奇异值分解与PKI结合的鲁棒图像认证方法

侯启槟[1] 杨晓帆[2] 王海涛[1] 王阳生[1]

(中科院自动化所 北京100080)[1] (重庆大学计算机学院 重庆400044)[2]

摘 要 在图像的传输过程中,恶意篡改导致图像内容变化,而压缩等正常处理也会引起图像质量下降。图像认证是通过检测图像中是否有被恶意篡改的部分,来验证图像内容的完整性,同时容忍压缩或噪声对原图像质量造成的影响。本文提出奇异值分解与公共密钥体系(PKI)相结合的图像认证方法。它选用奇异值分解结果作为特征,构成原图像的内容摘要,私钥加密后形成原图像的认证码,并由用户自定义特征匹配的阈值。密钥的交换,依赖于PKI。需要验证图像数据的完整性时,再次分解提取特征,并用公钥解密认证码,然后进行匹配,达到图像认证目的。通过与几种典型图像认证方法做比较,表明本方法具有更好的鲁棒性。

关键词 图像认证,奇异值分解,PKI

Robust Image Authentication Method Combing SVD and PKI

HOU Qi-Bin[1] YANG Xiao-Fan[2] WANG Hai-Tao[1] WANG Yang-Sheng[1]

(Institution of Automation,CAS,Beijing 100080)[1]

(College of Computer Science,Chongqing University,Chongqing 400044)[2]

Abstract The goal of image authentication is to verify the integrity of an image by discovering malicious manipulations. Small modifications such as lossy compression and noises do not affect the results of image authentication, whereas modifications that do modify the visual content make the altered image unauthentic. This is the main requirement for image authentication. In ths paper, we propose a robust authentication method combing singular value decomposition (SVD) and public key infrastructure (PKI). This proposed method uses SVD as basic means to extract invariable, compression-tolerant features of image block, and PKI as the way to exchange keys. Experiment results show that our method can verify the authenticity and the integrity of picture. Furthermore, the proposed method has more robust ability distinguishing compression and Gaussian noise from deliberate tampering than several other methods, which is presented in the compare test results.

Keywords Image authentication, Singular value decomposition (SVD), PKI

1. 引言

面对多媒体数据易于复制和修改的特性,多媒体数据的购买者往往要求认证所得到数据的完整性。多媒体数据的完整性包括两个方面,一是所接受的图像或视频确实来自内容提供者;二是图像或视频的视觉内容没有被改变。图像或视频内容的认证,与传统密码学[1]中认证的概念不同之处是,图像质量的变化(图像文件的大小)并不影响认证结果,但篡改图像内容将直接导致认证的失败。若干年来,图像认证的研究已经形成两大类:数字水印[2]和数字签名[3]。数字水印侧重于数字图像的版权保护,它是在图像中嵌入标识或随机数。当需要证明版权时,从其中检测出所嵌入的标识或随机数。一般认为,所有权水印的校验是基于原始图像的认证[3]。

在某些情况下,需要进行基于目标图像的认证。譬如,图像的购买者试图了解他所得到的图像是否确实来自合法的出售者,并且其内容保持着完整性。出售者与购买者之间的数字签名机制能够帮助解决这个问题,执行过程如下:出售者对原始图像的内容处理,产生尺寸较小的内容摘要。随后,出售者对该内容摘要用自己的私钥加密,生成数字签名。该数字签名与对应的图像一起发送给购买者。购买者用相同的方法所接收图像的内容摘要,并用出售者的公钥对数字签名解密。如果新得到的摘要与解密的摘要可以匹配,所购得图像的完整性与所有权便得到认证[3]。在这个过程中,由于原始图像往往是经过压缩以后再传送给购买者,有损压缩导致图像质量下降。如果使用传统密码学的Hash函数分别处理原始图像和质量下降后的图像,由于它们的比特位之间存在差异,将产生不同消息摘要,导致合法图像不能通过认证。

