

基于 HVS 的自适应鲁棒视频水印算法^{*})

陈光喜 成彦

(桂林电子科技大学数学与计算科学学院 桂林 541004)

摘要 视频水印与静止图像比较,包含有运动部分,具有变化的特性。根据人眼对快速运动的物体敏感度会下降的特点及 DCT 域视觉掩蔽模型,提出一种自适应的鲁棒视频水印算法。算法首先通过比较 I 帧与相邻帧的同一区域的 DCT 直流系数,得到一个快速运动图像块集,然后在图像块集中根据视觉掩蔽模型计算图像块的 DCT 中频系数,选择区域自适应嵌入水印。仿真实验表明,该水印算法不仅能很好地抵抗 MPEG 压缩,而且能较好地协调不可见性和鲁棒性,并能抵抗一定的几何攻击。

关键词 数字水印, DCT 变换, 人眼视觉特性

Robust Adaptive Video Watermarking Algorithm Based on HVS

CHEN Guang-xi CHENG Yan

(College of Computing Sciences and Mathematics, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract A robust adaptive Video Watermarking algorithm based on Human Visual System (HVS) was proposed. By comparing the direct current DCT coefficients of the I frame and its neighbor frame of the video, a fast moving image block set was chosen. The new algorithm embedded watermarking bits adaptively in proper place for some image block, which was decided based on HVS. Experiments show that the proposed algorithm is robust against MPEG compression, also the attacks of RST.

Keywords Digital watermarking, Video watermarking, Discrete cosine transform (DCT), Human visual system (HVS)

1 引言

数字媒体和计算机网络的迅速发展为社会生活提供了巨大便利。与模拟媒体相比,数字媒体具有明显优势;其质量远远优于相应的模拟形式,由于可以精确访问媒体信息的各个离散成分,对信息的编辑变得轻而易举,可以实现对媒体信息的完全一致的精确复制,方便在网络上进行传输。然而数字媒体的这些优势,尤其对数字信息的完全一致的精确复制,对传统的版权保护方法也提出了严峻的挑战,如何对数字作品的版权进行有效的保护,成为当今版权保护中亟待解决的问题。

作为数字作品版权保护的一种有效的方法,数字水印技术就是在这个背景下提出的。数字水印是指嵌入在原始数据中的由版权所有者确定的版权信息,用以证实数据的所有权归属。

由于视频信号比静止图像多出一维空间,通常认为视频序列具有更大的空间来嵌入水印。但是大量实验表明,嵌入水印很容易引起视频质量的下降,而且会产生在静止图像水印中不会出现的其它问题。

人类视觉的敏感度在各种环境下是不同的,例如对于快速运动的物体人眼敏感度会有所下降。基于此,本文提出了一种基于人类视觉系统的鲁棒自适应视频水印算法,有以下特点:(1)水印的嵌入选择在 I 帧中进行;(2)水印嵌入由当前帧与前/后帧同一位置的 DCT 变换后得到的直流系数差值与

某一阈值的比较所确定的快速区域中;(3)利用 Watson^[1] DCT 域视觉掩蔽模型在快速区域中自适应嵌入水印。

算法同时考虑帧内与帧间的关系,根据快速分区思想对原始视频图像块进行分类,就人类视觉系统来说,掩蔽阈值受到多种因素的影响,因此本算法根据 Watson 视觉模型在快速区域中选择松弛度最大的区域嵌入水印,使得水印嵌入位置自适应于视频图像的内容。算法将人类视觉特性运用到数字水印技术中,从而使得嵌入的水印具有很好的不可见性和鲁棒性。

