

# 基于视频跟踪的体感游戏设计

张英烈<sup>1</sup>, 夏梦<sup>2</sup>, 陈临强<sup>1</sup>

ZHANG Yinglie<sup>1</sup>, XIA Meng<sup>2</sup>, CHEN Linqiang<sup>1</sup>

1. 杭州电子科技大学 图形图像研究所, 杭州 310018

2. 杭州电子科技大学 计算机学院, 杭州 310018

1. Institute of Graphics and Image, Hangzhou Dianzi University, Hangzhou 310018, China

2. College of Computer Science, Hangzhou Dianzi University, Hangzhou 310018, China

**ZHANG Yinglie, XIA Meng, CHEN Linqiang. Design of somatosensory games based on video tracking. Computer Engineering and Applications, 2015, 51(17): 44-47.**

**Abstract:** This paper designs a somatosensory game about a panda's adventure in bamboo forest. The design of somatosensory game includes game design and body moving extraction and analysis. In the part of game design, it discusses from scene design, role design and rule design. In the part of body moving extraction and analysis, human object is extracted by the average background method. Human's face and hands are identified by the Bayesian color model. In the following, the tracking of human's face and hands is achieved by Kalman filtering. It analyzes the moving of face and hands to control game-role's movement, so that to achieve the control of game-role with body action.

**Key words:** somatosensory games; video tracking; moving object detection; Bayesian color model; Kalman filtering

**摘要:**设计了一款熊猫竹林冒险体感游戏,体感游戏的设计主要包括游戏设计和肢体运动提取分析两部分。游戏设计从场景设计、角色设计和规则设计三方面进行论述。肢体运动提取分析采用平均背景建模法建立背景模型,提取人体目标,根据贝叶斯肤色模型识别人的手脸,再利用 Kalman 滤波实现对手脸的跟踪,分析手脸运动控制游戏人物运动,从而实现用肢体动作操控游戏人物。

**关键词:**体感游戏;视频跟踪;运动目标检测;贝叶斯肤色模型;Kalman 滤波

**文献标志码:**A **中图分类号:**TP391 **doi:**10.3778/j.issn.1002-8331.1309-0115

## 1 引言

近年来,随着游戏产业的日趋发展,传统的电子游戏已经无法满足人们对于娱乐的需求,于是,体感游戏<sup>[1]</sup>作为娱乐性和交互性更强的新一代大众游戏应运而生。体感游戏是一种通过玩家的肢体动作变化来操作游戏目标的新型电子游戏,它突破了传统的以键盘、手柄、游戏操纵杆等特定设备为输入的操作方式,通过人体姿态、动作手势等肢体动作实现对游戏的控制和操作,极大地提高了游戏沉浸感,极大地丰富了玩家的体感体验。而计算机视觉、人机交互和虚拟现实等技术的迅速发展和成熟也为体感游戏提供了强大的技术支持,使得体感游戏有着极为广阔的发展前景。

目前体感游戏代表平台有美国微软开发的 Xbox360、

日本任天堂公司的 Wii<sup>[2]</sup>、日本索尼的 PlayStation Move 和国内亿旗体感公司的 17vee 体感平台等。体感游戏也是国内外学者研究的热点问题,芬兰技术研究中心的 Virku<sup>[3]</sup>项目将体感技术和自行车健身系统结合起来,实现了用户通过大屏幕进行虚拟环境漫游的运动体验, Ping Pong Plus<sup>[4]</sup>是 MIT 开发的将乒乓球运动和体感技术相结合的系统,日本国家信息通讯技术研究所设计实现的 EnergyBrower<sup>[5]</sup>和微软的 DanceYourWork<sup>[6]</sup>,也都是较为成功的有体感体验的虚拟健身游戏系统。谢文正<sup>[7]</sup>将划船健身器同游戏相结合实现了一种基于体感技术的划船健身器游戏,耿彬<sup>[8]</sup>实现了利用 Google Earth 地图匹配游戏地形的跑步机体感游戏,李子楠<sup>[9]</sup>通过对手势的识别和颜色的跟踪实现了手势控制的跳舞机,李波<sup>[10]</sup>

