doi:10.3969/j.issn.1000-436x.2014.04.018

数字可视媒体取证

冯春晖[1]，徐正全[1]，郑兴辉[1]，蒋力[2]
（1. 武汉大学 测绘遥感信息工程国家重点实验室，湖北 武汉 430079；2. 郑州大学 信息工程学院，河南 郑州 450001）

摘 要：对数字可视媒体取证技术的来源及概念进行了介绍，从原理上详细介绍了具有代表性的二次压缩取证及篡改取证算法。在此基础上，对现有取证算法的相关性以及取证技术中存在的问题进行了一定深度上的讨论，并提出了可视媒体取证发展的新思路。
关键词：固有特征；二次压缩检测；篡改检测；反取证
中图分类号：TN919.82 **文献标识码**：A **文章编号**：1000-436X(2014)04-0155-11

Digital visual media forensics

FENG Chun-hui[1], XU Zheng-quan[1], ZHENG Xing-hui[1], JIANG Li[2]
(1. State Key Laboratory of Information Engineering in Surveying,
Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China;
2. School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China)

Abstract: The origin and concept of digital visual media forensics is introduced, representative double compression detection and forgery detection algorithms are introduced in detail afterwards. On this basis, the relationships between different forensic algorithms and existing problems of visual media forensic techniques are discussed. Perspectives of the development of digital visual media forensics are put forward in the end.
Key words: intrinsic features; double compression detection; forgery detection; anti-forensics

1 引言

可视媒体（visual media）又称视觉媒体，是多媒体信息中的主要类型。视觉是人类认知世界最重要的手段之一，人类接受信息的 80%以上来自视觉。视觉对象包括文字、图像、视频和数字几何等，本文的研究对象就是其中的图像和视频，并称之为可视媒体[1]。

现今世界上到处充斥着数字化的可视媒体，而诸如 Photoshop 和 Premiere 等图像和视频编辑软件也日益成熟并具有简易的操作性，使得对数字可视媒体的篡改变得简单和普遍。而篡改后的图像或视频可能对大众产生误导，甚至带来严重的后果。例如，在法庭上，经过伪装的照片或监控录像可能伪造出犯罪嫌疑人不在现场的假象或者其他虚假场景，影响案件的判断；新闻机构如果发布虚假的照片或视频，则会造成错误的或煽动性的舆论导向；在保险机构，投保人可能会伪造出物品损坏的照片以骗取保险赔偿。诸如此类的可能性使得人们迫切需要一种能够对可视媒体的可信性进行有效鉴别的技术。

早期，科学家们运用数字水印技术鉴别可视媒体的真伪。数字水印技术需要在可视媒体产生时嵌入附加信息，作为后期检测的依据。但由于大多数可视媒体获取设备不具有嵌入附加信息的功能，使得数字水印的应用受到了很大的限制。这种情况下，数字可视媒体取证（digital visual media forensics）技术应运而生。其于 2002 年左右发端[2]，并在过去 10 年间迅速发展，成为一个理论基础和实现手段较为多样、实现思想较为发散的交叉型研究领域。取证

收稿日期：2013-04-21；**修回日期**：2013-11-30
基金项目：国家重点基础研究发展计划（“973”计划）基金资助项目（2006CB303104, 2011CB302204）
Foundation Item: The National Basic Research Program of China (973 Program)(2006CB303104, 2011CB302204)

技术不需要在可视媒体中嵌入附加信息，也不依赖于特殊的设备或参考文件，任何人拍摄的任一图像或视频，均可利用取证技术分析和鉴定。

“取证”一词在牛津词典中有 2 个含义。一是与警察破案时所作的科学测试相关，二是与法庭上法官和律师等对案件的分析相关。在取证过程中，执法人员需利用所有可能的物证，如 DNA、指纹等，还原案件发生的原貌；而数字可视媒体取证的概念与之相似，其通过对可视媒体中固有特征（intrinsic features）的分析，还原可视媒体的获取及后期处理的历史[2]，以确定可视媒体的可信度[3]。

取证算法需要分析的固有特征，即取证所需的物证，包括像素相关性、压缩效应、抽象统计特征等。与数字水印不同的是，这些特征是在可视媒体拍摄和后期处理过程中自然产生的，并非人为嵌入的附加信息。文献[4,5]认为取证技术包括主动取证和被动取证，主动取证包括数字水印及数字签名技术，被动取证为仅分析可视媒体固有特征以鉴定其可信度的技术。但根据牛津词典中对取证一词的定义，物证并非取证者事先加入，而是在案件发生时自然产生，所以取证不应包含主动之义，因为它本身就是被动地利用既有信息进行分析的[2,6,7]。

一般来说，一个完整的取证算法包括两部分，首先是提取有效特征，其次是利用有效特征生成检测算法，完成可信度的判别。其中提取出具有良好区分能力的特征是取证算法的核心。因此以下介绍取证算法时，重点在于对有效特征的讨论。

所谓对可视媒体历史的还原可分为 3 个部分。首先是还原可视媒体的来源，如获取设备的制造商及型号，判断可视媒体是否为某一特定设备所产生[7]，或检测视频和图像为直接拍摄自然景物还是由二次投影所得；其次是还原压缩历史，检测可视媒体是否经历了二次或多次的压缩编码；最后是还原后期处理及各种篡改操作。经由这三类取证手段，就能对可视媒体的真实性有一个整体和准确的判断。

如前所述，现今社会对可视媒体可信度鉴别的需求较大且较为迫切，可视媒体取证技术作为可能的解决方案之一具有特殊的优势。然而到现在为止其发展尚未完全成熟，仍有较大的探索和完善的空间。

篡改取证是取证算法中最为重要和应用最广泛的部分，因此本文试图在概括地介绍可视媒体取证技术的基础上，详细讨论其中的篡改取证算法。

2 可视媒体二次压缩检测

一般地，压缩格式的图像或视频的篡改都需要在解码后的空间域进行，然后再存储为压缩格式。也就是说，篡改会导致二次（或以上）的压缩。因此可以将二次压缩检测视为篡改取证的预取证，将二次压缩过程视为可视媒体可能经过篡改的依据。

JPEG 格式是一种应用最为广泛的静止图像压缩格式，其编码过程会影响图像的某些固有特征，如像素相关性等；但它同时又在图像中引入了一些新的特征，如 DCT 系数的变化、块效应等。取证算法正是利用这些新引入的特征来进行压缩检测。

相比于图像，视频文件因其数据量巨大，绝大多数都需要存储为压缩格式。而无论是早期的 MPEG-2 还是较新且已普及的 H.264，这些复杂的压缩编码也在视频数据中留下了各种可被取证算法利用的特征。

