

## 基于熵模型的高透明性自适应视频水印算法<sup>\*</sup>

李智<sup>+</sup>, 陈孝威

(贵州大学 计算机科学与技术学院, 贵州 贵阳 550025)

### Adaptively Imperceptible Video Watermarking Algorithm Using Entropy Model

LI Zhi<sup>+</sup>, CHEN Xiao-Wei

(College of Computer Science and Technology, Guizhou University, Guiyang 550025, China)

+ Corresponding author: E-mail: lizhigzu@163.com; gzu@vip.sina.com

**Li Z, Chen XW. Adaptively imperceptible video watermarking algorithm using entropy model. *Journal of Software*, 2010,21(7):1692-1703.** <http://www.jos.org.cn/1000-9825/3569.htm>

**Abstract:** An adaptively imperceptible video watermarking algorithm using entropy model is proposed. This algorithm works by first combining the Human Visual System (HVS) with block-matching techniques to obtain motion-related information. It then uses entropy model to statistically analyze the obtained information, and eventually establish a set of non-linear formulas based on HVS and motion information. Based on the contents of video frames, this set of non-linear formulas can adaptively calculate the maximum strength of every block. Experimental results show that using entropy model and non-linear formulas can significantly improve watermarking imperceptibility, effectively resist common attacks for video watermarking, and consequently achieve higher robustness.

**Key words:** human visual system; block motion estimation of video; largest embedding strength of watermarking; model of entropy

**摘要:** 提出一种基于视频运动估计熵模型的自适应视频水印算法.该算法将人类视觉系统(human visual system, 简称 HVS)与视频分块运动估计(block motion estimation of video)相结合,获取视频序列帧中与运动相关的视频运动信息,然后利用熵模型对视频序列帧中的运动信息进行统计,从而得到一组基于视频序列帧间运动信息与人类视觉屏蔽特性相结合的非线性计算公式.利用该组计算公式,可以根据视频帧的内容自适应地计算每个方块的水印最大嵌入强度.实验结果表明,熵模型与非线性公式的引入较大幅度地提高了视频水印的透明性,并且能够有效地抵抗常见的针对视频水印的攻击,具有较高的安全性和鲁棒性.

**关键词:** 人类视觉系统(HVS);视频分块运动估计;水印最大嵌入强度;熵模型

中图法分类号: TP309      文献标识码: A

近年来,随着国际互联网(Internet)技术的迅猛发展和各种视频编码标准的颁布,媒体服务商可以向用户提

---

\* Supported by the Science and Research Foundation of Guizhou Province of China under Grant No.20052109 (贵州省科研基金项目); the Governor Foundation of Guizhou Province of China under Grant No.200714 (贵州省省长专项资金); the Science and Technology Foundation of Guizhou Province of China under Grant No.20102257 (贵州省科学技术基金)

Received 2007-11-12; Revised 2008-03-11; Accepted 2008-12-29

供各式各样的多媒体服务.但与此同时,对这些多媒体资源进行非法篡改、复制和发行也比较容易.因此,如何对多媒体产品进行版权保护的问题日益凸现.数字水印再次成为解决多媒体产品版权保护问题的重要技术手段.视频水印的研究已成为信息安全和多媒体领域研究的热点课题<sup>[1-10]</sup>,但视频水印算法的研究远远滞后于静止图像水印算法的研究,主要原因有:(1) 多年来,众多学者在静态图像的视觉空间掩蔽效应领域获得了很多研究成果,从而直接推动了静态图像水印算法的研究;(2) 视频序列本身的一些固有特点,如数据冗余过大、运动区域与非运动区域分布不平衡等,也使得视频水印研究有相当的难度.而综合时间域(temporal domain)与空间域在内的更为精确的人眼视觉模型尚未完全建立,还很少有文章提出一种能够自适应地计算视频空间和时间掩蔽效应的视频水印方法.

在文献[1]中,Liu 等人提出对视频中每帧进行多层离散小波变换,利用视频中运动和纹理两大特性,选择纹理丰富并且有运动物体存在的区域,获得较高的水印嵌入强度.但是,文中水印嵌入量的计算只与纹理和运动相关,并没有考虑物体的运动速度、运动方向等因素.文献[2]中,Echizen 等人提出利用两帧间的运动矢量长度(length)和物体变形(deformation)判断视频序列中的物体是否处在运动状态.对于非运动状态的物体,利用静态图像水印算法中有关视觉阈值的计算方法获取水印嵌入强度;而对于运动物体,则利用固定系数 $\alpha$ 乘以静态物体的视觉阈值作为运动物体的水印嵌入强度.文中虽然使用帧间的运动特征对物体的运动与否进行判断,但是在计算视频水印的嵌入强度时,并没利用这些运动特征参与到水印嵌入强度的计算中,即没有根据视频序列中的不同运动内容自适应地获取水印的嵌入强度.通过文献[2,3]以及我们前期所作研究发现,运动速度与运动方向等帧间运动特性对人类视觉会产生一定的影响,当同时对运动速度较低的物体与静止物体嵌入相同量的水印信号时,比较后发现,运动速度较低的物体中所嵌入的水印更容易被人眼所察觉.文献[4]提出计算视频序列帧间的运动矢量,并将这些运动矢量进行量化,对量化后的运动矢量采用求熵的统计方法获取一个统计值,通过熵值来判断视频序列中的运动特性.根据每个视频序列不同的运动特性,自适应地计算该视频序列的水印嵌入强度,但是对运动矢量进行量化会引入一定量的误差,从而影响水印信号的有效检测和正确提取.文献[5]提出几何不变量——基于圆区域内的统计特征不变性,依据离散余弦变换域(discrete cosine transform,简称 DCT)系数特性,通过自适应地对 DCT 域系数进行量化,从而嵌入有意义的水印.但是,水印的嵌入与检测需要同步信息的参与,对于帧平均、帧交换等针对视频水印所特有的攻击实验,并未进行鲁棒性实验结果的阐述.文献[6]采用三维离散小波变换算法在时间和空间两个尺度上对视频序列帧进行小波变换,选取合适的频率子带添加水印信息.文中虽然利用离散小波变换良好的时空局部特性进行水印信号嵌入,但是没有充分考虑人眼视觉系统对于视频帧的掩蔽特性,对水印嵌入强度是否与视频帧的内容有关,未进行详细的分析和介绍.文献[7]提出计算图像帧内每个宏块运动矢量的特征矢量,利用该特征矢量与待嵌入的水印比特流计算得到一个特征值,根据特征值选取几个水印嵌入待选方案中的一个,水印信息的嵌入是通过对运动矢量的微小调整将水印信息嵌入到该宏块的运动矢量上.但是,文章中并没有考虑和分析视频运动矢量对人眼视觉特性的影响以及对水印嵌入容量的关系.在我们的前期工作中发现,只是简单地在视频序列的运动矢量上添加水印信息存在如下问题:首先,在运动矢量上添加水印只能通过对运动矢量的微小调整来实现,因而水印嵌入强度较少,抗攻击能力有限;其次,在运动矢量上直接嵌入水印会造成运动矢量速度大小和方向的变化,使得视频的动态视觉质量下降.

