

基于压缩域编码长度的视频显著性检测

张兆丰 吴泽民 杜 麟 胡 磊

(中国人民解放军理工大学通信工程学院 南京 210007)

摘 要 生物学研究表明,人会明显地注意视频中的运动目标。为模拟该特性并快速完成视频显著图的计算,提出一种压缩域时空显著度检测方法(Temporal-Spatial Saliency in Compress Domain model, TS2CD)。分别利用 H. 264 视频中对宏块的残差编码长度和运动矢量编码长度模拟人眼的显著性刺激强度,从而得到视频显著特征。通过线性的加权融合算法,综合两种编码长度得到的空域显著图和时域显著图,得到最终的视频显著图。在 3 个公开数据库上的实验表明,TS2CD 算法是当前性能最优的方法。

关键词 视频显著性,压缩域,时空显著度,编码长度

中图分类号 TN919.85 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.10.056

Video Saliency Detection Based on Compressed Domain Coding Length

ZHANG Zhao-feng WU Ze-min DU Lin HU Lei

(College of Communications Engineering, PLA University of Science and Technology, Nanjing 210007, China)

Abstract Biology studies show that people will pay much attention to the moving object when they watch a video. In order to simulate this feature and detect the salient region rapidly, we proposed the temporal-spatial saliency in compress domain model(TS2CD). By respectively using H. 264 residual coding length and motion vector coding length, we simulated the salient stimulus intensity and then got video saliency features. Finally, we used the linear weighted fusion algorithm to get the final video saliency maps. Experimental results on three public datasets demonstrate that our model outperforms state-of-the-art methods.

Keywords Video saliency, Compress domain, Temporal-Spatial saliency, Coding length

1 引言

视觉显著性正受到人们越来越多的关注。作为图像视频的预处理过程,显著性检测被广泛应用于计算机视觉的各个领域,例如目标跟踪^[1]、目标识别^[2]、视频压缩编码^[3]和视频摘要^[4-5]。目前,大部分视频显著性检测算法基于像素域^[6-8],这类模型在进行显著性检测时必须将压缩的视频进行解码,然后提取相关的视觉特征。这些过程导致基于像素域的视频显著性检测算法的计算过程复杂、耗时严重。为了解决上述问题,基于压缩域的视频显著性检测算法受到人们重视,它只需要抽取编码数据中的部分信息来形成视频的频域特征,大幅降低了运算时间。Fang 等人^[9]提取码流信息中的 DCT 系数及运动矢量(MV),然后利用“中心-周围”差异原则对每个特征分别获得时域和空域显著性区域,再利用融合算法得到最后的显著图。Khatoonabadi 等人^[10]通过视频压缩后的比特数来衡量显著性,并结合马尔可夫随机场来获得更好的效果。

目前,基于压缩域的视频显著性检测还存在两个问题:

1)相对于像素域算法,压缩域算法的检测精度普遍较低,可以

应用于眼动点识别^[11]以提取最高显著度的区域,而在整帧视频的显著性检测上还需进一步提高精度;2)对不同场景的普适性较弱。此外,基于压缩域的显著性检测算法目前仍处于研究初期,可对比算法相对较少。

针对以上两个问题,受文献^[10]利用视频压缩后的数据比特数来衡量显著性的启发,本文提出利用视频编码长度和运动矢量编码长度分别构造显著性信息测度。利用频域空时特征融合的方法,将通过两种特征得到的显著图融合成最终的显著图。在 3 个公开数据集上的实验表明,本文的视频显著性检测模型有较强的鲁棒性,其检测性能与目前的主流算法相比具有明显优势。

2 相关工作

视觉注意机制是人类视觉系统中的重要一环,这种有选择性的注意机制通常被认为是由刺激驱动或者任务驱动的,在显著性检测中对应为自底向上和自顶向下的检测模型。自 Treisman 等人提出著名的特征整合理论(Feature Integration Theory, FIT)以来,人们从心理学^[12-13]、生物学^[14]以及计算机

到稿日期:2016-09-19 返修日期:2016-12-22 本文受国家自然科学基金(61501509)资助。

张兆丰(1991—),男,硕士生,主要研究方向为图像处理,E-mail:zhangzhaofeng5@163.com;吴泽民(1973—),男,博士,副教授,主要研究方向为信息融合、图像处理;杜 麟(1992—),男,硕士生,主要研究方向为视频传输保障;胡 磊(1985—),男,博士,讲师,主要研究方向为压缩感知、视频信息处理。

