

# 立体视觉研究的现状与进展<sup>\*</sup>

游素亚 徐光祐

(清华大学计算机系 北京 100084)

**摘要** 立体视觉是计算机视觉中的一个重要分支,一直是计算机视觉研究的重点和热点之一,在20多年的发展过程中,逐渐形成了自己的方法和理论。本文系统地评述了立体视觉研究的现状与进展,分析比较了各种技术的优缺点和适用范围,并着重论述了立体视觉的核心——立体匹配的研究进展,总结分析了目前立体视觉研究所存在的主要问题和今后的发展方向。

**关键词** 立体视觉,立体匹配,视觉机理。

## 1 引言

获取空间三维场景的距离信息是计算机视觉研究中最基础的内容。目前,获取距离信息的方法和技术很多,学术思想非常活跃,新技术、新方法不断涌现。立体视觉是计算机被动测距方法中最重要的距离感知技术,它直接模拟了人类视觉处理景物的方式,可以在多种条件下灵活地测量景物的立体信息,其作用是其它计算机视觉方法所不能取代的,对它的研究,无论是从视觉生理的角度还是在工程应用中都具有十分重要的意义。

立体视觉的开创性工作是60年代中期开始的。美国MIT的Robert完成的三维景物分析工作,把过去的二维图象分析推广到了三维景物,这标志着立体视觉技术的诞生,并在随后的20年中迅速发展成一门新的学科。特别是70年代末,Marr创立的视觉计算理论对立体视觉的发展产生了巨大影响,现已形成了从图象获取到最终的景物可视表面重建的完整体系,在整个计算机视觉中已占有越来越重要的地位。经过20多年的研究,立体视觉在机器人视觉、航空测绘、军事应用、医学诊断及工业检测中

的应用越来越广,研究方法从早期的以统计相关理论为基础的相关匹配,发展到具有很强生理学背景的特征匹配,从串行到并行,从直接依赖于输入信号的低层处理到依赖于特征、结构、关系和知识的高层次处理,性能不断提高,其理论正处在不断发展与完善之中。为了充分反映出这一研究领域的研究动态,本文采用由面到点的方法,以各种技术的核心思想为线索,对目前立体视觉研究的现状加以评述,指出各种方法的优缺点和值得注意的问题,最后指出了立体视觉研究存在的问题和发展方向。

## 2 立体视觉

立体视觉的基本原理是从两个(或多个)视点观察同一景物,以获取在不同视角下的感知图象,通过三角测量原理计算图象象素间的位置偏差(即视差)来获取景物的三维信息,这一过程与人类视觉的立体感知过程是类似的。一个完整的立体视觉系统通常可分为图象获取、摄像机定标、特征提取、立体匹配、深度确定及内插等6个大部分<sup>[1]</sup>。

立体图象获取的方式很多,主要取决于应用的场合和目的。例如,为了研究人类视觉的立体融合机

\* 国家自然科学基金资助。

收稿日期:1996.05.29

制,一般采用由计算机产生的随机点立体图象对<sup>[2]</sup>;在航空测量领域,通常是利用飞行器携带的高性能摄像机沿航向序列地摄取图象\*;再如,在显微立体分析中,通过旋转扫描电镜的样品台,同样可获得物体的立体图象对<sup>[4]</sup>。在获取立体图象时不但要满足应用要求,而且要考虑视点差异、光照条件、摄像机性能以及景物特点等因素的影响,以有利于立视计算。

摄像机标定是为了确定摄像机的位置、属性参数和建立成象模型,以便确定空间坐标系中物体点同它在图象平面上象点之间的对应关系。对于理想的平行视点模型,其核线与图象扫描线重合,这一特点可大大降低立体匹配的复杂性。对于一般存在旋转和位移的成象系统,都可通过投影变换等手段得到理想模型。摄像机标定方法可分为两大类,第一类是直接估计摄像机位置、光轴方向、焦距等参数;第二类是通过最小二乘法拟合,确定三维空间点映射为二维图象点的变换矩阵。建立一个有效的摄像机模型,除了能够精确地恢复出空间景物的三维信息外,还有利于解决立体匹配问题。

特征提取是为了得到匹配赖以进行的图象特征,由于目前尚没有一种普遍适用的理论可运用于图象特征的提取,从而导致了立体视觉研究中匹配特征的多样性。目前,常用的匹配特征主要有点状特征、线状特征和区域特征等。一般来讲,大尺度特征含有较丰富的图象信息,在图象中的数目较少,易于得到快速的匹配,但它们的定位精度差,特征提取与描述困难。而小尺度特征数目较多,其所含信息较少,因而在匹配时需要较强的约束准则和匹配策略,以克服歧义匹配和提高运算效率。良好的匹配特征应具有可区分性、不变性、稳定性、唯一性以及有效解决歧义匹配的能力。