## 1 视觉屏蔽(visual masking)特性

视觉屏蔽特性是指心理学家和计算机学家通过对人类视觉成像现象和规律的研究和实验,获取能够提高图像可视质量、以视觉感受为基础的高效算法<sup>[11]</sup>.本文首先将视频序列分割为帧(frame),选取每帧中的亮度通道  $Y$  作为水印嵌入的载体,将亮度通道  $Y$  分割为一系列互不重叠的  $16 \times 16$  的图像宏块  $M(x,y)$ ,  $x,y$  为该图像宏块左上角在视频帧内的位置坐标.以宏块为单位,利用视频标准中的块匹配算法(block-matching techniques)计算出基于视频时间序列(temporal axis)的多种运动特性,如速度大小(velocity)、运动方向、变形尺度(deformation)等,并以这些运动特性为研究基础,计算出基于视频时间序列的人类视觉屏蔽阈值( $JND_t$ ).同时,将视频序列中的帧视为静态图像,利用文献[12]中的算法计算静态图像的人类视觉屏蔽阈值( $JND_s$ ),然后,利用非线性公式将静

态图像的人类视觉屏蔽阈值与视频时间序列的人类视觉屏蔽阈值相结合,获取视频水印嵌入算法中每个图像方块的水印嵌入强度( $JND$ ).最后,将亮度  $Y$  通道分割为一系列互不重叠的  $8 \times 8$  的图像方块  $B(x,y)$ ,  $x,y$  为该图像方块左上角在视频帧内的位置坐标.对相邻的两个图像方块进行一层二维离散小波变换(discrete wavelet transform,简称 DWT),选择两图像方块相对应的 3 个细节子带( $HL,HH,HL$ )进行水印的嵌入.本文算法的流程图如图 1 所示.

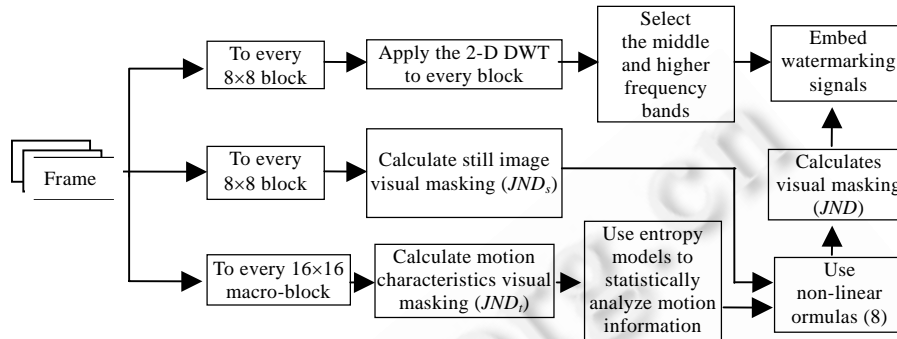


Fig.1 Flow chat of the algorithm

图 1 算法流程图

### 1.1 基于静态图像的人类视觉屏蔽阈值计算

为了保证视频水印的不可见性,我们在考虑视频序列的空间尺度和运动特性对人类视觉屏蔽特性影响的同时,充分考虑静态图像中人类视觉屏蔽特性<sup>[11-14]</sup>.对人类视觉系统中如下几个对于静态图像的视觉屏蔽特性进行了分析和计算:(1) 基于纹理的视觉屏蔽特性;(2) 基于边缘的视觉屏蔽特性;(3) 基于亮度的视觉屏蔽特性;(4) 基于小波域中不同频率的视觉屏蔽特性.通过综合这几个视觉屏蔽特性,计算获取每帧中静态图像方块  $B(x,y)$  的视觉屏蔽阈值  $JND_s(x,y)$ <sup>[12]</sup>.在实验中,我们对  $JND_s(x,y)$  进行归一化处理获得  $JND'_s(x,y)$ .为简化符号,以下用  $JND_s(x,y)$  表示归一化的  $JND'_s(x,y)$ .

### 1.2 图像宏块 $M(x,y)$ 的运动信息检测

视频序列不同于静态图像,在视频播放中,不仅存在大量的静态背景,还同时存在许多运动物体.在研究基于视频时间序列的人类视觉屏蔽阈值计算之前,应当首先将视频中的运动物体从静态的背景中分离出来.因此,需要首先对图像宏块  $M(x,y)$  内的运动信息进行检测.

我们使用两个相邻视频序列帧中图像宏块对应像素点间灰度值的差  $Dif(x,y)$  及标准方差  $Var(x,y)$  来表征视频序列帧内的图像宏块  $M(x,y)$  的运动特征.由于只选取亮度通道作为水印嵌入的载体,所以在以下的叙述中都只考虑亮度通道  $Y$ .

设  $aveY_c$  和  $aveY_p$  分别表示当前帧和前一帧的亮度平均值,  $Y_c(x,y)$  和  $Y_p(x,y)$  分别表示当前帧与前一帧亮度通道内坐标为  $(x,y)$  的像素值,则当前帧与前一帧像素偏离其平均值的差  $Dif_{frame}(x,y)$  为

$$Dif_{frame}(x,y) = ((Y_c(x,y) - aveY_c) - (Y_p(x,y) - aveY_p)) / 2 \quad (1)$$

用  $Dif(x,y)$  表征视频序列帧内每个图像宏块  $M(x,y)$  的运动特征,其计算见公式(2):

$$Dif(x,y) = \sum_{i=1}^{16} \sum_{j=1}^{16} |Dif_{frame}(x+i, y+j)| \quad (2)$$

式中,  $i,j$  是指像素在该图像宏块内的位置坐标.通过计算  $Dif(x,y)$  和每个图像宏块的标准方差  $Var(x,y)$ , 可以判断和检测图像宏块的运动信息.当  $Dif(x,y)$  与  $Var(x,y)$  的值都比较大时,该对应的图像宏块包含有较多的运动信息;当  $Dif(x,y)$  与  $Var(x,y)$  中任意一个的值较小时,则对应的图像宏块内具有较少的运动信息.因此,对视频序列进行训练可以获取两个阈值  $Dif_t$  与  $Var_t$ , 用来判断图像宏块内是否存在运动.即,当  $Dif(x,y) \geq Dif_t$  且  $Var(x,y) \geq Var_t$  时,

则对应的图像宏块存在有较多的运动信息;否则,该图像宏块内不存在或者存在较少的运动信息.我们在进行基于视频时间序列的人类视觉屏蔽阈值计算时,只针对包含有较多运动信息的图像宏块进行计算.本文中  $Dif_i$  与  $Var_i$  的值都必须根据视频序列的内容进行自适应的计算训练获得.

