

基于单视觉原理实时提取运动物体三维信息的一种方法

洪海 姜伟 李金宝

(黑龙江大学计算机科学系, 哈尔滨 150080)

摘要 针对多视觉方式中对应点匹配及系统复杂性等问题, 采用单视觉方式, 应用专用特征点匹配器实时进行特征点的匹配, 运用刚体运动模型、摄像机模型、物体几何形状模型以及旋转矩阵的正交性, 给出了一种实时提取运动物体平移、旋转等6个自由度信息的新方法。

关键词 单视觉, 实时, 运动参量提取

1 引言

计算机视觉实用化的关键在于系统的简易性、实时性和可靠性。基于视觉原理获取运动物体三维信息主要是通过单视觉方式和多视觉方式。多视觉方式的对应点匹配问题尚未得到圆满的成果^[1,2]。单视觉方式结构简单, 不存在多幅画面对应点匹配问题, 因此易于实现实时化; 如果利用物体的几何形状模型, 也能够同时获取三维物体的形状、姿态和位置信息^[3~9]。

本文采用单视觉方式, 充分利用对象物体几何形状上的约束条件以及透视投影变换数学模型中的约束条件, 克服了多视觉方式对应点匹配上的致命弱点, 保证了系统的实时性、简易性和可靠性。该方法的技术关键在于:

(1) 如何建立空间物体的运动模型;

(2) 如何实时地、可靠地识别和匹配画面上特征点

和模型上特征点间的对应关系;

(3) 如何实时地、正确地求解通过二维信息复原其三维信息的数学模型;

(4) 如何实时地分离出物体的三维运动参量。

2 特征点匹配器原理

本文给出的特征点匹配器, 不经过视频图象的数字化、存储等处理, 充分利用操纵器的几何形状信息, 直接从视频信号中实时实现特征点的提取和匹配, 保证了处理过程的实时性^[10]。特征点匹配器的结构如图1:

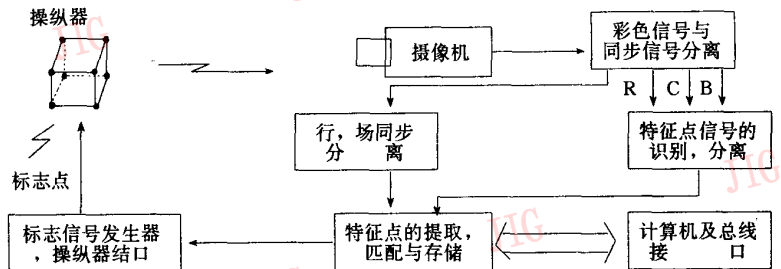


图1

Fig. 1

操纵器发出光信号,此信号通过场同步信号控制,点的位置确定是靠行扫描时间间隔内,产生 512 个脉冲的脉冲发生器,通过计数确定 X 坐标;Y 坐标通过行同步脉冲的计数来确定。特征点的识别靠灰度的阈值处理来实现。

3 数学模型及其解法

在本文方法中,定义如图 2 所示的坐标关系。 $X_m-Y_m-Z_m$ 是模型坐标系(与世界坐标系一致),空间物体模型定义在该坐标系上; $X_e-Y_e-Z_e$ 是视坐标系; X_p-Y_p 是图象坐标系,空间物体经透视投影变换生成的图象建立在该坐标系上

空间物体顶点可以用矩阵形式来描述^[11];对空间物体进行坐标变换可以转化为对矩阵运算来实现。

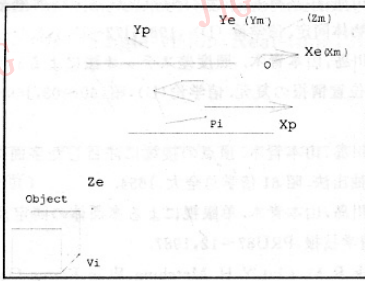


图 2 坐标关系

Fig. 2 Coordinate relation

定义变换矩阵 T 为如下形式:

$$[T] = [L][R][F][P] \quad (1)$$

其中 $[L]$ 是右手坐标系向左手坐标系转换矩阵。

$$[R] = [X][Y][Z]$$

矩阵 $[X]$ 、 $[Y]$ 、 $[Z]$ 是空间物体绕坐标轴 X 、 Y 、 Z 的旋转矩阵。矩阵 $[F]$ 是平移变换矩阵。矩阵 $[X]$ 、 $[Y]$ 、 $[Z]$ 、 $[F]$ 是图形学中的基础矩阵,这里不详细给出。

$$[P] = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -2 & 1/f \\ 0 & 0 & -f & 0 \end{bmatrix}$$

