

一种基于模糊信息粒化的视频时空显著单元提取方法

郎丛妍, 须 德, 李 兵

(北京交通大学计算机研究所, 北京 100044)

摘 要: 提出一种基于模糊信息粒化的视频时空显著单元提取方法, 为视频分析及检索等高层应用提供一个有效的内容表示模式. 本文首先提出了一种类相关的特征粒化方法, 粒化后的模糊粒特征简化了分类关系且在一定程度上解决了感知主观性问题, 因而通过简单的分类器可以有效地提取空域中具有高视觉感知显著度的区域(简称为显著区域); 其次, 通过对显著区域的时域一致性分析提取视频序列中时域连续的显著区域集合, 定义为时空显著单元. 提取的时空显著单元能作为一种较为通用的语义级内容表示模式. 实验结果分别从时域和空域两个方面验证了本文方法的有效性.

关键词: 显著区域; 模糊信息粒化; 内容表示; 视觉注意力机制

中图分类号: TN911. 73 **文献标识码:** A **文章编号:** 0372-2112 (2007) 10-2023-06

Using Fuzzy Information Granulation for Spatio-temporal Salient Unit Detection in Video Sequences

LANG Cong-yan, XU De, LI Bing

(Institute of Computer Science and Engineering, Beijing Jiaotong University, Beijing 100044, China)

Abstract: In this paper, we propose an approach to automatic extraction of spatio-temporal salient unit from video sequences. Firstly, a class related feature granulation algorithm is proposed, which can map original feature space to concept space based on fuzzy information granular. To detect spatial salient regions, segmented homogenous regions are classified according their prominent importance. Then, salient regions are tracked and the results of tracking are sequences of temporal coherent regions, called spatio-temporal salient unit. Experimental results verify efficiency of proposed approach in spatial and temporal aspects.

Key words: salient region; fuzzy information granulation; content representation; visual attention mechanism

1 引言

近十年来随着多媒体和网络技术的发展, 基于内容的视频检索和分析成为研究的热点问题. 视频信息的快速增长迫切需要一个高性能的内容表示模式, 该模式应不仅能有效地缩减无关信息而且能满足基于内容的视频信息操作和管理需求. 在已有的视频检索和分析系统中, 大多数实现方法^[1]是提取视频序列的底层视觉特征(如颜色、纹理以及运动等特征)进行全局相似度匹配. 然而, 视频信息语义内容的复杂性、模糊性以及主观性使得底层特征和高级语义内容之间的鸿沟问题没有很好地解决, 如何建立有效的视频内容表示模式是一个有待研究的重要课题.

从已发表的研究成果来看, 视频序列的内容表示方法可以分为三类: (1) 基于文本的描述方法. 对于视频信息内容手工或自动地进行文本标注. 手工标注工作量巨

大, 这种方法目前已经很少采用. 自动标注的实现依赖于图像或视频序列的语义内容分析, 底层特征和高层语义内容间的鸿沟是实现内容自动标注的一个最大的困难. (2) 基于关键帧的表示方法^[2]. 这是目前关于视频镜头的一个主要的内容表示模式. 该方法的主要思想是提取少量的但是能最大程度地表示视频序列主要内容的帧(关键帧)的集合来表示视频序列. 这种方法虽然能够建立一个压缩的内容表示模式, 但视频序列的内容分析仍然需要进一步的内容建模. (3) 基于对象的内容描述. 视频对象的提取是实现语义内容表示的一个理想的解决方法, 然而, 自然场景中的对象提取和跟踪仍然面临着许多困难, 如对象的遮挡和形变、背景干扰以及运动估计等, 所以, 视频对象提取目前不能实现自动的方法, 需要人的交互. 鉴于人类视觉系统对运动信息的敏感性, 目前研究较多的是运动对象的提取^[3]. 用运动对象来表示视频序列的内容在特定的应用中能建立有效的

内容表示模式, 如对运动轨迹的分析等。

针对现有方法的缺点, 本文以镜头为分析对象, 研究实现语义内容表示模式的自动提取方法, 使之不仅能有效地剔除无关信息, 同时还具有较高的语义性, 从而能进一步提高基于内容的视频检索和分析系统的性能。L. Itti 等人^[4]首先将注意力模型运用于自然图片的场景分析中, 为注意力机制在人工智能中的应用创建了一个好的开始点。虽然注意力机制在认知科学领域还存在分歧, 但是已有的研究成果对机器视觉的研究有着非常重要的指导。基于以上事实, 本文提出时空显著单元概念, 结合人的视觉感知特征, 提取一个镜头内的区域级语义内容——时空显著单元。