### 1.3 基于视频时间序列的人类视觉屏蔽阈值计算

#### 1.3.1 基于运动速度大小的视觉屏蔽阈值计算

视频水印与静态图像水印的最大区别在于视频序列中存在较多冗余的空间尺度与较多的运动信息.因此,必须将视频帧间(inter-frame)特性作为视觉屏蔽阈值计算的一个重要方面进行研究.我们采用前一节介绍的运动检测的方法,检测并选取每帧中存在较多运动信息的图像宏块进行后续计算.首先,利用 MPEG 中的块匹配算法,获取每个图像宏块的运动矢量  $MV(u,v)$ ;其中, $(u,v)$ 对应于 $(x,y)$ .然后,利用公式(3)计算出每个图像宏块的运动矢量长度,用于标识图像宏块内物体的运动速度的大小 $|MV(u,v)|$ .

$$|MV(u,v)| = \sqrt{(MV_x(u,v))^2 + (MV_y(u,v))^2} \quad (3)$$

式中, $MV_x(u,v)$ 和  $MV_y(u,v)$ 分别表示图像宏块的运动矢量  $MV(u,v)$ 在  $X$  轴方向和  $Y$  轴方向上的运动分量.如果  $|MV(u,v)|$ 较大,则该图像宏块内存在运动速度较快的物体,当把水印信号嵌入到该图像宏块时,能够获得较好的透明性;如果  $|MV(u,v)|$ 较小,则该图像宏块内存在运动速度较慢的物体.通过前期实验发现,在对低速运动物体进行水印信号嵌入时,很容易被人眼察觉,水印的透明性较差.因此,我们利用阈值  $V_i$  筛选出运动速度较低的一些图像宏块,对这些宏块不进行视频时间序列的人类视觉屏蔽阈值计算,而只计算其静态图像的人类视觉屏蔽阈值  $JND_s(x,y)$ .

#### 1.3.2 基于运动方向的视觉屏蔽阈值计算

由于运动物体包含运动大小和方向,在进行基于视频时间序列的人类视觉屏蔽阈值计算时,还应考虑物体的运动方向.本文用相邻两帧间所对应图像宏块的时间运动角度(temporal angle)来表征视频的帧间运动方向特性,用公式(4)计算获得当前帧图像宏块的时间运动角度.

$$\theta(u,v) = \arccos \frac{MV_c(u,v) \cdot MV_p(u,v)}{|MV_c(u,v)| \times |MV_p(u,v)|} \quad (4)$$

式中, $MV_c(u,v)$ 和  $MV_p(u,v)$ 分别是当前帧和前一帧对应位置图像宏块的运动矢量,“ $\cdot$ ”表示矢量计算的点乘.图像宏块的时间运动角度越大,说明该视频序列中包含有较多不同运动方向的物体.通过前期实验发现,包含较多不同运动方向物体的图像宏块适合嵌入水印.因此,我们使用  $\theta(u,v)$ 作为计算基于视频时间序列的人类视觉屏蔽阈值  $JND_t(u,v)$ 的一个分量.

#### 1.3.3 基于运动物体变形尺度的视觉屏蔽阈值计算

当视频序列中存在运动物体时,随着时间尺度的推移,运动物体会在每个帧内留下一定程度的变形(如膨胀、缩小、倾斜等).变形尺度的计算采用 MPEG 中的计算方法.具体如公式(5)所示:

$$D(x,y) = \sum_{i,j=0}^{15} |Y_c(x+i,y+j) - Y_p(x+k+i,y+l+j)|, \quad -15 \leq k,l \leq 15 \quad (5)$$

式中, $i,j$ 是指每个像素在该图像宏块内的位置坐标, $k,l$ 是图像宏块在相邻帧的搜索范围.

实验结果表明,变形尺度越大,嵌入的水印透明性就越好;变形尺度越小,嵌入的水印信号越容易被人眼察觉,水印的透明性较差.我们选择一个阈值  $D_i$ ,将图像宏块分为变形尺度较大和较小的两类.对变形尺度较大的图像宏块,计算视频时间序列的人类视觉屏蔽阈值  $JND_t(x,y)$ ;对变形尺度较小的图像宏块,只进行静态图像的人类视觉屏蔽阈值  $JND_s(x,y)$ 计算.

#### 1.3.4 基于视频时间序列的人类视觉屏蔽阈值计算

我们采用以下方法计算获得图像宏块的基于视频时间序列的人类视觉屏蔽阈值  $JND_t(u,v)$ <sup>[10]</sup>:

$$\text{if } Dif(x,y) \geq Dif_i \text{ and } Var(x,y) \geq Var_i \text{ and } |MV(u,v)| \geq V_i \text{ and } D(x,y) \geq D_i \\ JND_t(x,y) = D(x,y) \times \theta(u,v) \times |MV(u,v)|$$

else

计算每个图像宏块所包含的 4 个图像方块的  $j$  静态图像视觉阈值  $JND_s(x,y)$

end

在实验中,我们对  $JND_i(x,y)$  进行与静态图像的人类视觉屏蔽阈值计算相同的归一化处理获得  $JND'_i(x,y)$ . 为简化符号,以下用  $JND_i(x,y)$  表示归一化的  $JND'_i(x,y)$ .

## 2 熵模型与非线性公式结合获取视觉屏蔽阈值

### 2.1 视频的运动特性熵模型

自适应的水印嵌入算法是指利用人眼视觉系统的掩蔽特性,在载体中嵌入水印量达到最大值的情况下仍能保证水印具有透明性,即保证图像或者视频序列的可视质量不下降.本文采用信息论中熵的定义,并将熵的含义推广至视频序列中有关运动信息的统计.与文献[4]中所提及的算法不同,本文算法中没有对每个图像宏块的  $|MV(u,v)|$  和  $\theta(u,v)$  进行量化处理,不会产生由于量化所带来的误差影响.与文献[4]相比,用本文方法计算获取的视频帧的运动特性熵能够更真实地反映视频序列中的运动信息,从而更准确地获取视频序列的视觉屏蔽阈值的判断阈值. $|MV(u,v)|$  和  $\theta(u,v)$  的计算统计示意图如图 2 所示.

