

文章编号: 1002-0446(2001)05-0455-04

立体视觉中的三目匹配方法研究*

任 重 邵军力

(解放军理工大学 计算机与指挥自动化学院 南京 210016)

摘 要: 本文提出了利用金字塔图传播匹配的三目基元立体视觉方法. 该方法中匹配起始三元组的选择建立在透视投影的可靠基础之上, 而进行传播匹配可以加速计算, 并避免了可能的误匹配. 传播匹配利用立体图像组在二值边界图像金字塔图中的结构相似性, 而整个匹配结果的考察与认定是通过传播率的比较实现的. 该方法能够有效地克服双目立体视觉方法的某些本质局限性, 确保匹配结果精确可靠.

关键词: 立体视觉; 特征基元; 匹配

中图分类号: TP24 **文献标识码:** B

RESEARCH ON TRINOCULAR MATCHING METHODS FOR STEREO VISION

REN Zhong SHAO Jun-li

(Institute of Computer Science and Command Automation, PLAUST, Nanjing, 210016)

Abstract In this paper a trinocular stereo vision method using pyramidal graph propagation matching is proposed. The choice of the starting triplet in the algorithm is based on perspective projective geometry and matching through propagation greatly limits the searching scope. Therefore, it can accelerate the computing and avoid the possible false targets. Matching through propagation makes use of the structural similarities of stereo triplet images in pyramidal graphs of BLI's, whereas the overall examination and verification of matching results is done by a comparison of the propagation success ratio. This algorithm can effectively overcome some essential limitations of binocular stereo vision methods and assure the matching results precise and reliable.

Keywords: stereo vision, feature primitives, matching

1 引言(Introduction)

长期以来, 国际上立体视觉的研究主要集中在双目立体视觉上, 并已取得了大量的研究成果. 然而, 双目立体视觉有其自身的局限性, 如所有立体视觉方法均毫不例外地基于图像之间相似性的先验假设, 而这些假设在实际情况下一般仅有约束地、近似地成立; 双目立体视觉的精度依赖于景物的深度及摄像机之间的距离(基线长度), 基线越长, 精度越好. 然而, 当增大基线长度时, 图像之间的透视形变的差别也越来越大, 从而给利用相似性进行匹配带来了困难; 另外, 当特征线段与基极线(Epipolar line)较小时, 匹配精度较差, 且对于均匀分布、模式重复的景物难以匹配.

多目立体视觉的研究可以追溯到 Morenci 1979 年的工作^[1]. 但真正突破双目立体视觉局限的早期工作可见于文献[2]. 与双目系统不同, 三目系统寻求的是三元匹配, 而且这种匹配建立在透视投影的可靠基础之上. Yachida 指出, 采用边界点作为匹配特征, 在实际图像中错误的三元匹配仅在某些巧合下才出现^[2]. Ayache 采用多边形近似后的边界线段的中点作为特征点, 取得了较好的效果^[3]. Shen 建立了立体视觉伪目标问题的定量描述, 并提出了三维空间匹配特征基元的三目立体视觉方法, 在没有任何先验约束及后处理的情况下, 错误匹配率可以方便地降至 1% 以下^[4]. 本文所述的三目基元立体视觉方法包括: (1) 边界基元的提取和金字塔图的建立(2)

金字塔图三元匹配

2 边界基元的提取和金字塔图的建立^[5](**Extraction of edge primitives and construction of pyramidal graph**)

2.1 边界基元的提取

边界基元的提取应按如下规则进行:

a) 每个边界片断不应有分枝, 即若一个边界点是两边界公共点, 则将此边界点看作那两边界片断的一个端点.

b) 边界的起始点和终结点作为边界片断的端点.

c) 若一个边界起始和终止在两个端点, 且不通过其它端点, 则此边即为一个边界片断.

d) 若在边界片断的某一点的曲率有很明显变化, 就将此边界片断分成两个(多边形近似).

2.2 基元间连接关系的描述

经过以上分割基元, 得到了二值图像的所有基元, 将每个基元表达为金字塔图中的一个节点, 而它们之间的相互关系表达为图中的指针. 这里主要考虑两类关系, 即直接关系和间接关系. 直接关系包括上邻接和下邻接的关系. 若基元 B 是基元 A 的连通近邻并且其在图像中的位置高于 A , 则称 A 为 B 的一个下近邻, 而 B 则称为 A 的一个上近邻, 显然, 一个基元可以有不止一个上(下)近邻. 至于间接关系, 这里考虑到四种不同的关系, 即上左邻, 上右邻, 下左邻和下右邻. 它们以如下方式定义: 一个基元 A 的上近邻的所有下近邻被认为是与基元 A 间接地上邻接的, 因为这些间接上邻与基元 A 的连接是通过 A 的一个上近邻而实现的, 它们中那些位于 A 左面的基元被称为基元 A 间接地上左连接, 而其中离 A 最近者被称为基元 A 的上左邻. 类似的, 可以定义上右邻、下左邻、下右邻等间接连接关系. 从以上定义可以看出, 一个基元最多只有一个上左邻, 对于上右邻、下左邻、下右邻也是如此.

