

立体匹配算法研究综述^①陈 华^② 王立军 刘 刚^③

(燕山大学河北省测试计量技术及仪器重点实验室 秦皇岛 066004)

摘 要 双目立体视觉的主要任务是通过 2 台位于不同位置的摄像机拍摄同一场景得到 2 幅视图,然后计算视差并获取 3 维数据信息。立体匹配是双目立体视觉中的关键环节,它利用 2 幅或者多幅匹配后的图像获取视差进而得到深度信息。本文通过分析国内外对双目立体匹配算法的研究,对立体匹配算法进行介绍,以全局匹配算法和局部匹配算法为对象作详细阐述,最后对各种算法的特性进行比较,并作出总结与展望。

关键词 双目立体视觉;立体匹配;全局匹配算法;局部匹配算法

0 引 言

双目立体视觉系统是计算机视觉系统中最接近人眼的立体视觉系统,通过模仿人类视觉原理,得到包括深度信息在内的多种 3 维数据信息。Marr 的视觉理论立足于计算机科学,较为系统地揭示了用 2 维图像恢复 3 维物体形态的可能性和基本方法,是计算机视觉领域的重要基础。根据 Marr 视觉理论,Barnard 等人^[1]提出了完整的计算机立体视觉系统,包括图像获取、摄像机标定、特征提取、立体匹配、深度确定、内插 6 个环节。其中,立体匹配是最核心、最困难的问题,也是立体视觉领域研究的热点。

立体匹配,指的是通过 2 台参数一致的摄像机拍摄同一场景获得左右 2 幅视图,然后采用数学运算模型计算出左右 2 幅视图中的对应点以及视差,最后求取深度信息的过程^[2,3]。立体匹配领域的主要困难体现在遮挡区域、低纹理或无纹理区域、深度不连续区域和噪声几个方面,近年来对立体匹配的研究主要是引入不同的算法来达到更好的匹配效果。

本文主要介绍了双目立体匹配的简化模型、立

体匹配的约束条件、算法步骤、评价标准以及立体匹配算法的分类,并分别对全局匹配算法、局部匹配算法、稠密匹配算法和稀疏匹配算法的特性进行了阐述。

1 双目立体匹配研究内容

1.1 双目立体匹配简化模型

双目立体视觉系统可以看成是简化的人类视觉系统,它利用 2 个光轴平行配置,参数一致并且在同一平面的摄像机拍摄相同的物体,从而得到深度信息。如图 1 所示,假设以下 2 台摄像机之间处于绝

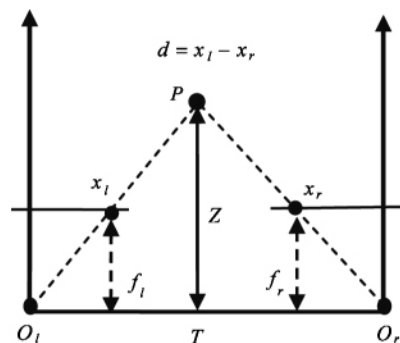


图 1 双目立体视觉简化模型

① 河北省自然科学基金(F2015203287)资助项目。

② 女,1972 年生,博士,副教授;研究方向:视觉测量,数字图像处理,模式识别等;E-mail: chenhua@ysu.edu.cn

③ 通信作者,E-mail: liugang@ysu.edu.cn

(收稿日期:2019-01-08)

对的平行和相对固定的位置,且点 P 在左右摄像机投影平面上处于同一极线。其中,用深度值 Z 来表示点 P 相对于摄像机的距离,点 P 为被摄物体表面上的一点。

图 1 中, T 表示左右摄像机光轴中心 O_l 和 O_r 之间距离, f_l 和 f_r 为左右摄像机的焦距, x_l 和 x_r 为点 P 在左右摄像机投影平面上的横坐标, Z 表示点 P 到相机光轴中心的距离。通过三角形相似性原理可以推导出深度值 Z , 公式如下:

$$\frac{T - (x_l - x_r)}{Z - f} = \frac{T}{Z} \Rightarrow \frac{fT}{x_l - x_r} = \frac{fT}{d} = Z \quad (1)$$

