

# 基于灰度段的立体匹配算法\*

周东翔, 蔡宣平, 孙茂印

(国防科学技术大学 电子科学与工程学院, 湖南 长沙 410073)

E-mail: dxzhou@nudt.edu.cn

http://www.nudt.edu.cn

**摘要:** 立体匹配一直是计算机视觉领域的一个中心研究问题. 为了得到适用于 IBR (image-based rendering) 技术中视图合成的比较精确的高密度视差图, 提出了一种基于灰度段的立体匹配算法. 该算法以灰度段作为匹配基元, 并将应用于语音识别的 DTW (dynamic time warping) 方法引入灰度段的匹配算法中. 相对于点基元, 灰度段基元覆盖的图像空间要大得多, 且不易受噪声、光度变化等因素的影响, 因此可以减少误匹配发生的几率, 更容易进行匹配, 比特征线段、二次曲线等匹配基元计算要简便得多. 实验结果表明, 该算法效果良好, 具有实用价值.

**关键词:** 立体匹配; 灰度段; 视差图; 动态时间规整

**中图法分类号:** TP391 **文献标识码:** A

基于图像的绘制技术 (image-based rendering, 简称 IBR) 已成为计算机图形学和虚拟现实领域的一个重要研究方向. 而基于立体视觉恢复景物的深度信息又为 IBR 提供了更方便、更优越的途径, 新视点的图像可以很容易地通过已知相邻视图合成产生. 立体视觉技术的关键是找出每对图像间的对应关系, 即解决立体匹配问题, 得到视差图 (disparity map), 这也是最困难的问题. 视差  $d$  与景物的深度  $z$  之间的关系为

$$d = BF \frac{1}{z}, \quad (1)$$

其中  $B$  为图像对的基线距离,  $F$  为相机的焦距. 从式 (1) 中可以看出, 如果两个相机的相对位置及焦距已知, 由视差图就能很容易地估计出场景中可见点的深度信息.

由于噪声、遮挡及透视失真等因素的影响, 视差估计是一个很困难的问题. 目前, 这个领域的工作可分为 3 类: 基于特征、基于区域和基于相位. 基于特征的方法主要是提取特征点、线或边缘等特征, 然后进行匹配. 因为这些特征对噪声不太敏感, 因此可以得到比较精确的匹配, 但特征匹配方法只能得到稀疏的视差图<sup>[1~5]</sup>. 很多研究人员通过插值或表面重建等方法来得到高密度的深度图, 这在大多数情况下都得不到很好的效果. 区域匹配方法可以得到高密度深度图<sup>[6~10]</sup>, 但可靠性比较差, 而且在遮挡区域、无纹理区域以及深度不连续处得不到正确的结果. 相位匹配方法一般只能得到景物的粗糙结构<sup>[11,12]</sup>, 而且在一些不定的傅立叶相位区域需要进行特殊处理.

在 IBR 的视图合成应用中, 我们希望得到既精确密度又高的深度信息. 为了满足这种需求, 我们在参考文献 [13] 中提出了一种基于特征约束的立体匹配算法, 获得了较好的效果. 在这种算法中, 边缘特征点的匹配比较关键, 对于边缘特征不显著或过于稀疏的图像对, 该算法有其适用的局

\* 收稿日期: 1999-11-30; 修改日期: 2000-12-05

基金项目: 高等学校骨干教师资助计划资助项目

作者简介: 周东翔 (1968—), 男, 湖南隆回人, 博士, 讲师, 主要研究领域为计算机视觉, 虚拟现实, 图形图像处理, 信息可视化技术; 蔡宣平 (1962—), 男, 福建龙溪人, 教授, 主要研究领域为计算机图形学, 计算机应用; 孙茂印 (1941—), 男, 河南巩县人, 教授, 博士生导师, 主要研究领域为图形图像处理, 信息可视化技术.

限性. 针对这种情形, 我们提出一种基于灰度段的立体匹配算法, 将图像中对对应极线上的图像信号按灰度信息分段, 以灰度段作为匹配基元, 利用 DTW(dynamic time warping)方法进行匹配. 以灰度段作为匹配基元, 比点基元覆盖的图像空间要大, 可以减少误匹配发生的几率, 更容易进行匹配, 而且这种方法比基于特征线段、二次曲线或任意平面代数曲线<sup>[14]</sup>的方法计算起来要简便得多. DTW 是应用于语音识别的一种有效方法<sup>[15]</sup>, 由于灰度段与语音段有惊人的相似性, 我们将 DTW 方法引入到灰度段的匹配算法中. 与其他立体匹配算法一样, 基于灰度段的立体匹配算法需要利用一些基本约束: 标准外极线约束、唯一性约束和匹配保序性约束.