视觉^[15-16]等不同的学科角度来研究人类的视觉注意机制。1998年,Itti等人^[15]根据当时已知的视觉注意理论研究成果,首次提出了针对图像数据的视觉显著性计算方法。该方法采用中心边缘算子,分别计算颜色、亮度、方向3个显著图来然后通过线性加权的方式叠加3个显著图,形成最终的显著图,以检测图像中视觉显著性对象的位置。虽然图像显著性检测算法都可以将检测结果直接应用于单帧的视频显著性区域检测,但是心理学研究表明,与颜色、亮度、方向和纹理等静态信号相比,人类视觉系统对运动信息也很敏感,单纯地把视频每一帧的图像显著性检测结果进行时间平均,将丢失重要的运动信息,因此其并不适用于视频显著性检测。

由于运动目标更能吸引人类视觉系统的注意,国内外许多学者对如何度量时间显著性进行了研究。Kim等人^[17]通过连续两帧之间的亮度差来获得时间梯度信息,然后计算中心区域和周边邻域的时间梯度差的绝对值的总和,从而获得时间显著性;Meur等人^[17]通过一种分层的块匹配方法,首先计算块的局部运动矢量,然后采用六参数仿射模型估计出全局运动矢量,最后利用全局运动矢量和局部运动矢量的差异估计出时间显著性;Zhu等人^[19]用三步搜索的块匹配方法来估计运动矢量,当存在相对运动时,采用归一化的运动矢量的幅值来形成时间显著图。块匹配的方法只能计算粗糙的运动矢量,搜索时容易达到局部最优。Zhou等人^[20]认为由于相机运动造成的背景平移在频域空间可以表示为相位差,利用傅里叶变换的性质和一些假设,可以根据当前帧重构出它的下一帧,而前景运动物体的运动不同于背景,时间显著图可以通过重构出的下一帧和真实的下一帧之间的差异计算出来;Bian等人^[21]通过傅里叶变换估计出两帧之间存在的全局运动,然后计算全局运动补偿后的两帧的亮度差,获得时间梯度信息,进而计算显著性。频率域检测方法的计算速度快,但对相机的运动方式有所限制,只适用于相机平移的情形,而且对噪声敏感。以上算法均有缺陷,帧差法只适用于相机固定和光照不变的条件,应用场景的局限性较大。采用光流法的模型大多只考虑运动矢量的幅值,而有时仅仅通过运动矢量的幅值很难区分前景物体和背景物体。更重要的是,这些基于像素域的模型的计算过程复杂、耗时严重,因此基于压缩域的视频显著性检测模型得到了更多的关注。

与像素域不同,基于压缩域模型无需对视频进行完全解码,而是从码流中提取DCT系数、运动矢量等相关数据。Fang等人^[9]综合考虑亮度、色度、纹理以及运动特征,于2014年提出了一种基于特征对比度的压缩域视频显著性检测模型。该模型在检测效率上相较于以往的像素域检测模型有大幅提升,引起了广泛的关注;但是在空域和时域显著图融合时,其对每一帧分配一个权重,未对帧内宏块加以区分,得到的最终显著图的精度较低。Khatoonabadi等人^[10]从全新的角度切入,认为视觉显著性取决于目标与周围的视觉刺激,用视频压缩后的比特数直接模拟这种刺激,在眼动点预测领域取得了较好的成果。

本文受文献^[10]的启发,提出了一种新的模拟视觉刺激的模型,简称为TS2CD。在视频压缩编码时,用分配给一个分割块的比特数来表示该块的压缩长度(Block Compression Length, BCL),用分配给一个运动矢量的比特数来表示该运

动矢量压缩长度(Motion Vector Compression Length, MVCL)。对于大的分割尺寸而言,运动矢量选择和分割类型只需要少量的比特,运动补偿残差在细节丰富区域中的能量将非常高。小尺寸分割的运动补偿残差能量低,需要较多比特来表征运动矢量和分割选择。由此可见,以上两个指标与视频运动信息有很强的相关性。然后,利用压缩域空时特征融合的方法将BCL和MVCL特征融合在一起得到最终的显著图。TS2CD模型示意图如图1所示。