立体匹配是立体视觉中最重要也是最困难的问题。当空间三维场景被投影为二维图象时,同一景物在不同视点下的图象会有很大不同,而且场景中的诸多因素,如光照条件,景物几何形状和物理特性、噪声干扰和畸变以及摄像机特性等,都被综合成单一的图象中的灰度值。因此,要准确地对包含了如此之多不利因素的图象进行无歧义的匹配,显然是十分困难的。对于任何一种立体匹配方法,其有效性有赖于3个问题的解决,即:选择正确的匹配特征,寻

找特征间的本质属性及建立能正确匹配所选特征的稳定算法。立体匹配的研究都是围绕着这三方面在工作,并已提出了大量各具特色的匹配方法。但是由于立视匹配涉及的问题太多,至今仍未得到很好的解决,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,降低实现的复杂程度和计算量,都需要进行更深入的探索和研究。有关立体匹配技术的详细论述将在下节给出。

已知立体成象模型和匹配视差后,三维距离的恢复是很容易的。影响距离测量精度的因素主要有摄像机标定误差、数字量化效应、特征检测与匹配定位精度等。一般来讲,距离的测量精度与匹配定位精度成正比,与摄像机基线长度成反比。增大基线长度可改善距离测量精度,但同时会增大图象间的差异,增加匹配的困难。因此,要设计一个精确的立视系统,必须综合考虑各方面的因素,保证各个环节都具有较高的精度。Behrooz<sup>[5]</sup>和 Blöstein<sup>[6]</sup>等对确定空间位置坐标点时的误差来源,以及图象量化效应对立体测量精度的影响进行了详细的分析,Gem<sup>[7]</sup>则对序列图象中的噪声影响提出了模型估计和唯一性结果的论证。

立体视觉的最终目的是为了恢复景物可视表面的完整信息,目前,无论是哪种匹配方法都不可能恢复出所有图象点的视差。对于一个完整的立视系统来讲,都必须进行最终的表面内插重建。在内插过程中,最重要的问题就是如何有效地保护景物面的不连续信息。Grimson 根据物体的表面物理性质和景物成象特性,提出了内插重建所必须满足的表面相容性原理,并建立了正则化重建算法<sup>[8,9]</sup> TerEopulos<sup>[10]</sup> 基于 Grimson 的重建原理,提出了由粗到细的多通道重建技术,以改善重建精度和加速收敛过程。Maitre<sup>[11]</sup>等根据空间景物的结构特点,提出了一种基于模型的内插重建算法,其原理是将图象分割成不同的结构区域,采用一次/二次曲面对分割区域进行最佳拟合,从而重建出不同区域的三维距离。此方法较好地保护了景物的不连续信息,适合于人工景物的处理<sup>[1,2]</sup>。

以上6个部分构成了立体视觉的完整体系,其中立体匹配是最重要也是最困难的部分。其存在的主要问题是,对存在特殊结构的景物,如平坦、缺乏纹理细节、周期性的重复特征等易产生假匹配;对方

\* 游素亚,双目立体视觉理论与方法的研究。见:华中理工大学博士论文集。1994年。

向平行于核线的特征匹\*配出现歧义;在录象机基线距离增大时,遮挡严重,能重建的空间点减少。为了解决这些问题,降低双目匹配的难度,自1986年以来出现了三目立体视觉系统,即采用3个摄像机同时摄取空间景物,通过利用第三目图象提供的信息来消除匹配的歧义性。Watanale<sup>[12]</sup>等采用相互垂直的三目视点,较好地解决了与核线平行特征的匹配歧义性。Ohta<sup>[13]</sup>提出的共线结构三目视觉系统;可有效地避免遮挡问题。Ayache<sup>[14]</sup>等提出了基于边缘线段的三目机器人视觉系统,Gurwitz<sup>[15]</sup>的文章详细地分析了三目立体视觉的性能。

### 3 立体匹配

立体匹配是寻求同一空间景物在不同视点下投影图象的象素间的一一对应关系。与普通的图象模板匹配不同,立体匹配是在两幅存在视点差异,几何、灰度畸变和噪声干扰的图象间进行的,不存在任何的标准模板,这无疑是十分困难的,更何况立体匹配还要求得到象素间的一一对应匹配。目前,立体匹配的研究基本上分为两个方向:①从理解人类视觉的立体融合机制出发,试图建立一种通用的人类双眼视觉计算模型;②从实际应用和要求出发,建立实用的立体视觉系统。