每个基元特征为它的长度、宽度、位置(用中线的中点表示)和中线斜率, 以基元作图的节点, 而以基元间直接和间接关系作为节点间的指针, 就形成了描述二值图像的一个图表示, 它表达了每个基元的几何特征以及基元间的结构关系.

2.3 金字塔图的建立

以二值边界图像作为金字塔图的 0 级图像, 用的 2×2 窗口进行逐级压缩, 依次得到 0 级、1 级、2 级、...、 n 级的图像, 每幅图像尺寸前级的四分

之一, 在 i 级图像中每个不相叠的 2×2 窗口对应于第 $i+1$ 级金字塔图的一个像素. 如果在 2×2 窗口中, 至少存在一个属于某基元的像素, 那么认为 $i+1$ 级对应于此窗口的像素是一个基元的像素. 这些图像构成了图像金字塔, 它不仅反映了基元在原始图像分辨率下的关系, 并且反映了它们之间新发现的关系, 即越来越全局性的关系.

3 用金字塔图进行三元匹配(**Triplet matching by pyramidal graphs**)

对左、右、下三幅边界图像分别提取边界基元, 左、右、下图的基元集合分别为: $\{L_1, \dots, L_x\}$ 、 $\{R_1, \dots, R_y\}$ 、 $\{D_1, \dots, D_z\}$, X 、 Y 、 Z 分别为其包含的基元数. 匹配过程如下所述:

(1) 选择匹配起始点并搜索其可能的匹配基元

首先, 在立体视觉图像组的三个金字塔图中的一个(设为左图)内选择一个节点, 其对应于一个特征基元. 为避免噪声影响, 基元长度和宽度应大于某一预定门限. 选好起始节点后, 在另二图中的一个(基极级方向与其相差较大)内寻找其可能的匹配节点, 因为基元方向与基极级方向相差越大, 利用基极线几何原理进行计算的精度越高. 不妨设 L_i 首先在右图上寻找匹配, 右图上的基元集合中符合以下约束的被认为是基元 L_i 的可能匹配.

a) 基极线约束: R_j 至少应该有一部分与 L_i 的一部分在相同的基极线区域中.

b) 像差约束: L_i 和 R_j 间的像差应该在一个根据景物深度的先验知识决定的某个范围中, 如果先验知识不能得到, 那就接受所有像差值.

c) 结构相似性: 从相互对应节点发出指针数应相似.

设 L_i 在右图上的可能匹配集为 $\{R_{i1} \dots R_{is}\}$, 可设法在此集合中找到基元 L_i 的一个匹配. 取出 R_{iu} ($u=1 \sim s$) 逐点计算 L_i 和 R_{iu} 之间的对应性, 对于 L_i 上的某点, 用基极线几何原理决定在 R_{iu} 上的对应点, 如果找不到对应点, 就检查 L_i 的下一点, 否则 L_i 上的这点与其对应点在三维空间就决定一点 P , 再将 P 投影到另一幅图像上(这儿即指下图像平面), 如果 P 的投影点不是下图像的边界点, 则检查 L_i 的下一点; 如果是边界点, 且属于基元 D_q , 则将 $\{L_i, R_{iu}, D_q\}$ 这一组可能的匹配因子 $N[L_i, R_{iu}, D_q]$ 加 1, ($1 \leq q \leq Q_i$, Q_i 为可能的匹配组数), 找出具有最大匹配因子的那组数, 设为 $[L_i, R_{iu}, D_j]$ ($1 \leq j \leq Q_i$). 分别求得这组基元的各基元被匹配像素占的百分比.