研究者们已经提出若干种基于内容的图像认证方法[3~10]来保证有损压缩图像的内容摘要与原始图像的内容摘要相匹配,同时抵抗更改原图内容的非法操作,其中文[3~5]比较有代表性。Der-Chyuan Lou[3]提出容忍压缩的图像认证机制的几个度量指标,并给出基于图像块的灰度平均值方法。Lou方法计算量较少,但容忍压引起图像质量变化的能力较低。Ching-Yung Lin[4]针对JPEG图像压缩,提出基于DCT系数编码的鲁棒认证方法,然而Lin方法对基于零树小波压缩的图像,并不鲁棒。Liehua Xie[5]提出产生近似图像消息认证码(AMAC)的方法,实验表明,Xie方法工作性能不可靠。

本文提出基于奇异值分解和PKI相结合的鲁棒图像认证方法。本文的第2部分将简介奇异值分解及其特性;第3部分,描述我们的基于奇异值分解和PKI结合的图像认证方法;第4部分给出实验结果,并对实验结果加以对比、分析;最后对本文加以总结。

2. 奇异值分解 SVD

近年来，奇异值分解 SVD 在图像处理、模式分析和数据压缩等许多方面都获得广泛应用[11]，它表示了图像的代数特征，在某种程度上，SVD 所抽取图像的代数特征[12]，同时拥有代数与几何两方面的不变性。

有关奇异值分解的定理如下：

定理1 矩阵的奇异值分解式为

$A=P\begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}Q^H$，其中 $A\in C^{m\times n}$，P、Q 为酉矩阵

$D=diag(d_1,d_2,\cdots,d_r)$，且 $d_1\geqslant d_2\geqslant\cdots\geqslant d_r>0$。$d_i(i=1,2,\cdots,r)$ 是 A 的奇异值，它是 AA^H 或 A^HA 的特征值 λ 的平方根，即 $d_i=\sqrt{\lambda}$。

定理2（奇异值的稳定性） 假设 $A,B\in C^{m\times n}$，A、B 的奇异值分别为 $d_1\geqslant d_2\geqslant\cdots\geqslant d_r$，$\beta_1\geqslant\beta_2\geqslant\cdots\geqslant\beta_r$，则 $|d_i-\beta_i|\leqslant\|A-B\|_2$，即矩阵值有小的扰动，奇异值的变化不大于扰动矩阵的2-范数。

如果将 A 看成是所选取的图像块，奇异值稳定性特性可以用来抑制由于图像压缩/解压或者传输过程所引起的噪声对原始图像的影响，从而增强了图像认证方法的鲁棒性。

3. 奇异值分解与 PKI 相结合的图像认证方法

本文提出将奇异值分解(SVD)与 PKI 相结合生成鲁棒图像认证码，用来进行图像认证的方法。该方法首先利用奇异值分解，寻求原图像中能够容忍压缩或噪声污染引起的失真，但对非法篡改敏感的特征。由这些特征构成原图的内容摘要，并进行加密处理。在需要对图像内容的完整性和图像的所有权校验时，提取验证图像的相应特征，并与解密后的原始特征比较，从而确认图像内容是否已被篡改。整个过程如图1所示。

