

## 多功能双水印算法

叶天语 钮心忻 杨义先

(北京邮电大学网络与交换技术国家重点实验室信息安全中心 北京 100876)

**摘要:** 该文针对单水印算法往往存在功能单一的问题, 利用奇异值的稳定性, 提出一种多功能双水印算法。先在图像分块的奇异值上嵌入鲁棒水印, 然后在含鲁棒水印图像的空域 LSB 嵌入脆弱水印, 并设计了判别恶意篡改和无意篡改的准则。实验不仅考察鲁棒水印抵抗攻击的鲁棒性, 而且还考察脆弱水印对鲁棒性的影响和篡改检测与定位的能力。实验结果表明: 鲁棒水印具备很强的抗攻击鲁棒性; 脆弱水印对篡改敏感, 而且篡改定位精确。因此算法具备版权保护和内容认证双重功能。

**关键词:** 奇异值分解; 双水印; 版权保护; 篡改检测; 篡改定位

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2009)03-0546-06

## A Multi-purpose Dual Watermark Algorithm

Ye Tian-yu Niu Xin-xin Yang Yi-xian

(Information Security Center, State Key Laboratory of Networking and Switching Technology,  
Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** The single watermark algorithms always merely have single function. In order to overcome the drawback, a multi-purpose dual watermark algorithm is proposed in the paper, taking advantage of the stability of singular values. It divides the original image into several blocks, and inserts robust watermark into their singular values. Then it inserts the fragile watermark into LSB in the space domain of previous watermarked image. Moreover, it designs a rule to distinguish malicious tamper from unintentional tamper. It tests not only robust watermark's robustness towards attacks, but also fragile watermark's influence towards robustness as well as the ability to tamper detection and localization. The experimental results show that the robust watermark possesses strong robustness to resist attacks, and that the fragile watermark is very sensitive to tamper and has good accuracy of tamper localization. Therefore it can achieve copyright protection and content authentication at the same time.

**Key words:** Singular Value Decomposition (SVD); Dual watermark; Copyright protection; Tamper detection; Tamper localization

### 1 引言

单水印算法往往存在功能单一的问题。鲁棒水印算法<sup>[1]</sup>拥有良好的鲁棒性, 具备抵抗攻击的能力, 可以实现版权保护功能, 但对篡改操作不够敏感, 篡改检测和定位能力差, 很难实现内容认证。脆弱水印算法<sup>[2]</sup>正好相反, 它对篡改操作非常敏感, 篡改检测和定位能力强, 容易实现内容认证, 但不具备抵抗攻击的能力。半脆弱水印算法<sup>[3]</sup>的性能介于两者之间, 但很难做到同时具备良好的鲁棒性和敏感的脆弱性。

在同一幅载体图像中嵌入鲁棒和脆弱双水印<sup>[4]</sup>, 存在水印嵌入顺序的问题。脆弱水印对图像的改动非常敏感, 容易遭到破坏, 而鲁棒水印具备较强的鲁棒性, 可以抵抗一定程度的外在干扰, 因此应该在载体图像先嵌入鲁棒水印后嵌入

脆弱水印, 这样既可以发挥鲁棒水印具备较强鲁棒性的优势, 又可以发挥脆弱水印对篡改敏感和容易实现篡改定位的优势。

文献[5]指出: 图像经奇异值分解后得到的奇异值能够表现出图像内在的代数特性而非视觉特性; 一幅图像的奇异值具有相当好的稳定性, 当图像受到轻微的扰动时, 它的奇异值不会发生剧烈的变化。如果把鲁棒水印嵌入在原始图像的奇异值, 脆弱水印嵌入在含鲁棒水印图像的空域, 那么只要嵌入脆弱水印的方法对含鲁棒水印图像构成的扰动足够微小, 就可以保证脆弱水印不会对算法的鲁棒性有很大的影响。另一方面, 脆弱水印的脆弱性又可以实现篡改检测和定位的目的。因此利用奇异值分解定理<sup>[5]</sup>和奇异值扰动定理<sup>[5]</sup>, 可以设计一个多功能双水印算法, 实现版权保护与内容认证双重功能。

