一种基于特征空间分解的非对称鲁棒水印方法

谭秀湖[1],刘国枝[1],王雪松[2]
(1.哈尔滨工程大学水声工程学院,黑龙江哈尔滨 150001;2.哈尔滨工程大学宣传部,黑龙江哈尔滨 150001)

摘 要: 针对数字图像的版权保护,提出一种基于特征空间分解的非对称数字水印方法,即对数字图像的特征空间分解,将水印嵌入到最少受图像改变影响的子空间,使嵌入的水印具有鲁棒性.同时,因水印嵌入矩阵(密钥)与提取矩阵(公钥)不同,可公开除密钥外所有其它信息.通过对特征子空间选取限制,算法具有高的检测概率和安全性,低的虚警概率.仿真得到的结果,表明了算法具有好的鲁棒性.

关键词: 特征空间分解;非对称水印;安全性;鲁棒性
中图分类号: TN391 **文献标识码:** A **文章编号:** 0372-2112 (2006) 11-1981-05

An Asymmetric Robust Watermarking Method Based on Feature Space Decomposing

TAN Xiu-hu[1], LIU Guo-zhi[1], WANG Xue-song[2]
(1. *Underwater Acoustic Engineering College, Harbin Engineering University, Harbin, Heilongjiang* 150001, *China*;
2. *Department of Publication, Harbin Engineering University, Harbin, Heilongjiang* 150001, *China*)

Abstract: To the problem of digital picture's copyright protection, we propose an asymmetric watermarking method based on feature space decomposing. Passing feature space decomposing of digital image, we obtain that the robustness of the approach lies in hiding a watermark in the subspace that is least susceptible to potential modification. In the same time, because we use different matrix operation to embed (secret key) and extract (public key) a watermark, we are able to release all information for the public, except the secret key. Through analysis and constraint the conditions of feature space, the algorithm we proposed can obtain a high detection probability and security, a low false alarm probability. The robustness of the asymmetric watermarking method is demonstrated by a kind of attacks through computer simulation.

Key words: feature space decomposing; asymmetric watermarking; security; robustness

1 引言

数字水印作为一种新兴技术,已广泛地应用于授权认证,版权保护等方面.然而,数字水印技术大多采用对称方式,即水印检测密钥与嵌入密钥相同,这就导致严重安全问题.因为密钥还可用于移去水印.在水印检测时,需出示私人密钥,且密钥一旦暴露,攻击者就能伪造、修改和移去水印,从而背离了版权保护目的,且在所有权鉴别时会出现无法判别版权真正所有者的问题.根据 Kirckhoffs 准则[10],一个安全系统应让攻击者知道算法中所有细节,除了密钥.非对称水印方法可满足 Kirckhoffs 准则,是解决水印安全问题一种有效方法.非对称水印采用两套钥匙,即密钥和公钥.密钥用于嵌入,公钥用于检测.公钥不能移去水印,而密钥是不公开,所以攻击者将无法伪造、修改和移去水印,从而达到版权保护目的,且密钥可鉴别版权所有者.一个非对称水印算法,除非拥有密钥,应能保证系统安全性,几种非对称的经典算法[1,2]已经给出.

文中提出非对称水印方法是基于特征空间分解.文献[6]中采用特征空间分解,先在时域将数字水印 W_0 嵌入数字图像 I,生成含水印图像 I_w.对 I 和 I_w 分别进行离散余弦变换 DCT2[7](或离散小波变换 DWT[8]),得系数矩阵 A 和 A_w.模仿攻击者对 I_w 进行各种操作(如旋转,模糊和锐化等),得到一组图像.在 DCT2 域中计算这组图像数学期望 A_g.将 A_w 和 A_g 按照矩阵列顺序展开为列矢量 cs(A_w) 和 cs(A_g) 后,再对 cs(A_w) 和 cs(A_g) 作差,得特征向量 α(α 为随机变量).计算 α 协方差矩阵,对协方差矩阵进行奇异值分解(SVD),得到酉矩阵(协方差矩阵为对称矩阵,SVD 分解后两个酉矩阵相同)为特征空间 Z,将 Z 分为两个子空间,X 空间和正交补空间 Y.X(水印嵌入空间)为 Z 中较小的特征值对应的特征向量张成,因而 X 为最不受 I 变化影响空间,使得嵌入水印具有鲁棒性.而 Y 为 Z 较大特征值对应空间,受 I 变化影响明显.

收稿日期:2005-10-10;修回日期:2006-03-25

2 构造非对称算法所涉及的线性代数

2.1 行满秩和列满秩矩阵的酉矩阵的存在性

为构造非对称嵌入水印方法,我们提出并证明以下的推论.