本文首先叙述利用 DTW 方法进行立体匹配的原理, 然后描述本文提出的灰度段立体匹配算法过程及其具体实现, 最后分析实验结果, 得出结论.

## 1 基于 DTW 的灰度段匹配原理

设  $A, B$  分别为左、右图像中对应像素行上的灰度段:

$$\begin{cases} A = a_1, a_2, \dots, a_M \\ B = b_1, b_2, \dots, b_N \end{cases} \quad (2)$$

如图 1 所示,  $A, B$  构成平面. 平面上的点阵  $c_k = (i, j)$  的序列表示  $A, B$  两灰度段中对应像素点的匹配路径.

$$F = c_1, c_2, \dots, c_k, \dots, c_L, \quad c_k = (i_k, j_k) \quad (3)$$

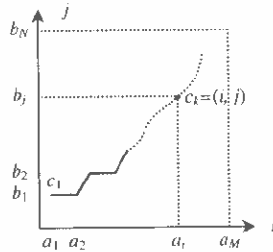


Fig. 1 Illustration of the matching path of pixels between gray-level segments  $A$  and  $B$

图 1 灰度段  $A$  和  $B$  的像素匹配路径示意图

若  $d(c) = d(i, j) = \|I(a_i) - I(b_j)\|$  ( $I(a_i), I(b_j)$  分别表示像素点  $a_i, b_j$  的灰度值) 表示像素点  $a_i$  和  $b_j$  的距离, 则可以用下式表示沿  $F$  的距离的总和:

$$D(F) = \frac{\sum_{k=1}^L d(c_k) w_k}{\sum_{k=1}^L w_k} \quad (4)$$

其中  $w_k$  表示与  $F$  相关的正加权系数. 该值越小, 则  $A, B$  的对应越好.

如果决定了  $w_k$ , 分母就变成了与  $F$  无关的常数, 式(4)就可以简化. 设  $W_k = (i_k - i_{k-1}) + (j_k - j_{k-1}), i_0 = j_0 = 0, w_k$  就是沿纵、横线轴的基本轨迹线, 称为市区化距离(city block distance)<sup>[15]</sup>, 则

$$\sum_{k=1}^L w_k = M + N. \quad (5)$$

这时, 式(4)就可以改写成下式:

$$D(F) = \frac{1}{M+N} \sum_{k=1}^L d(c_k) w_k. \quad (6)$$

设部分点阵  $c_1, c_2, \dots, c_k (c_k = (i, j))$  的和为  $g(c_k) = g(i, j)$ , 则动态规划的定型公式为

$$g(i, j) = \min \begin{bmatrix} g(i, j-1) + d(i, j) \\ g(i-1, j-1) + 2d(i, j) \\ g(i-1, j) + d(i, j) \end{bmatrix}. \quad (7)$$

因此设  $g(1, 1) = 2d(1, 1)$ ,  $j = 1$ , 一边变化  $i$ , 一边计算式(7), 然后增加  $j$ , 反复进行相同的计算, 使  $j = N$ . 由于存在遮挡的情况, 图像对中相匹配的灰度段的端点与端点不一定完全对应, 因此, 我们在应用 DTW 方法时, 必须将灰度段的始点与终点都放松. 在最后求取距离总和时, 我们采用如图 2 所示的策略, 从  $g(M, N)$  出发, 寻找  $g(m_k, N)$  (其中  $m_k = M, M-1, \dots, 1$ ) 中的第 1 个极小值  $g(m, N)$  以及  $g(M, n_k)$  (其中  $n_k = N, N-1, \dots, 1$ ) 中的第 1 个极小值点  $g(M, n)$ , 归一化后的距离为

$$D(F) = \min(g(m, N), g(M, n)) / (M + N). \quad (8)$$

这样, 选取左图像中某一像素行上的灰度段, 在按照上述方法求得的右图像中所对应的像素行上归一化距离最小的灰度段即为匹配的灰度段.

## 2 基于灰度段的匹配算法过程及其具体实现

前面我们叙述了基于 DTW 方法进行灰度段匹配的基本原理. 这一节我们主要描述如何将图像的每一行信息进行分段, 然后利用上述的 DTW 方法进行匹配的整个算法过程及其具体实现. 为了便于处理, 一般都需要先将图像对进行预处理, 也即对图像进行校正, 使得两幅图像的匹配点位于同一条扫描线上, 这也就是绝大多数立体匹配算法都要利用的标准外极线约束.

