相似视频片段的检测与定位方法研究

郭延明 谢毓湘 老松杨 白 亮
(国防科学技术大学信息系统工程重点实验室 长沙 410073)

摘 要 相似视频片段探测可以辅助网络视频检索、内容关联分析等方面的研究,具有重要的意义。重点研究了位置随机的相似视频片段的探测与定位问题,首先在视频结构化分析与关键帧提取的基础上,对不同视频进行相似关键帧探测。为保证探测的精度与效率,针对视频关键帧的特点,采用了FAST检测子和BRIEF描述子相结合的方法,利用关键帧的局部特征进行相似关键帧探测;其次提出了一种相似关键帧距离度量的方法,利用相似关键帧所在源视频的位置来构建相似关键帧距离矩阵,保留矩阵中距离较小的相似关键帧,将寻找相似视频片段的过程转化为寻找矩阵对应的连通图的过程。最后对算法进行了实验,结果表明,该方法可以有效地探测处于各个位置的相似视频片段。

关键词 FAST,BRIEF,相似关键帧,相似视频片段
中图法分类号 TP391.7 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2014.10.012

Detection and Location of Near-duplicate Video Clips

GUO Yan-ming XIE Yu-xiang LAO Song-yang BAI Liang
(The Key Laboratory of Information Systems Engineering, National University of Defense Technology, Changsha 410073, China)

Abstract The paper aimed to detect and locate multiple near-duplicate video clips at random locations. First, video structure analysis was carried out to extract the keyframes. Second, to ensure the accuracy and efficiency, a method which combines the advantages of FAST and BRIEF was proposed to find the near-duplicate keyframes (NDK) between the videos. Then, the paper put forward an algorithm to calculate the distance between NDKs, using the locations of the keyframes in the source video. In this way, we could get a distance matrix of NDKs. We reserved the close distance in the matrix by setting a distance threshold, and transformed the detection and location of multiple near-duplicate video clips to find the connected graphs that the matrix corresponds to. The experimental results show that the method can effectively detect and locate near-duplicate video clips at random locations.

Keywords FAST, BRIEF, Near-duplicate keyframe, Near-duplicate video clips

1 引言

近年来,随着网络技术、视频硬件处理技术和软件处理技术的快速发展和广泛应用,网络视频数据呈现出了爆炸式的增长,其中包含了大量的近重复视频片段,迅速地探测这些相似视频片段对于网络视频检索、内容关联分析、自动视频标注具有重要的意义。

目前,国内外对于视频相似性探测主要集中在两个方面:视频片段检索和近重复视频探测。视频片段检索[1-3]是指在大量的视频集合中寻找含有某一视频片段的视频集合,一般所要检索的视频片段长度较小,常用的检索方法是选择比视频片段稍大的移动窗口在时间较长的视频序列上移动,记录窗口内的视频片段与所要检索的视频片段的相似度,选择相似度较高的片段作为最终的检索结果。

近重复视频探测是指对内容和时间长度相同或者相似的视频进行探测,其突出的特点是所要比较的视频之间保持了基本的时序关系,常用的探测算法有两种:DTW方法[4]和构造视频特征曲线的方法[5-7]。DTW方法将所要比较的视频划分为视频帧序列,计算两个帧序列元素之间的距离,并找到满足帧序列距离之和最小时的视频帧之间的对应关系,如果帧序列距离之和小于一定的阈值,则两者为相似视频。构造特征曲线的方法一般是利用视频的全局或者局部特性,为每个视频帧计算一个特征值,将特征值连接起来构建视频的特征曲线,通过比较视频之间的特征曲线来判定视频是否相似。

上述方法都是基于视频之间保持了基本的时序关系而开展的,当两段视频包含多个相似视频片段并且相似视频片段的位置未能满足基本的时序关系时,上述方法无法有效地对这些随机位置处的相似视频片段进行探测与定位。本文重点研究相似关键帧位置随机时的相似视频片段探测,如图1所示[8],相同颜色代表了相似视频片段,连线代表了相似视频片段之间的对应关系。

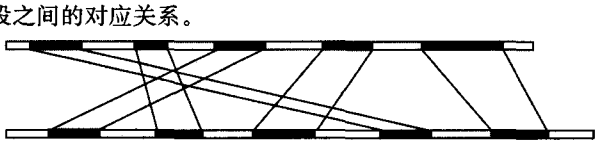

图1 位置随机的相似视频片段

到稿日期:2013-06-14 返修日期:2013-08-09 本文受国家自然科学基金(61201339),中国十二五规划(40901040202)资助。
郭延明(1989—),男,硕士,主要研究方向为多媒体信息系统,E-mail:nudtguoyanming@163.com;谢毓湘(1976—),女,副教授,主要研究方向为多媒体信息系统。

