

视觉不变量与机器视觉研究

宋克欧 黄凤岗

(哈尔滨工程大学计算机与信息科学系, 哈尔滨 150001)

易英辉

(哈尔滨工程大学水声工程系, 哈尔滨 150001)

摘要 综述了视觉不变量用于机器视觉研究的新趋向——不变量视觉, 概括了视觉不变量理论的主要内容和应用, 以及所作的实际研究。

关键词 不变量, 机器视觉, 方面图, 三维运动目标识别与跟踪, 极指数栅格处理

1 引言

Marr 的视觉计算理论目前仍然是机器视觉的主要理论框架, 它在理论研究和实际应用中存在很多困难和争论^[1]。大多数机器视觉研究都着眼于通过计算, 从二维视觉数据中提取深度信息进行目标三维描述和识别。然而, 很多人认为视觉感知识别功能的本质不在于此, 而是直接感知目标形态所含的某些不变的结构特征。从 80 年代末, 人们把在数学和物理学中成功运用的不变量理论引进了机器视觉研究, 初步形成了视觉不变量理论和应用的新框架 (Towards a new framework)^[2], 文献[2~6]包含了视觉不变量理论的基本内容, 并且反映出这种新理论体系的自然形成过程。

视觉不变量理论认为: **几何结构特征** (不连续性—边缘和纹理细节) 是机器视觉的基础, **不变量** 是几何结构的本质描述, 因此, 视觉不变量应是机器视觉的研究核心。从数学意义上讲, 视觉不变量是目标几何结构在某些变换群条件下保持不变的函数形式。**视觉目标最重要的几何结构是由一些不变量来确定的, 反过来, 每一个不变量都在几何结构分析中起关键作用 (Klein-Weyl 命题)**。在机器视觉领域中, 主要关心的是如何从已经映射成的二维图象逆向生成三维空间中的目标描述, 这种描述应便于计算机理解, 但不一定满足人类感观上的死硬要求。不变量理

论主要涉及同维空间之间的映射, 目前大都以平面到平面的映射为基础, 对于平面目标、由平面构成的目标、以及部份表面由富含特征的平面构成的目标, 视觉不变量理论的应用是非常成功的。由目标的二维图象建立其三维描述是视觉不变理论和应用最困难也是最有魅力的课题, 近年来的研究表明, 任意形状三维目标的仿射不变量和射影不变量可以从不同视角的多个二维图象中求得, 人们正满怀兴趣地期待未来的发展, 以便解决三维曲面目标难以捉摸的描述难题^[2]。

2 视觉不变量理论和应用

视觉不变量理论的核心内容是与视觉有关的重要变换群, 欧几里德变换群、仿射变换群、射影变换群以及拓扑变换群是重要的逐级包含的变换群, 它们具有封闭性、恒等性、可逆性和结合性。线性变换群可用齐次矩阵表示, 矩阵待定参数的数目称为该群的自由度, 自由度反映群的复杂性和通用性, 自由度大则通用性强, 同时也趋于复杂。对于曲面目标, 描述变换群微分特性的李群以及与之相关的轨道也非常重要, 它们在描述群的复杂度以及构成独立不变量上都非常有用。

视觉不变量理论最重要也是最诱人的优点和应用在于它从功能上更接近人的视觉本质。生理视觉可从不同距离不同角度轻而易举地感知和识别形状

剧烈变化(包括遮挡)的目标,其关键是人的视觉感知知识以不变特征为基础。以视觉不变量为基础的机器视觉比较容易从功能上与人的视觉贴近,它易于克服由于观察角度变化给机器视觉造成的困难,它既不介意目标平移、比例缩放和旋转造成的形状相似变化,也不介意目标滚动造成的投影形态变化;它无需任何有关摄像机内外参数的先验知识和校准;它能较容易地获取目标模型,易于构成大型模型库并可以高效检索;它可以从两个观测角构成三维目标形状的几何描绘子,此描绘子不要求观测点之间的运动知识就可以用来识别目标,以及产生其它观测角的目标投影;它可以从两个观测角探测未知目标的三维几何结构和刚性,以及用于运动分割,例如,从不同角度计算 5 个目标点集的平面投影不变量,如果数值不同,则 5 个点构成一种三维结构,否则为一个二维平面结构;它可以冗余地表示目标全局结构特征,当发生部份遮挡时,往往可以从剩余部份测出表示全局的结构不变信息。以上这些优点和应用正是当今视觉理论苦于应付的恼人之处。

3 视觉不变量理论的应用研究

当人从远到近观察多个任意形状的三维运动目标时,要依靠某些线索,通过头、眼和身体的运动注视(跟踪)和识别各个目标。以不变量为基础,模拟人的视觉行为方式去设计视觉系统,可以全面带动机器视觉研究。

