

基于三维 Gabor 变换的视频水印算法*

张立和⁺, 伍宏涛, 胡昌利

(北京邮电大学 信息安全中心, 北京 100876)

A Video Watermarking Algorithm Based on 3D Gabor Transform

ZHANG Li-He⁺, WU Hong-Tao, HU Chang-Li

(Information Security Center, Beijing University of Posts and Telecommunication, Beijing 100876, China)

+ Corresponding author: E-mail: zhanglihe@yeah.net, <http://www.bupt.edu.cn>

Received 2003-04-01; Accepted 2003-05-27

Zhang LH, Wu HT, Hu CL. A video watermarking algorithm based on 3D Gabor transform. *Journal of Software*, 2004,15(8):1252~1258.

<http://www.jos.org.cn/1000-9825/15/1252.htm>

Abstract: Watermarking technique is a method by hiding copyright information into covering signals to discourage unauthorized copying. Because the profiles of two-dimension Gabor base functions are similar to those of human visual cortical cell receptive field and the middle frequency of visual channels has octave relationship, a video watermarking algorithm is proposed based on spatio-temporal multi-channel model using 3D Gabor transform in this paper. Experimental results indicate that the Gabor domain watermarks have greater robustness and imperceptibility.

Key words: Gabor transform; visual channel; pyramid expansion; digital watermark

摘要: 水印技术是把版权信息隐藏于载体信号中达到版权认证的一种版权保护技术.利用 Gabor 基函数波形类似人视觉皮层简单细胞的感受野波形的特性,结合视觉通道中心频率具有对数频程关系的特点,从视觉系统时空多通道模型角度出发,提出一种三维塔式 Gabor 变换视频水印算法.实验表明,该水印算法具有较好的鲁棒性和不可感知性.

关键词: Gabor 变换;视觉通道;塔式分解;数字水印

中图法分类号: TP309

文献标识码: A

计算机网络和多媒体技术的迅速发展,使得在数字产品版权保护方面具有巨大商业潜力的水印技术日益为学术界和商业界所关注.静态图像和运动图像序列都是用人眼来观察,因而针对图像或图像序列水印的不可感知性也是由人眼来衡量,充分利用人视觉系统的特性是解决视频水印不可感知性的重要手段之一.现在很多

* Supported by the National Natural Science Foundation of China under Grant Nos.60073049, 90204017 (国家自然科学基金); the National Grand Fundamental Research 973 Program of China under Grant No.G1999035804 (国家重点基础研究发展规划(973)); the National High-Tech Research and Development Plan of China under Grant No.2002AA143041 (国家高技术研究发展计划(863))

作者简介: 张立和(1976—),男,山西朔州人,博士,讲师,主要研究领域为信息安全,信息伪装,数字水印;伍宏涛(1973—),男,博士生,工程师,主要研究领域为密码学,信息伪装,数字水印;胡昌利(1977—),男,工程师,主要研究领域为信息安全,信息伪装,数字水印.

已发表的关于视频水印的文章虽然也考虑了视觉特性,但大多数都是把视频看作独立二维图像序列处理,水印嵌入到人眼不敏感的图像帧最低比特位、中高频分量或者是用视觉敏感函数来调制水印的强度等.本文则根据心理物理学的视觉模型把水印信息隐藏在视频中,进一步提高了水印的性能.

人类视觉的多通道和多分辨率特征首先由 Campbell 和 Robson 提出^[1],并得到了心理物理学实验的证实.基于 Gabor 滤波器是窄带带通滤波器,有明显的方向选择和频率选择特性,能在空域和频域同时达到最优的联合分辨率,尤其是因为二维 Gabor 基函数波形类似于绝大多数哺乳动物的视觉皮层简单细胞的感受野波形^[2]的特性,Daugman 于 1980 年首先提出视觉通道的 Gabor 滤波模型^[3].研究表明,动物视觉系统的几种不同机制都可以用不同形式的 Gabor 变换来表示,比如视觉的空间选择特性用与位置相关的 Gabor 采样表示,视觉皮层相邻单细胞感知域重叠特性用过采样的 Gabor 滤波表示^[4],视觉系统的频率响应的对数频程特性用塔式 Gabor 变换来表示等等.总之,Gabor 滤波感知模型的提出不仅为研究动物视觉的空间特性提供了一种有效的方式,而且也图像处理提供了一种重要的手段.正是由于 Gabor 变换的上述优点,它已经被迅速广泛地应用到图像压缩、纹理分析和图像分割等领域中^[5-7].