推论 设矩阵 W_1,$W_1\in\mathbf{R}^{L\times g}$,($a$)当 $L\leqslant g$ 且 W_1 秩 rank(W_1) = L,即 W_1 为行满秩,则对任取 $L\times L$ 酉矩阵 U,总存在 $g\times g$ 酉矩阵 V 使下式成立

$$UW_1 = W_1V^{\mathrm{T}} \tag{1}$$

(b)当 $L\geqslant g$ 且 W_1 秩 rank(W_1) = g,即 W_1 为列满秩,则对任取 $g\times g$ 酉矩阵 V,总存在 $L\times L$ 酉矩阵 U 使下式成立

$$W_1V = U^{\mathrm{T}}W_1 \tag{2}$$

且只当 $L-g=\pm1$ 或 $L=g$ 时,式(1)中酉矩阵 V 和式(2)中酉矩阵 U 唯一存在.

证明:设 W_1 秩 rank(W_1) = L,对 W_1 进行 SVD 分解,得下式

$$W = U_wS_wV_w^{\mathrm{T}} \tag{3}$$

$U_w\in\mathbf{R}^{L\times L}$,$V_w\in\mathbf{R}^{g\times g}$,都为酉矩阵,对角矩阵 $S_w\in\mathbf{R}^{L\times g}$,因 rank($W_1$) = L,所以 S_w 主对角线为大于零正数,即行满秩.任取酉矩阵 $U_{w1}\in\mathbf{R}^{L\times L}$,设 $U_{w1}=(a_1,a_2,\cdots,a_l)$,$a_i$($i=1,\cdots,L$)为 $L\times1$ 列向量,对于 $V_{w1}^{\mathrm{T}}=(b_1,b_2,\cdots,b_L,\cdots,b_g)$,其中 b_i($i=1,\cdots,g$)为 $g\times1$ 列向量.设 V_{w1}^{T} 列向量满足下式

$$\begin{cases} b_i=(a_i^{\mathrm{T}},0\cdots,0)_{g\times1}^{\mathrm{T}}, & \text{如果 } 1\leqslant i\leqslant L \\ b_j^{\mathrm{T}}b_i=0(j\neq i),\ b_j^{\mathrm{T}}b_j=1, & \text{如果 } L\leqslant j\leqslant g,1\leqslant i\leqslant g \end{cases} \tag{4}$$

则下式成立

$$U_{w1}S_w = S_wV_{w1} \tag{5}$$

再进行简单矩阵变换,式(5)成立等价于任取酉矩阵 U,则存在酉矩阵 V^{T},使下式成立

$$UW_1 = W_1V^{\mathrm{T}} \tag{6}$$

同理,可以证明当 W_1 秩 rank(W_1) = g,有式(2)成立.(证毕)

2.2 矩阵行列展开定理的引用

设矩阵 $X_1=(x_{i,j})_{m\times n}$ 将 X_1 按照下式子展开,

$$\mathrm{cs}(X_1)=(x_{11}x_{21}\cdots x_{m1}x_{12}x_{22}\cdots x_{m2}\cdots x_{mn})^{\mathrm{T}} \tag{7}$$

为 X_1 列展开.任取矩阵 $B\in\mathbf{R}^{m\times m}$,$C\in\mathbf{R}^{n\times n}$,由文献[3],有下式成立

$$\mathrm{cs}(BX_1C)=(C^{\mathrm{T}}\otimes B)\mathrm{cs}(X_1) \tag{8}$$

$$\mathrm{cs}(BX_1)=(E_{n\times n}\otimes B)\mathrm{cs}(X_1) \tag{9}$$

$\otimes$ 为 Kronecker 积运算.

2.3 行满秩和列满秩矩阵的构造

设矩阵 W,$W\in\mathbf{R}^{L\times g}$,如果 $L\leqslant g$,则 W 秩 rank(W)$\leqslant L$,令 W_1 等于

$$W_1 = W + b\times E_{L\times g} \tag{10}$$

其中 b 为正常数,$E_{L\times g}$ 主对角线为 1,其它元素为 0 的 $L\times g$ 矩阵,调节 b,使 rank(W_1) = L.当 $L\geqslant g$,用同样方法使 rank(W_1) = g.

