**视频突变与渐变镜头检测实践**

院系：国家网络安全学院

作者：2019300003041 李潇楠

指导教师：蔡波

摘 要

本次技术报告的目的是通过经典的颜色直方图等算法，对视频进行突变和渐变镜头的检测。

关键词：视频 颜色直方图

目录

[1 研究背景 2](#_Toc11378)

[2 实验数据 3](#_Toc1946)

[3 突变镜头检测 4](#_Toc17160)

[3.1 基于颜色直方图检测 4](#_Toc3519)

[3.2 基于时空切片检测 6](#_Toc30989)

[4 渐变镜头检测 8](#_Toc10086)

[4.1 基于颜色直方图与滑动窗口检测 8](#_Toc5086)

[4.2 基于颜色直方图的方差 9](#_Toc1890)

[4.3 基于时空切片和滑动窗口检测 10](#_Toc16109)

[5 参考资料 12](#_Toc17529)

**1 研究背景**

随着多媒体技术的应用越来越广泛，基于内容的视频检索技术被提出。镜头检测技术是低层视频结构化分析中的一项关键技术，好的镜头边界检测技术一定能为视频结构化分析打下坚固的基础，在视频检索中起着重要的作用；镜头检测使更高层的语义视频处理成为可能，镜头监测意味着为视频的结构化表示提供基础。在镜头转换时，视频数据会发生一系列的变化，镜头边界检测方法，目前已经有了许多研究。而镜头边界可以分为突变镜头和渐变镜头，根据各自不同的特征需要使用不同的检测方式。