由式(1)可知,深度值 Z 和视差 d 之间为反比关系,所以通过匹配得出视差 d 后便可确定深度信息 Z 。

1.2 立体匹配的约束条件

立体匹配的约束条件又称为匹配准则,实际操作中,需要通过加入一些约束条件来提高算法精度,实现接近理想状态的双目立体视觉系统。

1.2.1 极线约束

考虑到系统中的 2 台摄像机的图像平面难以保持完全平行,如图 2 所示,极线约束就是通过校正使点 P 在左右摄像机平面的映射点 x_l 和 x_r 的纵坐标相同。极线校正可以降低匹配点的搜索难度,进而提高运算速度。

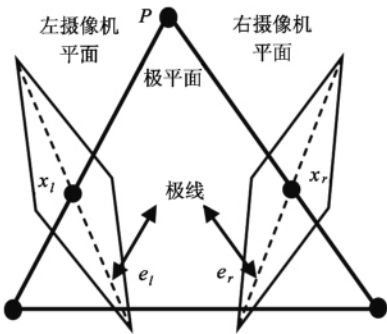


图 2 极线约束示意图

1.2.2 顺序一致性约束

考虑到像素点的方位问题,大多数情况下,认为位于相应极线上的所有左右匹配点以一致的顺序排列在极线上。

1.2.3 唯一性约束

考虑到匹配过程中可能出现图像中同时存在多

个像素相同的情况,设任一视点在投影图像上的像素点具有唯一性,且左右 2 幅投影图像上的像素点仅能一一对应。

1.2.4 连续性约束

考虑到物体表层的起伏状态和表层到摄像机的距离变动会影响到视差变化,假设被拍摄物体表面光滑,其视差变化连续,并将图像上相邻像素点也认为是连续的。在利用该约束条件的时候应该注意到不能超过图像的边界^[4]。

1.2.5 左右一致性约束

考虑到遮挡问题的影响,认为左右 2 幅图像中的像素点具有一一对应的匹配关系。

1.3 立体匹配的算法步骤

Scharstein 和 Szeliski^[5]对现有的立体匹配算法进行了概括比较,将立体匹配算法概括为 4 个步骤:匹配代价的计算、匹配代价的聚合、求取视差和视差优化、视差细化。

1.3.1 匹配代价的计算

匹配代价是匹配图像之间对应像素点的相似程度的表征。常用来计算匹配代价的方法有灰度差的平方、灰度差的绝对值和互相关函数^[6]。近年来,为了提高鲁棒性,常用 Census 变换和 Rank 变换计算匹配代价。

具体方法为:首先通过在水平极线上依次移动目标图像上的像素点,直到像素点数等于视差的最大搜索范围,运算后得到匹配图像之间的颜色代价差异,然后按同理得到匹配图像之间的梯度差异,最后赋予两者权重来进行计算。

1.3.2 匹配代价的聚合

局部匹配算法中通常会用到匹配代价聚合来提高算法的精度。该步骤的基本思想是在色彩一致性假设的前提下,结合初始匹配代价,利用匹配点周围信息重新计算。

具体方法为:首先以一个像素为中心,选择一个 $n \times n$ 的窗口,然后赋予窗口内的像素点不同的权值进行求和。

1.3.3 求取视差和视差优化

局部匹配算法中,视差的计算通常采用 WTA (winner-take-all) 准则,即选取最优匹配。

全局匹配算法则是构造能量函数,通过求取能量函数最小化来确定最优视差。能量函数一般表示为

$$E(d) = E_{\text{data}}(d) + E_{\text{smooth}}(d) \quad (2)$$

其中, $E_{\text{data}}(d)$ 为表示匹配像素的一致性程度的数据项, $E_{\text{smooth}}(d)$ 为用来约束匹配点与周围像素点具有一致视差的平滑项。

全局匹配算法相较于局部匹配算法来说,虽然精确度有所提高,但是两者都无法避免出现误匹配的情况,所以对于视差还要进一步进行优化。

常用的优化算法包括左右一致性检测法、RANSAC 数学模型优化和曲面拟合等。

1.3.4 视差细化

对于细节要求不高的区域,一般常用 WTA 准则或者能量函数最小化的方法。但对于一些要求视差值达到亚像素级(非整数视差)的场合,一般采用曲线拟合或者插值法来细化视差。

1.4 立体匹配的评价标准

通常情况下,会利用统一的测试图像来对比得出各算法的好坏。在立体匹配领域, Middlebury 图片库^[7]是研究人员使用率较高的测试图像。

1.4.1 误匹配率

误匹配率,指的是误匹配点占总像素点数的比率。

$$B = \frac{1}{M} \sum_{(x,y) \in A} [|d_c(x,y) - d_r(x,y)| > \delta] \quad (3)$$