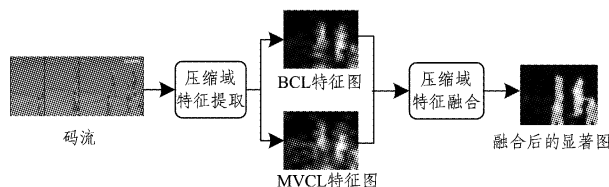


图1 TS2CD模型示意图

本文第3节将对BCL和MVCL特征进行详细介绍;第4节将详细说明压缩域空时融合算法;最后,在3个公开数据集上本文将本文所提的TS2CD与其他先进算法进行比较。

3 压缩域特征的提取和显著性计算

3.1 BCL和MVCL特征提取

近年来,视频压缩编码技术发展迅猛。最新的压缩编码标准HEVC已发布官方标准,其压缩效率较H.264有了明显提升,但是其由于高复杂度等限制还未得到广泛应用。H.264压缩标准从2003年正式发布到现在已经有十余年的历史,无论是在压缩效率还是压缩速度方面都很成熟。因此,目前网络视频和广播电视视频的主流压缩编码标准仍为H.264,而且本文中提取的特征可以扩展到包括HEVC的其他视频编码标准。鉴于实际应用,本文选择H.264标准。为了表述方便,本文使用像素块大小为 16×16 的宏块。

利用Ffmpeg进行视频解码,直接从码流中提取每一个宏块和运动矢量编码需要的比特数。该过程仅需要将码流分组的头部进行解码,其余的数据部分则无需解码。相较于传统的像素域算法,这种部分解码的压缩域特征提取极大地降低了计算复杂度,提升了检测效率。压缩域特征的计算公式如式(1)所示:

$$F_L = \frac{L - \min(L)}{\max(L) - \min(L)} \quad (1)$$

其中, $L \in \{BCL, MVCL\}$ 。对于每一帧视频,按式(1)将每一个宏块和运动矢量所需要的比特数 L 映射到 $[0, 1]$ 区间上, $F_L \in \{F_{BCL}, F_{MVCL}\}$, F_{BCL} 表示BCL特征值, F_{MVCL} 表示MVCL特征值, $F_L = 0$ 代表在所有宏块或运动矢量编码过程中分配了最少的比特数;与之相反, $F_L = 1$ 代表分配了最多的比特数。 $\max$ 和 $\min$ 函数表示帧内所有宏块中提取的最大比特数和最小比特数。

3.2 检测能力验证

为了验证本文提出的特征对显著区域的检验能力,测试了3个公开数据库(SegTrack^[22], NTT^[23]和 VideoSal^[24])的视频序列。利用人工标注的真值,比较显著区域和非显著区域的特征值,最后将特征值归一化到 $[0, 1]$ 区间。测试结果如图2所示,图2(a)为BCL特征采样散点图,图2(b)为BVCL特征采样散点图。图中每一个点代表视频序列的一帧,横坐

标 x 为非显著区域平均特征值,纵坐标 y 为显著区域平均特征值,计算方法如式(2)所示:

$$\begin{cases} x = \frac{1}{N} \sum_{i \in \text{bg}} F_L^i \\ y = \frac{1}{M} \sum_{i \in \text{fg}} F_L^i \end{cases} \quad (2)$$

其中, F_L^i 表示第 i 个宏块的特征值;bg(background)表示背景,即非显著区域;fg(foreground)表示前景,即显著区域; N 表示非显著区域宏块个数; M 表示显著区域宏块个数。

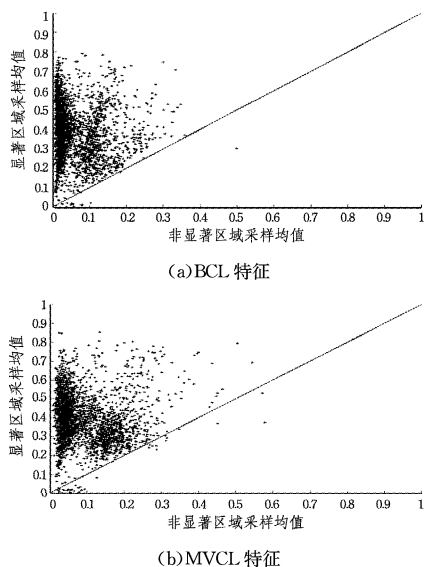


图2 显著-非显著区域特征采样散点图

图2表明显著区域的特征值明显高于非显著区域的特征值,即BCL和MVCL特征能够很好地反应视频的显著性。

为了更深入地研究BCL和MVCL特征的效果,分别对两种不同特征值在显著区域和非显著区域的概率分布加以统计。对于一个视频序列,利用已知的真值将每一个视频帧划分为显著区域和非显著区域,分别统计特征值在显著区域和非显著区域出现的概率。为了便于统计,将 $[0,1]$ 特征值取值范围平均划分成20个区间。图3中每一个点代表特征值在该区间出现的概率。最后,采用高斯分布来拟合特征概率分布。