2 MPEG 帧图像类型、快速区域的判定与 Watson 视觉模型

2.1 MPEG 帧图像类型

根据运动估计方向的不同,编码图像可以分为 I 帧(Intra frame)、P 帧(Predicted frame)和 B 帧(Bidirectional frame)。I 帧图像是利用图像自身的相关性压缩,提供压缩数据流中的随机存取点,采用基于 DCT 的编码技术,编码不需要其他帧的图像参考。这些帧图像为译码器提供随机存取点,是预测图像(P)和双向预测图像(B)帧的参考图像,所以压缩率不高。P 帧图像则是参考过去的帧内图像或者过去预测得到的图像用运动补偿预测技术来进行编码,通常作为进一步的参考,所以编码效率高。B 帧图像在预测时,既可使用前一个图像做参照,也可使用下一个图像作为参照或者同时使用前后两个图像作为参照图像(双向预测),它的压缩率最高。由

^{*})本文工作得到国家自然科学基金(编号 10501009 和 10661005)的支持。陈光喜 硕士生导师,副教授,博士,主要研究方向为数字水印与信息隐藏;成彦 硕士研究生,主要研究方向为数字水印。

此可以知道如果将水印嵌入到 I 帧,就能很好地抵抗 MPEG 压缩。

2.2 快速区域的判定

对于由 L 帧图像组成的一段视频,设每帧大小为 $M \times N$,令 V_l 表示视频中的第 l 帧其中 $l=0,1,\dots,L-1$ 。首先将当前帧 V_l 分割成互不重叠的大小为 8×8 的图像块 $V_{l,i}$,这里 $V_{l,i}$ 表示第 l 帧中的第 i 个图像块, $i=0,1,\dots,[M \times N/64-1]$;接着对每一图像块进行 DCT 变换,得到 DCT 系数块 $\hat{V}_{l,i}$ 。将 DCT 系数按“之”字形顺序排列,其中 $\hat{V}_{l,i,0}$ 表示直流系数,以相同的方法对 $l-1$ 帧或者 $l+1$ 帧进行 DCT 变换(当 l 帧为最后帧时采用 $l-1$ 帧其他采用 $l+1$ 帧),将 l 帧 DCT 系数块与 $l-1$ 帧、 $l+1$ 帧中对应的 DCT 系数块进行比较,得到直流系数的差值为 $\delta_{DC}(l,i)$:

$$\delta_{DC}(l,i) = |\hat{V}_{l,i,0} - \hat{V}_{l+1,i,0}|$$

设定一阈值 T_M ,若 $\delta_{DC}(l,i)$ 小于 T_M ,则该图像块 $V_{l,i}$ 被划分为慢速区域;否则 $V_{l,i}$ 被划分为快速区域^[2,3]。

2.3 Watson 感知模型

Watson 感知模型^[1]利用灵敏度、掩蔽性以及 Pooling 的思想对图像之间的 JND(Just noticeable difference)值做出估计。主要思想为:对一幅图像在经过分块 DCT 之后所得到的每个系数的可感知性进行评估,综合成感知距离 $D_{wat}(p, p^e)$,其中 p 是原始图像, p^e 是原始图像失真后所得到的某种版本的图像。模型给出的视觉感知误差门限 η_{ij} 的含义是:如果分块 DCT 系数中的某一项发生了改变,有时会使得图像产生可觉察的差异。使图像产生最小可觉察差异的阈值 η_{ij} 称为松弛度:

$$\eta_{ij} = \max\{t_{l,i,j}^e, |\hat{p}_{l,i,j} - \hat{p}_{l,i,j}^e| \cdot \Psi_{ij} \cdot (t_{l,i,j}^e)^{1-\Psi_{ij}}\}$$

此处 Ψ_{ij} 是 0 到 1 之间的一个常数,具体针对每一个频率系数其值会有所不同; $\hat{p}_{l,i,j}$ 为分块 DCT 系数, $t_{l,i,j}^e$ 为亮度掩蔽阈值:

$$t_{l,i,j}^e = t_{l,i,j} (\hat{p}_{l,0,0} / \ell)^\phi$$

其中 ϕ 是一个常数,建议取值为 0.649; $\hat{p}_{l,0,0}$ 代表原始视频图像的第 l 个图像块的直流系数; ℓ 为一个常数,代表所期望的图像像素值; $t_{l,i,j}$ 为块中的灵敏系数。