### 2.1 分段

我们利用像素点的灰度信息对图像的像素行进行分段, 设  $f(x, y)$  为  $y$  像素行上第  $x$  个像素点的灰度值, 则在每一段内需满足下述条件:

$$\left. \begin{aligned} |f(x_i, y) - M| < T, \quad i = 1, \dots, w, \\ M = \frac{1}{w} \sum_{i=1}^w f(x_i, y). \end{aligned} \right\} \quad (9)$$

其中  $T$  为某一阈值,  $w$  为灰度段的长度.

### 2.2 段的匹配及像素点的视差计算

#### (1) 段的匹配过程

按照上述方法, 我们对左、右图像中相应的像素行进行分段, 下面的关键问题是如何进行段的匹配以及进行像素点的视差计算.

首先, 我们来介绍段的匹配过程. 我们选取长度最长的段先进行匹配, 如图 3 所示. 这样做不仅可以尽可能地避免匹配过程中的遮挡情况, 而且可以缩小剩余段的匹配搜索范围. 我们在前面已对匹配的原理进行了详细的论述. 剩余段的匹配在该段的左、右部分分别进行, 也同样先选长度最长的段进行匹配. 像这样一直循环下去, 直到所有灰度段都完成匹配过程为止.

#### (2) 像素点的视差计算

设左图像中某像素行的  $A$  段与右图像中对应像素行上的  $B$  段匹配, 两段的长度不一定相等, 也就是说两段的像素点不一定一一对应, 那么如何找到最佳的匹配路径来完成像素点的视差计算

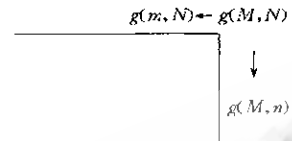


Fig. 2 Illustration of searching the minimum accumulated distance  
图2 搜索最小距离之和示意图

呢? 本文的算法采取如下的策略.

参照第 1 节对匹配原理的描述, 如果我们已经分别找到了第 1 个极小值  $g(m, N)$  及  $g(M, n)$ , 不失一般性, 设  $g(m, N) < g(M, n)$ , 则  $(m, N)$  为末端匹配点对, 也就是说, 左图像 A 段的第  $m$  个像素点与右图像 B 段的第  $N$  个像素点(端点)匹配. 从末端匹配点对出发, 逐步向前搜索其余的匹配点对, 直到始端点为止. 如图 4 所示, 遵循以下处理原则:

$$\left. \begin{array}{l} \text{if } \left\{ \begin{array}{l} g(i-1, j-1) < g(i-1, j) \text{ and} \\ g(i-1, j-1) < g(i, j-1) \end{array} \right\} \text{ or } \\ \left\{ \begin{array}{l} g(i-1, j-1) = g(i-1, j) \text{ and} \\ g(i-1, j-1) = g(i, j-1) \end{array} \right\} \end{array} \right\} \text{ 选取 } (i-1, j-1) \\ \text{if } \left\{ \begin{array}{l} g(i-1, j) \leq g(i-1, j-1) \text{ and} \\ g(i-1, j) < g(i, j-1) \end{array} \right\} \text{ 选取 } (i-1, j) \\ \text{else 选取 } (i, j-1) \end{array} \right\} \quad (10)$$

Fig. 3 Illustration of the matching processing between gray-level segments  
图3 灰度段的匹配过程示意图

Fig. 4 Illustration of finding the corresponding points  
图4 匹配点对的选取示意图

按照上述处理原则, 我们可以得到最佳匹配路径(如图 5 所示), 像素点的视差也就可以计算出来, 即为匹配点对的  $x$  坐标差值.

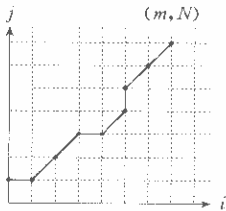


Fig. 5 Illustration of the best-matching path between the corresponding gray-level segments  
图5 对应灰度段的最佳匹配路径示意图

### (3) 纵向数据拟合校正及插值

按照上述方法已经可以得到左图像或右图像的高密度视差图, 但有可能存在少数的误匹配及没有匹配的像素点, 这时我们采取数据拟合方法进行校正和插值.

按照式(9)的原理, 选取某一阈值, 在纵向对图像的每一列进行分段, 然后对每一段的视差值进行数据拟合. 对简单的图像可以采用直线拟合.

