

doi: 10.3969/j.issn.1006-1576.2011.09.025

一种改进的自适应关键帧提取方法

陈青华¹, 程恭², 谢晓方¹, 唐江³

(1. 海军航空工程学院兵器科学与技术系, 山东 烟台 264001; 2. 海军装备部, 北京 100841;
3. 海军航空工程学院研究生管理大队, 山东 烟台 264001)

摘要: 针对关键帧提取方法中关键帧数目固定, 不能根据镜头自适应调整的缺点, 将帧间差法与运动变化信息熵方法相结合, 提出一种自适应关键帧提取方法。根据镜头的变化确定提取的关键帧数目, 采用运动信息的变化熵来提取关键帧, 构建关键帧提取的数学模型, 并通过实验进行验证。实验结果表明: 该算法明显优于等间隔关键帧提取算法, 提取的关键帧能较完整地表现序列图像的运动过程, 更有利于对视频内容的理解。

关键词: 关键帧; 自适应提取; 运动变化信息熵

中图分类号: TP391 **文献标志码:** A

An Improved Algorithm of Adaptive Key Frame Extraction

Chen Qinghua¹, Cheng Gong², Xie Xiaofang¹, Tang Jiang³

(1. Dept. of Ordnance Science & Technology, Naval Aeronautical & Astronautical University, Yantai 264001, China;
2. Equipment Department of the Navy, Beijing 100841, China;
3. Administrant Brigade of Postgraduate, Naval Aeronautical & Astronautical University, Yantai 264001, China)

Abstract: Based on the problem that key frame number can not adaptively adjusted according to the lens, the paper proposed an adaptively key frame extraction method based on frame difference and the motion information entropy. The method determined the number of extracted key frame according to changes of the lens and extracted key frame based on the maximal motion and changing information entropy. A mathematical model for key frame extraction was constructed. And the theory of this method was introduced in detail. Many contrasting experiment have demonstrated that this method is obviously better than the key-fame extraction by equal interval. By this method the changing process of sequence is represented perfectly. It provides better perception and understanding for the video.

Keywords: key frame; adaptive extraction; motion and changing information entropy

0 引言

在基于无线数据链的视频传输系统中, 带宽通常是变化的, 从而导致视频主观质量的大幅度变化。通过提取关键帧降低帧率, 提高每帧的目标码率是行之有效的方法。

关键帧提取算法大多以视频检索或视频摘要为目的, 并没有考虑关键帧的后续处理以及其他目的的应用。但对某些具体的应用, 如敏感目标的识别, 在提取出关键帧后还需要进一步分析处理, 这就要求所提取出的关键帧具有较好的静态表达能力, 处于镜头过渡期间或者运动剧烈的帧显然不适合^[1]。因此, 笔者在基于运动变化信息熵的关键帧提取的研究思路^[2]的基础上, 通过将帧间差法与运动变化信息熵方法相结合, 提出了一种改进的自适应关键帧提取方法。

1 关键帧提取方法

关键帧是用于描述一个镜头的关键图像帧, 它

反映了一个镜头的主要内容, 如图 1^[2]。

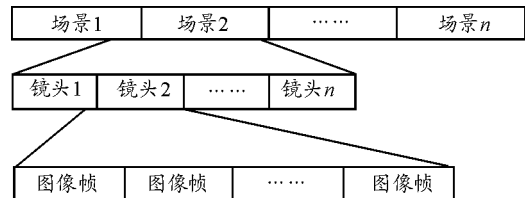


图 1 视频数据结构图

在同一镜头中, t 时刻的图像帧和时刻 $t+1$ 的图像帧在视觉特征和内容上差别不大, 如果对视频的每一帧都进行处理, 工作量会很大, 且将产生大量的冗余操作。关键帧作为架通视频与图像处理的桥梁, 其图像的表达能力和提取速度对后续的处理有重要的影响。根据其机理的不同分为如下 3 种:

1) 等时间间隔采样的方法

该方法以固定的时间间隔抽取视频帧, 算法简单并且计算复杂度低, 但没有考虑视频帧内容, 对于时间短、含义丰富的视频段容易遗漏关键帧, 而对于时间长、基本静止、含义少的视频段却会提取

收稿日期: 2011-04-22; 修回日期: 2011-06-14

基金项目: 海军“十一五”预研项目”

作者简介: 陈青华(1979—), 女, 山东人, 博士研究生, 讲师, 从事武器系统建模与仿真研究。

到多幅关键帧, 造成冗余, 不利于后续的处理。

2) 基于帧间内容变化的方法

该方法根据各视频帧之间的内容变化程度提取关键帧, 用于实时或在线环境, 能较完整地表达视频内容^[3-6]。但是对于两帧间变化程度的衡量所使用的度量标准并不是唯一的, 例如, 文献[4]应用了各帧的颜色直方图差; 文献[5]则使用相邻帧对应图像块的累计能量差作为度量标准。若选取的第一个关键帧恰是镜头变换的过渡帧, 那关键帧的内容表示能力将下降。另外, 如何设置适应于各种视频类型的阈值是这类方法的难点。

