第 37 卷 增刊 (I)　　东南大学学报 (自然科学版)　　Vol 37 Sup(I)
2007 年 9 月　　JOURNAL OF SOUTHEAST UNIVERSITY (Natural Science Edition)　　Sept 2007

基于内容的图像 Hash算法及其性能评估

叶 卫国　　韩水华

(厦门大学管理学院, 厦门 361005)

摘要: 介绍了一个有效的图像鉴别系统设计应具备的 4种特征: 鲁棒性、易碎性、安全性和篡改可定位性, 综述主要的图像内容 hash算法以及常用的 4种信息安全机制. 进而分析和讨论了现有算法所存在的性能折中和安全问题, 并对目前流行的图像篡改检测算法通过实验进行深入评估和分析. 实验表明与 hash属性相关的 4种性能是互相冲突的: 鲁棒性要求在小的扰动下保持良好的健壮性, 易碎性则要求尽可能降低感知不相似输入的碰撞可能性, 安全性则可能阻止鲁棒性的获得. 从安全的角度看易碎性和安全性是非常重要的, 一个 hash算法应通过某种程度的性能折中来实现上述互为冲突的目标.

关键词: 图像 hash; 基于内容的鉴别; 性能折中; 安全机制

中图分类号: TP391　　**文献标识码**: A　　**文章编号**: 1001-0505(2007)增刊 (I)-0109-05

Performance evaluation for content-based image authentication

Ye Weiguo　　Han Shuihua

(School of Management, Xiamen University, Xiamen 361005, China)

Abstract The four features required to design an effective authentication scheme are first introduced, i e, robustness, fragility, security and localization. And some algorithms and frequently used security mechanisms are also reviewed. Then the performance trade-offs and related security issues among existing technology are discussed and analyzed. Finally an in-depth experimental analysis on the methods of detecting image tampering is presented. The results show that the four desirable performance-related hash properties conflict with one another: robustness demands small perturbations to keep itself, whereas fragility requires minimization of collision probabilities for perceptually distinct inputs, and perfect randomization generally would stand in the way of achieving robustness. Fragility and security are very important, and it is desired for the hash algorithm to achieve these conflicting goals to some extent by trade-offs.

Key words: image hash; content-based authentication; performance trade-off; security mechanism

在现实应用中, hash函数主要用作数字签名, 通过对所发送消息的鉴别, 使得消息接收方可以验证它来源的可靠性. 虽然传统的用于消息完整性验证的数据鉴别技术已趋于成熟, 而图像内容的鉴别问题还处于早期的发展阶段, 还有很多问题有待探索[1]. 例如, 在过去几年中, 人们已经提出过很多各不相同的算法, 但我们很难断定哪一种方法对保证图像完整性或更一般的多媒体文档完整性方面是最有效的. 因此, 有必要通过综述相关文献, 来理解这个问题的本质, 找出可能的研究问题点, 探索新的研究领域, 并对不同算法的相对性能进行评估. 本文的主要目的是, 对基于内容的图像鉴别方法的现状和所面临的问题进行评述, 提供对算法优缺点以及鲁棒和安全等方面的相对性能评估.

1　图像 hash函数的性能需求

假定 $H(\cdot)$ 是一个图像 hash函数, 图像 I 为输入参数, 可以得到一个输出 hash值 $v=H(I)$. 对于图像鉴别来说, hash函数必须满足一个鉴别系统的设计需求. 一些常用的图像 hash性能测量指标包括鲁棒性、易碎性、安全性和篡改可定位性等.

收稿日期: 2007-07-20
基金项目: 福建省新世纪优秀人才计划资助项目.
作者简介: 叶卫国 (1970—), 男, 博士, 讲师, yewg@xmu.edu.cn.

1)鲁棒性. 当同一个密钥用在一组感知相似性的图像上, 最后得到 hash值相似时, 就认为这个 hash函数 $H(\cdot)$ 是满足鲁棒性的.

2)易碎性. 当同一个密钥用在一组感知不相似性的图像上, 最后得到的 hash值不相似时, 就认为这个 hash函数 $H(\cdot)$ 是满足易碎性的.

3)安全性. 一个密码学 hash函数也必须满足 3个基本安全需求:

①单向性. 给定一个输出值 y, 必存在一个图像 x 满足 $H(x)=y$, 但是无法反推得到图像 x', 满足 $H(x')=y$.