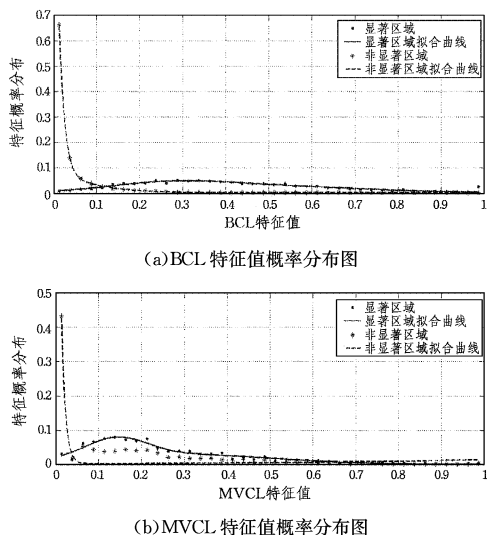


图3 特征值概率分布图

图3(a)和图3(b)表明,对于非显著区域,绝大部分特征值都非常小,几乎为零;而对于显著区域,特征值概率分布明显与非显著区域不同。

为了更好地证明BCL和MVCL特征的检测能力,基于t检验文献[25]的方法,利用显著区域采样和非显著区域采样进行双样本检测。t检验采用t分布理论来推导差异发生的概率,从而比较两个平均数的差异是否显著。其零假设是BCL和MVCL特征值在显著区域和非显著区域中没有差异,即显著区域采样的特征值和非显著区域采样的特征值具有相同的均值。本文的显著-非显著区域特征采样统计结果如表1所列。本文利用显著区域采样和非显著区域采样进行双样本检测,其零假设为两个采样总体的均值是相同的。

表1 显著-非显著区域特征采样统计结果

测试视频 序列	总帧数	百分比/% (BCL 特征)	P值 (BCL 特征)	百分比/% (MVCL 特征)	P值 (MVCL 特征)
DO01_013	73	100	10^{-22}	100	10^{-22}
DO01_014	85	100	10^{-26}	100	10^{-26}
DO01_030	85	97.65	10^{-23}	100	10^{-23}
DO01_055	47	100	10^{-15}	95.74	10^{-12}
DO02_001	63	100	10^{-19}	85.71	10^{-9}
M07_058	56	87.50	10^{-9}	83.93	10^{-7}
VWC102T	91	98.90	10^{-26}	98.90	10^{-26}
Cheetah	28	100	10^{-9}	100	10^{-9}
Girl	20	100	10^{-7}	100	10^{-7}
RC2	18	100	10^{-6}	100	10^{-6}
Ball2	73	100	10^{-22}	100	10^{-22}
Cross	1648	99.58	≈ 0	99.33	≈ 0
V13	320	100	10^{-97}	100	10^{-97}
Input2	106	100	10^{-32}	100	10^{-32}
Mc2	80	100	10^{-25}	98.75	10^{-23}

表1包含了测试视频所有帧中显著区域特征值高于非显著区域特征值的百分比以及P值。对于一般的假设检验,如果P值小于0.05,或者为了更加严谨,P值小于0.01,则可以拒绝零假设。对于所有测试视频,文中以更加严谨的显著性水平即P值小于 10^{-5} 拒绝零假设。可以得出结论:BCL和MVCL特征对显著性区域有很强的检测能力。因此直接利用特征值的大小模拟显著程度,分别将BCL和MVCL特征值按比例从 $[0,1]$ 区间映射到 $[0,255]$ 区间,从而得到BCL特征显著图和MVCL特征显著图,如图4所示。

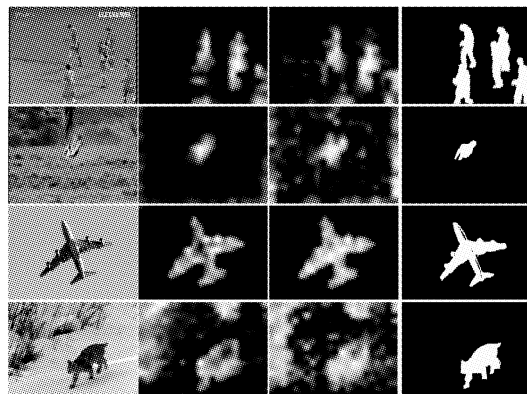


图4 BCL和MVCL特征图

4 压缩域空时特征融合算法

利用 BCL 特征和 MVCL 特征得到的显著图都可以较为直观地检测视频的显著性区域,然而不同的特征具有不同的优点,BCL 特征对空域显著性的检测效果更好,而 MVCL 特征对目标运动更加敏感,对时域显著性的检测效果更好,为了更好地利用不同特征的优点,借鉴空时显著图融合^[9]的方法,提出一种压缩域空时特征检测方法(TS2CD),将两种特征得到的显著图融合成最终的显著图。最终的显著图融合方法如式(3)所示。