2 时空显著单元提取

视频序列在时空域中包含了大量的冗余信息或与感知无关的信息。视觉注意机制(Visual attention)是人类感知外界信号的一个重要过程, 能够对视频序列的大量信息进行选择性处理。信息的选择处理结果使一部分区域的信息处于眼睛中央凹区域位置, 这部分信息由视觉系统进一步加以处理因而要比其它区域的信息具有更高的重要度, 即高显著度。图像中具有高显著度的像素集合称为显著区域(Salient region)。因为视频序列具有时序性, 所以本文提出时空显著单元(Spatio-temporal salient unit)的概念。它是对视频镜头进行时空域分割, 得到的连续时域的显著区域集合。

2.1 空间显著区域提取

2.1.1 预处理

本文首先将视频帧进行颜色同质区域分割, 以分割后的同质区域为单位分析显著度。均值漂移聚类方法^[5]能实现无参数的聚类分析, 在图像场景的应用中得到了很好的分割效果, 所以本文采用该方法实现颜色同质区域的划分, 颜色特征为 LUV 空间的三个分量值。

2.1.2 同质区域特征提取

(1) 区域对比度(CSr)

对比度是影响感知注意力一个重要因素, 可以计算为:

$$CSr(R_i) = \sum_{m=1}^{n_r} |I^*(R_i) - I^*(R_m)| \quad (1)$$

其中, $I^*(R_i)$ 表示区域 R_i 亮度的均值, $I^*(R_m)$ 是区域 R_i 的第 m 个相邻区域的亮度均值, n_r 是和区域 R_i 相邻的区域个数。

(2) 纹理密度(OGr)

区域的纹理特征通过 Gabor 滤波得到, 首先对图像的灰度图分别计算 $F = \{2.00, 5.03, 8.00\}$ 三个中心频率, $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ 四个方向的 Gabor 滤波, 累加

四个方向子图并经归一化得到方向显著图 \bar{I} , 对原图中预处理后的每个同质区域按下式计算其纹理密度:

$$OGr(R_i) = \frac{\bar{V}_{\text{pixel}}(R_i)}{N_{\text{pixel}}(R_i)} \quad (2)$$

其中, $N_{\text{pixel}}(R_i)$ 表示原图中区域 R_i 包含的像素总数, $\bar{V}_{\text{pixel}}(R_i)$ 表示 \bar{I} 图中区域 R_i 包含的像素点的值的和。

(3) 区域形状显著值(SIr)

区域 R_i 的形状显著值通过区域边缘像素个数和区域像素总数的比值 $Slr(R_i)$ 来计算:

$$Slr(R_i) = \frac{N_{\text{edge}}(R_i)}{N_{\text{pixel}}(R_i)} \quad (3)$$

其中, $N_{\text{edge}}(R_i)$ 表示区域 R_i 包含的边缘像素点数, 通常比值高的区域显著度高。

(4) 运动显著值(MIr)

运动信息对描述视觉内容具有重要的作用, 因而运动一致性度量是区域显著性分析的一个重要因素。首先计算区域中运动矢量的幅度均值, 记为 $IM(R_i)$ 。文献^[6]中定义的运动一致性特征在视频内容变化分析方面取得了较好的实验结果。该方法计算每个块分别在时域和空域上一个领域范围内的 8-bin 运动矢量方向直方图, 采用熵作为直方图的分布测度, 该测度可以用来描述运动一致性, 记为 $MC(R_i)$ 。则本文区域运动显著值计算为:

$$Mir(R_i) = IM(R_i) \cdot MC(R_i) \quad (4)$$

2.1.3 基于模糊信息粒化的特征空间映射

由于显著度是一个主观概念, 以往用精确数学方法并不能有效地分析显著度, 所以本文首次尝试利用模糊粒度理论将原始特征空间影射到概念级模糊粒特征空间, 对粒化后的特征进行分类实现显著区域提取。根据 2.1.1 和 2.1.2, 每个视频帧经过预处理后, 包含的同质区域作为原始特征空间中的一个样本点, 下面介绍本文提出的类相关特征粒化算法。

