

文章编号: 1001-0920(2008)07-0721-09

双目立体匹配算法的研究与进展

白 明, 庄 严, 王 伟

(大连理工大学 信息与控制研究中心, 辽宁 大连 116024)

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. 首先综合介绍了立体匹配算法的研究概况, 论述了双目立体匹配算法中各种约束的核心概念和适用范围; 然后重点归纳分析了立体匹配算法的分类及其发展过程中的各种演化算法, 对其关键技术进行了剖析和比较, 并总结了目前存在的主要难题和可能的解决途径; 最后对该领域存在的问题和技术发展趋势进行了分析和讨论.

关键词: 立体视觉; 双目立体匹配; 匹配约束

中图分类号: TP24

文献标识码: A

Progress in binocular stereo matching algorithms

BAI Ming, ZHUANG Yan, WANG Wei

(Research Center of Information and Control, Dalian University of Technology, Dalian 116024, China.

Correspondent: WANG Wei, E-mail: wangwei@dlut.edu.cn)

Abstract: Stereo matching has long been one of the central research problems in computer vision. Firstly, some current studies on binocular stereo matching technology are introduced. Then the concepts and applicable situations of the matching constraints are presented. Especially, the categories and the improved algorithms in the process of the development are emphatically analyzed. The key techniques are compared and addressed, and some major problems and possible solutions are also summarized. Finally, the focus and some existing problems of the current research in this area are pointed out, and the prospective technique development trends are proposed.

Key words: Stereo vision; Binocular stereo matching; Matching constraints

1 引 言

立体视觉技术是移动机器人视觉导航领域的关键技术之一. 在未知环境中导航, 快速、准确地识别障碍物信息, 规划出可行路径, 要求立体视觉导航系统的匹配算法应具有较高的准确性和实时性. 然而双目立体匹配一直是三维场景结构信息获取的研究热点之一, 其基本原理是从两个视点观察同一景物以获取立体像对, 匹配出相应像点, 从而计算出视差并获得三维信息. 特别是 20 世纪 80 年代, 美国麻省理工学院的 Marr 提出了一种视觉计算理论并应用在双目匹配上, 使两张有视差的平面图产生具有深度的立体图形, 从而奠定了双目立体视觉发展的理论基础.

由于匹配的不确定性, 立体匹配问题始终没有统一的解决方式, 对此许多研究人员作了大量研究.

Barnard^[1]回顾了立体重构的基本原理, 对主要方法作了综述, 并给出算法性能的评价标准. Dhond^[2]对大量立体匹配算法进行了总结, 介绍了分级处理思想, 并提出使用三目约束减少立体匹配的不确定性. Faugeras^[3]对在非摄像机校正条件下立体视觉的研究进行了详细综述. Koschan^[4]对动态立体和主动立体、早期的遮挡和无纹理问题及实时立体视觉实现进行了总结与探讨. Scharstein^[5]对由两帧图像产生稠密视差的立体匹配算法进行了详细分类和评价, 提出现有立体匹配算法一般由 4 个相对独立的模块组成: 匹配代价计算、代价/支持聚合、视差计算/最优化以及视差校正, 认为当前各种匹配算法是模块相对变动的结果. 通过对 4 种模块不同算法之间的性能比较, 可得出变动因素对算法性能的具体影响, 并进一步依据其分类法则构建了对立体匹配

收稿日期: 2007-03-12; 修回日期: 2007-06-11.

基金项目: 国家自然科学基金项目(60605023, 60775048); 高等学校博士学科点专项科研基金项目(20060141006).

作者简介: 白明(1979—), 男, 河北邢台人, 博士生, 从事移动机器人立体视觉导航的研究; 王伟(1955—), 男, 辽宁鞍山人, 教授, 博士生导师, 从事预测控制、智能控制等研究.

算法性能进行评价的通用软件平台. Brown^[6]从几何计算的角度概述了对应点匹配算法,重点分析了遮挡问题的匹配解决方法,并论述了典型应用中实时立体视觉的实际平台.

本文系统地归纳了基于图像几何和利用场景物体的各种约束的概念及其适用场合,着重分析了各类双目视觉立体匹配算法的发展历程,比较了其优缺点,并总结了其存在的主要难题和可能解决的途径.最后讨论了双目立体匹配研究所面临的主要问题和技術发展趋势.

2 基于图像几何和场景的约束

对于两幅从不同角度获取的同一场景的图像而言,传统的特征点搜索方法是在一幅图像上选择一个特征点,然后在第2幅图像上搜索对应的特征点. Poggio 将 Marr 的思想进一步发展为早期视觉不适定性问题,他认为立体视觉匹配问题可以表述为 Hadamard 意义下的不适定性问题,这一表述体现了约束假设的重要性.为减少匹配不确定性,提高匹配效率,众多研究人员应用了约束假设.这些约束可分为两类:1)基于图像几何的约束,例如位置、朝向以及图像获取方式;2)基于场景的约束,例如场景中物体的性状和相互关系.大多数文献仅仅依托某些约束,并未对其明确地定义.本文对目前较活跃的匹配算法中常用到的约束进行总结,并进一步说明了约束的适用范围.

2.1 基于图像几何的约束

基于图像几何的约束的依据主要是图像获取过程中的几何学和光度测定学原理.本文将这类约束分为以下5个部分:

1) 极线约束:设空间中一点 P ,在左(右)像平面 $I_l(I_r)$ 上的投影为 $P_l(P_r)$,则 P_l 点在像平面 I_r 上的对应点一定位于极线上.这样便将潜在的二维搜索空间降为一维.