$$S = \gamma_s S_s + \gamma_m S_m \quad (3)$$

其中, S 表示视频帧的最终显著图, S_s 为 BCL 特征得到的空域显著图, S_m 为 MVCL 特征得到的时域显著图, γ_s 和 γ_m 是每个部分的权重。权重参数由空域显著图和时域显著图的特征决定。一个相对准确的显著图的显著区域应该是小而紧凑的,而不应该分散在帧内各处。在时域显著图中显著区域如果是小而紧凑的,那么可以认为当前帧的运动对比度较强。在这种情况下,可以给 γ_m 赋予大的权重值,即与空域相比,时域显著图对最终显著图具有更大的贡献力。同样地,当空域显著图中的特征对比度较强时,应该给 γ_s 赋予大的权重值。

采用空间变化来衡量空域显著图和时域显著图的特征对比度。给定一个空域显著图或者时域显著图 S_k ($S_k \in \{S_s, S_m\}$),采用 Otsu 阈值^[26]的方法进行处理。空间变化计算公式如式(4)一式(6)所示:

$$v_k = \frac{\sum_{(i,j)} \sqrt{(i - S_{i,e})^2 + (j - S_{j,e})^2} \times S_k(i,j)}{\sum_{(i,j)} S_k(i,j)} \quad (4)$$

$$S_{i,e} = \frac{\sum_{(i,j)} i \times S_k(i,j)}{\sum_{(i,j)} S_k(i,j)} \quad (5)$$

$$S_{j,e} = \frac{\sum_{(i,j)} j \times S_k(i,j)}{\sum_{(i,j)} S_k(i,j)} \quad (6)$$

其中, v_k 表示显著图的空间变化。 $S_k(i,j)$ 为 S_k 中位置 (i,j) 的显著性值。 $S_{i,e}$ 和 $S_{j,e}$ 分别代表在水平和垂直方向上的显著性区域的空间期望值。空间显著图或时间显著图的空间变化可以通过式(4)一式(6)获得。显著图中,小而紧凑的显著区域的空间变化值比扩散的显著区域空间变化值小。因此,设置 $\gamma_k = 1/v_k$, $\gamma_k \in \{\gamma_s, \gamma_m\}$,使用 γ_k 来决定空域显著图或时域显著图在最终显著图所占的权重。最终将权重系数归一化后得到 $\gamma_s = v_m / (v_s + v_m)$, $\gamma_m = v_s / (v_s + v_m)$ 。

5 实验评价

在 3 个公开数据库 (SegTrack^[22], NTT^[23] 和 VideoSal^[24]) 上对本文所提出的视频显著性检测模型进行评价。其中,NTT 数据库的目标单一,背景相对较为简单,检测难度相对较小;SegTrack 数据库背景复杂,是显著性检测领域应用最为广泛的数据库;VideoSal 数据库综合了多目标、摄像机运动、目标形变、复杂背景等情况,是比较全面的测试数据库。

采用目前显著性检测最通用的评价指标(准确率-召回率曲线(P-R 曲线)和平均绝对误差(MAE)^[27])将所提 TS2CD 模型与最经典或者最先进的算法进行比较。其他不同方法的结果由作者提供或通过可运行代码得到。

5.1 参数和评价标准

P-R 曲线:通过设定阈值是对显著图进行二值化,且阈值范围为 0~255,可以对显著图进行二值化,并将二值化后的结果与真值做比较。通常情况下,良好的显著性检测结果不仅要求高准确率,也要求高召回率。

MAE:为了更加全面地评价不同模型的检测性能,同时采用 MAE 作为第二种评价标准。MAE 以像素点为单位计算显著图和真值之间的平均差异,并将其归一化到 [0,1] 区间内。

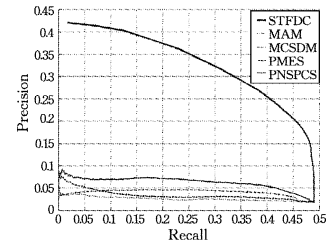
$$MAE = \frac{1}{W \times H} \sum_{x=1}^W \sum_{y=1}^H \|S(x,y) - G(x,y)\| \quad (7)$$