图像或视频二次压缩编码历史取证的研究成果较多，按其所利用的固有特征可分为以下几类。

2.1 利用 DCT 系数变化特征检测对准的二次压缩

所谓对准的（aligned）二次压缩，指的是前后两次压缩编码的宏块划分完全重合。二次压缩是否对准会使可视媒体压缩系数和压缩效应等产生不同的变化。本节讨论的均为对准的二次压缩检测。

2.1.1 利用 DCT 直方图模式变化检测不同量化矩阵二次压缩

当一幅 JPEG 图像经不同量化矩阵二次压缩后，重新解码得到的量化后 DCT 系数直方图就会出现周期性的空值和峰值，如图 1 和图 2 所示。这种周期模式被定义为双压缩效应(double quantization effect)[8]。Popescu 和 Farid[9]对双压缩效应的产生原因做了详细分析；Lukáš 和 Fridrich[10]提出了 3 种利用双压缩效应来估计首次压缩量化矩阵的方法，并利用估计结果来区分图像是否经过二次压缩。

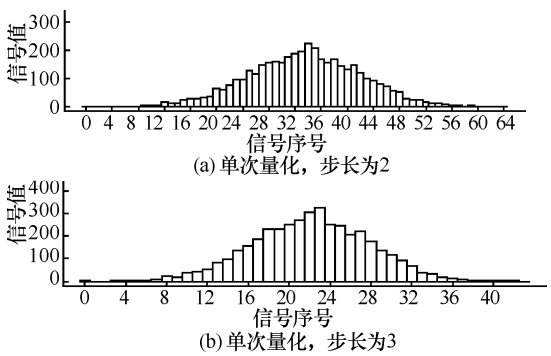


图 1 单次量化后的直方图

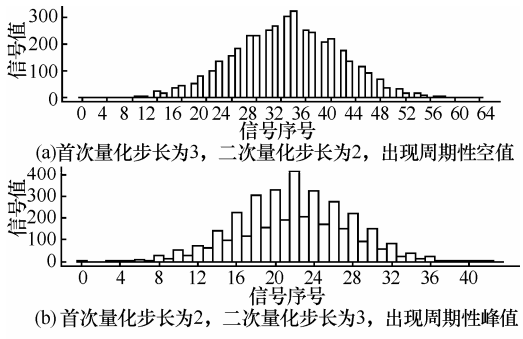


图 2　二次量化后的直方图

Wang 和 Farid[11]通过检测 MPEG 视频中 I 帧的双压缩效应来判断视频是否经过二次压缩；Su 和 Xu[12]提出，恒定比特率的视频文件中，帧内各宏块的量化因子是变化的，从而导致双压缩效应的不稳定。但此时 I 帧 DCT 系数直方图会呈现出稳定的凸起（convex）模式，根据这种模式可以判定 MPEG-2 视频是否经过了恒定比特率的二次压缩。

2.1.2　利用 Benford Law 检测不同量化矩阵二次压缩

除上述 2 种变化特征外，Li 和 Shi[13]根据二次压缩前后 DCT 系数首位数字的分布是否符合对数分布的 Benford's Law[14]来区分 JPEG 图像是否经过二次压缩。文章指出，单次压缩的 JPEG 系数[注1]首位数字分布与广义的 Benford's Law

$$p(d)=N\lg\left(1+\frac{1}{s+d^{q}}\right),\ d\in\{1,2,\cdots,9\} \qquad (1)$$

基本一致（其中 $p(d)$为概率分布，N 为标准化因子，s 和 q 用来控制不同质量因子 JPEG 系数的分布）；而二次压缩的 JPEG 系数则会明显违背 Benford's Law 的分布趋势，如图 3 所示。

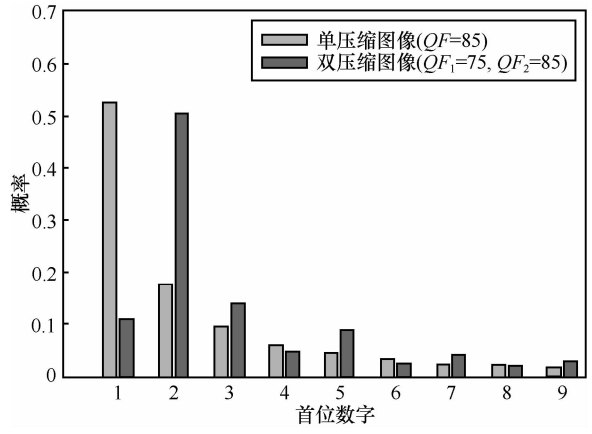


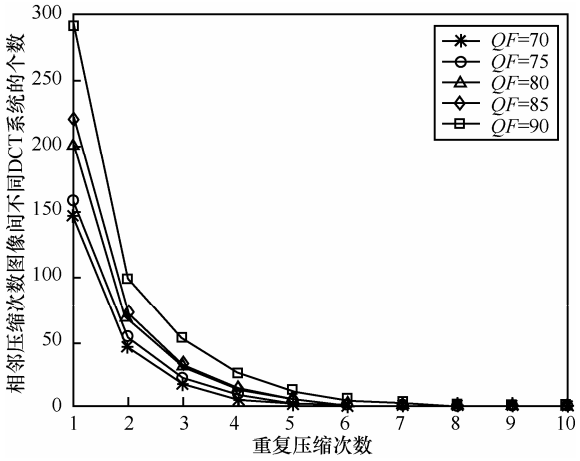
图 3　随机选取单压缩 JPEG 图像（量化因子 85）与二次压缩图像（首次量化因子 75，二次量化因子 85）JPEG 系数的首位数字分布

注1：文献[13]将量化后 DCT 定义为 JPEG 系数。

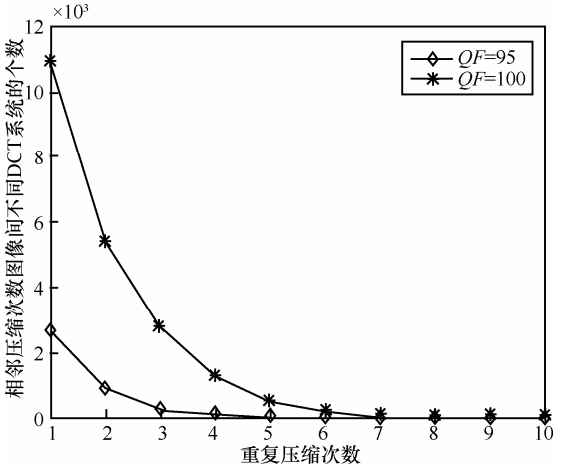
Chen 和 Shi[15]同样利用了 JPEG 系数与 Benford's Law 的符合度判定 MPEG 视频文件是否经过了二次压缩。

