

基于视觉显著模型的音视频交叉水印算法

汪新园 潘洋绪

(河海大学计算机与信息学院 南京 211100)

摘要: 提出一种基于 Itti 视觉显著模型的音视频交叉水印算法。首先利用音频码流的低通幅值特性生成水印,并嵌入到视频码流中;其次解码视频码流的关键帧 I 帧,获取 YUV 序列上的亮度(Y)分量,并对其进行显著性分析;最后分块对非显著区域进行 DCT 变换,并通过量化索引调制嵌入水印。实验结果表明,该算法将嵌入水印后的关键帧平均 PSNR 值提高了 1.678 5 dB,缓解了水印不可见性和鲁棒性的矛盾,实现了音视频码流之间的交叉保护。

关键词: 交叉水印;视觉显著模型;离散余弦变换;量化索引调制

中图分类号: TP391TN911.72 **文献标识码:** A **国家标准学科分类代码:** 510.40

Audio and video cross watermarking algorithm based on visual saliency model

Wang Xinyuan Pan Yangxu

(Department of Computer and Information, Hohai University, Nanjing 211100, China)

Abstract: This paper proposed a simple audio and video cross watermarking algorithm based on Itti's visual saliency model. First of all, it utilized the low-pass amplitude characteristic of audio streaming to generate watermarking, embedded into the video streaming; Secondly, decoding the keyframe, I frame, of video streaming to obtain the Y component of YUV sequence, remarkably analyzing it; Finally, non-remarkable area is blocked into DCT, embedded into watermarking by quantization index modulation. The experimental results show that the average PSNR of key frame, in this algorithm, is lifted by 1.678 5 dB, which relieved the contradiction of watermark invisibility and robustness and achieved the cross-protection between the audio and video streams.

Keywords: joint-watermarking; visual saliency model; DCT; quantization index modulation(OIM)

0 引言

近年来,随着多媒体技术和计算机网络的不断发展,人们正从不同方面,不同层次深入研究数字水印,水印载体正从图像向音频、视频甚至多维发展,音视频交叉水印随之产生,可以解决某一媒体数据水印算法不能保证其包含的另一种媒体数据安全问题。作为一种鲁棒性和不可见性兼备的信息隐藏技术,音视频交叉水印可以很好的保护各种多媒体数字资源的版权^[1]。

相较于单一媒体数据水印算法,音视频交叉算法的研究目前还处于初始阶段,有效成熟的算法较少。Dittmann^[2]提出一种相对简单的交叉水印算法,首先生成两个序列,一个以 1 递增,一个以 1 递减,然后将递增的序列嵌入到视频信号中,递减的序列嵌入到音频信号中,最后通过对所有水印数据相加,看其和是否为 0,来验证音视频之间的同步性。林晓丹^[3]为了提高算法的抵抗能力,提出

一种半脆弱水印算法,提取音频数据的梅尔频率倒谱系数作为音频特征信息并将其嵌入到视频码流中,完成对音频内容篡改的检测^[4]。

由此可见,现有的音视频交叉水印算法主要利用音频(视频)码流的特征信息生成水印文件,嵌入到视频(音频)码流中,实现音视频码流之间的交叉保护。

1 基于视觉显著模型的音视频交叉水印算法

视觉显著模型是一种用计算机来模拟人类视觉注意力系统模型,在一幅图像中提取人眼注意力感兴趣的区域,对计算机而言就是该图像的显著区域^[5]。检测人们感兴趣区域就是利用视觉显著模型得到图像中显著度较高的区域。

Itti 视觉显著模型属于自底向上控制策略的预注意机制,是由 Itti 等人提出的比较经典的视觉显著模型之一。

Itti 视觉显著模型主要分为特征图提取和显著图生成

两部分,主要框架如图1所示。

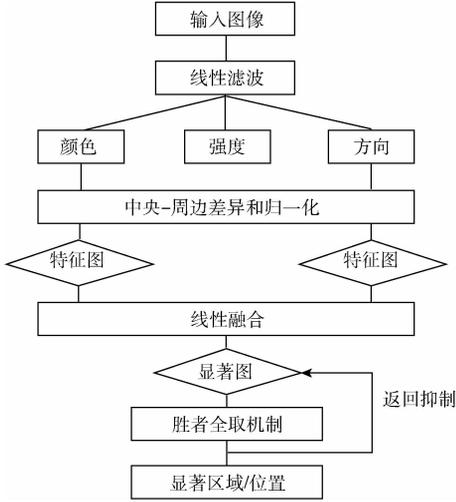


图1 Itti视觉显著模型框架

该模型基本思想如下:

- 1) 通过线性滤波对图像提取颜色特征、强度特征和方向特征;
- 2) 利用高斯金字塔模型对这3个特征分量进行多尺度采样;
- 3) 对采样得到的多尺度的特征数据进行中央-周边差异计算和归一化处理,得到12张颜色特征图、6张亮度特征图、24张方向特征图;
- 4) 将这些特征图结合归一化处理,分别形成颜色、亮度、方向关注图,3个特征关注图线性融合得到图像综合的视觉显著区域信息^[6]。
- 5) 在已有显著图中选取显著值最高的点作为视觉第一焦点,并以此焦点为圆点,固定长度为半径的圆形区域划分为第一视觉显著区域;
- 6) 采用返回抑制机制,抑制当前显著区域,在剩余显著区域中重复上诉方法寻找第二视觉显著区域;
- 7) 以此类推,实现视觉显著区域的划分选择工作。

2 水印嵌入与提取

2.1 水印嵌入

水印嵌入具体步骤如下:

步骤1:输入原始视频,将其分解为音频码流和视频码流,获取数据,解码YUV序列,获取视频关键帧I帧以及同一时刻解码出的音频帧。

步骤2:利用音频帧低通幅值特性生成待嵌入水印文件,水印嵌入方法如式(1)所示。

设原始载体数据为 $S = \{S(1), S(2), \dots, S(n)\}$, 待嵌入水印为 $W = \{W(1), W(2), \dots, W(n)\}$, 量化步长为 λ , 嵌入水印后的数据为 $S'(i)$ 。

$$S'(i) = \begin{cases} \left\lfloor \frac{S(i) + 0.5\lambda}{2\lambda} \right\rfloor \times 2\lambda + \frac{\lambda}{2}, & W(i) = 0 \\ \left\lfloor \frac{S(i) - 0.5\lambda}{2\lambda} \right\rfloor \times 2\lambda + \frac{3\lambda}{2}, & W(i) = 1 \end{cases}$$

$$S'(i) = S'(i) + 2\lambda, S'(i) < 0 \quad (1)$$

式中 $\lfloor \cdot \rfloor$ 表示向下取整。

步骤3:采用Itti视觉显著模型对视频关键帧进行显著性分析,通过注意焦点转移策略选取前3个视觉显著区域。

步骤4:将 $M \times N$ 大小的I帧进行 8×8 分块,对关键帧I帧亮度(Y)分量的DCT域嵌入水印,得到可嵌入水印的分块,对非显著区域分块进行DCT变换(离散余弦变换),求得DCT直流系数 $DCT_0(m, n)$ 。

步骤5:采用二进制奇偶量化调制算法^[7]嵌入水印,平衡鲁棒性和不可见性,选取量化步长 $\lambda(40)$ 均匀量化亮度分量DCT直流系数,使其满足:

$$W(i) = \text{floor}\left(\frac{DCT_0(m, n)}{\Delta}\right) \bmod 2 \quad (2)$$

若满足式(2),则调整后的DCT系数为:

$$DCT_w(m, n) = \text{floor}\left(\frac{DCT_0(m, n)}{\lambda}\right) \times \lambda + \frac{\lambda}{2} \quad (3)$$

否则根据系数在量化区间的位置,将系数调为:

$$DCT_w(m, n) =$$

$$\begin{cases} \text{floor}\left(\frac{DCT_0(m, n)}{\lambda}\right) \times \lambda - \frac{\lambda}{2} & \text{if } DCT_0(m, n) - \text{floor}\left(\frac{DCT_0(m, n)}{\lambda}\right) \times \lambda < \frac{\lambda}{2} \\ \text{floor}\left(\frac{DCT_0(m, n)}{\lambda}\right) \times \lambda + \frac{3\lambda}{2} & \text{if } DCT_0(m, n) - \text{floor}\left(\frac{DCT_0(m, n)}{\lambda}\right) \times \lambda \geq \frac{\lambda}{2} \end{cases} \quad (4)$$

