



## 基于动态帧的视频关键帧提取算法研究

李玉峰

(天津科技大学理学院, 天津 300457)

**摘要:** 针对基于内容的视频检索系统,提出了一种关键帧提取算法.为了提高算法的保真度和压缩比,首先构造了动态帧,它的每个像素对应一个像素代表灰度集合,该集合中的元素可以最大限度地代表镜头中相应像素的灰度值,然后根据镜头中每一帧与该镜头动态帧之间的距离来确定关键帧.为了验证本算法的有效性,选取了大量视频与 TMOF 算法及 SKF 算法进行比较.结果表明:该算法具有较高的准确性和可靠性,保真度和压缩比均高于其他两种算法.

**关键词:** 视频检索; 镜头; 关键帧提取; 动态帧

中图分类号: O244; TP391      文献标志码: A      文章编号: 1672-6510(2009)04-0069-04

## Research on Video Key Frame Extraction Algorithm Based on Dynamic Frame

LI Yu-feng

(College of Science, Tianjin University of Science & Technology, Tianjin 300457, China)

**Abstract:** A key frame extraction algorithm was presented according to content-based video retrieval system. At first, a dynamic frame was constructed to improve the fidelity and compression ratio of the algorithm. Each pixel in dynamic frame corresponds to a representative gray set, which means the elements in the set represent the furthest corresponding pixel gray value in the shot. Then the distance between each frame and the dynamic frame was calculated to determine the key frames. The algorithm was tested on comprehensive videos, and compared with both TMOF algorithm and SKF algorithm. Experiments show that the key frame extraction algorithm is more accurate and reliable, and the fidelity and compression ratio are higher than that of the other two algorithms.

**Keywords:** video retrieval; shot; key frame extraction; dynamic frame

视频因为可以表达大量不同层面的信息而被广泛应用,如何对这些海量视频进行管理,使用户能快速、准确地检索到他们需要的视频,是当前信息处理领域的一个研究热点和难点.基于内容的视频检索正是在这一需求下产生的.关键帧提取是基于内容视频检索的重要步骤,关键帧反映了镜头的主要内容,它的使用大大减少了视频索引的数量.

当前,主要有以下几种关键帧提取算法:(1)基于采样的关键帧提取算法(SKF),该算法每隔一段时间从视频中提取一帧作为关键帧<sup>[1]</sup>,其优点是算法实现简单,缺点是视频的采样间隔很难和视频内容分布一

致;(2)基于时域最大值的關鍵帧提取算法(TMOF)<sup>[2]</sup>,该算法提取的关键帧反映了镜头中出现频率较高的内容,而没有反映在镜头中出现时间较短,但有重要意义的对象;(3)基于聚类的关键帧提取算法,该类算法首先对镜头中的所有帧聚类,如果类足够大,则将其作为关键帧类,选择距离类中心最近的一帧作为关键帧<sup>[3-4]</sup>,该算法还有多种改进算法,例如在聚类算法中采用自适应的阈值<sup>[5]</sup>,改进的 C 模糊聚类算法<sup>[6]</sup>,非参数密度估计算法<sup>[7]</sup>等,这类算法最大的缺点是计算量相对较大;(4)基于运动信息的关键帧提取算法,该算法首先通过对光流分析计算运动量,然后将运动

收稿日期: 2009-02-16; 修回日期: 2009-05-13

基金项目: 天津科技大学科学研究基金资助项目(20080204)

作者简介: 李玉峰(1976—),男,山西大同人,讲师, tjulyf@126.edu.cn.

量局部最小值对应的帧作为关键帧<sup>[8]</sup>, 实验中发现局部最小值的选取往往不准确; (5) 基于 MPEG 宏块统计特性的关键帧提取算法, 该算法利用帧中宏块编码方式的不同生成帧差, 通过对帧差的度量提取关键帧<sup>[9]</sup>, 这类算法不需要解压视频, 可以进行实时关键帧提取, 但是与解压域算法相比, 在压缩域可以利用信息不多, 关键帧提取不够准确.

本文提出了基于动态帧的关键帧提取算法, 首先确定动态帧中每个像素的代表灰度集合, 然后通过计算视频中每一帧和动态帧之间的距离来确定关键帧, 并进行了实验研究.

### 1 动态帧及其构造方法

动态帧中某一像素的代表灰度集合记录了最有可能代表该像素的灰度集合, 像素  $(i, j)$  的代表灰度集合  $g(i, j)$  具有如下特征: (1)  $g(i, j)$  中各灰度值能代表镜头中像素  $(i, j)$  处更多的灰度值; (2)  $g(i, j)$  中各相邻灰度值的差别较大.