为比较原始图像 p 和产生失真的图像 p^e ,先计算相应的 DCT 系数之间的差值:

$$\Omega_{kij} = \hat{p}_{l,i,j}^e - \hat{p}_{l,i,j}$$

其中 $\hat{p}_{l,i,j}^e$ 和 $\hat{p}_{l,i,j}$ 分别表示 p^e 和 p 对应分块的 DCT 系数。然后分别用相应的松弛度 η_{ij} 对这些差值进行比较调整,从而获得第 k 块第 i, j 个频率分量所对应的感知距离 d_{kij} (单位为 JND):

$$d_{kij} = \frac{\Omega_{kij}}{\eta_{ij}}$$

得到图像总体失真的一个估计量,即感知距离:

$$D_{wat}(p, p^e) = (\sum_{l,i,j} |d_{kij}|^\theta)^\frac{1}{\theta}$$

本文算法采用 Watson 感知模型的一个重要原因就是它与 MPEG 压缩的方法以及快速区域的判定都用到了相同的过程:先将图像分成互不重叠的 8×8 图像块,然后进行 DCT 变换。这样就能很好地减少复杂度,实现简洁快速的程序计算。

3 水印的嵌入算法

本文的水印嵌入过程是在 I 帧中的分块 DCT 域内进行的,算法在判定出快速运动区域块集后,以 Watson 提出的 DCT 域视觉掩蔽为基础,同时借鉴了文献[4]的抖动调制技术。本算法的水印嵌入的流程图如图 1 所示。

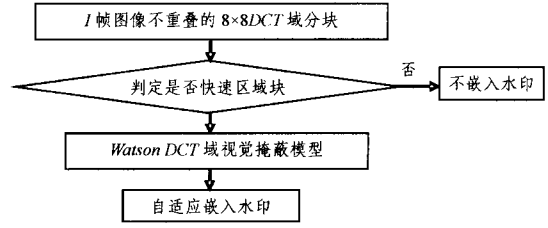


图 1

先对待嵌入的水印信息进行处理:利用线性反馈移位寄存器生成的伪随机序列对一维水印序列进行随机置乱,以提高嵌入水印的鲁棒性,其中伪随机序列发生器的结构和种子可由水印序列的长度和用户的密钥决定。这样就得到随机置乱的水印信息一维序列:

$$W = \{\omega_i, 0 \leq i \leq n\}.$$

现在将待嵌入水印的视频帧 V 分成互不重叠的 8×8 图像块,记为 $v_l(i, j)$, $l=0,1,\dots,n-1$,即有

$$V = \bigcup_{l=0}^{n-1} v_l(i, j), 0 \leq i, j \leq 8$$

对每一个图像块进行 DCT 变换,得到 DCT 系数块 $\hat{v}_l(i, j)$ 。将 DCT 系数按“之”字形顺序排列,其中 $\hat{v}_l(0,0)$ 表示直流系数。按照快速区域判定法得到一个图像块集合,表示为 $P = \{v_{m_1}(i, j), v_{m_2}(i, j), \dots, v_{m_q}(i, j)\}$

为了更好地折中不可见性和鲁棒性,我们选择在中频区域嵌入水印。假设 $\hat{v}_{m_k}(x, y)$ 表示选取的 DCT 中频系数,根据 Watson DCT 域视觉掩蔽模型,分别计算每一个 8×8 图像块中 DCT 系数 $\hat{v}_{m_k}(x, y)$ 所对应的松弛度 $S_{m_k}(x, y)$,并按照从大到小顺序排列。选取与待嵌入的水印信息相同数目的松弛度最大的图像块,并确定这些图像块和经过置乱的水印信息位 ω_i 之间一一对应的关系。根据文献[5]中所论述的抖动调制思想对 DCT 中频系数 $\hat{v}_{m_k}(x, y)$ 进行量化处理。假设量化处理后的结果为 $\hat{v}_{m_k}^w(x, y)$,则有

$$\hat{v}_{m_k}^w \in \begin{cases} \text{I, 如果 } \omega_i = 1 \\ \text{II, 如果 } \omega_i = 0 \end{cases}$$