$$P[L_i] = \frac{N[L_i, R_{iu}, D_j]}{M[L_i]}$$

$$P[R_{iu}] = \frac{N[L_i, R_{iu}, D_j]}{M[R_{iu}]}$$

$$P[D_j] = \frac{N[L_i, R_{iu}, D_j]}{M[D_j]}$$

$M[L_i], M[R_{iu}], M[D_j]$ 分别表示 L_i, R_{iu}, D_j 所包含的像素个数, 只要这三个百分比中至少有一个大于门限 T , 这组匹配被作为候选匹配组, 否则认为 L_i 与 R_{iu} 不匹配. 类似的, 再将 L_i 与 $\{R_{i1} \dots R_{is}\}$ 中其它基元依次匹配, 共找到具有最大匹配因子的 s 个值, 再根据这 s 个值找出全部候选匹配组, 并从中找出最大值, 其表示那组基元被认为是包含 L_i 的最可能匹配, 设为 $[L_i, R_{iv}, D_k] (1 \leq v \leq s, 1 \leq k \leq Q_i)$.

(2) 利用金字塔图传播匹配, 并进行全局验证.

取一个候选匹配组(首先为 $[L_i, R_{iv}, D_k]$) 作为三图的基节点, 传播从基节点开始, 通过从它们发出的一系列指针得以实现. 以候选匹配组 $[L_i, R_{iv}, D_k]$ 的上近邻指针为例, 设 U_L 为 L_i 的一个上近邻, $U_r(1), \dots, U_r(m)$ 为 R_{iv} 的全部上近邻, $U_d(1), \dots, U_d(n)$ 为 D_k 的全部上近邻, 当且仅当 $[U_L, U_r(\alpha), U_d(\beta)] (1 \leq \alpha \leq m, 1 \leq \beta \leq n)$ 满足如下条件时, 匹配组才被接受:

- (a) $U_L, U_r(\alpha), U_d(\beta)$ 之间满足基极线约束.
- (b) $U_L, U_r(\alpha), U_d(\beta)$ 之间满足像差连续性约束 (即从 $[U_L, U_r(\alpha), U_d(\beta)]$ 之间计算得到的像差值与从 $[L_i, R_{iv}, D_k]$ 之间计算得到的像差值足够接近).
- (c) 若 $U_L, U_r(\alpha), U_d(\beta)$ 中原已有匹配, 则必须在 $U_L, U_r(\alpha), U_d(\beta)$ 间的几何相似性大于原匹配间几何相似性时, 才被接受为一个新匹配, 并拒绝原结果.
- (d) $U_L, U_r(\alpha), U_d(\beta)$ 间的几何相似性应在 $(1 \leq \alpha \leq m, 1 \leq \beta \leq n)$ 范围内最优.

一旦发现新的匹配, 即将传播过程的成功数加 1. 类似地, 还可以匹配此基节点的其它邻节点. 每找到一组新匹配节点又可将其作为一组新的基节点继续传播匹配下去, 直到无新匹配产生.

定义:

$$\text{传播率} = \frac{\text{传播匹配成功数}}{\text{参加传播过程的节点总数}}$$

对所有候选节点重复上述过程, 取得最大匹配传播成功率且该成功率大于某一预置门限的匹配被认为最终接受的匹配结果. 并且, 所有在这一匹配过程中, 传播成功的匹配节点组均被接受.

(3) 金字塔图中其它节点的匹配

对于尚未匹配的节点, 重复(1)、(2)过程, 以求

新的匹配, 直到不能发现新的匹配结果为止.

4 由匹配点求对应空间点坐标(Calculating the coordinates in 3-D space by matched pixels)

即已知两匹配点为 $I_1(i_1, j_1), I_2(i_2, j_2)$, 求对应空间点 $P(P_x, P_y, P_z)$ 设两摄像机参数为:

1 号摄像机: $\overline{a_{01}}, \overline{h_{01}}, \overline{V_{01}}, i_{01}, j_{01}, \overline{C_1}$

2 号摄像机: $\overline{a_{02}}, \overline{h_{02}}, \overline{V_{02}}, i_{02}, j_{02}, \overline{C_2}$

$\overline{C_k}, \overline{a_{0k}}, \overline{h_{0k}}, \overline{V_{0k}}, i_{0k}, j_{0k} (k=1 \sim 2)$ 分别为摄像机光学中心、校准后的光轴、相邻像素间行与列方向矢量以及像平面中心在行、列方向的偏移量, $\|a_{0k}\| = 1^{[4]}$.