本文利用心理物理学研究成果:Gabor 基函数波形类似人视觉皮层简单细胞的感受野波形的特性,视觉通道中心频率对数频程特性和视觉方向选择特性,从时空多通道模型角度出发,提出一种基于三维塔式 Gabor 变换的视频水印算法.该算法模型把视频序列看作三维信号,区别于那些把视频序列当作独立的二维图像序列,然后分别用图像水印算法把水印隐藏在每帧图像的视频水印算法,而且该模型尽可能充分地体现了人类视觉的各种生理特性.实验表明,该水印算法具有较强的鲁棒性和较好的不可感知性.本文第 1 节讲述三维塔式 Gabor 变换,并与普通 Gabor 变换进行比较,阐明其频谱的对数分布关系,用以仿真视觉通道中心频率对数频程特性.第 2 节讲述视觉系统的时空相互作用.第 3 节具体讲述视频水印隐藏和提取算法以及 Gabor 变换系数的修改原则.第 4 节给出实验结果.最后给出本文的结论.

1 三维 Gabor 变换

如果 $s(k, l, t)$ 表示大小为 $L_1 \times L_2$ 的图像序列,那么该图像序列的 Gabor 基函数展开式为

$$s(k, l, t) = \sum_{m_1=0}^{M_1-1} \sum_{n_1=0}^{N_1-1} \sum_{m_2=0}^{M_2-1} \sum_{n_2=0}^{N_2-1} \sum_{m_3=0}^{M_3-1} \sum_{n_3=0}^{N_3-1} c_{m_1 n_1 m_2 n_2 m_3 n_3} \tilde{g}_{m_1 n_1 m_2 n_2 m_3 n_3}(k, l, t) \quad (1)$$

其中,

$$\begin{aligned} \tilde{g}_{m_1 n_1 m_2 n_2 m_3 n_3}(k, l, t) &= \tilde{g}(k - m_1 \Delta M_1, l - m_2 \Delta M_2, t - m_3 \Delta M_3) \exp \left[2\pi j \left(\frac{n_1 k}{N_1} + \frac{n_2 l}{N_2} + \frac{n_3 t}{N_3} \right) \right] \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} g(k - m_1 \Delta M_1 - m L_1, l - m_2 \Delta M_2 - n L_2, t - m_3 \Delta M_3 - i L_3) \exp \left[2\pi j \left(\frac{n_1 k}{N_1} + \frac{n_2 l}{N_2} + \frac{n_3 t}{N_3} \right) \right], \end{aligned}$$

$g(k, l, t)$ 为三维窗函数.时频采样间隔 $(\Delta M_i, \Delta N_i)$ 满足 $\Delta M_i M_i = \Delta N_i N_i = L_i, i = 1, 2, 3$.

Gabor 展开系数:

$$c_{m_1 n_1 m_2 n_2 m_3 n_3} = \sum_{k=0}^{L_1-1} \sum_{l=0}^{L_2-1} \sum_{t=0}^{L_3-1} \tilde{s}(k, l, t) \tilde{g}_{m_1 n_1 m_2 n_2 m_3 n_3}^*(k, l, t) \quad (2)$$

其中,

$$\begin{aligned} \tilde{g}_{m_1 n_1 m_2 n_2 m_3 n_3}(k, l, t) &= \tilde{\gamma}(k - m_1 \Delta M_1, l - m_2 \Delta M_2, t - m_3 \Delta M_3) \exp \left[2\pi j \left(\frac{n_1 k}{N_1} + \frac{n_2 l}{N_2} + \frac{n_3 t}{N_3} \right) \right] \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \gamma(k - m_1 \Delta M_1 - m L_1, l - m_2 \Delta M_2 - n L_2, t - m_3 \Delta M_3 - i L_3) \exp \left[2\pi j \left(\frac{n_1 k}{N_1} + \frac{n_2 l}{N_2} + \frac{n_3 t}{N_3} \right) \right], \end{aligned}$$

$\gamma(\cdot)$ 为 Bastiaans 解析法^[8]中定义的辅助函数.

