

基于方面图技术的三维运动目标识别

张海燕 宋克欧

(哈尔滨工程大学计算机科学与技术学院, 哈尔滨 150001)

王东木

(北京仿真中心, 北京 100854)

摘要 三维目标在不同的视点下呈现不同的姿态, 所得的二维视图也不尽相同, 因此三维目标识别是一个较为复杂的问题. 为此提出了通过图象序列和图象序列之间的转移关系, 根据胜者为王的原则来识别三维目标的方法. 该方法采用极指数栅格技术和傅立叶变换相结合得到目标的轮廓不变量; 用神经网络结合方面图技术, 通过识别运动目标图象序列来识别三维运动目标, 实现了一个目标识别系统. 实验结果证明, 此方法可以有效地用于三维运动目标的识别.

关键词 模式识别(520·2040) 极指数栅格技术 极对数坐标变换 轮廓不变量 方面图

中图法分类号: TP391.41 **文献标识码**: A **文章编号**: 1006-8961(2003)08-0866-06

3D Moving Object Recognition Based on Aspect Graphs

ZHAGN Hai-yan, SONG Ke-ou

(College of Computer Science & Technology, Harbin Engineering University, Harbin 150001)

WANG Dong-mu

(Beijing Simulation Center, Beijing 100854)

Abstract The poses of 3D object in different viewpoints are different, so the 2D digital images in different viewpoints are different. Therefore, the 3D object recognition is difficult. In this paper, the approach to recognize 3D moving object by image sequence and the transition relationship of the images in terms of the principle of "the winner is the king" is produced. Polar-exponential grid technique maps the image in Cartesian coordinate into polar-log coordinate. The profile of object in polar-log coordinate has translation, rotation and scale invariant by Fourier transform and it can be used as invariant to recognize object. In this paper, the profile invariant is obtained by polar-exponential grid technique and Fourier transform, and 3D moving object is recognized by neural network and aspect graphs. An object recognition system is carried out. The experimental results show that this method is effective to recognize 3D moving object.

Keywords Pattern recognition, Polar-exponential grid technique, Polar-log coordinate transform, Profile invariant, Aspect graphs

0 引言

目前, 三维目标信息的获取主要是通过通过对以任意角度观测的二维数字图象(侧面视图)的分析来得到的. 由于三维目标在不同视点下呈现不同的姿态, 所以三维目标的识别是计算机视觉中一个较为复杂的课题. 人类通过视觉进行三维目标的识别, 是对三维目标在视网膜上的二维视图进行分析和归类, 必要

时, 要利用三维目标的视图转移关系. 在计算机视觉中, 对三维目标的识别就是转换成对其二维视图的识别, 需要提取对目标的平移、旋转、缩放具有不变性的特征. Hu 最早提出了矩的概念^[1], 推导了一系列基本性质, 证明了矩的旋转、缩放、平移不变性. 以后人们将仿射不变量^[2], 透视不变量^[3], 拐点不变量^[4]等用于目标识别. Dudani 等将不变矩理论用于飞机的识别^[5], 杜亚娟用归一化的 4 阶矩来识别三维飞机^[6]均获得了良好的效果. 但是矩不变量的计算量大. 极指

基金项目: 导弹控制系统仿真国防科技重点实验室基金项目(51474040201CB0101)

收稿日期: 2002-05-16; 改回日期: 2003-03-14

数栅格采样和极对数坐标变换技术于 1979 年提出^[7],由于其中心清晰,周围模糊的成像特点,且变换本身对图象进行了极大的压缩,因此与传统的压缩标准结合可得到很高的压缩比^[8];由于变换具有旋转不变性和缩放不变性,因此可用于目标跟踪^[9];在人脸检测和跟踪研究中,这种技术的应用也获得了很好的效果^[10]。极指数栅格技术模拟人眼特性进行极指数栅格采样,然后进行极对数坐标映射。映射到极对数坐标下的图象尺寸减小,在此图象上进行识别,减少了计算量,缩短了计算时间。极指数栅格技术的具体介绍详见文献^[11]。方面图是一种图结构,它既含有观察目标三维结构的静态信息(节点),也含有与观察者相对运动的动态信息(弧),可以表示三维目标多个视点下的侧面图及各侧面图之间的转移关系。本文提出了通过图象序列和它们之间的转移关系来识别三维目标的方法。该方法利用极指数栅格技术和傅立叶变换相结合得到目标的轮廓不变量;用神经网络聚类结合方面图识别三维运动目标。实验证明,该方法可有效地识别三维目标。