**2 实验数据**

实验数据是从视频网站下载和截取的、时长约为一分半的视频，并提取出2199帧。

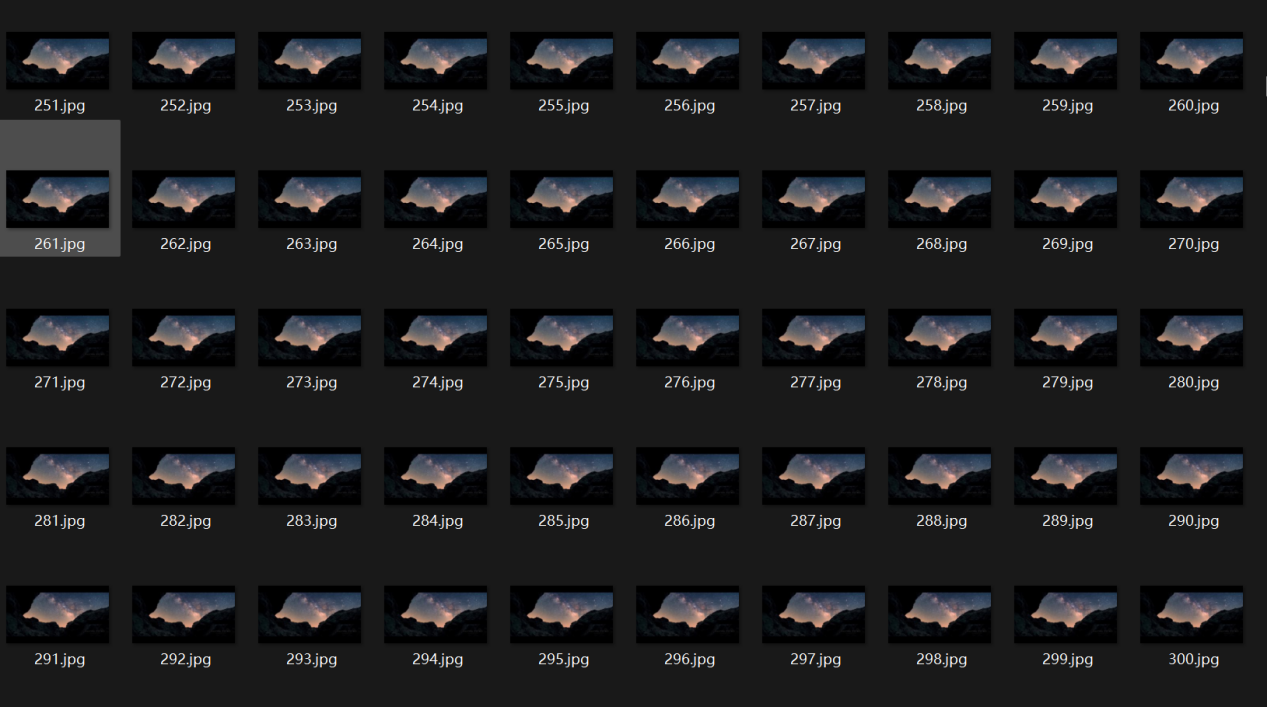


图 1 实验数据

经过人工每10帧检查，发现该视频有6个突变镜头和3个渐变镜头：

突变 360-370 550-560 720-730 920-930 1050-1060 1860-1870；

渐变 30-40附近 1380-1390附近 1570-1580附近。

**3 突变镜头检测**

**3.1 基于颜色直方图检测**

颜色直方图的计算同实验2，先将图像的RGB值转换为HSV，再进行量化。

之后构造一维特征矢量G=9H’+3S’+V’，其中整数G的取值范围为[0,71]，对图像绘制72bin一维直方图。

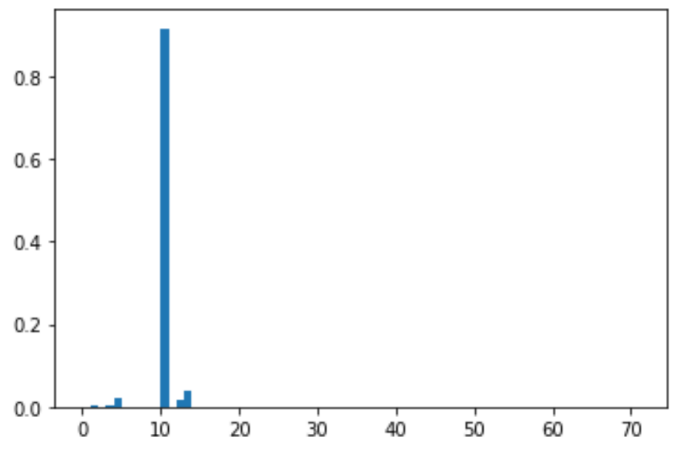


图 2 颜色直方图

对视频每10帧计算直方图，对第10~2190帧的相邻10帧的直方图根据如下公式，求欧几里得距离（论文中为差值，由于直方图已做归一化，这里用距离更科学一些，实际上两种方法描绘出的折线图差别不大）。