步骤6:当该帧所有分块嵌入后进行逆DCT变换,得到嵌入水印的视频关键帧。

步骤7:重复步骤2~6,直至将一个完整水印嵌在关键帧中,完成整个视频解码。

步骤8:将音频码流和包含水印的视频码流进行重新编码输出视频。

流程图如图2所示。

2.2 水印提取

水印提取过程与嵌入过程类似。首先输入包含水印的视频载体,解码音视频码流数据,锁定视频关键帧和相应时间的音频帧;其次对视频关键帧的亮度分量采用Itti视觉显著模型进行显著性分析,取前3个视觉显著区域;接着对非显著区域进行 8×8 分块DCT变换,获取DCT直流系数 $DCT_0'(m, n)$ 。根据式(5)提取当前帧嵌入的水印:

$$W(i) = \text{floor}\left(\frac{DCT_0'(m, n)}{\lambda}\right) \bmod 2 \quad (5)$$

最后为了验证水印提取的完整性,对比嵌入水印数据,选择误差率最小的部分作为提取的水印。

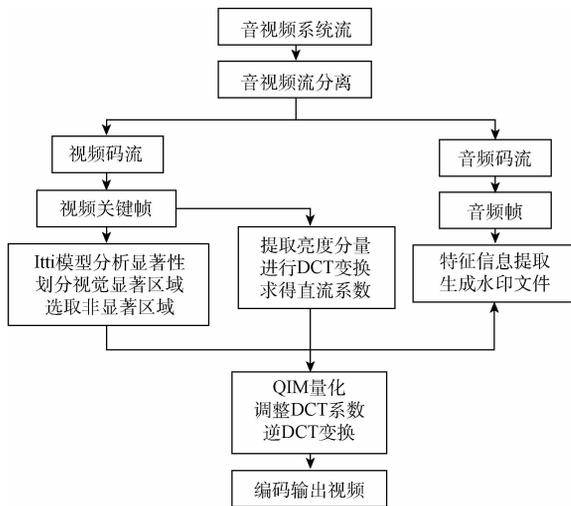


图 2 水印嵌入流程

3 实验结果分析

3.1 不可见性分析

在音视频交叉水印算法中采用视觉显著模型是为了提高水印的不可见性^[8], PSNR(峰值信噪比)是评价水印算法不可见性的重要标准,如式(6)所示。

$$PSNR = 10 \times \lg \frac{255^2}{MSE} \quad (6)$$

式中:图像像素范围为 0~255, MSE 是原图像与处理图像之间的均方误差。

PSNR 的单位是 dB,值越大代表失真越少,不可见性越高。当 PSNR>30 dB 时,人眼就很难发现视频质量的下降。如图 3 所示,其 PSNR 值为 42.397 1 dB,水印的嵌入对视频视觉效果影响很小。