2) 唯一性约束:图像上每一个像素点最多只能与另一幅图像上的一个像素点对应.另外,出现两个或以上点在一个图像上的投影点相同,但在另一个图像上的投影却作为分开的点,如自遮挡问题.

3) 几何相似性约束:两幅图像中发现的特征的几何特性相差不大.例如区域或轮廓,线段的长度或方向.此项是基于特征的匹配算法^[7]的核心.

4) 光度测定学相容性约束:由于光源、表面法线和摄像机位置之间角度的关系,在两幅图像中对对应点的灰度值不可能完全相同但相似.反之,具有相同或相似灰度值的图像点未必是空间同一点的投影.相容性是对应像点一致性的测度,目前文献中采用的相容性测度有:相关函数测度、协方差函数测

度、相关系数测度、差平方和测度以及差绝对值和测度,其中相关系数测度对灰度的线性畸变有较强的鲁棒性.区域匹配算法^[8]和通过定义能量函数得到稠密视差图的算法^[9]几乎都依托此约束. Jordan 进一步将此约束扩展到彩色光度测定学相容性约束,指出在两幅彩色图像中,匹配点在相应的色彩通道上有相似值.

5) 透视投影约束:大部分匹配算法依据透视投影建立图像模型,仅少数文献采用正投影模型.

2.2 基于场景的约束

基于场景的约束的依据主要是利用典型场景中物体的某些共同性质.这类约束可分为以下8个部分:

1) 视差光滑性约束:除图像中物体边缘外,视差在图像中几乎所有地方变化缓慢.即假设空间中同一物体表面上两点相邻,则它们分别在左右图像上的投影都是相邻的.但位于空间不同物体上的点,该约束则不成立.实际上,图像扫描线优化算法本质上是在该约束条件下的基于动态规划匹配算法.

2) 轮廓视差约束:轮廓对应点在满足视差光滑性约束的同时,在左右图像中都应落在边缘处.

3) 特征相容性约束:仅具有相同物理起源的点才能匹配.以物理起源分类,图像中边缘可分为物体面边缘、投影的阴影边缘、遮挡边缘和镜面性边缘.其中:前两种边缘在匹配时满足此约束,后两种边缘因随视点改变而变化不能用于解决对应问题.边缘匹配是场景三维重建的常用技术,但关键取决于边缘分类. Lee 通过利用边缘分类的色彩信息得到了较好的匹配结果.

4) 顺序约束:对于相似深度的表面,对应匹配点一般以相同的次序落在极线上.在动态规划算法中,应用此约束能减少遮挡产生的误匹配^[10].但对场景中深度相差较大的窄物体,此约束不再满足.

5) 互对应约束:首先从左图像点 P_l 开始搜索,得到匹配点 P_r ;反过来,从图像点 P_r 开始搜索,若没能找到 P_l ,则匹配不可靠,应排除.此约束有助于排除由于遮挡和噪声等原因而产生的误匹配点.用该约束检测遮挡优于许多其他方法,但耗时较长是其主要缺点.文献[8]提出一种虽依据该约束但时间复杂度较小的稠密算法.

6) 相位约束:两幅图像中对应点的局部相位是相等的.该约束是基于相位匹配算法^[11]的核心.尤其对图像中高频噪声有一定的抑制作用,并且相位匹配算法无需亚像素插值或亚像素特征提取就能达到亚像素级精度.

7) 视差范围约束:人为指定两幅图像中对对应点

的视差小于事先定义的阈值. 该约束是限制搜索对应点长度和可靠性^[10]的有效工具, 可同时提高效率和准确性.

8) 视差梯度范围约束: Burt 从生理学调查中发现, 人类视觉只能融合视差梯度不超过某个范围的立体图像. 如果某些点的视差梯度过大, 则这些点应是遮挡点或误匹配点. 文献[12]以该约束集成到目标函数最小化过程为例, 证明了所提出的连续优化下能量函数最小化框架的有效性.

3 双目立体匹配算法

从各具特色的双目立体匹配算法框架来看, 算法有效性主要依赖 3 个因素, 即选择准确的匹配基元、寻找相应的匹配准则和构建能够准确匹配所选基元的稳定算法. 但由于立体匹配涉及诸多因素, 至今仍未得到很好的解决. 一般情况下, 可通过算法所使用的匹配特征类型、约束假设、匹配准则以及优化或搜索方法将其分类. 针对已提出的多种立体匹配方法, 本文认为构建一个完整的立体匹配算法应从 4 个方面着手: 1) 广义匹配基元与特征参数选择; 2) 匹配方法及策略; 3) 匹配代价优化标准; 4) 视差提炼方式. 本文将立体匹配算法分为两类: 基于局部约束的算法和基于全局约束的算法.

3.1 基于局部约束算法

基于局部约束算法利用兴趣点周围的局部信息进行计算, 涉及信息量较少, 相应的计算复杂度较低, 大多实时性平台借鉴了此算法的思想. 但其对噪声较敏感, 对无纹理区域、视差不连续区域和遮挡区域匹配效果不理想. 本文将局部算法分为 3 类: 区域匹配算法、特征匹配算法和相位匹配算法.

3.1.1 区域匹配算法

本质上, 基于光度测量学不变性原理的区域匹配算法, 常以邻域窗的灰度为匹配基元, 以相关程度作为判别依据, 可以得到较稠密的视差图. 本文从灰度相关性和非参数变换思想两个方面进行分类.