文献[8]对相似关键帧位置随机时的相似视频片段探测进行了详细的讨论，通过构建相似关键帧关联图将寻找视频相似片段的过程转化为寻找密度图的过程。该方法在应用中取得了较好的效果，但是该文献的相似关键帧探测难以达到利用局部特征进行探测的精度，同时该方法需要利用两个阈值进行相似片段的合并与拆分，阈值的选取对最终的探测效果影响较大，并且阈值不好确定。为进一步提高探测的准确性，本文在相似视频片段的探测时，采用基于局部特征的相似关键帧探测方法，并提出了一种新的相似关键帧距离度量方式。

当前研究中广泛使用的图像局部特征的基础之一是2004年David G. Lowe提出的SIFT算法[9]，该算法通过在尺度空间和位置空间中探测特征点，并利用图像的局部区域特征为每个特征点创建128维的描述符向量，解决了尺度变化、光照变化和旋转变化等所带来的问题，在图像匹配领域得到了广泛的应用。Bay[10]等人在SIFT算法的基础上，利用积分图像代替高斯图像，提出了SURF算法，加快了处理进程。近年来，为进一步提高图像匹配的效率，一些学者对基于局部特征的检测算法和描述算法开展了大量的研究，提出了一系列行之有效的算法，如FAST检测算法[11]、BRIEF描述算法[12]、BRISK描述算法[13]和ORB算法[14]等。如何对这些算法进行分析整合，得到一种适用于大容量视频数据处理的方法，是本文需要解决的一个问题。

本文相似视频片段探测方法的具体流程如图2所示。

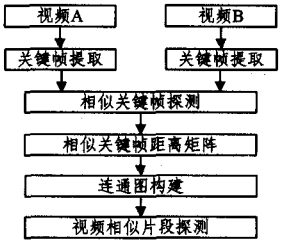

图2 视频相似片段探测框架

该方法首先在视频结构化分析以及关键帧提取的基础上，进行相似关键帧的探测。然后，基于相似关键帧在源视频中分别所处的位置，构建相似关键帧的距离矩阵，矩阵中的元素代表相似关键帧对之间的距离。最后，设定一个阈值，选择矩阵中数值较小的元素构成多个连通子图，每个连通子图对应一个相似视频片段。

2 基于局部特征的相似关键帧探测方法

相似关键帧的识别是相似视频片段探测的基础，为提高相似关键帧识别的精度，本文采用了基于局部特征的相似关键帧探测方法。

近几年，基于相对值比较的特征点检测与描述算法吸引了研究者广泛的关注，其中最具有代表性的是FAST检测算法和BRIEF描述算法，FAST检测算法利用图像的局部灰度值进行特征点检测，具有较高的可重复性和抗噪性；BRIEF描述算法则通过比较特征点附近像素值来构建描述向量，具有较强的独特性。同时，两种算法都具有较快的速率。由于视频所包含的关键帧一般数量巨大，为提高效率，本文采用FAST与BRIEF相结合的方法进行相似关键帧探测。

2.1 FAST特征点检测算法

FAST算法是一种基于图像局部像素灰度值的快速特征点检测算法。该算法的核心思想是：如果在像素点 p 周围，半径为 r 的离散圆周上至少有连续 n 个连续的像素同时比像素点 p 的灰度值亮或暗 ε_d 以上个单位，则认为 p 是一个特征点。

当 $r=3$，$n=12$ 时，利用FAST算法进行特征点检测的算法步骤如下：

1. 选择像素点 p 周围的16个像素点，如图3所示，将 p 点周围的一像素点 x 与 p 的相对位置记为 $p\to x$，并根据式(1)将周围像素点划分为3个子集：

$$S_{p\to x}=\begin{cases}d, & I_{p\to x}\leqslant I_p-t\\ s, & I_p-t<I_{p\to x}<I_p+t\\ b, & I_p+t\leqslant I_{p\to x}\end{cases} \tag{1}$$

其中，t 为灰度差阈值，$S_{p\to x}$ 值为 d 时表示周围点比中心点暗，值为 s 时表示周围点与中心点具有相似灰度值，值为 b 时表示周围点比中心点亮。