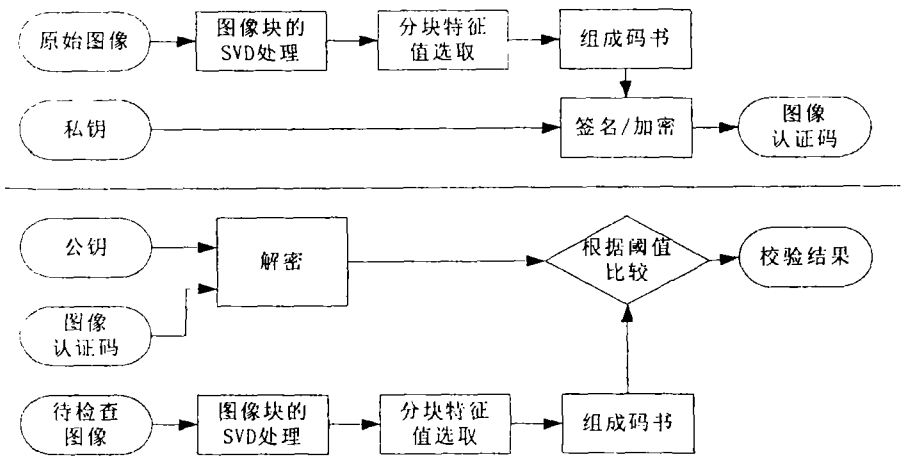


图1 SVD 与 PKI 相结合进行图像认证流程图

3.1 基于奇异值的特征提取与认证码的生成

通过 SVD，一个图像块矩阵可以表示为：$A=USV^T$，其中奇异值矩阵 S 由大到小分别对应原图的低、高频系数。将原始图像分解成若干个图像块，与篡改图像内容相比，正常操作（包括小于某个比率的压缩和引入噪声）并不会对每个图像块低频系数对应的奇异值产生大的影响，因此可以将其奇异值作为不变特征提取。将提取的各个图像块特征排列，构成矢量，形成图像内容摘要，并用私钥加密后得到认证码。设原始图像为 $X=\{x_{ij}\}_{m\times n}$，图像的所有者已从认证中心得到一对密钥：私钥 KR 与公钥 KU。从原始图像提取不变特征并产生认证码的算法描述如下：

算法 SVD_PKI_COMBINED_IMAGE_MAC_BORN

输入：原始图像 X，所有者的私钥 KR

输出：基于原图像内容的认证码 T

第一步：图像大小预处理。将原始图像 $X=\{x_{ij}\}_{m\times n}$ 分成相互不覆盖、大小均等 $a\times b$ 的图像块 $Y_r=\{y_{ij}\}_{a\times b}$，其中 $r=1,2,\cdots,R$；$R=(m/a)\times(n/b)$。

第二步：图像块特征提取。对每个图像块分别进行如下处理：

图像块共有 R 个；

FOR I=1 to R

①对第 I 个图像块像素值组成的矩阵 A 奇异值分解 USV^T；

②从奇异值矩阵选取适当的元素作为不变特征 φ_i；

③映射 $F:\Phi\rightarrow[0,1]^k$ 将不变特征 φ_i 映射到一个长为 k 的二进制序列 β_i

END FOR

第三步：认证码生成。将 R 个长为 k 的二进制序列 β_i 合并成图像内容摘要 $S=\bigcap_{i=1}^{R}\beta_i$，用私钥 KR 加密后生成认证码 $T=ENC_{KR}(S)$

原始图像 X 通过网络或其他媒介进行传输之前，这了提高传输效率，需要被压缩。在传播过程中，图像也可能被噪声污染。但是，这些操作没有改变原图像内容，仅仅影响图像质量。用 X' 表示质量受损的原图。原始图像的认证码 T 一般以两种方式伴随 X'。一种是 T 以标签的方式附加在 X' 后，用 X"=(X',T) 表示；另一种是 T 作为水印用数字隐藏算法嵌入到 X' 中，用 X"=X'·T 表示。

3.2 图像认证处理与特征匹配

图像的认证过程，如图1的下半部分所示。它由三个部分组成。首先，待认证的图像 X" 采用特征提取同样的方法，计算 R 个图像块的不变特征值 λ_i。随后，原图像的认证码 T 被公钥 KU 解密，恢复出图像内容摘要 $S=DEC_{KU}(T)$，并按照逆映射 $F':[0,1]^k\rightarrow\Phi$ 获得原始特征值 φ_i。第三步，根据选定的阈值，逐一比较抽取的特征值 λ_i 和原始特征值 φ_i 的差值，判断图像内容是否已经被篡改。如果所有的差值都大于阈值，说明

所接受的图像是可信的，通过验证；否则，就意味着图像内容已被更改，或者图像质量受到严重损坏。

在获取原始图像的不变特征时，通过确定它所能容忍的最大压缩，或最严重噪声，可以计算图像验证时选取阈值的大小，如图2所示。设原始图像的不变特征值为 φ；所能容忍的最大质量受损图像的不变特征值为 $\hat{\varphi}$；选取阈值 $V_{th}=\max(\text{abs}(\varphi-\hat{\varphi}))$。

