

一种基于图的近重复视频子序列匹配算法

刘红

(第二军医大学网络信息中心, 上海 200433)

摘要: 为了解决近重复视频检测中的效果和效率问题,提出了一种基于图的近重复视频子序列匹配算法。将基于关键帧特征的相似性查询结果构建成匹配结果图,进而将近重复视频检测转换成一个在匹配结果图中查找最长路径的问题。该算法有三个主要优势:a)它能在众多杂乱的匹配结果中找到最佳的匹配序列,有效剔除了某些假“高相似度”匹配带来的噪声,因而能在一定程度上弥补底层特征描述力的不足;b)由于它充分考虑和利用了视频序列的时序特性,具有很高的近重复视频定位准确度;c)它能自动检测出匹配结果图中存在的多条离散路径,从而能一次性检测出两段视频中可能存在多段近重复视频的情形。提出的算法不仅提高了检测的准确度,而且提高了检测效率,取得了良好的实践效果。

关键词: 图;近重复视频;子序列匹配

中图分类号: TP391;TP301.6

文献标志码: A

文章编号: 1001-3695(2013)12-3857-06

doi:10.3969/j.issn.1001-3695.2013.12.089

Graph-based near-duplicate video subsequence matching algorithm

LIU Hong

(Network Information Center, Second Military Medical University, Shanghai 200433, China)

Abstract: In order to improve the effectiveness and efficiency of near-duplicate video detection, this paper proposed the graph-based video subsequence matching algorithm. The algorithm constructed a matching results graph from the similarity search results based on the key frame features, and then converted the problem of near-duplicate video detection into the problem of finding the longest path in the matching results graph. The method has three main advantages: a) Graph-based method could find the best matching sequence in many messy match results, which effectively excluded false “high similarity” noise and compensated the limited description of image low level visual features. b) The graph-based method took fully into account the spatiotemporal characteristic of video sequence, and had high location accuracy. c) The graph-based sequence matching method could automatically detect the discrete paths in the matching result graph. Thus, it could detect more than one near-duplicate video. The proposed algorithm not only improves the detection accuracy, but improves the efficiency of detection, achieves good practical effect.

Key words: graph; near-duplicate video; video subsequence matching

0 引言

视频的一个非常重要的内在特征是它的时间特性,即视频由时间方向上连续的视频帧所组成。因此,基于内容的视频检索或者近重复视频的判断通常需要通过比较视频帧集合的相似性来判断两段视频之间的相似性。虽然在原理上,基于内容的近重复视频检测与基于内容的视频检索是基本一致的,但在实际应用中,两者的实现方式还是存在较大的差别。基于内容的视频检索侧重于查询视频与目标视频在整体上的相似度,即更多的是一种“一对一”的关系;而对于近重复视频检测而言,其相似性检索可能会面对更多的匹配形式,这些匹配形式能够被概括为“一对一”“一对多”和“多对多”三种形式^[1]。图1对这三种形式进行了说明。因为近重复视频以何种形式出现完全依赖于目标视频的具体应用,如果近重复视频是经过复杂的编辑、修改等操作,其检测工作将是一个典型的视频子序列匹配问题。在复杂的近重复视频检测中,目标子序列的长

度、出现的位置以及出现的频率都是未知的(如图1(c)所示的“多对多”的形式),这种形式的视频子序列匹配任务将比普通的视频检索任务更为复杂。

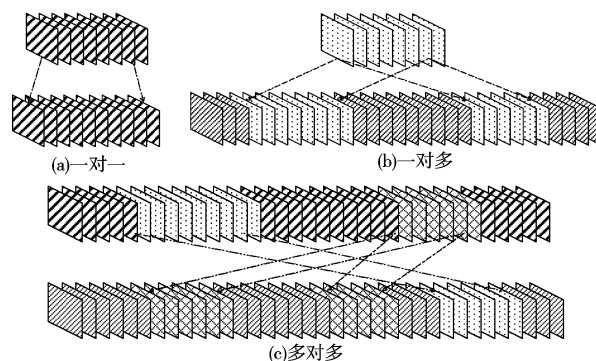


图1 三种视频子序列匹配情形

目前,绝大多数基于内容的近重复视频检测方法主要侧重于两段完整视频的拷贝检测^[2-5]或者部分近重复视频检测^[1,6-9]。这些方法尽管在各自的应用场景下取得了较好的

收稿日期: 2013-01-22; 修回日期: 2013-04-19

作者简介: 刘红(1975-),男,湖南涟源人,高级工程师,硕士,博士,主要研究方向为多媒体信息处理与检索(liuhong 2007@fudan.edu.cn)。