灰度相关性算法的研究主要集中在匹配代价聚合的窗口构建上, 主要包括: 固定窗算法、多窗口关联算法和自适应窗算法. 本文认为算法性能依赖于窗口的两个性质: 其一, 窗内像素点最好具有同一深度, 不同深度的像素点应尽量少; 其二, 窗内应包含足够多易于识别计算的灰度变化.

自适应算法思想是以灰度值和深度值作为自变量构造尺度函数, 进行窗的选择. 通过比较不同尺寸窗口的性能, Veksler^[13]提出不同形状尺寸窗口的适用范围, 但只是对有限的窗口形状进行评价, 且需用户指定窗口测量计算的参数. Wang^[14]根据分割区域动态选择窗口的形状和大小的方法得到了较好

效果, 但对于复杂纹理图像, 精确的色彩分割是个难题. Yoon 和 Kweon^[15]利用像素间光度学和几何学关系调配窗口中像素的权值, 本质上类似于分割的思想, 但却回避了分割难题. 此方法不依赖初始视差估计, 同时在深度不连续和连续区域得到精确匹配, 但其高时间计算复杂度问题有待解决.

基于局部非参数变换思想是为了使后期匹配更具有鲁棒性. 典型算法是 Rank transform 和 Census transform. Rank transform 思想类似滤波器, 通过对待匹配点与特征窗内各点灰度差, 定义灰度等级. Zabih^[16]把等级变换算法推广到 Census transform, 通过编码像素间的等级信息保留其空间分布, 用加权平均编码信息匹配图像. 由于这种方法依据灰度排序方式, 提高了算法的鲁棒性, 但同时增加了窗口图像信息的维度, 计算更加复杂.

区域匹配算法的一个突出缺点是, 对无纹理区域常常由于相关函数变化的锐度不够以及难于保留深度不连续特性, 不能取得精确的匹配结果.

3.1.2 特征匹配算法

本质上, 基于几何不变性原理的特征匹配算法克服了区域匹配算法对深度不连续和无纹理区域敏感的缺点, 特别是由于特征基元的统计特性和数据结构的规则性, 使其适合于硬件设计, 备受学者关注. 按匹配基元的特征可分为全局特征和局部特征两大类. 全局特征包括多边形和图像结构等, 多与下文中的全局算法混合使用; 局部特征算法常用点、边缘、线段、小区域或局部能量等图像信息作为匹配基元, 对噪声不敏感, 可以得到较精确的匹配.

近年来, 众多应用要求稠密视差图, 人们开始研究一些新型算法. 这些算法^[7]的共同点是先对图像进行加工, 对潜在的新型特征进行匹配, 使计算复杂度和匹配效果得到不同程度的改进. 本文主要讨论较受关注的两类算法: 分级特征匹配和基于分段的特征匹配.

分级特征匹配算法^[7]的思想是根据图像结构特点提取边缘特征, 由低层次到高层次特征构建特征分组. 首先用高层次的特征进行粗糙匹配, 并指导后续匹配; 然后对没有匹配的区域用低层次的特征进行再匹配; 最终算法将所有信息融合, 产生的视差图将满足全局一致性约束条件. 这样逐层分级、由粗到精的策略显著减少了搜索空间, 从而减少了计算复杂度.

基于分段的特征匹配算法^[17]的思想是, 首先采用多路分割思想^[18]把图像分割成不同的段, 对同一区域像素加以相同标号; 然后迭代计算两幅图像对应段的仿射转换参数, 合并每次迭代后具有相同参

数的段.因这类算法基于在同一段中视差连续的假设,故减少了整体深度不连续的敏感性,保证了每段视差光滑,但如何解决提高分段质量与减少计算代价之间的矛盾,是提高其性能应首要考虑的问题.

特征匹配主要存在两点不足及其改进措施:其一,图像特征的稀疏性决定特征匹配只能得到稀疏视差图,要得到稠密视差需附加较为复杂的插值过程.因此常规的特征算法往往难以达到预期效果,结合各算法优点的混合特征算法值得关注. Kim^[19]采用分级思想有效融合了区域匹配、特征匹配和能量最小化 3 种算法.其二,特征匹配结果的性能紧密依赖于特征提取的精度. Prince^[20]摒弃了建立提取特征先验模型的步骤,而用局部能量法识别多方向的亚像素特征,检测出多种类型的特征用于匹配,这种非梯度算子不会放大高频噪声而引起特征的虚假提取.

3.1.3 相位匹配算法

基于傅立叶平移定理的相位匹配算法的本质是,对带通滤波后的时/空-频域定位性的基元信号相位信息进行处理而得到像对间的视差.与上述两类方法相比,作为匹配基元的相位信息本身反映了信号的结构信息,能有效抑制图像的高频噪声和畸变,且适于并行处理,可获得亚像素级精度的致密视差.

究其本质,相位匹配就是寻找局部相位相等的对应点,因此出现了窗内相位差异积分最小法和以加权相位差绝对值作为代价函数的动态规划法.本文主要讨论目前常用的两类算法:相位差-频率法^[11]和相位相关法^[21].

相位差-频率法利用空间域上的平移与频率域上的相位平移之间的比例关系计算对应点处的视差值.徐奕^[11]利用双正交小波基构造滤波器,采用多尺度匹配策略的动态规划技术,得到最优的匹配路径.相位相关法根据带通信号之间的相关度确定相位重叠的匹配位置. Muquit^[21]提出对任意形状的表面进行重构的算法,采用多尺度搜索策略对孤立点进行检测和校正,得到可靠的亚像素级精度.