是透视投影变换矩阵, f 是摄像机焦距 ($f > 0$), 该矩阵是根据摄像机模型定义的。

设一个空间物体 V_m 有 N 个顶点, $[V_m]$ 为顶点描述矩阵。对 $[V_m]$ 进行 T 变换:

$$[V_m][T] = [V_p] \quad (2)$$

$[V_p]$ 是一个 $N \times 4$ 的矩阵,元素为空间物体 V_m 的顶点经过 $[T]$ 变换后在投影平面 P 上所成二维画面的特征点坐标。

在本文方法中,用摄像机采集 V_p 数据,空间物体 $[V_m]$ 已知,根据式(2)逆向求解矩阵 $[T]$, $[T]$ 是一个 4×4 的矩阵。

设

$$[T] = \begin{bmatrix} -r_{11} & -r_{12} & -2r_{13} & r_{13}/f \\ -r_{21} & -r_{22} & -2r_{23} & r_{23}/f \\ r_{31} & r_{32} & 2r_{33} & -r_{33}/f \\ -X_0 & -Y_0 & -2Z_0 - f & Z_0/f \end{bmatrix} \quad (3)$$

$[T]$ 中有 12 个独立元素。由于透视投影变换是不可逆的,所以由 $[V_p]$ 和 $[V_m]$ 求解 $[T]$ 是一个非线性的过程。

本文给出的方法,物体模型是一个立方体,在物体处于某一状态,对 $[T]$ 进行正规化处理后的 $[T']$ 进行整理并分离各个分量,得到一个含有 12 个未知量的 11 维方程组。由 11 维方程组求解 12 个互相独立的未知量,通常的最优化方法速度慢,不能做到实时,并且无法保证解的收敛性和解的唯一性,因此,不能使用在实际应用系统中。

我们根据模型的物理意义,把这类问题化解为线性或简单的代数方程,给出了此类方程在各种条件下的解。

4 实验结果

以上给出了该方法的理论依据,下面,通过模拟实验来检验该方法的正确性和可行性。在模拟实验中,首先,对特征点匹配器的各项技术指标在自然条件下进行了实际测试,结果如下:

- (1) 实现了特征点的提取与匹配。
- (2) 特征点的提取、匹配、存储时间小于 1/25 秒。
- (3) 特征点提取的分辨率优于 500×300 。
- (4) 实现了特征点匹配器与计算机之间的实时数据传输。

在算法测试方面,根据矩阵 $[T]$ 各个分量的意义,通过输入 α 、 β 、 γ 、 X_0 、 Y_0 、 Z_0 6 个运动参量确定 $[T]$,并根据式(2),求解 $[V_p]$,得到一个二维画面 D ;反过来通过 $[V_m]$ 和 $[V_p]$ 求解出 $[T']$,也得到一幅二维画面 D' ,通过 $[T]$ 和 $[T']$ 的具体

实验数据比较,及画象 D 和画象 D' 的重合情况的比较以及运行时间来分析系统的准确性、精确度及实时性。

给定参量: $\alpha=21.3, \beta=10.6, \gamma=23.1, X_0=2.492, Y_0=1.578, Z_0=4.137$, 根据这些参数求得的变换矩阵 $[T]$ 如下:

```

-0.90652  -0.384795  -0.347296  4.3412e-002
0.307496  -0.88368    -0.705847  8.8231e-002
0.289253  -0.266536    1.83879   -0.229849
-2.492    -1.578         -12.274    1.03425

```

根据 $[V_m]$ 和 $[V_p]$ 求解的 $[T]$ 如下:

```

-0.90652  -0.384795  -0.347295  4.34119e-002
0.307495  -0.883681  -0.705853  8.8232e-002
0.289253  -0.266536    1.83879   -0.229849
-2.492    -1.578         -12.274    1.03425

```

画象 D, D' 及二者的重合图 D'' 如下:

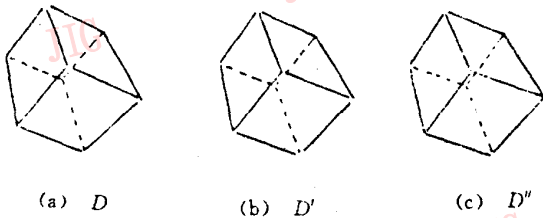


图3 画象 D, D' 和二者的重合图 D''