**作者简介:**张英烈(1990—),男,硕士研究生,研究领域为图像处理;夏梦(1993—),女,研究领域为图像处理;陈临强(1963—),男,教授,研究领域为视频图像处理、计算机视觉。E-mail:326997612@qq.com

**收稿日期:**2013-09-10 **修回日期:**2013-11-13 **文章编号:**1002-8331(2015)17-0044-04

**CNKI 网络优先出版:**2014-01-26, <http://www.cnki.net/kcms/doi/10.3778/j.issn.1002-8331.1309-0115.html>

实现了一种基于传统健身单车和网络游戏相结合的健身单车体感游戏。但这些体感游戏硬件设备要求严格,配置环境复杂,游戏成本高,不利于体感游戏的普及和推广。本文利用普通摄像头设计了一款基于视频跟踪的熊猫竹林冒险的体感游戏,极大降低了体感游戏的硬件要求和成本,使玩家能够在任意一台装有摄像头的电脑上上进行体感体验,在游戏娱乐的同时也能锻炼身体。

## 2 系统整体流程

体感游戏的设计主要包括游戏设计和肢体运动提取与分析两部分。游戏设计主要完成对游戏的构思,包括场景设计、角色设计、规则设计等。肢体运动提取与分析主要完成对人体目标检测、肢体目标识别、肢体目标跟踪和肢体动作分析,主要涉及平均背景建模法、贝叶斯肤色模型和Kalman滤波等技术。

本文设计的体感游戏是通过普通摄像头实时采集玩家动作,经过视频处理技术完成手和脸的位置识别和实时跟踪,然后分析手和脸的运动并转换为对游戏人物的实际控制数据,从而控制游戏人物做出与玩家运动相应的动作,并根据游戏规则判断游戏是否结束,更新游戏状态,直到游戏结束。具体的流程如图1所示。

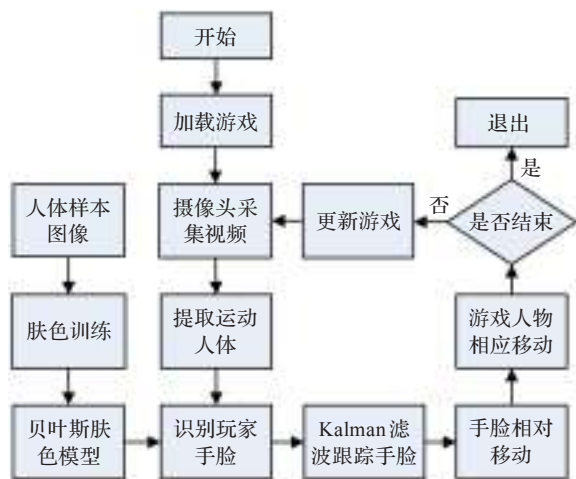


图1 系统流程图

## 3 游戏设计

### 3.1 游戏介绍

熊猫竹林冒险是一款模仿 Temple Run 的体感游戏,操作方便简单有趣,主要锻炼玩家的反应速度和敏捷度。游戏中,一只熊猫在竹林里迷路,玩家要用身体控制熊猫,帮助它躲避各种危险,并收集尽可能多的金币,完成冒险。

### 3.2 场景设计

游戏场景以瀑布和竹林为背景,以竹子为中景,以熊猫为近景。其中有左、中、右三根竹子供熊猫左右移动。竹子上随机产生坏竹节和金币,右侧的计分板,显示熊猫得分。计分板上的小熊猫个数代表熊猫的生命值。具体场景图如图2所示。



图2 游戏场景图

### 3.3 角色设计

游戏的主角是一个熊猫。由3dsmax建模设计得到的熊猫效果图如图3所示。熊猫动画的制作是通过将3dsmax预设的骨骼系统应用到熊猫模型上,并利用动作捕捉数据驱动这套骨骼,从而使熊猫完成向上爬,左移,右移,跳起的动作。熊猫动作效果图如图4所示,依次为熊猫向上爬,跳起,左移,右移的动作效果。