相位匹配算法主要存在两个问题及其改进措施:其一,因带通输出信号的幅度太低而带来相位奇点问题. Fleet 和 Jepson^[22]详尽地分析了相位奇点产生的原因并实现了对其有效检测,利用邻域视差信息解决奇点问题.其二,由于相位匹配算法的收敛范围与带通滤波器的波长有关,需考虑相位卷绕问题.徐彦君^[23]采用低冗余度的质数序列作为 Gabor 滤波器组的波长,实现了高效的尺度自适应选择算法,将相位匹配算法的收敛范围扩大,解决了相位卷

绕问题.

3.2 基于全局约束算法

基于全局约束算法利用对应扫描线或整个图像数据信息进行计算,着重解决图像中不确定区域的匹配问题,能达到全局最优解.全局最优算法的本质是将对应点的匹配问题转化为寻找某一能量函数的全局最优问题,通常跳过代价聚合步骤,直接计算视差值.这类算法的核心环节包括:1) 能量函数构造方法;2) 能量函数优化求解策略.文献[5]评价了各种优化策略的性能,指出动态规划算法在满足对应点顺序约束的条件下能最快地实现全局最优搜索;基于二维马尔可夫过程的图割方法性能最好,但耗时太长.本文把全局算法分为 4 类:动态规划算法、图割算法、人工智能算法和其他全局算法.

3.2.1 动态规划算法

基于动态规划算法的本质是在左右图像对应扫描线上寻找最小匹配代价路径的过程.常在视差空间图中应用连续性约束和顺序约束,由全局代价函数决定匹配代价,规划出的路径由趋于具有最小代价的匹配点集合构成.动态规划算法能有效降低计算复杂度.假设扫描行上有 N 个像素,视差搜索范围为 D ,则求解所有像素视差的计算复杂度为 $O(D^N)$,而利用动态规划的计算复杂度仅为 $O(ND^2)$.文献[24]进一步把计算复杂度降为 $O(ND\log D)$.另外,动态规划算法为局部无纹理区域提供了全局支持,从而提高了匹配精确度.

动态规划算法的最大局限是不能有效融合水平和垂直方向连续性约束.单点像素的错误匹配将会影响同一扫描线上后续像素的匹配,不仅使视差图上带有明显的条状瑕疵,而且造成匹配精度较低. Ohta^[25]提出了以边缘线为匹配特征的动态规划匹配算法,首次把动态规划的搜索空间扩大到三维. Cox^[26]引入水平方向不连续和垂直方向不连续最小的约束来获得高密度视差图,改善了动态规划匹配效果,计算复杂度仅为 $O(ND)$.另外,当场景中存在较窄的前景物体时,顺序性约束将很难满足,常会出现误匹配,并且决策遮挡像素合适的匹配代价也是一个难题.

目前涌现出具有良好前景的几种改进动态规划算法. Bobick^[27]提出了利用事先确定的正确匹配点作为匹配控制点的修正算法,在动态规划过程中对寻优过程进行指导,从而减轻了视差图上的条纹瑕疵. Gong^[10]另辟蹊径,引入了可靠度的概念,对采用双向动态规划得到的匹配路径进行可信度计算和阈值比较,得到了光滑可靠的视差图.值得一提的是, Lei^[28]将动态规划搜索推广到树形结构,在二维

区域树上定义能量函数以寻找最优匹配路径,将区域算法融合到动态规划的框架中,充分发挥二者的优点,得到了满意的性能。

3.2.2 图割算法

为解决动态规划算法不能充分融合水平和垂直方向连续性约束问题,将匹配问题看成是利用这些约束在图中寻求最小割问题。根据 Ford 和 Fulkerson 的最大流-最小割理论,这样最小割问题便可转化为最大流问题来处理。

图割算法的基本思想是将立体匹配问题转化为一种能量函数的形式,根据能量函数构造合适的图,求其最小割(最大流)。Cox^[29]利用图割算法求解立体匹配时,设计的能量函数首次考虑了遮挡情况,并与文献[26]中动态规划算法的性能进行了比较。图割算法因需要多次迭代,一般比动态规划算法的计算复杂度高,但能有效地去除条纹瑕疵。

最近研究主要体现在新型构图框架和能量最优化算法上。图割算法的特性表现在:其一,构造能量函数用于图割求解时,其本身必须满足一定的约束条件,Kolmogorov^[9]对可用的能量最小化函数类型进行了总结归纳;其二,不同的能量函数有着不同的构图方式。一个值得注意的工作是,Boykov^[30]提出了能量函数优化的两种构图标号函数: - swap move 和 expansion move,并证明了其扩张算法所获得的局部小和全局小相差一个已知的常数;而交换算法可以处理更一般的能量函数。文献[31]进一步将标号集扩展到二维向量,并重新定义了能量函数。Boykov^[32]提出一种改进的路径扩张算法,通过把算法应用于图像恢复、图像分割、立体匹配和运动检测方面,表明该算法优于常用的求最大流(最小割)的 Dinic 的算法和 push-relabel 算法。

3.2.3 人工智能算法

Marr 提出的重要理论之一是立体视觉的不确定性,而且立体匹配问题可以看作是代价函数的优化问题。智能算法的基本思想是在约束条件下,首先建立以基元匹配的相似性条件和相容性条件为基础的最优准则;然后进行点集的最优匹配搜索,使各基元的相似性和相容性达到总体最优。人工智能算法对最优问题的解决有其独特的性能,本文主要讨论两种智能搜索算法:基于神经网络的匹配算法和基于遗传算法的匹配算法。

(1) 基于神经网络的匹配算法

该算法的本质是根据所构建网络的形式,通过迭代学习算法将匹配代价函数及其匹配固有约束转化成能量最小化寻优过程,网络的动态变化过程是多个约束的极小化实现过程。如前所述,立体匹配是