其中, W 和 H 分别表示图像的宽和高, $S(x,y)$ 和 $G(x,y)$ 分别表示 (x,y) 坐标位置上的显著值和真实值。该评价指标可以很好地反应显著图和真值之间的相似程度。

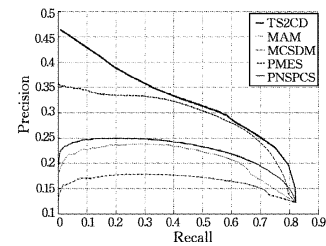
5.2 与其他典型算法的比较

为了证明本文所提的 TS2CD 模型的有效性,在 SegTrack,NTT 和 VideoSal 公开数据库上进行了测试,并与其他先进算法进行了比较。需要注意的是,基于像素域的视频显著性检测算法由于耗时过长,实时性较差,因此选取的对比算法都是近年来比较优秀的基于压缩域的视频显著性检测算法,并且可以获得其 Matlab 的代码。

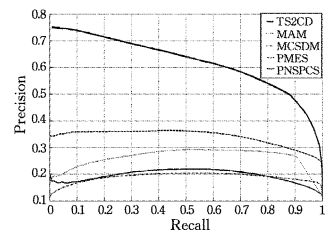
如图 5 所示,在 3 个公开数据库上,相较于其他算法,通过本文所提出的特征得到的显著图已经具有较高的准确率和召回率。利用本文融合算法融合后的最终显著图在所有数据库上都取得更高的准确率和召回率。不仅如此,表 2 列出的 MAE 值也表明了本文的显著性检测模型得到的最终显著图与真值更为相似。



(a) SegTrack 数据集



(b) NTT 数据集



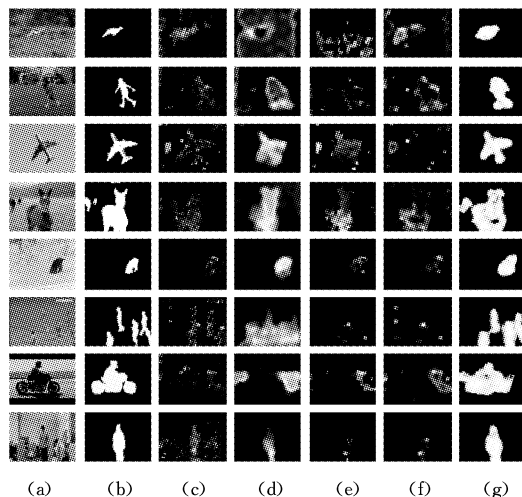
(c) VideoSal 数据集

图 5 不同算法在不同数据集上的 P-R 曲线

表 2 不同算法的 MAE 值

	TS2CD	PNSP ^[9]	PMES ^[28]	MAM ^[29]	MCSDM ^[30]
SegTrack	0.053	0.065	0.183	0.093	0.071
NTT	0.165	0.176	0.258	0.185	0.176
VideoSal	0.119	0.188	0.219	0.157	0.166

为了更直观地进行结果之间的比较,图 6 给出了不同模型检测得到的显著图。从图 6 中不难发现,本文所提模型的检测结果优于其他先进模型。



注:(a)为视频帧原图;(b)为真值;(c)~(f)分别是 PNSP, PMES, MAM, MCSDM 检测得到的显著图;(g)为本文 TS2CD 模型得到的显著图

图 6 视频显著性检测结果示意图

值得一提的是,本文所提的 TS2CD 不仅在检测效果上优于其他模型,而且耗时较少,基本可以满足实时性要求。用一台配置为 2.53GHz i3 处理器、内存为 4GB 的 PC 测试不同算法所需时间,如表 3 所列。

表 3 每帧视频平均处理时间

算法	PNSP	PMES	MAM	MCSDM	TS2CD
时间/ms	890	575	779	15	22

本文所提的基于压缩域的视频显著性检测模型的耗时较少,可以达到 40 帧/秒左右的处理速度,基本上满足了大多数计算机视觉应用的要求。个别算法例如 MCSDM 的处理速度与本文算法相近,但是其检测精度太低,基本上只能完成对显著区域边缘的检测。本文所提模型在检测精度和检测效率上都优于现有的算法。