经实验验证对于显著区域分类问题对函数形式不敏感, 所以粒化算法的关键问题是给定模糊函数的形式后确定隶属函数的参数。本文初始隶属函数采用高斯函数形式:

$$\mu(x) = \exp[-(x - \bar{x})^2 / 2\sigma^2] \quad (5)$$

因为高斯函数形式具有较好的通用性和连续性, 符合本文的特征描述特性。粒化算法主要分两步, 首先从数据分布特性来建立初始的粒化特征表示, 然后在给定训练集中根据后验概率分布情况对初始粒化结果进行优化, 找到最优粒化形式。

从数据分布方面考虑, 评价模糊集合的划分标准可通过该集合包含信息的泛化性和语义性两方面, 在文献^[7]中定义了模糊集合的性能指标(Performance Irr

dex) Q 为:

$$Q = E(\tilde{A}) / Sp(\tilde{A}) \quad (6)$$

其中, $E(\tilde{A}) = \sum_{k=1}^n |\tilde{A}(x_k)|^2$, $Sp(\tilde{A}) = \text{card}(\text{supp}(\tilde{A}))$.

$E(\tilde{A})$ 是模糊集合 \tilde{A} 的能量值, 表示隶属函数的泛化性; $Sp(\tilde{A})$ 表示模糊集合 \tilde{A} 的语义性度量, $\text{supp}(\tilde{A})$ 是模糊集合 \tilde{A} 的支持集. 通过最大化指标 Q 确定细粒度 (fine-grain) 的初始隶属函数参数 σ 值.

由于不同的特征对分类的重要度不同, 所以下一步需要结合分类性能对每个特征的粒化结果分别进行优化处理. 设特征 f_i 模糊粒化后的结果为 G_i , $G_i = \{g_{i1}, \dots, g_{im_i}\}$ ($i = 1, \dots, n$), m_i 表示第 i 个特征包含的隶属函数个数. 本文的显著区域提取属于两分类问题, 即每帧的同质区域经过分类后由显著区域和非显著区域组成, 类别集合为 $\{C_{CR}, C_{\overline{SR}}\}$. 本文根据后验概率的分布偏离程度作为优化函数分别对每个特征的初始粒化结果进行优化, 使最后得到的模糊粒特征空间能够最佳地区分区域的显著度. 相对熵能很好地度量两个概率分布的偏离程度, 考虑到对称性因素, 根据相对熵定义判别熵作为优化函数, 其定义如下:

$$W(P_A, P_B) = D(P_A, P_B) + D(P_B, P_A) \quad (7)$$

$$D(P_A, P_B) = \sum_{x \in X} P_A \log \left(\frac{P_A}{P_B} \right) \quad (8)$$

其中, $P_A = p(C_{SR} | G_i)$, $P_B = p(C_{\overline{SR}} | G_i)$. 判别熵的计算需要确定三个密度 $P(C_{SR})$, $P(G_i | C_{SR})$, $P(G_i | C_{\overline{SR}})$, 对给定的训练集, 假设 $P(C_{SR}) = P(C_{\overline{SR}})$, 类条件密度可以从训练集中估计得到, 对训练样本的每个特征计算 m_i 个 bin 的模糊直方图, 每个 bin 的直方图值通过属于同一类中样本的隶属度求和计算:

$$\text{His}(G_i)_j = \sum_{x \in G_i} \mu_j(x), j = 1, \dots, m_i \quad (9)$$

$\mu_j(x)$ 表示模糊集合 g_j 中特征点 x 的隶属度, 归一化后的直方图近似类条件密度.