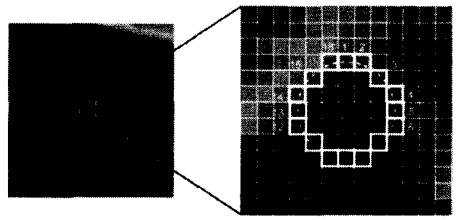

图3 特征点探测模板示意图

2. 依据像素点熵值的正负，利用ID3算法[15]判定候选像素是否为角点。算法在上述3个集合中迭代运算，直到一个子集的熵值是0为止。

3. 对检测出来的角点进行非最大值抑制，建立评估函数：

$$E=\sum_{x\in J}|I(x)-I(p)| \tag{2}$$

其中，J 为中心点 p 周围8个像素点的集合，对3×3邻域内的所有特征点按式(2)计算，保留 E 值最大的点作为最终角点。

相对于其他常用的特征点检测算法，FAST检测子在保证可重复性和抗噪性的同时，提高了检测速度，但其缺陷是缺乏尺度不变性。本文借鉴SIFT算法的思想，通过引入高斯尺度空间来克服该缺陷。

2.2 BRIEF描述算法

BRIEF描述算法的原理[16]是采用一组相同的模板在特征点周围采集像素点，并通过比较周围像素点对的大小来得到一个二进制序列，然后将二进制序列按字节划分为维度，得到特征点的描述符。BRIEF描述符的具体构建步骤如下：

1. 在特征点周围选取 $S\times S$ 大小的图像块 p，在图像块 p 上定义测试函数 η：

$$\eta(p;x,y)=\begin{cases}1, & p(x)<p(y)\\ 0, & \text{otherwise}\end{cases} \tag{3}$$

其中，$p(x)$ 和 $p(y)$ 分别是图像块中像素点 $x=(u_1,v_1)$ 和 $y=(u_2,v_2)$ 处的灰度值。

2. 选择一个包含 n_d 个像素对的模板，对模板上的所有像素对 (x,y) 进行比较，得到BRIEF描述符：

$$f_{n_d}(p)=\sum_{1\leqslant i\leqslant n_d}2^{i-1}\eta(p;x_i,y_i) \tag{4}$$

根据文献[12]，取 $n_d=256$ 即可达到识别速度与精度的平衡，而模板的选取则是在 $G(\mu=0,\sigma^2=S^2/25)$ 的高斯分布

下采样得到，以字节为单位划分维度，每个特征点可以获得256/8=32维的描述符。由于使用单个像素的灰度值容易受到噪声的影响，在进行像素灰度值的比较时，算法采用了高斯平滑的方式对特征点周围的图像进行预处理。

利用BRIEF算法可以快速提取特征点的描述符，并且由于BRIEF描述子为二进制序列，因此特征点描述符的匹配具有较小的计算复杂度，可以达到较快的速率。但是，BRIEF描述子对方向的旋转十分敏感，当图像旋转超过30度时，描述符匹配的精度会急剧下降。

2.3 FAST与BRIEF相结合的相似关键帧探测方法

对于视频而言，相似关键帧之间的差异一般是由于光照、拍摄角度以及后期编辑手段等方面的不同造成的，相似关键帧之间很少有大角度的旋转。基于该特点，本文结合FAST检测算法与BRIEF描述算法速率快的优势，提出了一种将两者相结合的相似关键帧探测方法。

针对FAST检测子的不足，本文引入高斯尺度空间[17]，通过高斯卷积核 $G(x,y,\sigma)$ 对原始图像 $I(x,y)$ 进行平滑，得到图像的尺度空间表示 $L(x,y,\sigma)$，即：

$$L(x,y,\sigma)=G(x,y,\sigma)\otimes I(x,y) \tag{5}$$

其中，$\otimes$ 代表卷积操作，σ 为尺度因子。

然后对尺度空间下的每一层图像利用FAST检测算法进行特征点检测，并利用BRIEF描述算法对特征点进行描述，对特征点描述子进行匹配，根据匹配的结果识别相似关键帧。

与该方法最相近的方法是ORB算法[14]，但是由于忽略了特征点方向的计算，该算法比ORB算法具有更高的效率。

3 基于相似关键帧的视频相似片段探测

为了探测视频之间的相似片段，在探测出相似关键帧之后需要融合相似关键帧所在视频的位置信息，将位置接近的相似关键帧所代表的镜头合并为相似视频片段。本文借鉴文献[8]的思想，提出了一种新的相似关键帧之间的距离度量方式，之后利用矩阵所对应的连通图得到了最终的相似视频片段。具体过程如下：