如果窗函数 $g(k, l, t)$ 是时空可分离的,即有 $g(k, l, t) = g_1(k, l)g_2(t)$, 则根据 Bastiaans 解析法,辅助函数也是可分离的,

$$\gamma(k,l,t) = \gamma_1(k,l)\gamma_2(t),$$

且 $g_i(\cdot)$ 和 $\gamma_i(\cdot)(i=1,2)$ 满足双正交关系^[9].

为了阐明视觉系统的方向选择性,在时空分离的情况下用极坐标 (r, θ) 表示式(1)和式(2):

$$r = \sqrt{k^2 + l^2}, \quad \theta = \begin{cases} \tan^{-1}(l/k), & k \geq 0 \\ \tan^{-1}(l/k) + \text{sign}(l) \cdot \pi, & k < 0 \end{cases}$$

$$s(k,l,t) = \sum_{m_r, n_r, m_\theta, n_\theta, m_t, n_t} c_{m_r, n_r, m_\theta, n_\theta, m_t, n_t} g_1(r - m_r D_r, \theta - m_\theta D_\theta) g_2(t - m_t D_t) \exp(jr n_r W_r + j\theta n_\theta W_\theta + j t n_t W_t),$$

$$c_{m_r, n_r, m_\theta, n_\theta, m_t, n_t} = \sum_r \sum_\theta \sum_t s(r \cos \theta, r \sin \theta, t) \gamma_1^*(r - m_r D_r, \theta - m_\theta D_\theta) \gamma_2^*(t - m_t D_t) \exp(-j r n_r W_r - j \theta n_\theta W_\theta - j t n_t W_t).$$

由此可知, Gabor 基函数是由窗函数经过平移和调制生成的一系列函数. 如果时频采样网格 $(\Delta M_i, \Delta N_i)$ 在整个展开过程中保持不变, 那么这时 Gabor 基函数的功率谱在频域是等宽分布的. 视觉研究的结果表明, 人的视觉在不同频段范围内的感觉响应是不同的, 并且这些频段的中心频率具有对数频程的关系^[10]. 因此, 希望 Gabor 基函数的功率谱按照对数频程分布, 这种非均匀采样的 Gabor 基函数称为塔式 Gabor 基函数. 图 1 和图 2 分别清晰地显示出二维可分离塔式 Gabor 基函数空间频域在笛卡尔坐标系^[11]和对数极坐标系中塔式结构的划分^[2]. 如图 2 所示, 对数极坐标下更能体现出 Gabor 滤波的方向性. 研究表明, 人眼对不同方向 θ 上空间频率变换的敏感程度是不同的, 这样可以有选择地选取那些敏感程度低的系数作为水印嵌入点.

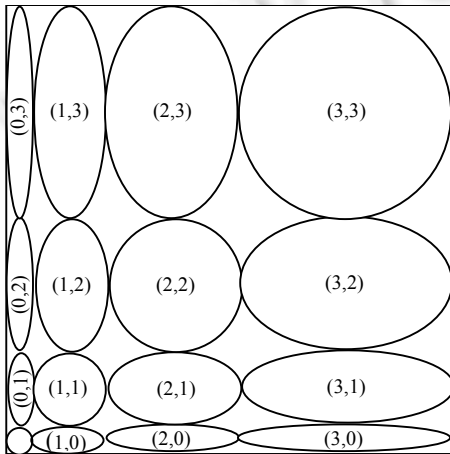


Fig.1 Cartesian spatial-frequency partition
图 1 笛卡尔坐标系下空间频域划分

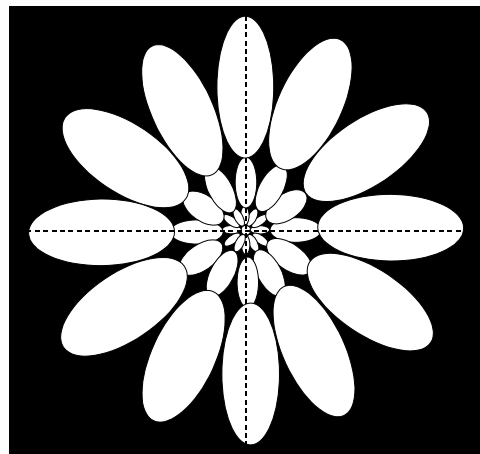


Fig.2 Log-Polar spatial-frequency partition
图 2 对数极坐标系下空间频域划分

为了形象地说明塔式基函数的频域分布情况,我们将 $(0,0)$ 阶基函数 $g_{00}(\cdot)$ 的频谱进行了平移,如图 3 所示. 与图 4 比较可看出,塔式基函数的频谱是按照对数频程分布的,均匀采样基函数的频谱是等宽的.