其中, M 为整个图像的像素数, $d_c(x,y)$ 为计算出的视图差, $d_r(x,y)$ 为真实的视图差, δ 为误差容许值,一般情况下对于正整数范围内的视差标注取 1。

1.4.2 峰值信噪比

峰值信噪比(peak signal to noise ratio, PSNR), 主要衡量视差图像整体上和真实图像的不同。

$$PSNR = 10 \times \log_{10} \left(\frac{MAX_i^2}{MSE} \right) \quad (4)$$

$$MSE = \frac{1}{N} \sum_{(x,y)} \|R(x,y - O(x,y))\|^2 \quad (5)$$

其中, R 表示视差图像, O 表示真实图像, N 为像素点的总数。

2 立体匹配算法分类

立体匹配算法的种类有很多,其划分标准也有很多,目前主要可以概括为 2 大类:第 1 类是根据算法运行时的约束范围划分;第 2 类是根据生成的视差图划分。

根据第 1 类划分标准,立体匹配算法可以分为全局立体匹配算法和局部立体匹配算法 2 种。

全局立体匹配算法,基于图像本身和邻域内的像素点信息。其主要思想是将多个像素点视差值求解归一化到同一个能量框架下,即将立体匹配问题转化为求解能量函数最小化问题^[8]。该算法精度较高,但算法效率较低,难以在实际工程中应用。

局部立体匹配算法,基于图像的灰度信息。在匹配过程中,不可忽略的前提是待匹配中心点与邻域点的视差深度一致,且物体表面与摄像机平面平行。该算法运算虽较为简单,但在低纹理区域和不连续区域的匹配效果不好。

根据第 2 类划分标准,立体匹配算法可以分为稠密立体匹配算法和稀疏立体匹配算法 2 种。

稠密匹配算法,指的是能够生成包含所有像素点的视差值的稠密视差图的方法。传统稠密匹配算法为窗匹配法,利用局部信息,根据窗口内灰度的相似性来确定是否匹配。该方法操作简单,但窗口的大小过程比较复杂,结果不稳定。

稀疏匹配算法,指的是生成只包含被选择的像素点视差值的稀疏视差图的方法。通常先提取特征点,然后使用局部匹配算法加以匹配,得到稀疏匹配点。此法运算快,但容易造成视差值的缺失。

2.1 全局立体匹配算法

全局匹配算法的本质是构建能量函数并求解能量最小化,通常需要借助马尔科夫随机场来进行求解。传统的全局匹配算法有:模拟退火(simulated annealing)法^[9]、mean-field annealing 法^[10]等,后来又陆续提出了动态规划法、图割法(graph-cuts)^[11-12]、置信度传播法(belief-propagation)^[13]以及协同优化法(cooperative-optimization)^[14]。本文将着重对动态规划法、图割法和置信度传播法进行介绍。

2.1.1 基于动态规划的立体匹配算法

动态规划算法基于极线约束,通过依次寻找极线上匹配点的最小代价路径来求解最优值。具体流程可描述为以下 4 步:

(1) 以像素点的行方向为横坐标,视差值为纵坐标,依次将匹配过程分为 k 个阶段,每个 x 坐标点对应一个阶段。

(2) 将立体匹配中各个阶段用不同的状态表示。共 3 种状态:仅存在于左视图中且无匹配点、仅存在于右视图中且无匹配点、在左右视图中相互对应。

(3) 确定状态转移方程。

(4) 求最优解。首先按顺序对各阶段的 3 种状态的代价和依次进行计算,然后根据最小代价和确定最优路径。

Liu 等人^[15]提出了一种点对点的动态规划匹配方法,但这种算法在匹配阶段执行了大量运算,后来又在原有基础上改进了 Forward-working 算法,缩短了计算时间,但最终视差精度比较低。罗嗣卿等人^[16]在前者基础上改进算法,利用快速自适应权重累积策略累积匹配成本,通过 2 维有序表结构加快计算速度,采用方向滤波的处理方法减少条纹现象。Li 等人^[17]为了解决条纹现象,提出了先提取特征点进而用最近邻搜索算法进行特征点匹配的方法。Leung 等人^[18]为提高匹配速度,提出了一种快速匹配方案,即采用迭代动态规划的方法,对行和列分别进行迭代,从而比较快速地达到匹配效果。Hu 等人^[19]为了解决图像边缘的匹配问题,采用了单向四连接树的方法,并在此基础上融合了快速动态规划法,不但提高了匹配精度,同时也降低了运算复杂度。