3) 基于运动分析的方法

上述 2 类方法都没有考虑运动特征。Wolf 等人提出了一套基于光流分析的算法规则^[7], 该方法通过光流分析来计算镜头中的运动量, 在运动量达到局部最小值处选取关键帧, 它反映了视频数据中的静止。因为需要计算每一帧的光流场, Wolf 的关键帧选取算法需要较长的计算时间, 并只能得到固定数量的关键帧, 不能由用户随意设定关键帧的数量。

2 关键帧提取的数学模型

笔者在研究现有算法的基础上, 提出了改进的最大运动变化信息熵的关键帧提取技术, 该算法先求出镜头内提取的关键帧数目, 然后根据运动信息的变化熵来提取关键帧。基于最大变化信息熵关键帧提取算法的基本思想是: 首先对视频序列中的各个视频帧进行帧间运动分析, 利用运动场的总体能量来描述各帧的运动变化信息, 然后对运动变化信息进行归一化处理, 得到序列中各帧图像的运动变化信息的概率分布, 在此基础上, 采用信息熵来度量各个视频帧所具有的变化信息量, 最后按照信息熵最大化准则提取关键帧。

如图 2 所示, 关键帧提取技术基本的研究任务是从经过镜头分割的视频序列中提取出一些能够代表视频内容的静止图片, 从而以较少的关键帧序列来表达整个视频序列的内容, 相关描述如下:

- 1) 定义一段视频序列 V , 其中视频帧数为 N , 即 $V = \{F_1, F_2, \dots, F_N\}$, 其中 F_1, F_2, \dots, F_N 为视频序列帧号。
- 2) 确定不同镜头关键帧提取数目 N'_i , 整个视频序列共有 s 个镜头。
- 3) 定义最大变化信息熵提取关键帧准则 $F(F'_k)$, 使提取关键帧能很好地表达视频序列的内容。

- 4) 定义输出视频序列 V' , $V' = \{F'_1, F'_2, \dots, F'_{N'}\}$, $F'_1, F'_2, \dots, F'_{N'} \in V$ 。

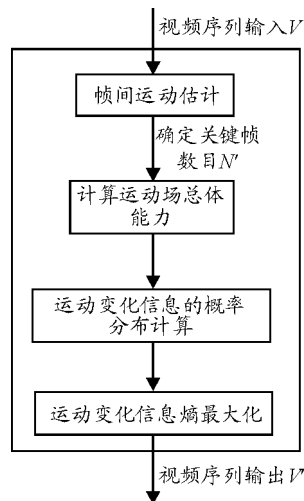


图 2 关键帧提取描述框

2 自适应关键帧提取的流程

2.1 关键帧数目的确定

对每个镜头选取的关键帧数目, 该算法视其内容变化的程度而定, 对于镜头 K_i 内容的变化, 可以使用在一镜头内 K_i 内的平均帧间差 $\overline{d_{ki}}$ 来衡量:

$$\overline{d_{ki}} = \frac{\sum_{k=2}^{N_{ki}} dk}{N_{ki}} \quad (1)$$

式 (1) 中, N_{ki} 表示第 i 个镜头 K_i 内帧的数目, 由于平均 $\overline{d_{ki}}$ 反映了镜头内容变化的程度, 因此可以根据 $\overline{d_{ki}}$ 的大小在所有镜头的累计平均帧间差中所占的比重来确定其所选用的关键帧数目, 这就避免了当目标变化不大的镜头, 选取的关键帧数目过多的问题。

设整个视频序列共有 s 个镜头, 对它们逐个计算平均帧间差, 给内容变化大的镜头分配关键帧数目应该比内容变化小的多, 给定一个关键帧数目 N , 则每个镜头头分配到的关键帧数目公式如下:

$$N_{\text{keyframe}}(K_i) = \max(1, N \times \overline{d_{ki}} / \sum_{j=1}^s \overline{d_{kj}}) \quad (2)$$

其中 \max 是为了保证每个镜头至少有一个关键帧。

2.2 最大运动变化信息熵提取关键帧

假设每一帧图像 B 被分成个 b 个子块, 每一帧子块的集合位于每一子块中像素位移均相同, 所以

第 i 块的速度矢量为 (MV_{xi}, MV_{yi}) ，则当前帧的 $F_k (F_k \in V = \{F_1, F_2, \dots, F_k, \dots, F_N\})$ 的运动场总能量为：

$$Eg_v(F_k) = \sum_{i=1}^b [(MV_{xi})^2 + (MV_{yi})^2] \quad (3)$$

从提取过程可以看出，在视频序列运动变化较为剧烈的各帧中，关键帧提取比较密集，相反，在视频序列运动变化较缓慢的帧中，关键帧的提取较分散，关键帧反映了镜头的动态性，因此提取的关键帧能够很好地还原视频的内容。