动态帧中像素  $(i, j)$  处代表灰度集合的构造方法如下:

(1) 计算镜头中每一帧的各像素  $(i, j)$  的灰度值, 第  $n$  帧中像素  $(i, j)$  的灰度值计算如式 (1) 所示.

$$f_n(i, j) = 0.3 \times R_n(i, j) + 0.59 \times G_n(i, j) + 0.11 \times B_n(i, j) \quad (1)$$

其中,  $R_n(i, j)$ ,  $G_n(i, j)$ ,  $B_n(i, j)$  分别为该像素的红、绿、蓝分量. 为了加快计算速度, 同时考虑到一个镜头持续时间一般不会超过 20 min (1 min 不超过 3 000 帧), 所以直方图纵坐标未归一化, 纵坐标为横坐标对应的灰度级在该像素出现的次数, 直方图  $H_b(i, j)$  的计算如式 (2) 所示<sup>[10]</sup>.

$$H_b(i, j) = \sum_{n=0}^{N-1} \delta(f_n(i, j) - b), \quad 0 \leq b \leq B \quad (2)$$

式中:  $N$  是该镜头中帧的总个数;  $B$  是直方图的颜色灰度级个数. 为加快计算速度, 颜色灰度级设为 128 级.  $\delta(\cdot)$  为冲激函数. 图 1 和图 2 分别为图 3 中的视频帧序列在 (243, 71) 像素的灰度直方图和代表灰度集合.

(2) 按照直方图中像素的灰度值进行排序, 如果第  $m$  个灰度值  $g_m(i, j)$  与相邻灰度值差的绝对值  $d_m(i, j) = |g_m(i, j) - g_{m+1}(i, j)|$  很小, 则选取它们中较大的一个灰度值作为代表, 而将另一个灰度值去掉.

最后得到的灰度值可能有多个, 这些灰度值作为像素  $(i, j)$  的代表灰度集合  $g(i, j)$ , 如图 2 所示, 就存

在两个代表灰度值. 通过以上步骤, 可以得出整个镜头在各个像素的代表灰度集合, 这样, 动态帧的每个像素对应于它的代表灰度集合.

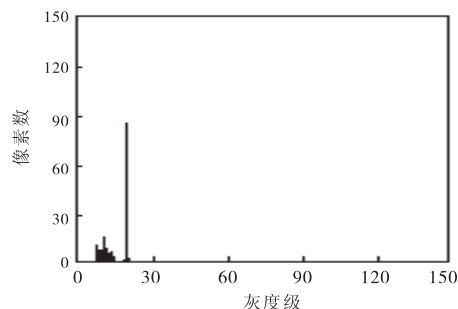


图 1 像素灰度直方图

Fig.1 Gray histogram of a pixel

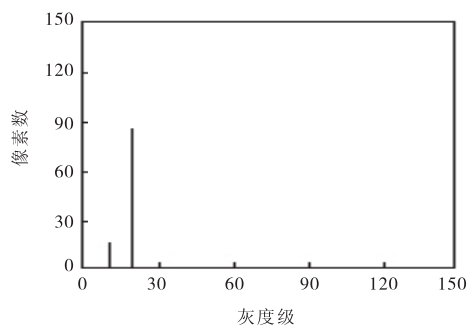


图 2 像素的代表灰度集合

Fig.2 Representative gray set of a pixel

### 2 关键帧提取算法

计算镜头所有帧和动态帧之间的距离, 帧间距离的定义如式 (3) 所示:

$$D(f_i, f_k) = \sum_{x=0}^{W-1} \sum_{y=0}^{H-1} \min |f_i(x, y) - f_k(x, y)| \quad (3)$$

式中:  $f_i$  为镜头的第  $i$  帧;  $f_k$  为镜头对应的动态帧;  $W$  为帧的宽度;  $H$  为帧的高度. 由此得到整个镜头与动态帧之间的距离曲线  $D(f_i, f_k)$ , 曲线上的最大值表明了所对应的帧与动态帧的差别较大, 这是由于出现了新的对象引起的, 最小值通常代表了视频的背景, 所以取曲线上的极值点所对应的帧作为关键帧.

镜头的关键帧提取算法如下:

- (1) 读入镜头当前帧, 根据式 (1) 计算每个像素的灰度值, 同时根据式 (2) 修改各像素在灰度直方图中对应的值, 并将修改后的值存储到三维数组中;
- (2) 判断当前帧是否是镜头的最后一帧, 如果不是, 则将下一帧设为当前帧, 转向 (1), 否则转向 (3);
- (3) 根据上节介绍, 计算该镜头的动态帧;
- (4) 根据式 (3) 计算关键帧;
- (5) 算法结束.