首先定义相似关键帧对 (r_i,t_j) 与 (r_m,t_n) 的时序距离 L，该距离与相似关键帧所在源视频的位置有关：

$$L=f[d(r_i,r_m),d(t_j,t_n)] \tag{6}$$

其中，r_i 和 r_m 分别代表视频1中的第 i 个和第 m 个关键帧，t_j 和 t_n 分别代表视频2中的第 j 个和第 n 个关键帧。

本文将视频视作二维空间，定义相似关键帧对之间的欧氏距离为最终的时序距离 L，两个处于同一视频中的关键帧 r_i 与 r_m 之间的欧氏距离定义为两个关键帧之间的序号差，即：

$$d(r_i,r_m)=|m-i| \tag{7}$$

$$L=f[d(r_i,r_m),d(t_j,t_n)]=\sqrt{d^2(r_i,r_m)+d^2(t_j,t_n)}$$
$$=\sqrt{|m-i|^2+|n-j|^2} \tag{8}$$

然后，利用上文所探测到的 N 对相似关键帧以及相似关键帧之间的距离，构建两个视频之间 $N\times N$ 的相似关键帧距离矩阵 C，矩阵的元素代表相应关键帧对之间的距离，即 c_{ab} 代表第 a 对相似关键帧 (r_i,t_j) 与第 b 对相似关键帧 (r_m,t_n) 之间的距离。

视频相似关键帧之间的对应关系存在如图4显示的3种情形。

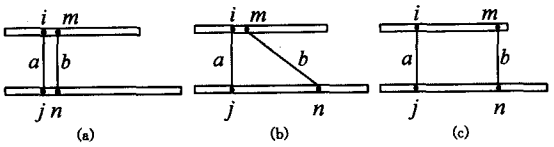

图4 相似关键帧之间的对应关系

图4(a)表示相似关键帧对所涉及的关键帧在原始视频中距离均较小，相应的 c_{ab} 值较小；图4(b)表示相似关键帧对所涉及的关键帧在原始视频中距离一个较大、一个较小，相应的 c_{ab} 值较大；图4(c)表示相似关键帧对所涉及的关键帧在原始视频中距离均较大，相应的 c_{ab} 值较小。

视频之间的相似片段通常是由连续的几个镜头构成，在关键帧层面表示为图4中(a)所表现的形式，对应的相似关键帧对之间的距离 c_{ab} 应该较小，所以本文设定一阈值 T，对矩阵 C 中的每个元素 c_{ab} 做如下处理：

$$c_{ab}=\begin{cases}1, & \text{if } c_{ab}<T\\ 0 & \text{otherwise}\end{cases} \tag{9}$$

经过处理后，矩阵中保留的都是距离较近的相似关键帧对。

相似视频片段的探测具有传递性，即若第 a 对相似关键帧与第 b 对和第 c 对相似关键帧分别同属一组相似视频片段，则第 b 对相似关键帧与第 c 对相似关键帧也同属一组相似视频片段。经过上述的矩阵处理，矩阵中值1代表相应的相似关键帧对距离较近，属于同一组相似视频片段。将 $N\times N$ 的矩阵对应为 N 个节点，当矩阵中某一元素 c_{ab} 为1时，第 a 个节点与第 b 个节点之间存在连边，这样可以得到矩阵 C 所对应的连通图。根据相似视频片段的传递性，矩阵 C 所对应的每个连通图都对应一个相似视频片段，相似视频片段的起止帧由所对应的相似关键帧的起止位置决定。

将寻找相似视频片段的过程转化为寻找矩阵 C 对应的连通图的过程，提高了相似视频片段探测的可行性，而且计算复杂度较小，能够取得较好的效果。

4 实验和结果分析

本文实验可分为相似关键帧探测和相似视频片段探测两部分。

4.1 相似关键帧探测实验

本文所使用的数据来源于哥伦比亚大学DVMM实验室提供的测试数据库[18]，包括600幅从Trecvid数据中选择的视频关键帧，其中150对为相似关键帧，其他150对为非相似关键帧，关键帧的内容涉及军事、政治、体育、风景等各个方面。