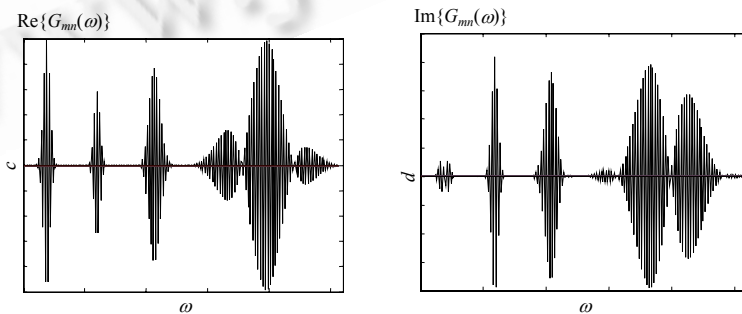


Fig.3 Pyramid Gabor base functions
图 3 塔式 Gabor 基函数

总之,塔式 Gabor 变换在频率域和空间域都表现出良好的局部化特性.通过选取适当的参数,就能够构造一

组具有不同调制频率和调制方向的 Gabor 滤波器模拟人类视觉系统分析视频、图像的特征.

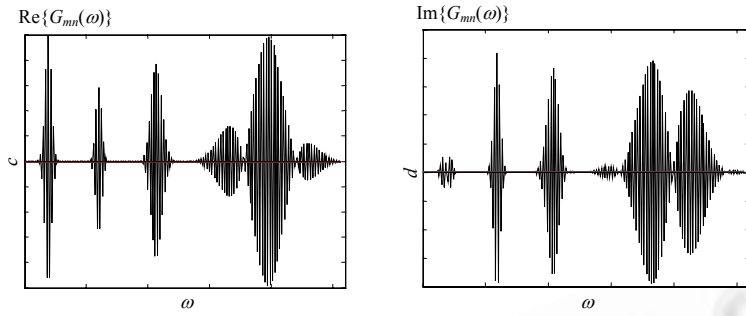


Fig.4 Uniform sampling Gabor base functions

图 4 均匀采样 Gabor 基函数

2 时空相互作用

如果用时空分离的 Gabor 基函数滤波器组仿真人类视觉系统,那么这个仿真系统在时间域和空间域的作用是独立的.事实上,人视觉系统的时间感知和空间感知作用是互相制约、不可分离的.因此,需要对前面提到的仿真系统进行修正.文献[12]的研究表明,用时空不可分的对比度敏感函数(CSF)作为滤波器的增益可以模拟视觉系统时空相互作用.如果把人类视觉系统作为在一定激励下产生一定输出的黑箱,这个黑箱就可以用一组滤波器来表示,而滤波器的系统函数就是用 CSF 加权的塔式 Gabor 基函数,每一个基函数对应于一个视觉通道,这就是视觉系统的多通道时空模型(如图 5 所示).该模型考虑了视觉的空间特性、时间特性及其相互作用、通道中心频率按对数频段分布特性和方向选择性.

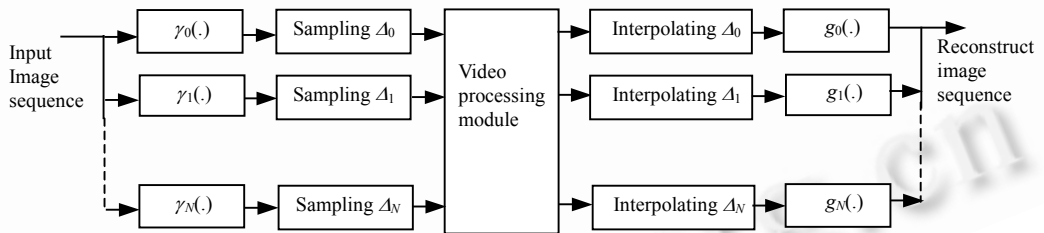


Fig.5 Visual multi-channel model

图 5 视觉多通道模型

图 5 前端的分析滤波器组 $\gamma_i(\cdot)$ ($i = 0, 1, \dots, N$) 为 Gabor 变换中的辅助函数,采样后的输出是 Gabor 变换系数,而后端的综合滤波器 g_i ($i = 0, 1, \dots, N$) 就是相应的 Gabor 基函数.