动态规划算法的优点包括时间复杂度低,能够很好地解决边缘区域和弱纹理区域的匹配问题,但缺点是最后呈现的视差图上条纹现象十分明显。

2.1.2 基于图割的立体匹配算法

图割算法是对视差范围内图像像素的离散标号求最优解,即能量最小化问题的理想求解方法之一。其主要思想是将匹配问题转化为最大流/最小割来实现,通过选择合适的能量函数来构造网络图,进而

对网络图求最小割。

如图 3 所示, $G = (V, E)$ 表示一个双终端有向图, V 为点集, E 为边集, $s, t \in V$ 为源点和汇点。定义容量函数 $c: E \rightarrow R^+, R^+$ 为非负实数,则图 G 和函数 c 构成网络 $N = \{G, s, t, c\}$ 。而网络流则是定义在边集 E 上的一个非负函数,表示为 $flow = \{flow(v, w)\}$ 。所谓最大流问题即在求解网络 G 上使流量 f 达到最大的 $flow$ 。

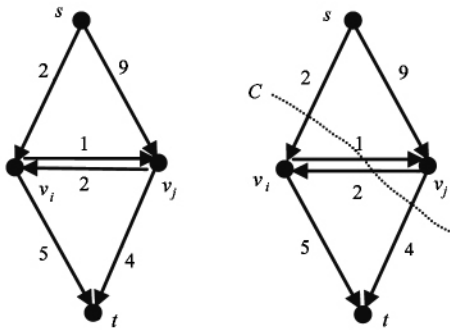


图 3 双终端图及图割

一个割 C 把点集 V 分为集合 S 和 T , 切割的代价是指从集合 S 到集合 T 的所有边的流量之和。要实现最小割,也就是使流量和最小。Kolmogorov 等人^[20]证明了网络最大流和最小割的等价性,所以网络的最小割可以通过计算最大流得到。

吴伟仁等人^[21]将图割算法运用到解决立体匹配的问题中,并提出图割法能够解决视差不连续问题的观点。Boykov 与 Kolmogorov^[22]通过构造能量函数以及改进的最大流方法,取得了良好的致密视差图,但该方法空间和时间复杂度均较高。Hong 等人^[23]为了提高效率,提出基于无重叠视差区域分割的匹配方法,用分割块的能量最小化取代了像素级的能量最小化,但生成的视差图边缘有毛刺现象。Bleyer 等人^[24]基于低尺度分割将图像分割成超像素形式,从而减少图割算法生成节点。Lempitsky 等人^[25]通过将图像进行最优分割,然后对分割后的图像进行分别计算,再进行结合,确保准确率的同时,提高了运行速度。

图割算法精度高,收敛速度快,能很好地解决光照变化和弱纹理区域的低匹配率问题,缺点是计算时间较长,实时性差。

2.1.3 基于置信度传播的立体匹配算法

置信传播算法最早在 1982 年由 Pearl^[26] 提出,该算法与图割法基本思想类似,需要先构造网络。马尔科夫随机场是置信度传播算法的基石,主要通过消息传输机制来实现能量函数最小化。

图 4 所示为一个马尔科夫网络, x_s 为隐藏结点,表示视差; y_s 为可见节点,表示像素在不同视差下匹配代价的向量。

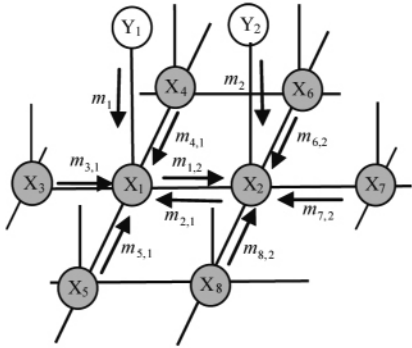


图 4 马尔科夫网络示意图

置信度传播算法在网络中传输信息,并且通常由“最大积”和“和积”2 种方式来对信息进行更新。设 $m_{st}(x_s, x_t)$ 为 x_s 传递到 x_t 的信息, $m_s(x_s, y_s)$ 为 y_s 传递到 x_s 的信息, $b_s(x_s)$ 为 x_s 的置信度,三者均为 1 维向量,则“最大积”的算法可以表示为