本文利用文献[19]提到的查准率和查全率对本文算法进行了评估：

$$\text{查准率}=\frac{\text{所匹配的正确相似关键帧数目}}{\text{所匹配的相似关键帧数目}} \tag{10}$$

$$\text{查全率}=\frac{\text{所匹配的正确相似关键帧数目}}{\text{所有相似关键帧数目}} \tag{11}$$

为验证本文所提算法的效率和精度，对该数据集分别利用本文所提算法、未引入高斯尺度空间的FAST算法和

BRIEF 算法、ORB 算法、SIFT 算法和 SURF 算法进行实验，得到的最终结果如表 1 所列。

表 1　不同算法的相似关键帧探测对比

方法	查准率	查全率	时间(s)
本文算法	96.6%	94.7%	6.4
FAST+BRIEF	95.8%	47.8%	5.2
ORB	95.2%	92.3%	8.3
SIFT	98.6%	98.5%	191.2
SURF	97.3%	96.7%	28.9

从实验结果可以看出，对于视频相似关键帧的探测，当未引入高斯尺度空间时，单纯利用 FAST 和 BRIEF 算法的探测结果大都是没有尺度变化的相似关键帧，虽然具有较高的准确率，但是查全率较低；相对于 ORB 算法，本文所提算法虽然忽视了旋转因素，但是在处理视频帧时反而取得了较好的查准率和查全率，而且处理时间也有所减少，这是由于 ORB 算法利用图像矩来求取特征点主方向存在误差，对于常用的能够保持旋转不变性的算法，特征点主方向的失准会直接改变特征描述子的提取区域，进而影响算法的性能；对于 SIFT 算法和 SURF 算法，虽然相似关键帧探测效果优于本文所提算法，但是时间开销较大。综合考虑，本文所提出的引入高斯尺度空间并将 FAST 与 BRIEF 相结合的方法比较适合视频相似关键帧的探测。

4.2　相似视频片段探测实验

本文选定两段关于 2010 年“××事件”的视频为研究对象，视频长度约为 10 分钟，包含多个处于不同位置的相似视频片段。

首先，通过对视频进行基于局部特征的关键帧提取，最终得到视频 A(包含 87 个关键帧)、视频 B(包含 96 个关键帧)，然后进行相似关键帧探测，得到两个视频，共包含有 25 对相似关键帧，这些相似关键帧所代表的镜头所在原视频中的位置如下：

P_A=[6,7,8,9,10,17,19,20,21,28,30,31,33,34,36,45,47,48,50,72,73,74,76,79,81]

P_B=[61,63,64,67,70,25,26,28,30,10,11,13,14,16,17,40,42,43,44,80,81,83,86,87,88]

由该位置向量可以看出，相邻的相似关键帧所对应的源视频中的位置不一定相邻，并且相似关键帧在两段视频中并没有保持严格的时序关系。本文设定相似关键帧距离阈值 T=5，对视频的相似关键帧距离矩阵进行处理并得到处理后的矩阵所对应的连通图，如图 5 所示。

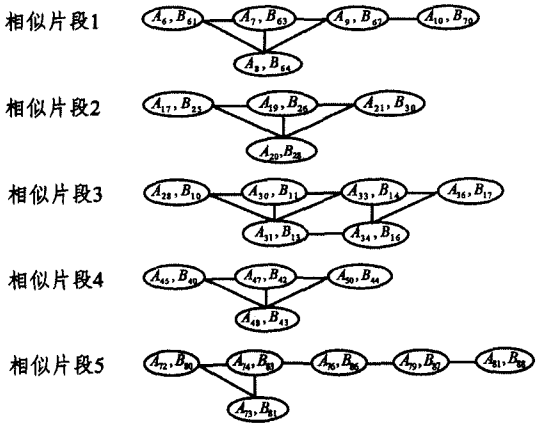

图 5　相似视频片段所对应的连通图

每个连通图对应一个相似视频片段，相似片段的起止位置分别为对应的相似关键帧的起止位置。由实验可以看出视频 A 与视频 B 共存在 5 个相似视频片段，如表 2 所列。

表 2　相似视频片段对应的起止关键帧序号

相似视频片段序号	视频 A 中的起止关键帧	视频 B 中的起止关键帧
1	6—10	61—70
2	17—21	25—30
3	28—36	10—17
4	45—50	40—44
5	72—81	80—88