### 3 实验结果与分析

根据文献[10]和文献[11]给出的保真度和压缩比的定义,镜头中帧  $f_i$  到该镜头关键帧集合  $f_k$  的距离可以用式(4)计算:

$$d(f_i, f_k) = \min\{D(f_i, f_k(j)) | j = 1, 2, \dots, N\} \quad (4)$$

式中:  $N$  为镜头的关键帧个数. 镜头  $S_i$  的所有帧到关键帧集合的距离最大值为

$$D(S_i, f_k) = \max\{d(f_i, f_k) | i = 0, 1, \dots, n\} \quad (5)$$

将距离最大值  $D(S_i, f_k)$  归一化为  $d(S_i, f_k)$ , 保真度由式(6)计算<sup>[10-11]</sup>.

$$Fidelity(S_i, f_k) = 1 - d(S_i, f_k) \quad (6)$$

通常情况下,一段视频包括多个镜头,例如在普通的故事片中,大约有 600~1 200 个镜头<sup>[12]</sup>,一段视频上的关键帧应该反映该段视频的主要内容,由式(6)可以计算出每个镜头的保真度,整段视频的保真度为该视频所包含的所有镜头的视频保真度的平均值<sup>[10-11]</sup>,如式(7)所示.

$$Fidelity = \frac{1}{N_s} \sum_{i=1}^{N_s} Fidelity(S_i, f_k(i)) \quad (7)$$

式中:  $N_s$  为整段视频包含的镜头个数;  $S_i$  为第  $i$  个镜头;  $f_k(i)$  为从第  $i$  个镜头提取的关键帧集合.

对同样的保真度,整段视频的压缩比越低,说明关键帧提取效果越好<sup>[10-11]</sup>,整段视频的压缩比定义如式(8)所示.

$$CRatio = 1 - \frac{N_k}{N_T} \quad (8)$$

式中:  $N_k$  为从整段视频提取的关键帧数;  $N_T$  为整段视频包含的总帧数. 实验中选择了多个 MPEG-1 格式的视频片段,表 1 分别为对视频 1 新闻片段和视频 2 MTV《国乐》片段的实验结果. 这两个视频片段的内容差别较大,新闻片段中镜头内容变化相对简单,每个镜头持续时间也较短;MTV 片段中相机和物体运动较多,镜头持续时间相对较长. 实验中选取的部分视频序列如图 3 所示,提取的对应关键帧如图 4 所示.

表 1 实验结果比较

Tab.1 Experimental result comparison

算法	视频 1		视频 2	
	保真度	压缩比	保真度	压缩比
TMOF	0.870	0.992	0.820	0.991
SKF	0.780	0.997	0.620	0.993
本算法	0.912	0.997	0.876	0.994

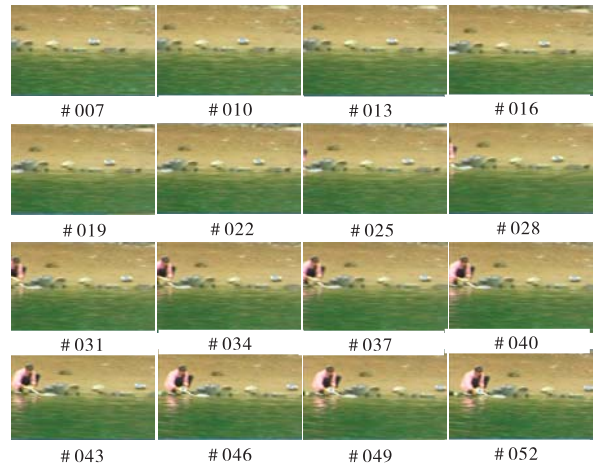


图 3 视频帧序列

Fig.3 Video frame sequence



图 4 视频帧序列对应的关键帧

Fig.4 Key frame of video frame sequence

由表 1 可以看出,SKF 算法的保真度比 TMOF 算法和本算法低,TMOF 算法在视频 1 得到的保真度较高,对于视频 2,由于同一个镜头内容变化较大,特别是视频中包含很多运动对象,所得的保真度较低,本算法所得到的保真度和压缩比均高于 TMOF 算法和 SKF 算法.

### 4 结论

本文提出了一种基于动态帧的视频关键帧提取算法,由于算法中各像素的代表灰度集合进行了相应的处理,提取结果与人的视觉具有良好的一致性,且算法的保真度和压缩比较高,可以满足基于内容视频检索系统的需要.当然,实验中发现该算法还存在由于大物体的快速运动而产生冗余帧问题,今后的实验将在算法中融入运动补偿,以提高算法的有效性.

### 参考文献:

[1] Hammound R, Mohr R. A probabilistic framework of selecting effective key frames from video browsing and indexing [C]//Proceedings of International Workshop on Real-Time Image Sequence Analysis. Oulu, Finland: IEEE Computer Society, 2000: 79-88.