式中H和D分别代表第i帧的直方图和距离。将所有距离绘制折线图。

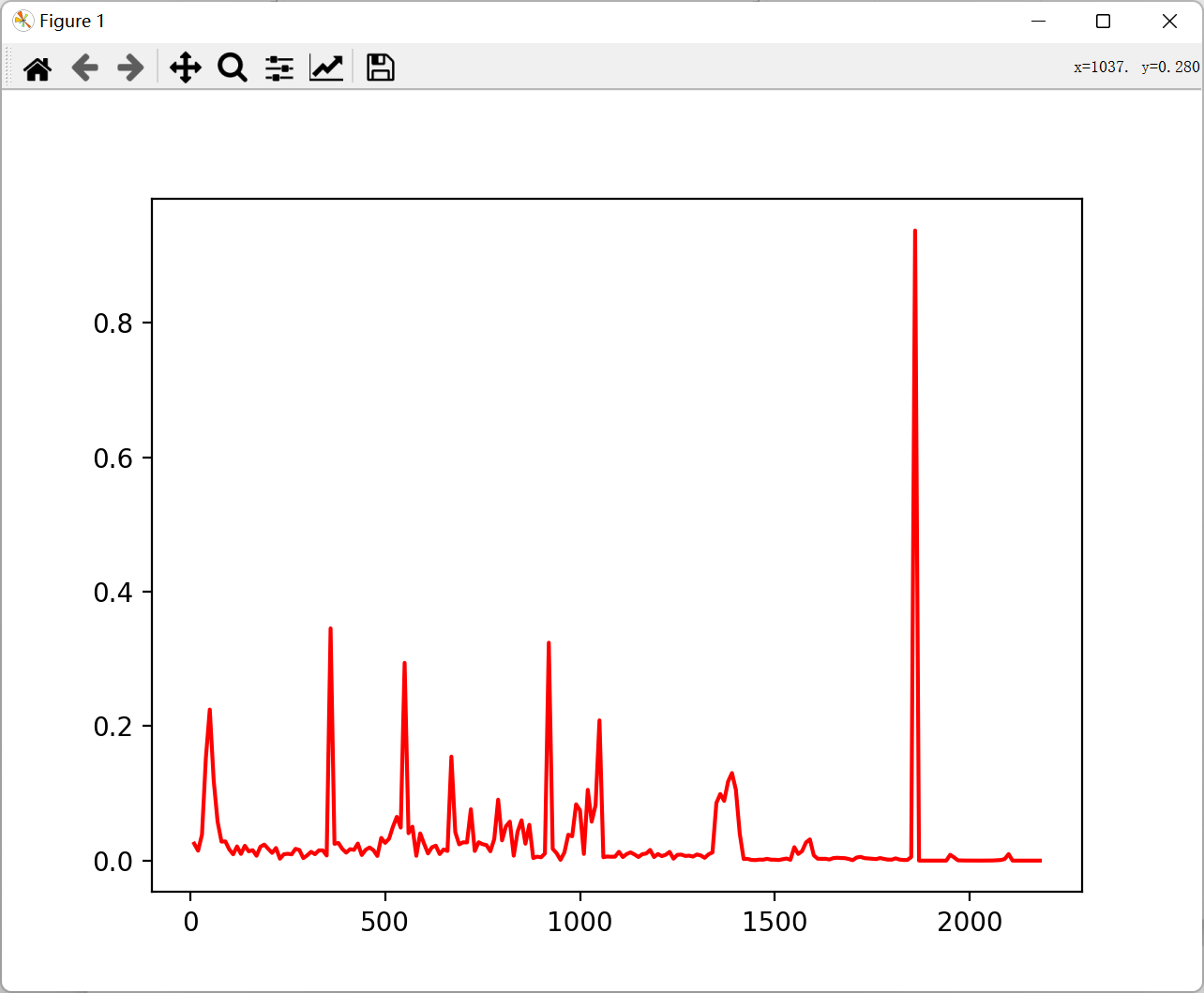


图 3 相邻帧颜色直方图距离折线图

标注蓝色位置为突变镜头，橙色渐变镜头。

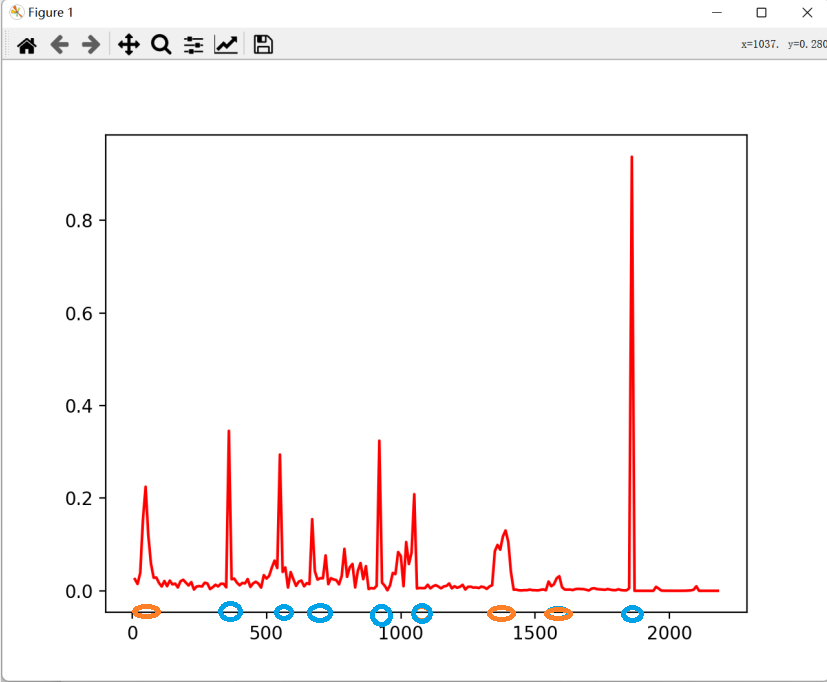


图 4 标注颜色直方图

调整阈值T=0.15，若D>T则视为突变镜头，检测到突变镜头有：40,50,**360,550,670,920,1050,1860**（帧），突变镜头全部检测到了，也有少量渐变镜头混入其中。猜想将帧与帧提取的间隔缩小，这个问题可以解决，但因为要计算直方图，使得计算量增大了许多，而且这个方法本身有缺点，只关注了每帧全部像素的颜色，没有关注帧间像素的位置关系。

**3.2 基于时空切片检测**

所谓切片就是从连续的视频图像序列的同一个位置提取出的一行（列）像素组合而成的一幅二维图像。如果将视频看作是一个(x,y,t)三维图像序列，其中(x,y)为图像维，t为时间维，则视频的时空切片可以看作是由时间维与图像维构成的一幅二维图像。