即水印比特 ω_i 的信息由 DCT 系数 $\hat{v}_{m_k}^w(x, y)$ 所属的 I 区域集合或者 II 区域集合来表示,最后进行 DCT 逆变换,从而得到含水印的视频图像 $V^w(i, j)$ 。此处要注意量化步长的大小选择,本文采用文献[5]所给出的量化步长取值范围。

水印嵌入算法步骤总结如下。

步骤一:对待嵌入的水印信息进行预处理,利用线性反馈移位寄存器生成的伪随机序列对一维水印进行随机置乱,其中伪随机序列发生器的结构和种子由水印序列的长度和用户的密钥决定。

步骤二:将某一 I 帧与其相邻帧进行不重叠的 8×8 分块,并进行 DCT 变换,根据快速区域的判定原则,得到一个图像块集。

步骤三:对上述图像块集中每一图像块进行 DCT 变换,根据 Watson 提出的 DCT 域视觉掩蔽模型选择图像块并在该图像块的 DCT 域中频区域来嵌入水印,整个水印的嵌入过程充分考虑原始载体的局部特征,使得水印嵌入的位置、强度随着局部特征的变化而变化,即实现水印嵌入的自适应性。

4 水印的提取

该算法水印的提取需要原始视频的参与。首先对原始视频的 I 帧以及其后一帧进行不重叠的 8×8 分块操作并进行 DCT 变换,然后根据快速区域的判断法得到 I 帧中的图像块集 P ,其次根据 DCT 域视觉掩蔽模型计算中频系数 $\hat{v}_{m_k}^{\wedge}(x, y)$ 对应的松弛度并根据松弛度的大小进行排序,标记出对应松弛度最大的 n 个图像块,与含水印视频的 I 帧中找到相应的水印嵌入的位置,判断含水印的 DCT 系数 $\hat{v}_{m_k}^{\wedge}$ 属于 I 区域集或者 II 区域集。水印的提取算法可表示为:

$$\omega_i = \begin{cases} 1, & \text{如果 } \hat{v}_{m_k}^{\wedge}(x, y) \in I \\ 0, & \text{如果 } \hat{v}_{m_k}^{\wedge}(x, y) \in II \end{cases}$$

根据嵌入算法中所使用的伪随机序列,按照相应的排序规则对提取出来的水印序列进行逆排序,从而恢复出水印信息。

5 仿真实验结果

本实验的程序用 Matlab 与 VC++ 实现。实验中采用 Container 和 Mobile 视频为例,对算法进行测试。在测试中,对感知距离的参数 θ 设置为 4,量化步长设置为 25,试验中仅考虑视频序列的亮度分量。

图 2 和图 3 给出了原始视频图像以及相应的含水印视频图像,采用 Watson 模型中的感知距离来评价和控制含水印视频图像质量。Container 在嵌入水印后的峰值信噪比 PSNR 为 42.78, Mobile 在嵌入水印后的峰值信噪比 PSNR 为 41.56,对应的感知距离为 5.72 和 5.28,这里感知距离的单位为 JND。

为验证水印对于 MPEG 压缩的鲁棒性,实验采用 MPEG 工作组提供的标准编码器来实现 MPEG-2 压缩编码以及相应的解码过程,表 1 列出了经过 MPEG-2 攻击后的水印提取结果。仿真实验表明,该算法对 MPEG 压缩有较好的鲁棒性。进一步验证水印对于几何攻击的鲁棒性,利用 Stirmark 软件来模拟实际应用中所产生的几何攻击,主要包括旋转、尺度变换、剪切、长宽对比调整、弯曲等等。在算法测试过程中,通过剪裁或者填充方法将经过攻击的含水印图像恢复成原始大小,然后进行水印提取。表 2 列出经过 Stirmark 攻击后的水印提取结果。仿真实验结果表明,该算法对各种几何攻击具有较好的鲁棒性。