2.1.3　利用 DCT 系数变化检测相同量化矩阵二次压缩

当篡改者用相同的量化矩阵对图像进行二次 JPEG 压缩时，DCT 系数直方图的模式不会产生明显的改变，但 DCT 系数本身仍然发生着不易察觉但极具规律性的变化。Huang 和 Huang 等[16]提出，在多次同量化矩阵压缩下，相邻压缩次数的图像间不同 DCT 系数的个数呈递减的趋势，如图 4 所示。利用这种特征可检测同量化矩阵的二次 JPEG 压缩。



(a) 量化系数以 5 为增量，从 70 到 90



(b) 量化系数分别为 95 和 100

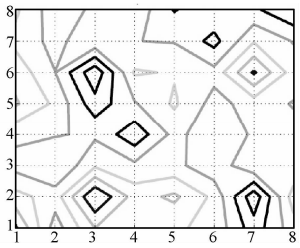
图 4　多次同量化矩阵 JPEG 压缩下 DCT 系数变化规律统计

2.2　利用块效应检测未对准的二次压缩

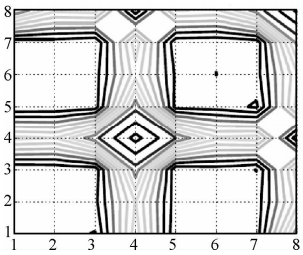
以上对图像和视频对准的二次压缩检测均取得了较好的结果，但当首次编码的分块被打散再错位分块压缩（reshuffle）后，也即二次压缩未对准

时，DCT 系数的规律性变化也会被破坏。当图像篡改者任意将原图剪切掉若干行或若干列后再次压缩，宏块划分完全对准的可能性只有 3.125%[17]，因此对准的二次压缩取证在很大可能性上会失效。针对这种情况，Luo 和 Qu[17]提出了一种基于块效应（blocking effect）[18]的检测 JPEG 图像剪切后重压缩的取证算法。其首先计算出一种可以表征块效应的特征矩阵 BACM（blocking artifact characteristics matrix），并提出，单次压缩图像的 BACM 等高线图是规则的，而经过剪切再压缩的图像则会呈现不规则的扭曲，如图 5 所示。利用 BACM 计算出多维特征向量，训练支持向量机，以区分待检测图像是否为剪切后再次压缩所得。由于文献[17]检测的是不同量化矩阵压缩下的 JPEG 图像，Liu[19]提出了一种基于块效应的检测同量化矩阵未对准二次压缩的取证算法。

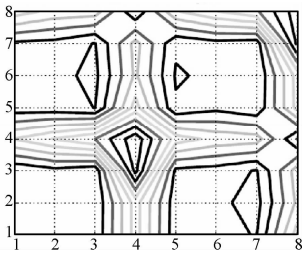
(a)原图 Lena



(b)未压缩的 Lena 图像不存在块效应



(c) 单次 JPEG 压缩 Lena 图像的 BACM 等高线为较规律的垂直和水平线条



(d) 剪切后再次压缩 Lena 图像的 BACM 等高线图在图 5(c)的基础上有明显的扭曲

图 5 单次压缩与剪切后再压缩块效应比较

可以说，基于块效应的取证算法与基于 DCT 系数变化的算法在检测可视媒体二次压缩上是互为补充的。

3 可视媒体篡改检测

可视媒体的篡改定义为可能会造成原始媒体信息理解错误或丢失的恶意编辑[20]。现实中对可视媒体的恶意篡改事件可谓层出不穷。可以说，篡改检测是应用最为广泛的一类取证方法。本文在介绍具体的篡改检测算法前先给出常见的图像及视频的篡改方法。

3.1 可视媒体常见篡改方法

3.1.1 常见图像篡改方法

1) 复制移动（copy-move）：复制图像中的部分区域，并移动到同一图像的不同位置上。这种篡改常伴有旋转缩放等几何变换以及羽化融合等过程，使复制区域能自然融入新的背景中。一些图像编辑软件中的工具，如纹理修补（in-painting）等也属于复制移动的范畴。

2) 剪切粘贴（cut and paste）：剪切图像中部分区域，粘贴到另外一幅图像上。这种篡改手段也需要与几何变换及羽化融合等操作相结合；有时还需要调整粘贴区域的光照，因为不同图像中的光照条件未必一致。

3) 文献[21,22]将复制移动及剪切粘贴共同归类于拼接操作（splicing）。

4) 各种图像增强操作，如改变亮度、对比度、颜色饱和度和清晰度等。

3.1.2 常见视频篡改方法

视频由多帧图像组成，因此所有的图像篡改手段都可以用于视频篡改。此外，视频篡改还包括帧的删除或增加，改变码率以及图像组（GOP，group of pictures）结构等。

以下按可揭示图像及视频篡改的不同固有特征的顺序，对可视媒体篡改检测进行介绍。

3.2 利用像素相关性检测篡改

可视媒体的采集及修改过程通常会使相邻像素产生特殊的相关性。一般由像素相关性形成的特征对低质压缩产生的噪声以及几何变换较为敏感。

3.2.1 像素克隆

复制移动操作在本质上即为像素的克隆。由于复制移动篡改一般要对复制区域进行几何变换和羽化融合等后续处理，因此取证算法不能仅仅计算各区域内像素灰度是否相同，而是需要提取对相关操作顽健的特征。Fridrich 和 Soukal 等[23]将分块量化后的 DCT 系数作为对融合操作和压缩噪声顽健的特征，检测复制粘贴篡改。相似的文献还有文献[24~26]等。但这些方法均对几何变换不具有顽健性。Huang 和 Guo 等[27]利用 SIFT （scale invariant feature transform）算法从图像中提取出在缩放和旋转等操作下保持不变的特征，达到了检测几何变换下的复制移动篡改的目的。

Wang 和 Farid[28]将视频序列分段，并提取出各个子序列在空间和时间上的相关性，具有较强相关性的序列被认为是经过复制移动篡改的。此方法对 MPEG 压缩顽健，且具有较高的检测效率。

3.2.2　CFA 插值

在可视媒体的获取过程中，光线通过镜头后，需通过颜色滤镜阵列（CFA，color filter array）后才能使感光元件感光，如图 6 所示。感光元件上的每个像素仅对一种色光感光。最后生成的彩色图像中每个像素上其他 2 种颜色均是通过 CFA 插值算法得到的。这种插值算法使整个图像中的像素具有了线性或非线性的相关性。当图像被篡改后，这种相关性很可能被破坏掉。Popescu 和 Farid[29]通过计算图像中由 CFA 插值形成的像素相关性是否一致来检测图像的篡改。此算法不能抵抗图像篡改后重新进行 CFA 插值的伪装手段，同时对 JPEG 压缩噪声也比较敏感。