性能,但是当面对如图 1(c)中的“多对多”的复杂情形时,这些方法难免会遇到检测瓶颈。本文在总结这些方法各自优缺点的基础上,提出了一种基于图的视频子序列匹配方法。该方法将基于视频关键帧特征相似性查询结果转换为匹配结果图,进而将近重复视频检测转换成了一个在匹配结果图中查找最长路径的问题。

1 基于图的视频子序列匹配

在基于内容的近重复视频检测中,可以利用视频帧的空间特征来度量视频在内容上的相似性^[10-12],同时时间特性是视频的一个本质特性。近重复视频检测最终是在视频帧序列上的匹配和定位。本文中,视频空间信息上的相似性通过单帧图像特征的相似性查询来完成^[13,14],本文研究的视频子序列匹配方法则侧重于对视频空间信息和时间信息的结合,实现对近重复视频子序列的判别和定位,即解决一个近重复视频出现在什么位置和出现了多少次的问题。通常情况下,近重复视频在目标视频的出现位置和次数是很难预测的,因此,基于单帧图像特征得到的相似性匹配结果(即空间信息匹配结果)需要一个有效的模型将其转换成时间序列上的匹配结果。查询视频和参考视频在 1 维时间线上某个点上的匹配可以用一个匹配节点来描述,而一段视频序列的匹配可以用一串点来描述,这很自然地联想到通过图模型来解决视频子序列匹配问题。图转换用一种自然直观的方式被广泛地用于描述复杂的结构^[15],本文将视频在空间信息上的匹配结果转换成匹配结果图,进而将近重复视频的定位问题转换成在匹配结果图中查找关键路径问题。其具体流程如图 2 所示。其核心步骤包括:视频关键帧抽取;基于关键帧的特征匹配;根据匹配结果构建匹配结果图;匹配路径检测;匹配路径过滤;近重复子序列定位。

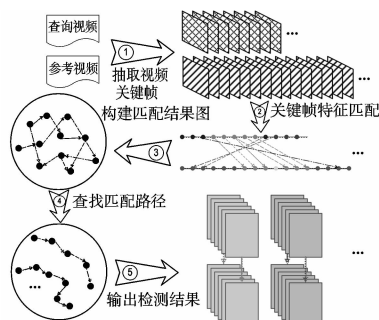


图 2 基于图的视频子序列匹配流程

为了便于接下来讨论,本文对相关概念先作如下定义:

$Q = \{q_1, \dots, q_{|Q|}\}$ 和 $R = \{r_1, \dots, r_{|R|}\}$ 分别表示查询视频和参考视频关键帧集合。其中, q_i 表示查询视频的第 i 帧, r_j 表示参考视频的第 j 帧, $|Q|$ 表示查询视频帧的数目, $|R|$ 表示参考视频帧的数目。

$G = (N, E)$ 表示匹配结果图。其中, N 表示匹配节点集合, E 表示边集合。

$N = \{M_{i,j} | 1 \leq i \leq |Q|, 1 \leq j \leq |R|\}$ 。节点集合中的节点 $M_{i,j}$ 表示查询视频帧 q_i 和参考视频帧 r_j 之间存在一个满足指定约束条件的匹配。最常用的约束是查询视频帧 q_i 和参考视频帧 r_j 之间的相似度大于某个指定的阈值 T 。另外,也可通过 KNN 来指定,即每帧查询帧 q_i 返回前 k 个最近邻的参考视频帧。

$E = \{\langle M_{i,j}, M_{l,m} \rangle\}$ 。 $\langle M_{i,j}, M_{l,m} \rangle$ 表示节点 $M_{i,j}$ 和 $M_{l,m}$ 之间

存在一条边。

$\text{sim}(q_i, r_j)$ 表示查询视频帧 q_i 和参考视频帧 r_j 之间的相似度。

$P = \{p_1, \dots, p_n\}$ 表示图 G 所有匹配路径的集合。

$\vec{Q} = \{q_m, \dots, q_n\}, 1 \leq m < n \leq |Q|$ 和 $\vec{R} = \{r_m, \dots, r_n\}, 1 \leq m < n \leq |R|$ 分别表示查询视频子序列和参考视频子序列。

$\text{sim}(p_j)$ 表示匹配路径 p_j 中对应两段视频子序列之间的相似度。

$S_Q = \{a_1, \dots, a_L\}$ 和 $S_R = \{b_1, \dots, b_L\}$ 分别表示视频子序列匹配路径中查询视频对应的帧 ID 集合和参考视频对应的帧 ID 集合,即表示在该匹配路径中,查询视频的第 a_i 帧与参考视频的第 b_i 帧,它们之间是一一对应的关系, L 表示匹配路径的长度。