结束语 本文提出了一种新颖的基于压缩域的视频显著性检测模型。该模型同时考虑了视频编码长度和运动矢量编码长度,从根本上模拟在观察视频时显著区域对人眼的刺激。构建了两个可以用来衡量显著性的特征,并且证明了其检测的能力;不仅如此,在融合理论的框架下,采用模拟空域显著图融合的方法,将两个特征得到的显著图很好地结合起来,并最终融合成具有更高准确率和召回率的显著图。而且本文所提模型的检测速度较快,每秒大约可以处理 40 帧。在 3 个公开数据库上的实验表明,相较于其他现有的模型,本文所提出的模型能够获得更好的结果。但是值得注意的是,基于压缩域的视频显著性检测算法起步较晚,目前仍处于探索阶段,算法相对较少,仍有很多可以改进之处。

参考文献

- [1] LEE E, GU E, PARK K. Effective small target enhancement and detection in infrared images using saliency map and image intensity[J]. Optical Review, 2015, 22(4): 659-668.
- [2] QIN Y, LH H C, XU Y Q, et al. Saliency detection via cellular automata[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 110-119.
- [3] XU M, ZHANG H L. Saliency detection with color contrast based on boundary information and neighbors[J]. Visual Computer, 2015, 31(3): 355-364.
- [4] MA Y F, HUA X S, LU L, et al. A generic framework of user attention model and its application in video summarization[J]. IEEE Transactions on Multimedia, 2005, 7(5): 907-919.
- [5] GHOSH J. Discovering important people and objects for egocentric video summarization[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Washington, 2012: 1346-1353.
- [6] WANG W G, SHEN J B, PORIKLI F. Saliency-aware geodesic video object segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 3395-3420.
- [7] CHANG J, WEI D L, III J. A video representation using temporal superpixels[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013: 2051-2058.
- [8] WANG Y L, ZHANG Q, LI B X. Efficient unsupervised abnormal crowd activity detection based on a spatiotemporal saliency detector[C]//Proceedings of the IEEE Computer Society Winter Conference on Applications of Computer Vision. 2016: 1-9.
- [9] FANG Y M, LIN W S, CHEN Z Z, et al. A video saliency detection model in compressed domain[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2014, 24(1): 27-38.
- [10] KHATOONABADI S H, VASCONCELOS N, BAJIC I V, et al. How many bits does it take for a stimulus to be salient? [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015: 5501-5510.
- [11] KHATOONABADI S H, BAJIC I V, SHAN Y F. Compressed-domain correlates of human fixations in dynamic scenes[J]. Multimedia Tools & Applications, 2015, 74(22): 10057-10075.
- [12] TREISMAN A M, GELADE G. A feature-integration theory of attention[J]. Cognitive Psychology, 1980, 12(1): 97-136.
- [13] PALENICHKA R, LAKHSSASSI A, PALENICHKA M. Visual attention-guided approach to monitoring of medication dispensing using multi-location feature saliency patterns[C]//Proceedings of the IEEE International Conference on Computer Vision Workshop. 2015: 461-468.
- [14] NINASSI A, MEUR O L, CALLET P L, et al. Task impact on the visual attention in subjective image quality assessment[C]//Proceedings of the European Signal Processing Conference. 2015: 1-5.
- [15] ITTI L, KOCH C, NIEBUR E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pat-