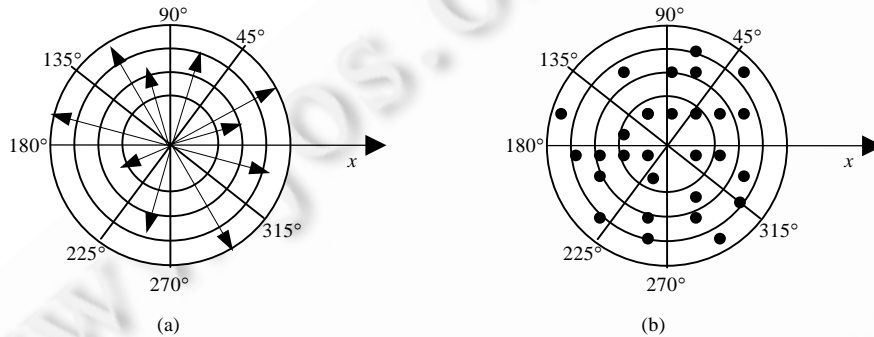


Fig.2 Model of entropy

图 2 熵模型

如图 2(a)所示,各个同心圆的半径大小代表各个图像宏块的运动矢量的运动速度大小,该矢量与  $X$  坐标轴之间的夹角则代表了各个图像宏块的运动矢量的运动方向.图 2(b)与图 2(a)类似,只是用黑点来标识各个视频序列帧的所有图像宏块的运动矢量.然后,对视频序列帧的所有图像宏块的运动矢量的  $|MV(u,v)|, \theta(u,v)$  进行熵模型的统计.具体计算如公式(6)所示:

$$H_f(|MV(u,v)|) = p(|MV(u,v)|) \log \frac{1}{p(|MV(u,v)|)}, H_f(\theta(u,v)) = p(\theta(u,v)) \log \frac{1}{p(\theta(u,v))}, f = 1, 2, \dots, N \quad (6)$$

式中  $f=1, 2, \dots, N$  表示视频序列中帧的数目,  $p(|MV(u,v)|)$  和  $p(\theta(u,v))$  分别表示视频序列中所有  $|MV(u,v)|$  和  $\theta(u,v)$  的出现概率,  $H_f(|MV(u,v)|), H_f(\theta(u,v))$  分别表示第  $f$  帧的  $|MV(u,v)|$  和  $\theta(u,v)$  的熵.在计算得到视频序列中每一帧的  $H_f(|MV(u,v)|), H_f(\theta(u,v))$  后,将所有帧的  $H_f(|MV(u,v)|)$  和  $H_f(\theta(u,v))$  分别投射到直方图上,对直方图进行分析计算,分别获取关于运动速度和运动方向的判断阈值  $T_v(\min), T_v(\max), T_\theta(\min), T_\theta(\max)$ . 再通过计算各帧的  $H_f(|MV(u,v)|)$  和  $H_f(\theta(u,v))$  获得整个视频序列帧的运动特性熵  $H(|MV(u,v)|, \theta(u,v))$ :

$$H(|MV(u,v)|, \theta(u,v)) = \frac{\sum_{f=1}^N H_f(|MV(u,v)|, \theta(u,v))}{N} \quad (7)$$

### 2.2 计算视觉屏蔽阈值

利用上一节计算得到的判断阈值  $T_v(\min), T_v(\max), T_\theta(\min), T_\theta(\max)$ , 对视频序列的运动特点进行分类描述和计算. 根据  $T_v(\min), T_v(\max)$ , 将视频中的运动速度分为 3 种情况: 当  $H(|MV(u, v)|) \leq T_v(\min)$  时, 表明该视频序列运动速度较为缓慢; 当  $T_v(\min) < H(|MV(u, v)|) \leq T_v(\max)$  时, 表明该视频序列运动速度居中; 当  $H(|MV(u, v)|) > T_v(\max)$  时, 表明该视频序列的运动速度较快. 用  $T_\theta(\min)$  和  $T_\theta(\max)$  将视频中的物体运动的方向进行分类, 得到 3 种情况: 当  $H(\theta(u, v)) \leq T_\theta(\min)$  时, 表明该视频中物体运动方向趋于一致; 当  $T_\theta(\min) < H(\theta(u, v)) \leq T_\theta(\max)$  时, 表明视频中物体运动方向存在少量的不一致; 当  $H(\theta(u, v)) > T_\theta(\max)$  时, 则表明该视频序列中物体的运动方向存在较多的不一致. 具体分类如图 3 所示.

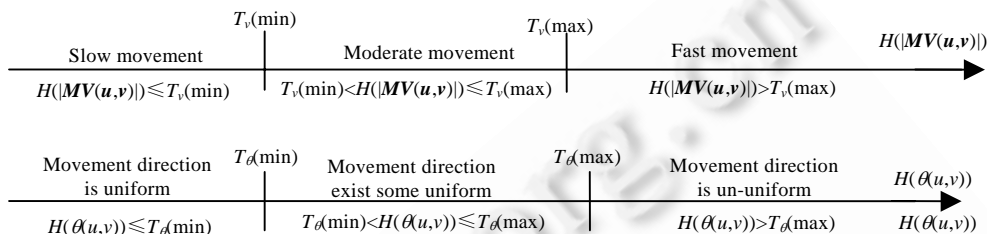


Fig.3 Classified the  $H(|MV(u, v)|), H(\theta(u, v))$

图 3  $H(|MV(u, v)|), H(\theta(u, v))$  分类

如果  $H(|MV(u, v)|) > T_v(\min)$  且  $T_\theta(\min) < H(\theta(u, v)) \leq T_\theta(\max)$ , 则表明该视频序列中具有速度居中和运动速度较快的图像宏块, 并且视频中物体的运动方向存在少量的不一致. 这时, 其视觉屏蔽特性阈值的非线性计算如公式(8)所示:

$$JND(x, y) = JND_s(x, y) + JND_t(x, y) - \alpha \times \min\{JND_s(x, y), JND_t(x, y)\}, 0 < \alpha \leq 1 \tag{8}$$

公式(8)表明, 该视觉阈值是由静态图像的视觉屏蔽阈值  $JND_s(x, y)$  与视频时间序列的人类视觉屏蔽阈值  $JND_t(x, y)$  两部分组成的. 非线性公式的引入不会产生静态图像与视频时间序列视觉阈值直接叠加. 通过对参数  $\alpha$  的控制, 可以保证在以该计算阈值作为水印的嵌入强度时, 不会引起视频视觉质量的下降.

