

探讨基于 Tsallis 熵的视频关键帧提取技术

杨振

(天津城市职业学院, 天津, 270000)

摘要: 关键帧是视频中的一组有限数量的帧的子集, 一个视频的关键帧序列能够合理地概括该视频信息, 从而减少过大的视频数据对生产生活带来的承载负重。本文讨论了基于 Tsallis 熵的 Jensen 距离公式——JTD 在视频关键帧提取中的应用。根据得到的差异性距离值 JTD, 首先检查子镜头边界, 进而从每个子镜头中抽取一帧作为该镜头的代表帧, 最终得到该段视频的关键帧序列。

关键字: 关键帧提取; Tsallis 熵; Jensen 不等式

The key frame extraction technology based on Tsallis entropy is discussed

Yang Zhen

(Tianjin city Career Academy, Tianjin, 270000)

Abstract: Key frame is a subset of a limited number of video frames and key frames of a video sequence can reasonably generalize the video information, thereby minimizing weight bearing on production and life bring large video data. This paper discusses Jensen distance formula based on Tsallis entropy of JTD application in video key frame extraction. According to the difference of the distance value JTD, first check the sub shot boundary, and then extracted from each sub shot frame as a representative of the lens frame, finally get the key frames of the video sequence.

Key words: key frame extraction; Tsallis entropy; Jensen inequality

0 引言

20 世纪 90 年代以来, 人类逐步步入了信息社会的时代。随着人们日常娱乐活动的增多, 数字视频得到人们的广泛应用。视频关键帧的提取, 使得一个视频得到有效地总结, 从而大量降低了视频存储的数据量。本文提出了利用 Tsallis 熵和 Jensen 距离的视频关键帧提取算法, 对视频关键帧提取技术的研究起到积极的推动作用。

1 基于 Tsallis 熵的关键帧提取理论基础

1.1 Tsallis 熵

Harvda 和 Charvat^[1] 给出了 Tsallis 信息熵的广义定义。随机变量 X 的 Tsallis 熵定义为:

$$T_{\alpha}(p) = \frac{1}{\alpha-1} \left(1 - \sum_{i=1}^n p_i^{\alpha} \right) \quad (\text{公式 1})$$

其中, $\alpha > 1$ 称为 Tsallis 熵指数。

1.2 Jensen 距离

对于一个在区间 $[a, b]$ 上的 f 是一个凸函数, 并且 x_1, x_2, \dots, x_n 在区间 $[a, b]$ 中, Jensen 距离(JD)^[2]:

$$JD(p_1, p_2, \dots, p_n) = f\left(\frac{1}{n} \sum_{i=1}^n p_i\right) - \frac{1}{n} \sum_{i=1}^n f(p_i) \quad (\text{公式 2})$$

成立, 并且 $0 \leq w_i \leq 1, \sum_{i=1}^n w_i = 1$ 。

1.3 基于 Tsallis 熵的 Jensen 距离

Tsallis 熵在其指数范围内, 具有凸函数的性质。对于具有

RGB 直方图分布差异的两个相邻帧 f_{i-1} 和 f_i , 它们之间的基于 Tsallis 熵的 Jensen 距离可以表示为:

$$JTD(f_{i-1}, f_i) = T\left(\frac{p_{f_{i-1}} + p_{f_i}}{2}\right) - \frac{T(p_{f_{i-1}}) + T(p_{f_i})}{2} \quad (\text{公式 3})$$

2 关键帧提取的原理过程

视频序列可以看成是一种分级模式: 视频序列 \rightarrow 子镜头 \rightarrow 帧^[3]。一般来说, 基于镜头的视频关键帧提取算法是对一小部分的视频帧作处理^[4]。

2.1 子镜头的检测与视频序列的分割

根据公式(3)可得到具有平滑效果的窗口均值化的 JTD 公式:

$$JTD_w(f_{i-1}, f_i) = \frac{1}{n_w} \sum_{j=i-\frac{n_w}{2}}^{j=i+\frac{n_w}{2}} JTD(f_{j-1}, f_j) \quad (\text{公式 4})$$

其中, n_w 是所做的窗口的大小。同时定义 $JTD(f_{i-1}, f_i)$ 与 $JTD_w(f_{i-1}, f_i)$ 的比值:

$$\beta(f_{i-1}, f_i) = \frac{JTD(f_{i-1}, f_i)}{JTD_w(f_{i-1}, f_i)} \quad (\text{公式 5})$$

2.2 子镜头的评估标准与分割

当一个视频中有较大的内容变化时, 需要在这个视频序列中分割子镜头, 这样可以提取出更加完全展示出这个视频序列的内容信息。此处定义帧 f_i 处的 JTD 梯度:

$$\delta(f_i) = JTD_w(f_i, f_{i+1}) - JTD_w(f_{i-1}, f_i) \quad (\text{公式 6})$$

(下转第 74 页)