复杂的非线性过程,而畸变和环境因素带来的非线性误差将通过网络学习被分散到各神经元之间的连接权值上,不仅能减少非线性因素的影响,而且降低了对系统的要求,训练好的网络几乎可以进行实时测量,获取物体的三维信息。

目前,利用二维 Hopfield 网络寻优的算法较多。Ruichek^[33]提出了利用 Hopfield 网络实时障碍物检测的立体匹配算法。文中通过边缘特征在 3 个约束下的对应关系构建目标函数,采用分级提取边缘的思想实现实时匹配。较多研究针对灰度图像进行匹配,而 Hua^[34]基于竞争-合作神经网络对彩色图像直接匹配。作者采用区域算法对 7 种色彩空间进行了对比实验,指出其中 3 种色彩空间利用自适应颜色特征能有效地改善匹配性能。

然而,二维 Hopfield 网络拓扑无法刻画整个 2D 视差图景演化过程。为满足整体最优的匹配,构造一个三维 Hopfield 网络拓扑是个有意义的工作,但至今未见公开报道。

(2) 基于遗传算法的匹配算法

该算法的本质是根据采用的匹配基元和基因构建策略,将问题的解编码到染色体中,考虑匹配约束构造适当的适应度函数,利用进化机制获得匹配最优解,是一种利用自然选择和进化思想在高维空间寻优的方法。本文认为此类算法性能差异主要体现在 4 个方面:1) 基因构建策略;2) 染色体编码方式;3) 匹配约束下适应度函数的构建;4) 进化机制的实施。

Wang^[35]直接以视差值编码,利用 4 个匹配约束构建目标函数,在没进行特征提取和后续插值的情况下得到平滑稠密视差图,减少了匹配模糊性和时间复杂度。为减少误匹配,提高视差图的精度,Gong^[36]引入基于邻域的分割思想,通过优化对应点的相容性和视差连续性,利用 Markov Random Field(MRF)表征邻域像素交互关系,构建适应度函数,在边缘处理上优于文献[37]的算法。Ruichek^[38]采用改进的整数编码机制,弥补了二进制编码匹配不确定性较大的不足,在 3 个加权约束下构建适应度函数,采用基于边缘的多级搜索机制,提高了匹配正确率和寻优效率,把基于遗传的立体匹配算法应用到实时障碍物检测中。

3.2.4 其他全局算法

除上述全局优化算法之外,众多学者还提出了一些其他算法,如协作算法、非线性扩散算法和置信度传播算法。

协作算法的灵感源于人类视觉计算模型,其所用的非线性迭代操作在整体行为上类似于全局算

法. Zitnick^[37]极具创意地改进了 Marr 原始协作算法.

非线性扩散算法的研究热点集中在提高遮挡问题的处理性能上. Scharstein^[39]采用多种非线性扩散模型,自适应地聚合匹配代价. Lee^[40]基于 MRF 模型提出了分级随机扩散算法,利用邻域的概率分布在 Gibbs 潜在空间迭代扩散,减少了计算负荷,提高了收敛速度,算法通用性较好.但对于较复杂的 Tsukuba 图像物体边缘估计不佳.

Sun^[41]把问题考虑成在 Markov 网络框架下的全局匹配问题,用 Bayesian 置信度传播算法,在整个图像区域得到较好的结果.文献[5]把此算法与图割算法作了细致的比较.由于传统算法传播单条消息的复杂度与标签数量是二次关系,而 Felzenszwalb^[42]提出的实时算法将其改进为线性关系,并且利用减少消息传播时间的分级思想将消息传播迭代次数降为常数,此算法能得到全局算法的精度和局部算法的效率.而 Klaus^[43]进一步利用文献[42]置信度传播算法优化分割后的平面标号,对彩色图像分割算法进行了改进,对每个单色平面采用自适应算法增加可靠匹配几率,并用4组数据进行了性能实验比较.

4 研究重点和技术发展趋势

4.1 研究重点

双目立体匹配问题是一个“病态”问题.双目立体匹配实施要考虑诸多因素,并以计算复杂度和稳定性等总体性能指标来衡量方案实施的可行性和有效性.因此,研究双目立体匹配算法要从系统化的角度来分析,研究也应从以下几个方面有所侧重:

1) 多种立体匹配算法有效融合.单一匹配算法往往存在局限性,而多种算法思想的有效融合能带来性能上的改善.特征匹配算法只能对特征点处建立稀疏视差;区域匹配算法计算量较大,对遮挡、无纹理和深度不连续等复杂图像得不到正确结果;梯度优化算法容易收敛到局部极小值,对初始值要求较高;动态规划算法在计算量不增加的前提下,不能很好地融合水平和垂直连续性约束;图割算法计算量大,实时性难以满足.因此,在完善单一匹配方法的同时,应系统地综合各种技术并加以应用.例如,将特征匹配的鲁棒性和区域匹配的致密性充分结合;将局部算法融合到全局算法框架中进行整体优化.

2) 研究新颖的特定匹配约束以及合理变通利用匹配约束.挖掘现实场景和任务的某些固有属性,将其表示成匹配所必须遵循的若干准则,用以提高系统的去歧义能力和计算效率,是提高匹配性能的

关键.此外,根据场景特点的不同,具体使用约束方式时应体现多样性.例如,在匹配算法中常利用匹配连续性约束将匹配视差限制在平滑解空间.变通此约束要考虑3个方面:如何有效地保护视差表面的不连续信息;如何避免出现局部最优的问题;如何检测匹配不可见的遮挡区域.文献[10]提出了新颖的可靠性约束和强、弱连续性约束,并被迭代过程有效融合.文献[15,44]把光度学和几何关系经过加权建立了一种新型约束,在此约束下得到引人注目的匹配效果.