如果不考虑处理速度, 可以采用高阶的多项式拟合. 这样不但可以校正误匹配像素点, 而且对没有获得匹配的像素点也进行了插值处理.

## 3 实验结果

我们对本文提出的算法在普通 PC 机上进行了实验验证. 实验原图像取自德国波恩大学计算机视觉研究组. 图 6(a)、(b)分别为左、右原图. 图 6(c)为最后的视差图, 图中灰度级低(亮度低)的地方表示视差小, 按照式(1), 也即距视点较远, 深度值大; 反之, 灰度级大(亮度高)的像素点表示视差大, 也即距视点较近, 深度值小. 图 6(d)是德国波恩大学计算机视觉研究组提供的真实视差图. 它是以分辨率  $2560 \times 2560$  制作, 然后压缩成分辨率  $256 \times 256$  在网上提供, 因此层次非常分明. 我们从分辨率  $256 \times 256$  的原图像得到的视差图自然不可能如此精细. 图 6(e)为该视觉研究组的 Fröhlinghaus 等人<sup>[11]</sup>利用相位匹配方法得到的视差图, 正如我们在前面分析的那样, 只能得到景物的粗糙结构. 图 6(f)是澳大利亚 Macquarie 大学 CSIRO 研究中心的 Sun Changming<sup>[3,9]</sup>利用区

域相关及动态规划方法得到的视差图. 通过比较可以明显看出, 我们提出的基于灰度段的立体匹配算法获得了较好的效果.

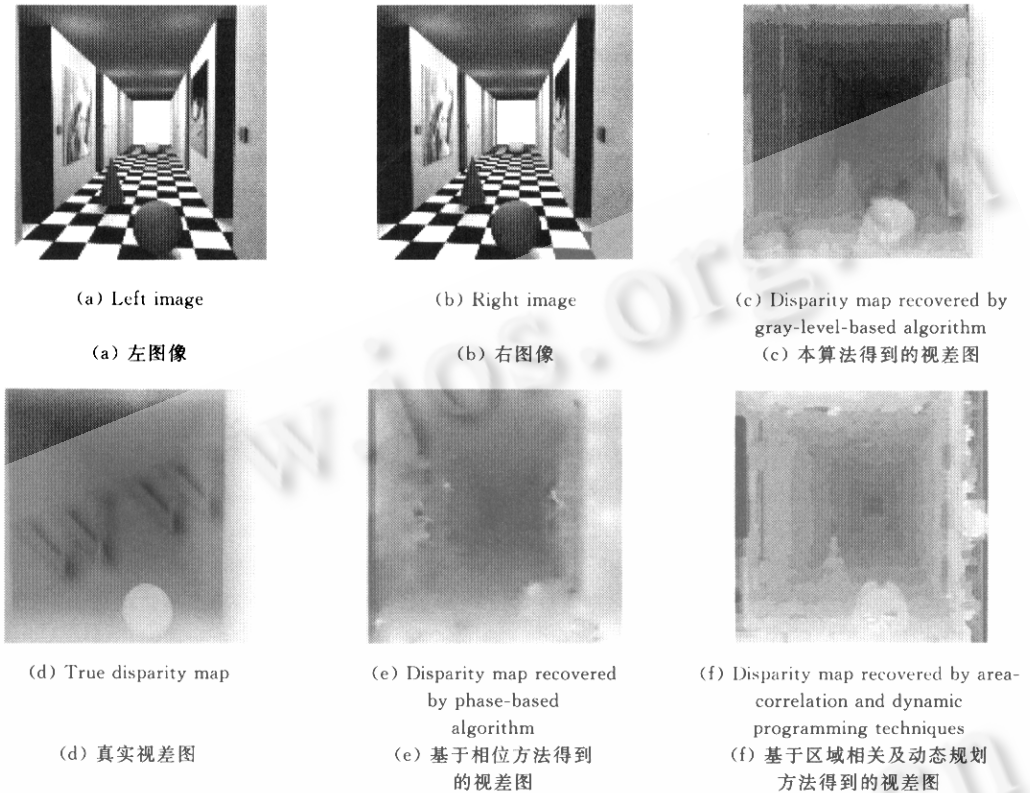


Fig. 6 The matching result for corridor images  
图 6 走廊图像的匹配结果

## 4 结论

本文在深入分析现有立体匹配算法的基础上, 提出了基于灰度段的立体匹配算法. 以灰度段作为匹配基元, 利用 DTW 方法进行匹配. DTW 是成功地应用于语音识别的一种有效方法, 我们将它引入到灰度段的匹配当中, 通过实验验证, 同样获得了良好的效果. 采用灰度段作为匹配基元的优势在于, 相对于点基元, 覆盖了更大的图像空间, 从而减少了误匹配发生的几率, 更容易进行匹配. 实验结果也表明, 该算法比采用点基元的区域相关匹配方法效果要好.