根据上述定义,视频子序列匹配的任务也可定义如下:在一个预定义的分值函数下,查找参考视频序列中与查询视频序列中相对应的最相似的部分,即查找两个视频子序列 $\vec{Q}$ 和 $\vec{R}$ 使之具有最高的相似度。

1.1 自适应双阈值法抽取视频关键帧

如图 2 所示的基于图模型的视频子序列匹配流程图,视频关键帧的抽取是基于内容的近重复视频检测工作的第一步。视频关键帧抽取的准确性和粒度很大程度上决定了近重复视频检测工作的质量。视频是由具有时序特性的帧序列组成的,考虑到人眼的视觉特性,相邻视频帧的相似度会很高,这就使得连续视频帧的视觉信息在时间方向上有很大的冗余性。如果在进行近重复视频检测时,对视频中的每一帧都进行特征的抽取,必然带来很多额外的不必要的计算。所以需要抽取有代表性的视频关键帧,以降低视频视觉信息的冗余度。这样,进行近重复视频检测时只需要匹配关键帧序列即可,从而大大减少了匹配过程中的计算代价。

目前,已经提出了许多视频关键帧的抽取方法^[16-23],这些方法由于应用场景的不同,其方法的侧重点也不同,因此,这些方法在拥有其自身固有优点的同时也难免暴露其在另一方面的缺点。例如,帧间连续比较方法^[18,24]有计算简单、直观、低计算复杂度等优势,但有关键帧数目不可控、相同镜头里可能会出现冗余信息等局限性;基于聚类的方法^[25]虽然具有能够使用通用的聚类算法,关键帧能反映镜头的全局特性等优点,但它不能自然地利用视频的时序特性,通常要通过一些特定设置去确保毗邻的视频帧可能被分配到相同的类;基于对象或者事件检测的方法虽然包含了重要的语义信息,但它与具体的应用之间存在强烈的关联和依赖性,而且这类方法具有高的计算复杂度。

本文研究的近重复视频检测具有明显的针对性,即检测两段视频是否相同或者接近相同,因此本文的方法并没有将研究重点放在视频的结构化分析上,而是将重点聚焦于两个要点之上:有效消除视频冗余帧提高检测效率;保持抽取关键帧的颗粒度确保检测结果的准确性。基于这两个目标,本文提出了一种自动双阈值的方法来消除视频冗余帧,并将视频分割成一个个视频片段,对每一个视频片段采用固定抽帧率的方法在每个视频片段中抽取三帧关键帧来代表该视频片段。该方法既有效消除了视频冗余帧,又保证了抽取关键帧的颗粒度,取得了良好的实践效果。

为了消除视频在视觉信息上的冗余性,提高视频子序列的

匹配速度,一种有效的解决方案是抽取一组远少于整个视频帧数目的关键帧来代表视频内容。视频关键帧抽取的通用做法是先对视频进行镜头分割,然后再在每一个镜头中抽取最能代表镜头的视觉内容的关键帧。目前,基于镜头分割的关键帧抽取方法已有很多成熟的算法^[17,18],在视频检索领域里也有很好的应用效果。但是就近重复视频检测应用而言,基于镜头分割的方法可能在两个方面存在一定的局限性:内容相近的镜头里存在冗余的视觉信息;当查询视频很短时(小于 5 s),关键帧抽取的粒度难以保证检测准确度。因此,本文并没有完全采取镜头分割的方法来抽取关键帧,而只是沿用了分割的思想,将视频分割成一段段在视觉内容上相似的视频片段,然后在每个视频片段中抽取固定数目的关键帧来代表视频片段。图 4 显示了提出的视频片段分割与常用的镜头分割方法之间的区别。为了应对视频视觉内容上的突变和渐变,本文采用类似于文献[26]中提出的双阈值方法。该方法有如下两个特点:a)使用了双阈值方法,一个阈值 T_h 用于检测帧视觉信息的突变,另一个阈值 T_l 用于检测帧视觉信息的渐变;b)两个阈值的大小并不是预先确定的,而是根据整个视频的视觉信息的变化自动确定阈值的大小。两个阈值 T_h 和 T_l 分别定义如下: $T_h = \mu + \alpha\sigma$, $T_l = \beta T_h$ 。其中, μ 和 σ 分别是整个视频帧基于颜色特征的均值和标准偏差;根据文献[16,26]的研究, α 建议设置为 5~6;根据文献[18]的研究, β 通常在 0.1~0.5 间选择。