第一种方法的特点是从视觉生理的角度揭示人类视觉的立体融合机制,从而建立通用的立视系统。虽然从目前的研究来看,要构造出如同人类视觉一样灵活的机器视觉系统还有待于进一步的研究,但目前已取得的一些研究成果,如Marr的立视匹配三大约束,零交叉感知基元以及多通道协同处理方式等,对立体视觉的发展起了巨大的推动作用。第二种方法则是针对不同的应用和要求,以建立更直接的、专用的和面向对象的立视系统为目的,通过强调场景和任务的约束等手段来降低视觉处理问题的难度,从而增加系统的实用性。

#### 3.1 立视模型

70年代出现的Marr立视计算理论<sup>[2]</sup>被认为是目前对人类立视过程最完整的论述,作为其理论重要组成部分的立视匹配三大约束(唯一性、相容性和

连续性)、零交叉匹配基元以及多通道协同处理方式对立体视觉的发展起了巨大的推动作用。1976年Marr基于其理论建立了模拟人类立视机制的匹配算法<sup>[16]</sup>,并由Grimson在计算机上实现<sup>[17]</sup>。该算法的核心是首次采用多分辨率 $\nabla^2G$ 零交叉作为匹配基元,利用零交叉分布特性和由粗到细引导的多通道协同匹配技术来解决匹配的歧义性和搜索空间大的问题。1985年Grimson对上述算法进行了改进<sup>[18]</sup>,在匹配中使用视差沿零交叉轮廓线连续的约束来保证匹配的一致性。分析表明,对于某特殊结构的景物,单纯的零交叉匹配基元并不能得到正确的解释<sup>[19]</sup>,为此,Mayhew<sup>[20]</sup>等以视觉生理学的观点对Marr的零交叉匹配基元和匹配约束进行了补充修正,提出 $\nabla^2G$ 的卷积峰也应是人类立视的匹配输入基元,并基于人类视觉首先观察重要的目标轮廓信息这一事实,提出了形状连续性约束,在一定程度上解决了Marr理论所存在的问题。即使如此,Jenkin<sup>[21]</sup>等证明了 $\nabla^2G$ 零交叉加卷积峰地同样不能正确解释所有的景物形式,而对于人类来说,对这样的景物的立体融合是十分自然的,因此,必定存在其它更有利于立视匹配的基元形式。图象相位是一种值得重视的信息,与幅度信息相比,相位包含了更多的图象信息,幅度信息确定图象的灰度变化,而频率相位却决定了图象的结构位置。事实上,许多已有的匹配基元,如零交叉、卷积峰等都对应着特定的相位值。早在70年代,Robson<sup>[22]</sup>就从生理学角度证明了人类视觉系统中的简单视细胞是以正交相位关系形式成对出现的,并且可以用一对实部与虚部相互正交的复数滤波器来模拟。JulesZ等也指出相位信息在人类解释景物象点之间的过程中起着非常重要的作用<sup>[23]</sup>,相位信息可能是人类立体融合中最理想的基元形式。基于此,Sanger<sup>[24]</sup>和Jenkin<sup>[21]</sup>等提出了基于Gabor滤波相位的立视系统。游素亚<sup>[25]</sup>等于1993年提出了采用Hardy小波变换相位为基元的匹配算法。该方法模拟了人类视觉细胞的正交相位关系,利用Hardy小波基函数将原始图象对表示成一系列具有良好空频域定位性的正交带通分量,以带通分量的相位信息作为匹配基元,从而将视差估计转化为小波相位差估计问题。虽然,目前并没有更多的生理学证据来证明相位信息就是人类立视融合的输出基元,但这里到少有两点是合理的,其一,

• 1) 巢京. 利用立体视觉方法重建数字地面模型的研究, 见: 全国第四届三遥技术年会论文集. 1992年  
2) 程伟. 三维表面重建系统的设计与实现. 见: 华中理工大学毕业论文集. 1993年

小波理论优良的多分辨率局部分析特性,是与人类视觉的多通道特性相类似的<sup>[26]</sup>,其灵活多变的小波基函数选择方式,使我们有可能寻找到更加符合视觉机制的图象信息表达形式,从而建立有效的立视模型;其二,相位信息包含了目前所共识的零交叉和卷积峰匹配基元,它对景物具有更强、更全面的解释能力。因此,无论从视觉机制角度还是在实际应用中,对它的进一步研究都是十分有意义的。

Griswold<sup>[27]</sup>等提出过一种新的基于视觉立体融合机制的双目视觉计算模型,其基本思想是将象平面分为两个区域;中心区域称为凹区,其余部分称为外视网膜区。对凹区的数据作高分辨率的分析处理,以精确地获得目标表面的信息,而对外视网膜区的数据只作粗略的低分辨率处理,获得一些全局性的结构信息,最终将二者融合,求得匹配视差。在算法的实验中,系用了神经网络机制。同时考虑视觉皮层细胞的各向异性特征。