Fig. 3 Image D, D' and overlay image D''

对比 $[T]$ 和 $[T']$, 误差是 10^{-3} 。而从 D'' 来看, D'' 和 D 及 D' 完全一致, D'' 没有重合的痕迹。

在 LX-486 计算机上, 给定以上参数, 进行了 1 000 次连续模拟实验, 测得系统运行时间为 0.038 秒/次, 其中算法运行时间为 0.027 秒/次, 二维画象的显示时间为 0.011 秒/次。可见, 本文方法是实时的。

以上实验结果表明, 从理论上讲, 本文方法是完全正确、实时的。

5 结束语

本文给出了运用单视觉原理实时提取运动物体三维运动参量的一种方法, 并通过实验验证了该方法的正确性。本文方法在刚体三维运动参量的提取、运动机器人对环境的理解及自身方位的确定、人一机三维信息交互装置等方面有着广阔的应用前景。今后, 作为本方法的实际应用, 将在人一机三信息交换接口方面做进一步研究。

参考文献

- 1 池内克史. 单眼视における2次元画象からの3次元情報の再構成. 情报处理, 昭58~12(24)12:1421~1427.
- 2 Roberts L G. Machine Perception of Three-dimensional Solids, Optical and Electro-optical Information Processing. J T Tippet, ed, MIT Press, Cambridge MA, 1965:157~159.
- 3 洪海, 川岛, 山本青木. 回转行列の正规直交性拘束を用いた多面体の物体同定. 信学论(D), 1989, J72~D, 3:404~415.
- 4 洪海, 川岛, 山本青木. 照度差ステレオ法による3次元物体の形状と位置情報の復元. 信学论(D), 昭, 60~03, J69~D, 3:427~433.
- 5 洪海, 川岛, 山本青木. 頂点の接続に注目した多面体の3次元位置情報抽出法. 昭61 信学总全大, 1654.
- 6 洪海, 川岛, 山本青木. 单眼视による多面体の同定及び位置の決定. 信学技报, PRU87~12, 1987.
- 7 Haralick R M, Chu Y H. Matching Wire Frame Objects from Their Two Dimensional Perspective Projections. Pattern Recognition, 1984, 17(6):607~618.
- 8 Sugihara K. An Algebraic Approach to Shape-from-image Problems. Artif Intell, 1984, 23:59~95.
- 9 白井良明. 3次元モデルに基づくロボットツヨン. 信学论(D), 昭60~44, J68~D, 4:449~455.
- 10 洪海, 李金宝等. 实时识别与匹配运动物体特征点图象处理器的研制. 机器人, 1995, 7: 91~96
- 11 洪海, 姜伟等. 单眼视觉实时提取物体三维信息的方法. 机器人, 1995, 7:111~118



洪海, 黑龙江大学教授, 黑龙江大学信息技术研究所副所长。1984~1989 年以中国政府派遣出国研究生的身份在日本北海道大学大学院相继获得硕士、博士学位。其后在日本 NTT Human interface 研究所等地做博士后及研究工作。主要从事计算机视觉、计算机图形学、CAD/CAM、多媒体技术等方面的研究。

A New Method for Obtaining Moving-object 3-D Information Based on Single-vision

Hong Hai, Jiang Wei, Li Jinbao

(Department of Computer Science Heilongjiang University, Haerbin 150080)

Abstract Because of it's difficult to match the object model feature points in Multi-Vision system. In this paper a method which obtains moving-object 3D-information based on single-vision be given. This method based on object-model, camera-model and object-shape model applies a kind of special image-processor which may realize the real time extraction, matching of the object model feature points readize real-time obtain moving-object 3D information.

Keywords Single-vision, Real-time, 3D information obtain

(上接 463 页)

A Multi-viewpoint Modeling Method Based on Stereo Vision

Dong Zaili, Wang Jiangan, Xu Xinping

(Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110015)

Abstract A multi-viewpoint modeling method based on stereo vision is presented in this paper. The local 3D information of objects are acquired by a stereo vision system at different viewpoints. The local models of an object are merged into a global model by matching feature points and using coordinate transformation. A spatial vector-based resolvent method of coordinate transformation has been developed.

Keywords Stereo vision, Model buiding, Coordinate transformation