1 不同视点下三维目标的表示

1.1 概述

一个三维目标能产生多少个侧面视图(简称侧面图),是由目标的复杂程度决定的。两个相隔较远的视点,得到的侧面图差别大,而两个相隔较近的视点得到的侧面图差别小。采用 ART(Adaptive Resonance Theory)神经网络区分两个不同视点的侧面图。侧面图的转移关系是在建立目标模型的时候,通过学习得到的,它是视图之间直接、间接的邻接,转移关系的数学表示构成转移矩阵。三维目标的多个侧面图以及侧面图间的转移关系可用方面图表示。

1.2 方面图

方面图最早是由 Koenderink 提出的^[12]。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得的二维投影侧面图象,节点间的弧代表观察目标与观察者之间相对运动造成的投影图象的转换,这是一种以观察者为中心的表示方法。方面图既含有观察目标三维结构的静态信息(节点),也含有与观察者相对运动的动态信息(弧),因此,被认为在机器视觉研究中具有巨大潜力。假设目标被一个球面包围,称其为目标方面球(aspect sphere),观察者可沿球面从各种角度获得目标的投影侧面图象,

在某个观察角范围内摄取的图象可能很相似,用其中使均方误差最小的一幅作代表,它就是方面图的一个节点,当观察者进入相邻的一个观察角范围内时,所得到的图象会有很大变化,这相当于方面图从一个节点转换到另一个节点。由方面图的定义,一个正方体的方面图可用图 1 表示。

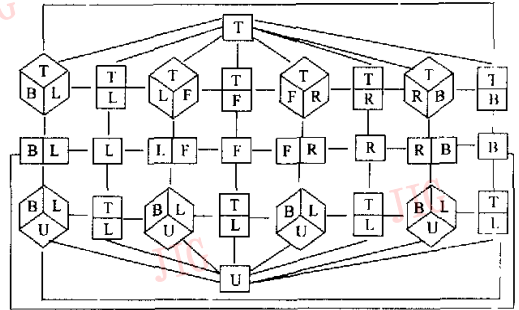


图 1 立方体的方面图

其中, T, U, F, B, L, R 分别表示立方体的上、下、前、后、左、右。

1.3 侧面图的获取

得到目标的侧面图是建立方面图的前提条件。如图 2 所示,将目标固定,以目标为中心,按一个圆球轨迹围绕目标进行观察,在不同视点得到目标的不同侧面图,同时得到它们之间的转换关系。例如摄像机在 B 点采样,那么,物体所进行的扭动也就相当于摄像机沿球面运动,下次进入摄像机的画面应是摄像机在 A、C、D、E 的角度所得。由于目标的对称性,且只使用目标轮廓特征,因此只考虑半球范围即可。

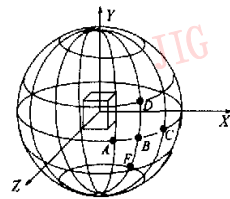


图 2 侧面图获取示意图

2 目标识别系统

2.1 系统概述

系统分为学习和识别两个过程。在学习过程中又分为学习一个目标和学习所有目标。学习一个目标的过程是将该目标的所有侧面图送入神经网络聚类,生成该目标的方面图,得到转移矩阵和典型图象

表示.学习所有目标的过程是将所有目标的典型图象送入神经网络聚类,得到聚类结果送目标模型库,供识别用.整个系统的框架结构如图3所示.