3) 能量函数构造及其最优求解策略是全局算法的核心问题.立体匹配问题是代价函数的全局寻优问题,根据匹配策略和优化标准构建反映匹配本质的能量函数.利用图像和场景信息构建合理能量函数及其保证全局寻优的执行策略,是提高这类算法性能的关键所在. Goulermas^[12]提出一种连续性问题的能量最小化框架,采用约束的非线性规划算法构建基函数,把场景约束和最优化过程融为一体.文献[45]提出基于能量函数的分割算法,通过判断区域能量处理过分割,从而达到全局最优.文献[46]进一步考虑了分割区域间遮挡性能,同时改善了有倾斜表面的自然场景的无纹理和不连续区域性能.

4) 纹理缺乏、特征缺失、遮挡和不连续的图像区域,是立体匹配研究中的一个难点和重点.对视差图进行修补无疑是进一步改进性能的方法之一.目前匹配修补算法中只用到了待补全点处的信息,更有效地使用已知视差信息是一个值得进一步研究的课题.近年来,人们对遮挡区域的检测和测量,乃至恢复遮挡区域的正确视差估计等方向进行了大量的研究,提出了若干遮挡问题的解决方案.其中主要有3种方法^[6]:降低算法对遮挡的敏感性,从而减轻遮挡的不良影响^[47];对遮挡进行检测并标示出遮挡的存在,虽然并不对其进行解决,但可以避免由其引起的错误^[48];为遮挡问题建立专门的几何模型,对于遮挡问题进行特别处理^[24,46].

5) 对匹配算法评价方法的研究.目前对立体匹配算法的评价还没有统一的方法.一般从以下4个角度评判各类匹配算法:精度:定量地表示正确匹配的程度;复杂度:设备及计算量等的代价;可靠性:表示排除总体分类误差的程度;通用性:对不同场景的匹配能力.对此,可以考虑3种评价思路:其一,最直接的评价是将匹配获得的视差结果与手工实测数据进行比较,有助于评价算法对单一具体问题的处理能力;其二,若手工实测数据不易获取,如果已知成像模型,则可利用预测误差来评价算法的特性;其三, Scharstein^[5]提出的相对独立模块

变动的评价分类法,采用 16 种误差统计参数分析评价每个模块变动因素对算法性能产生的具体影响以及对图像不同性质区域的性能。

4.2 技术发展趋势

虽然双目立体匹配技术经过数十年的研究,匹配算法不断丰富,但无论是从视觉生理的角度,还是从实际应用方面,完善现有的立体匹配技术还有很大空间。根据当前双目立体匹配算法的研究状况,技术发展存在如下趋势:

1) 完善双目立体匹配总体理论框架。立体匹配经过多年的研究已出现大量各具特色的研究成果,但 Marr 提出的理论框架还有待进一步的完善和改进。特别是在复杂场景中,如何提高算法的去歧义匹配和抗干扰能力,如何降低计算复杂度和运算量,仍需要进行更加深入的探索和研究。

2) 立体匹配与重建相结合。对于一个完整的立体视觉系统而言,匹配与内插重建过程并不是两个不相关的独立模块,它们之间存在着信息的反馈。匹配结果约束着内插的重建;反过来,重建结果又引导着正确的匹配。基于这一思想,Hoff 等提出了一种匹配、重建一体化的方法,在得到匹配的同时,用一次/二次曲面拟合技术进行内插,重建出完整的视差表面,并以此为信息不断修正匹配结果,循环往复,最终得到精确的匹配视差。这种将匹配与重建视为一体的方法是值得重视的,它不但能生成更为完整的视差场,而且可以为匹配提供更多的信息。

3) 充分挖掘人工智能算法的发展空间,建立基于知识、规则、模型的智能化匹配算法。作为有效求解不确定性问题的最优化工具,人工智能算法越来越多地应用于解决立体匹配问题,它们提供了各类求解复杂系统优化问题的通用框架,以此建立有目的和面向任务的智能化匹配算法,在实际应用中有很强的鲁棒性。另外,对于匹配左右图像亮度不一的立体对应对,如 birch tree 图像,也是智能匹配算法设计的一个挑战。

4) 建立实时、精确以及稳定的实用立体匹配算法。许多立体匹配算法需要处理的数据量较大,相应地需要存储空间较大,不易实现快速处理。随着大规模集成电路和专用硬件处理器的飞速发展,从实际应用和要求出发,致力于建立适用专门领域和面向对象的实用立体匹配算法是很有意义的研究课题。研究工作可以考虑算法从以前的串行处理向并行处理发展,采用并行流水线机制和专用的信号处理器件,增强立体匹配的实用性^[49]。

5) 立体匹配与认知科学紧密结合。已有的立体匹配方法大多与人类视觉极为相似。许多学者对研

究人类视觉计算模型比研究立体视觉系统更感兴趣,希望立体视觉更加自然化,更加接近生物视觉。认知科学长期将人类视觉作为主要研究对象,总结了许多规则,因此自然成为立体匹配研究学习的主要对象。值得一提的是,认知科学中的图形组织规则对立体匹配中的图像分割问题有着重要的指导意义。文献[45,50]提出了利用格式塔规则提取具有组织结构特征群作为高级基元进行分级的立体匹配算法。Desolneux^[51]全面指出了格式塔原理与计算机视觉的内在联系。