一个依靠仿射不变量的平面目标识别实验系统已设计成功,视觉不变量的优点已获初步证实。随着视角的改变:直视 0° , 仰视 50° 和斜视 60° 3 个方向,不变量的变化(交比不变量)不超过 2.1%。图 1 示出一种目标形状随角度变化的情况。该系统要适应两种环境:中远距离情况和近距离情况。中远距离时,目标尺寸远小于到摄像机的距离,可近似为平面目标,投影图象通过仿射变换形成,是一种没有内部纹理的黑影,信息完全包含于轮廓之中。这类目标最适合用方面图表示法进行三维描述和识别。在目标迫近视觉系统时,环境变得艰难而复杂,投影图象通过射影变换形成,信息包含于完整的几何结构之中,此时应对三维视觉不变量开展全面研究,以显示其优越性。

方面图是一种图结构,图的节点代表从观察空间某个角度范围获取的二维图象,节点参数可以是

目标的不变量特征、姿态以及运动参数,节点间的弧代表目标与观察者之间相对运动造成的投影图象转换,这是一种以观察者为中心的目标表示法。假设目标被一个球面包围,观察者沿球面从各种角度获取投影图象并建立图象之间的转换关系,通过节点参数的变化就可以获得完整的目标三维描述。图 2 示出某种飞机的 5 个序列二值图,单个二值图可能具有多义性,无法判断飞机的种类,然而,多个序列判断以后,属于某种飞机的证据积分远大于其它机种,从而完成识别,同时可以对飞机运动状态进行描述。

目前还存在一种以模拟人类视神经细胞分布的极指数栅格视觉研究体系,它极善于模仿人类对目标的视觉搜索跟踪行为,并可以把复杂的二维图象匹配转换成一维波形匹配,同时对目标的方位、姿态进行估计。与视觉不变量理论结合起来将会使视觉系统更具人类生理视觉的特性。前述视觉实验系统的目标搜索跟踪和匹配功能就是依据极指数栅格视觉完成的。

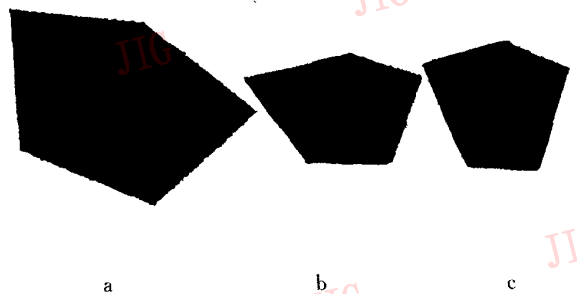


图 1 a: 模板图象 b: 45° 仰视图象 c: 45° 斜视图象
Fig. 1 a: The template image. b: 45° Bootom views image. c: 45° sidelong image.

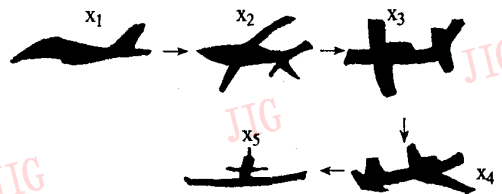


图 2 利用方面图由侧面图转换序列进行三维目标描述和识别。

Fig. 2 3D object description and recognition by using aspect graph and its transformation sequence.

x_1 可属于 A, B, C 三种飞机, x_2 可属于 A, B 两种飞机, x_3 属于 B 种飞机, x_4 属于 B, C 两种飞机, x_5 属于 A, B 两种飞机, 累计证据分: A 类 2.35, B 类 3.69, C 类 1.35。结果: 判决目标 X 属于 B 种飞机

参考文献

- 1 吴立德, 李介谷, 边肇祺等. 模式识别与人工智能. 论文与报告 1992, 5(4): 261~292.
- 2 Weiss I. Projective invariants of shapes. In Proceedings of Image Understanding Workshop, Cambridge, Ma, 1988, 2: 1125 ~

1134.

- 3 Kanatani K. Group-Theoretical Methods in Image Understanding. Berlin: Springer Verlag, 1990.
- 4 Forsyth D, Mundy J L, Zisserman A, et al. Invariant descriptors for 3D object recognition and pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(10): 971~991.
- 5 Mundy J L, Zisserman A. Geometric Invariance in Computer Vision. London, MIT Press, Cambridge, Ma, 1992.
- 6 Reiss T H. Object recognition using algebraic and differential invariants. Signal Processing, 1993.



宋克欧 哈尔滨工程大学计算机与信息科学系教授。主要研究领域包括图象处理、模式识别、计算机视觉水声信号特征提取和识别等。

Visual Invariants and Machine Vision Research

Song Keou, Huang Fenggang, Yi Yinghui

(Department of Computer and Information Sciences, Haerbin Engineering University, Haerbin 150001)

Abstract A new tendency to develop machine vision, invariant vision, is summarized in this paper. The paper has given major contents and applications of vision invariants in broad outline, as well our practical research engaged and engaging in they.

Keywords Invariants, Machine vision, Aspect graph, 3D moving object recognition and tracking, Polar exponential grid processing