(1) 统一分配所有初始信息;

(2) 迭代更新信息;

$$m_{st}^{i+1}(x_s, x_t)$$

$$\leftarrow \kappa \max_{x_s} \psi_{st}(x_s, x_t) m_s^i(x_s) \prod_{x_k \in N(x_s) \setminus x_t} m_{ks}^i(x_s) \quad (6)$$

(3) 计算可信度。

$$b_s^i(x_s) \leftarrow \kappa m_s(x_s) \prod_{x_k \in N(x_s)} m_{ks}(x_s) \quad (7)$$

$$x_s^{MAP} = \arg\max_{x_k} b_s(x_k)$$

2003 年, Sun 等人^[27] 将置信扩展算法应用到立体匹配中, 将能量函数与视差概率结合并引入图像分割作为局部约束取得了很好的效果。Ahmadzadeh 等人^[28] 利用多核系统平台解决了该算法运行时间长的问题。Felzenszwalb 等人^[29] 针对优化有环图结构和像素的标号范围, 提出使用多层置信扩展, 在线性时间复杂度内得到了算法结果。Yi 等人^[30] 根据

层次置信扩展原理, 针对立体匹配中的遮挡问题进行了效果良好的检测。刘忠艳等人^[31] 通过采用分水岭分割图像的方式有效地对图像进行了过分割处理, 实现了基于图像块置信传播的匹配算法, 克服了传统立体匹配中的精度低和速度慢的问题。

置信度传播算法提高了在低纹理区域匹配的精度, 但算法时间复杂度非常高。

2.2 局部立体匹配算法

局部匹配算法出现较早, 其特点是简便快捷、容易实现, 很多时候是不错的选择方案。具体来讲, 根据选取的匹配基元的不同, 局部匹配算法又可以分为基于区域、基于特征、基于相位的匹配算法。

2.2.1 基于区域的立体匹配算法

基于区域的立体匹配算法是最早开始研究的局部匹配算法, 其基本原理是: 首先以左视图中某个像素点为中心, 选取一个窗口, 然后根据相似性条件在对应右视图中寻找最相似的窗口, 左右视图中的对应窗口内的中心像素点即为匹配点。

目前选取窗口的方法大致分 2 种, 一是从多窗口中选择最优, 二是固定窗口的大小和尺寸, 对窗口内的像素增加距离和颜色权重, 即自适应权重方法^[32]。该算法利用颜色权重和几何空间距离权重去描述聚合窗口内不同点对中心点匹配代价的影响。这种方法对像素邻域信息具有很强的抗歧义性, 可以得到与全局算法相媲美的视差图。但聚合窗口选取过大时, 算法效率将变得比较低, 实时性差。

曾凡志等人^[33] 在自适应窗口的基础上, 采用 8 个相同的窗口运用并行处理的方法, 根据图像的平滑情况从 8 个方向选择合适的区域, 该算法解决了在低纹理区域容易造成误匹配的情况, 并且运行时间与传统时间差别不大。Wang 等人^[34] 将区域作为匹配的基元, 通过引入区域之间的合作与竞争, 采用协作优化模式来最小化匹配成本。

基于区域的立体匹配算法较为成熟、运算简单, 但在弱纹理区域容易出现误匹配, 容易受光照、遮挡等影响, 而且窗口的选择也是一个难点。

2.2.2 基于特征的立体匹配算法

基于特征的立体匹配算法, 以图像特征为匹配

基元进行匹配,图像特征包括图像的点、线段、边缘或者局部能力,这些特征必须具有唯一性和再现性,并且有物理意义。该算法的主要流程为:首先进行特征信息的提取,然后根据相似性准则实现左右视图中特征的匹配,获取视差信息。

国内外研究人员对图像特征的提取和匹配这一领域展开了很多研究,并收获了不错的成果,如 SIFT 特征提取算法^[35]、基于特征点的图像配准技术的研究^[36]等。郭龙源等人^[37]将 SIFT 算法引入到边缘特征点的立体匹配中,提高了视差的准确性。狄红卫等人^[38]通过将图像的特征点用 WTA 匹配,非特征点只是简单验证的方式,降低了计算成本。

