

基于视频内容广告植入图片的匹配研究^①

刘 海, 曾东海

(广东科学技术职业学院 软件学院, 珠海 519090)

摘 要: 基于视频内容植入广告系统目的是在用户观看视频时, 能实时获取与视频内容相关商品的广告咨询. 给出了系统解决方案, 并分析了问题难点, 提出了解决视频广告匹配问题难点的处理算法: 先简单分割, 然后通过直方图百分率方式二次匹配过滤后, 得到分割后的图像, 再利用改进的 SIFT 算法进行精确匹配, 从而确定当前视频图片中是否存在广告图片相似度高的商品信息. 实验表明, 基于视频内容的广告植入系统中图片匹配算法在常用视频中, 算法实时性能好, 匹配率高.

关键词: 视频; 广告植入; SIFT; 颜色直方图

Image Matching with Advertisement of Content-Based Video

LIU Hai, ZENG Dong-Hai

(Department of Computer Engineering and Technology, Guangdong Institute of Science and Technology, Zhuhai 519090, China)

Abstract: The purpose of video advertisement system is real-time obtainment of advertisement consultation of relative goods showed in video when users watching video. The contents of this research are the solution of system, analysis of difficult points, and solution of the difficult process on image matching with video ads. Firstly, it segments image simply. Then, it filters unmatched image twice by percentage of histogram method. At the last, it matches the image after segmenting precisely by simplified SIFT. We can receive the information whether the goods is the most similar to the image from video ads. Experiments show that this video advertisement system is real-time and possessing high matching rate on certain condition.

Key words: video; advertising implantation; SIFT; color histogram

1 背景

目前网络视频广告一般有两种方式: 第一种是视频开播前强制播放广告, 或者视频暂停时强制播放广告, 这种广告的广告内容与视频内容无关联性; 第二种是根据视频介绍, 视频标题等关键字投放关联广告, 这种广告咨询和视频内容有一定的关联性, 但是关联性不是太大. 美国研究机构 Forrester Research 的研究表明, 因投放的广告和视频内容缺乏关联性, 导致 80% 的网络视频观众表示对放置在视频之前或之后的贴片广告很讨厌, 从实际投放效果来看, 这种单项、粗放式、侵扰性的广告形式, 与当今开放的、民主的、以用户为中心的、注重用户体验的网络环境不协调,

无法得到用户的认可而使得传播效果不佳, 也将挫伤广告客户的投放积极性. 为此, 广告与视频内容的紧密结合将是今后网络视频广告发展的方向之一.

基于视频内容植入广告并友好展示给用户是一种新广告的创新方式(比如, 某视频中一位明星穿一件时装, 如果能够对该明星所穿的时装做广告就很有意义), 目前国内外文献未发现与该创新方式的相关研究, 本文就基于视频内容植入广告技术做初步研究, 以期到达抛砖引玉的目的. 文章后续内容如下组织: 第 2 节概述系统的解决方案; 第 3 节讲述广告图像与视频源画面匹配分割以及匹配的处理算法第 4 节简述实验环境和实验效果.

^① 基金项目: 广东自然科学基金(S2013010012920)

收稿时间: 2014-01-02; 收到修改稿时间: 2014-03-03

2 解决方案

基于视频内容植入广告系统分为两部分: 后台和前台. 后台是用于广告制作商或者视频制作方在视频中植入广告图片, 这些广告图片来自视频本身画面中的某些商品; 前台是用户观看视频时, 当用户点击感

兴趣的商品或者用户暂停视频时, 系统比较当前视频画面中的商品和后台植入的各张广告图片的相似程度, 当相似程度达到一定阈值, 则展示最相似的广告图片的相应资讯信息, 若无相似广告图片则展示默认广告图片. 系统解决方案的流程如图 1 所示:

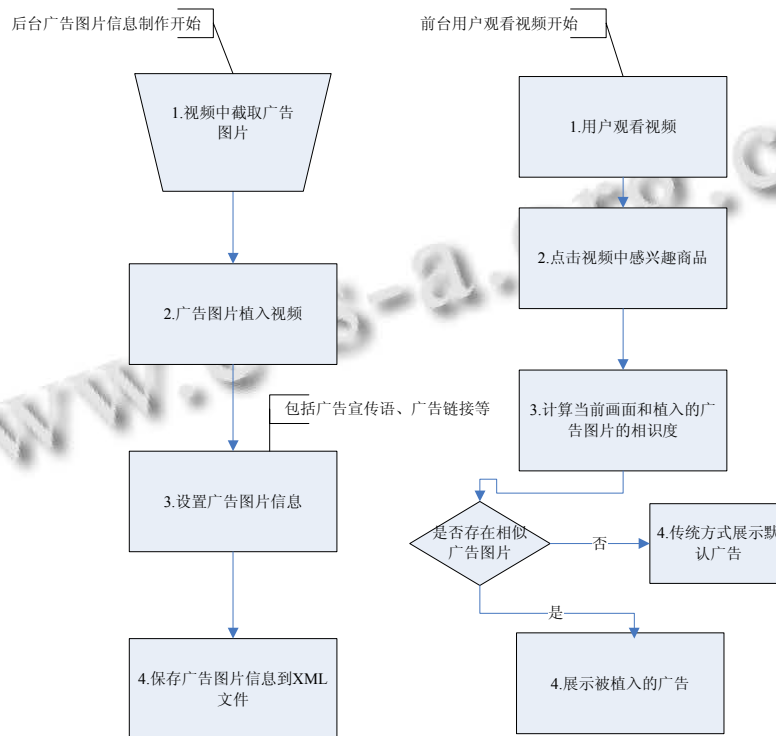


图 1 系统流程图

系统处理的关键难点在于: 前台用户观看视频当点击感兴趣画面时, 进程 3 中如何“计算当前画面和植入的广告图片的相似度”。

3 广告图片与视频源画面图像匹配算法

目前在图像匹配方面有一些比较好的算法, 比如 SIFT 算法^[7]、SURF 算法^[2]等, SIFT 算法的优点匹配率高, 具有平移, 旋转, 缩放不变性等良好特点. 但是 SIFT 算法由于特征算子维数大, 算法时间效率不高, 难以符合视频播放中实时弹出相应的广告咨询的应用. 如果直接使用 2004 LOWE 所述 SIFT^[7]算法匹配广告图片 (大概视频画面的 1/8 大小) 和视频画面 (720*480px), 响应时间在 20 s 以上, 且实验表明, 当两张需匹配的图片大小比例差异逐步增大时, SIFT 算法的匹配准确率逐步下降. 文献^[10]使用 Gabor 滤波器对 SIFT 描述符进行改进, 使其更适用于目标识别

等精确度要求较高的场景中, 但算法的计算复杂度增加. 文献^[11]基于 SIFT 局部特征的贝叶斯方法匹配目标物体, 在尺度变换、旋转等情况下最低识别正确率达 92%以上, 但是在本文图片规格下, 难以实现实时 (8s, “8s”定律指出, 用户对页面最大的忍耐限度是 8 秒钟) 应用.

约定后台截取广告图片为图 A (图 3 所示), 图 A 为 $W_A * L_A$ 像素, 用户观看当前视频图片为图 B (图 2 所示), 图 B 为 $W_B * L_B$ 像素, 设定图 A 和图 B 的特点为:

- 1) 广告图片即图 A, 比较完整的包含某商品.
- 2) 视频图片图 B 的大小是广告图片图 A 大小的 4 倍以上.

由于目前高清和标清视频格式每一帧图像的像素值一般在 720*480 以上, 直接采用 SIFT 效果较好的图像匹配算法处理时, 在实时处理速度上慢 (实验效果用户点击视频, 广告弹出时间介于 20s-30s 之间), 用户

体验不好. 为此处理思路是, 先把图 B 分割成和图 A 大小一致的图片, 然后通过颜色直方图百分率多次匹配, 粗略过滤掉相似度较小的图片, 留下相识度最大的图片, 最后通过简化 SIFT 算法精确匹配, 成功后展示匹配相识度最大的广告图片, 否则展示默认广告图片(传统广告方式).