## References:

- [1] Grimson, W. E. L. Computational experiments with a feature based stereo algorithm. IEEE Transactions on PAMI, 1985, 7(1):17~34.
- [2] Terzopoulos, D. The computation of visible surface representation. IEEE Transactions on PAMI, 1988, 10(4):417~438.
- [3] Baker, H. H., Binford, T. O. Depth from edge and intensity based stereo. In: Proceedings of the 7th Joint Conference on Artificial Intelligence. 1981. 631~636.
- [4] Hoff, W., Ahuja, N. Surfaces from stereo; integrating feature matching, disparity estimation, and contour detection. IEEE Transactions on PAMI, 1989, 11(2):121~136.
- [5] Ohta, Y., Kanade, T. Stereo by intra- and inter-scanline search using dynamic programming. IEEE Transactions on

- PAMI, 1985,7(2):139~154.
- [6] Lan, Zhong-dan, Roger, M. Robust matching by partial correlation. Technical Report, 2643, INRIA, 1995.
- [7] Cochran, S., Medioni, G. 3-D surface description from binocular stereo. IEEE Transactions on PAMI, 1992,14(10):981~994.
- [8] Sun, Chang-ming. A fast stereo matching method. In: Proceedings of Digital Image and Vision Computing: Techniques and Applications. Auckland, New Zealand; Massey University, 1997. 95~100. <http://www.cmis.csiro.au/IAP/RecentProjects/StereoEE02.htm>.
- [9] Sun, Chang-ming. Multi-Resolution stereo matching using maximum-surface techniques. In: Proceedings of Digital Image and Vision Computing: Techniques and Applications. 1999. 195~200. <http://www.cmis.csiro.au/IAP/RecentProjects/StereoEEO2.htm>.
- [10] Faugeras, O., Hotz, B., Mathieu, H., *et al.* Real time correlation-based stereo; algorithm, implementation and applications. Technical Report, 2013, INRIA, 1993.
- [11] Fröhlinghaus, T., Buhmann, J.M. Regularizing phase-based stereo. In: Proceedings of the International Conference on Pattern Recognition (ICPR'96). 1996. 451~455. <http://www-dbv.informatik.uni-bonn.de/abstracts/fröhlinghaus.icpr96.html>.
- [12] Crespi, B., Alex, G., Cozzi, analog computation for phase-based disparity estimation; continuous and discrete models. Machine Vision and Applications, 1998,11(2):83~95.
- [13] Zhou, Dong-xiang, Cai, Xuan-ping, Sun, Mao-yin. A feature-constrained stereo matching algorithm. In: Proceedings of the 6th International Conference of CAD/CG'99. Shanghai; Wen Hui Publishers, 1999.
- [14] Li, L., Ma, S.D. 3D pose estimation from a N-degree planar curve in two perspective views. In: Proceedings of the 13th ICPR. 1996. <http://www.cs.cornell.edu/Lili/publications.html>.
- [15] Rabiner, L.R., Schafer, R.W. Digital Processing of Speech Signals. Englewood Cliffs, NJ, Prentice Hall, Inc., 1978.

## A Stereo Matching Algorithm Based on Gray-Level Segments\*

ZHOU Dong-xiang, CAI Xuan-ping, SUN Mao-yin

(School of Electronic Science and Engineering, National University of Defence Technology, Changsha 410073, China)

E-mail: dxzhou@nudt.edu.cn

<http://www.nudt.edu.cn>

**Abstract:** Stereo matching has long been one of the central research problems in computer vision. A stereo matching algorithm is proposed based on gray level segments in order to recover dense depth map for the application of view synthesis in IBR (image-based rendering). This algorithm uses gray-level segments as elements to be matched, and introduces the principle of DTW (dynamic time warping), which applied for speech recognition, into the matching processes on the segments. Covering more image space than point elements, gray-level segments are less sensitive to noise and photometric variations and easier to be matched, and can reduce the false matching. Moreover, gray-level segments require less computational resources than feature-based line segments and conics. Experimental results show that this algorithm is effective and of great value to use.

**Key words:** stereo matching; gray-level segment; disparity map; dynamic time warping

\* Received November 30, 1999; accepted December 5, 2000

Supported by the National Sustentation Foundation for College Skeleton Teachers