### 2 鲁棒水印嵌入过程

(1)  $m \times n$  大小原始图像进行  $8 \times 8$  分块, 记为 **block1**。

2007-09-27 收到, 2008-04-03 改回

国家自然科学基金(90604022, 60473016), 北京市自然科学基金(4062025)和国家 973 计划项目(2007CB311203)资助课题

(2)对每个 $8 \times 8$ 分块进行 SVD, 得到  $\mathbf{block1} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ , 其中  $\mathbf{U}$  和  $\mathbf{V}$  分别为左右正交矩阵,  $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_8 \geq 0$  为对角阵  $\mathbf{\Sigma}$  的 8 个奇异值, 符号“T”表示转置。(3)对鲁棒水印  $\mathbf{rw}$  进行 Arnold 置乱, 置乱次数作为密钥 key1, 提取鲁棒水印时用到。  $\mathbf{rw}_j$  为第  $j$  比特鲁棒水印。(4)用密钥 key2 生成两个服从均匀分布且相关性很小的随机数序列。每个序列的长度为 7, 再将它们转化为  $\{-1,1\}$  二值序列, 分别记为 **sequence1**, **sequence2**。(5)用下列公式将鲁棒水印嵌入后 7 个奇异值。鲁棒水印嵌入公式如下: 如果  $\mathbf{rw}_j = 1$ , 则  $\sigma'_i = \sigma_i + \alpha \times \text{sequence1}(i-1)$ ; 如果  $\mathbf{rw}_j = 0$ , 则  $\sigma'_i = \sigma_i + \alpha \times \text{sequence2}(i-1)$ 。其中  $\alpha$  为嵌入强度因子,  $\sigma'_i$  为嵌入鲁棒水印后的奇异值,  $i = 2, 3, \dots, 8$ 。之所以不将水印嵌在第一个奇异值, 是为了防止方块效应发生。不同的嵌入鲁棒水印后的奇异值之间有可能改变最初的大小关系。在提取鲁棒水印时, 对含水印图像进行 SVD, 产生的奇异值是按照递减顺序排列的, 因此应该对上式得到的 7 个  $\sigma'_i$  进行递减排序, 并且用密码文件 key3 记录递减排序后原始随机数序列各分量的位置。密码文件在提取鲁棒水印时用到。(6)将  $\sigma'_i$  代回奇异值分解公式, 得到含鲁棒水印图像。

在上述鲁棒水印嵌入过程中, 总共涉及到 3 个密钥 (key1、key2 和 key3), 提取鲁棒水印时缺一不可。因此, 算法具备较好的安全性。而且 Arnold 置乱有利于算法抵抗剪切攻击。

### 3 脆弱水印产生及嵌入过程

(1)含鲁棒水印图像进行  $2 \times 2$  分块。每个分块记为  $\mathbf{block2} = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}$ ,  $x_i$  为像素值。(2)各分块像素的 LSB 置 0。此时得到的  $2 \times 2$  分块记为  $\mathbf{block2}' = \begin{bmatrix} x'_1 & x'_2 \\ x'_3 & x'_4 \end{bmatrix}$ ,  $x'_i$  是  $x_i$  的 LSB 置 0 的结果。(3)各分块进行 SVD, 计算奇异值的范数并取整。即  $\text{Norm1} = \sqrt{\sum_{i=1}^2 (\sigma_i)^2}$ ,  $\text{Norm1}' = \text{floor}(\text{Norm1})$ , Norm1 为分块 2 个奇异值的范数, Norm1' 是范数 Norm1 向  $-\infty$  取整数的结果。(4)通过 Norm1' 的 8 个位平面异或运算产生脆弱水印。过程如下:  $\mathbf{fw}_1 = b_1 \oplus b_2 \oplus b_3$ ;  $\mathbf{fw}_2 = b_2 \oplus b_3 \oplus b_4$ ;  $\mathbf{fw}_3 = b_4 \oplus b_5 \oplus b_6$ ;  $\mathbf{fw}_4 = b_6 \oplus b_7 \oplus b_8$ ;  $\mathbf{fw} = \begin{bmatrix} \mathbf{fw}_1 & \mathbf{fw}_2 \\ \mathbf{fw}_3 & \mathbf{fw}_4 \end{bmatrix}$ 。其中  $b_i$  是 Norm1' 的第  $i$  比特位,  $\mathbf{fw}$  是此分块产生的脆弱水印,  $\mathbf{fw}_i$  是第  $i$  比特脆弱水印。至此, 完成脆弱水印的产生过程。(5)脆弱水印嵌入到对应分块像素 LSB 上。此时得到的  $2 \times 2$  分块记为  $\mathbf{block2}'' = f(\mathbf{block2}', \mathbf{fw}) = \begin{bmatrix} x''_1 & x''_2 \\ x''_3 & x''_4 \end{bmatrix}$ 。  $f$  代表水印嵌入操作, 即将  $\mathbf{fw}_i$  嵌入  $x'_i$  的 LSB 得到  $x''_i$ 。(6)各  $\mathbf{block2}''$  分块重组含鲁棒和脆弱水印图像。