以第 3 个相似视频片段为例，对应到原始视频中关键帧的对应关系如图 6 所示。通过对其他的类似视频进行处理，可以得到类似的结果。由实验可以看出，本文所提出的算法取得了较好的效果。

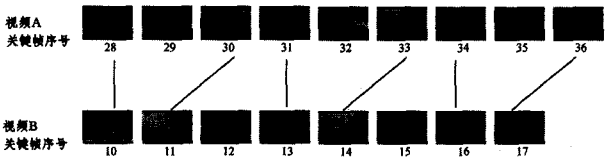

图 6　相似视频片段中的关键帧对应图

结束语　本文通过定义相似关键帧之间的距离，提出了一种有效的相似视频片段检测与定位方法。该方法实施的基础是准确的相似关键帧探测，针对视频相似关键帧之间旋转较小的特点，本文采用了 FAST 检测子与 BRIEF 描述子相结合的方法进行相似关键帧探测，提高了相似关键帧探测的精度和效率。实验结果表明，利用本文所提出的方法可以准确检测出在视频中处于任何位置的相似视频片段。

参 考 文 献

[1] Chen L, Chua T S. A match and tiling approach to content-based video retrieval[C]// Int. Conf. on Multimedia and Expo. 2001: 417-420

[2] Kashino K, Kurozumi T, Murase H. A quick search method for audio and video signals based on histogram pruning[J]. IEEE Transactions on Multimedia, 2003, 5(3): 348-357

[3] Chiu C, Wang H. A novel video matching framework for copy detection[C]// Proc. of the 21th IPPR Conference on Computer Vision, Graphics and Image Processing (CVGIP'2008). 2008

[4] Myers C S, Rabiner L R. A comparative study of several dynamic time-warping algorithms for connected word recognition[J]. The Bell System Technical Journal, 1981, 60(7): 1389-1409

[5] 曹政，卢宝丰，朱明. 一种快速相似视频检索方法[J]. 信息与控制，2010，39(5)：635-639

[6] 陈秀新，贾克斌，邓智批，等. 一种新的鲁棒性视频复制检测算法[J]. 北京工业大学学报，2011，37(5)：691-696

[7] Liu F, Dong D G, Miao X, et al. A fast video clip retrieval algorithm based on VA-File[C]// Electronic Imaging 2004. International Society for Optics and Photonics, 2003: 167-176

[8] Chen T, Jiang S, Chu L, et al. Detection and location of near-duplicate video sub-clips by finding dense subgraphs[C]// Proceedings of the 19th ACM international conference on Multimedia. ACM, 2011: 1173-1176

[9] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International journal of computer vision, 2004, 60(2): 91-110

(下转第 71 页)

的共享子空间表征 u_{n+1} 的计算复杂度方面，本文提出的 LSPSA 的在线扩展的复杂度由 $O((n+1)^3)$ 降为 $O(n^2 \times p)$。这里，我们进一步通过检索实验验证 LSPSA 的在线扩展的有效性。图 3 给出了多特征组合为(fac,fou)且检索数据规模为 2000 时的检索性能。从中可以看出，随着返回列表范围的变化，在线扩展与在线非扩展在检索性能方面取得了非常相似的检索结果。但在检索效率方面，图 4 表明在线扩展的效率是在线非扩展效率的近 2100 倍。从图 4 中也可看出，随着样本量增大，在线扩展算法的计算效率优势将会更加明显。

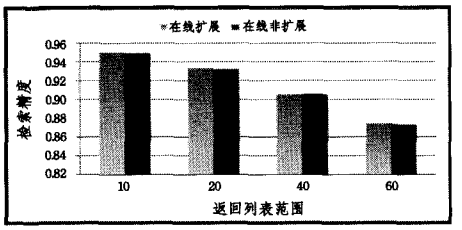

图 3　在线扩展和在线非扩展情况下检索性能比较

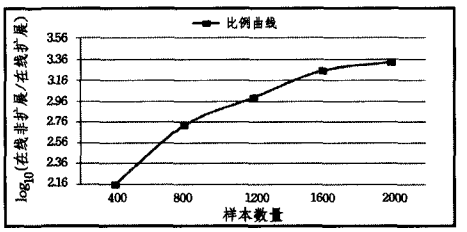

图 4　在线扩展和在线非扩展情况下检索效率比较

结束语　针对多视角分析问题，通过在 MORP 模型的基础上引入图约束，本文提出了一种局部结构保持的共享子空间分析(LSPSA)算法，它在把多视角数据映射到一个共享子空间的同时，保持原始多视角特征空间与共享子空间中的数据具有相近的局部几何结构关系，从而避免多视角数据在共享子空间的过拟合问题。此外，针对新测试样本的在线非扩展问题，本文还提出了一个图模型逼近的在线扩展方法，使得获取新测试样本在共享子空间表征的复杂度大大降低。