3 非对称水印的方法

原始图像 I,$I\in\mathbf{R}^{m\times n}$ DCT2 变换为 A,$A\in\mathbf{R}^{m\times n}$,$\mathrm{cs}(A)\in\mathbf{R}^{mn\times1}$.将 Z 分为两个子空间,空间 X 和正交补空间 Y.设矩阵 P,$P\in\mathbf{R}^{mn\times Lg}$($L\leqslant m$,$g\leqslant n$),为 X 空间一组单位正交基.P_1 构成密钥,即水印嵌入矩阵,P_2 和 $\mathrm{cs}(W_1)$ 为公钥,即水印检测矩阵.

3.1 水印的嵌入和提取

水印 W_0(设 $L\leqslant g$) DCT2 变换为 W,$W\in\mathbf{R}^{L\times g}$.$W$ 通过式(10)变换得 W_1,(满足 $UW_1V=W_1$),$W_1\in\mathbf{R}^{L\times g}$,将 W_1 通过下式嵌入到 X

$$\mathrm{cs}(A_w)=\mathrm{cs}(A)+P\times\mathrm{cs}(U_1W_1V) \tag{11}$$

$U_1\in\mathbf{R}^{L\times L}$ 为酉矩阵,且满足

$$U_2\times U_1 = U \tag{12}$$

根据式(8),则式(11)可变换为

$$\begin{aligned}\mathrm{cs}(A_w)&=\mathrm{cs}(A)+P\times(V^{\mathrm{T}}\otimes U_1)\mathrm{cs}(W_1)\\&=\mathrm{cs}(A)+P_1\times\mathrm{cs}(W_1)\end{aligned} \tag{13}$$

$$P_1 = P\times(V^{\mathrm{T}}\otimes U_1) \tag{14}$$

为使水印有好的鲁棒性,式(10)中 b 应选尽可能小;同时为获得高信噪比(SNR),应调节正数 k_1,使水印嵌入强度 $k_1\|\mathrm{cs}(W_1)\|$ 尽可能大,但不应使 $k_1\|\mathrm{cs}(W_1)\|$ 对图像视觉效果有明显影响.再通过 DCT2 重构,产生含水印图像 I_w.公钥($\mathrm{cs}(W_1)$,P_2)用于检测,P_2 为

$$P_2 = P\times(E_{g\times g}\otimes U_2)^{\mathrm{T}} \tag{15}$$

设获得图像特征为 A_w^*,可通过下式用公钥做水印检测

$$\begin{aligned}P_2^{\mathrm{T}}(\mathrm{cs}(A_w^*)-\mathrm{cs}(A))&=(E_{g\times g}\otimes U_2)\times P^{\mathrm{T}}(\mathrm{cs}(A_w^*)-\mathrm{cs}(A))\\&=(E_{g\times g}\otimes U_2)^{\mathrm{T}}\times P^{\mathrm{T}}\times P\times\mathrm{cs}(U_1W_1^*V_1)\\&=(E_{g\times g}\otimes U_2)\times E_{lg\times lg}\times\mathrm{cs}(U_1W_1^*V_1)\\&=(E_{g\times g}\otimes U_2)\times\mathrm{cs}(U_1W_1^*V_1)\\&=\mathrm{cs}(U_2U_1W_1^*V_1E_{g\times g})=\mathrm{cs}(W_1^*)\end{aligned} \tag{16}$$

再做相关判决.尽管攻击者能分析出 Z 划分,既 X 空间和 Y 空间,得到矩阵 P.但他无法从公钥 P_2 和 $\mathrm{cs}(W_1)$ 及 P 计算出 P_1,从而无法移除嵌入水印,因而保证嵌入水印安全性.

3.2 检测概率和虚警概率

检测函数使用相关判决.被检测图像 I_g^*,对 I_g^* 进行 DCT2 变换得 A_g^*,将 A_g^* 通过 P_2^{T} 投影,得

$$P_2^{\mathrm{T}}\times\mathrm{cs}(A_g^*-A_g)=c\times\mathrm{cs}(W_1)+n \tag{17}$$

c 为向量 I_g^*-I 通过 P_2^{T} 投影系数,$n\in\mathbf{R}^{Lg\times1}$ 为误差.设判决函数满足如下条件,

$$\begin{cases}\|n\|\ll\|\mathrm{cs}(W_1)\|,\\ \quad\text{如果 } \mathrm{cs}(A_g^*-A_g)\text{通过 } P_2 \text{ 投影后含有 } \mathrm{cs}(W_1)\\ |c|\to0,\\ \quad\text{如果 } \mathrm{cs}(A_g^*-A_g)\text{通过 } P_2 \text{ 投影后不含有 } \mathrm{cs}(W_1)\end{cases} \tag{18}$$