在脆弱水印产生及嵌入过程中, 利用每个分块的范数自适应产生脆弱水印, 因此脆弱水印序列与每个分块像素联系在一起。这有利于增强脆弱水印的篡改敏感性。

### 4 鲁棒水印提取过程

(1)遭攻击的含鲁棒水印和脆弱水印图像进行  $8 \times 8$  分块, 记为  $\mathbf{block3}$ 。(2)对每个  $8 \times 8$  分块进行 SVD, 得到  $\mathbf{block3} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ ,  $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_8 \geq 0$  为  $\mathbf{\Sigma}$  的 8 个奇异值。(3)将后 7 个奇异值按顺序组成一个序列, 记为 **sequence3**。(4)根据密钥 key2 产生嵌入水印时用到的随机数序列 **sequence1**, **sequence2**。(5)根据密码文件 key3 调整 **sequence1**、**sequence2** 中各个分量的顺序, 计算调整后的 **sequence1**、**sequence2** 与 **sequence3** 的相关性。采用相关检测提取鲁棒水印, 即: 计算  $c1 = \text{corr2}(\mathbf{sequence3}, \mathbf{sequence1})$  和  $c2 = \text{corr2}(\mathbf{sequence3}, \mathbf{sequence2})$ ; 如果  $c1 > c2$ , 则  $\mathbf{rw}'_j = 1$ , 反之  $\mathbf{rw}'_j = 0$ 。其中  $\mathbf{rw}'_j$  为提取的第  $j$  比特鲁棒水印,  $\text{corr2}$  计算两个序列相关系数。(6)根据密钥 key1 对提取的鲁棒水印进行 Arnold 逆置乱, 完成鲁棒水印的提取过程。

### 5 脆弱水印提取过程及篡改检测与定位

(1)遭篡改的含鲁棒水印和脆弱水印图像进行  $2 \times 2$  分块。一个  $2 \times 2$  分块记为  $\mathbf{block4} = \begin{bmatrix} y_1 & y_2 \\ y_3 & y_4 \end{bmatrix}$ 。(2)提取各分块的 LSB, 记为  $\mathbf{l} = \begin{bmatrix} l_1 & l_2 \\ l_3 & l_4 \end{bmatrix}$ , 其中  $l_i$  代表  $y_i$  的 LSB。(3)各分块像素的 LSB 置 0。此时得到的  $2 \times 2$  分块记为  $\mathbf{block4}' = \begin{bmatrix} y'_1 & y'_2 \\ y'_3 & y'_4 \end{bmatrix}$ 。(4)各分块进行 SVD, 计算奇异值的范数并取整。即  $\text{Norm2} = \sqrt{\sum_{i=1}^2 (\sigma_i)^2}$ ,  $\text{Norm2}' = \text{floor}(\text{Norm2})$ 。(5)通过 Norm2' 的 8 个位平面异或运算产生脆弱水印。过程如下:  $\mathbf{fw}'_1 = b'_1 \oplus b'_2 \oplus b'_3$ ;  $\mathbf{fw}'_2 = b'_2 \oplus b'_3 \oplus b'_4$ ;  $\mathbf{fw}'_3 = b'_4 \oplus b'_5 \oplus b'_6$ ;  $\mathbf{fw}'_4 = b'_6 \oplus b'_7 \oplus b'_8$ ;  $\mathbf{fw}' = \begin{bmatrix} \mathbf{fw}'_1 & \mathbf{fw}'_2 \\ \mathbf{fw}'_3 & \mathbf{fw}'_4 \end{bmatrix}$ 。其中  $b'_i$  代表 Norm2' 的第  $i$  比特位,  $\mathbf{fw}'$  代表此分块提取的脆弱水印,  $\mathbf{fw}'_i$  代表第  $i$  比特水印。至此, 完成了脆弱水印的提取过程。(6)将提取的各分块 LSB 与提取的脆弱水印进行比较。若完全一致, 即  $\mathbf{fw}'_i = l_i$  ( $i = 1, 2, 3, 4$ ), 则说明此分块没有发生篡改; 若有一比特不一致, 则代表此分块发生篡改。(7)对发生篡改的分块做标记来定位篡改的位置。