Yeshurun<sup>[28]</sup>等人在研究了灵长目动物大脑皮层对双目图象刺激响应的结构关系的基础上,提出了一种倒滤波的局部非线性匹配模型,它提供了灵长目动物大脑皮层中有对左右眼刺激响应呈柱状交织现象的一种解释。整个算法是一种并行的迭代形式,其特性与灵长目动物立体视学的生理限制和解剖特性相一致。这种模型为进一步研究人类视觉的立体感知机理提供了一条新思路。Olson<sup>[29]</sup>等已将此算法用于双目机器人视觉,取得了较好的效果。

在过去的几十年里,基于多频率通道的图象分析方法受到了极大的重视,这一方法有着深刻的生理学、心理学背景。它是以人类视觉系统早期信息处理的多通道滤波理论为基础的,Marr的多通道协同处理方式即与这一思想相一致。但这一方法未能全面地反映出视觉的多通道感知特性。生理学实验证明,人类视觉的感知通道不仅是多频率通道,而且是多方向通道,人类具有很强的方向鉴别能力。即在视觉皮层中存在一定数量的方向敏感性视神经元,它们只对各自方向区间内的图象信息有响应,一旦偏离相应的方向区间,就会大大减弱甚至不能引起该神经元的响应。基于这种考虑,文献<sup>[30]</sup>将这种与视觉多频率通道特性同样重要的多方向通道特性引入立体视觉领域,并建立了实现算法,在一定程度上补充丰富了Marr的立视理论,而且多频率通道特性与多方向通道特性的结合,将有助于建立更加完善的立视系统。

### 3.2 匹配算法

在目前的立体视觉研究中,绝大多数都是针对各自的应用目的和要求,以建立实用的立视系统为目的。立体匹配实质上是在匹配准则下的最佳搜索问题,许多数学中的最优化技术都可应用于立体匹配。灰度相关是一种常用的匹配方法,自40年代相关技术被应用于航空立体测量以来,相关匹配方法一直在立体视觉的研究和应用中起着非常重要的作用,其基本思想是以统计的观点将图象看成是二维信号,进而采用统计相关的方法寻找信号间的相关匹配,正因为如此,相关匹配又被称为信号匹配。它是图象统计相关理论直接在立体匹配中的拓展,并没有很多视觉生理上的解释。由于相关法是直接对图象像素进行匹配,匹配结果不受特征检测精度和密度的影响,因而可得到很高的定位精度和密集的视差表面。但这种对图象灰度统计特性的过分依赖,导致了匹配对景物表面结构、光照反射和成象几何十分敏感,而且区域灰度相关的假设意味着相关窗口内的像素都具有相同的视差,这对包含有深度不连续的区域来讲,是不成立的。因此,在空间景物表面缺乏足够的纹理细节以及成象失真较大(如视点间距离过大)的场合,相关匹配法难以应用。

与灰度相关法不同,基于特征的匹配方法则是选择地匹配能表示景物自身特性的特征,通过更多地强调空间景物的结构信息来解决歧义性问题。Baker<sup>[31]</sup>和Arnold<sup>[32]</sup>等提出以边缘点为匹配特征的匹配算法,并在匹配时首次采用了动态规划技术,将左右图象同一扫描线上的匹配特征,分别等效为动态规划的阶段与状态,通过在二维规划平面上搜索最佳路径得到最优的匹配。Ohta<sup>[33]</sup>等提出了以边缘线为匹配特征的动态规划算法,这相当于把动态规划的搜索空间扩大到三维,首先利用二维规划得到边缘点的最佳匹配代价,然后利用此代价在三难规划空间中搜索边缘线的匹配,这样不但保证了匹配在同一扫描线内是最佳的,而且在扫描线间也是全局最优的。实验表明,此算法对多种类型的景物都能得到较好的匹配,缺点是计算量大,实现复杂。在Llogd<sup>[34]</sup>提出的算法中,综合利用了动态规划和概率型的松弛标记技术,算法以边缘点为匹配基元,首先采用动态规划技术搜索相关扫描线上的最佳匹配路径集合,然后以扫描线间的边缘视差连续为约束,利用松弛算法从已得到的匹配路径集合中寻找正确匹配。这种方法也可以保证匹配在扫描线间的