使用相关测试函数

$$\begin{aligned}&\mathrm{sim}(\mathrm{cs}(W_1),P_2^{\mathrm{T}}(\mathrm{cs}(A_g^*)-\mathrm{cs}(A)))\\&=\mathrm{sim}(\mathrm{cs}(W_1),P_2^{\mathrm{T}}(\mathrm{cs}(A_g^*-A_g)+\mathrm{cs}(A_g-A)))\\&=\frac{|\mathrm{cs}(W_1)^{\mathrm{T}}(P_2^{\mathrm{T}}(\mathrm{cs}(A_g^*-A_g)+\mathrm{cs}(A_g-A)))|}{\|\mathrm{cs}(W_1)\|\ \|P_2^{\mathrm{T}}\times(\mathrm{cs}(A_g^*-A_g)+\mathrm{cs}(A_g-A))\|}\end{aligned} \tag{19}$$

检测概率:如 I_g^*-I 通过 P_2^{T} 投影后含有 W_1,由式(18),得

$$\mathrm{sim}(\mathrm{cs}(W_1),P_2^{\mathrm{T}}(\mathrm{cs}(A_g^*-A_g)))$$

$$= \mathrm{sim}(\mathrm{cs}(\boldsymbol{W}_1), ((1+c)\mathrm{cs}(\boldsymbol{W}_1) + n))$$

$$\approx \frac{|1+c|\,|\mathrm{cs}(\boldsymbol{W}_1)^{\mathrm{T}}\mathrm{cs}(\boldsymbol{W}_1)|}{|1+c|\,\|\mathrm{cs}(\boldsymbol{W}_1)\|\,\|\mathrm{cs}(\boldsymbol{W}_1)\|} = 1 \tag{20}$$

虚警概率 如 $\boldsymbol{I}_g^* - \boldsymbol{I}$ 通过 $\boldsymbol{P}_2^{\mathrm{T}}$ 投影后不含有 $\boldsymbol{W}_1$，由式(18)，得

$$\mathrm{sim}(\mathrm{cs}(\boldsymbol{W}_1), \boldsymbol{P}_2(\mathrm{cs}(\boldsymbol{A}_g^* - \boldsymbol{A}_g))) = \mathrm{sim}(\mathrm{cs}(\boldsymbol{W}_1), (c \times \mathrm{cs}(\boldsymbol{W}_1) + n))$$

$$\approx \frac{|c|\,|\mathrm{cs}(\boldsymbol{W}_1)^{\mathrm{T}}\mathrm{cs}(\boldsymbol{W}_1)|}{|c|\,\|\mathrm{cs}(\boldsymbol{W}_1)\| + \|\mathrm{cs}(\boldsymbol{W}_1)\|\,\|n\|} \approx 0 \tag{21}$$

3.3 $\boldsymbol{U}_2$ 选取注意事项

为使攻击者很难破译算法，即从公钥推导或计算出密钥(实际上，除了公认 RSA 算法外，其它算法通过一些有效的方法，密钥都有被破译可能，对 $\boldsymbol{U}_2$ 加以限制，是增加攻击者破译密钥难度)．不使 $\boldsymbol{U}_2$ 满足下式

$$\boldsymbol{P}_2 = d \times \boldsymbol{P}_1 \tag{22}$$

d 为常数

$$\boldsymbol{P} \times (\boldsymbol{E}_{g\times g} \otimes \boldsymbol{U}_2)^{\mathrm{T}} = d \times \boldsymbol{P} \times (\boldsymbol{V}^{\mathrm{T}} \otimes \boldsymbol{U}_1) \tag{23}$$

将 $\boldsymbol{P}^{\mathrm{T}}$ 左乘式(23)，得

$$(\boldsymbol{E}_{g\times g} \otimes \boldsymbol{U}_2)^{\mathrm{T}} = d \times (\boldsymbol{V}^{\mathrm{T}} \otimes \boldsymbol{U}_1) \tag{24}$$

将 $\mathrm{cs}(\boldsymbol{E}_{L\times g})$ 右乘上式，再由式(9)，得

$$\boldsymbol{U}_2\boldsymbol{E}_{L\times g} = d \times \boldsymbol{U}_1\boldsymbol{E}_{L\times g}\boldsymbol{V} \tag{25}$$

4 安全分析

对信息隐藏和嵌入算法，安全性是至关重要，是性能评价重要组成部分．算法从 Projection 攻击，盲攻击(在仿真中给出)，拷贝攻击，Oracle 攻击和掩蔽攻击进行分析．