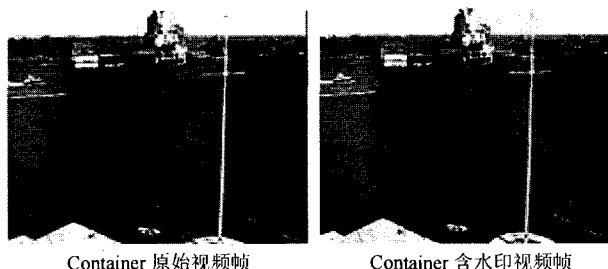
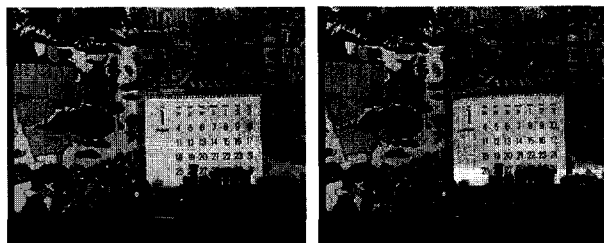


图 2 Container 视频的实验对比



Mobile 原始视频帧

Mobile 含水印视频帧

图 3 Mobile 视频的实验对比

表 1 经过 MPEG-2 攻击后的水印提取结果

MPEG-2 编码率	对应的 压缩比	提取水印信息的正确率(%)	
		Container	Mobile
6mbps	12.5	100	100
5mbps	14.7	100	100
4mbps	18.3	100	97.3
3mbps	24.4	95.3	90.6
2mbps	36.4	90.8	85.6

表 2 经过 Stirmark 攻击后的水印提取结果

Stirmark 攻击		提取水印信息的正确率(%)	
		Container	Mobile
旋转 角度	1°	100	100
	5°	98.5	100
	10°	90.2	92.3
缩放 比例	0.5	100	95.7
	0.8	92.5	89.4
弯曲 水平翻转	1.2	85.4	72.3
		100	99.5
		100	100

结束语 本文在基于人类视觉系统的基础上提出了一种自适应的鲁棒视频水印算法,算法的特点表现在三个方面:一方面在对视频帧进行不重复的图像分块后利用人眼的视觉特性,对帧与帧之间相同图像块位置图像判断其是否属于快速区域,并以视觉模型来自适应嵌入水印,以此来很好地协调不可见性和鲁棒性;第二方面从 MPEG 的压缩方面考虑把水印嵌入 I 帧;最后从抗几何攻击的方面考虑将水印嵌入 DCT 域的中频区域。所有这些都使得该算法能很好地抵抗 MPEG 压缩和几何攻击的同时能协调不可见性和鲁棒性。由于该算法在检测时还是需要原始视频的参与,因此实现算法的盲检测是今后工作的努力方向。

参考文献

- [1] Watson A B. DCT Quantization Matrices Visually Optimized for Individual Images[J]//Proceeding of SPIE, Human Vision, Visual Processing, and Digital Display IV, 1993, 1913:286-291
- [2] Lu Z M, Ge Q M, Niu X M. Rubust Adaptive Video Watermarking in the Spatial Domain[C]//The 5th International Symposium on Test and Measurement (ISTM' 2003). Shenzhen, China, June, 2003:1875-1880
- [3] Lu Z M, Ge Q M, Niu X M. Oblivious Video Watermarking Scheme with Adaptive Embedding Mechanism[C]//The Second International Conference on Machine Learning and Cybernetics, Xi'an, China, November, 2003:2876-2881
- [4] Chen B, Wormell G W. Digital Watermarking and Information Embedding Using Dither Modulation[J]//Proceeding of IEEE, Multimedia Signal Processing, 1998:273-278
- [5] 葛庆鸣. 数字视频水印算法研究[D]. 硕士学位论文. 哈尔滨:哈尔滨工业大学, 2003