### 6 篡改检测虚警概率和漏警概率分析

虚警概率是指没有发生篡改而检测器却报告发生篡改的概率。当含双水印图像的某个分块没有发生篡改, 提取的脆弱水印将与提取的 LSB 完全相同。因此, 算法的虚警概率

$P_{FA} = 0$ 。

漏警概率是指发生篡改而检测器却报告没有发生篡改的概率。在一个被篡改的 $2 \times 2$ 分块,对于提取的脆弱水印

$$\mathbf{fw}' = \begin{bmatrix} \mathbf{fw}'_1 & \mathbf{fw}'_2 \\ \mathbf{fw}'_3 & \mathbf{fw}'_4 \end{bmatrix}$$

和提取的LSB  $\mathbf{l} = \begin{bmatrix} l_1 & l_2 \\ l_3 & l_4 \end{bmatrix}$ 来说,将 $\mathbf{fw}'_i$

和 $l_i$ 看成是二值随机变量,那么 $P(\mathbf{fw}'_i = 0) = P(\mathbf{fw}'_i = 1) = 1/2$ 和 $P(l_i = 0) = P(l_i = 1) = 1/2$ 成立。因此,由全概率公式可以知道:对于某个 $i$ , $P(\mathbf{fw}'_i) = P(l_i) = 1/2$ 。要判定没有遭到篡改,必须使得 $\mathbf{fw}'_i = l_i (i = 1,2,3,4)$ 成立。因此,一个被篡改的 $2 \times 2$ 分块被判定为没有遭到篡改的概率为 $2^{-4}$ 。假设遭到篡改的区域涉及到 $b$ 个 $2 \times 2$ 分块,那么 $b$ 个分块中有 $a (a \leq b)$ 个分块被判定为没有遭到篡改的概率服从二项分布 $P(a|b) = \binom{b}{a} \left(\frac{1}{16}\right)^a \left(1 - \frac{1}{16}\right)^{b-a}$ 。假设遭到篡改的区域涉及到 $b$ 个分块, $b$ 个分块的漏警概率为 $P_M = 2^{-4b}$ 。例如,当 $b = 5$ 时,漏警概率为 $9.5367 \times 10^{-7}$ ,非常小。对于恶意篡改来说,篡改的图像块太少是无法达到目的的,因此算法完全可以检测出恶意篡改。

### 7 实验结果

#### 7.1 不可见性

用 $256 \times 256$ 大小的256灰度级图像Lena(图1)作为原始载体图像。根据已知的理论,水印信息嵌入在图像空域的LSB,对图像视觉效果影响很细微。因此,双水印算法的不可见性主要取决于含鲁棒水印图像的不可见性。实验中嵌入强度因子 $\alpha$ 取为50,含双水印图像(图2)与原始图像之间的PSNR为36.5961dB,满足不可见性要求。



图1 原始载体图像Lena



图2 含双水印图像

#### 7.2 抵抗攻击的鲁棒性

把在脆弱水印嵌入过程产生的脆弱水印还原成 $256 \times 256$ 大小,见图3。根据算法设计的目的,鲁棒水印用于检验算法抵抗攻击的鲁棒性,以实现版权保护。因此,用原始鲁棒水印与提取的鲁棒水印之间的相关系数NC衡量算法的鲁棒性。图4是原始鲁棒水印Hand,图5是没有受到攻击时提取的鲁棒水印,两者之间的相关系数为0.9630。相关系数之所以没有达到1,是由于在含鲁棒水印图像中又嵌入了脆弱水印,对鲁棒水印的提取有细微的影响。