4.1 Projection 攻击

Projection 攻击者，伪造一个 $\boldsymbol{A}_g^a$，使 $\boldsymbol{A}_g^a$ 中不含有嵌入水印．其攻击方法是寻找 $\boldsymbol{A}_g^a$ 满足下式来实现

$$\min \|\mathrm{cs}(\boldsymbol{A}_g^a - \boldsymbol{A})\| < \|\mathrm{cs}(\boldsymbol{A}_w - \boldsymbol{A})\| \tag{26}$$

上式右侧等于计算下式

$$E((\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}} \times \boldsymbol{\alpha})(\boldsymbol{\alpha}^{\mathrm{T}} \times \mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})))$$

$$= \mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}E(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}})\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V}) \tag{27}$$

$$\boldsymbol{\alpha} = \mathrm{cs}(\boldsymbol{A}_g) - \mathrm{cs}(\boldsymbol{A}_w) \tag{28}$$

下面就根据先前对 Z 划分证明攻击者攻击是不可能实现．对 $\boldsymbol{\alpha}$ 协方差矩阵进行 SVD 分解，由式(27)得下式，

$$\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}E(\boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathrm{T}})\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V}) = \mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}\boldsymbol{U}_s\boldsymbol{\Sigma}\boldsymbol{U}_s^{\mathrm{T}}$$

$$\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V}) = \sum_{i=1}^{mn} \sigma_i^2(\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}u_i)^2 \tag{29}$$

其中 $u_i(i=1,\cdots,mn)$ 为对应特征向量．矩阵 $\boldsymbol{\Sigma}$ 特征值 $\sigma_1^2(i=1,\cdots,mn)$ 对应着 $\boldsymbol{\alpha}$ 中的各维的方差，代表对嵌入水印的图像改变而引起其特征空间中变化量的方差．由于已经按照特征值大小，将特征空间分化为 X 空间和 Y 空间，根据划分原则，有

$$\begin{cases} u_i \in \boldsymbol{X}, & \text{如果对应的 } \sigma_i^2 \leqslant \varepsilon \\ u_i \in \boldsymbol{Y}, & \text{如果对应对 } \sigma_i^2 \geqslant \varepsilon \end{cases} \tag{30}$$

正数 ε 为选取门限，$\boldsymbol{U}_s$ 为酉矩阵．已经选定 $\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V}) \in \boldsymbol{X}$，由酉矩阵正交性质可知，如果 $u_i \in \boldsymbol{Y}$，则与 $\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})$ 正交为零，因而，由式(29)可得

$$\sum_{i=1}^{mn} \sigma_i^2(\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}u_i)^2 = \sum_{i=1}^{mn} \sigma_1^2\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}u_iu_i^t\mathrm{cs}$$

$$(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V}) = (\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V}))\sum_{i=k+1}^{mn}\sigma_i^2 \tag{31}$$

其中 k 为常数，$mn-k$ 为 X 空间维数．因此由式(30)，选取的特征值已经为最小，所以 $\sum_{i=k+1}^{mn}\sigma_i^2$ 为最小，且 $\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})^{\mathrm{T}}\mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})$ 为常数．因而，通过特征空间划分和选取，获得不等式式(26)的右侧函数已经取最小值，因而攻击者无法找到使式(26)成立的 $\boldsymbol{A}_g^a$．所以，攻击者将无法通过 Projection 攻击方法．

4.2 拷贝攻击

攻击者发动拷贝攻击，方法是将一件合法含水印 x_1^w 及一件不含水印 x_2，先对 x_1^w 使用水印去除攻击得原始作品近似版本 $\tilde{x}_1$．通过下式，从含水印作品中减去所估计原始作品来估计所嵌入水印模式

$$\tilde{w}_a = x_1^w - \tilde{x}_1 \tag{32}$$

把通过估计得水印添加到不含水印作品 x_2

$$x_2^w = x_2 + \tilde{w}_a \tag{33}$$

对于本文中算法，假定攻击者已成功估 X 出空间，Y 空间和矩阵 $\boldsymbol{P}$．由于攻击者先估计 $\boldsymbol{I}$ 和相应 $\mathrm{cs}(\boldsymbol{A})$，因而拷贝攻击实际为估值问题，$\mathrm{cs}(\boldsymbol{A})$ 和 $\boldsymbol{P} \times \mathrm{cs}(\boldsymbol{U}_1\boldsymbol{W}_1\boldsymbol{V})$ 为随机变量．由于算法使用 DCT2 变换，而 DCT2 系数一般对应低频和中频，准确说并不服从高斯分布，通常服从 GGD 或 Cauchy 分布(这仅从能量角度定性分析拷贝攻击，不需要精确估计值，因而还用 Gaussian 模型．假定 $\mathrm{cs}(\boldsymbol{A}) \sim N(0, \boldsymbol{C}_A)$，$\boldsymbol{P}_1\mathrm{cs}(\boldsymbol{W}_1) \sim N(0, a^2\boldsymbol{E}_{mn\times mn})$，由文献[9]，可得 $\boldsymbol{P}_1\mathrm{cs}(\boldsymbol{W}_1)$ 估计值 $\hat{\boldsymbol{P}}_3$ 为