全局一致性,且计算量比Ohta的三维规划搜索小。

松弛技术也是解决立视匹配的有效手段。首先利用简单的相似准则求得初始匹配,然后根据匹配特征间利用的几何和视差约束,利用松弛迭代技术增强初始匹配结果,使正确匹配的概率上升,从而达到解决歧义匹配的目的。在Barnard<sup>[35]</sup>提出的匹配算法中,采用了Moravec算子的“兴趣点”<sup>[36]</sup>作为匹配基元,以视差连续性约束为松弛迭代准则,通过迭代处理得到视差连续意义下的最佳匹配。虽然此算法只能得到很离散的匹配点,但它对核平面几何不作任何要求,允许有垂直视差,且可应用于运动图象分析。在Kim的零交叉匹配<sup>[37]</sup>,Medioni的直线段匹配<sup>[38]</sup>,Nasrabadi的曲线段匹配<sup>[39]</sup>以及Lee的分割区域匹配算法中<sup>[40]</sup>都采用了松弛迭代技术。松弛方法已在立体匹配中得到了广泛的应用,原则上讲,无论选择什么形式的匹配基元,都可构造出相应的松弛迭代算法。

立体匹配是典型的视觉“病态”计算问题,Poggio<sup>[41]</sup>认为初级视觉处理中的许多“病态”计算问题都可以统一在正则化框架下,通过加入一定的约束来使“病态”解转化为“良态”。在立体匹配领域,一般是利用匹配连续性约束将匹配视差限制在平滑解空间里。在这里,存在两个得解决的问题,其一是如何有效地保护视差表面的不连续信息,Terzopoulos<sup>[42]</sup>提出了含有不连续信息的通用正则化模型,通过引入一个可控制平滑度的权函数来保护解的不连续信息。在Terzopoulos, March及Yokoye<sup>[43,44]</sup>的正则化匹配算法中都采用了这一方式。其二是如何避免出现局部最小的问题,在Barnard<sup>[45]</sup>提出的方法中,采用由粗到细的多尺度引导策略和模拟退火技术来解决这一问题。正则化匹配方法的最大优点是可以全局最优地生成密集的视差场。

立体视觉的最终目的是为了恢复景物可视表面的完整信息,而特征匹配方法只能得到离散特征点的视差,需要进行较多的内插处理。对于一个完整的立视系统来讲,不能断然地将匹配与内插重建过程分为两个不相关的独立模块,它们之间应该存在着很多的信息反馈,匹配结果约束内插重建,重建结果引导正确匹配。基于这一思想,Hoff<sup>[46]</sup>等提出了匹配、重建一体化的方法,在得到匹配的同时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。在他们的实现算法中,内插采用一次/二次曲面拟合技术,同时具有检测视差表面不连续的能力,整个算法是一种由粗

到细引导的多分辨率结构。实验结果表明,此方法具有处理复杂场景的能力。类似的还有Olsen<sup>[47]</sup>,Terance<sup>[48]</sup>和McLauchlan<sup>[49]</sup>方法等。这种将匹配与重建视为一体的重建方法是值得重视的,它不但能生成完整的视差场,而且可以为匹配提供更多的信息,如不连续性、遮挡等,提高了匹配的可信程度。

近年来,随着科学的发展,许多新理论已被引入立体匹配领域,典型的如神经网络分析法和小波分析理论。Poggio提出的正则化模型是一个三层前馈网络,其价值函数的最小化可用Hopfield网络来求解。同时为了避免出现局部最小化问题,可采用模拟退火(Simulated Annealing)技术。而且目前已有的许多算法都可转化为神经网络实现,如当判决价值函数为二次型时,其Euler方程是一阶微分方程组,可用阻容网络求解。动态规划算法可映射多级神经网络等。但从目前的研究来看,这些算法并未反映出人类视觉神经感知网络的本质,真正构造模仿生物视觉系统的立视匹配方法还有待于进一步的研究。小波理论是近年来兴起的一种崭新的时(空)频域分析理论,被认为是傅立叶分析发展的新阶段,它突破了傅立叶分析中没有任何分辨力的缺陷,可以对指定频带和时(空)段内的信号成分进行分析。由于良好的时(空)频域局部化特性,小波分析可以准确地定位信号的瞬变特征,而通过多分辨率变换,又可对信号进行“切片式”分析。这一思想与视觉中由粗到细的多分辨率分析过程是一致的,而且小波变换的多分辨率分解特性,更加符合人类的视觉特性,更加适合于视觉信息的处理<sup>[26]</sup>。小波理论已被成功地应用到立体匹配领域<sup>[50,51]</sup>,并证明它有助于解决立体匹配问题。例如小波分析的局部化特性有利于充分挖掘可利用的图象信息,增强立体匹配解决歧义性的能力,在文献<sup>[50]</sup>中所提出的方法中正是利用了这一特性,将点匹配转化为可信度更高的串匹配。小波理论的另一优点是它允许灵活地选择基函数,以满足不同的应用,这使得我们有可能寻找到更加符合视觉机制的图象信息表达形式,从而建立更有效的立视模型。