(a) 水印嵌入前



(b) 水印嵌入后

图 3 水印的嵌入对视频视觉效果的影响

将本文算法与不采用视觉显著模型的普通算法^[9]做对比,得到 9 个嵌入水印关键帧的 PSNR 值,结果如图 4 所示。

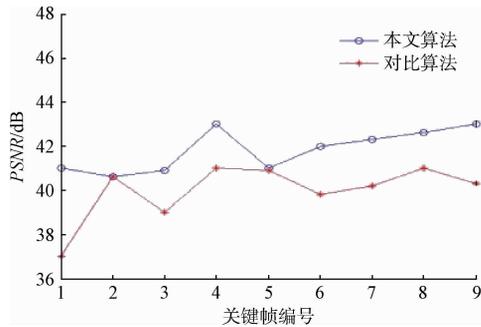


图 4 嵌入水印关键帧 PSNR 值对比

实验结果表明本文算法每个嵌入帧的 PSNR 值均高于对比算法,平均 PSNR 值对比算法提高了 3.921%,具有良好的不可见性。

3.2 鲁棒性分析

鲁棒性是指嵌入水印的视频帧遭受各种攻击之后,恢复原始水印的能力。在水印为二进制比特流的情况下, BER(误码率)是用来衡量鲁棒性的重要指标。

$$BER = \text{水印提取错误位数} / \text{嵌入水印位数} \quad (7)$$

BER 值越小,提取出来的水印错误率越低,水印算法鲁棒性越高^[10]。

将本文算法和对比算法(不采用视觉显著模型的算法)嵌入水印后得到的视频帧施加 MPEG-2 压缩和外部攻击双重攻击,计算 BER 值,如表 1 所示。

表 1 视频双重攻击后嵌入帧的平均 BER 值对比

攻击类型	本文算法 BER 值/%	对比算法 BER 值/%
乘性噪声强度 0.005	0.868	6.944
乘性噪声强度 0.02	5.208	12.523
高斯噪声方差 0.01	0.347	6.251
高斯噪声方差 0.04	6.163	12.011
椒盐噪声密度 0.005	2.431	7.732
椒盐噪声密度 0.03	9.681	14.672
饱和度幅度 +10	0.954	7.135
饱和度幅度 -10	0.954	6.283
均值滤波 3 * 3	6.295	9.752
均值滤波 5 * 5	32.382	40.685
无攻击	0	5.545

实验结果表明本文算法能够抵抗大多数攻击,包括乘性噪声、高斯噪声、椒盐噪声、饱和度、均值滤波等,具有较强的鲁棒性。

4 结 论

本文提出一种基于 Itti 视觉显著模型的音视频交叉算法。将视觉注意机制应用到音视频交叉水印中,在水印嵌入时,首先利用音频码流的低通幅值统计特性生成水印;然后获取视频关键帧的亮度分量;最后借助 Itti 视觉显著模型,将水印嵌入到非显著区域的 DCT 直流系数上,保证嵌入水印对视频影响在视觉失真范围之内^[11]。实验结果表明,本文算法具有较好的不可见性和鲁棒性,同时水印与音频特性相结合,增加了水印的安全性,为后续音视频交叉水印的研究提供参考。

参考文献

- [1] HU H T, HSU LY. Robust, transparent and high capacity audio watermarking in DCT domain [J]. *Signal Processing*, 2015, 109(C): 226-235.
- [2] 惠巧娟,李洪安,陆焱. 一种基于小波变换和人类视觉系统的图像压缩算法[J]. *电子测量与仪器学报*, 2016, 30(12): 1838-1844.
- [3] 王俊. 一种基于 FFMPEG 的 MPEG-2 视频水印算法[J]. *电子测量技术*, 2016, 39(6): 85-88.
- [4] 赵杰. 一种基于 DCT 量化的视频水印算法[J]. *电子测量技术*, 2016, 39(6): 72-75.
- [5] PATEL M J, BORISAGAR K R. A review on audio watermarking techniques[J]. *Journal of Open Source Developments*, 2014, 1(2): 1-5.
- [6] BISWAS S N, HASAN T, DASGUPTA S, et al. Compressed video watermarking technique [C]. *Instrumentation and Measurement Technology Conference (I2MTC)*, 2013 IEEE International. IEEE, 2013: 1790-1794.
- [7] 闻新,张婉怡,王嘉轶,等. 基于视觉感知的全参考图像质量评价算法[J]. *电子测量与仪器学报*, 2016, 30(11): 1780-1789.
- [8] 潘洋绪,刘猛. 基于 DCT 系数的量化音频水印[J]. *电子测量技术*, 2015, 12(38): 46-48.
- [9] 胡棚,施惠娟. 基于视觉掩蔽模型的视频水印算法[J]. *电子测量技术*, 2014, 12(37): 36-40.
- [10] 吴岳,施惠娟. 基于 FFMPEG 的视频水印系统[J]. *电子设计工程*, 2013, 21(23): 185-187.
- [11] LEI B, SOON Y, TAN E L. Robust SVD-based audio watermarking scheme with differential evolution optimization [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2013, 21(11): 2368-2378.

作者简介

汪新园,1992 年出生,硕士研究生,主要研究方向为音视频交叉水印。

E-mail: 1797331294@qq.com

潘洋绪,1990 年出生,硕士研究生,主要研究方向为信号与信息处理、视频水印。

E-mail: pyangxuhhu@163.com