图 2 视频源画面 图 B(841*906 pix)



图 3 广告图片 图 A(290*148pix)

详细算法处理步骤如下:

1) 图 B(图 2 所示)简单分割成图 A(图 3 所示)的 1/4 大小:

把图 B(图 2 所示)分割成 $NW_C = 2W_B / W_A$, $NL_C = 2L_B / L_A$ 即 $NW_C * NL_C$ 整数份, 图 B 多余的边缘去除(图像边缘具有商品信息的概率很小), 设定分隔后的图像为记为图 $C_{ij}(0 \leq i \leq NL_C, 0 \leq j \leq NW_C)$, 则容易计算出分隔后的图 C_{ij} 是图 A 的 1/4 大小, 图 B(图 2)分割 C_{ij} 后效果图 4 所示:

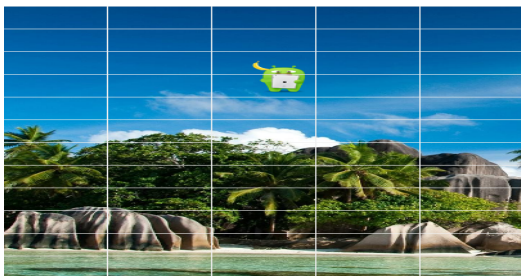


图 4 简单分割后(5*12 子图)

2) 颜色直方图百分率粗略匹配筛选:

对于给定的图像 $I[x,y]$, 它的直方图^[3] $h_c[m]$ 可以由公式(1)给出:

$$h_c[m] = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \begin{cases} 1, & Q_c(T_c I(x,y))=m \\ 0, & \text{其他} \end{cases} \quad (1)$$

其中, X,Y 分别为图像的宽度和高度, m 为图像像素点的某一灰度值($0 \leq m \leq 255$), $Q_c(T_c I(x,y))$ 点(x,y)为灰度化后的颜色值.

颜色直方图定义了图像或图像中区域的颜色分布, 颜色直方图保证了图像的平移、旋转不变性.

定义. 颜色直方图百分率 $h_c[m]\%$:

$$h_c[m]\% = h_c[m] / \sum_{x=0}^{255} h_c[x] \quad (2)$$

其中 $h_c[m]$ 是图像的颜色直方图.

颜色直方图百分率保证了图像的缩放、平移、旋转不变性.

灰度化后计算图 A 的颜色直方图百分率, 计算图 C_{ij} 的颜色直方图百分率.

3) 图像合并成合图 A 大小一致:

根据绝对值距离, 图 A 和所有图 C_{ij} 根据颜色直方图百分率逐一匹配, 匹配值 k 属于[0,2]区间, 当 $k < 1$ 时留取图 C_{ij} . 如果 $C_{ij}, C_{i+1j}, C_{ij+1}, C_{i+1j+1}$ 匹配 k 值有 2 个小于 1 时, 则 4 图组合成新图 D_{ij} , 效果如图 5 所示.



图 5 粗略匹配合并图 C_{ij} 后效果

4) 再次颜色直方图百分率匹配筛选: 把图 C_{ij} 与图 A 再次使用颜色直方图百分率进行匹配, 进一步过滤相似度低的图片因为每次最多弹出一幅广告图片, 为此这时只需留下相似度最高的图, 记为图 D, 效果

如图 6 所示。



图 6 颜色直方图二次匹配后图 D

5) 然后通过简化的 SIFT 算法过程精确比对, 进一步确定图 D 和图 A 是否相似。

原 SIFT 算法主要包括以下 4 步骤^[8]。

- i. 尺度空间极值检测。
- ii. 利用关键点邻域像素的梯度方向分布特性, 为每个关键点指定方向参数, 使算子具备旋转不变性
- iii. 生成 SIFT 特征向量, 将坐标轴旋转到特征点方向, 以保证旋转不变性图
- iv. 特征匹配

由于原第 3 步骤中耗时几乎占用了整个 SIFT 算法的 80%, 而且第 2 步骤和第 3 步骤做了类似直方图的统计工作, 所以我们可以考虑合并第 2 步骤和第 3 步骤并以适当降低匹配正确率为代价简化算法, 提高时间效率。