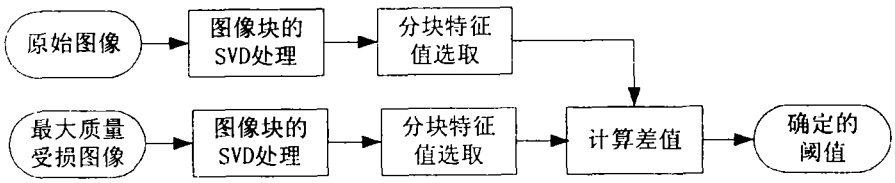
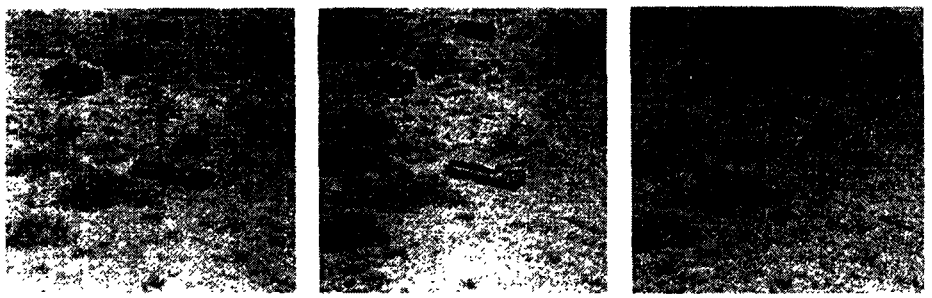


图2 阈值选取方法

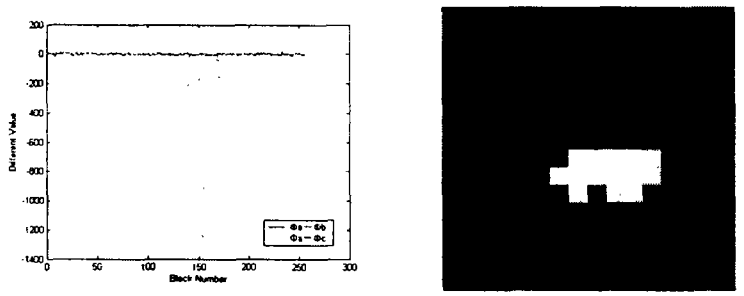
4. 实验结果

4.1 一个算例

以256×256的256色灰度 CarAPC 图，来直观说明本文方法的有效性。将原图3(a)分成16×16小块，对每个图像块对应灰度值矩阵奇异值分解，获得特征值 Φ_A。图3(b)、图3(c)分别为 JPEG 压缩(质量因子=40)的图像、更改内容后的图像。以相同的方法计算图3(b)、图3(c)的特征值 Φ_B、Φ_C。图3(d)为 $\Phi_B-\Phi_A$ 与 $\Phi_C-\Phi_A$ 变化曲线。阈值 $V_{th}=\max(\text{abs}(\Phi_B-\Phi_A))=18$。$\Phi_C-\Phi_A$ 曲线表明图3(c)的内容已经被篡改，并指出被篡改部分在图像中的具体位置，如图3(e)所示。



(a) 原始图像　(b) JPEG 压缩图像(QF=40)　(c) 内容篡改后的图像



(d) 特征值差异变化曲线　(e) 由特征值差异变化曲线确定篡改位置

图3 一个算例(a)原始图像，(b)JPEG 压缩图像(质量因子=40)，(c)内容篡改后的图像，(d)特征值差异变化曲线，(e)由特征值差异变化曲线确定篡改位置