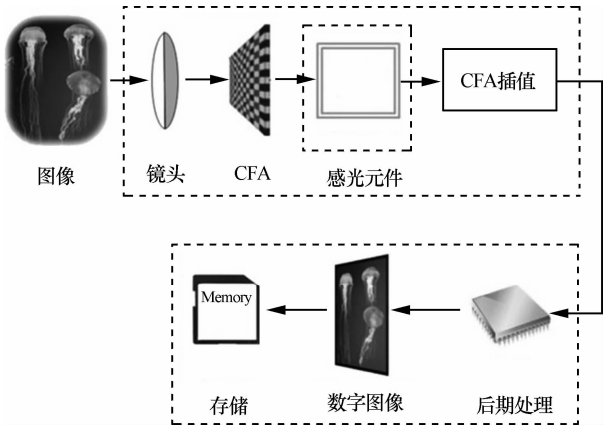


图 6　标准数字图像获取流程

3.2.3　重采样

篡改操作在破坏像素间原有相关性的同时，也会引入新的相关性。对图像进行剪切粘贴操作时，一般要对粘贴区域进行缩放和旋转等几何变换，达到与被粘贴图像中的景物比例一致的目的。而几何变换的实现需要在重采样后进行插值，这样就引入了新的像素相关性。以一维离散信号 $X=\{a,b,c\}$ 进行上采样为例，将 X 与矩阵

$$\boldsymbol{A}=\begin{bmatrix}1 & 0 & 0\\ 0.5 & 0.5 & 0\\ 0 & 1 & 0\\ 0 & 0.5 & 0.5\\ 0 & 0 & 1\end{bmatrix}$$

相乘，插值后的信号 Y 为

$$Y=\boldsymbol{A}X^{\mathrm{T}}=\begin{bmatrix}1 & 0 & 0\\ 0.5 & 0.5 & 0\\ 0 & 1 & 0\\ 0 & 0.5 & 0.5\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}a\\ b\\ c\end{bmatrix}=\begin{bmatrix}a\\ 0.5a+0.5b\\ b\\ 0.5b+0.5c\\ c\end{bmatrix}$$

可知，插值后信号的奇数行均为其相邻行之和的 1/2。Popescu 和 Farid[30]探讨了图像中由于重采样而引入的像素间相关性，并使用 EM（expectation/maximization） 算法得到图像像素间相关性的概率图，以定位篡改区域。这种检测算法受低质 JPEG 压缩形成的噪声影响较大，亦不能抵抗篡改后重新插值的伪装手段。

3.2.4　去隔行算法

视频文件在产生过程中也会形成特殊的像素相关性。Wang 和 Farid[31]指出，摄像机一般将一帧图像分成两场拍摄，一场在 t 时间拍摄，另一场在 t+1 时间拍摄。隔行扫描的视频将两场简单地结合在一起，从而会产生梳状效应，如图 7 所示；为减少梳状效应，非隔行扫描的视频利用去隔行算法对相邻行的像素进行插值，从而引入像素间的相关性。当某帧图像被篡改后，这种相关性即被破坏，从而可以据此定位篡改区域。

(a) 隔行扫描视频帧中产生的梳状效应

(b) 非隔行视频应用隔行算法，去除了梳状效应

图 7　隔行扫描视频与非隔行视频帧比较

3.3　利用压缩特征检测篡改

压缩编码形成的特征不仅可以检测二次压缩，更重要的是可以应用于篡改检测。

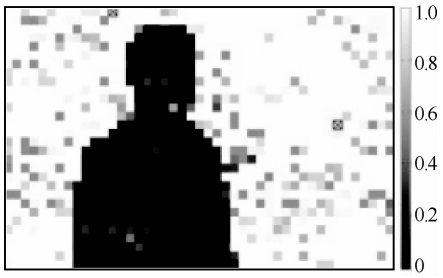
3.3.1　DCT 系数变化特征

在由两幅 JPEG 图像拼接形成的篡改图像中，

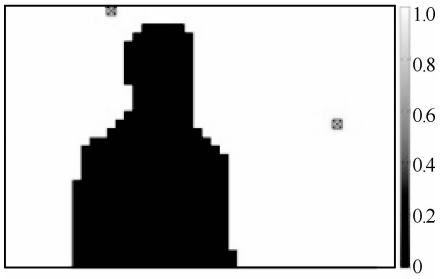
伪装区域（粘贴区域）与非伪装区域均经历了二次压缩，但由于粘贴操作的随机性，伪装区域二次压缩对准的概率较小[17]，由 2.2 节中的讨论可知，此时伪装区域不具有双压缩效应；而非伪装区域因两次压缩宏块划分完全重合而具有双压缩效应。在这个前提下，Lin 和 He 等[8]提出了一种对图像融合等操作顽健的自动定位伪装区域的剪切粘贴检测算法。

Wang 和 Farid[32]对视频篡改中的双压缩效应进行了进一步的探究。其理论前提与文献[8]相同，即 I 帧中非伪装区域具有双压缩效应，伪装区域无双压缩效应。其通过建模得到双压缩宏块的 DCT 系数的边缘分布 $P_{q1}(z)$(式 2)，然后得到待检测 I 帧中具体宏块的系数分布 $P(z)$。计算二者距离 D(式 3)，得出相应宏块为对准型二次压缩的概率。最后得出 I 帧中所有宏块为双压缩宏块的概率分布图，经过滤波去除误检区域，得到伪装区域。如图 8 所示。

(a) 将人物合成于静止背景的图像帧，背景具有双压缩效应，而前景无此效应



(b) 各宏块为双压缩宏块的概率



(c) 运用中值滤波后得到的篡改检测结果

图 8 帧内篡改检测过程

$$P_{q_1}(z)=\sum_x P_{q_1}(x)P_{q_1}(z\mid x)$$

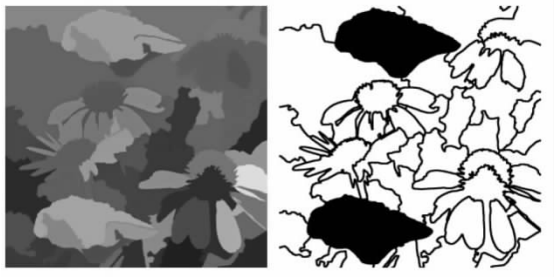
$$=\sum_x P_{q_1}(x)\int_{(z-0.5)q_2}^{(z+0.5)q_2} N(y;xq_1,\sigma)\mathrm{d}y \qquad (2)$$

$$D(P(z),P_{q_1}(z))=\sqrt{\sum_z \frac{(P(z)-P_{q_1}(z))^2}{s^2(z)}} \qquad (3)$$