如果  $H(|MV(u, v)|) < T_v(\min)$  或是  $H(\theta(u, v)) < T_\theta(\min)$ , 则表明该视频序列中运动速度缓慢, 或是物体的运动方向趋于一致. 对于这种情况, 我们将视频序列考虑为静态帧, 从而只考虑静态图像的人类视觉屏蔽特性, 其视觉屏蔽特性阈值的计算如公式(9)所示:

$$JND(x, y) = \beta \times JND_s(x, y), 0 < \beta \leq 1 \tag{9}$$

如果  $H(\theta(u, v)) \geq T_\theta(\max)$ , 则表明该视频序列中存在较多的运动方向不一致的物体. 对于这种情况, 我们更多地考虑视频时间序列的人类视觉特性. 因此, 其视觉屏蔽特性阈值的计算如公式(10)所示:

$$JND(x, y) = \gamma \times JND_t(x, y), 0 < \gamma \leq 1 \tag{10}$$

为了更好地控制和获取水印的最大嵌入强度, 我们用一个水印强度控制参数  $\delta$  对  $JND(x, y)$  进行强度控制, 即  $JND'(x, y) = \delta \times JND(x, y)$ .  $\delta$  的取值越大, 表明可以插入的水印容量越大. 为简化符号, 以下用  $JND(x, y)$  表示  $JND'(x, y)$ .

## 3 水印嵌入与提取算法

### 3.1 水印嵌入算法

在本文的水印嵌入算法中, 选用没有经过 MPEG 压缩的视频序列, 其颜色格式为 YCbCr. 可以根据需要选择需要嵌入水印信号的帧. 如果是对视频序列每帧进行水印嵌入, 则在提取水印时, 可以从任何一帧中提取. 在水印嵌入时, 选择待嵌入信号帧的亮度通道 Y, 采用上一节方法计算获得该帧中各个  $8 \times 8$  图像方块  $B(x, y)$  的视觉屏蔽阈值  $JND(x, y)$ , 我们以该视觉屏蔽阈值作为每个图像方块进行水印嵌入时的最大嵌入强度.

本文采用如下步骤在小波域内实现视频水印的嵌入:

(1) 水印信号生成. 设  $w_s = \{w_s(p) | 1 \leq p \leq m_w \times n_w, w_s(p) \in \{0,1\}\}$  为二值水印序列, 我们使用  $l_w$  标记水印长度, 其中,  $l_w = m_w \times n_w$ . 使用密钥  $key$  生成与水印长度  $l_w$  相一致的混沌序列对  $w_s$  进行置乱;

(2) 按行扫描顺序依次选择视频帧内亮度通道  $Y$  中两个相邻的  $8 \times 8$  图像方块  $B_1(x_1, y_1)$  和  $B_2(x_2, y_2)$  (其中,  $(x_1, y_1)$  和  $(x_2, y_2)$  分别为  $B_1$  和  $B_2$  左上角像素在该帧中的(行,列)坐标), 分别对这两个图像方块进行一层二维离散小波变换. 为简化符号, 以下用  $B_1$  和  $B_2$  分别表示  $B_1(x_1, y_1)$  和  $B_2(x_2, y_2)$ ;

(3) 利用两个相邻图像方块中相对应的 3 个小波细节子带系数( $HL, LH, HH$ )进行水印嵌入, 其中, 各个图像方块  $B_1$  和  $B_2$  的水印最大嵌入强度值分别是该图像方块的视觉屏蔽阈值  $JND(x_1, y_1)$  和  $JND(x_2, y_2)$ . 通过调整两个对应细节子带系数值与视觉屏蔽阈值  $JND(x_1, y_1)$  与  $JND(x_2, y_2)$  之间的均值大小关系, 自适应地嵌入水印信号;

(4) 3 个小波细节子带系数( $HL, LH, HH$ )的水印嵌入算法相同. 下面以  $HL(u, v)$  为例, 介绍水印嵌入算法.

$$\text{设 } HL_1 = \sum_{u_1=1}^4 \sum_{v_1=1}^4 HL_1(u_1, v_1), HL_2 = \sum_{u_2=1}^4 \sum_{v_2=1}^4 HL_2(u_2, v_2)$$

$$dif = HL_1 - HL_2$$

$$\text{if } w(p) == 1$$

$$\text{if } dif > \overline{JND}$$

$$\begin{cases} HL_1^*(u_1, v_1) = HL_1(u_1, v_1) \\ HL_2^*(u_2, v_2) = HL_2(u_2, v_2) \end{cases} \quad (11)$$

else

$$\begin{cases} HL_1^*(u_1, v_1) = HL_1(u_1, v_1) + \frac{HL_1(u_1, v_1)}{HL_1} \times \left( \frac{JND(x_1, y_1) - dif}{2} \right) \\ HL_2^*(u_2, v_2) = HL_2(u_2, v_2) - \frac{HL_2(u_2, v_2)}{HL_2} \times \left( \frac{JND(x_2, y_2) - dif}{2} \right) \end{cases} \quad (12)$$

end

$$\text{else } w(p) == 0$$

$$\text{if } (-dif) > \overline{JND}$$

$$\begin{cases} HL_2^*(u_2, v_2) = HL_2(u_2, v_2) \\ HL_1^*(u_1, v_1) = HL_1(u_1, v_1) \end{cases} \quad (13)$$

else

$$\begin{cases} HL_2^*(u_2, v_2) = HL_2(u_2, v_2) + \frac{HL_2(u_2, v_2)}{HL_2} \times \left( \frac{JND(x_2, y_2) - dif}{2} \right) \\ HL_1^*(u_1, v_1) = HL_1(u_1, v_1) - \frac{HL_1(u_1, v_1)}{HL_1} \times \left( \frac{JND(x_1, y_1) - dif}{2} \right) \end{cases} \quad (14)$$

end

end

设  $\overline{JND} = (JND(x_1, y_1) + JND(x_2, y_2)) / 2 > 0$  为  $B_1$  和  $B_2$  的视觉屏蔽阈值的平均值,  $HL_1(u_1, v_1)$ ,  $HL_2(u_2, v_2)$  和  $HL_1^*(u_1, v_1)$ ,  $HL_2^*(u_2, v_2)$  分别表示  $B_1$  和  $B_2$  图像方块水印嵌入前、后的小波细节子带系数.