2.1.4 基于贝叶斯分类的显著区域提取

粒化后的特征空间有效地简化了模糊特征和类别之间的关系, 所以我们采用简单的分类器实现对显著区域的分类. 分类目的是对每帧的区域根据其显著度分配显著值, 提取显著度高的区域. 区域原始特征 $F = (f_1, f_2, f_3, f_4)$ ($n = 4$) 经模糊粒化后对应的模糊特征表示为 $G = (\{g_{11}, \dots, g_{1m_1}; \dots, g_{n1}, \dots, g_{nm_n}\})$. 为了预测区域的类别我们需要计算后验概率 $\text{Pr}(C_k | f_1, \dots, f_n)$ $k = SR, \overline{SR}$. 假设特征具有独立性, 由于粒化算法针对每个特征分别进行, 所以粒化后不改变其独立性, 根据贝叶斯理论后验概率可以计算为:

$$\text{Pr}(C_k | G) = \frac{\text{Pr}(C_k) \prod_{i=1}^n P(G_i | C_k)}{p(G)} \quad (10)$$

类条件密度 $P(G_i | C_k)$ 通过训练样本集中模糊粒度直方图估计得到. 根据最大后验概率(MAP) 确定区域的类标号, 区域 R_i 满足 $P(C_{SR} | G_i) > P(C_{\overline{SR}} | G_i)$ 属于显著区域. 另外, 如果一个区域其所有相邻区域均为显著区域, 则合并该区域和相邻区域作为一个显著区域.

2.2 时空显著单元提取

假设在一个时间窗(设时间窗长度 $l = 5$ 帧)内各帧内容相似, 采用 2.1.4 节中介绍的显著区域提取方法对时间窗的第一帧提取显著区域, 窗内后续各帧的显著区域通过区域跟踪得到. 根据显著度能量值确定每个时间窗的窗内时空显著单元. 如果相邻时间窗的窗内显著单元相似, 则合并成一个时空显著单元.

以下阐述区域跟踪和合并显著单元的方法. 时间窗内的第一帧中的每个显著区域在该时间窗内后续各帧中进行区域相似度匹配, 即根据式(11)(12)计算相似度和显著性能量值:

$$\text{Sim}(SR_i^t, SR_j^s) = \frac{d(SR_i^t, SR_j^s)}{1 + d_{ij}^\alpha} \quad (11)$$

$$E(SR_t) = \sum_{j=1}^l \text{Sim}(SR_t^i, SR_j^s) \cdot e^{-\alpha} \quad (12)$$

式中 $d(SR_i^t, SR_j^s)$ 是帧 f_t 和帧 f_s 中两个显著区域的特征向量的欧氏距离, 区域特征包括 24 维颜色直方图(HSV 颜色空间, H 分为 8 份, S 分为 3 份). d_{ij} 表示两个区域中心点间的距离, 参数 α 是惩罚因子权重系数, 取经验值 $\alpha = 0.65$. 满足 $E_{R_i} > T_e$ 的显著区域构成该时间窗的窗内显著单元 ($T_e = 0.5$).

经过时间窗内的一致性分析, 得到每个时间窗的窗内显著单元之后, 需要对时域相邻且包含相似显著区域的窗内显著单元进行合并. 时间窗合并是依据各时间窗的首帧中的显著区域相似性, 合并后的显著单元其时域分割点能够对应内容变化不连续点, 实验中也验证了时空显著单元的时域分割性能. 考虑到不同时间窗中由于遮挡或噪声的影响, 各自包含的显著区域可能出现多匹配情况, 即后续帧中可能会有多个区域对应前一帧中的一个显著区域, 我们通过 kolmogorov-Smirnov(K-S) 检验^[8] 确定如何进行时间窗合并. SR_s^{t+1} 表示 $\{SR_j^{t+1}, SR_k^{t+1}, SR_j^{t+1} \cup SR_k^{t+1}\}$ 中一个显著区域, 假设区域 SR_i^t 和 SR_s^{t+1} 属于同一个时空显著单元 $STR_k^{(t, t')}$, 零假设 H_0 ; 备择假设 H_1 表示两个区域不属于同一个时空显著单元, 统计量, 记为 $D(SR_i^t, SR_s^{t+1})$.

$$D(SR_i^t, SR_s^{t+1}) = \max_x |F_1(x) - F_2(x)| \quad (13)$$

其中, $F_1(x) = \int_{-\infty}^x p_1(x) dx$, $F_2(x) = \int_{-\infty}^x p_2(x) dx$.