参 考 文 献

[1] Estellers V, Gurban M, Thiran J. On Dynamic Stream Weighting for Audio-Visual Speech Recognition [J]. IEEE Transactions on Audio, Speech, and Language Processing, 2012, 20(4): 1145-1157

[2] Li Hao-jie, Wang Xiao-hui, Tang Jin-hui. Combining global and local matching of multiple features for precise item image retrieva[J]. ACM/Springer Multimedia System Journal, 2013, 19(1): 37-49

[3] Kalamaras I, Mademlis A, Malassiotis S, et al. novel framework for retrieval and interactive visualization of multimodal data[J]. Electronic Letters on Computer Vision and Image Analysis, 2013, 12(2): 28-39

[4] Hotelling H. Relations between two sets of variates [J]. Biometrika, 1936, 28(3/4): 321-377

[5] 王惠文. 偏最小二乘回归的线性与非线性方法[M]. 北京：国防工业出版社，1999

[6] Akaho S. A kernel method for canonical correlation analysis [C]// IMPS 2001 International Meeting of Psychometric Society. Feb. 2007

[7] 彭岩，张道强. 半监督典型相关分析算法[J]. 软件学报，2008，19(11)：2822-2832

[8] Rosipal R, Trejo L J. Kernel partial least squares regression in reproducing kernel Hilbert space[J]. Journal of Machine Learning Research, 2001(2): 97-123

[9] 王珏，周志华，周傲英. 机器学习及其应用[M]. 北京：清华大学出版社，2006

[10] Sharma A, Kumar A, Daume H III, et al. Generalized MultiviewAnalysis: A Discriminative Latent Space[C]// IEEE Conference on Computer Vision and Pattern Recognition. 2012: 2160-2167

[11] Chen Yao-nan, Lin H T. Feature-aware label space dimension reduction for multi-label classication[C]// Advances in Neural Information Processing Systems. 2012, 25: 1538-1546

[12] Yu Shi-peng, Yu Kai, Tresp V, et al. Multi-Output regularized feature projection[J]. IEEE Conference on knowledge and data engineering, 2006, 18(12): 1600-1613

[13] http://www.ics.uci.edu/~mlearn/MLSummary.html

[14] Liu Nan, Zhao Yao, Zhu Zhen-feng, et al. Exploiting Visual-Audio-Textual Characteristics for Automatic TV Commercial Block Detection and Segmentation[J]. IEEE Transaction on Multimedia, 2011, 13(5): 961-973

(上接第 56 页)

[10] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up robust features[M]. Computer Vision-ECCV 2006. Springer Berlin Heidelberg, 2006: 404-417

[11] Rosten E, Drummond T. Machine learning for high-speed corner detection[M]// Computer Vision-ECCV 2006. Springer Berlin Heidelberg, 2006: 430-443

[12] Calonder M, Lepetit V, Strecha C, et al. BRIEF: binary robust independent elementary features[M]// Computer Vision-ECCV 2010. Springer Berlin Heidelberg, 2010: 778-792

[13] Leutenegger S, Chli M, Siegwart R Y. BRISK: Binary robust invariant scalable keypoints[C]// 2011 IEEE International Conference on Computer Vision (ICCV). IEEE, 2011: 2548-2555

[14] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF[C]// 2011 IEEE International Conference on Computer Vision (ICCV). IEEE, 2011: 2564-2571

[15] Quinlan J R. Induction of decision trees[J]. Machine learning, 1986, 1(1): 81-106

[16] 邓黎，王春鸿，饶长辉. 基于二进制稳健独立基元特征的扩展目标快速检测[J]. 中国激光，2012，39(B06)：327-331

[17] 郭莉莎，等. 基于多尺度 FAST-9 的图像快速匹配算法[J]. 计算机工程，2012，38(12)

[18] http://www.ee.columbia.edu/ln/dvmm/researchProjects/FeatureExtraction/NearDuplicateByParts/INDDetection.html, 2004

[19] Tan H K, Ngo C W, Chua T S. Efficient mining of multiple partial near-duplicate alignments by temporal network[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(11): 1486-1498