#### 4 立体视觉研究所存在的问题和发展方向

虽然立体视觉经过20多年的研究,已经有了很大的发展。但无论是从视觉生理的角度,还是从实际应用方面来看,现有的立体视觉技术还处在十分不

成熟的阶段。这不仅仅涉及到技术上的原因,而且更多地在于人类对自身视觉机理还不十分了解。人类视觉系统具有惊人的分析理解能力,但人类是如何精选、获取和分析理解视觉知识的,至今还未充分搞清楚。立体匹配作为立体视觉的核心,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,以解决存在严重灰度失真、几何畸变(透视、旋转、缩放等)、噪声干扰、特殊结构(平坦区域、重复相似结构等)及遮挡景物的匹配问题,如何建立更有效的图象表达形式和立视模型,以便更充分地反应景物的本质属性,为匹配提供更多的约束信息,降低立体匹配的难度。因此,立体视觉问题的彻底解决还有待于对人类自身视觉机理的深入研究。

虽然现有的立体视觉系统还存在着许多问题,要构造出类比于人类视觉的通用立视系统也不是近期内可以达到的,但经过一、二十年的研究,立体视觉已经有了很大的发展,从早期的以统计相关理论为基础的相关匹配,发展到具有很强生理学背景的特征匹配,从串行到并行,从直接依赖于输入信号的低层处理到依赖于特征、结构、关系和知识的高层次处理,性能不断提高,其理论正处在不断发展与完善之中。从当前的发展来看,其发展方向可归纳如下:

(1)以全面的观点将立体视觉系统的各个模块联系起来,充分挖掘内在信息。

(2)综合利用各种技术,多传感器信息融合,包括综合使用不同的匹配方法和主动式、被动式单目视觉技术等,提高系统的性能。

(3)由传统的双目视觉向多目视觉、完全静态的视觉向动态视觉发展,通过增加信息输入降低视觉计算的难度。

(4)向智能化发展,建立基于知识的、模型的和规则的立体视觉方法。

(5)算法向并行化发展,采用并行流水线机制和专用的信号处理器件,增强系统的实用性。

(6)强调场景与任务约束,针对不同的应用目的,优化选择各部分,建立有目的的和面向任务的立视视觉系统。

立体视觉作为一门多学科的交叉科学,正吸引着大批包括视觉生理、心理、物理、数学以及计算视觉等多种学科的研究人员运用不同的技术手段对之进行的研究,它不但具有重要的实用价值,而且对促进人类视觉机理的研究,彻底揭开人类视神经

系统的奥秘具有非常重要的意义。可以相信,随着众多研究工作者的努力,立体视觉的研究和应用必将取得丰硕的成果。

## 参考文献

- 1 Barnard S, et al, Computational stereo. ACM Computing Surveys, 1982,14;553~572.
- 2 Marr D. Vision. W. H. Freeman and Compay, 1982.
- 3 国际摄影测量与遥感学会第十六届大会综述,测绘学报, 1988, (6)
- 4 Boyde A. Photogrammetry of stereo pair SEM image using separate measurement from the two images. SEM/I, Illinois Institute of Technology Research, 1974;101~110.
- 5 Behrooz K P. Minimizataion of the quantization error in camera calibration. proc. of image Understand workshop, 1987; 671~680.
- 6 Blostein S D. et a1. Error analysis in stereo determination of 3-Dpoint positions. IEEE Trans. on PAMI, 1987, 9;752~765.
- 7 Gem-Sun J Young. 3D motion estimation using a sequence of noisy image: Models Estimation and Uniquence Resul. IEEE Trans. PAMI, 1990, 12(8).
- 8 Grimson W E L. Surface Consistency constrsints in Vision. CVGIP, 1983, 24;28~51.
- 9 Grimson W E L. A computation theory of visual surface interpolation. Phil. Trans. Roy. Soc. London, 1982, B298;395~427.
- 10 Terzopoulos D. The Computer of Visible-Surface Representations. IEEE Trans. on PAMI, 1988, 10(40);417~438.
- 11 Maitre H, et a1. Using Models to Improve Stereo Reconstruction. IEEE Trans. on PAMI, 1992, 149(2);269~277.
- 12 Watanale M, et a1. Improved Depth Map by Right-Angled Trinocular Stereo. proc. of Int. Conf. Pattern Recognition, 1986, (10);519~521.
- 13 Ohta Y, et a1. Stereo by two-level dynamic programming. Proc. of 9th IJCAI, 1985;1120~1124.
- 14 Ayache N, et a1. Trinocular stereo Vision for robotics. IEEE Trans. on PAMI, 1991, 13(1);73~85.
- 15 Gurwitz E, et a1. More on the Benfit of a Third Eye Machine Stereo Perception. Proc. of Int. Conf. Pattern Recognition, 1986;966~968.
- 16 Marr D, et a1. Cooperative computation of stereo disparity. Sci., 1976, 194;283~287.
- 17 Grimson W E L. A Computer implementation of a theory of human stereo vision. Phil. Trans. Roy. Soc. London, 1981, B292; 217~253.
- 18 Grimson W E L. Computational experiments with a feature based on stereo algo-rithm. IEEE Trans. on PAMI, 1985, 7(1);17~33.
- 19 Ohzawa I, et a1. The Binocular Organization of Simple Cells in the Cat's Visual Cortex. J. Neurophysiology, 1986, 56(1);221~242.
- 20 Mayfew J E W, et a1. Psychophysical and computational study