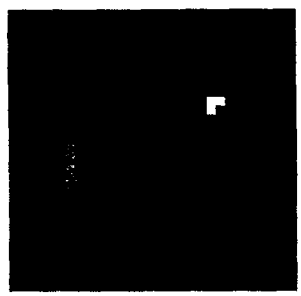
4.2 与其他典型图像认证方法比较

4.2.1 Der-Chyuan Lou 等图像块灰度均值法　在文[3]中，Der-Chyuan Lou 等采用图像块灰度平均值作为图像认证的依据，将每个8×8图像块的灰度值平均并四级量化。每个量化值用2bits 表示，整幅图像的认证码共2×(256/8)×(256/8)=2048bits。实验表明，Lou 方法在阈值=2.5时可以发现图4(b)已经被篡改，并且能够指出内容改变的位置，如图4(d)所示。但是对于 EZW(Embedded Zero-tree Wavelet)方法压缩的图4(c)(bpp=0.5，PSNR=32.9dB)，图像内容并未被篡改，却不能通过 Lou 方法的验证。如图4(e)所示，Lou 方法指出图4(c)的内容已经被篡改，这与实际情况不符，是错误的验证结果。

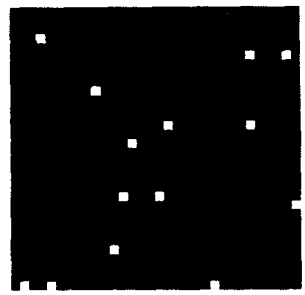
本文方法对于这两幅图像能够很好区分，如图4(f)所示，在阈值 $V_{th}=22$时，图4(c)通过验证，图4(b)不能通过验证。图4(g)指出图4(b)内容被篡改的位置。与 Lou 方法相比，本文提出的方法鲁棒性较好，能够容忍较大压缩比对原图像质量造成的影响。

4.2.2 Ching-Yung Lin 等区分 JPEG 压缩与恶意篡改图像内容的鲁棒图像认证方法　在文[4]中，Ching-Yung Lin 等先对每个8×8图像块 DCT 变换，得到全部图像块对应的 DCT 系数集合 P。将不同的 DCT 系数块 P_p 和 P_q 按照任意确定的映射函数 W 组成互不重叠的系数对 $P_q=W(P_p)$，它满足 $P_p\cap P_q=\Phi$，$P_p\cup P_q=P$。再对变换系数进行 zig-zag 排序，用 F_p、F_q 表示系数块对应的系数矢量。计算各个系数矢量的差 $\Delta F_{p,q}=F_p-F_q$。对每个 $\Delta F_{p,q}(v)$(其中 $v=1,2,\cdots$，图像块总个数/2)，确定循环次数 N、选择每次循环所选取的 $\Delta F_{p,q}(v)$ 中低频系数差的个数 b_n 以及固定的门限制 k_n，逐个比较，产生认证码。在进行图像认证时，考虑到压缩量化引入的噪声，

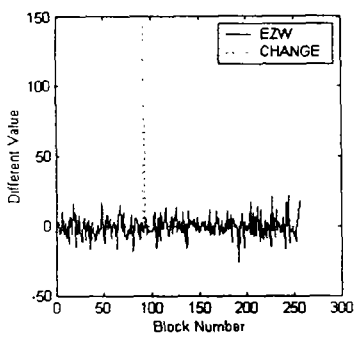
(a) 原始图像　(b) 内容篡改图像　(c) EZW 压缩图像 (bpp=0.5, PSNR = 32.9 dB)



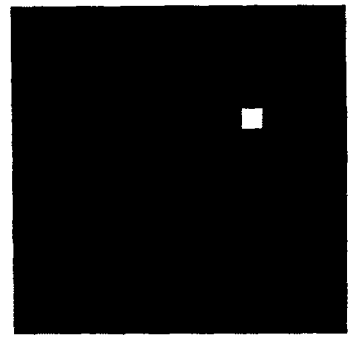
(d) Lou 方法正确指出(b)篡改位置(阈值=2.5)



(e) Lou 方法错误指出(c)篡改位置(阈值=2.5)