6) 从理解人类视觉的立体融合机制出发,建立一种通用的人类双目匹配计算模型。尽管匹配研究已取得一些成果,但在理论上和技术上都存在着很多问题。例如,如何选择合理的匹配特征,以克服匹配准确性与恢复视差全面性之间的矛盾;如何选择有效的匹配准则和算法结构,以解决存在畸变、噪声、特殊结构及遮挡的匹配问题;如何建立更有效的图像表达形式和立视模型,以便更充分地反映景物的本质属性,为匹配提供更多的约束信息,降低立体匹配的难度。就目前的研究而言,要构造出与人类视觉一样灵活、精确的匹配系统,还有待于对人类自身机制的深入研究。

5 结 论

立体匹配技术是一门交叉学科,有着广阔的工程应用前景,它与认知心理学和视觉神经系统科学都密切相关。本文主要综述了双目立体匹配算法问题,在归纳了立体匹配中关键约束条件的基础上,总结了较为典型的基于局部信息的算法和基于全局信息的算法,并讨论了当前双目立体匹配技术的研究重点和发展趋势。目前,立体匹配算法的设计只是依靠视觉推理的统计理论知识解决视觉不适定问题。应该指出,立体匹配问题的进一步研究面临着理论和实践两方面的挑战——视觉行为的定量理论以及自然场景和认知行为的复杂性,因此立体视觉匹配算法的设计还存在许多未知的理论依据。随着研究的不断深入和完善以及人类视觉机理的不断揭示,立体匹配技术的研究将给计算机视觉、机器人、图形图像处理以及生物医学等研究领域注入新的活力。

参考文献(References)

- [1] Barnard S T, Fischler M A. Computational stereo [J]. ACM Computing Surveys, 1982, 14(4): 553-572.
- [2] Dhond U R, Aggarwal J K. Structure from stereo — A review [J]. IEEE Trans on Systems, Man and Cybernetics, 1989, 19(6): 1489-1510.
- [3] Faugeras O. What can be seen in three dimensions with an uncalibrated stereo rig[C]. Proc of the 2nd European

- Conf on Computer Vision. Santa Margherita Ligure, 1992: 563-578.
- [4] Koschan A. What is new in computational stereo since 1989: A survey of current stereo papers [R]. Berlin: Technical University of Berlin, 1993.
- [5] Scharstein D, Szeliski R. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms [J]. *Int J of Computer Vision*, 2002, 47(1): 7-42.
- [6] Brown M Z, Burschka D, Hager G D. Advances in computational stereo [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2003, 25(8): 993-1008.
- [7] Venkateswar V, Chellappa R. Hierarchical stereo and motion correspondence using feature groupings[J]. *Int J of Computer Vision*, 1995, 15(3): 245-269.
- [8] Mühlmann K, Maier D, Hesser J, et al. Calculating dense disparity maps from color stereo images, an efficient implementation[J]. *Int J of Computer Vision*, 2002, 47(1-3): 79-88.
- [9] Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2004, 26(2): 147-159.
- [10] Gong M, Yang Y H. Fast unambiguous stereo matching using reliability-based dynamic programming [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2005, 27(6): 998-1003.
- [11] 徐奕, 周军, 周源华. 基于小波及动态规划的相位匹配[J]. *上海交通大学学报*, 2003, 37(3): 388-392. (Xu Y, Zhou J, Zhou Y H. Phase matching based on wavelet transform and dynamic programming[J]. *J of Shanghai Jiaotong University*, 2003, 37(3): 388-392.)
- [12] Goulermas J Y, Liatsis P, Fernando T. A constrained nonlinear energy minimization framework for the regularization of the stereo correspondence problem [J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2005, 15(4): 550-565.
- [13] Veksler O. Fast variable window for stereo correspondence using integral images [C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. Madison, 2003: 556-561.
- [14] Wang L, Kang S B, Shum H Y. Cooperative segmentation and stereo using perspective space search [C]. *Proc of Asian Conf on Computer Vision*. Jeju Island, 2004: 366-371.
- [15] Yoon K J, Kweon I S. Locally adaptive support-weight approach for visual correspondence search[C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. San Diego, 2005: 924-931.
- [16] Zabih R, Woodfill J. Non-parametric local transforms for computing visual correspondence [C]. *Proc of the 3rd European Conf on Computer Vision*. Stockholm, 1994: 150-158.
- [17] Birchfield S, Tomasi C. Multiway cut for stereo and motion with slanted surfaces [C]. *Proc of Int Conf on Computer Vision*. Greece, 1999: 489-495.
- [18] Boykov Y, Veksler O, Zabih R. Markov random fields with efficient approximations [C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. Santa Barbara, 1998: 648-655.
- [19] Kim H, Sohn K. Hierarchical disparity estimation with energy-based regularization [C]. *Proc of Int Conf on Image Processing*. Barcelona, 2003: 373-376.
- [20] Prince J D, Eagle R A. Weighted directional energy model of human stereo correspondence [J]. *Vision Research*, 2000, 40(9): 1143-1155.
- [21] Muquit M A, Shibahara T, Aoki T. A high-accuracy passive 3D measurement system using phase-based image matching [J]. *IEICE Trans on Fundamentals of Electronics, Communications and Computer Sciences*, 2006, E89-A(3): 686-697.
- [22] Fleet D J, Jepson A D. Stability of phase information [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1993, 15(12): 1253-1268.
- [23] 徐彦君, 杜利民, 侯自强. 基于相位的尺度自适应立体匹配方法[J]. *电子学报*, 1999, 27(7): 38-41. (Xu Y J, Du L M, Hou Z Q. A scale-adaptive phase-based stereo matching method [J]. *Acta Electronica Sinica*, 1999, 27(7): 38-41.)
- [24] Birchfield S, Tomasi C. Depth discontinuities by pixel-to-pixel stereo [C]. *Proc of IEEE Int Conf on Computer Vision*. Bombay, 1998: 1073-1080.
- [25] Ohta Y, Kanade T. Stereo by intra- and inter-scanline search using dynamic programming [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1985, 7(2): 139-154.
- [26] Cox I J, Hingorani S L, Rao S B, et al. A maximum likelihood stereo algorithm [J]. *Computer Vision and Image Understanding*, 1996, 63(3): 542-567.
- [27] Bobick A F, Intille S S. Large occlusion stereo [J]. *Int J of Computer Vision*, 1999, 33(3): 181-200.
- [28] Lei C, Selzer J, Yang Y H. Region-tree based stereo using dynamic programming optimization [C]. *Proc of IEEE Conf on Computer Vision and Pattern Recognition*. New York, 2006: 378-385.
- [29] Cox I J, Roy S. A maximum-flow formulation of the N -camera stereo correspondence problem [C]. *Proc of the 6th Int Conf on Computer Vision*. Bombay, 1998: 492-499.
- [30] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts [J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2001, 23