3.3.2 块效应

由压缩编码产生的块效应也可以用来检测图像或视频的篡改。由文献[17]可知，非对准二次压缩区域中的块效应表征为不规则的扭曲状，而对准的二次压缩区域上的块效应则比较规则。Barni 和 Costanzo 等[33]利用这种规律，提出了基于块效应的图像篡改检测算法。此算法首先将图像分割成同质区域（如图 9 所示），再用文献[17]的检测方法分析各同质区域的压缩特性，最后定位篡改区域。这种基于语义（semantic）的取证算法在很大程度上降低了误检率和运算复杂度。

(a) 原始图像 (b)篡改图像



(c) 同质区域由相同的颜色表示 (d) 篡改检测输出

图 9 基于块效应的图像篡改检测过程

Luo 和 Wu 等[34]根据 MPEG 文件中帧间编码图像的块效应变化来检测帧的删除或增加操作以及图像组结构的变化。其提出，P 帧和 B 帧中的块效应受多种因素的影响，如参考帧中块效应的传播，不同量化矩阵和不同运动补偿算法的影响等。当某些帧被删除后，前后两次压缩的块效应相互叠加，并随被删除帧的个数而变化。根据其变化曲线即可

检测相应的篡改。

3.3.3　帧间预测误差

Wang 和 Farid[11]探讨了视频序列删除若干帧后产生的两类特征。首先是空间压缩特征，即 I 帧中因二次量化产生的双压缩效应；其次是时间压缩特征，当删除序列中的若干帧以后，来自不同图像组的帧组合至同一图像组，从而导致视频序列的运动残差均值（mean motion error）产生周期性变化。这种周期性反映在傅立叶频谱上即为中频上明显的峰值。此方法的缺陷在于不能检测整数倍图像组的删除，并且需要人为观察傅里叶频谱的峰值特征[35]。

3.4　利用抽象统计特征检测篡改

利用抽象统计特征的取证与其他取证算法不同，其大多没有直接在理论上证明篡改对可视媒体数据的具体影响，而是利用一些抽象的，被实验证明具有良好区分性质的统计特征（statistical feature），如频域系数的统计矩等，形成多维特征向量并训练分类器，以区分篡改图像与自然图像。

设一维信号 $x(t)$的傅里叶变换为 $Y(w)$，其功率谱 $P(w)=Y(w)Y^*(w)$可用来对信号的频率成分进行分析。文献[36]的作者对语音信号拼接对频率域统计特征的影响进行了分析，其指出，信号经过拼接操作后，功率谱并没有相应的变化，反映频率域高阶相关性的双阶谱 $B(w_1,w_2)=Y(w_1)Y(w_2)Y^*(w_1+w_2)$却会发生变化。当平滑信号经过拼接后，会产生一定的非连续性，表现在双阶谱振幅的增加以及相位的偏移。Ng 和 Chang[21]进一步将此原理应用于图像信号的拼接检测。由于双阶谱不受白噪声的干扰，所以此类方法对压缩编码噪声等具有顽健性。

Shi 和 Chen 等[22]利用更复杂的统计特征来捕捉拼接对图像连续性和周期性等方面的影响，并在与文献[37~39]使用相同的拼接检测评估图像库[40]的条件下得到了更高的检测率（92%）。其首先得到图像像素的二维阵列和多尺寸 DCT 系数块的二维阵列，然后利用特征方程统计矩和马尔可夫转移概率求得多维特征向量，组成具有区分拼接图像和自然图像功能的自然图像模型。由于单一尺寸的块 DCT 系数特征很难捕捉复杂多变的拼接操作，文中对多种尺寸的 DCT 系数分块进行了特征提取，实验证明，此类特征大幅度提高了的检测率。

Farid 和 Lyu[41]计算了小波系数的两类统计特征，第一类是小波分解各层子带系数的统计矩，包括均值、方差、偏度和峰度；第二类是所谓的高阶统计特征，用于捕捉不同子带上小波系数的相关性，用预测误差 E（式(6)）表示。设 $V_i(x,y)$为 i 阶垂直子带上的小波系数，则相邻层系数对其幅值的线性预测为

$$\begin{aligned}&|V_i(x,y)| = w_1|V_i(x-1,y)| + w_2|V_i(x+1,y)| + \\ &w_3|V_i(x,y-1)| + w_4|V_i(x,y+1)| + \\ &w_5|V_{i+1}(x/2,y/2)| + w_6|D_i(x,y)| + w_7|D_{i+1}(x/2,y/2)|\end{aligned} \tag{4}$$

其中，w_i 为加权值。式（4）的矩阵形式为

$$\boldsymbol{V} = \boldsymbol{Q}\boldsymbol{w} \tag{5}$$

其预测误差 E 可表示为

$$E = \mathrm{lb}(\boldsymbol{V}) - \mathrm{lb}(|\boldsymbol{Q}\boldsymbol{w}|) \tag{6}$$

计算小波系数在各阶不同方向上的相关性，得到 72 维特征向量，并训练分类器，可达到同时区分多种篡改图像的目的。

抽象统计特征往往反映了可视媒体信息中比较本质的变化与相关性，其不受加性噪声等较弱干扰的影响，并具有同时检测多种篡改操作的潜力。

3.5　利用噪声模式检测篡改

在图像和视频的获取过程中，由于感光元件的不完美性等原因，每幅图像或视频序列的每一帧均会带有固定的噪声模式。而篡改操作则会对这些噪声产生破坏。Chen 和 Fridrich 等[42]利用最大似然估计检测图像各个块中由相机传感器产生的噪声模式 PRNU（photo-response non-uniformity noise），以定位图像的篡改区域。

与文献[42]不同的是，Hsu 和 Hung 等在文献[43]中定义的噪声同时包含了传感器模式噪声和视频帧中的高频分量等。其首先计算每一帧与自身去噪版本的差值，得到噪声误差，再以宏块为单位计算相邻帧噪声误差的相关性 r（如图 10 所示），当 r 发生突变的时候，即可认为相应宏块被篡改。