基于特征的立体匹配算法匹配速度快,对光照变化和噪声等具有鲁棒性,但往往只能获得稀疏视差图,如果想要稠密的立体图则需要进行相应的差值算法。

2.2.3 基于相位的立体匹配算法

相位匹配方法是由 Schaffalitzky 等人^[39]提出的第 3 种方法,该算法的核心思想是在左右视图中匹配的像素点之间的局部相位是相等的,首先选取局部相位为匹配基元,再依次加入合理的约束条件进行图像匹配。因此可根据相位分析法来进行立体匹配。

常见的相位分析法有相位相关法^[40]和相位差-频率法^[41]。相位相关法的理论基础是傅立叶变换的平移定理,并归一化相移后的谐波函数,再经过傅立叶反变换获取相应位置的脉冲函数,从而求出图像对应的位移值。相位差-频率法基于局部频率变化相似的条件,根据相位差异,对局部频率进行定位,确定视差的搜寻趋势。

基于相位的立体匹配算法具有较强的抗畸性而且对噪声不敏感,能较好地反映出图像的结构等信息,处理速度快,但是等相位点的寻找是一个难点。

2.3 稠密立体匹配算法

稠密立体匹配算法是通过计算机得到 2 幅或者多幅图像间尽可能多的对应点,其主要思想是利用局部信息,根据窗口内灰度的相似性来确定是否匹配。

该方法运算速度快,但在纹理稀疏区域和重复

纹理区域容易出现误匹配,而且窗口大小的选择也是一个难题。

2.4 稀疏立体匹配算法

稀疏立体匹配算法,其算法流程是首先提取特征点,然后使用局部匹配算法并加以匹配,得到稀疏匹配点。

这种算法的主要问题在于不能在计算中使用光滑约束,导致一些沿物体表面的视差跳跃,容易在边缘出现错误,得到的视差图质量较低。

3 各种立体匹配算法比较

3.1 定性比较

表 1 为各类算法特性的定性比较。

表 1 各类算法特性比较	
算法名称	算法特性
基于动态规划的立体匹配算法	实时性好,能解决低纹理和边缘区域的问题,条纹现象明显
基于图割的立体匹配算法	实时性差,能解决条纹现象以及光照变化和低纹理区域问题
基于置信度传播的立体匹配算法	实时性差,能很好地处理低纹理区域
基于区域的立体匹配算法	实时性好,容易受光照、遮挡等变化影响
基于特征的立体匹配算法	实时性好,抗畸性好,适用于几何特征明显的图像
基于相位的立体匹配算法	实时性好,能很好地反映图像结构,视差精度可达亚像素级
稠密匹配算法	实时性差,纹理稀疏区域和重复纹理区域容易出现误匹配
稀疏匹配算法	实时性好,容易造成视差缺失

3.2 量化比较

本文选取 Middlebury^[7]网站上的 2 组理想图像 Tsukuba 和 Venus,表 2 和表 3 给出了几种算法在 3 个定量参数下的量化评价结果,即 all(所有区域误匹配百分比)、nonocc(非遮挡区域误匹配百分比)和 disc(不连续区域误匹配百分比),误差阈值限定为 1 个像素。

表格中呈现出各算法的误匹配百分比,误匹配点占总像素点数的比率。误匹配百分比越大,则说

明出现错误匹配的像素点数越多,匹配效果越差。

表 2 Tsukuba 各类算法量化对比

算法名称	all	nonocce	disc
基于动态规划的立体匹配算法	5.04	4.12	12.00
基于图割的立体匹配算法	4.12	1.11	9.39
基于置信度传播的立体匹配算法	1.37	8.71	5.79
自适应权重算法	1.85	1.38	6.90
基于梯度相似的局部适应算法	2.63	2.26	8.99
基于特征的立体匹配算法	6.57	8.43	24.10
基于相位的立体匹配算法	6.53	4.26	15.40

表 3 Venus 各类算法量化对比

算法名称	all	nonocce	disc
基于动态规划的立体匹配算法	11.00	10.10	21.00
基于图割的立体匹配算法	3.44	1.79	8.75
基于置信度传播的立体匹配算法	0.21	0.10	1.44
自适应权重算法	1.19	0.71	6.31
基于梯度相似的局部适应算法	1.39	0.99	4.92
基于特征的立体匹配算法	5.19	4.20	13.50
基于相位的立体匹配算法	8.16	6.71	26.40