- tern Analysis & Machine Intelligence, 1998, 20(11): 1254-1259.
- [16] ESCALERA S, RADEVA P, PUJOL O. Complex salient regions for computer vision problems[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2007; 1-8.
- [17] KIM W, JUNG C, KIM C. Spatiotemporal saliency detection and its applications in static and dynamic scenes[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2011, 21(4): 446-456.
- [18] MEUR O L, THOREAU D, CALLET P L, et al. A spatio-temporal model of the selective human visual attention[C]// Proceedings of the IEEE International Conference on Image Processing. 2005; 1188-1191.
- [19] ZHU Y P, JACOBSON N, PAN H, et al. Motion-decision based spatiotemporal saliency for video sequences[C]// Proceedings of the International Conference on Acoustics, Speech & Signal Processing. 2011; 1333-1336.
- [20] ZHOU B L, HOU X D, ZHANG L Q. A phase discrepancy analysis of object motion[C]// Proceedings of the Asian Conference on Computer Vision. 2010; 225-238.
- [21] BIAN P, ZHANG L M. Biological plausibility of spectral domain approach for spatiotemporal visual saliency[C]// Proceedings of the International Conference on Neuro-Information Processing, Auckland. 2008; 251-258.
- [22] LI F X, KIM T, HUMAYUN A, et al. Video segmentation by tracking many figure-ground segments[C]// Proceedings of the IEEE International Conference on Computer Vision. 2013; 2192-2199.
- [23] AKAMINE K, FUKUCHI K, KIMURA A, et al. Fully automatic extraction of salient objects from videos in near real time[J]. Computer Journal, 2010, 55(1): 3-14.
- [24] HUANG C R, CHANG Y J, YANG Z X, et al. Video saliency map detection by dominant camera motion removal[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2014, 24(8): 1336-1349.
- [25] KREYSZIG E. Introductory mathematical statistics principles and methods[J]. Technometrics, 1971, 13(4): 922-925.
- [26] OHTSU N. A threshold selection method from gray-level histograms[J]. IEEE Transactions on Systems Man & Cybernetics, 1979, 9(1): 62-66.
- [27] BORJI A, CHENG M M, JIANG H, et al. Salient Object Detection: A Survey[J]. Eprint Arxiv, 2014, 16(7): 3118.
- [28] MA Y F, ZHANG H J. A new perceived motion based shot content representation[C]// Proceedings of the IEEE International Conference on Image Processing. 2001; 426-429.
- [29] MA Y F, ZHANG H J. A model of motion attention for video skimming [C]// Proceedings of the IEEE International Conference on Image Processing. 2002; 129-132.
- [30] LIU Z, YAN H B, SHEN L Q, et al. A motion attention model based rate control algorithm for H. 264/AVC[C]// Proceedings of the IEEE/ACIS International Conference on Computer and Information Science. 2009; 568-573.
-
- (上接第 295 页)
- [13] ZHANG L, HU Q, DUAN J, et al. Multi-label Feature Selection with Fuzzy Rough Sets [M]// Rough Sets and Knowledge Technology. Springer International Publishing, 2014; 121-128.
- [14] DUAN J, HU Q H, ZHANG L J, et al. Feature Selection for Multi-Label Classification Based on Neighborhood Rough Set [J]. Journal of Computer Research and Development, 2015, 52(1): 56-65. (in Chinese)
段洁, 胡清华, 张灵均, 等. 基于邻域粗糙集的多标记分类特征选择算法[J]. 计算机研究与发展, 2015, 52(1): 56-65.
- [15] SPOLAOR N, CHERMAN E, MONARD M. Using ReliefF for multi-label feature selection[C]// Conferencia Latinoamericana de Informática. 2011; 960-975.
- [16] SPOLAOR N, CHERMAN E, MONARD M, et al. A comparison of multi-label feature selection methods using the problem transformation approach [J]. Electronic Notes in Theoretical Computer Science, 2013, 292: 135-151.
- [17] SPOLAOR N, CHERMAN E, MONARD M, et al. ReliefF for multi-label feature selection[C]// 2013 Brazilian Conference on Intelligent Systems (BRACIS). IEEE, 2013; 6-11.
- [18] REYES O, MORELL C, VENTURA S. Scalable extensions of the ReliefF algorithm for weighting and selecting features on the multi-label learning context [J]. Neurocomputing, 2015, 161: 168-182.
- [19] LI J H, FU J F, JIANG W J, et al. Feature Selection Method Based on MRMR for Text Classification[J]. Computer Science, 2016, 43(10): 225-228. (in Chinese)
李军怀, 付静飞, 蒋文杰, 等. 基于 MRMR 的文本分类特征选择方法[J]. 计算机科学, 2016, 43(10): 225-228.
- [20] SUN Y. Iterative RELIEF for feature weighting: algorithms, theories, and applications [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(6): 1035-1051.
- [21] GILAD-BACHRACH R, NAVOT A, TISHBY N. Margin based feature selection-theory and algorithms [C]// Proceedings of the Twenty-first International Conference on Machine Learning. ACM, 2004; 43.
- [22] TSOU MAKAS G, VLAHAVAS I. Random k-label sets: An ensemble method for multi-label classification [C]// European Conference on Machine Learning. 2007; 406-417.
- [23] ZHANG M, PEÑA J, ROBLES V. Feature selection for multi-label naive Bayes classification [J]. Information Sciences, 2009, 179(19): 3218-3229.
- [24] ZHANG M, ZHOU Z. ML-KNN: A lazy learning approach to multi-label learning [J]. Pattern Recognition, 2007, 40(7): 2038-2048.
- [25] FRIEDMAN M. A comparison of alternative tests of significance for the problem of m rankings [J]. The Annals of Mathematical Statistics, 1940, 11(1): 86-92.
- [26] DUNN O. Multiple comparisons among means [J]. Journal of the American Statistical Association, 1961, 56(293): 52-64.