(f) 本文方法对(b)、(c)验证结果



(g) 本方法正确指出(b)被篡改位置,(c)通过验证

图4　本文方法与 Lou 方法的比较

选择容忍度 τ。本实验选用 $N=1, b_1=10, k_1=0$ 生成 Lin 方法的认证码，图像验证时选用 $\tau=16$。分别将原图像进行了基于 DCT 的 JPEG 压缩(质量因子 QF=40，PSNR=32.7dB)，如图5(b)；基于 EZW 的压缩(bpp=0.8，PSNR=36.1dB)，如图5(c)；JPEG2000(bpp=0.5)压缩，如图5(d)；篡改图像内容(右眼被改变)，如图5(e)；加入高斯噪声(方差=0.01)，如图5(g)。实验结果表明，Lin 方法能够在一定程度容忍 JPEG 压缩对图像质量造成的影响，图5(b)顺利通过验证，如图(g)所示；而遭受恶意篡改的图5(e)不能通过 Lin 方法的认证，验证结果如图5(h)所示。但是，Lin 方法不能容忍 EZW 压缩(图5(c))、JPEG2000压缩(图5(d))以及高斯白噪声(图5(f))引起的图像质量下降。虽然这几幅图像内容与原图相比，没有变化，但它们不能通过 Lin 方法的认证，如图5(i)所示。

采用本文提出的方法，未曾篡改原图像内容的图5(b)、图5(c)、图5(d)以及图5(f)虽然存在质量下降，但可以通过验证，而被篡改图像5(e)不能通过验证。验证结果如图5(j)所示，图5(k)根据验证结果指出图5(e)中被篡改的位置。

与 Lin 方法相比，本文提出的方法正确验证篡改图像的同时，不仅容忍基于 DCT 变化的 JPEG 压缩引起的质量下降，而且对基于 EZW 算法和 JPEG2000压缩的图像具有鲁棒性，另外，也能忍受高斯白噪声一定程度的污染。

结束语　本文将奇异值分解与 PKI 相结合，提出一种图像认证方法，它容忍压缩或噪声对原图像质量造成的影响，但不允许内容遭到非法更改的图像通过验证。对比实验结果表明，与几种典型图像认证方法相比，该方法选用奇异值分解结果作为特征，构成原图像的内容摘要，并由用户定义特征匹配的阈值，可以适应各种不同的压缩方法，同时保持对恶意篡改的敏感性。因此本文提出的方法，有益于解决图像数据的真实性和完整性问题。

参 考 文 献

1 Schneier B. Applied Cryptography. John Wiley & Sons, Inc. 1996

2 Wong P W, Memon N. Secret and public image watermarking schemes for image authentication and ownership verification. IEEE Trans. On. Image Processing, 2001, 10(10): 1593～1601

3 Lou D-C, Liu J-L. Fault resilient and compression tolerant digital signature for image authentication. IEEE Trans. On. Consumer Electronics, 2000, 46(1): 31～39

4 Lin C-Y, Chang S-F. A robust image authentication method distinguishing JPEG compression from malicious manipulation. IEEE Trans. On. Circuit System & Video Technology, *Feb. 2001*

5 Xie L, Arce G R, Graveman R F. Approximate Image Message Authentication Codes. IEEE Trans. On. Multimedia, *2001*, 3(2): 242～252