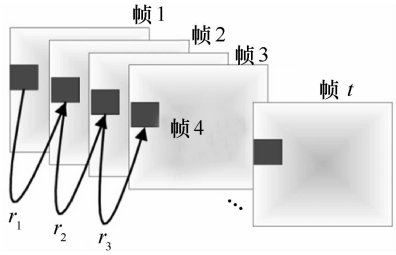


图 10　时域相邻帧各个块的噪声误差相关性 r 示意

文献[44]利用辐照依赖噪声（irradiance- dependent noise）的均一性检测静止场景视频的篡改。

3.6 利用物理特征检测篡改

基于景物物理特征的检测技术，通过对光照、投影、空间几何等场景信息的估计和测度，利用其在不同图像区间之间的差异来判断图像的完整性[5]。由于剪切粘贴图像内伪装区域较难与周围景物的光线条件保持一致，文献[45,46]的作者分别对单幅图像单一光源的方向进行二维及三维建模，根据图像各区域光照方向的不一致性检测图像合成篡改。文献[47]对实际景物中的多光源光照环境进行建模，用于检测复杂光照条件下图像的合成。文献[48,49]通过分析物体的几何特征是否符合自然透视规律来检测拼接操作。

4 可视媒体取证技术发展趋势

通过以上讨论可知，现有的可视媒体取证算法从多种角度对可视媒体的获取和后期处理历史进行了较全面的分析，但它同时也存在着一些问题。

首先，由于现有取证算法绝大多数只针对常规的图像或视频篡改操作进行检测，当熟知取证算法的篡改者将伪装操作留下的痕迹进行去除或掩盖后，取证算法则会失效。这种试图对可视媒体的伪装痕迹进行隐藏的算法称为反取证算法(anti-forensics)[50]。反取证算法的研究起步较晚，已发表的成果也较少：Stamm 和 Liu[51]在不同量化表二次 JPEG 压缩图像的 DCT 系数中加入噪声，消除了 DCT 系数直方图中存在的特殊模式，从而使基于双压缩效应的取证算法失效；Feng 和 Xu[52]通过替换 JPEG 图像篡改区域的 DCT 系数，避免了篡改过程中的全局二次压缩，去除了篡改区域以外的不同量化矩阵或相同量化矩阵的二次压缩效应；文献[53]和文献[54]分别提出了隐藏重采样痕迹和伪造新的 CFA 插值相关性的反取证算法；Cao 和 Zhao[55]提出了一种使对比度增强操作不被检测的反取证算法。事实上，每种取证算法均可能存在攻破它的反取证算法。可以说取证与反取证是一种相互促进的关系，反取证的发展会催生更具顽健性的取证算法，顽健的取证算法又会推动生成更复杂有效的反取证算法。但反取证的研究到目前为止还很不充分，因此可抵抗多种攻击算法的取证算法自然也难以发展。

其次，不同取证算法之间不能直接比较性能的高低。这主要是因为缺少具有公信力的评价基准及标准测试数据库。相比之下，同属媒体安全领域的数字水印技术已拥有相对成熟的若干套评测系统[56~58]，其中包括一系列常见的水印攻击算法。如果取证算法与反取证算法相继成熟，则能发展出用于测试取证算法抗攻击性的评价系统。此外，取证算法还需要有公用的篡改图像或视频测试数据库，以比较各种算法在相同测试条件下的检测率。文献[22,37~39]利用了同一个拼接图像数据库[40]进行了无攻击条件下的图像拼接检测，分别得到了 82%、80%、72%和 92%的检测率。但现有的篡改图像尤其是视频数据库在数据量和内容的多样性上还非常有限。

第三，实际中伪装图像或视频大多不是单一操作手段的产物，而是各种伪装手段反复叠加后形成的。这些操作相互作用，很可能会影响相关特征的提取。而现存的取证算法大都只能检测单一的或是互不干扰的操作过程[59]；而且，在很多情况下，如法庭或新闻机构对可视媒体进行取证时，需要对某一幅图像或一段视频的真伪进行准确的判定，而现有取证算法的检测率均没有达到 100%，其检测结果也就不能作为完全可信的依据。由以上两点可知，现有取证算法尚不能应用于实际。

对于数字可视媒体取证技术的发展方向，本文有以下几点思考。

1) 取证工具箱的思想。Fridrich [23]提出了取证工具箱（FTS, forensic tool set）的思想。即按检测目的将取证算法分类，建成一个取证工具集，其中每一种算法可能并不完全有效，但将集合中的取证工具联合起来使用，则能对可视媒体的来源及可信度有一个较准确的判断。持相似想法的还有数字图像取证之父 Hany Farid[60]。实际上，很多取证算法往往都需要与另一种算法互为补充，才能得到某个完整的功能。例如基于 DCT 系数变化的二次压缩取证与基于块效应的二次压缩取证等。如果能将适用范围有限的算法有机地结合，则可以生成功能更强的取证算法。

2) 重点发展基于抽象统计特征的取证算法。由于抽象统计特征往往更接近于可视媒体信号的内在规律及本质，因此可以在有效地反映图像或视频信号相关变化的同时，不受加性噪声等弱干扰的影响；另外，这类方法可以同时检测多种篡改操作，其内在原因在于其具抓住了自然形成的数字可视媒体的综合特征，使之可以与多种篡改手段下形成的非自然可视媒体相区分。从长远角度看，这种取证算法比针对单一篡改方法的取证具有更强的效

用。另外，由于统计特征的抽象性，此类取证算法更不容易产生相应的反取证算法。

3) 继续发展反取证算法。通过研究反取证算法，可视媒体安全的研究者可以获知现有取证算法的缺陷，生成取证算法评价基准以量化取证算法的可信度[35]，促进更具顽健性的取证算法的生成。

4) 继续发展视频取证算法。由于视频编码复杂性以及视频数据的多维性，视频中可被取证利用的特征比图像更加丰富。相比于图像，在多帧的视频序列中提取的特征更具稳定性。现今对视频取证的研究主要集中在 MPEG 格式上，而 H.264 格式的取证算法甚少。

5) 与语义相结合。将人类对可视媒体内容的认知融合在取证算法中，可以降低误检率和提高检测效率，并可能具有区分单纯的视效增强操作和恶意篡改操作的能力[2]。

5　结束语

本文对数字可视媒体取证技术的来源、概念以及应用进行了介绍，对具有代表性的压缩历史取证算法和篡改历史取证算法进行了归类与比较，并从原理上进行了较为详细的讨论。

可视媒体取证算法利用图像和视频中固有的特征，对其获取设备、后期处理以及伪装过程进行还原。这些特征丰富多样，如由压缩产生的 DCT 系数变化、块效应；由 CFA 插值及去隔行算法等引入的像素间相关性；反映可视媒体内在规律的抽象统计特征以及噪声模式等。各种取证算法之间往往存在着互补的关系，对取证算法的综合运用可以得到更全面和准确的检测结果。

当今社会对可视媒体可信度鉴别的需求量较大且极为迫切，可视媒体取证技术具有无需嵌入附加信息、适用范围较广等优点，但其发展尚未成熟，距大规模实际应用仍存在着一定距离。需要继续对反取证算法以及多种手段有机结合的取证算法进行研究，以促进更加顽健和适用于实际应用的取证技术的产生。

参考文献：

[1] 徐正全, 徐彦彦. 可视媒体信息安全[M]. 北京: 高等教育出版社, 2012.
XU Z Q, XU Y Y. Visual Media Information Security[M]. Beijing: Higher Education Press, 2012.