图3 熊猫效果图



图4 熊猫动作效果图

### 3.4 规则设计

(1)控制方式:游戏时熊猫自行向上移动,玩家通过肢体动作来控制熊猫左移、右移、跳起获取金币和避开坏竹节。

(2)得分方式:熊猫爬过一个竹节得1分,获取一个金币得10分,以此累加直至游戏结束。

(3)难度控制:当得分达到一定数值,坏竹节个数增多和爬行速度加快,以提高游戏难度。

(4)游戏结束:熊猫有三条生命,碰到坏竹节损失一条生命,直至游戏结束。

## 4 肢体运动提取与分析

对采集到的视频进行处理获取肢体运动信息是体感游戏实现的关键,本文使用平均背景法建立背景模

型,获取运动人体目标,再利用贝叶斯肤色模型识别人的手脸等肢体部位,并进行标识,然后使用 Kalman 滤波进行实时跟踪,最后通过分析肢体目标位置判定肢体运动,并对游戏人物进行相应的操作。

#### 4.1 平均背景法提取运动人体目标

常用的运动目标检测方法有光流法、帧差法、背景差分法。其中,光流法运算复杂,一般很难满足实时处理的要求<sup>[11]</sup>。帧差法会使目标内部产生空洞,无法得到完整的前景目标<sup>[12]</sup>。背景差分法检测目标速度快,检测准确<sup>[13]</sup>,但背景建模技术是其性能好坏的关键。平均背景建模法运算简单,计算速度快,适用于室内环境,因此本文采用平均背景建模法建立背景模型,提取运动目标。

平均背景建模法的基本思路是计算每个像素的平均值和标准差作为它的背景模型。主要步骤如下:

(1) 累积背景图像和每一帧图像差值的绝对值,通常需要学习 30 至 1 000 帧,本文学习 100 帧视频图像建立背景模型。

平均像素值公式为:

$$A(x, y) = \frac{1}{n} \sum_{i=1}^n |f_i(x, y)| \quad (1)$$

式中,  $n$  为总帧数,表示第  $i$  帧图像在  $f_i(x, y)$  处的像素值。

像素的平均差分公式为:

$$Diff(x, y) = \frac{1}{n} \sum_{i=1}^n |f_i(x, y) - f_{i-1}(x, y)| \quad (2)$$

式中,  $n$  为总帧数,  $f_{i-1}(x, y)$  和  $f_i(x, y)$  为连续两帧图像。

(2) 学习完 100 帧图像后,计算每一个像素的均值和平均差分,建立背景的统计模型。同时设置高低阈值,将背景像素和前景像素区别开。根据实验测试结果,将高、低阈值分别设置为 4 和 3,即每一帧图像的绝对差大于平均值的 4 倍的像素和绝对差值小于平均值 3 倍的像素认定为前景。

(3) 背景模型建立后,对每一帧中的像素值进行判定,其公式如下:

$$B_i(x, y) = \begin{cases} 1, & |f_i(x, y) - A_i(x, y)| > Diff_i(x, y) \times \alpha, \text{ or} \\ & |f_i(x, y) - A_i(x, y)| < Diff_i(x, y) \times \beta \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

式中,  $\alpha$  和  $\beta$  分别为高、低阈值,在阈值范围内的像素值取“1”,否则取“0”,在非零区域即可检测到运动目标。

#### 4.2 贝叶斯肤色模型识别肢体目标

人体肤色在 YCbCr 颜色空间的分布相对集中,肤色之间的差异主要由亮度决定,而与颜色分量无关<sup>[14]</sup>。利用 YCbCr 颜色空间亮度与色度分离的特点将图像像素分为肤色和非肤色像素两类,能有效提高肤色检测的速度和正确性。建立肤色模型是为了建立一个对图像像素点是否属于肤色的判定规则,利用贝叶斯决策能有效提高肤色判定的准确性。贝叶斯决策方法是先根据系