- (11): 1222-1239.
- [31] 王年, 范益政, 鲍文霞, 等. 基于图割的图像匹配算法[J]. 电子学报, 2006, 34(2): 232-236.
(Wang N, Fan Y Z, Bao W X, et al. An images matching algorithm based on graph cuts [J]. Acta Electronica Sinica, 2006, 34(2): 232-236.)
- [32] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2004, 26(9): 1124-1137.
- [33] Ruichek Y. Multilevel and neural-network-based stereo-matching method for real-time obstacle detection using linear cameras [J]. IEEE Trans on Intelligent Transportation Systems, 2005, 6(1): 54-62.
- [34] Hua X J, Yokomichi M, Kono M. Stereo correspondence using color based on competitive-cooperative neural networks[C]. Proc of the 6th Int Conf on Parallel and Distributed Computing, Applications and Technologies. Denver, 2005: 856-860.
- [35] Wang B, Chung R, Shen C. Genetic algorithm-based stereo vision with no block-partitioning of input images [C]. Proc IEEE Int Symposium on Computational Intelligence in Robotics and Automation. Kobe, 2003: 830-836.
- [36] Gong M, Yang Y H. Genetic-based stereo algorithm and disparity map evaluation [J]. Int J of Computer Vision, 2002, 47(1-3): 63-77.
- [37] Zitnick L C, Kanade T. A cooperative algorithm for stereo matching and occlusion detection [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2000, 22(7): 675-684.
- [38] Ruichek Y, Issa H, Postaire J G, et al. Towards real-time obstacle detection using a hierarchical decomposition methodology for stereo matching with a genetic algorithm[C]. Proc of the 16th IEEE Int Conf on Tools with Artificial Intelligence. Boca Raton, 2004: 138-147.
- [39] Scharstein D, Szeliski R. Stereo matching with nonlinear diffusion [J]. Int J of Computer Vision, 1998, 28(2): 155-174.
- [40] Lee S H, Kanatsugu Y, Park J I. MAP-based stochastic diffusion for stereo matching and line fields estimation[J]. Int J of Computer Vision, 2002, 47(1-3): 195-218.
- [41] Sun J, Zheng N N, Shum H Y. Stereo matching using belief propagation[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2003, 25(7): 787-800.
- [42] Felzenszwalb P F, Huttenlocher D P. Efficient belief propagation for early vision[C]. Proc of IEEE Conf on Computer Vision and Pattern Recognition. Washington, 2004: 261-268.
- [43] Klaus A, Sormann M, Karner K. Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure[C]. Proc of the 18th Int Conf on Pattern Recognition. Hong Kong, 2006: 15-18.
- [44] Wang L, Liao M, Gong M, et al. High-quality real-time stereo using adaptive cost aggregation and dynamic programming[C]. 3rd Int Symposium on 3D Data Processing, Visualization and Transmission. North Carolina, 2006: 798-805.
- [45] Birchfield B, Natarajan B, Tomasi C. Correspondence as energy-based segmentation [J]. Image and Vision Computing, 2007, 25(8): 1329-1340.
- [46] Deng Y, Yang Q, Lin X, et al. Stereo correspondence with occlusion handling in a symmetric patch-based graph-cuts model[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2007, 29(6): 1068-1079.
- [47] Binaghi E, Gallo I, Marino G, et al. Neural adaptive stereo matching [J]. Pattern Recognition Letters, 2004, 25(15): 1743-1758.
- [48] Bleyer M, Gelautz M. Graph-cut-based stereo matching using image segmentation with symmetrical treatment of occlusions[J]. Signal Processing: Image Communication, 2007, 22(2): 127-143.
- [49] Gong M, Yang Y H. Real-time stereo matching using orthogonal reliability-based dynamic programming[J]. IEEE Trans on Image Processing, 2007, 16(3): 879-884.
- [50] Ambrosio G, Gonzalez J. Extracting and matching perceptual groups for hierarchical stereo vision [C]. Proc of the 15th Int Conf on Pattern Recognition. Barcelona, 2000: 542-545.
- [51] Desolneux A, Moisan L, Morel J M. Gestalt theory and computer vision[R]. Paris: CMLA, 2002.