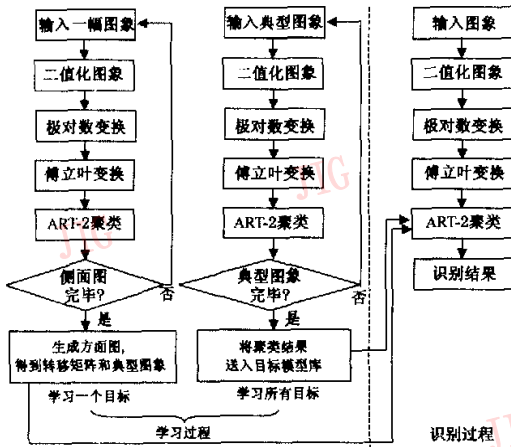


图3 系统结构图

2.2 实验方案

2.2.1 学习过程

(1) 获取侧面图

将摄像机固定,目标只在右半球范围内运动,实验中,目标为两个飞机模型,分别在水平、竖直平面运动180°(水平平面的运动范围是0°~180°,竖直平面的运动范围是-90°~90°),在水平平面每隔30°拍摄一幅图象,在竖直平面每隔15°拍摄一幅图象,每个目标得到79幅图象,图象的分辨率为384×284,转移关系如图4所示.

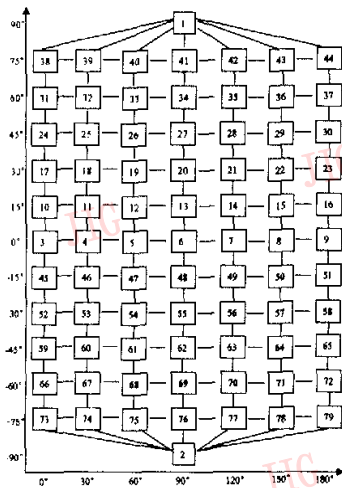


图4 目标的转移关系图

实际中每一水平面上的0°和180°之间也直接相连.

(2) 特征提取

对于获得的三维目标侧面图,二值化后进行极对数坐标变换,得到目标轮廓,这是一个与目标的位置、比例无关的一维向量.对它做一维傅立叶变换,得到轮廓不变量,作为三维目标的二维侧面视图的特征向量.极对数坐标变换如式(1)所示^[11],变换后的图象大小是128×128.

$$u = k_u \ln \left(\frac{r}{r_{\min}} \right)$$

$$v = \theta = k_v \arctan \left(\frac{y}{x} \right) \tag{1}$$

其中,(u, v)是像素在极对数坐标系下的坐标,(x, y)是像素在直角坐标系下的坐标, k_u, k_v 分别为距离分辨率和角度分辨率常数, $r = \sqrt{x^2 + y^2}$ 为像素距变换中心的长度, r_{\min} 是变换的最小半径.

(3) 神经网络聚类

用ART-2神经网络对由上一步得到的特征向量进行聚类.特征向量输入后,经过计算得到输出 r .与预先设定的阈值 ρ ($0 < \rho < 1$)进行比较,若 $r > \rho$,则说明输入与网络中存储的某一类相似,将输入归入该类;若 $r \leq \rho$,则说明输入与网络中存储的每一类都不相似,开辟一个新类.关于ART-2神经网络的具体内容参见文献[13].

(4) 得到每个目标的方面图及其转移矩阵

在学习一个目标的过程中,将每个目标的所有侧面图都送入ART-2神经网络聚类完毕,即得到该目标典型的侧面图表示,将此典型的侧面图作为方面图的节点.侧面图之间的转移关系根据图4作如下规定:同一水平平面上的侧面图直接相连;不同水平平面上的侧面图与竖直平面距离它最近的侧面图相连.由此关系得到每个目标的方面图以及各侧面图间的转移关系,生成转移矩阵.转移矩阵的权值规定为:相邻侧面图之间转换证据分为1,不相邻侧面图之间的转换按 $(0.85)^{n-1}$ 记分, n 为最近转换路径经过的节点数.飞机1、飞机2的典型图象如图5、图6所示.

飞机1、飞机2的方面图如图7、图8所示.

飞机1、飞机2的特征矩阵如表1、表2所示.