②防碰撞性. 给定任一个图像 x, 不可能计算找出另一个不同的图像 $x' <> x$, 使得 $H(x)=H(x')$.

③钥匙保密性. 在钥匙未知时, hash值应该不容易被伪造或被估计, 因为图像的 hash值是由密钥产生的.

除此以外, 还有些技术特征需要考虑在内: 定位被篡改的区域以及 hash值长度等.

2 研究现状综述

图像鉴别技术以及图像 hash算法近几年来受到学术界和产业界的广泛关注, 已有很多不同的算法提出. 这些算法大致上可以分成 4类: ①基于图像统计特性的方法; ②基于图像变换系数对关系的方法; ③基于原始图像特征描述的方法; ④基于低层图像特征提取的方法.

2.1 基于统计学特性的方法

第 1类方法利用了图像块直方图的均值、方差和高次惯量等统计不变性属性. Schneider等[2]建议挑选图像块的灰度统计属性如均值和方差, 生成图像 hash. 他们推断这种统计特性对于小的图像扰动应具有很好的鲁棒性. 这一方法的主要缺陷是人们可以很容易地修改一副图, 但仍保持直方图不变. 这对任何依赖于灰度统计特性方法的安全性带来严重损害. Venkatesan等[3]利用图像小波变换, 开发出一种基于各子波统计矢量的图像 hash算法. 他们经过观察发现这种统计值如粗粒度子波均值和其他细粒度子波的方差, 在经历绝大多数的内容不变性图像处理变换后仍保持不变性. 虽然小波系数的统计量远比灰度统计量更鲁棒, 但他们并没能很好捕捉到内容的改变, 特别是那些恶意篡改后的特征. 因而, 一些可以获得更好几何不变性的新方法如基于惯量标准化[4]、Radon变换[5]和 Zernike惯量[6]被采纳. 这些方法所存在的主要缺陷是它们无法抵制剪切的攻击, 因为内容的丢失将直接导致惯量的改变.

2.2 基于关系对的方法

第 2类方法利用了诸如 DCT系数[7]、连续小波系数对[8]等不变性关系. Lin等[7]提出一种基于图像分块离散余弦变换的数字签名方法. 这一方法探究了不同块图像同一位置的 DCT系数对之间的不变性关系, 这种关系即使在 DCT系数被量化后仍保持不变, 因而它可以有效区分恶意操作和 JPEG压缩. 这样构造出的特征值对 JPEG压缩很鲁棒, 但是对其他一些感知不明显的修改却非常脆弱. Lu等[8]提出一种用于图像鉴别的结构化数字签名方法. 他们观察到在小波分解的子波中, 父子节点尽管是不相关的, 但是他们却是统计相关的. 特别地, 他们发现在连续尺度上的小波系数的差值对于几种内容不变性操纵大体上维持不变, 由此识别出这种父子对, 进而将它编码到数字签名中. 然而, 这种方法对于全局变换 (如小角度旋转和扭曲) 以及局部的几何扭曲很敏感, 实际上, 它们并没有引起图像的感知明显性变化.

2.3 基于粗略特征描述的方法

第 3类方法利用图像粗略特征对感知的显著性, 任何大一点的修改都将导致整个图像内容的明显差异. 被采纳的粗略特征主要包括: 低频 DCT系数[9]、二值 DC子波 (低分辨率的小波系数)[10]、奇异值分解 (SVD)的最强奇异矢量[11]、Fourier-Mellin变换的旋转不变性系数[12]. Fridrich等[9]提出一种基于图像的可视化 hash函数构造方法. 他们的想法是对低频 DCT系数大的修改会构成图像内容的显著改变, 为使这一过程依赖于密钥, 他们通过一把私钥将输入图像映射成零均值的随机平滑模式, DCT基矢量被低频 DCT系数所取代. 其得到的 hash具有对滤波操作的稳定性, 但是对于几何扭曲效果不佳, 并且它无法做到无冲突. MihçAk等[10]采纳迭代计算将图像块小波变换计算得到直流子波 (低分辨率的小波系数)二值化来生成特征矢量, 该方法具有对一般灰度级图像操作的不变性. 最近, Kozat等[11]提出通过保留图像块的

奇异值分解(SVD)中最大的奇异矢量和值, 来构造图像 hash的方法. Swaminathan等[12]提出通过保留Fourier-Mellin变换中的旋转不变性矢量来构造图像. 这些基于粗略特征描述的方法对感知不明显的图像修改, 显示出很好的鲁棒性, 但他们对于局部的篡改或内容改变仍表现得很脆弱.