显然

$$\overline{C_1 I_1} \times \overline{C_1 P} = \overline{0} \tag{1}$$

$$\overline{C_2 I_2} \times \overline{C_2 P} = \overline{0} \tag{2}$$

而

$$\overline{C_1 I_1} = u_1 [\overline{a_{01}} + (j_1 - j_{01}) \overline{h_{e1}} + (i_1 - i_{01}) \overline{V_{e1}}]$$

$$\overline{C_2 I_2} = u_2 [\overline{a_{02}} + (j_2 - j_{02}) \overline{h_{e2}} + (i_2 - i_{02}) \overline{V_{e2}}]$$

其中 $\overline{h_{ek}} = \overline{h_{0k}} / (\overline{h_{0k} h_{0k}})$
 $\overline{V_{ek}} = \overline{V_{0k}} / (\overline{V_{0k} V_{0k}})$ u_k 为系数, $k=1, 2$ 由此可得 $\overline{C_1 I_1}, \overline{C_2 I_2}$ 单位向量, 令其分别为

$$\overline{Y_1}: [\mathcal{Y}(1, 1), \mathcal{Y}(1, 2), \mathcal{Y}(1, 3)]^T$$

$$\overline{Y_2}: [\mathcal{Y}(2, 1), \mathcal{Y}(2, 2), \mathcal{Y}(2, 3)]^T$$

故(1), (2) 可变为

$$\overline{Y_1} \times (\overline{P} - \overline{C_1}) = \overline{0} \tag{3}$$

$$\overline{Y_2} \times (\overline{P} - \overline{C_2}) = \overline{0} \tag{4}$$

将(3), (4) 式展开得

$$\mathcal{Y}(1, 2)P_x - \mathcal{Y}(1, 1)P_y = \mathcal{Y}(1, 2)C_{1x} - \mathcal{Y}(1, 1)C_{1y}$$

$$\mathcal{Y}(1, 3)P_x - \mathcal{Y}(1, 1)P_z = \mathcal{Y}(1, 3)C_{1x} - \mathcal{Y}(1, 1)C_{1z}$$

$$\mathcal{Y}(1, 3)P_y - \mathcal{Y}(1, 2)P_z = \mathcal{Y}(1, 3)C_{1y} - \mathcal{Y}(1, 2)C_{1z}$$

$$\mathcal{Y}(2, 2)P_x - \mathcal{Y}(2, 1)P_y = \mathcal{Y}(2, 2)C_{2x} - \mathcal{Y}(2, 1)C_{2y}$$

$$\mathcal{Y}(2, 3)P_x - \mathcal{Y}(2, 1)P_z = \mathcal{Y}(2, 3)C_{2x} - \mathcal{Y}(2, 1)C_{2z}$$

$$\mathcal{Y}(2, 3)P_y - \mathcal{Y}(2, 2)P_z = \mathcal{Y}(2, 3)C_{2y} - \mathcal{Y}(2, 2)C_{2z}$$

以上方程组含有三个未知数, 经整理可得如下形式

$$AP = B \tag{5}$$

其中 A 为 3×3 对称阵, P 为 $(P_x, P_y, P_z)^T$ B 为 3×1 常数矩阵

$$a_{11} = \mathcal{Y}^2(1, 1) + \mathcal{Y}^2(2, 1) - 2$$

$$a_{12} = a_{21} = \mathcal{Y}(1, 1)\mathcal{Y}(1, 2) + \mathcal{Y}(2, 1)\mathcal{Y}(2, 2)$$

$$a_{13} = a_{31} = \mathcal{Y}(1, 1)\mathcal{Y}(1, 3) + \mathcal{Y}(2, 1)\mathcal{Y}(2, 3)$$

$$a_{22} = \mathcal{Y}^2(1, 2) + \mathcal{Y}^2(2, 2) - 2$$

$$a_{23} = a_{32} = \mathcal{Y}(1, 2)\mathcal{Y}(1, 3) + \mathcal{Y}(2, 2)\mathcal{Y}(2, 3)$$

$$a_{33} = \mathcal{Y}^2(1, 3) + \mathcal{Y}^2(2, 3) - 2$$

$$b_1 = \mathcal{Y}(1, 1)\overline{C_1Y_1} - C_{1x} + \mathcal{Y}(2, 1)\overline{C_2Y_2} - C_{2x}$$

$$b_2 = \mathcal{Y}(1, 2)\overline{C_1Y_1} - C_{1y} + \mathcal{Y}(2, 2)\overline{C_2Y_2} - C_{2y}$$

$$b_3 = \mathcal{Y}(1, 3)\overline{C_1Y_1} - C_{1z} + \mathcal{Y}(2, 3)\overline{C_2Y_2} - C_{2z}$$

只要 A 非奇异, 由方程(5)可唯一地确定 $I_1(i_1, j_1)$ 和 $I_2(i_2, j_2)$ 在三维空间对应点 P 的坐标.

