, 0)^T, (-3.3, 1\,376, 0)^T, (-38.6, -16.9)^T]$$

本文算法是利用 Matlab 5.3 编程实现的, 从结果上看, 用该算法可以完成所选特征点的射影重构和欧氏重构, 从而为进一步研究真实场景和虚拟物体的无缝合成以及场景漫游打下了良好的基础.

本文的目的是为了尝试利用遗传算法来解决计算机视觉领域中的一些优化问题, 以寻求更为鲁棒的算法. 通过模拟数据实验, 本算法的全局收敛概率, 即全局收敛次数与实验次数(本实验共进行 10 次) 的比值为 0.9, 这是可以接受的. 当对应点噪声较大时, 本算法的迭代时间虽较长, 但通过模拟数据实验发现, 与已有算法相比, 本算法的鲁棒性较好, 如文献[5] 中提出的通过基础矩阵及极点来计算射影深度的算法, 当图象匹配平均误差大于 0.5 个象素时, 则计算误差较大, 而本算法, 当图象匹配平均误差大于 1.5 个象素以上时, 仍然可以获得较为满意的结果, 其原因是基础矩阵, 尤其是极点对匹配误差非常敏感, 由于本算法不涉及基础矩阵和极点, 故鲁棒性较好.

优化问题是计算机视觉领域中一个非常重要的问题, 也是一个难题. 本文算法除有待进一步改进之外, 今后需进一步研究其他优化算法, 以期获得更好的结果.

参 考 文 献

1 Frangois G. Hierarchical visual perception with calibration[R]. Report of research, INRIA, France, 1996, No: 3002.

2 Tomasi C, Kanade T. Shape and motion from image streams under orthography: a factorization method[J]. International Journal of Computer Vision, 1992, 9(2): 137~154.

3 Poelman C, Kanade T. A paraperspective factorization method for shape and motion recovery[A]. In: Proc. ECCV'3[C], Stockholm, 1994, 2: 97~108.

4 Christy C, Horaud R. Euclidean reconstruction: from paraperspective to perspective[A]. In: Proc. ECCV'4[C], Cambridge, 1996, 2: 129~140.

5 Sturm P, Triggs B. A factorization based algorithm for multi-image projective structure and motion[A]. In: Proc. ECCV'4[C], Cambridge, 1996, 2: 709~720.

6 Toshio U, Fumiaki T. A factorization method for projective and Euclidean reconstruction from multiple perspective views via iterative depth estimation[A]. In: Proc. ECCV'5[C], Freiburg, 1998, 1: 296~310

7 刘勇, 康立山, 陈毓屏. 非数值并行算法——遗传算法[M]. 北京: 科学出版社, 1995.

梁　栋　1963 年生, 1985 年毕业于安徽大学电子工程与信息科学系, 1990 年在该校获硕士学位. 现为安徽大学副教授, 在读博士. 主要研究领域为计算机视觉、图象处理和模式识别. 已发表学术论文 20 多篇.

王　年　1966 年生, 1986 年毕业于安徽大学电子工程与信息科学系, 1998 年在该校获硕士学位, 现为安徽大学副教授. 主要研究领域为计算机视觉、图象处理等, 已发表学术论文 10 篇.

刘春梅　1978 年生, 硕士研究生. 主要研究领域为计算机视觉、模式识别和人工智能. 已发表学术论文 3 篇.

韦　穗　1946 年生, 现任安徽大学副校长、教授、博士生导师. 主要研究领域为计算机视觉、图象处理、模式识别和人工智能等. 主持并完成过多项国家和省部级科研项目, 已发表学术论文 20 多篇.

《中国图象图形学报》第三届编辑委员会换届通知

各位专家, 学者, 研究人员:

《中国图象图形学报》编委会将于今年换届, 为了把《中国图象图形学报》办成高水平的学术刊物, 本刊主编徐冠华曾指出, 没有一流的人才, 办不成一流的学报. 在此, 趁《学报》编委会换届之机, 本刊欲从广大科技工作者和专家、学者中评聘编委, 希望热心科技事业, 有奉献精神的专家、学者踊跃自荐或推荐.

条件: 有深厚的学术修养和较高的学术水平;

有丰富的科研经验和敏锐的学术洞察力;

了解学术研究的前沿, 有创新思想和能力;

能完成编委担负的工作任务(约稿, 审稿等) .

有意担任编委者, 如没收到自荐表, 请来函、来电或发电子邮件(主题写明索取自荐表) 索取自荐表. 自荐表返回截止日期为 2002 年 7 月 15 日.

JIG