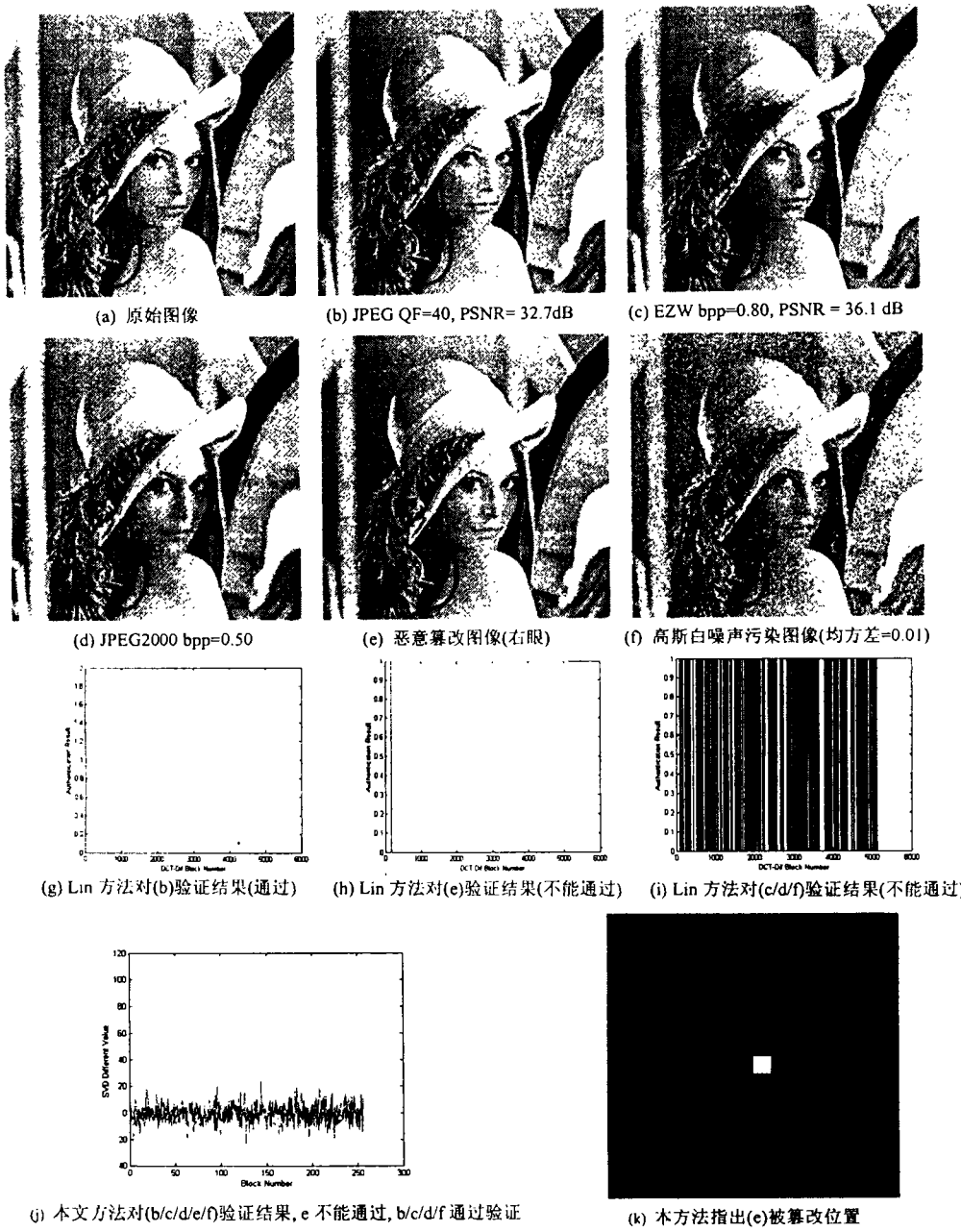


图5 本文方法与 Lin 方法的比较

6 Sun Q, Chang S-F, Maeno K, Suto M. A new semi-fragile image authentication framework combining ECC and PKI infrastructures. ISCAS 2002

7 Venkatesan R, Koon S-M, Jakubowski M H, Moulin P. Robust image hashing. ICIP2000

8 Maeno K, Sun Q, Chang S-F, Suto M. New semi-fragile image authentication watermarking techniques using random bias and non-uniform quantization. EI2002

9 Wu C W, On the design of content-based multimedia authentication systems. IEEE Trans. On. Multimedia, 2002, 4(3): 385～393

10 Schneider M, Chang S-F. A robust content based digital signature for image authentication. ICIP96

11 Klema V C. The singular value decomposition: Its computation and some applications. IEEE Trans. On. Automatic Control, 1980, 25: 164～176

12 Hong Z. Algebraic feture extraction of image for recognition. Pattern Recognition, 1991, 24(3): 211～219