另外,表格中提到的自适应权重算法和基于梯度相似的局部适应算法均为局部匹配算法。

因为 2 种划分标准下立体匹配算法有重叠,所以表格中仅涉及到第 1 类划分标准下的各类立体匹配算法的量化结果。

4 结 论

本文主要对双目立体视觉的模型、立体匹配技

术相关原理内容以及当前立体匹配算法的分类进行了介绍,并对全局立体匹配算法和局部立体匹配算法现阶段的研究作出了较为详细的阐述。最后,分别通过定性比较和实验将几种立体匹配算法进行对比,分析出各类算法的优缺点。

立体匹配算法正逐步从实验室走向现实应用,但是纵观各种算法,立体匹配还是存在以下问题:(1) 对于简单的场景已经达到较高的精确度,但在复杂场景中,精确度仍有待提高;(2) 速度快的立体匹配算法精度一般达不到理想状态,而精度高的立体匹配算法往往运算量大,速度较慢。目前还没有一种算法能够兼顾算法性能和算法效率来完美地解决立体匹配问题。要实现图像之间既快速又准确的匹配,并且找出一个具有普适性的算法,仍然有很长的路要走。

参考文献

[1] Barnard S T, Fishler M A. Computational stereo[J]. *ACM Computing, Surveys*,1982,14(4):553-572

[2] 刘志花. 面向立体视觉的匹配算法研究[D]. 北京:中国科学院研究生院,2011:19-28

[3] Lim K B, Kee W L. Geometrical-analysis-based algorithm for stereo matching of single-lens binocular and multi-ocular stereo vision system[J]. *Journal of Electronic Science and Technology*, 2012, 10(2):107-112

[4] 贾云得. 机器视觉[M]. 北京:科学出版社,2000:15-48

[5] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. *International Journal of Computer Vision*, 2002, 47(1-3):7-42

[6] Cyganek B, Borgosz J. A comparatative study of performance and implementation of some area-based stereo algorithms[C]//*Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns*, Berlin, Germany, 2001:709-716

[7] Scharstein D, Blasiak A. Middlebury stereo evaluation-version 2[EB/OL]. <http://vision.middlebury.edu/stereo/eval/>; Middlebury College, 2015

[8] Ke Z, Lu J, Lafruit G. Cross-based local stereo matching using orthogonal integral images[J]. *IEEE Transactions*

- on Circuits and Systems for Video Technology*, 2009, 19 (7):1073-1079
- [9] Geman S, Geman D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, 6(6):721-741
- [10] Geiger D, Girogi F. Parallel and deterministic algorithms from MRFs: surface reconstruction and integration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, 13(5):4011-412
- [11] Kolmogorov V, Zabih R. Computing visual correspondence with occlusions using graph cuts[C]//International Conference on Computer Vision, Vancouver, Canada, 2001:508-515
- [12] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization on via graphcuts[J]. *Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(11):1221-1239
- [13] Klaus A, Sormann M, Karner K. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure[C]//18th International Conference on IEEE, Hong Kong, China, 2006:15-18
- [14] Wang Z F, Zheng Z G. A region based stereo matching algorithm using cooperative optimization[C]//Conference on Computer Vision and Pattern Recognition, Anchorage, USA, 2008: 1-8
- [15] Liu C L, Jaeger S, Nakagawa M. Online recognition of Chinese characters: the state-of-the-art[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, 26(2):198-213
- [16] 罗嗣卿. 基于动态规划的快速立体匹配算法[J]. *计算机工程*, 2015, 41(11):224-231
- [17] Li K, Wang S, Yuan M, et al. Scale invariant control points based stereo matching for dynamic programming [C]// International Conference on Electronic Measurement & Instruments, Beijing, China, 2009: 769-774
- [18] Leung C, Appleton B, Sun C. Iterated dynamic programming and quadtree subregioning for fast stereo matching [J]. *Image & Vision Computing*, 2008, 26(10):1371-1383
- [19] Hu T B, Qi B J, Wu T, et al. Stereo matching using weighted dynamic programming on a single-direction four-connected tree[J]. *Computer Vision & Image Understanding*, 2012, 116(8):908-921
- [20] Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? [J]. *European Conference on Computer Vision*, 2002, 26(2):147-159
- [21] 吴伟仁, 王大轶, 邢琰, 等. 月球车巡视探测的双目视觉里程算法与实验研究[J]. *中国科学:信息科学*, 2011, 41(12):1415-1422
- [22] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision[J]. *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, 2001, 26(9):1124-1137
- [23] Hong L, Chen G. Segment-based stereo matching using graph cuts[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, USA, 2004: 1-8
- [24] Bleyer M, Rother C, Kohli P, et al. Object stereo joint stereo matching and object segmentation[C]//Conference on Computer Vision and Pattern Recognition, Providence, USA, 2011: 3081-3088
- [25] Lempitsky V, Rother C, Blake A. Logcut-efficient graph cut optimization for markov random fields[C]//International Conference on Computer Vision, Rio de Janeiro, Brazil, 2008: 1-8
- [26] Pearl J. Reverend Bayes on inference engines: a distributed hierarchical approach[C]//Proceedings of the 2nd National Conference on Artificial Intelligence, Pittsburgh, USA, 1982:133-136
- [27] Sun J, Shum H Y, Zheng N N. Stereo matching using belief propagation[J]. *European Conference on Computer Vision*, 2003, 25(7):787-800
- [28] Ahmadzadeh A, Madani H, Jafari K, et al. Fast and adaptive BP-based multi-core implementation for stereo matching[C]//11th IEEE/ACM International Conference on Formal Methods and Models for Codesign, Portland, USA, 2013: 135-138
- [29] Felzenszwalb P F, Huttenlocher D P. Efficient belief propagation for early vision[J]. *International Journal of Computer Vision*, 2006, 70(1):41-54
- [30] Yi D, Yang Q, Lin X Y, et al. A symmetric patch-based correspondence model for occlusion handling[C]//10th IEEE International Conference on Computer Vision, Beijing, China, 2005:1316-1322