密度函数分别通过区域像素亮度的高斯分布估计得到, 即 $p_1(x) = N(x, \mu_1, \sigma_1)$, $p_2(x) = N(x, \mu_2, \sigma_2)$. 根

据定义的统计量, 假设检验表示为:

$$\begin{cases} H_0: D(SR_i^t, SR_s^{t+1}) = 0 \\ H_1: D(SR_i^t, SR_s^{t+1}) > 0 \end{cases} \quad (14)$$

3 实验研究

3.1 特征粒化实验结果

以对比度为例, 给出粒化过程, 7 个初始隶属函数的中心点集合为: $\{0.081, 0.152, 0.213, 0.233, 0.564, 0.816, 0.985\}$, 图 1 表示横轴为参数 σ , 纵轴 (a) 图中为能量值, (b) 图中是性能指标值. 图中分别对应以上中心点集合的 7 个隶属函数参数 σ 变化时相应的能量值变化曲线 (a) 和性能指标曲线 (b).

能量值曲线和性能指标曲线重叠越多表示对应的隶属函数重叠度越大, 从结果看出, 由于初始粒化结果中各隶属函数间存在很大重叠度, 例如图 1 中两曲线中都有两组曲线分布基本相同, 则表示对应的隶属函数重叠度很大. 经过类相关的优化合并后, 当隶属函数个数等于四时对应最大判别熵值 $W_4 = 0.783$, 相应的四个中心点为 $\{0.081, 0.211, 0.564, 0.860\}$. 对比度特征经过粒化后表示为四个模糊集合, 粒化结果能达到最好的分类性能. 下一节详细介绍了基于粒化后特征空间的显著区域提取结果.

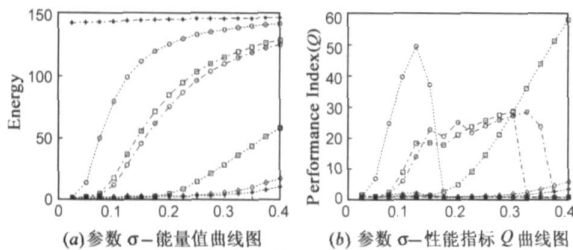


图 1

3.2 空间显著区域提取结果

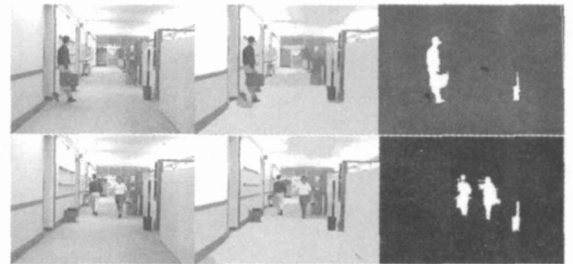
本文对时空显著单元提取方法分别从空域和时域分割性能评价. 由于提取的时空显著单元是时间段内各帧中显著区域的集合, 所以其空域分割性由显著区域体现. 以下是显著区域的实验结果以及与现有显著区域提取方法的比较, 时域分割性能在 3.3 节中讨论.

在图像预处理中, 我们采用参数为 $h_s = 7, h_t = 6.5$ 进行颜色同质区域划分. 图 2 是 hall monitor 视频序列和电影 Friends 中两个时间窗首帧的提取结果.

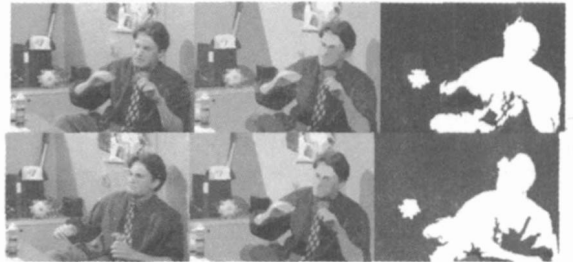
视频序列 hall monitor 常用于测试视频对象提取, 虽然是静态背景, 但由于光照变化影响, 背景部分包含了許多由对比度变化引起的噪声区域. 图 2(a) 是该序列的第 31 和 141 帧的显著区域提取结果, 从结果可看出, 提取的显著区域能较好地对应显著高的区域, 同时去除了背景中光照变化的区域. 图 2(b) 中的显著区域提取结果包含一个背景中静态物体 (橘黄色花饰).

因为该物体在镜头中具有高的颜色对比度且在整个镜头中的显著度连续性较高, 所以这个结果是符合视觉感知特性的.