$$\hat{\boldsymbol{P}}_3 = a^2\boldsymbol{E}_{mn\times mn}(\boldsymbol{C}_A + a^2\boldsymbol{E}_{mn\times mn})^{-1}\mathrm{cs}(\boldsymbol{A}_w) \tag{34}$$

将 $\hat{\boldsymbol{P}}_3$ 嵌入到另外一幅图像 I_Y 在特征空间的特征中，通过下式实现水印拷贝

$$\mathrm{cs}(A_{YW}) = \mathrm{cs}(\boldsymbol{A}_Y) + \hat{\boldsymbol{P}}_3 \tag{35}$$

由于 $\|a^2\boldsymbol{E}_{mn\times mn}\|$ ($\|\cdot\|$ 为矩阵 Frobonius 范数，文献[5])和 $\|\boldsymbol{C}_A\|$ 分别表示嵌入水印和原始图像能量，而水印的能量远远小于原始图像能量，$\hat{\boldsymbol{P}}_3$ 协方差矩阵范数为

$$\begin{aligned} \|\mathrm{cov}(\hat{\boldsymbol{P}}_3, \hat{\boldsymbol{P}}_3)\| &= \|a^4E((\boldsymbol{C}_A + a^2\boldsymbol{E}_{mn\times mn})^{-1}\mathrm{cs}(\boldsymbol{A}_w)\mathrm{cs}(\boldsymbol{A}_w)^{\mathrm{T}} \\ &\quad ((\boldsymbol{C}_A + a^2\boldsymbol{E}_{mn\times mn})^{-1})^{\mathrm{T}})\| \\ &= \|a^4(\boldsymbol{C}_A + a^2\boldsymbol{E}_{mn\times mn})^{-1}\boldsymbol{C}_A \\ &\quad (\boldsymbol{C}_A + a^2\boldsymbol{E}_{mn\times mn}{}^{-1})\| \\ &\approx \frac{\|a^2\boldsymbol{E}_{mn\times mn}\|}{\|\boldsymbol{C}_A\|}\|a^2\boldsymbol{E}_{mn\times mn}\| \\ &\ll \|a^2\boldsymbol{E}_{mn\times mn}\| \end{aligned} \tag{36}$$

对 $\hat{\boldsymbol{P}}_3$ 进行检测判决，相当于式(18)中投影后不含有 $\boldsymbol{W}_1$ 的情况，即虚警事件，所以拷贝攻击成功概率将非常低．

4.3 Oracle 攻击和掩蔽攻击

进行 Oracle 攻击，攻击者需要知道水印嵌入和检测密钥，因此 Oracle 攻击，对于对称水印方式十分有效．而对于非对称方式，由于攻击者无法知道嵌入密钥，因而无法发起有效攻击．对于掩蔽攻击，通过控制水印嵌入的强度(部分 3.1)，来

避免攻击者通过掩蔽效应发动攻击.

5 算法鲁棒性能分析

设获得一个含噪声图像,其在特征空间的特征为 A'_w,则

$$\mathrm{cs}(A'_w) = \mathrm{cs}(A_w) + \mathrm{cs}(n_w) \tag{37}$$

用 P_2^{T} 对式(37)投影.得

$$P_2^{\mathrm{T}}\mathrm{cs}(A'_w) = P_2^{\mathrm{T}}\mathrm{cs}(A_w) + P_2^{\mathrm{T}}\mathrm{cs}(n_w) \tag{38}$$