简化 sift 算法合并原 2 和 3 步骤更改为:

i. 在原来第一步获得特征关键点之后, 直接以特征点为中心采用圆形窗体确定需要统计的邻域范围, 窗口尺寸采用 Lowe 推荐的 $9\sigma \times 9\sigma$ ^[8], 这样圆形窗口半径取 4.5σ , 在该圆形窗体内统计 12 个梯度方向。

ii. 将 12 个梯度方向规一化, 从而保证光照不变性。假设 M 是特征点的特征向量, 即 $M = (m_1, m_2, \dots, m_{12})$, 归一化后得到特征向量 N:

$$N = M / \sqrt{\sum_{i=1}^{12} m_i^2} = (n_1, n_2, n_3, \dots, n_{12}) \quad (3)$$

iii. 向左循环移动整个向量序列, 直到最大的梯度方向统计量移动到向量的第一个元素, 即若 $n_1 = m_a \times \{n_i, n_i \in N\}$, 最终特征向量形成 $N = (n_1, n_2, n_3, \dots, n_{12})$, 保证了旋转不变性。

4 实验效果

该算法过程由于把视频图像分割成更小图像进行处理,

并简化了 SIFT 算法, 从而使得 SIFT 特征描述子数据大量减少, 匹配时间急剧递减, 实时性能好。实验采用普通 PC 机, 内存 2G, 处理器 2.20GHz。视频图片帧大小 $720 \times 480px$, 后台截取并植入 2-3 幅广告图片(广告图片数量影响响应时间), 广告图片大小不一, 大概在 $240 \times 160px$ 左右。实验视频 5 个, 当用户点击暂停视频后, 一幅广告视频弹出的响应时间在 $2s - 4s$ 之间, 多幅时间相应增加(表 1 中平均响应时间是多幅广告图像的响应时间)。广告弹出准确率平均在 80% 左右。图 7 是实验软件运行效果之一——视频《AA 制生活》广告弹出效果。实验视频弹出广告效果本文算法对比 SIFT 算法具体统计数据见表 1。



图 7 《AA 制生活》广告弹出效果

表 1 视频广告匹配实验统计信息表

视频名称	广告商品 1	广告商品 2	广告商品 3	实验视频画面次数	平均响应时间 (ms)		匹配准确率	
					本文算法	SIFT	本文算法	SIFT
AA 制生活	白色衣服	红色衣服		6	6512	26213	0.833	0.833
浪漫满屋 1	花色衣服	黑色衣服	粉色衣服	8	9576	33234	0.75	0.75

浪漫 满屋 10	黑 色 衣 服	灰 色 衣 服		6	5548	25215	0.667	0.833
野 生 动 物	白 马	鸟		6	5049	25200	0.833	0.833
时 装 秀	绿 色 衣 服	蓝 色 外 套		10	6346	27544	0.9	0.9

从表中可以看出,本文算法在实时性方面远远优于 SIFT 算法,在匹配准确率方面同 SIFT 算法相当。

参考文献

- 于丽莉,戴青.一种改进的 SIFT 特征匹配算法.计算机工程 2011,(1):210-212.
- 毛运柳,黄东军.基于 SURF 视频分割的视频水印算法.计算机工程,2010,10:241-243.
- 何金国.数字图像处理实训教程.北京:清华大学出版社,2008.
- 章毓晋.基于内容的视觉信息检索.北京:科学出版社,2003.
- 王开.Visual C++数字图象获取处理及实践应用.北京:北京人民邮电出版社,2003.
- 张建东,苏鸿根.基于内容的图像检索关键技术研究.计算机工程,2004,30(14):119-121.
- Lowe DG. Distinctive image features from scale-invariant key-points. International Journal of Computer Vision, 2004, 60(2): 91-110.
- Lowe DG. Object recognition from local scale-invariant features. International Conference on Computer Vision. 1999. 1150-1157
- Moreno P, Bernardino A, Victor SJ. Improving the SIFT descriptor with smooth derivative filters. Pattern Recognition Letters, 2009, 30(1): 18-26.
- Chang L, Duarte MM, Sucar LE, et al. A Bayesian approach for object classification based on clusters of SIFT local features. Expert Systems with Applications, 2012, 39: 1679-1686.
- Gurjal P, Kunnur K. Real Time hand gesture recognition using SIFT. International Journal of Electronics and Electrical Engineering, 2012, 2: 19-33.