本文对双水印算法抵抗攻击的鲁棒性进行检验,相关系数见表1的“算法1相关系数”一栏。为节省篇幅,只将每

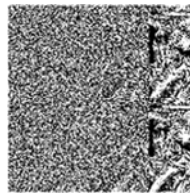


图3 脆弱水印



图4 原始鲁棒水印Hand



图5 提取的鲁棒水印

表1 鲁棒性检验

攻击类型	椒盐噪声			高斯噪声			中值滤波		
	强度			方差			窗口大小		
参数	0.03	0.04	0.05	0.05	0.1	0.2	[2,2]	[3,3]	[4,4]
算法1 相关系数	0.9291	0.9350	0.9381	0.9695	0.9794	0.9862	0.7675	0.7175	0.6581
算法2 相关系数	0.9320	0.9350	0.9411	0.9598	0.9695	0.9794	0.7597	0.7175	0.6552
攻击类型	低通滤波			JPEG 压缩					
	标准差			质量因子					
参数	0.3	0.4	0.5	90	80	70	60	50	40
算法1 相关系数	0.9598	0.9320	0.8250	0.9630	0.9504	0.9535	0.8763	0.8137	0.7210
算法2 相关系数	0.9663	0.9350	0.8273	0.9663	0.9567	0.9535	0.8815	0.8027	0.7175
攻击类型	剪切			直方图均衡	调整对比度		旋转		
	面积				范围		角度		
参数	左上角 1/16	左上角 1/8	左上角 1/4	64	[0.4,1]	[0.2, 0.8]	5	30	
算法1 相关系数	0.9338	0.8968	0.8591	0.9261	0.9535	0.9567	0.8447	0.7635	
算法2 相关系数	0.9371	0.9002	0.8626	0.9203	0.9535	0.9598	0.8456	0.7610	



图 6 算法 1 提取的鲁棒水印

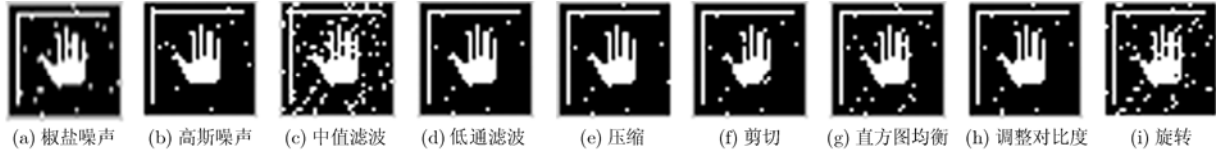


图 7 算法 2 提取的鲁棒水印

种攻击第 1 种情况提取的鲁棒水印列举出来，见图 6。根据鲁棒性要求设定一个阈值  $th$ ，当  $NC > th$ ，认为算法满足鲁棒性要求；阈值  $th$  越大，鲁棒性要求越高。算法设定  $th = 0.65$ 。由表 1 和图 6 可知，双水印算法对各种攻击都具有很强鲁棒性。

### 7.3 脆弱水印对鲁棒性的影响

本文的算法是在同一幅载体图像同时嵌入鲁棒水印和脆弱水印，因此不仅要考虑各种攻击对鲁棒性的影响，而且还要考虑嵌入脆弱水印对鲁棒性的影响。将本文的双水印算法记为算法 1，仅嵌入鲁棒水印的算法记为算法 2。算法 2 的抗攻击实验结果见表 1 “算法 2 相关系数”一栏和图 7。根据表 1 中的“算法 1 相关系数”一栏与“算法 2 相关系数”一栏的对比，或者根据图 10 与图 11 的对比，可以知道：(1) 后嵌入的脆弱水印对仅嵌入鲁棒水印的算法的鲁棒性有一定的影响，但双水印算法的抗攻击能力仍然还是很强；(2) 对于某些攻击而言，后嵌入的脆弱水印对仅嵌入鲁棒水印的算法的鲁棒性起增强作用，对于另外一些攻击则起削弱作用，但是不管起增强作用还是削弱作用，这些作用的影响都很微小。这正好验证了奇异值的稳定性，说明本文双水印算法“利用奇异值的稳定性，先在图像分块的奇异值嵌入鲁棒水印，然后在含鲁棒水印图像的空域嵌入脆弱水印，既发挥鲁棒水印具备较强鲁棒性的优势，又发挥脆弱水印对篡改敏感和容易实现篡改定位的优势”这一思路是合理的。