将式(38)带入到式(19),得

$$\begin{aligned}&\mathrm{sim}(\mathrm{cs}(W_1), P_2^{\mathrm{T}}(\mathrm{cs}(A'_w - A)))\\ &= \frac{\mathrm{cs}(W_1)^{\mathrm{T}}((1+c)\mathrm{cs}(W_1) + n + P_2^{\mathrm{T}}\mathrm{cs}(n_w))}{\| \mathrm{cs}(W_1) \| \, \| (1+c)\mathrm{cs}(W_1) + n + P_2^{\mathrm{T}}\mathrm{cs}(n_w) \|}\\ &\approx \frac{\mathrm{cs}((1+c)W_1 + n)((1+c)\mathrm{cs}(W_1) + n + P_2^{\mathrm{T}}\mathrm{cs}(n_w))}{\| (1+c)\mathrm{cs}(W_1) + n \| \, \| (1+c)\mathrm{cs}(W_1) + n + P_2^{\mathrm{T}}\mathrm{cs}(n_w) \|}\\ &= \cos\theta\end{aligned} \tag{39}$$

其中 $\mathrm{cs}(n_w)$ 为图像中噪声特征,θ 为向量 $\mathrm{cs}((1+c)\mathrm{cs}(W_1) + n)$ 与 $P_2^{\mathrm{T}}\mathrm{cs}(n_w)$ 夹角,$\theta \in (0, \pi)$.根据三角函数知识,有下式成立

$$\theta \leqslant \frac{\| P_2^{\mathrm{T}}\mathrm{cs}(n_w) \|}{\| (1+c)\mathrm{cs}(W_1) + n \|} \tag{40}$$

因此只需使 $\dfrac{\| P_2^{\mathrm{T}}\mathrm{cs}(n_w) \|}{\| (1+c)\mathrm{cs}(W_1) + n \|}$ 值尽可能小,就能保证算法具有好的鲁棒性能表现.当噪声的向量次序发生改变时,将影响到式(40)中 θ 值,从而对算法的鲁棒性产生影响.但由于算法中,水印嵌入空间受图像改变带来的误差的影响已经为最小,噪声可以看为对图像的一种操作,因而,当噪声向量次序发生改变,对算法的鲁棒性能影响为有限.

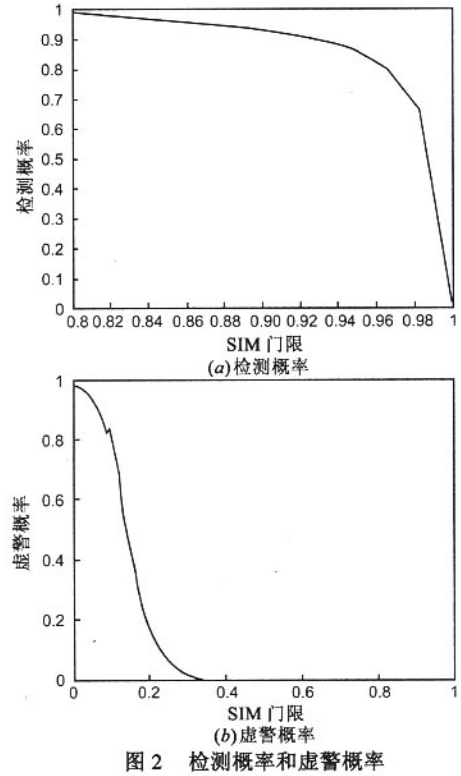
6 仿真试验

参照文献[6]中的实验方法,实验中选取的宿主图像的像素都为 256×256,包括人物,动物,如 Lena,猴子等.对每一幅嵌入水印的图像进行图像操作,如顺,逆时针旋转 $\pm 1°$, $\pm 2°$, $\pm 3°$,中值滤波,加噪等(均来自于 Matlab 工具箱和 Adobe Photoshop 和 Stirmark benchmark 4.0),共获得 129 幅改变后图像.对这 129 幅图像分别进行 DCT2 变换后,计算这 129 幅图像的数学期望.从这 129 幅图像中任取一幅,与求得的数学期望作差值后,取差值矩阵的左上角的 32×32 个数值(即垂直和水平方向都为 1 ~ 32),共包含 1024 个不同频率,将其进行列展开,因而我们选定的特征空间的维数为 1024 维.计算 1024 维特征空间协方差矩阵,并对其协方差矩阵进行 SVD 分解,得到的酉矩阵为 1024 维.在酉矩阵 1024 维空间中,选取对应特征值小的 900 维特征向量为 X 空间(X 空间不能太小,否则密钥容易被破译和攻击).选取 30×30 的熊猫 Panda 图像作为将嵌入的水印,首先对 Panda 进行 DCT2 变换,再经过行满秩和列满秩矩阵的构造后嵌入到空间,形成了含水印图像如图 1 所示.