(5) 学习所有目标

将上一步得到的每个目标的典型侧面图经过预处理,提取特征后,送入ART-2神经网络聚类,生成聚类结果,存入目标模型库.两个飞机的典型图象

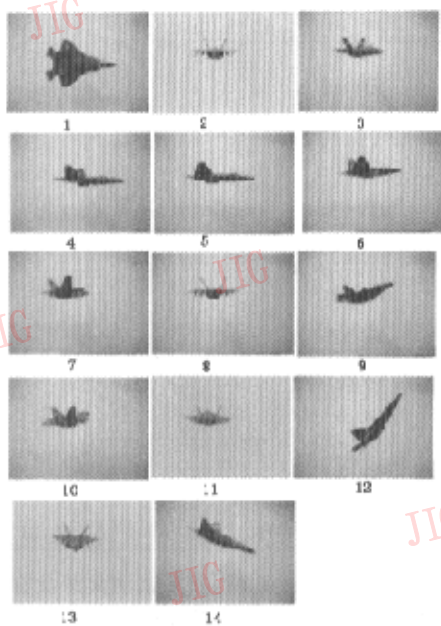


图 5 飞机 1 的典型图象

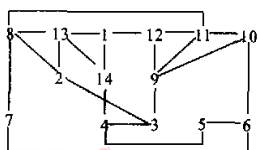


图 7 飞机 1 的方面图

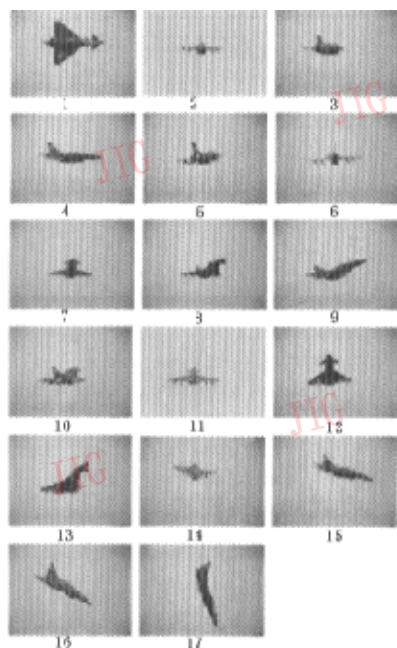


图 6 飞机 2 的典型图象

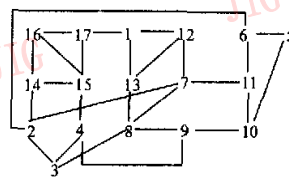


图 8 飞机 2 的方面图

表 1 飞机 1 的转移矩阵

序号	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2	0.85												
3	0.723	1											
4	0.85	0.85	1										
5	0.723	0.723	0.85	1									
6	0.614	0.723	0.723	0.85	1								
7	0.723	0.85	0.723	0.723	0.85	1							
8	0.85	1	0.85	0.723	0.723	0.85	1						
9	0.85	0.85	1	0.85	0.723	0.85	0.723	0.85					
10	0.723	0.723	0.85	0.723	0.85	1	0.85	0.85	1				
11	0.85	0.85	0.85	0.723	0.723	0.85	0.85	1	1	1			
12	1	0.723	0.85	0.723	0.614	0.723	0.723	0.85	1	0.85	1		
13	1	1	0.85	0.85	0.723	0.723	0.85	1	0.723	0.723	0.85	0.85	
14	1	0.85	0.85	1	0.85	0.723	0.723	0.85	0.723	0.614	0.723	0.85	1

表 2 飞机 2 的转移矩阵

序号	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1																
2	0.723															
3	0.723	1														
4	0.723	0.85	1													
5	0.614	0.85	0.85	1												
6	0.614	1	0.85	0.85	1											
7	0.85	1	0.85	0.723	0.723	0.85										
8	0.85	0.85	1	0.85	0.723	0.723	1									
9	0.723	0.723	0.85	1	0.85	0.723	0.85	1								
10	0.614	0.723	0.723	0.85	1	0.85	0.85	0.85	1							
11	0.614	0.85	0.723	0.723	0.85	1	1	0.85	0.85	1						
12	1	0.85	0.723	0.614	0.614	0.723	1	0.85	0.723	0.723	0.85					
13	1	0.723	0.85	0.723	0.614	0.614	0.85	1	0.85	0.723	0.723	1				
14	0.723	1	0.85	0.85	0.723	0.85	0.85	0.723	0.723	0.614	0.723	0.723	0.614			
15	0.85	0.85	0.85	1	0.85	0.723	0.723	0.723	0.85	0.723	0.614	0.723	0.723	1		
16	0.85	0.85	0.723	0.85	0.723	0.723	0.723	0.614	0.723	0.614	0.614	0.723	0.723	1	1	
17	1	0.723	0.723	0.85	0.723	0.614	0.723	0.723	0.723	0.614	0.614	0.85	0.85	0.85	1	1