统的先验概率密度预测系统的后验概率密度,然后用得到的系统观测值对后验概率密度进行修正,最终得到最优决策。因此本文采用基于 YCbCr 颜色空间的贝叶斯肤色模型对肢体目标进行识别<sup>[15]</sup>。建立肤色模型主要步骤如下:

(1) 采集不少于 100 张的不同背景、不同人物的图像作为样本图像。本文使用 400 张样本图像以提高肤色模型精度。

(2) 为每张样本图像进行二值化制作二进制掩码图像,其中白色区域为肤色区域,黑色区域为非肤色区域。

(3) 将样本图像从 RGB 空间转换到 YCbCr 空间,仅考虑肤色在 Cb 分量和 Cr 分量的贝叶斯分布。

(4) 利用 YCbCr 图像和二进制掩码图像建立 Cr 分量肤色直方图、Cb 分量肤色直方图、Cr 分量非肤色直方图和 Cb 分量非肤色直方图,得到肤色模型。

利用训练得到的肤色模型对图像的像素点进行肤色判断主要是对像素点是肤色的概率进行阈值比较,该阈值为根据实验效果设定的经验阈值,本文中阈值设定为 0.8。当像素点是肤色的概率大于阈值时,认定此像素点为肤色,即可识别出人体目标中的脸部和手部,然后对识别结果进行形态学开运算即先腐蚀后膨胀处理,得到手脸目标并进行标识。

#### 4.3 Kalman 滤波实现肢体目标跟踪

Kalman 滤波具有计算量小、存储量低、准确度高、适合于实时处理等优点<sup>[16]</sup>,能够较为准确地预测出跟踪目标的下一位置,大大减少了目标匹配的计算量,有效提高了跟踪速度。Kalman 滤波包括预测和更新两个阶段,在预测阶段,滤波器利用上一状态的估计值对当前状态做出估计。

系统状态预测:

$$X(k|k-1) = AX(k|k-1) + BU(k) \quad (4)$$

式中,  $A$  是状态转移矩阵,  $B$  是控制矩阵,  $X(k|k-1)$  是利用上一状态预测的结果,  $X(k-1|k-1)$  是上一状态最优的结果,  $U(k)$  为现在状态的控制量。

协方差预测:

$$P(k|k-1) = AP(k-1|k-1)A^T + Q \quad (5)$$

式中,  $P(k|k-1)$  是  $X(k|k-1)$  对应的协方差,  $P(k-1|k-1)$  是  $X(k-1|k-1)$  对应的协方差,  $A^T$  表示  $A$  的转置矩阵,  $Q$  是系统过程的协方差。

在更新阶段,滤波器再利用当前状态的系统观测值对预测阶段的估计值进行修正优化,以获得更精确的当前状态的估计。

系统状态更新:

$$X(k|k) = X(k|k-1) + Kg(Z(k) - HX(k|k-1)) \quad (6)$$

式中,  $Kg$  为卡尔曼增益,  $Z(k)$  是  $k$  时刻的测量值。

卡尔曼增益:

$$Kg(k) = P(k|k-1)H^T / (HP(k|k-1)H^T + R) \quad (7)$$

式中,  $R$  是测量过程的协方差,  $H$  是测量矩阵。

状态协方差更新:

$$P(k|k) = (I - Kg(k)H)P(k|k-1) \quad (8)$$

式中,  $I$  为 1 的矩阵。

跟踪时首先对每一个肢体目标分配一个 Kalman 滤波器,并分别初始化,然后利用 Kalman 滤波器预测肢体目标在当前帧的位置,再利用贝叶斯肤色模型对预测位置进行检验和匹配,然后对预测位置进行校正,反复迭代,不断跟踪肢体目标,并记录跟踪到的位置信息。