本文提出的自适应双阈值法抽取关键帧的原理如图 3 所示。图中是一段连续的视频帧, C_{1f} 表示片断 1 的第一帧; C_{1l} 表示片断 1 的最后一帧, C_{2f} 表示片断 2 的第一帧, T_h 表示突变阈值, T_l 表示渐变阈值,则视频片断 1 和视频片断 2 必须满足下列两种情况中的任意一种:a)视频片断 1 的最后一帧与视频片断 2 的第一帧之间的相似度 $\text{sim}(C_{1l}, C_{2f}) < T_h$; b)视频片断 1 的第一帧与视频片断 2 的第一帧之间的相似度 $\text{sim}(C_{1f}, C_{2f}) < T_l$ 。通过自适应双阈值分割法可以将连续的视频帧分割成一段段时间上连续的、视觉上相近的视频片断,对于每个视频片断抽取三帧来代表,这三帧分别为片断的第一帧、关键帧和最后一帧。其中关键帧由与平均帧(该片断所有帧特征的均值)最相似的帧来充当。关键帧用于视频序列匹配,而第一帧和最后一帧主要用于精确定位拷贝位置 and 进行辅助匹配,沿时间方向给分割后的视频片断分配连续的 ID 号。

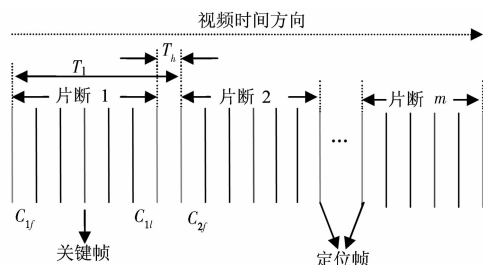
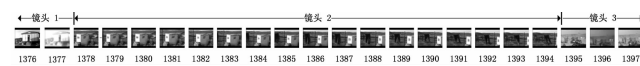


图 3 自适应双阈值消除视频冗余帧和抽取关键帧

图 4 说明了本文提出的自适应双阈值消除视频冗余帧的方法与普通的基于镜头分割方法的区别。图 4 中, (a) 是使用镜头分割方法得到的结果, (b) 是使用自动双阈值消除视频冗余帧得到的分割结果。在图 4 中, 视频片段 1 的最后一帧(ID 1377)与视频片段 2 的第一帧(ID 1378)之间的距离(基于颜色特征之间的距离)大于突变阈值 T_h , 视频片段 2 的第一帧(ID 1378)与视频片段 3 的第一帧(ID 1385)之间的距离大于突变阈值 T_l 。因此, 基于本文提出的自适应双阈值分割法, 镜头 2

能够被更进一步地分割成片段 2 和片段 3, 具有更小的粒度, 本文称之为视频子镜头或者视频片段。



(a) 镜头分割结果



(b) 自适应双阈值分割结果

图 4 自适应双阈值分割法与镜头分割法结果对比

1.2 构建匹配结果图

根据图 2 所示的匹配流程, 构建匹配结果图的前期工作还包括: 抽取视频关键帧特征; 查询视频关键帧特征和参考视频关键帧特征之间的相似度查询。这两项研究工作请参见文献[27,28], 这里不再赘述。匹配结果图的构建是建立在查询视频关键帧特征和参考视频关键帧特征之间的相似度查询结果之上的。对于查询视频关键帧集合 $Q = \{q_1, \dots, q_{|Q|}\}$ 和参考视频关键帧集合 $R = \{r_1, \dots, r_{|R|}\}$, 计算查询视频帧 q_i 和参考视频帧 r_j 之间的相似度 $\text{sim}(q_i, r_j)$, 对于查询视频中的每一关键帧 q_i , 根据相似度值的大小返回参考视频中的前 k 帧最相似的关键帧。 k 的大小由参考视频分割后的片断个数来确定, 即 $k = \alpha n$, α 是一个比例因子, 通常取 0.05。该匹配方法有效地避免了需要确定一个刚性的阈值来决定匹配结果返回的缺陷。图 5 是一个匹配结果示意图, 图中的前 8 列数据取自两段真实视频的匹配结果。匹配时, 查询视频帧集合 $Q = \{q_1, \dots, q_{|Q|}\}$ 中的每一帧 q_i 按时间排序, 并对参考视频帧集合 $R = \{r_1, \dots, r_{|R|}\}$ 中的每一帧 r_j 进行相似性检索, 匹配结果按相似度从大到小排序。以图中的第一列为例, 则有 $\text{sim}(q_1, r_{26}) \geq \text{sim}(q_1, r_{27}) \geq \text{sim}(q_1, r_{28}) \geq \dots$ 。取前 k 个返回结果, 图中 $k=5$ 。图 5 中的每一个匹配可以转换为匹配结果图中的一个节点 $M_{i,j}$, 从而可得到匹配结果图中的每一个节点; 得到图节点之后, 下一步就是要确定两个节点之间是否存在有向边。为了描述图中两个节点是否可达, 本文先作了两个定义:

a) 节点时间方向一致性。对两节点 $M_{i,j}$ 和 $M_{l,m}$, 如果满足 $(i-l)(j-m) > 0$, 则认为两节点具有时间方向一致性。

b) 节点时间跳跃度。节点 $M_{i,j}$ 和 $M_{l,m}$ 之间的时间跳跃度定义为

$$\Delta t_{l,m}^{i,j} = \max(|t_i - t_l|, |t_j - t_m|) \quad (1)$$

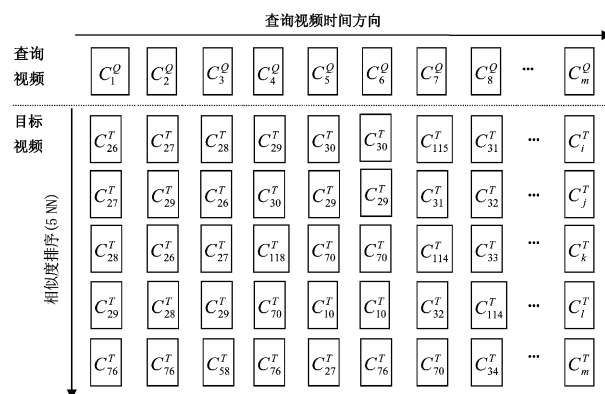


图 5 查询视频与参考视频匹配结果示例

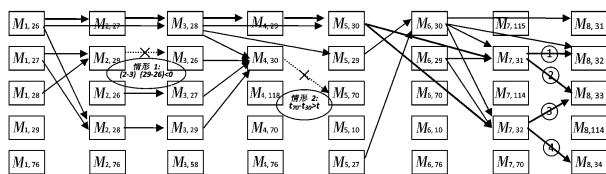
说明: 匹配时, 查询视频 QC 中的每一个 C_i^Q (按时间排序) 分别与目标视频 TC 中的所有类进行比较, 匹配结果按相似度从大到小排序, 取前 k 个返回结果, 图中 $k=5$ 。

匹配结果图中的两个节点之间如果存在边, 则必须同时满

足如下两个条件:

- 两节点必须满足时间方向一致性;
- 两节点的时间跳跃度 $\Delta t < \tau$ (τ 为时间跳跃度阈值)。

条件 a) 表示图中两节点代表的查询视频的时间方向与目标视频的时间方向必须一致, 这样做似乎也是合理的, 因为视频序列是一个时间序列, 拷贝视频和被拷贝视频在时间方向必定是一致的。时刻增大的方向即为节点间有向边的方向。条件 b) 则表示图中两节点所代表的两个匹配结果在时间方向上的跳跃不能超过某个阈值, 否则认为两个匹配结果之间不存在关联。根据上述方法和条件, 图 5 的匹配结果就可以转换成一个匹配结果图, 很明显, 这样生成的匹配结果图一定是一个有向无环图。图 5 中的前 8 列数据的匹配结果, 转换成匹配结果图后如图 6 所示。本文也做了一个对比实验来获取一个最优的时间跳跃度阈值, 如图 7 所示, 图中的 T1 ~ T10 分别代表 10 种不同的视频拷贝类型 (详情请参见文献 [29]), 实验数据显示一个相对合适时间跳跃度阈值应该在 10 ~ 29 s 间。另一方面, 本文也对自适应双阈值法分割视频后得到的视频片段的时间长度进行了统计。如图 8 所示, 图中的数据显示视频片段的时间长度大多集中在 10 ~ 30 s 间, 这也从另一个侧面验证了时间跳跃度阈值选择在 10 ~ 30 s 之间的合理性。



匹配结果在时间序列进行重新排列和提炼,有效地整合了视频的空间特性和时间特性。该方法具有如下主要优点:

a) 能从无序的视频空间特征匹配结果中查找最优匹配子序列。图9使用了极具挑战性的拷贝类型即“画中画”拷贝类型,来说明基于图方法的优点;虚线框A中包含时间方向的一些查询视频帧;虚线框B中则是查询视频帧根据相似性查询返回的参考视频帧集合(参考视频帧按相似度从高到低排列)。然后,基于图的子序列匹配方法能自动从这些匹配结果中查找到一条最优的匹配路径。在图9中,箭头代表的路径即是本文基于图方法得到的两段视频最终的匹配结果。