3 水印的隐藏和提取

用 $c(k_1, k_2, k_3)$ 表示视觉模型滤波器组的输出,即三维 Gabor 变换系数. k_1, k_2 和 k_3 分别代表空间 x 轴、空间 y 轴和时间轴.高时间频率的分量(k_3 值比较大)表示视频中运动激烈的部分,低时间频率的分量(k_3 值比较小)表示视频中变化缓慢的部分.在 k_3 一定时, k_1 和 k_2 可以看作对静态图像进行二维 Gabor 变换的情形. k_1 和 k_2 小的分量表示每帧图像的低频区域, k_1 和 k_2 大的分量表示高频区域.MPEG2 视频压缩算法采用运动预测编码压缩时间冗余.简单地说,就是把相邻几帧图像中相同的部分只保留一份传输,不相同的部分进行差分预测,只传输误差,这样就在时间方向上大大压缩了数据.对于空间上的冗余,MPEG2 标准采用类似于 JPEG 的压缩算法,先对图像进行 DCT 变换,将系数量化后再熵编码,这样做的目的是保留占有整个图像能量绝大部分的低中频系数,去掉高频系数,从而减少数据量.由此可以看出,视频中变换比较缓慢的地方以及图像高频部分在压缩时要丢失一些信息,为了使水印能够抵抗视频压缩,水印嵌入点选择为运动激烈(k_3 值大)且图像低中频(k_1 和 k_2 值大)的分量.考虑

到视觉系统的方向选择性, $c(k_1, k_2, k_3)$ 必须位于敏感度低的方向 θ 上.

设系数 c_i 是根据上述条件选择的水印嵌入点, Q 为预先确定的常数, 水印信息(二进制序列)经伪随机码扩频调制生成 w_i . 如果 w_i 与 $|c_i|/Q$ 取整后奇偶性相同, 那么 c_i 保持不变; 如果 w_i 与 $|c_i|/Q$ 取整后奇偶性相反, 那么 $|c_i| = |c_i| + Q$.

由于图像是实信号, Gabor 变换系数 $c_{m_1 n_1 m_2 n_2 m_3 n_3}$ 关于参数 (n_1, n_2, n_3) 具有共轭对称特性, 如果 $|c_{m_1 n_1 m_2 n_2 m_3 n_3}| = |c_{m_1 n_1 m_2 n_2 m_3 n_3}| + Q$, 那么必须使 $|c_{m_1 - n_1 m_2 - n_2 m_3 - n_3}| = |c_{m_1 - n_1 m_2 - n_2 m_3 - n_3}| + Q$, 这样经过逆变换后仍是实信号.

检测水印时根据 $|c_i|/Q$ 取整后的奇偶性来确定 w_i 是 0 还是 1. 若 $|c_i|/Q$ 取整后为奇数, 则 $w_i = 1$; 否则 $w_i = 0$.

对于扩频调制的水印算法, 扩频码同步是个关键问题. 检测和隐藏时的扩频码必须同步, 否则水印是检测不出来的. 人们提出很多解决同步问题的办法, 比如对水印信息编码加一个同步标志或者在隐藏水印的同时再隐藏一个模板, 这个模板可以记录载体经过的一系列处理, 从而保证检测时扩频码的同步^[13]. 本文提出的算法在隐藏和检测时都需要将视频码流解压, 这样, 只要解压参数(GOP、码率等)设置成和隐藏时视频的参数一样, 就可以达到同步的效果了. 也就是说, 在后面针对删帧和视频压缩鲁棒性的实验中, 在检测水印时解压视频时帧率和码率分别都设置成原来视频的帧率和码率.