(共 31 幅, 1~14 为飞机 1 的序号, 15~31 为飞机 2 的序号) 共得到 28 类, 聚类结果如表 3 所示, 其中类别 1、10、14 各包含两个侧面图, 这些侧面图有多义性。若要使每个类别中, 都只含有一个侧面图, 即所有的侧面图没有多义性, 则 ART-2 神经网络中的阈值 ρ 应设成较大的值。

表 3 聚类结果

类别	序号	类别	序号	类别	序号	类别	序号
1	1,15	8	8	15	16	22	25
2	2	9	9	16	17	23	26
3	3	10	10,19	17	18	24	27
4	4	11	11	18	20	25	28
5	5	12	12	19	21	26	29
6	6	13	13	20	22	27	30
7	7	14	14,23	21	24	28	31

2.2.2 识别过程

假设 X, Y, Z 是模型库中的目标, 系统观察其中一个, 从侧面图看上去像 X , 像 Y , 又像 Z , 无法判断到底是哪个目标, 只好给待选者各记一些证据分, 这就是侧面图的多义性; 当通过相对运动获得待识别目标的又一个侧面图时, 如果看上去像 X , 又像 Y , 并肯定不是 Z , 则只给 X 和 Y 再记一些证据分; 如果以后连续获得几个侧面图象都像 X , 则候选者 X 连连得分, 最后按得分多少认定该目标是 X , 这就是胜者为王准则。在观察和识别运动目标时, 可以连续获取几个侧面图, 这样有利于准确地识别目标。对于静态目标只能获取一个侧面图, 如果该侧面图

具有多义性, 则无法判断; 如果该侧面图没有多义性, 虽然可以判断, 但显得证据不够充分, 可靠性不高, 此时需要视觉系统通过自身运动获取多个侧面图, 得到更多证据, 因此该系统更善于识别运动目标, 也就是通过相对运动获取目标的更多的三维信息, 以提高目标的识别精度。

系统通过获得的多个侧面图来识别三维运动目标。与学习过程类似, 将待识别侧面图序列预处理, 提取特征, 送 ART-2 神经网络聚类。在本系统中, ART-2 神经网络的阈值设为 0.8, 当输入经过计算后得到的 $r > 0.8$, 则进行证据分的计算; 若 $r \leq 0.8$, 则说明该目标不是目标库中的目标, 不进行证据分的计算。规定侧面图入选证据分为 0.5, 然后通过目标的转移矩阵权系数的计算, 得到相应的证据分, 根据胜者为王准则, 判断其属于哪一类目标。识别中, 假设拍摄到的图象序列如图 9 所示。

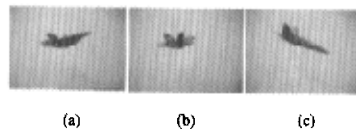


图 9 待识别的目标

经过神经网络聚类得到: 图 9(a) 属于第 9 类 (序号 9), 图 9(b) 属于第 10 类 (序号 10, 19), 图 9(c) 属于第 14 类 (序号 14, 23), 通过转移矩阵得到飞机 1 的证据分为 $0.5 + 1 + 0.614 = 2.114$, 飞机 2 的证据分为 $0.5 + 0.85 = 1.35$, 故该目标是飞机 1。

2.3 小结

(1)系统通过图象序列识别三维运动目标,并且利用了运动目标之间状态转移关系,能有效识别目标。

(2)系统中,输入 3 幅图象基本就可以得到正确的识别结果,在 CPU 为 P III 733 的计算机上进行处理需要花费 5.05s。若将本系统用于实时处理,则需要 CPU 更快的计算机。