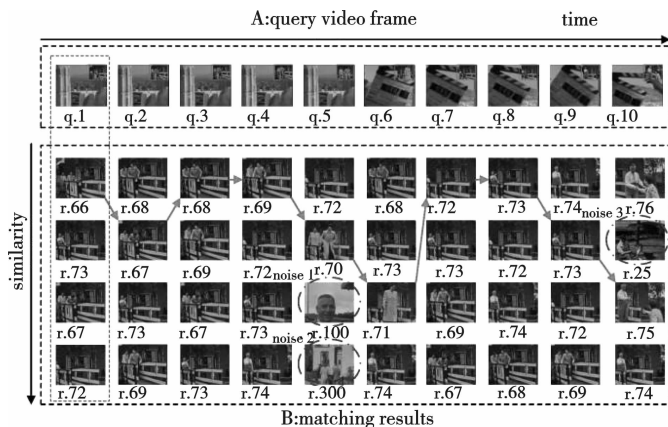


图9 图方法能从匹配结果中自动查找最优的路径

b) 能自动剔除视频空间特征匹配结果中的一些噪声。在检测过程中,在基于视频空间特征的匹配结果中往往会存在一些噪声(如图9中所示的 noise 1、2、3)。基于图的子序列匹配方法能利用时间方向一致性和时间跳跃度阈值两个约束条件自动剔除这些噪声。

c) 能自适应视频帧率的改变。图10中查询视频的帧率是图9中查询视频帧率的2倍,实验结果表明基于图的匹配方法能有效应付不同帧率的拷贝视频。

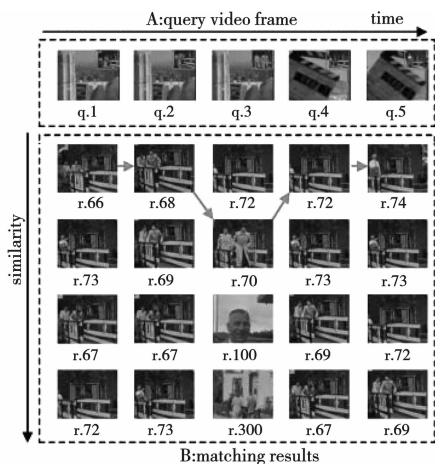
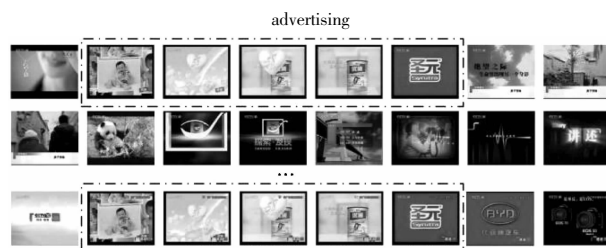


图10 图方法能自适应视频帧率的改变

d) 能同时检测两段视频中包含的多个近重复视频片段。在实际应用中,近重复视频片段可能会在一段视频中出现多次。例如,图11显示了同一段广告在不同时段被多次播放。图12中的实验结果表明,基于图的视频子序列匹配方法能一次性检测出该视频中包含的多段近重复视频。同时,这也是有实际应用价值的,如广告商想知道电视台是否按合同在规定的时间内和频率播放了指定的广告,就可以通过基于图的方法来进行自动检测,而不必通过人工观看视频的方式来确认。