[2] JUDITH A R, WIEM T, DUGELAY J L. Digital image forensics: a booklet for beginners[J]. Multimedia Tools and Applications , 2011, 51(1):133-162.

[3] CHUANG W H, SU H, WU M. Exploring compression effects for improved source camera identification using strongly compressed video[A]. 2011 18th IEEE International Conference on Image Processing[C]. Belgium, 2011.1953-1956.

[4] FARID H. Image forgery detection[J]. IEEE Signal Processing Magazine, 2009, 26(2): 16-25.

[5] 曹刚, 赵耀, 倪蓉蓉. 多媒体内容认证[J]. 中国计算机学会通讯, 2011, 7(2): 38-44.
CAO G, ZHAO Y, NI R R. Multimedia content authentication[J]. China Computer Federation Communications, 2011, 7(2): 38-44.

[6] LIAN S G, KANELLOPOULOS D, RUFFO G. Multimedia information system security[J]. Informatica, 2009, 33(1): 3-24.

[7] XU G S, SHI Y Q. Camera model identification using local binary patterns[A]. 2012 IEEE International Conference on Multimedia and Expo[C]. Melbourne, 2012.392-397.

[8] LIN Z C, HE J F, TANG X O, *et al*. Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis[J]. Pattern Recognition, 2009, 42(11):2492-2501.

[9] POPESCU A C, FARID H. Statistical tools for digital forensics[J]. Information Hiding, 2005, 3200: 128-147.

[10] LUKAS J, FRIDRICH J. Estimation of primary quantization matrix in double compressed JPEG images[A]. Digital Forensic Research Workshop[C]. Cleveland, 2003.

[11] WANG W H, FARID H. Exposing digital forgeries in video by detecting double MPEG compression[A]. MM&Sec '06 Proceedings of the 8th Workshop on Multimedia and Security[C]. New York, 2006.37-47.

[12] SU Y T, XU J Y. Detection of double compression in MPEG2 videos[A]. 2010 2nd International Workshop on Intelligent Systems and Applications[C]. Wuhan, China, 2010.1-4.

[13] LI B, SHI Y Q, HUANG J W. Detecting doubly compressed JPEG images by using mode based first digit features[A]. 2008 IEEE 10th Workshop on Multimedia Signal Processing[C]. Queensland, 2008. 730-735.

[14] FU D D, SHI Y Q, SU W. A generalized Benford's law for JPEG coefficients and its applications in image forensics[A]. Security, Steganography and Watermarking of Multimedia Contents IX[C]. San Jose, 2007. 6505.

[15] CHEN W, SHI Y Q. Detection of double MPEG compression based on first digit statistics[J]. Digital Watermarking, 2009, 5450:16-30.

[16] HUANG F J, HUANG J W, SHI Y Q. Detecting double JPEG compression with the same quantization matrix[J]. IEEE Transactions on Information Forensics and Security, 2010, 5(4): 848-856.

[17] LUO W Q, QU Z H, HUANG J W, *et al*. A novel method for detecting cropped and recompressed image block[A]. IEEE International Conference on Acoustics, Speech and Signal Processing[C]. Vancouver, 2007.II-217-II-220.

[18] REEVE H C, LIM J S. Reduction of blocking effects in image coding[J]. Optical Engineering, 1984, 23(1): 34-37.

[19] LIU Q Z. Detection of misaligned cropping and recompression with the same quantization matrix and relevant forgery[A]. The 3rd International ACM Workshop on Multimedia in Forensics and Intelligence[C]. Scottsdale, 2011.25-30.

[20] 秦运龙. 压缩域视频被动取证研究[D]. 上海大学, 2010.
QIN Y L. The Research of Passive Video Forensics in Compressed Domain[D]. Shanghai University, 2010.

[21] NG T T, CHANG S F. A model for image splicing[A]. 2004 International Conference on Image Processing[C]. Singapore, 2004.1169-1172.

[22] SHI Y Q, CHEN C H, CHEN W. A natural image model approach to splicing detection[A]. MM&Sec '07 Proceedings of the 9th Workshop on Multimedia and Security[C]. New York, 2007.51-62.

[23] FRIDRICH J, SOUKAL D, LUKAS J. Detection of copy-move forgery in digital images[A]. Proceedings of Digital Forensic Research Workshop[C]. 2003.

[24] POPESCU A C, FARID H. Exposing Digital Forgeries by Detecting Duplicated Image Regions[R]. Dartmouth College, Hanover, NH, Tech Rep TR2004-515, 2004.

[25] LI G.H, WU Q, TU D, *et al*. A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD[A]. 2007 IEEE International Conference on Multimedia and Expo[C]. Beijing, China, 2007.1750-1753.

[26] LANGILLE A, GONG M L. An efficient match-based duplication detection algorithm[A]. The 3rd Canadian Conference on Computer and Robot Vision[C]. Quebec, 2006.64.

[27] HUANG H L, GUO W Q, ZHANG Y. Detection of copy-move forgery in digital images using SIFT algorithm[A]. IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application[C]. Wuhan, China, 2008.272-276.

[28] WANG W H, FARID H. Exposing digital forgeries in video by detecting duplication[A]. MM&Sec '07 Proceedings of the 9th Workshop on Multimedia & Security[C]. Dallas, 2007. 35-42.

[29] POPESCU A C, FARID H. Exposing digital forgeries in color filter array interpolated images[J]. IEEE Transactions on Signal Processing, 2005, 53(10):3948-3959.

[30] POPESCU A C, FARID H. Exposing digital forgeries by detecting traces of re-sampling[J]. IEEE Transactions on Signal Processing, 2005, 53(2):758-767.

[31] WANG W H, FARID H. Exposing digital forgeries in interlaced and deinterlaced video[J]. IEEE Transactions on Information Forensics and Security, 2007,2(4):238-258.

[32] WANG W H, FARID H. Exposing digital forgeries in video by detecting double quantization[A]. MM&Sec '09 Proceedings of the 11th ACM Workshop on Multimedia and Security[C]. New York, 2009. 39-48.