为了更好地验证显著区域提取方法的有效性, 本文实现了文献 [9] 的方法, 以此方法作实验对比主要有两个原因: (1) 该文献是最近比较新的一种图像显著区域提取方法, 该方法提取的显著区域作为图像的内容表示模式来进行图像检索; (2) 文献 [9] 提出的方法是采用 SVM 分类器在原始区域特征空间中分析区域显著度, 而这点能和本文中的粒化特征形成好的比较. 图 3 是两种方法的比较结果.

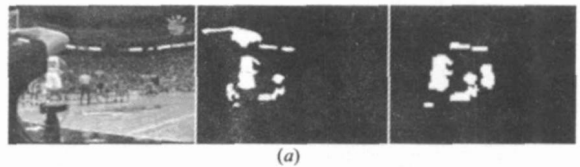


(a) 《hall monitor》中 31# 和 141# 帧: 原图、同质区域分割结果、显著区域分割结果



(b) 《Friends》中连续两帧: 原图、同质区域分割结果、显著区域分割结果

图 2 显著区域提取结果



(a)



(b)

图 3 文献 [9] 和本文的显著区域提取两组结果比较: 每组结果从左到右分别为原图、文献 [9] 显著区域提取结果和本文显著区域提取结果

从比较结果可看出, 对于噪声区域尤其体育视频中的场外区域, 由于其颜色等视觉特征具有高的显著性, 所以采用文献 [9] 中的方法会误分类成显著区域, 而本文算法基于中层特征分析, 除视觉特征对比度外, 其形状、纹理密度等因素结合时间一致性的分析在去除场外的噪声区域有良好的性能. 对于复杂场景中的区域, 我们的算法能更

加符合感知度的分析。此外,由于文献[9]中的方法采用原始特征空间,支持向量机对于显著区域邻近的区域易出现误分类,我们的提取算法更加精确,区域具有很好的轮廓性,更接近实际物体

本文显著区域提取算法的误分类主要存在于连续出现的噪声区域,显著区域间连续出现的噪声区域往往也同样具有高的特征值,所以其显著度较高。图 4 中两个主要显著区域中间的背景区域误分类为显著区域



图 4 《Forrest Gump》中视频帧显著区域提取结果

3.3 时空显著单元时域分割性能评价

实验数据分别选自电影(《Forrest Gump》、《Friends》)、篮球和足球体育视频和监控测试序列(《hall monitor》)三种类型的视频序列。图 5 是《hall monitor》中提取的四个时空显著单元,图中分别用每个显著单元的首末帧和相应的显著区域表示。从实验结果看出,时空显著单元的时域边界能较好地对应镜头内容变化的不连续点。

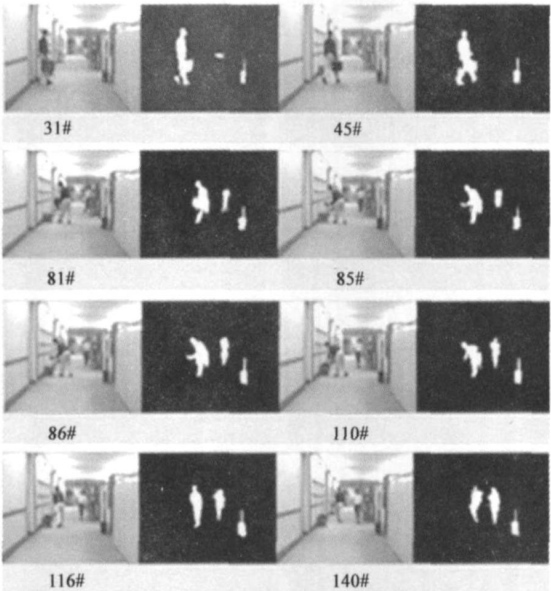


图 5 《hall monitor》序列中四个时空显著单元提取结果:每组分别为该单元的首帧和末帧及其对应的显著区域

由于感知本身具有一定的主观性,目前对镜头中内容变化检测还没有标准的客观评价方法。Zack 等人^[10]为检测划分最小语义单元(smallest meaningful unit)提供了一个较为客观的评价方法,我们采用该评价方法作为显著单元时域分割性能的评价方法。该方法让不同的观测者对短时序列中感知到的最小内容单元的边界处做判断然后综合所有测试者的感知划分结果,作为评价依据。依照该方法,本文实验中邀请 10 位测试