对实验视频的1/4处做垂直切片，显示它的纹理。

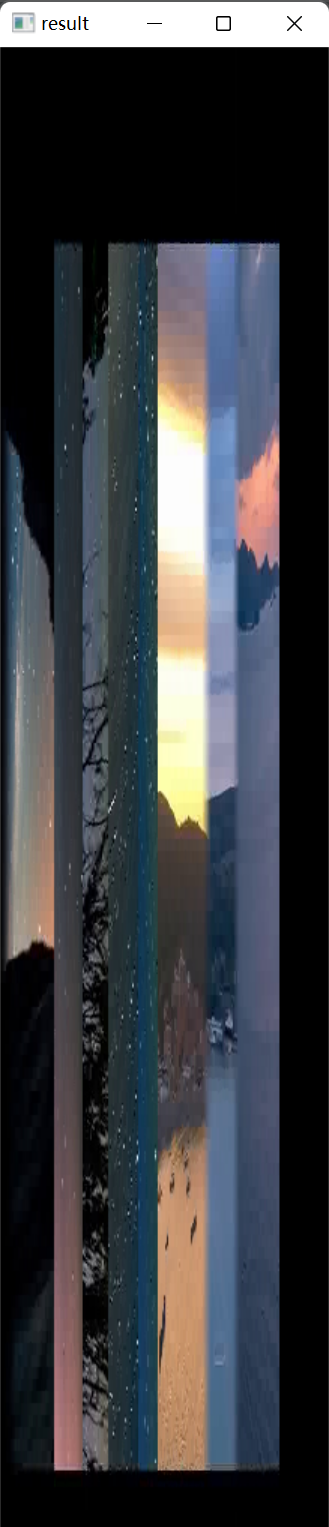


图 5 每10帧拼接

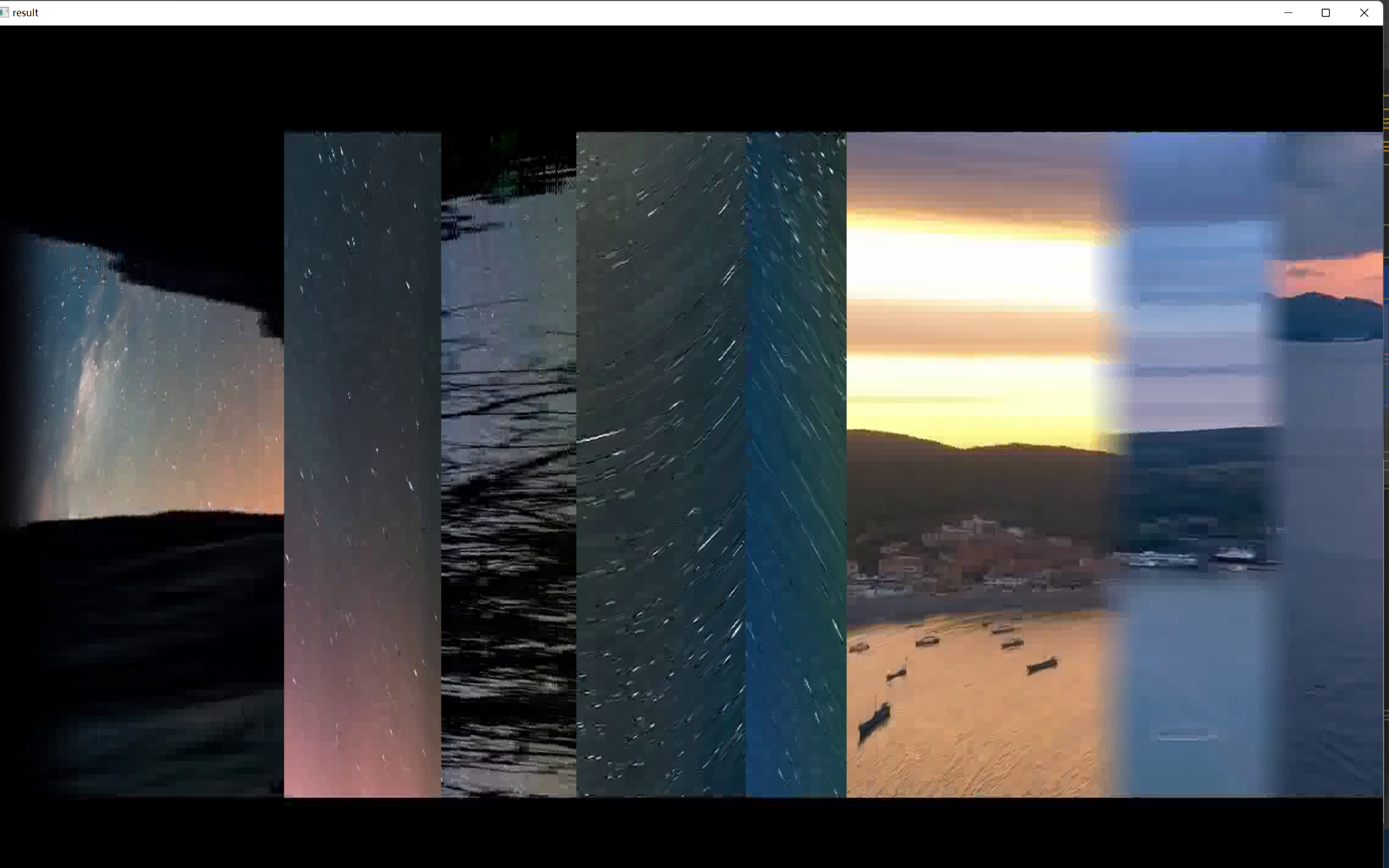


图 6 每1帧拼接

直观上很容易从时空切片的纹理找到突变和渐变镜头。对于突变镜头的处理，可以使用最近邻像素匹配法。

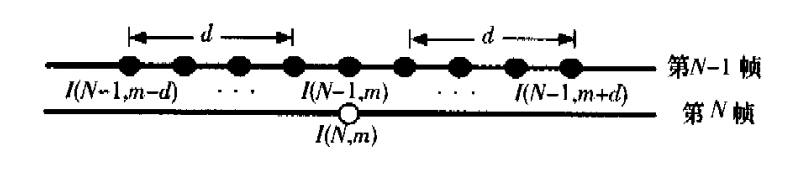


图 7 最近邻像素匹配法

对第N帧目标列的第m个像素，匹配第N-1帧同一列的第m-d到m+d的最多2d+1个像素，如果没有相近的像素，则认为镜头在这个点发生了突变。原论文没有描述怎样算“相近”，这里通过尝试，定义“相近”是RGB三个通道的差值都不超过10。设邻域大小d=10，对视频的每一帧统计突变点个数。