[33] BARNI M, COSTANZO A, SABATINI L. Identification of cut & paste tampering by means of double-JPEG detection and image segmentation[A]. 2010 IEEE International Symposium on Circuits and Systems[C]. Paris, 2010.1687-1690.

[34] LUO W Q, WU M, HUANG J W. MPEG recompression detection based on block artifacts[A]. Security, Forensics, Steganography, and Watermarking of Multimedia Contents X[C]. San Jose, 2008. 68190X-1-68190X-12.

[35] STAMM M C, LIN W S, LIU K J R. Temporal forensics and anti-forensics for motion compensated video[J]. IEEE Transactions on Information Forensics and Security, 2012, 7(4): 1315-1329.

[36] FARID H. Detecting Digital Forgeries Using Bispectral Analysis[R]. AI Lab Massachusetts Institute of Technology, Tech Rep AIM-1657, 1999.

[37] CHEN W, SHI Y Q, SU W. Image splicing detection using 2-D phase congruency and statistical moments of characteristic function[A]. Proc SPIE, Electronic Imaging, Security, Steganography, Watermarking of Multimedia Contents IX[C]. San Jose, 2007.

[38] FU D D, SHI Y Q, SU W. Detection of image splicing based on hilbert-huang transform and moments of characteristic functions with wavelet decomposition[A]. Digital Watermarking, Proceeding of 5th International Workshop on Digital Watermarking[C]. Jeju Island, 2006.177-187.

[39] NG T T, CHANG S F, SUN Q B. Blind detection of photomontage using higher order statistics[A]. IEEE International Symposium on Circuits and Systems[C]. Vancouver, 2004.688-691.

[40] Columbia image splicing detection evaluation dataset[EB/OL]. http://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet/AuthSplicedDataSet.htm, 2007.

[41] FARID H, LYU S W. Higher-order wavelet statistics and their application to digital forensics[A]. CVPRW '03, Conference on Computer Vision and Pattern Recognition Workshop[C]. Madison, 2003.94.

[42] CHEN M, FRIDRICH J, GOLJAN M, *et al*. Determining image origin and integrity using sensor pattern noise[J]. IEEE Transactions Information Forensics Security, 2008, 3(1):74-90.

[43] HSU C C, HUNG T Y, LIN C W, *et al*. Video forgery detection using correlation of noise residue[A]. Proceedings of IEEE Workshop Multimedia Signal Processing[C]. Queensland, 2008.170-174.

[44] KOBAYASHI M, OKABE T, SATO Y. Detecting forgery from static-scene video based on inconsistency in noise level functions[J]. IEEE Transactions on Information Forensics and Security, 2010,5(4): 883-892.

[45] JOHNSON M K, FARID H. Exposing digital forgeries by detecting inconsistencies in lighting[A]. ACM Multimedia and Security Workshop[C]. New York, 2005.1-10.

[46] JOHNSON M K, FARID H. Exposing digital forgeries through specular highlights on the eye[A]. The 9th International Workshop on Information Hiding[C]. Saint Malo, 2007.311-325.

[47] JOHNSON M K, FARID H. Exposing digital forgeries in complex lighting environments[J]. IEEE Transactions Information Forensics Security, 2007, 3(2): 450-461.

[48] JOHNSON M K, FARID H. Metric Measurements on a Plane from a Single Image[R]. Dept Comput Sci, Dartmouth College, Tech Rep TR2006-579, 2006.

[49] JOHNSON M K, FARID H. Detecting photographic composites of people[A]. International Workshop on Digital Watermarking[C]. Guangzhou, China, 2007.

[50] STAMM M C, LIN W S, LIU K J R. Forensics anti-forensics: a decision and game theoretic framework[A]. 2012 IEEE International Conference on Acoustics, Speech and Signal Processing[C]. 2012.1749-1752.

[51] STAMM M C, LIU K J R. Anti-forensics of digital image compression[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3):1050-1065.

[52] FENG C H, XU Z Q, ZHENG X H. An anti-forensic algorithm of JPEG double compression based forgery detection[A]. Proceedings of the 2012 4th International Symposium on Information Science and Engineering[C]. Shanghai, China, 2012.159-164.

[53] GLOE T, KIRCHNER M, WINKLER A. Can we trust digital image forensics[A]. Proceedings of the 15th International Conference on Multimedia[C]. New York, 2007.78-86.

[54] KIRCHNER M, B¨OHME R. Synthesis of color filter array pattern in digital images[A]. Proceedings of Electronic Imaging: Media Forensics and Security[C]. San Jose, 2009.72540.

[55] CAO G, ZHAO Y, NI R R, *et al*. Anti-forensics of contrast enhancement in digital images[A]. Proceedings of ACM Multimedia and Security Workshop[C]. Roma, 2010.25-34.

[56] PETITCOLAS F A P, ANDERSON R J, KUHN M G. Attacks on copyright marking systems[A]. David Aucsmith (Ed), Information Hiding, Second International Workshop, IH'98[C]. Portland, 1998. 219-239.

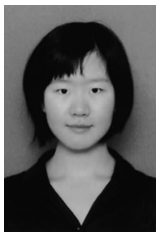
[57] Certification of watermarking techniques[EB/OL]. http://www.certimark.org/, 2002.

[58] SOLACHIDIS V, TEFAS A, NIKOLAIDIS N, *et al*. A benchmarking protocol for watermarking methods[A]. 2001 IEEE International Conference on Image Processing[C]. Thessaloniki, 2001.1023-1026.

[59] BESTAGINI P, FONTANI M, MILANI S, *et al*. An overview on video forensics[A]. The 20th European Signal Processing Conference[C]. Bucharest, 2012.1229-1233.

[60] http://www.fourandsix.com/about-us/hany-farid-phd-chief-technology-officer.html[EB/OL]. 2012.

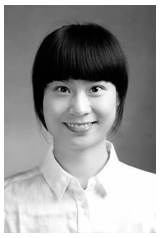
作者简介：



冯春晖（1985-），女，黑龙江哈尔滨人，武汉大学博士生，主要研究方向为多媒体信息安全、多媒体信息处理。

徐正全[通信作者]（1962-），男，湖北英山人，博士，武汉大学教授，主要研究方向为多媒体信息处理、多媒体网络通信、多媒体信息安全。E-mail:xuzq@whu.edu.cn。

郑兴辉（1984-），男，河南新密人，武汉大学博士生，主要研究方向为影像增强、影像分割。



蒋力（1985-），女，广西柳州人，博士，郑州大学讲师，主要研究方向为多媒体信息安全、图像识别。