

基于弹幕的视频高潮探测分析

高旭

(同济大学 软件学院, 上海 201804)

摘要:在当今快节奏的生活中,大多数人只会用零散的时间来浏览视频。因此,他们倾向于看一些自己感兴趣的高潮剪辑。同时,视频中的弹幕数据越来越庞大,并且在一定程度上,每个时间戳下的弹幕数量代表了该时刻的视频受观众喜爱的程度以及视频是否为高潮的相关性程度。分析这些弹幕数据可以捕获视频的高潮区间。因此,以此分析视频并对其进行切片可以区分视频的高潮部分和低潮部分。针对视频高潮探测这种无标签数据的分析场景,基于弹幕数据,使用弹幕数量这一数据特征并设定移动阈值进行分析获得了比传统聚类方法更好的实验结果。

关键词:视频高潮;弹幕数量;移动阈值;聚类

中图分类号:TP311 **文献标识码:**A

文章编号:1009-3044(2020)06-0209-02



开放科学(资源服务)标识码(OSID):

1 概述

随着人们生活节奏的加快,他们会更倾向于在一些闲暇的片段时间浏览视频的精彩片段。这能很大程度上减少观看视频的时间成本,同时也能大致了解视频的主要内容。该文基于弹幕数量这一个能够在某种程度上反映视频片段是否为高潮的数据特征进行了分析研究,采用该方法只需要很少的计算资源和时间需求便能够获得相较于传统的聚类方法更好的实验效果。

2 背景

目前,也有一些基于文本数据的分析研究工作。比如一种探测视频片段主题的研究方法^[1]。利用弹幕来区分视频的边界^[2]。根据体育赛事中的评论数量来探测比赛的高潮时段^[3]。目前,大部分文本类的研究工作是关于微博以及商品评论等领域,而对视频中的弹幕文本的分析较少。

本论文考虑使用LDA或者word2vec先将文本数据转换成词向量再进行聚类。最终的实验结果显示采用弹幕数量这一特征进行分析比聚类获得了更好的实验效果。

3 数据预处理

此论文使用的是从哔哩哔哩网站爬取的两部电影的弹幕数据,分别是唐伯虎点秋香和楚门的世界。这两部电影也分别代表了两种不同风格的电影。爬取的弹幕数据的格式如下表1所示。其中唐伯虎点秋香一共有大约24000条弹幕,楚门的世界有大约11000条弹幕。

表1 弹幕数据格式

时间戳	用户ID	弹幕ID	弹幕
55.9	D53205e5	24016808	音乐好听
348.1	c65ced76	23931064	经典要来了
5084.6	a92e9036	23945747	秋香真美

由于弹幕中包含许多网络用语,比如“233”“+1”等。它们分别是大笑和赞同的意思,因此我们需要用中文去替换他们。

4 高潮探测分析

4.1 基于弹幕数量的高潮探测

该方法通过每个视频片段里的弹幕数量这一特征并构建移动阈值来进行高潮探测。构建移动阈值是因为之前的探测方法通常会定义一个固定的阈值来探测高潮,因此阈值的制定在一定程度上决定了探测的效果。另外,一种阈值在一种数据集上效果良好,但不一定在另一种数据集上依然表现良好^[8]。因此,我们采用构建移动阈值这一灵活的方法来解决高潮探测这一问题。该方法用到的一些参数如下所示:

每个片段的弹幕数量用 N_i 表示,在第 i 个片段下的移动阈值用 T_i 表示, T_i 由公式(1)计算得出。

$$T_i = \alpha * (\text{mean}_i + x * \text{std}_i) \quad (1)$$

其中, mean_i 表示从第一个片段到第 i 个片段的弹幕数量均值。 std_i 表示从第一个片段到第 i 个片段的弹幕数量标准差。 mean_i 和 std_i 由公式(2)和公式(3)计算得出。 α 和 x 是超参数。

$$\text{mean}_i = \frac{N_{f_1} + N_{f_2} + \dots + N_{f_i}}{i} \quad (2)$$

$$\text{std}_i = \sqrt{\frac{1}{i-1} \sum_{n=1}^i (N_{f_n} - \text{mean}_i)^2} \quad (3)$$

确定某一个片段 i 是否为高潮片段的算法过程如下所示:

- (1) 对每个片段 i 计算它的 mean_i , std_i 和 T_i 。
- (2) 如果 $N_i > T_i$,判断该片段为高潮。
- (3) 如果 $N_i \leq T_i$,判断该片段为非高潮。

下图是对利用弹幕数量并构建移动阈值来探测视频高潮这一方法的图形解释。

收稿日期:2019-11-08

作者简介:高旭(1994—),男,江苏扬州人,硕士,学生,主要研究方向为数据挖掘。

本栏目责任编辑:唐一东

万方数据

多媒体技术及其应用

209

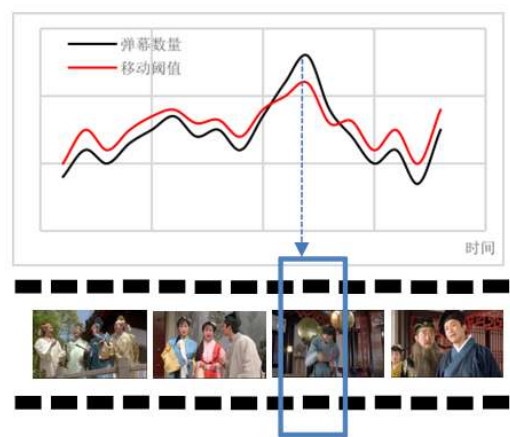


图 1 高潮探测示意图

4.2 基于聚类算法的高潮探测

除了以上的研究方法,目前主流的机器学习方法也能处理这类无标签的数据,例如一些常规的聚类算法K-means,DN-BSCAN和AHC。这些方法的实验结果将作为比较对象与我们的方法进行对比。

5 实验评估

我们选择了唐伯虎点秋香和楚门的世界这两部电影的弹幕作为我们的数据集,下面的表2,表3展示了收集的两部影片的实际高潮片段。其中片段范围是指该播放时间内所属的之前所划分好的片段区间。表里的片段范围将用来评价我们实验结果的好坏。

表 2 唐伯虎点秋香的实际高潮片段

序号	播放时间	片段范围
1	12:40 ~ 14:20	76 ~ 83
2	22:30 ~ 24:30	135 ~ 144
3	28:20 ~ 29:15	170 ~ 173
4	30:00 ~ 33:30	180 ~ 198
5	34:00 ~ 35:15	204 ~ 209
6	37:40 ~ 38:50	226 ~ 230
7	39:30 ~ 41:35	237 ~ 247
8	48:05 ~ 50:50	288 ~ 302
9	54:10 ~ 56:35	325 ~ 337
10	68:25 ~ 70:40	410 ~ 421
11	81:00 ~ 83:30	486 ~ 498

表 3 楚门的世界的实际高潮片段

序号	播放时间	片段范围
1	12:40 ~ 13:20	76 ~ 77
2	32:50 ~ 24:30	197 ~ 204
3	40:15 ~ 42:30	241 ~ 252
4	44:10 ~ 45:20	265 ~ 269
5	52:00 ~ 52:55	312 ~ 315
6	53:40 ~ 55:20	322 ~ 329
7	58:40 ~ 60:40	352 ~ 361
8	78:20 ~ 80:40	470 ~ 481
9	81:30 ~ 83:05	489 ~ 496
10	84:00 ~ 87:30	504 ~ 522
11	88:10 ~ 92:10	529 ~ 550
12	92:10 ~ 94:50	553 ~ 566
13	94:50 ~ 95:20	569 ~ 569

6 总结

整体而言,基于弹幕数量对电影视频进行高潮探测的实验方法,能够较为简便快速地达到找到高潮片段这一研究目标,它相较于常规的聚类算法有更好的实验结果,对于简便快捷地找到长视频中的高潮片段提供了一种研究思路。

参考文献:

[1] Wu B, Zhong E, Tan B, et al. Crowdsourced time-sync video tagging using temporal and personalized topic modeling[C]. Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014: 721-730.

[2] Xian Y, Li J, Zhang C, et al. Video Highlight Shot Extraction with Time-Sync Comment[C]. Proceedings of the 7th International Workshop on Hot Topics in Planetscale Mobile Computing and online Social Networking. ACM, 2015:31-36.

[3] Liang-Chi Hsieh, Ching-Wei Lee, Tzu-Hsuan Chiu. Live Semantic Sport Highlight Detection Based on Analyzing Tweets of Tweet[C].

【通联编辑:代影】