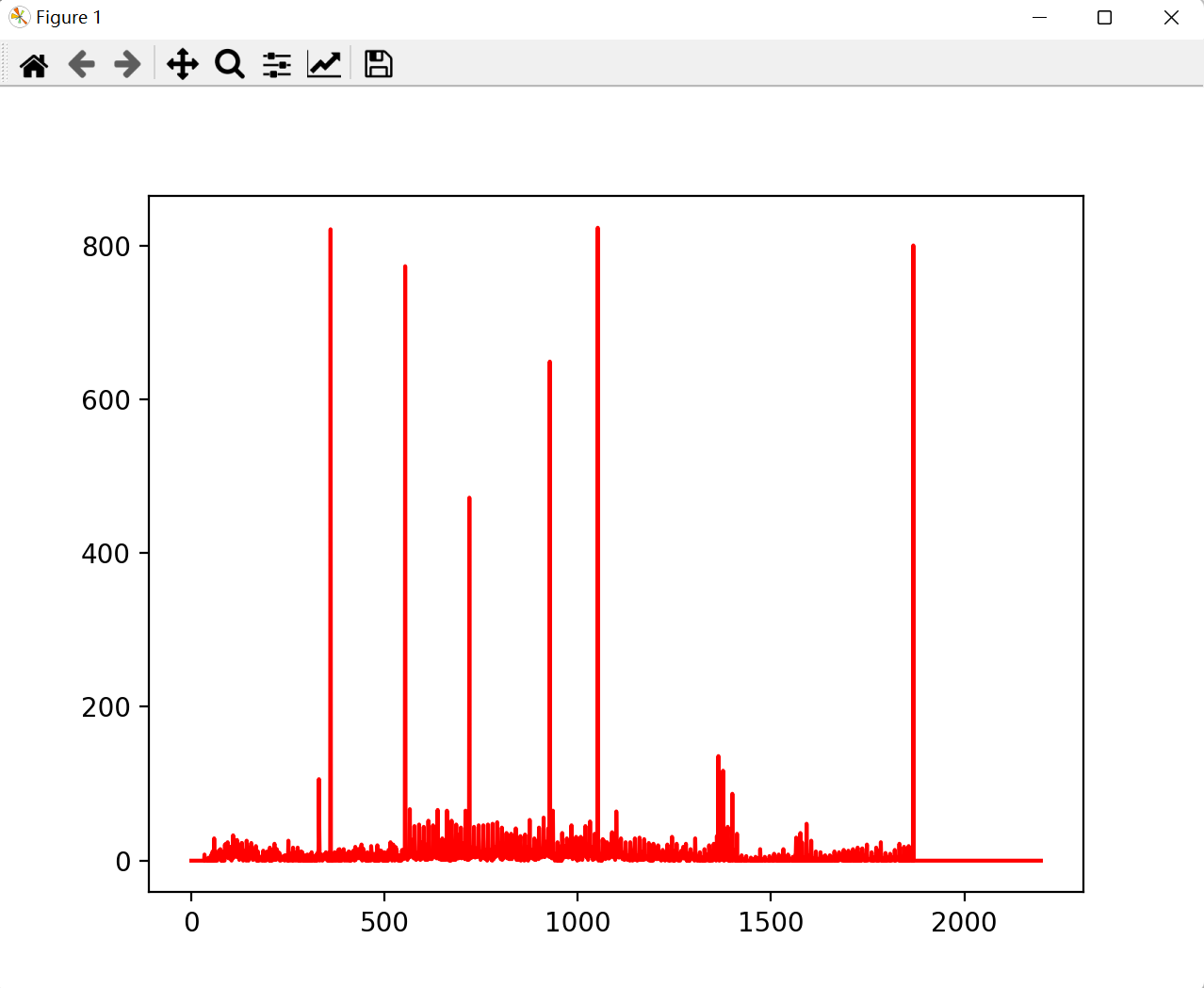


图 8 突变点数量折线图

检测到突变镜头分别位于**360,553,719,927,1051,1867**（帧），完全正确，且计算量比颜色直方图小很多，而且由于考虑了每个像素上一帧相同位置的邻域，不会因为镜头在斜向移动就误判为突变。后续也可以考虑选取垂直和水平的多个行或列来统计，综合结果再判断。