2 控制律的设计与仿真

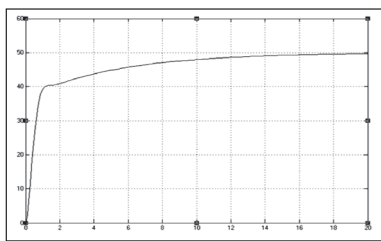


图1 PID控制下高度阶跃响应(A点)

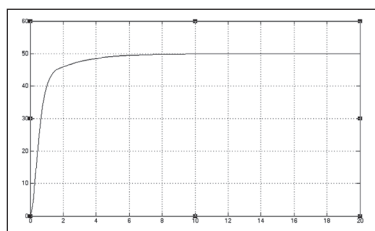


图2 PID控制下高度阶跃响应(B点)

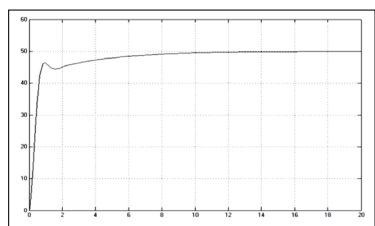


图3 PID控制下高度阶跃响应(C点)

在设计基于PID控制的无人机高度控制系统时,我们通常只需保持原来所设计的俯仰姿态回路不变,然后在此基础上设计基

于PID控制的高度保持/控制外回路就可以了。

如前文所述,我们仍以中空和高空域中的状态点A和C作为基准状态点,分别设计了基于常规PID和智能PID的高度控制器。然后将所设计好的控制器分别用于A点附近的B点和C点附近的D点。全部仿真结果分别如图1~4所示。

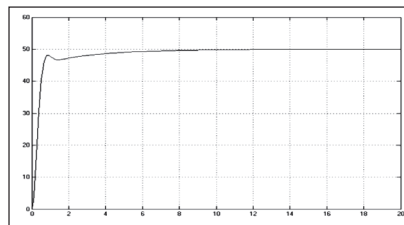


图4 PID控制下高度阶跃响应(D点)

3 结束语

本文以某型固定翼无人机为研究对象,主要研究了其飞控系统控制律的设计问题,分别设计了基于常规PID控制和智能PID控制策略的飞行控制律,并进行了大量的仿研究。仿真结果表明:基于经典PID控制律结构简单,容易实现,但这种传统的设计技术需要反复选择大量的设计参数,需要借助于大量的直观的经验,按照闭环回路来依次选择控制系统的结构和参数。

参考文献

- [1] 许陈元,李春涛. 无人机快速着陆控制律设计及仿真验证[J]. 计算机仿真, 2016,(07):141-146.
- [2] 曹南,苏媛. 简化模型下的舰载无人机着陆控制律设计与仿真[J]. 飞机设计, 2015,(02):11-14.

(上接第81页)

2.3 视频关键帧的提取

基于我们对一段视频分割得到的镜头和子镜头,根据每一个分割片段所表现出来的内容信息变化量的大小,选取其中的一帧作为关键帧。对于在视觉内容上较为平滑的子镜头,可以选取其中的任何一帧作为该子镜头的关键帧;对于在视觉上含有较大的内容信息变化量的子镜头,则选取与其周围帧的JTD之和达到最小值的帧作为该子镜头中的关键帧。

3 实验结果与分析

本文采用的测试视频来自于网站“The Open Video Project”^[5]。测试视频“BOR14_001”是一段由多种视频类型组合而成的视频。该视频含有1083帧,36秒时长。图1(a)(b)分别给出了Downsampling和JTD方法提取出来的关键帧序列。图中可以看到,使用JTD提取出的关键帧较完整和无冗余地表述了该段视频的每一个场景内容。

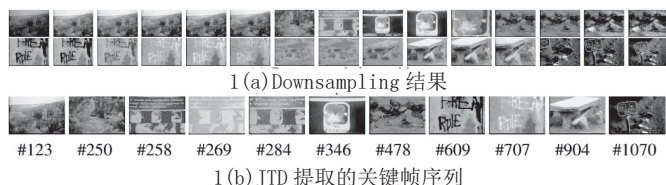


图1 对测试视频“BOR14_001”的关键帧提取实验

4 总结

关键帧的提取在视频信息检索技术中占有很重要的地位。本文讨论了基于Tsallis熵的Jensen距离公式——JTD在视频关键帧提取中的应用。实验结果显示,JTD关键帧提取技术可以根据视频内容选取较为完整和无冗余的视频关键帧序列,可以满足基于内容的视频检索系统的需求,这为当下存储代价较高的信息社会提供了许多便利。

参考文献

- [1] Harvda, J., Charv'at, Quantification method of classification processes, Concept of structural a-entropy, 1967, 30~35.
- [2] T. M. Cover, J. A. Thomas, Elements of Information Theory, 2nd Ed., San Francisco: Wiley-Interscience, 2006, 10~18.
- [3] A. Hanjalic, Shot-boundary detection: Unraveled and resolved?, IEEE Trans. Circuits Syst. Video Technol, 2010, 90~105.
- [4] B. T. Truong, S. Venkatesh, Video abstraction: A systematic review and classification, ACM T. Multim. Comput., 2007, 1~37.
- [5] <http://www.open-video.org/index.php>.