#### 4.4 肢体运动分析

分析肢体运动就是对一定间隔帧数的两帧对应的肢体目标中心点  $p_i(x, y)$  的横纵坐标的比较。当玩家左手向左移动,左手目标的中心点横坐标会变大,相应地玩家右手向右移动,右手目标中心点横坐标会变小,因此可依据图像中的肢体目标的移动来确定玩家的运动。根据实验测试效果,本文将间隔帧数  $NumDiff$  设置为 3,即每隔 3 帧就对肢体目标中心点横纵坐标进行比较。为了避免肢体的细微移动对肢体运动分析的干扰,需要设置一个移动阈值  $TH_{move}$ ,根据实验测试效果,本文将移动阈值设置为 20,即肢体目标中心点的横纵坐标变化超过 20 个像素才能认定肢体目标运动。根据设置的参数,肢体运动判定公式如下:

$$Result = \begin{cases} 1, & x_i(left) - x_{i-NumDiff}(left) > TH_{move} \\ 2, & x_{i-NumDiff}(right) - x_i(right) > TH_{move} \\ 3, & y_i(left) - y_{i-NumDiff}(left) > TH_{move}, \text{ and} \\ & y_i(right) - y_{i-NumDiff}(right) > TH_{move}, \text{ and} \\ & y_i(face) - y_{i-NumDiff}(face) > TH_{move} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

式中,  $x_i(left)$  和  $y_i(left)$  表示当前帧中左手目标中心点的横纵坐标,  $x_i(right)$  和  $y_i(right)$  表示当前帧中右手目标中心点的横纵坐标,  $y_i(face)$  表示当前帧中脸部目标中心点的纵坐标,  $i$  表示当前帧,  $i - NumDiff$  表示与当前帧间隔一定帧数的之前帧。

(1) 当左手目标横坐标变大且超过移动阈值,即左手向左移动,  $Result$  取“1”,即控制熊猫向左移动。

(2) 当右手目标横坐标变小且超过移动阈值,即右手向右移动,  $Result$  取“2”,即控制熊猫向右移动。

(3) 当手脸目标纵坐标都变大且超过移动阈值,即手和脸都向上移动,  $Result$  取“3”,即控制熊猫向上跳起。

(4) 否则熊猫不改变移动状态,继续向上移动。

#### 5 实验效果

本文的实验环境:主机为处理器为 i3 2.0 GHz,内存为 2 GB 的 PC 机,开发工具为 Visual C++ 6.0 和 OpenCV1.0。实验视频由 PC 机自带摄像头采集,视频

图像尺寸为  $320 \times 240$ 。

玩家在摄像头前做出不同的动作,经过一系列的处理,最终控制游戏中的熊猫做出相应的操作,部分游戏效果图如图 5 所示。其中(a1)为玩家的初始状态,红色方框是对识别到手脸目标的标识,(a2)为熊猫的初始状态,玩家没有做出特定动作前,熊猫保持自行向上爬行的状态,(b1)为玩家右手向右移动,(b2)表示当玩家右手向右移动时,熊猫受玩家控制向右移动,(c1)为玩家左手向左移动,(c2)表示当玩家左手向左移动时,熊猫受玩家控制向左移动。在实时性方面,系统平均运行速度约 27 帧每秒,由于对玩家的运动方向的确定需要一定帧数的计算,所以熊猫的移动相对玩家的运动有几帧的滞后,不过基本可以满足实时的需求。



图5 游戏效果图

#### 6 结束语

本文利用平均背景建模法、贝叶斯肤色模型和 Kalman 滤波等视频图像处理技术,实现了一款简单有趣的熊猫竹林冒险的体感游戏。实验表明,游戏的操作性和灵敏度都能够基本满足人们的需求。同时,无需其他特殊输入设备,普通的摄像头即可满足游戏的需求的特点,让人们在较低的硬件环境下就能体验到体感游戏带来的身体就是操纵器的全新感受,体会到前所未有的快乐。

#### 参考文献:

- [1] Falstein N.A grand unified game theory[C]//Game Developers Conference Proceedings, 1999: 229-239.
- [2] Schlömer T, Popponga B, Henze N, et al. Gesture recognition with a Wii controller[C]//Proceedings of the 2nd International Conference on Tangible and Embedded Interaction, USA, 2008: 11-14.

(下转 87 页)