**4 渐变镜头检测**

**4.1 基于颜色直方图与滑动窗口检测**

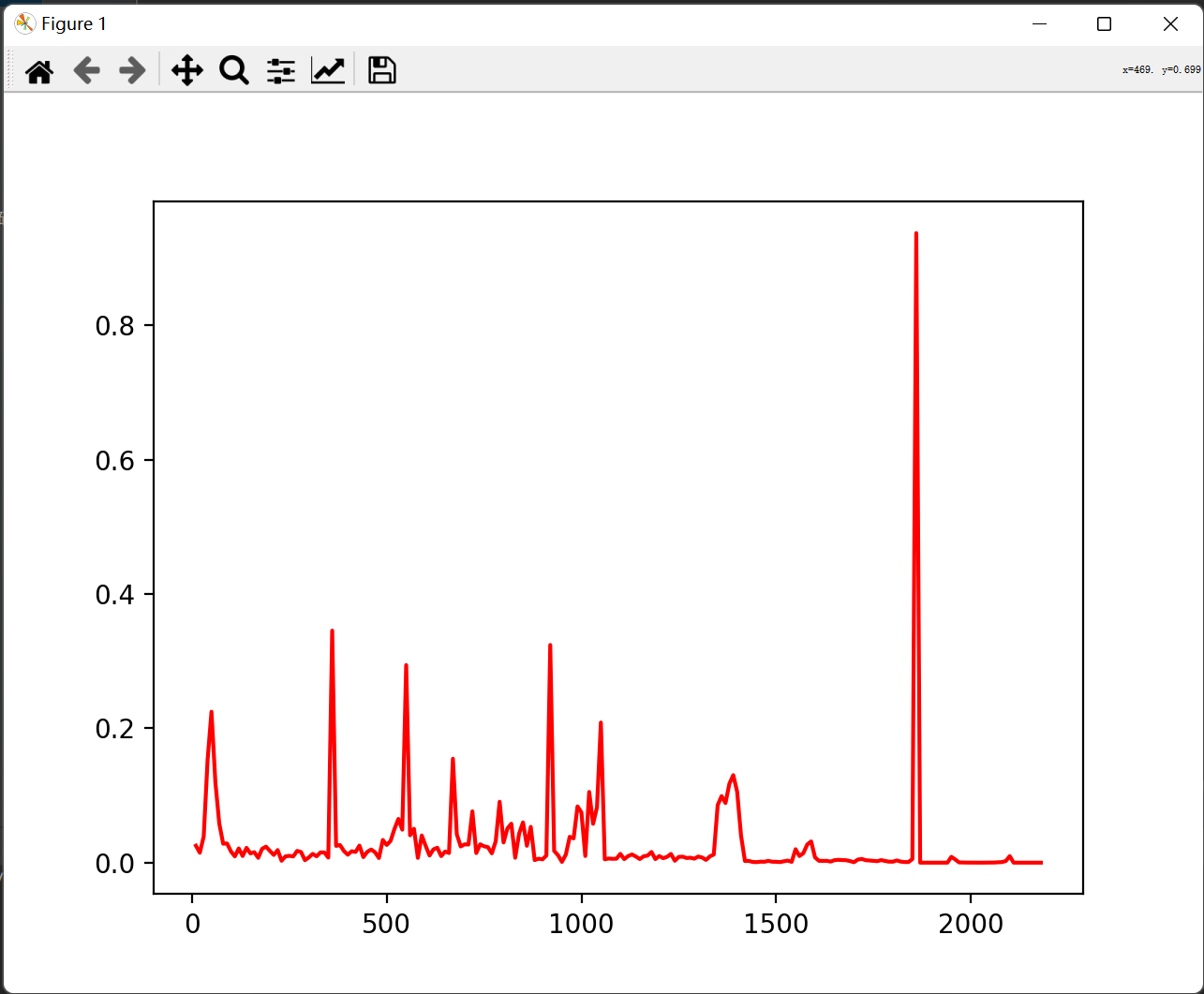


图 9 相邻帧颜色直方图距离折线图

文献[1]描述了通过颜色直方图和滑动窗口寻找渐变镜头的算法，流程如下：

（1）求颜色直方图和帧间距离，同突变镜头；设参数alpha=3，窗口大小为3；

（2）计算第一个窗口内3个距离的平均值，记为avg；

（3）向后滑动一个窗口，计算窗口内的最大值peek；

（4）如果peek<=alpha\*avg，将最后一个差值累计到前面的平均值里，goto（3）；

（5）如果peek>=beta\*avg，将窗口移动到最大值位置的右侧（说明此处是突变镜头），goto（2）

实验结果：

**40**,360,550,670,790,920,970,1050,**1350**,1860,1880,1920,2100（帧），算法很明显受到了突变窗口的影响，更多的关注是否出现峰值而没有把三角形曲线检测出来。

**4.2 基于颜色直方图的方差**

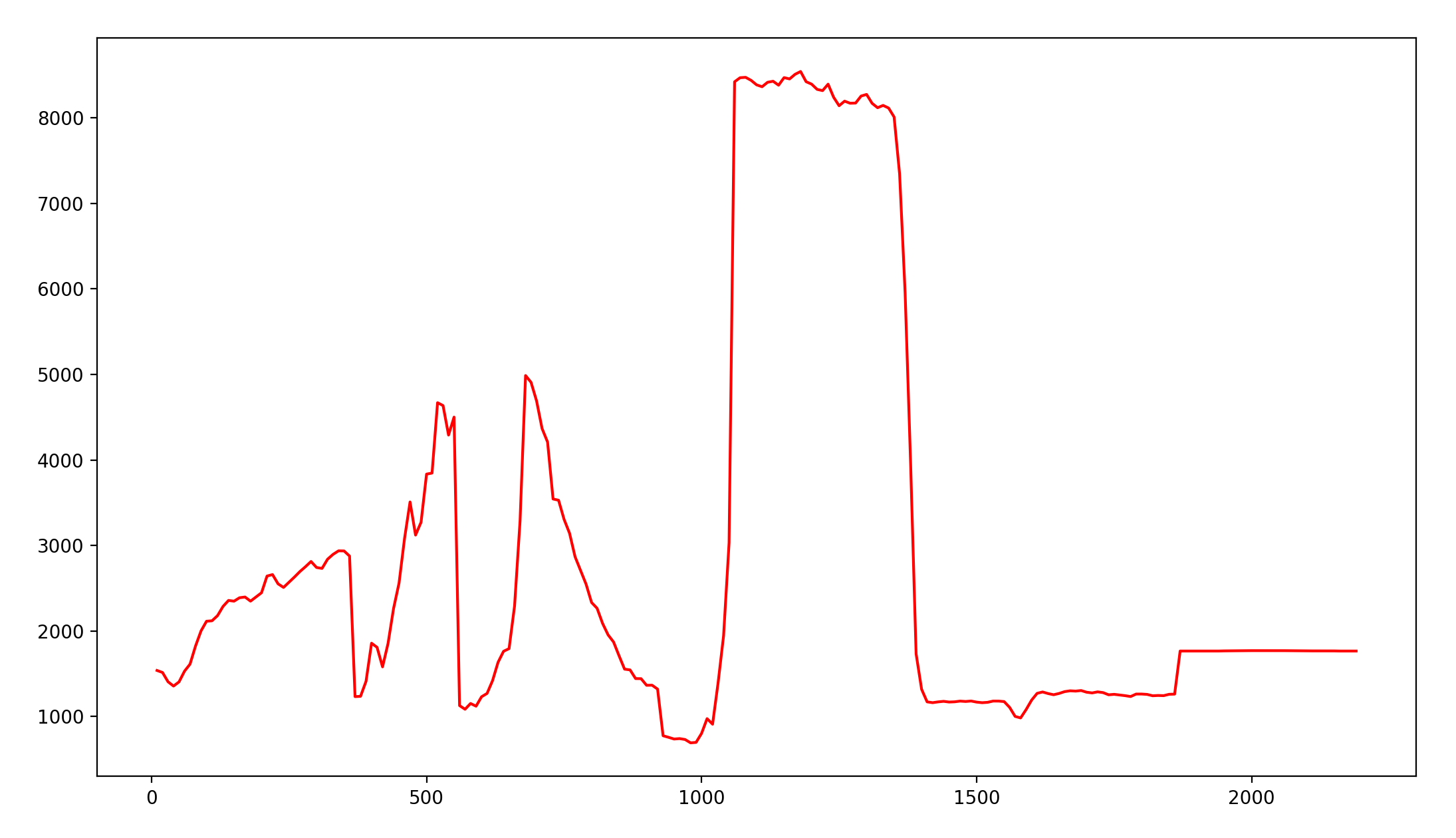