- [31] 刘忠艳,周波,张兴华,等. 一种基于置信度传播的立体匹配算法[J]. 自动化与仪器仪表, 2010(1):111-113
- [32] 殷虎. 基于图像分割的立体匹配算法研究[D]. 南京:南京航空航天大学自动化学院, 2010: 23-42
- [33] 曾凡志,鲍苏苏. 一种自适应多窗口的立体匹配算法[J]. 计算机科学, 2012, 39(s1):519-521
- [34] Wang Z F, Zheng Z G. A region based stereo matching algorithm using cooperative optimization[C]//Conference on Computer Vision and Pattern Recognition, Anchorage, USA, 2008: 1-8
- [35] 张洁玉. 图像局部不变特征提取与匹配及应用研究[D]. 南京:南京理工大学自动化学院, 2010: 25-70
- [36] 丁南南. 基于特征点的图像配准技术研究[D]. 中国科学院长春光学精密机械与物理研究所, 2012:29-50
- [37] 郭龙源,孙长银,杨万扣,等. SIFT 特征点引导的区域立体匹配算法[J]. 计算机工程与应用, 2013, 49(4):23-25
- [38] 狄红卫,柴颖,李逢. 一种快速双目视觉立体匹配算法[J]. 光学学报, 2009, 29(8):2180-2184
- [39] Schaffalitzky F, Zisserman A. Viewpoint invariant texture matching and wide baseline stereo [C] // International Conference on Computer Vision, Vancouver, Canada, 2001: 636-643
- [40] Hutter J L, Bechhoefer J. Calibration of atomic-force microscope tips[J]. *Review of Scientific Instruments*, 1993, 64(7):1868-1873
- [41] Schaffalitzky F, Zisserman A. Multi-view matching for unordered image sets, or ‘how do i organize my holiday snaps?’ [C] // European Conference Computer Vision, Berlin, Germany, 2002:414-431

A survey of stereo matching algorithms

Chen Hua, Wang Lijun, Liu Gang

(Key Laboratory of Measurement Technology and Instrumentation of Hebei Province,
Yanshan University, Qinhuangdao 066004)

Abstract

The main task of binocular stereo vision is to take 2 views by shooting the same scene from 2 cameras in different positions, then calculate the parallax and obtain the 3-dimensional data information. Stereo matching is a key link in binocular stereo vision. It uses 2 or more matched images to obtain parallax to further obtain depth information. This paper introduces researches on binocular matching algorithm at home and abroad, analyzes the stereo matching algorithm, and elaborates the global matching algorithm and local matching algorithm. Finally, the characteristics of various algorithms are compared, and the conclusions and prospects are made.

Key words: binocular stereo vision, stereo matching, global matching algorithm, local matching algorithm