4 实验结果

本文采用 Table Tennis, Mobile Calendar 和 Flower Garden 常用的标准 MPEG-2 视频序列进行实验. 这 3 段视频序列的参数见表 1.

Table 1 Video parameters
表 1 视频参数

Profile/Level	Main/Main
Video format	NTSC
Chroma format	4:2:0
Size	704×480
Bit rate	8M bits/s
I/P frame distance	3
Frames in GOP	15
Number of frames	180

图 6 是从这 3 段视频中取出其中一帧图像来说明水印的嵌入没有影响视频质量. 上面 3 幅图是未嵌入水印的原始图像帧, 下面 3 幅是相应的嵌入水印后的图像帧.



Fig.6 Visual effect comparison before and after embedding watermark

图 6 嵌入水印前后视觉效果比较

通过比较说明,利用 Gabor 变换的空间域和频率域的局部化特性、多方向性,再结合载体图像自身的特征,将水印嵌入到低中频和运动较激烈的区域,是不会引起人的觉察的.表 2 给出嵌入水印后视频帧平均峰值信噪比,从客观上来衡量视频的效果.

Table 2 Relation between Bitrate and PSNR
表 2 Bitrate 和 PSNR 关系

PSNR \ Bitrate	1.5	4.0	6.0	8.0	12.0	18.0	40.0
Video							
Table tennis	42.24	42.96	44.12	45.37	45.81	47.02	47.94
Flower garden	43.03	44.26	45.46	47.11	47.78	48.43	49.20
Mobile calendar	42.35	43.71	44.59	46.07	47.42	48.63	49.32

帧删除是专门针对视频水印的攻击方法.本文对嵌入水印后的视频序列每隔几帧删除一帧,然后检测水印,在丢帧率小于 30%的情况下,水印仍然能够检测出来(如图 7 所示).

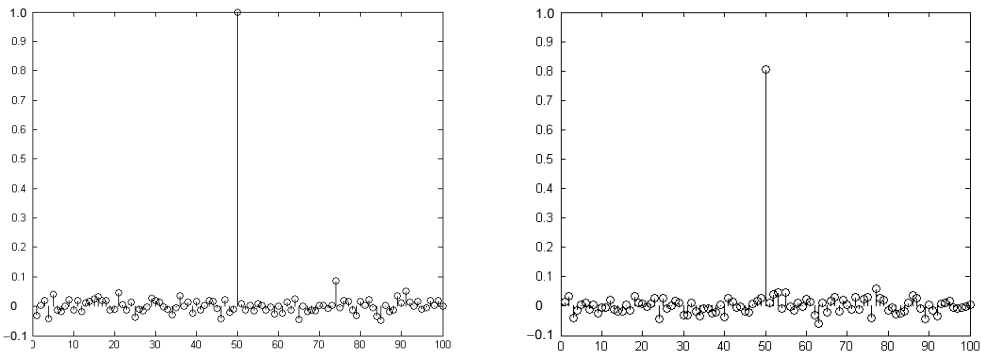


Fig.7 Watermark robustness to frame delete
图 7 水印对帧删除的鲁棒性

图 7 是将用 100 个不同伪随机序列提取的水印与原始水印相关而得出的实验曲线,横坐标为序列编号,纵坐标为相关值.其中第 50 个序列与嵌入水印时用的扩频调制序列相同.图 7 中的左图是对嵌入水印后的视频直接检测的结果,相关值等于 1;图 7 中的右图是经帧删除处理后检测的结果,可以看出,第 50 个序列的相关值仍远大于其他 99 个序列的检测相关值.

类似于 JPEG 压缩是针对静态图像水印的攻击方法,MPEG 压缩也是针对视频水印的攻击方法,比如将视频压缩码流解码后再编码,虽然视频效果未受多大影响,但是可能给水印带来毁灭性打击.实验表明,该算法对视频压缩具有一定的鲁棒性.图 8 给出了嵌入水印后视频经两种不同码率压缩后水印的相关检测结果.从图中可以明显看出,用第 50 个序列提取的水印信息就是当初嵌入的水印信息.