图 10 颜色直方图方差折线图

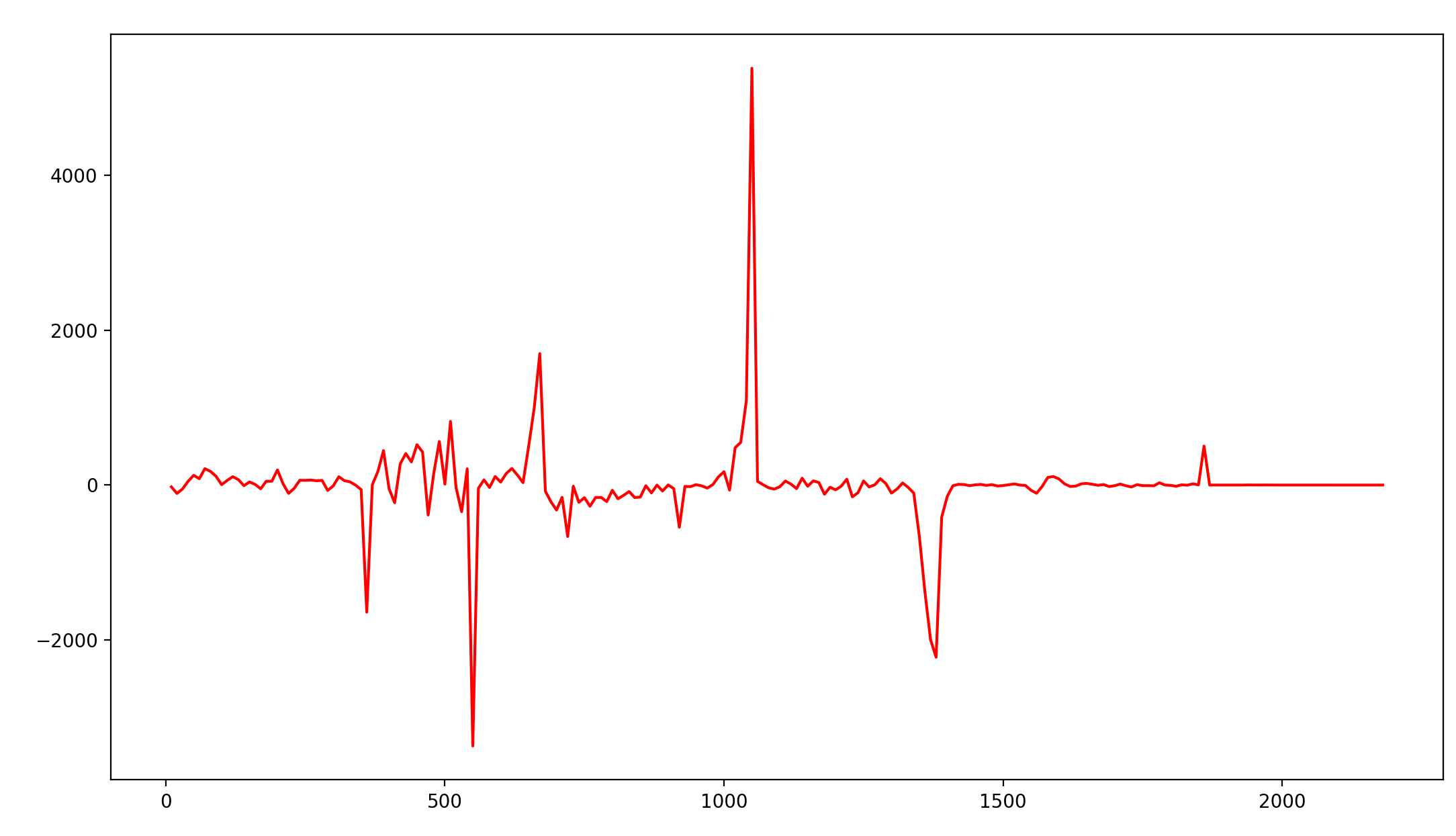


图 11 方差的差值折线图

文献[5]中描述了一种方法，通过颜色直方图的方差和相邻帧方差间的差值来寻找渐变镜头，尝试复现并画出方差和差分的折线图，但实际效果并不好，无法区分突变或渐变镜头。

**4.3 基于时空切片和滑动窗口检测**

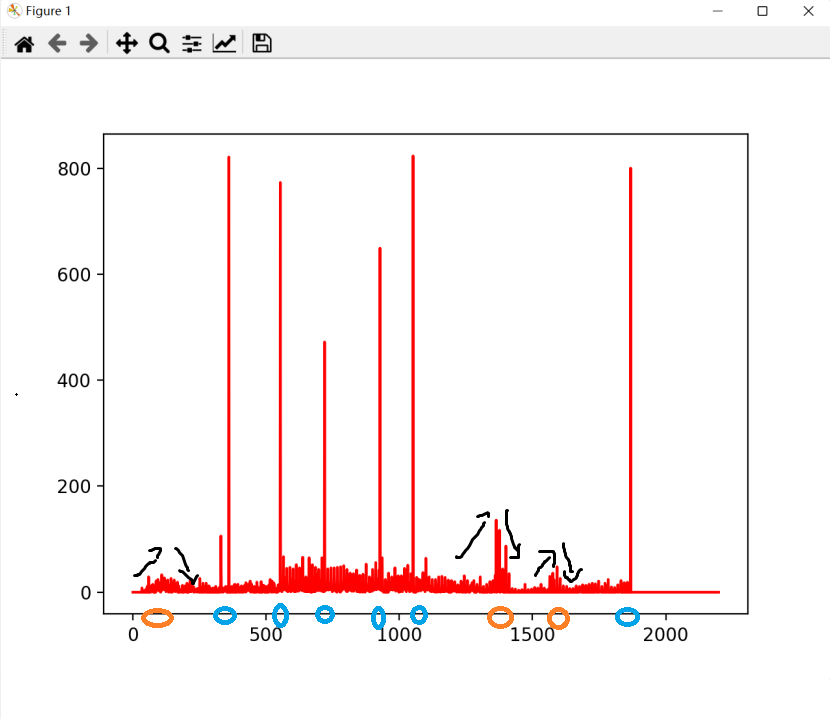


图 12 标注突变、渐变镜头的突变点数量折线图

观察3.2节中时空切片突变点数量的折线图，发现渐变镜头的位置往往表现为先上升后下降，比突变镜头的变化要缓慢；同时，渐变镜头的峰值明显小于突变镜头；因此尝试在时空切片的基础上，改进滑动窗口算法，寻找渐变镜头。

算法流程如下：

（1）对1/4列的位置求时空切片突变点数量，同3.2节；设参数alpha=10，beta=15，窗口大小为60；

（2）计算第一个窗口内60个距离的平均值，记为avg；

（3）向后滑动一个窗口，计算窗口内的最大值peek；

（4）如果peek<=alpha\*avg，将最后一个差值累计到前面的平均值里，goto（3）；

（5）如果peek>=beta\*avg，将窗口移动到最大值位置的右侧（说明此处是突变镜头），goto（2）

（6）如果alpha\*avg<peek<beta\*avg之间，则最大值处为渐变镜头的位置，将窗口移动到最大值位置的右侧，goto（2）

结果为**1375**,1471,1531,**1591**,1603,1783,1831（帧），有一定效果，受噪声干扰比较大。

**5 参考资料**

视频资源：

<https://www.bilibili.com/video/BV1Wt411h7Hx>

[1]周艺华,曹元大,张洪欣.一种通用的渐变镜头检测方法[J].计算机应用研究,2006(2):250-252.

[2]贾伟.基于颜色直方图的视频突变镜头边界检测[J].机械设计与制造工程,2014(5):67-70.

[3]王垚. 视频镜头分割及关键帧提取技术研究[D]. 北方工业大学, 2019.

[4]李勇,刘雨,唐波,等.一种基于时空切片的镜头突变检测方法[J].计算机工程与科学,2006(3):53-54,63.4

[5]杨碧天,王煦法.一种渐变镜头检测方法[J].计算机工程与应用,2005,41(20):37-39+59.