3 实验结果及分析

为了验证基于最大运动变化信息熵的关键帧提取技术的有效性，同时采用等间隔提取关键帧的方法进行实验，并将 2 种方法得到的实验结果进行对比。以 Intel Core(TM)2 Quad CPU Q8200 为例，CPU 主频 2.33 GHz 作为测试计算机，基于 MATLAB 9.0 选取红外热像仪采集视频片段提取关键帧：一组为 5 761 帧的 ship 测试序列，该序列运动变化较明显，其中从第 1 876 帧到 1 976 帧共 100 帧属于同一个镜头，提取的关键帧为 7 帧，运行时间为 1 443.46 ms。

图 3 显示了不同帧熵值变化概率分布情况，需要从 100 帧中提取出 7 关键帧号分别为 (10, 19, 28, 63, 73, 83, 99)。关键帧提取的结果如图 4。另外，采用等间隔关键帧提取方法所得到的采样结果如图 5，关键帧号分别为 (14, 28, 42, 56, 70, 84, 98)。

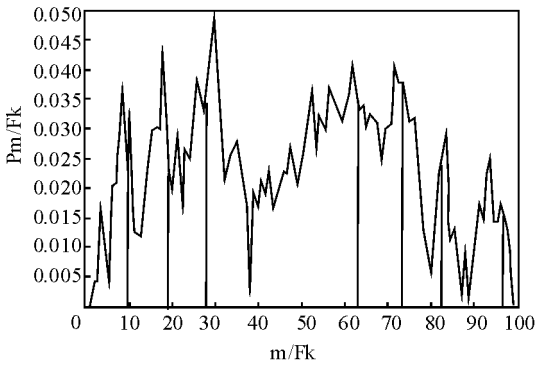


图 3 序列的运动变化信息概率分布

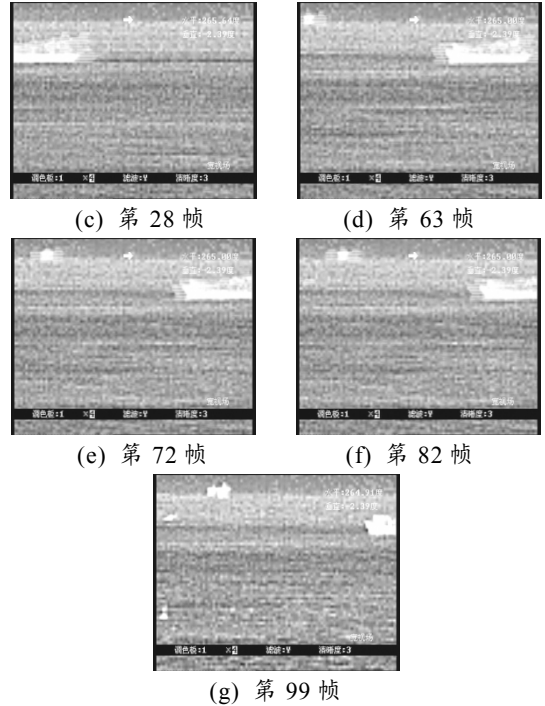
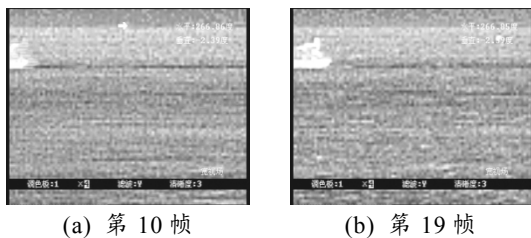


图 4 基于最大运动变化信息熵的关键帧提取的结果

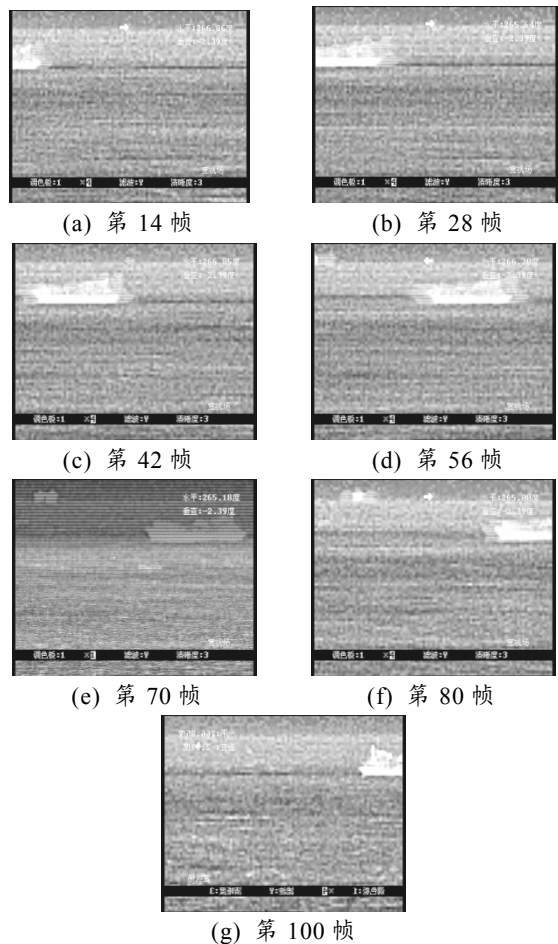


图 5 等间隔提取关键帧结果