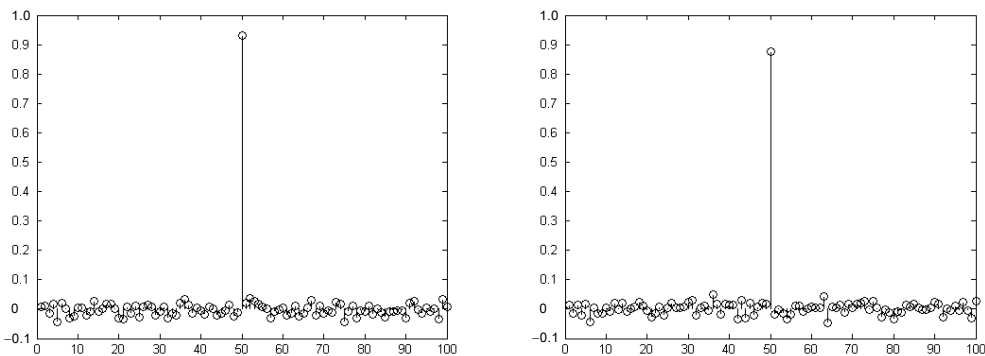


Fig.8 Watermark robustness to MPEG compression
图 8 水印对 MPEG 压缩的鲁棒性

5 结 论

本文把视频看作三维信号,隐藏水印在视频信号三维 Gabor 变换系数的幅度上,提出一种新的视频水印算法.这种水印基于 Gabor 变换的基函数波形和人视觉皮层简单细胞的感觉响应曲线相似的特性,能够很好地利用视觉系统的各种机制,表现出较好的不可感知性和鲁棒性.实验也表明,这种视频水印对帧删除和视频压缩具有良好的鲁棒性.该算法的缺点是计算复杂,运算量比较大,隐藏和提取水印时需要解码视频压缩码流.如果采用一些优化快速算法和高性能计算机是可以解决复杂性问题的.

致谢 在此,我们向对本文的工作给予支持和建议的同行,尤其是北京邮电大学信息安全中心杨义先教授、胡正名教授以及其他老师和同学表示感谢.

References:

- [1] Campell FW, Robson JG. Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 1968,197(2):551~556.
- [2] Daugman JG. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression. *IEEE Trans. on Acoustics, Speech and Signal Processing*, 1988,36(7):1169~1179.
- [3] Daugman JG. Two-Dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 1980,20(5):847~856.
- [4] Porat M, Zeevi YY. The generalized Gabor scheme of image representation in biological and machine vision. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1988,10(4):452~468.
- [5] Wang H, Yan H. Efficient implementation of Gabor transforms for image compression. *Electronics Letters*, 1992,28(9):870~871.
- [6] Bovik AC, Clark M, Geisler WS. Multichannel texture analysis using localized spatial filters. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1990,12(1):55~73.
- [7] Weldon TP, Higgins WE. Designing multiple Gabor filters for segmenting multi-textured images. *Optical Engineering*, 1999,38(9):1478~1489.
- [8] Bastiaans MJ. A sampling theorem for the complex spectrum, and Gabor's expansion of a signal in Gaussian elementary signals. *Optical Engineering*, 1981,20(4):594~597.
- [9] Wexler J, Raz S. Discrete Gabor expansion. *Signal processing*, 1990,21(3):207~221.
- [10] Daugman J. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by 2D visual cortical filters. *Journal of the Optical Society of America*, 1985,2(7):1160~1169.
- [11] Zhang XD, Bao Z. *Unbalanced Signal Analysis And Processing*. Beijing: National Defence Industry Press, 1998 (in Chinese).
- [12] van den Branden Lambrecht CJ. A working spatio-temporal model of the human visual system for image restoration and quality assessment applications. In: *IEEE, Signal Processing Society Staff. Acoustics, Speech, and Signal Processing. ICASSP'96. Conf. Proc. Atlanta: Georgia Institute of Technology, 1996. 2291~2294.*
- [13] Deguillaume F, Csurka G, Ruanaidh J, Pun T. Robust 3D DFT video watermarking. In: Wong PW, Delp EJ, eds. *IS&T/SPIE's 11th Annual Symp., Electronic imaging'99: Security and Watermarking of multimedia Contents. San Jose, 1999. 113~124.*

附中文参考文献:

- [11] 张贤达, 保铮. 非平稳信号分析与处理. 北京: 国防工业出版社, 1998.