advertising again

图11 同一段广告在不同时间点被播放

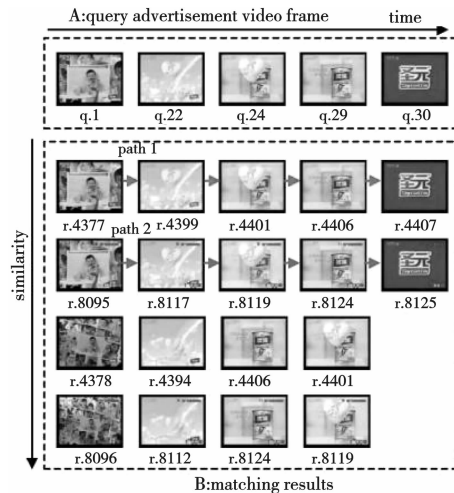


图12 图方法能同时检测两段视频中包含的多个近重复视频片段
path 1 和 path 2 表示同一段广告视频在不同时间播放时的检测结果

3 结束语

由于近重复视频检测过程中存在多种不确定因素,如目标视频中是否存在近重复视频、近重复视频片断有多长和近重复视频片断处于什么位置等,这使得近重复视频检测相比于通常的视频检索难度更大。目前,大多数近重复视频检测中的子序列匹配算法主要存在两个缺陷:a) 匹配过程中需要事先确定一个相似度阈值来决定返回的匹配结果,而对于不同的视频,要确定一个通用的阈值是很困难的;b) 穷尽法检测所有可能长度和可能位置的拷贝视频,其计算代价是很大的。为此,本文提出了基于图的视频子序列匹配算法,该方法巧妙地将视频序列匹配结果转换为匹配结果图,进而将视频拷贝检测转换成了一个在匹配结果图中查找最长路径的问题。基于图的方法有几个明显的优势:a) 基于图的方法能够在众多杂乱的匹配结果中找到一条最佳的匹配序列,有效地剔除了某些假“高相似度”匹配带来的噪声,因而它能弥补底层特征描述力不足的缺陷;b) 由于它充分考虑和利用了视频序列的时序特性,从而具有很高的近重复视频定位准确度;c) 基于图的序列匹配方法能够自动检测出多条离散的路径,因而它能批量定位两段视频中可能存在多段近重复视频的情形;d) 相比于穷尽检测方法,该方法在检测速度方面也具有明显优势。

参考文献:

- [1] CHEN Tian-long, JIANG Shu-qiang, CHU Ling-yang, *et al.* Detection and location of near-duplicate video sub-clips by finding dense subgraphs[C]//Proc of the 19th ACM International Conference on Multimedia. New York: ACM Press, 2011: 1173-1176.
- [2] LAW-TO J, BUISSON O, GOUET-BRUNET V, *et al.* Robust voting algorithm based on labels of behavior for video copy detection[C]//Proc of the 14th ACM International Conference on Multimedia. New