(a)原始图像 (256×256) (b)嵌入水印后的图像 (256×256) (c)提取的水印 (30×30)

图 1 实验中使用的图像

对含和不含水印的图像进行各种图形操作后,作相关检测,得到了检测概率和虚警概率如图 2 所示.从图 2,可以看到算法有较高的检测概率和低的虚警概率.同时,又对算法进行了拷贝攻击测试,即从图 1(b)估计出水印后,嵌入到另外一幅图像当中.实验的结果是,当门限取值范围为 0.5 ~ 0.6(门限值已经很低)区间时,检测概率仅为 0.04 左右,相当于虚警事件水平,这就证明了我们以前的分析,算法可以有效地抵御拷贝攻击.我们对含水印的图像进行了盲攻击和鲁棒性能测试,得到的实验结果(表 1).实验结果已证明了我们的理论分析.由于篇幅的关系,这里仅列出实验的几个主要指标图表.



(a)检测概率

(b)虚警概率

图 2 检测概率和虚警概率

表 1 鲁棒性能测试和盲攻击测试

攻击类型	Sim 均值	Sim 方差
高斯加噪(均值 0,方差 0.01)	0.9377	0.0216
salt & pepper 噪声(噪声密度 0.02)	0.9953	0.0033
Speckle 噪声(均值 0,方差 0.04)	0.9279	0.0314
Possion 噪声	0.9862	0.0014
中值滤波(medfilt2)	0.9988	0.0012
Jpeg 压缩	0.9978	0.0010
旋转($\pm 2°$)	0.9886	0.0085
马赛克攻击	0.9614	0.0189
锐化	0.9970	0.0037
拼贴(10)	0.9645	0.0091
模糊	0.9981	0.0024

7 结论

通过理论分析和实验结果，都表明了基于特征空间分解的非对称鲁棒水印方法是一种可行的水印方法，有较高的安全性和好的鲁棒性．但本算法作水印提取时，需要原始图像，未能做到盲提取．因而，水印的盲提取是我们目前正在研究的课题．

致谢　感谢顾艳丽，朱海峰，聂东虎的无私帮助．

参考文献：

[1] Teddy Furon, Pierre Duhamel. An asymmetric watermarking method[J]. IEEE Trans, Signal Processing, 2003, 51(4): 981 – 995.

[2] J Tzeng, W Hwang, I Chern. An asymmetric subspace watermarking method for copyright protection[J]. IEEE Trans, Signal Processing, 2005, 53(2): 784 – 792.

[3] 史荣昌．矩阵分析[M]．北京：北京理工大学出版社，1996.

[4] 戈卢布 G·H，范洛恩 C·F，袁亚湘．矩阵计算[M]．北京：科学出版社，2001.

[5] 张贤达．矩阵分析[M]．北京：清华大学出版社，2004.

[6] Jengnan Tzeng, Wen Liang Hwang, I Liang Chern. Enhancing image watermarking methods with/without reference images by optimization on second order statistics[J]. IEEE Trans, Image Processing, 2002, 11(7): 771 – 783.

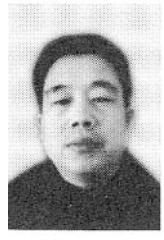
[7] IJ Cox, J Kilian, T Leighton. Secure spread spectrum watermarking for multimedia[J]. IEEE Trans, Image Processing, 1997, 6(12): 1673 – 1687.

[8] D Kundur, D Hatzinakos. Digital watermarking using multiresolution wavelet decomposition[A]. Proceeding of IEEE Conference on Acoustics, Speech and Signal processing[C]. Seattle Washington, USA: IEEE, 1998, 5. 2969 – 2972.

[9] S M Kay. Fundamentals of statistical signal processing: detection theory[M]. New Jersey: Prentice Hall, 1998.

[10] A Kerckhoffs. La cryptographie militaire[J]. Journal des Sciences Militaires, 1883, 9: 5 – 38.

作者简介：



谭秀湖　男，1971 年 8 月出生于吉林省辽源市，博士研究生，主要研究方向为信息隐藏和信息安全．E-mail：txhchinese@163.com

刘国枝　男，1944 年 12 月出生于河北省唐山市，于 1967 年毕业于原中国人民解放军军事工程学院海军工程系水声工程专业，哈尔滨工程大学研究员、博士生导师，获得国家科技进步奖一等奖一次，中国船舶工业总公司科技进步一等奖一项，中国船舶工业总公司科技进步二等奖三项，中国船舶工业总公司科技进步三等奖一项，黑龙江省高等学校实验室先进工作者，于 1995 年度获得国务院颁发的政府特殊津贴，1999 年被授予黑龙江省优秀科技工作者称号，2000 年评为全国优秀科技工作者．

E-mail：Liuguozhi@vip.sina.com