### 7.4 篡改检测与定位

**7.4.1 恶意篡改检测与定位实验** 恶意篡改一般指剪切-粘贴、叠加等操作。为了更直观地与含双水印图像进行对比，直接把篡改的区域定位在含双水印图像上，遭到篡改的区域用全黑色标识。

(a)从含双水印图像剪切一块，替换帽子上的一处(图 8)，其定位图像如图 9 所示。可见，此时定位出篡改的范围。(b)对含水印图像的嘴巴处每个像素加上 20(图 10)，其定位图像如图 11 所示。可见，此时定位非常精确。每个像素加上 20 虽然不改变篡改位置的 LSB，但改变了奇异值的范数，从而



图 8 篡改图像

图 9 定位图像



图 10 篡改图像图

图 11 定位图像

精武

图 12 jingwu

改变了由范数产生的脆弱水印，因此仍然可以定位出篡改的位置。(c)将  $32 \times 32$  二值图像 jingwu(由 0 和 1 组成的，“精武”两个字的像素值为 1，其他像素值为 0，见图 12)叠加到含水印图像左上角，篡改图像如图 13 所示。在图 13 中，二值图像 jingwu 像素值为 0 处没有发生篡改，像素值为 1 处发生了篡改。因为 lena 是 256 灰度级图像，所以将二值图像 jingwu 叠加到含双水印图像左上角后并不能分辨出“精武”两个字。定位图像见图 14。可以看出，即使只对含双水印图像做轻微的改动，但仍然可以定位出“精武”两个字。可见算法对篡改非常敏感，而且定位精确。(d)将  $64 \times 64$  灰度图像 kids(图 15)替换含双水印图像右边中间位置，篡改图像如图 16 所示。定位图像见图 17，可以定位出篡改的范围。(e)从原始载体图像的帽子上剪切一块替换含双水印图像的相同位置(图 18)，由于含水印图像与原始载体图像之间的视觉差异很小，所以并不能从图 18 中观察到篡改的区域。定位图像如图 19 所示，可以定位出篡改的范围。从以上篡改检测和定位实验可以看出：算法对篡改非常敏感，具有良好的篡改检测能力，可以定位到  $2 \times 2$  大小的分块。



## 8 结束语

本文利用奇异值的稳定性, 提出一种多功能双水印算法, 既发挥鲁棒水印具备较强鲁棒性的优势, 又发挥脆弱水印对篡改敏感和实现篡改定位的优势。先在图像分块的奇异值嵌入鲁棒水印, 然后在含鲁棒水印图像空域 LSB 嵌入脆弱水印, 并设计判别恶意篡改和无意篡改的准则。实验不仅考察鲁棒水印抵抗攻击的鲁棒性, 而且还考察脆弱水印对鲁棒性的影响和篡改检测与定位的能力。实验结果证明本文的算法具备版权保护和内容认证双重功能。

## 参 考 文 献

- [1] Lahouari G, Ahmed B, Mohammad K I, and Said B. Digital image watermarking using balanced multiwavelets. *IEEE Trans. on Signal Processing*, 2006, 54(4): 1519-1536.
- [2] Yuan H and Zhang X P. Multiscale fragile watermarking based on the gaussian mixture model. *IEEE Trans. on Image Processing*, 2006, 15(10): 3189-3200.
- [3] Schlauweg M, Pröfrock D, Palfner T, and Müller E. Quantization-based semi-fragile public-key watermarking for secure image authentication. In Proc. of SPIE, San Diego, California, USA, 2005, 5915: 41-51.
- [4] Sharkas M, ElShafie D, and Hamdy N. A dual digital-image watermarking technique. In Proc. of 3rd World Enformatika Conference (WEC'05), Istanbul, Turkey, 2005: 136-139.
- [5] 刘瑞祯, 谭铁牛. 基于奇异值分解的数字图像水印算法[J]. 电子学报, 2001, 29(2): 168-171.  
Liu R Z and Tan T N. SVD based digital watermarking method [J]. *Acta Electronica Sinica*, 2001, 29(2): 168-171.

叶天语: 男, 1982年生, 博士生, 研究方向为信息隐藏与数字水印.

钮心忻: 女, 1963年生, 教授, 博士生导师, 研究方向为信号与信息处理、信息隐藏、数字水印等.

杨义先: 男, 1961年生, 长江学者特聘教授, 博士生导师, 研究方向为密码学、计算机网络与信息安全等.