者对视频序列中内容变化点也即最小内容单元边界处标记,对标记结果进行直方图统计,局部峰值点对应的帧号认为是视觉心理学中视频序列感知过程中的最小内容单元的边界,序列包含子最小内容单元数记为 N_d ,本文方法中镜头包含的显著单元数记为 N_r 。表 1 是对实验数据进行准确度(Precision)和查全率(Recall)的统计结果。

表 1 时空显著单元时域分割评价结果

序列	帧数	N_d	N_r	Recall	Precision
Hall monitor	217	12	15	0.92	0.73
Friends	331	18	22	0.94	0.77
Forrest Gump	349	11	9	0.82	1.00
Basketball Games	196	9	12	1.00	0.75
Soccer Games	236	11	12	0.91	0.83

由实验结果可以看出,显著单元在分割最小内容单元的精度能够达到较好的满意度。由于体育节目尤其足球场地背景的唯一因素,本文方法中显著区域提取的准确率较高,所以对应显著单元时域分割性能较好。电影《Forrest Gump》我们选取的镜头是背景复杂的序列,由于显著区域中存在一些噪声区域所以对应显著单元提取容易产生漏检情况,也就是我们提取的显著单元包含了多个最小内容单元。

4 结论

本文提出了一种时空显著单元提取算法,提取的时空显著单元能够有效地表示视频内容。首先在提取空域显著区域过程中,类相关的特征粒化算法结合分类性能和数据分布两个方面将原始特征空间映射到概念级特征空间,从而采用简单的分类器完成显著区域提取。其次,考虑到显著区域跟踪过程中受噪声、光照变化等影响,本文提出的时间窗级区域跟踪和合并方法能在一定程度上解决不连续性和遮挡问题。实验结果表明:时空显著单元在时空域中具有很好的语义内容表征性。下一步的工作,主要是通过优化粒化特征空间的分类规则解决这一问题。

参考文献:

- [1] Sameer Antani, Rangachar Kasturi, Ramesh Jain. A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video[J]. Pattern Recognition, 2002, 35(4): 945-965.
- [2] H J Zhang, J Wu, D Zhong, S Smoliar. An integrated system for content based video retrieval and browsing [J]. Pattern Recognition, 1997, 30(4): 643-658.
- [3] A S Ogale, C Femuller, Y A Ioimono s. Motion segmentation using occlusions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(6): 988-992.

- [4] Laurent Liti, Christof Koch, Ernst Niebur. A model of saliency based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11): 1254– 1259.
- [5] D Comaniciu, P Meer. Mean shift: a robust approach toward feature space analysis[J]. IEEE Trans Pattern Analysis Machine Intelligence, 2002, 24(5): 603– 619.
- [6] Y F Ma, H J Zhang. A model of motion attention for video skimming[A]. In Proc of ICIP (2002) [C]. New York: IEEE Press, 2002. 129– 132.
- [7] W Pedrycz, George Vukovich. On elicitation of membership functions[J]. IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans, 2002, 32(6): 761– 767.
- [8] R V on Mises. Mathematical Theory of Probability and Statistics [M]. New York: Academic Press, 1964.
- [9] Byoung Chul Ko, Soo Yenong Kwak, Hyeran Byun. SVM-based salient regions extraction method for image retrieval [A]. In Proc of the 17th International Conference on Pattern Recognition(ICPR' 04) [C]. Cambridge, UK: IEEE Press, 2004. 977– 980.
- [10] J Zacks, B Tversky, G Iyer. Perceiving, remembering, and communicating structure in events[J]. Journal of Experimental Psychology: General, 2001, 130(1): 29– 58.

作者简介:



郎丛妍 女, 1978 年出生于内蒙古呼和浩特, 2006 年获北京交通大学计算机应用技术专业博士学位, 主要研究方向为视觉感知信息表示和视频内容分析, 现为北京交通大学讲师。
E-mail: cylang@bjtu.edu.cn



须 德 男, 1944 年出生于江苏常州, 北京交通大学教授、博士生导师, 主要研究方向为数据库系统及应用和多媒体信息处理。
E-mail: dxu@bjtu.edu.cn