在上面所述的水印嵌入算法中, 我们首先将两个相邻图像分块所对应的小波细节子带系数之和进行相减, 获得其差值  $dif$ . 然后分别在  $w(p)=1$  和  $w(p)=0$  两种情况下比较  $dif$  与  $\overline{JND}$  的值, 调整相应的小波系数.

当  $w(p)=1$  时, 如果  $dif > \overline{JND}$ , 则  $HL_1 > HL_2$ , 对  $B_1$  和  $B_2$  中的小波系数不作修改, 如公式(11)所示. 因为在这种情况下进行水印嵌入, 会引起图像分块的视觉质量下降, 水印的透明性会受到影响; 如果  $dif \leq \overline{JND}$ , 则利用公式(12)进行水印信号嵌入, 可以数学证明公式(15)成立:

$$\sum_{u_1=1}^4 \sum_{v_1=1}^4 HL_1^*(u_1, v_1) - \sum_{u_2=1}^4 \sum_{v_2=1}^4 HL_2^*(u_2, v_2) = \overline{JND} > 0 \quad (15)$$

公式(15)表明,两相邻图像方块之间的水印嵌入强度为相邻图像方块的视觉阈值的平均值(即水印最大嵌入强度的平均值  $\overline{JND}$ ),所以,当  $w(p)=1$  时,本文算法可以保证  $HL_1^* > HL_2^*$ .

同理,当  $w(p)=0$  时,本文算法可以保证  $HL_2^* > HL_1^*$ .

### 3.2 水印提取算法

本文所述的水印提取算法无需原始视频序列的参与,只需产生混沌序列的密钥  $Key$  和水印信号的长度  $l_w$ ,在对视频序列中每一帧嵌入水印时,可以从任何一帧开始进行水印信号的提取.本文算法步骤如下:

(1) 将已嵌入水印信号的视频序列分割为帧,选取每帧亮度通道  $Y'$  进行水印信号的提取检测;

(2) 按行扫描顺序依次选择视频帧内亮度通道  $Y'$  中两个相邻的  $8 \times 8$  图像方块  $B_1$  和  $B_2$ ,分别对这两个图像方块进行一层二维离散小波变换,利用两个相邻图像方块中相对应的 3 个小波细节子带系数( $HL^*$ ,  $LH^*$ ,  $HH^*$ )进行水印提取.由于 3 个小波细节子带系数( $HL^*$ ,  $LH^*$ ,  $HH^*$ )的水印提取算法相同,我们以  $HL^*(u,v)$  为例说明本文的水印提取算法.视频帧内相邻图像方块对 ( $B_1$  和  $B_2$ ) 的对数为  $num=(M_f \times N_f)/(8 \times 8 \times 2)$ ,因此该帧中该子带嵌入了  $num$  比特的水印信号,即嵌入了长度为  $l_w$  的水印信号  $N$  次,其中,  $N=\lfloor num/l_w \rfloor$ .通过公式(16)可以提取出  $N$  个水印信号版本  $W'(s,p)$ ;

$$w'(s,p) = \begin{cases} 1, & \text{if } \sum_{u_1=1}^4 \sum_{v_1=1}^4 HL_1^*(u_1, v_1) \geq \sum_{u_2=1}^4 \sum_{v_2=1}^4 HL_2^*(u_2, v_2) \\ 0, & \text{if } \sum_{u_2=1}^4 \sum_{v_2=1}^4 HL_2^*(u_2, v_2) > \sum_{u_1=1}^4 \sum_{v_1=1}^4 HL_1^*(u_1, v_1) \end{cases}, 1 \leq s \leq N, p = 1, 2, \dots, l_w \quad (16)$$

(3) 利用对水印信号进行按列统计求和的方法.通过公式(17),从  $N$  个水印信号版本中提取出需要解调的水印信号  $W''(p)$ ;

$$w''(p) = \begin{cases} 1, & \text{if } \sum_{s=1}^N w'(s,p) \geq \frac{N}{2}, p = 1, 2, \dots, l_w \\ 0, & \text{else} \end{cases} \quad (17)$$

(4) 利用密钥  $key$  生成与水印长度  $l_w$  相一致的混沌序列对  $W''(p)$  进行解调,得到恢复后的水印信号  $W^*(p)$ .

## 4 视频水印透明性

为评价本文算法的有效性,本节着重讨论该自适应的视频水印是否真正具有较高的透明性.本算法的基本实验步骤如下:(1) 从 ITE(image information and television engineers)标准视频序列中选取 6 个标准视频序列,每个视频序列的长度都是 150 帧,大小为  $720 \times 480$ ,颜色的格式为 YCbCr;(2) 选取视频序列每一帧中的亮度通道  $Y$  作为水印的嵌入载体.对每帧亮度通道的  $8 \times 8$  方块进行一层二维离散小波变换;(3) 选取小波变换后的中高频子带进行水印的嵌入.接受测评的视频序列共有 6 组,如图 4 所示.



Fig.4 Video sequence used in the experiments

图 4 进行水印嵌入的视频序列



#### 4.1 客观测评

为了评估水印的透明性,我们选择 PSNR 值作为客观评判的尺度.根据实验.本文第 2.2 节中的各个参数为: $\alpha=0.6, \beta=0.4, \gamma=0.3$ .在计算各个视频序列的 PSNR 时,先计算每个帧的 PSNR 值,然后再求平均得到整个视频序列的 PSNR,表 1 中记录了各个视频序列的平均 PSNR 值(单位:dB).

**Table 1** PSNR of the videos  
表 1 各个视频序列的 PSNR 值

Video	$\delta$			
	0.3	0.5	0.7	1.0
Whaleshow	47.54	46.982	45.33	44.36
Europeanmarket	46.10	45.17	44.25	42.917
Horsrace	52.352	52.086	51.805	51.37
Picturecut	44.46	44.16	43.22	41.302
Church	45.93	45.37	44.80	43.86
Brustleave	44.98	43.68	43.16	42.010

从表 1 可以看出,多数用于测试的视频序列帧的 PSNR 值均在 45dB 左右.而在文献[3]中,PSNR 值在 45dB 左右时, $\delta=0.02$ (文献[3]使用 $\alpha$ 来控制水印的嵌入强度,与本文中的 $\delta$ 意义相同).因此与文献[3]中的算法相比,本文算法能够嵌入更多量的水印信号,并能保持水印的透明性.由于 PSNR 只能对嵌入水印和视频与未嵌入水印的视频进行理论上的比较,而不能真实表现嵌入水印后的视频是否具有基于时间上的视觉失真、是否具有较好的透明性,因此,我们还对嵌入水印后的视频序列进行了主观测评.