5 由空间点求在第三个像平面上的投影点 (Projection of the point in 3-D space unto the third image)

即已知 $P(P_x, P_y, P_z)$ 和摄像机参数, 求 $I_3(i_3, j_3)$.

设

$$\overline{C_3P} = \lambda \overline{C_3I_3}$$

$$= \lambda'[\overline{a_{03}} + (j_3 - j_{03})\overline{h_{e3}} + (i_3 - i_{03})\overline{v_{e3}}]$$

式中 $\lambda' = \lambda u_3$, u_3 为系数, 两边点乘 $\overline{a_{03}}$, 得

$$\lambda' = \overline{a_{03}} \overline{C_3P}$$

两边点乘 $\overline{h_{e3}}$ 得

$$\overline{c_3p} \overline{h_{e3}} = \lambda' (j_3 - j_{03}) \overline{h_{e3}} \overline{h_{e3}}$$

$$j_3 - j_{03} = \overline{c_3p} \overline{h_{e3}} / (\|\overline{h_{e3}}\|^2 \lambda')$$

$$j_3 = \overline{c_3p} \overline{h_{03}} / (\overline{a_{03}} \overline{c_3p}) + j_{03}$$

同理, 两边点乘 $\overline{v_{e3}}$, 得

$$i_3 = \overline{c_3p} \overline{v_{03}} / (\overline{a_{03}} \overline{c_3p}) = i_{03}$$

这样, I_3 点的坐标 (i_3, j_3) 已经求出, 即

$$\begin{cases} i_3 = \overline{c_3p} \overline{v_{03}} / (\overline{a_{03}} \overline{c_3p}) + i_{03} \\ j_3 = \overline{c_3p} \overline{h_{03}} / (\overline{a_{03}} \overline{c_3p}) + j_{03} \end{cases}$$

(上接第 399 页)

5 结论(Conclusion)

本文提出的机器人路径搜索方法是一种动态的局部搜索方法: 引入模糊概念对环境空间进行模糊建模, 参照物体的位置和运动信息构造二维隶属函数; 然后通过模糊综合评价, 对各个方向进行综合考察, 得到搜索结果. 通过实验表明该方法有较好的实时性能, 能够有效的解决动态环境中机器人避碰和路径搜索问题.

6 结论(Conclusion)

本文提出了利用金字塔图传播匹配的三目基元立体视觉方法. 该方法中匹配起始三元组的选择建立在透视投影的可靠基础之上, 而进行传播匹配大大限制了搜索范围, 因而可以加速计算, 并避免了可能的误匹配. 传播匹配利用立体图像组在二值边界图像金字塔图中的结构相似性, 而整个匹配结果的考察与认定是通过传播率的比较实现的. 该方法能够有效地克服双目立体视觉方法的某些本质局限性, 确保匹配结果精确可靠.

参考文献 (References)

- 1 Moravec H P. Visual mapping by a robot ROVER, Proc. 6th IJ-CAI, 1979, 598- 620
- 2 Yachida M, Kitamura Y, Kimachi M. Trinocular vision: new approach for correspondence problem, Proc. 8th ICPR, Paris, 1986
- 3 Ayache N, Lustman F. Fast and reliable passive stereo vision using three cameras, Int. Workshop on Ind. Appli and Mach. Int., Tokyo, 1987
- 4 Shen J, Castan S, Zhao J. A new passive measurement method by trinocular stereo vision, Industrial Metrology, 1990
- 5 Shen J, Castan S, Zhao J. Stereo vision by pyramidal BLI graph matching, Proc. SPIE'87, Cannes, 1987

作者简介:

任重 (1963-), 男, 博士生. 研究领域: 人工智能, 计算机视觉, 网络管理等.

邵军力 (1938-), 男, 教授, 博士生导师. 研究领域: 人工智能, 专家系统, 指挥自动化理论与技术等.

参考文献 (References)

- 1 Philip Kilby, Patrick Prosser, Paul Shaw. Guided Local Search for the Vehicle Routing Problem. 2nd International Conference on Metaheuristics-MIC97
- 2 Borenstein J, Koren Y. The Vector Field Histogram-Fast Obstacle Avoidance for Mobile Robots. IEEE Journal of Robotics and Automation, 1991, 7(3): 278- 288
- 3 袁曾任, 高明. 在动态环境中移动机器人导航和避碰的一种新方法. 机器人, 2000, 22(2)

作者简介:

庄晓东 (1977-), 男. 研究领域: 人工智能及其应用.

孟庆春 (1958-), 男. 研究领域: 智能控制及其应用.