[2] Kin W S, Kin M L, Guoping Q. A new key frame representation for video segment retrieval[J]. IEEE Trans on

Circuits and Systems for Video Technology, 2005, 15(9):1148-1155.

[3] Zhuang Y, Rui Y, Huang T S. Adaptive key frame extraction using unsupervised clustering[C]// Proceedings of International Conference on Image Processing. Washington D C, USA: IEEE Computer Society, 1998:866-870.

[4] Ferman A M, Tekalp A M. Two-stage hierarchical video summary extraction to match low-level user browsing preferences[J]. IEEE Trans on Multimedia, 2003, 5(2): 244-256.

[5] 王方石, 须德, 吴伟鑫. 基于自适应阈值的自动提取关键帧的聚类算法[J]. 计算机研究与发展, 2005, 42(10):1752-1758.

[6] 张婵, 高新波, 姬红兵. 视频关键帧提取的可能性 C 模式聚类算法[J]. 计算机辅助设计与图形学学报, 2005, 17(9):2040-2045.

[7] 陈卓夷. 基于非参数密度估计聚类的关键帧提取方法[J]. 计算机科学, 2007, 34(4):119-121.

[8] Wolf W. Key frame selection by motion analysis [C]// Proceedings of IEEE International Conference on Acoustic, Speech and signal processing. Washington D C, USA: IEEE Computer Society, 1996:1228-1231.

[9] 朱映映, 周洞汝. 一种从压缩视频流中提取关键帧的方法[J]. 计算机工程与应用, 2003, 20(18):13-15.

[10] Sun Z H, Fu P, Xiao J, Meng D. A feature distance based algorithm for video segmentation[C]// Proceedings of the 7th IASTED International Conference on Computer Graphics and Imaging. Kauai Hawaii, USA: ACTA Press, 2004:406-410.

[11] Chang H S, Sull S, Lee S U. Efficient video indexing scheme for content-based retrieval [J]. IEEE Trans on Circuits and Systems for Video Technology, 1999, 9(8): 1269-1279.

[12] 章毓晋. 基于内容的视觉信息检索[M]. 北京: 科学出版社, 2003.

(上接第 68 页)

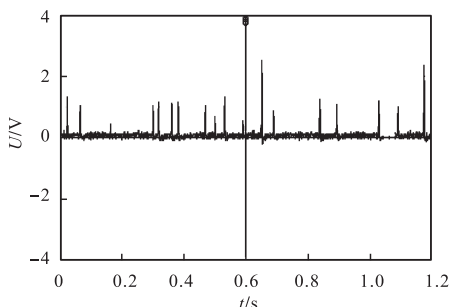


图4 混合大米的示波器波形  
Fig.4 Waveform of mixed rice

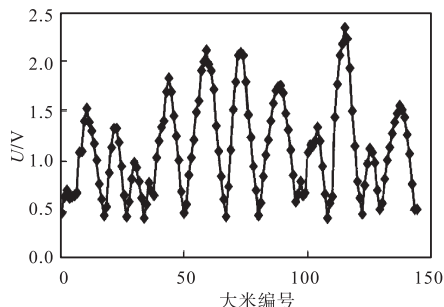


图5 电压峰值分布  
Fig.5 Peak value distribution of voltage when working

### 5 结论

利用双因素方差的检验方法对色选机的输出电压信号进行统计分析, 结果表明: 光源类型的影响显

著, 大米的下落速度和反光板角度对硅光电池输出的影响不显著.

由于首先通过双因素方差分析确定了光源, 只需通过较少对比实验就可确定适宜的反光板角度和下降速度, 能较快地进行针对硅光电池的实验分析, 减少了重复实验的次数.

实验表明, 采用波长范围为 470 ~ 420 nm 的蓝光光源、3 m/s 的下落速度、2 V 的阈值, 能够得到较高的色选精度.

### 参考文献:

[1] Peterson C J, Shelton D R, Martin T J. Grain color stability and classification of hard white wheat in the US[J]. Euphytica, 2001, 119:101-106.

[2] Ram M S, Dowell F E, Seritz L, et al. Development of standard procedures for a simple, rapid test to determine wheat color class[J]. Cereal Chem, 2002, 79(2):230-237.

[3] 宋俊杰. 统计分析[M]. 天津:南开大学出版社, 1986.

[4] 王忠. 我国谷物色选机应用技术现状及发展趋势[J]. 农机化研究, 2006(4):23-25.

[5] 郑力新. 大米色选机信号处理算法与实现[J]. 福建工程学院学报, 2006(8):45-47.

[6] 何静, 关榆君, 张银蒲, 等. 大米色选中的光源选择与模式分类方法研究[J]. 农机化研究, 2007(6):90-93.