#### 4.2 主观测评

我们按照 ITU-R BT.500-10.所要求的方法对嵌入水印后的视频序列进行主观透明性测评,在这次测评中,共有 10 人接受测评,他们都是很熟悉数字图像处理和数字水印的研究者.在进行主观测评时,接受测评的人员对所看到的每组视频序列进行主观评价,并根据视频的可视质量等级进行打分.5 分制的图像主观质量等级打分标准见表 2.

**Table 2** Standard of the subjective evaluation

表 2 图像主观质量等级打分标准

Score	Quality	Feeling
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slight annoying
2	Poor	Annoying
1	Bad	Very annoying

最终的测评结果如图 5 所示.

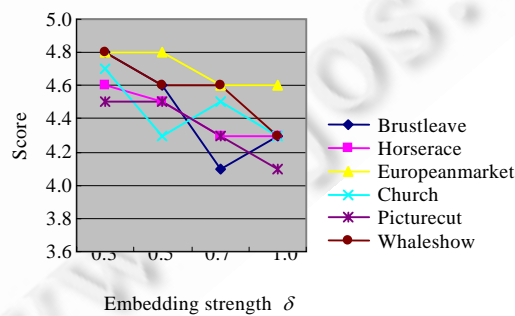


Fig.5 Result of subjective evaluation

图 5 主观评测结果

根据主观测试的结果可以发现,当嵌入强度控制参数 $\delta$ 在 0.3~1.0 区间内时,所有的测试视频均表现出较好的透明性,其主观测试的得分均在 4.5 分左右,基本上感觉不到水印的存在.因此,本文所提出的算法除了具有较好的客观评测结果外,在主观评测实验中也显示出较好的评测结果.

## 5 鲁棒性测试

在测试鲁棒性的实验中,我们选取了帧长为 150 帧大小为 720×480 的视频序列 *brustleave.yuv* 作为鲁棒性测试的视频.在进行水印嵌入时,使用本文第 2.2 节与第 3.1 节中所述的算法,算法中各个参数的值分别为  $\alpha=0.6$ ,  $\beta=0.4$ ,  $\gamma=0.3$ .水印信号在每个视频帧中被重复嵌入 135 次,嵌入总量为 8 100 bit.对嵌入水印后的视频序列分别进行帧裁剪、添加噪声、帧平均、帧交换和 MPEG 压缩等攻击,然后使用本文第 3.2 节中所述的水印提取算法,分别提取视频序列中前 50 帧的水印信号.对结果进行统计得到最后的一维二值水印信号  $W^*(p)$ .最后,将提取的水印信号  $W^*(p)$  与原始的水印信号  $W(p)$  进行相似度(SIM)的计算.其中, $\oplus$ 表示“异或”.计算公式如下:

$$SIM = \frac{\sum_p W^*(p) \oplus W(p)}{\sum_p W(p) \oplus W(p)} \quad (18)$$

### 5.1 帧平均、帧交换

帧平均是视频水印的一个特殊的攻击方式,它利用多帧的平均值破坏视频序列中的动态成分.实验中,定义帧平均模式  $Y_k(x,y)=1/3 \times (Y_{k-1}(x,y)+Y_k(x,y)+Y_{k+1}(x,y))$ ,  $k=2,3,\dots,49$ ,  $k$  代表视频序列帧的数目.通过对进行帧平均攻击后的测试视频序列进行水印的提取,实验结果表明,无论参与水印提取的帧数取值多少,其检测出来的水印 SIM 值均为 1.

帧交换是将视频的一些帧相互交换,并且帧交换不会显著影响视频的视觉效果.实验中建立的交换模式  $Y_k(x,y)$  与  $Y_{k-1}(x,y)$  互换,  $k=2,3,\dots,50$ .通过对进行帧交换攻击后的测试视频序列进行水印的提取,实验结果表明,无论参与水印提取的帧数取值多少,其检测出来的水印 SIM 值均为 1.

### 5.2 MPEG压缩

MPEG 压缩是对视频序列的最基本的攻击方式,并且是一种无意攻击方式.在 MPEG-1, MPEG-2 的压缩处理过程中,对视频的压缩率均为 200kb/s.然后,对进行攻击后的测试视频序列进行水印的提取,将检测帧数从 1 帧依次增加为 50 帧.通过记录相应的帧所检测出的水印 SIM 值可以发现,根据帧数的不同,其检测出来的水印 SIM 值也有相应的变化,其中,在执行 MPEG-1 压缩时,其检测出的水印最大 SIM=0.542,最小 SIM=0.410;执行 MPEG-2 压缩时,其检测出的水印最大 SIM=0.487,最小 SIM=0.403.文献[3]中所采用的压缩率均为 4Mbit/s,在这样的压缩率下检测出来的水印最大 SIM 为 0.30,最小 SIM 为 0.05.

### 5.3 帧裁剪

视频帧裁剪是将视频序列中的每一帧都进行了裁剪攻击,用每一帧的裁剪比例来衡量视频的破坏程度.在一定的裁剪比例下,从攻击后的视频序列中重构水印图像.通过实验可以发现,裁剪每帧 12%~26%的信息时,可以完全无误地提取出水印信号,即 SIM=1;当我们剪裁掉视频序列中每帧 46%的信息时,提取出的水印信号 SIM=0.98.

### 5.4 噪声攻击

随机噪声攻击是在视频序列中加入一定量的噪声,从而破坏或除掉作为版权标志的水印信息.我们在视频序列中的每一帧中都加入了一定量的随机噪声,从被污染过的视频中重构水印图像.通过实验发现,在对视频序列添加 20%的椒盐噪声时,其 SIM=1;添加 30%的椒盐噪声时,其 SIM=0.97;添加 40%的椒盐噪声时,其 SIM=0.88.对视频序列添加方差  $\sigma^2$  为 0.1 的高斯噪声时,其 SIM=1;添加方差  $\sigma^2$  为 0.2 的高斯噪声时,其 SIM=0.97;添加方差  $\sigma^2$  为 0.3 的高斯噪声时,其 SIM=0.96.