(3)若在学习过程中,将 ART-2 神经网络的阈值 ρ 设为较大的值,聚类结果中各类别只有一个侧面图,则在识别过程中,一幅图象通过神经网络聚类,其输出 $r > \rho$,就能识别该目标;若学习过程中,聚类结果中有的侧面图有多义性,则在识别过程中,一幅图象通过神经网络聚类,其输出 $r > \rho$,并且其侧面图没有多义性(即通过检索目标模型库,该图象不在类 1,10,14),就能识别该目标。在上述两种情况下,若神经网络的输出 $r \leq \rho$,则该目标不是模型库中的目标。这说明只有一幅图象时,神经网络聚类的输出 $r > \rho$,且该图象没有多义性,也能识别目标。

(4)若拍摄到的目标虽有小部分被遮挡,但是在极对数坐标下的目标轮廓没有太大的变化,如图 10 所示的图象序列,经过实验证明,能够正确识别。

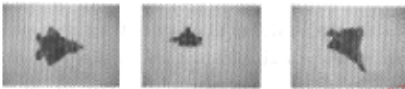


图 10 小部分遮挡的特识别目标

3 结论

该方法适用于运动目标识别,对于静态目标,则需要摄像机是可以运动的,从而可拍摄到静态目标的多个侧面图进行识别。该方法也有局限性,在学习和识别过程中需要三维目标的多个侧面图,三维目标越复杂,则侧面图越多,所占用的系统空间就越大,因此在实际中,应正确处理侧面图数量和系统空间的关系,这也是今后需要进一步研究的问题。

参考文献

- 1 Hu M K. Visual pattern recognition by moment invariant[J]. IRE Transactions on Information Theory, 1962, 8(2): 179 ~ 187.
- 2 Arbert K, Snyder W E. Application of affine invariant fourier descriptors to recognition of 3D objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(7):

- 640~647.
- 3 董敏,徐正伟,吴成柯. 基于透视不变量的平面曲线的识别[J]. 模式识别与人工智能, 1995, 8(2):108~113.
- 4 Nirwan A, Edward J D. On detecting dominant points [J]. Pattern Recognition, 1991,24(5):441~445.
- 5 Wallace T P, Wintz P. An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors[J]. Computer Graphics and Image Processing, 1980, 13(2): 99 ~ 126.
- 6 杜亚娟,张宏荣,潘泉. 基于矩特征的三维飞机目标识别[J]. 数据采集与处理, 2000, 15(3): 390~394.
- 7 Carl F R. Weirman, George Chaikin. Logarithmic spiral grids for image processing and display [J]. Computer Graphics and Image Processing, 1979,11(2):197~226.
- 8 吴晖. 仿生视觉理论及其在图像压缩中的应用[D]. 哈尔滨:哈尔滨工程大学,1998.
- 9 Bailey J G. Messner R A. Log-polar mapping as a preprocessing stage for an image tracking system[A]. In: Proceedings of SPIE Intelligent Robots and Computer Vision; Seventh in a Series [C], Cambridge, Massachusetts, USA,1988, 1002: 15~22.
- 10 Jurie F. A new log-polar mapping for space variant imaging application to face detection and tracking [J]. Pattern Recognition, 1999,32(5):865~875.
- 11 张海燕,宋克欧,王东木. 图象序列运动目标特征点对应的极指数栅格方法[J]. 中国图象图形学报, 2003,8A(5):527~532.
- 12 Koenderink J J, Van Doorn A J. The singularities of the visual mapping[J], Biological Cybernetics, 1976,24(1):51~59.
- 13 郑君里,杨行峻. 人工神经网络[M]. 北京:高等教育出版社,1992.



张海燕 1976 年生,博士研究生,1998 年获山东师范大学理学学士学位,2001 年获哈尔滨工程大学工学硕士学位。主要研究方向为计算机视觉、图象处理、模式识别等。



宋克欧 1940 年生,哈尔滨工程大学计算机科学与技术学院教授。主要研究方向为图象处理、模式识别、计算机视觉、水声信号特征提取、石油测井曲线特征提取和识别等。



王东木 1944 年生,航天科工集团北京仿真中心总工程师,研究员。主要研究方向为飞行器自动控制系统设计、系统仿真、分布交互仿真、虚拟现实技术、试验学等。