- towards a theory of human stereopsis. *AI*, 1981, 17: 349~385.
- 21 Jenkin M, et al. The measurement of binocular disparity in computational process in human vision. Ablex, New Jersey, 1988.
  - 22 Robson J G. Receptive fields; neural representation of the spatial and intensive attributes of the visual images. *Handbook of Perception*, Academic press, New York, 1985, 5: 81~116.
  - 23 Julesz T B. *Stereoscopic Vision*. *Vision Res.*, 1986, 26: 1601~1612.
  - 24 Sanger T D. Stereo disparity computation using Gabor filter. *Biol. Cybern.*, 1988, 59: 405~480.
  - 25 游素亚, 柳健, 万发贯, 基于小波变换相位的立体匹配系统, 信息与控制, 1994, (5).
  - 26 Mallat S. Multifrequency Channel Decompositions of Images and Wavelet Models. *IEEE Trans. on ASSP*, 1989, 37 (12): 2091~2110.
  - 27 GriswoId N C, et al A new Stereo Vision model based upon the binocular fusion concept. *CVGIP*. 1988, 41: 153~171.
  - 28 Yeshurun, et al. Cepstral filtering on columnar image architecture; a fast algorithm for binocular stereo segmentation. *IEEE Trans. on PAMI*, 1989, 11: 759~767.
  - 29 Olson T J, et al. Real-time Vergence Control for Binocular Robots. *IJCV*, 1991, 7(1): 67~89.
  - 30 游素亚, 柳健, 万发贯. 一种方向通道的立体匹配算法. 模式识别与人工智能, 1994.
  - 31 Baker H H, et al. Depth from edge and intensity based stereo. *Proc. 7th IJCAI*, 1981: 631~636.
  - 32 Aronld R D. Automated Stereo Perception. *Stanford University Tech. Rep.*, STAN-CS-83-961, 1983.
  - 33 Ohta Y, et al. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Trans. on PAMI*, 1985, 7: 139~154.
  - 34 Lloyd S A, et al. A Parallel Binocular Stereo Algorithm Utilized Dynamic Programming and Relaxation Labelling. *CVGIP*, 1987, 39: 202~225.
  - 35 Barnard S, et al. Disparity analysis of images. *IEEE Trans. on PAMI*, 1980, 2: 333~340.
  - 36 Moravec H P. Visual mapping by a robot rover. *proc. of 6th IJ-CAI*, 1979: 589~600.
  - 37 Kim Y C, et al. Finding range from stereo image. *Proc. of IEEE Conf. on CVPR*, 1985: 289~294.
  - 38 Medioni G, et al. Segment-Based Stereo Matching. *CVGIP*, 1985, 31(1): 2~18
  - 39 Nasrabadi N M. A Stereo Vision Technique Using Curve-Segments and Relaxation Matching. *IEEE Trans. on PAMI*, 1992, 14(2): 566~572.
  - 40 Lee H J, et al. Region Matching and Depth Finding for 3-D Objects in Stereo Aerial Photographs. *Pattern Recognition*, 1990, 23(2): 81~93.
  - 41 Poggio T, et al. Computational vision and regularization theory. *Nature*, 1985, 317: 314~319.
  - 42 Terzopoulos D. Regularization of Inverse visual Problems Involving Discontinuities. *IEEE Trans. on PAMI*, 1986, 8(4): 413~424.
  - 43 March R. A regularization model for stereo vision with controlled continuity. *Pattern Recognition*, 1989, 10: 259~263.
  - 44 Yokoye N. Surface reconstruction by regularizing a stereo correspondence problem. TR-90-15, *Electrotechnical Laboratory*, Tsukuba, Japan, 1990.
  - 45 Barnard S T. Stochastic stereo matching over scale. *IJCV-3*, 1989: 17~32
  - 46 Hoff W, et al. Surface from stereo: Integrating feature matching, disparity estimation, and contour detection. *IEEE Trans. on PAMI*, 1989, 11(2): 121~136.
  - 47 Olsen S I. Stereo Correspondence by Surface Reconstruction. *IEEE Trans. on PAMI* 1990: 309~315.
  - 48 Terance E B. Analysis of Two New Stereo Algorithms. *Proc. of 2th ICCV*, 1988: 118~122.
  - 49 McLauchlan P F, et al. Stereoscopic recovery and description of smooth textured surfaces. *Image and Vision Computing*, 1991, 9(1): 20~26.
  - 50 游素亚, 柳健, 万发贯, 一种基于小波变换的多分辨率立体匹配算法, 华中理工大学学报 1994, (5).
  - 51 Sheng Zhong, et al, Improved Stereo Correspondence Algorithm Based on Wavelet Transform. *ACCV'93*, Japan, 1993: 268~271