本文提出的水印算法的鲁棒性实验结果与参考文献中的算法鲁棒性实验结果的比较见表 3。

**Table 3** Comparison of the experiment of robust  
**表 3** 鲁棒性实验对比表

	Averaging of frames	Exchanging of frames	MPEG-1	MPEG-2	Cropping of frames (% square)	Adding pepper and salt noise (% strength)	Adding Gauassing noise ( $\sigma^2$ )
Proposed method	1.0	1.0	0.54~0.41 (200kb/s)	0.49~0.40 (200kb/s)	1~0.98 (12%~46%)	1~0.88 (20%~40%)	1~0.96 (0.1~0.3)
In Ref.[2]	No mentioned	No mentioned	No mentioned	0.36 (2Mb/s)	No mentioned	No mentioned	No mentioned
In Ref.[3]	No mentioned	No mentioned	0.3~0.05 (4Mb/s)	No mentioned	No mentioned	No mentioned	No mentioned

## 6 总 结

本文提出一种具有较高透明性的自适应的视频水印算法,该算法首先研究视频序列独有的帧间和时间特性,挖掘基于视频序列速度、运动方向、变形尺度等方面的人眼屏蔽特性,然后利用熵模型对视频内容中基于视频时间序列的人眼屏蔽特性作进一步的判断和细化,并利用非线性公式将基于视频时间序列的视觉屏蔽阈值与基于静态图像的视觉屏蔽阈值相结合,最终获得基于视频内容的人眼视觉屏蔽特性阈值作为水印的最大嵌入强度.通过主观和客观的评测实验可以表明,本文算法具有较高的透明性.鲁棒性实验结果也表明,本文所述的算法在保持较高透明性的前提下,对于帧裁剪、帧平均、帧交换、添加噪声和 MPEG 压缩等方面的攻击都能有效地进行水印的提取.因此,本文算法具有较高的安全性和透明性.

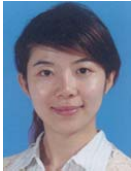
## References:

- [1] Liu HM, Huang JW, Xiao ZM. An adaptive video watermarking algorithm. In: Proc. of the IEEE Int'l Conf. on Multimedia and Expo. Tokyo: IEEE Computer Society, 2001. 190–193. <http://www.computer.org/portal/web/csdl/proceedings/i#4>
- [2] Echizen I, Yoshiura H, Fujii Y, Tezuka S. Use of motion estimation to improve video watermarking for MPEG encoders. In: Ton K, Cox IJ, eds. Proc. of the 2nd Int'l Workshop on Digital Watermarking (IWDW 2003). Seoul: Springer-Verlag, 2003. 184–199.
- [3] Agung IW, Sweeney P. Video scene characteristics detection to improve digital watermarking transparency. IEE Proc. of Vision, Image and Signal Processing, 2004,151(2):146–152.
- [4] Kim SW, Suthaharan S. An entropy masking model for multimedia content watermarking. In: Proc. of the 37th Hawaii Int'l Conf. on System Science. Hawaii: IEEE Computer Society, 2004. 1–6. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1265437>
- [5] Lou OJ, Wang XH, Wang ZX. Research on quantization-based robust video watermarking technique against geometrical attacks. Journal of Computer Research and Development, 2007,44(7):1212–1218 (in Chinese with English abstract).
- [6] LI Y, Gao XB, Ji HB. A 3D wavelet based spatial-temporal approach for video watermarking. In: Proc. of the Int'l Conf. on Computational Intelligence and Multimedia Applications. Xi'an: IEEE Computer Society, 2003. 260–265. <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1238135>
- [7] Zhu ZJ, Jiang GY, Yu M, Wu XW. A new adaptive algorithm for digital video watermarking. Journal of Zhejiang University (Science Edition), 2003,30(1):41–45 (in Chinese with English abstract).
- [8] Yang LS, Guo ZM. An adaptive video watermarking scheme based on the middle-frequency energy relationship between adjacent frames. Journal of Software, 2007,18(11):2863–2870 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/18/2863.htm> [doi: 10.1360/jos182863]
- [9] Gao Q, Li RH, Liu LS. Blind video watermarking algorithm based on correlations between neighboring frames. Journal on Communications, 2006,27(6):43–48 (in Chinese with English abstract).
- [10] Li Z, Yamamoto S, Takashima Y, Chen XW. Based on motion characteristics to calculate the Adaptive embedding tolerance for imperceptible video watermarking. In: Wang GP, ed. Proc. of the 2007 10th IEEE Int'l Conf. on Computer-Aided Design and Computer Graphics. Beijing: IEEE Computer Society, 2007. 192–198.

- [11] Podilchuk CI, Zeng WJ. Image-Adaptive watermarking using visual models. IEEE Journal Selected Areas in Communications, 1998,16(4):525-539. [doi: 10.1109/49.668975]
- [12] Li Z, Chen XW. Content-Based adaptive robust public watermarking algorithm. Journal of Computer Application, 2005,25(9): 2148-2150 (in Chinese with English abstract).
- [13] Nakamura T, Ogawa H, Tomioka A, Takashima Y. Improved digital watermark robustness against translation and/or cropping of an image area. IEICE Trans. on Fundamentals, 2000,E83-A(1):68-76.
- [14] Yang HF, Chen XW. A robust image-adaptive public watermarking technique in wavelet domain. Journal of Software, 2003,14(9): 1652-1660 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/14/1652.htm>

#### 附中文参考文献:

- [5] 楼偶俊,王相海,王钰旋.抗几何攻击的量化鲁棒视频水印技术研究.计算机研究与发展,2007,44(7):1212-1218.
- [7] 朱仲杰,蒋刚毅,郁梅,吴训威.一种视频图像的自适应水印新算法.浙江大学学报, 2003,30(1):41-45.
- [8] 杨列森,郭宗明.基于帧间频能量关系的自适应视频水印算法.软件学报,2007,18(11):2864-2870. <http://www.jos.org.cn/1000-9825/18/2863.htm> [doi: 10.1360/jos182863]
- [9] 高琦,李人厚,刘连山.基于帧间相关性的盲视频数字水印算法.通信学报,2006,27(6):43-48.
- [12] 李智,陈孝威.基于内容自适应小波域公开数字水印算法.计算机应用,2005,25(9):2148-2150.
- [14] 杨恒伏,陈孝威.小波域鲁棒自适应公开水印技术.软件学报,2003,14(9):1652-1660. <http://www.jos.org.cn/1000-9825/14/1652.htm>



李智(1977—),女,贵州贵阳人,博士,副教授,主要研究领域为多媒体技术,数字水印,虚拟现实.



陈孝威(1945—),男,教授,博士生导师,CCF高级会员,主要研究领域为多媒体技术,数字水印,虚拟现实.