- York: ACM Press, 2006: 835-844.
- [3] LAW-TO J, CHEN Li, JOLY A, *et al.* Video copy detection: a comparative study[C]//Proc of the 6th International Conference on Image and Video Retrieval. 2007: 371-378.
- [4] XIE Qing, HUANG Zi, SHEN Heng-tao, *et al.* Efficient and continuous near-duplicate video detection[C]//Proc of the 12th International Asia-Pacific Web Conference. 2010.
- [5] YEH Mei-chen, CHENG Kwang-ting. Video copy detection by fast sequence matching[C]//Proc of ACM International Conference on Image and Video Retrieval. New York: ACM Press, 2009.
- [6] HUA Xian-sheng, CHEN Xian, ZHANG Hong-jiang. Robust video signature based on ordinal measure[C]//Proc of International Conference on Image Processing. 2004: 685-688.
- [7] SHEN Heng-tao, SHAO Jie, HUANG Zi, *et al.* Effective and efficient query processing for video subsequence identification[J]. *IEEE Trans on Knowledge and Data Engineering*, 2009, 21(3): 321-334.
- [8] TAN H K, NGO C W, CHUA T S. Efficient mining of multiple partial near-duplicate alignments by temporal network[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2010, 20(11): 1486-1498.
- [9] TAN H K, NGO C W, HONG R, *et al.* Scalable detection of partial near-duplicate videos by visual-temporal consistency[C]//Proc of the 17th ACM International Conference on Multimedia. New York: ACM Press, 2009: 145-154.
- [10] HAMPAPUR A, BOLLE R M. Comparison of distance measures for video copy detection[C]//Proc of IEEE International Conference on Multimedia and Expo. 2001: 737-740.
- [11] HAMPAPUR A, HYUN K H, BOLLE R. Comparison of sequence matching techniques for video copy detection[C]//Storage and Retrieval for Media Databases. 2002: 194-201.
- [12] WU Xiao, NGO C W, HAUPTMANN A, *et al.* Real-time near-duplicate elimination for Web video search with content and context[J]. *IEEE Trans on Multimedia*, 2009, 11(2): 196-207.
- [13] CHEN Li, STENTIFORD F W M. Video sequence matching based on temporal ordinal measurement[J]. *Pattern Recognition Letters*, 2008, 29(13): 1824-1831.
- [14] KIM C, VASUDEV B. Spatiotemporal sequence matching for efficient video copy detection[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2005, 15(1): 127-132.
- [15] EHRIG H, TAENTZER G. Computing by graph transformation: a survey and annotated bibliography[R]. [S. l.]: Leiterder Fach Bibliothek Informatik, 1996.
- [16] ZHANG Hong-jiang, KANKANHALLI A, SMOLIAR S W. Automatic partitioning of full-motion video[J]. *Multimedia Systems*, 1993, 1(1): 10-28.
- [17] SHAHRARAY B, GIBBON D C. Automatic generation of pictorial transcripts of video programs[C]//Multimedia Computing and Networking. 1995: 512-518.
- [18] ZHANG H J, WU J H, SMOLIAR S W. System for automatic video segmentation and key frame extraction for video sequences having both sharp and gradual transitions[US, 5635982 A[P]]. 1997-06-03.
- [19] DIRFAUX F. Key frame selection to represent a video[C]//Proc of International Conference on Image Processing. 2000: 275-278.
- [20] SZE K W, LAM K M, QIU Guo-ping. A new key frame representation for video segment retrieval[J]. *IEEE Trans on Circuits Systems for Video Technology*, 2005, 15(9): 1148-1155.
- [21] TRUONG B T, VENKATESH S. Video abstraction: a systematic review and classification[J]. *ACM Trans on Multimedia Computing, Communications, and Applications*, 2007, 3(1): 1-37.
- [22] KUCUKTUNC O, GUDUKBAY U, UIUSOYA O. Fuzzy color histogram-based video segmentation[J]. *Computer Vision and Image Understanding*, 2010, 114(1): 125-134.
- [23] DOUZE M, GAIDON A, JEGOU H, *et al.* INRIA-LEAR's video copy detection system[EB/OL]. (2008). <http://www-nlpir.nist.gov/projects/tvpubs/tv8.papers/inria-lear.pdf>.
- [24] ZHANG Xu-dong, LIU Tie-yan, LO K T, *et al.* Dynamic selection and effective compression of key frames for video[J]. *Pattern Recognition Letters*, 2003, 24(9): 1523-1532.
- [25] GUIL N, GONZÁLEZ-LINARES J M, CÓZAR J R, *et al.* A clustering technique for video copy detection[C]//Proc of the 3rd Iberian Conference on Pattern Recognition and Image Analysis. Berlin: Springer-Verlag, 2007: 452-458.
- [26] FURHT B, SMOLIAR S W, ZHANG Hong-jiang. Video and image processing in multimedia systems[M]. Boston: Kluwer Academic Publisher, 1995.
- [27] LIU Hong, LU Hong, WEN Zhao-hui, *et al.* Gradient ordinal signature and fixed-point embedding for efficient near-duplicate video detection[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2012, 22(4): 555-566.
- [28] 刘红. 近重复视频检测算法研究[D]. 上海: 复旦大学, 2012.
- [29] Final list of transformations[EB/OL]. (2008). <http://www-nlpir.nist.gov/projects/tv2008/active/copy.detection/final.cbcd.video.transformations.pdf>.
- [30] LIU Hong, LU Hong, XUE Xiang-yang. A segmentation and graph-based video sequence matching method for video copy detection[J]. *IEEE Trans on Knowledge and Data Engineering*, 2013, 25(8): 1706-1718.
- [31] XUE Xiang-yang, ZHANG Wei, GUO Yue-fei, *et al.* Fudan university at TRECVID[EB/OL]. (2008). <http://www-nlpir.nist.gov/projects/tvpubs/tv8.papers/fudan.pdf>.

(上接第 3856 页)

- [6] SCHECHNER Y Y, NARASIMHAN S G, NAYAR S K. Instant dehazing of images using polarization[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Kauai, HI: IEEE Computer Society, 2001: 1-325.
- [7] SHWARTZ S, NAMER E, SCHECHNER Y Y. Blind haze separation[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2006: 1984-1991.
- [8] SONKA M, HLANACE V, BOYLE R. 图像处理、分析与机器视觉[M]. 艾海舟, 武勃, 等译. 2 版. 北京: 人民邮电出版社, 2003.
- [9] TAN R. Visibility in bad weather from a single image[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2008: 2347-2354.
- [10] 禹晶, 李大鹏, 廖庆敏. 基于物理模型的快速单幅图像去雾方法[J]. *自动化学报*, 2011, 37(2): 143-149.
- [11] FATTAL R. Single image dehazing[J]. *ACM Trans on Graphics*, 2008, 27(3): 1-9.
- [12] HE K, SUN J, TANG Xiao-ou. Single image haze removal using dark channel prior[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2009: 1956-1963.
- [13] LEVIN A, LISCHINSKI D, WEISS Y. A closed form solution to natural image matting[C]//Proc of IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Computer Society, 2006: 67-68.