(上接 24 页)

#### 四、组织机构

大会程序委员会主席: 徐光佑 教授 清华大学计算机系 100084  
 电话: (010)62784141 (010)62782406 E-mail: xyg-dcs@mail. tsinghua. edu. cn  
 大会组织委员会主席: 谈正 教授 西安交通大学电信学院 710049  
 电话: (029)3268714 (029)3269680(夜间) E-mail: tanzh@xjtu. edu. cn  
 宋云娴 教授 西安空军工程学院 710038 电话(029)4397655

# State of the art and Future the development of stereo vision

You Suyu, Xu Guangyou

(Dept. of Computer Science & Engineering, Tsinghua University, Beijing, 100084)

**Abstract** Stereo vision, an important branch of computer vision, is one of the widest research area in Computer Vision. In previous decades, a lot of characteristic approaches and theories were progressively proposed in this area. In this paper, the state of the art of stereo vision is expounded systematically, and the advantages and disadvantages of some important approaches are discussed, with an emphasis on stereo matching. The key problem in stereo vision is expounded comprehensively. Finally, the existing problems and future development in stereo vision are discussed.

**Keywords** Stereo Vision, Stereo Matching, Visual Mechanism

## 第六届全国多媒体技术学术会议 NCMT'97 征文通知

由中国图象图形学会(多媒体专业委员会)、中国计算机学会(多媒体专业委员会)和香港多媒体与图象计算学会联合主办的第六届全国多媒体技术学术会议定于 1997 年 10 月下旬在西安召开。具体由西安交通大学、空军工程学院承办。会议将为从事多媒体技术的研究开发和工程技术人员以及制造厂商提供交流科学研究、科技开发经验与成果及推广应用的场所。

这次会议就多媒体技术基础理论、多媒体应用支撑平台、分布式多媒体技术及多媒体应用技术展开广泛的交流和讨论。将重点讨论多媒体通讯、多媒体数据库及多媒体应用等热点问题。本次年会还将组织多媒体作品评议、产品展示及上述热点问题的专题讨论会。大会程序委员会将从论文中评选出一批高质量论文,向国内核心学术刊物推荐出版,并设立优秀文奖。

### 一、征文范围

1. **多媒体技术基础理论**:多媒体信息处理、多媒体信息压缩、多媒体信息模型、多媒体信息获取/显示和接口技术、实时信号处理技术、可视化技术、动画技术、虚拟现实
2. **多媒体应用支撑平台**:多媒体著作工具、多媒体系统软件平台、多媒体数据库、多媒体应用标准、超文本与超媒体、多媒体应用系统体系结构、多媒体 ASIC 设计与应用
3. **分布式多媒体技术与多媒体通信**:多媒体电子邮件、多媒体网络技术、分布式多媒体支撑平台、协同工作和群件技术
4. **多媒体应用技术**:多媒体信息系统、多媒体在计算机辅助教育中的应用、多媒体家庭教育和娱乐系统、商用多媒体系统、多媒体医疗信息系统、Internet 上的多媒体应用

### 二、征文要求

1. 反映在多媒体及有关技术领域中的技术和应用研究成果。
2. 未在其他会议和刊物上发表过(来稿请注明论文所属征文大类)。
3. 在本次征文中,程序委员会对高质量论文采用直接向重要期刊(中国图象图形学报、软件学报等)推荐、刊载的办法。
4. 文稿需按 word 或 wps 格式,于 A4 纸排版。
5. 每篇来稿篇幅不超过 7000 字,并含有中英文题目、摘要和关键词。

### 三、征文日期

征文截止日期:1997 年 7 月 20(收到日期) 录取通知日期:1997 年 8 月 10 提交正式论文截止日期:1997 年 9 月 10(收到日期) 来稿请寄:西安交通大学电信学院(邮编:710049) 谈正教授 电话:(029)3268714 传真:(029)3237910(下转 23 页)