2.4 基于低层图像特征的方法

第 4类方法利用了低层图像特征如边或特征点[13]包含了图像的基本内容的事实. 但这种方法也有很大的局限性, 对诸如放大、高量化、分辨率下调等一些感知不明显的修改很敏感, 鲁棒性较差. 最近, Monga等[14]基于人类视觉系统的特性——突触细胞对于强鲁棒性的特征如角、高曲率点等响应很强烈, 提出一个对特征点感知的图像 hash框架. 他们首先采用突触小波特征检测算法提取出重要的图像特征, 然后, 通过迭代计算, 锁定对感知不显著的扰动具有强不变性的图像特征点集合. 由于特征点检测器对内容篡改的内在敏感性, 这一方法可以获得很好的鲁棒性.

3 图像 hash算法的性能评估

3.1 鲁棒性和易碎性的评估

图像 hash的鲁棒性和易碎性已经被文献所广泛讨论. 然而, 仍缺乏一种客观的评价方法和结果. 在表1中, 基于相关的文献和定性化的评价, 给出了有关鲁棒性和易碎性的相对性能列表.

表 1　图像 hash算法的鲁棒性和易碎性比较

图像 Hash算法		内容保持不变						内容改变
		高压缩	调低分辨率	旋转	比例放大	局部扭曲	剪切	
基于统计的方法	灰度直方图					√	√	
	小波统计特性	√	√		√	√	√	
	惯量	√	√	√	√	√		√
基于关系的方法	DCT系数对	√						√
	小波系数对	√	√		√		√	√
基于粗略特征的方法	低频 DCT系数	√	√					√
	低分辨率的小波系数	√	√		√		√	√
	不变性的 FFT	√	√	√	√		√	√
	最大的 SVD值	√	√	√	√		√	√
基于低层特征的方法	边							√
	特征点							√
	突触小波系数	√	√	√	√	√	√	√

从表 1可以看出, 采纳低层特征的"突触小波"方法鲁棒性最好, 其次是基于最大的 SVD值、不变性的FFT和基于惯量统计的方法. 一些基于如边、特征点、DCT系数对和灰度直方图方法对大多数内容不变性的图像处理表现敏感.

3.2 安全机制评估

为防止伪造攻击, 必须考虑在 hash生成过程融入一种安全机制. 一般地, 现有的用于保护图像 hash免受攻击的安全机制可以大致分成 4种主要的类型 (见图 1): ①无密钥, 直接投影变换; ②先投影变换, 再通过密钥随机化; ③先用密钥对图像随机化, 再投影变换; ④用密钥控制变换的随机化.

方法 1　直接提取对常规的图像处理操作具有不变性的特征, 然后生成 hash值[2, 3, 7, 13], 如图 1(a)所示. 由于提取的特征都是大家所熟知的, 单独采用这些特征容易被伪造攻击者识破[15].

方法 2　先对图像作鲁棒变换, 再由密钥对所提取的特征进行随机化[14, 16], 如图 1(b)所示. 但这仅维持有限的安全性, 因为主要的鲁棒性值并不受钥匙保护, 攻击者可能生成一个内容截然不同的全新图像, 但同时能使特征值保持不变[17, 18].

方法 3　先由密钥对图像块进行随机化组合, 然后做相关的变换[10-12, 15], 如图 1(c)所示. 但因所采纳的小波、傅里叶变换均属正交变换, 通过统计学习, 攻击者仍可以获得鲁棒 hash函数的空间聚类特征, 进而生成一个内容截然不同的全新图像[19]. 此外, 这些变换还不具备定位功能, 无法实现对文档的篡改定位.

方法 4　先随机生成变换基函数, 然后将图像进行投影变换. 如果基函数未知, 敌方很难预测最后的投影变换结果. Maltin等[20]采纳高斯和曲线小波作为基函数, 然后实现 Rendlet变换, 这种变换的随机性

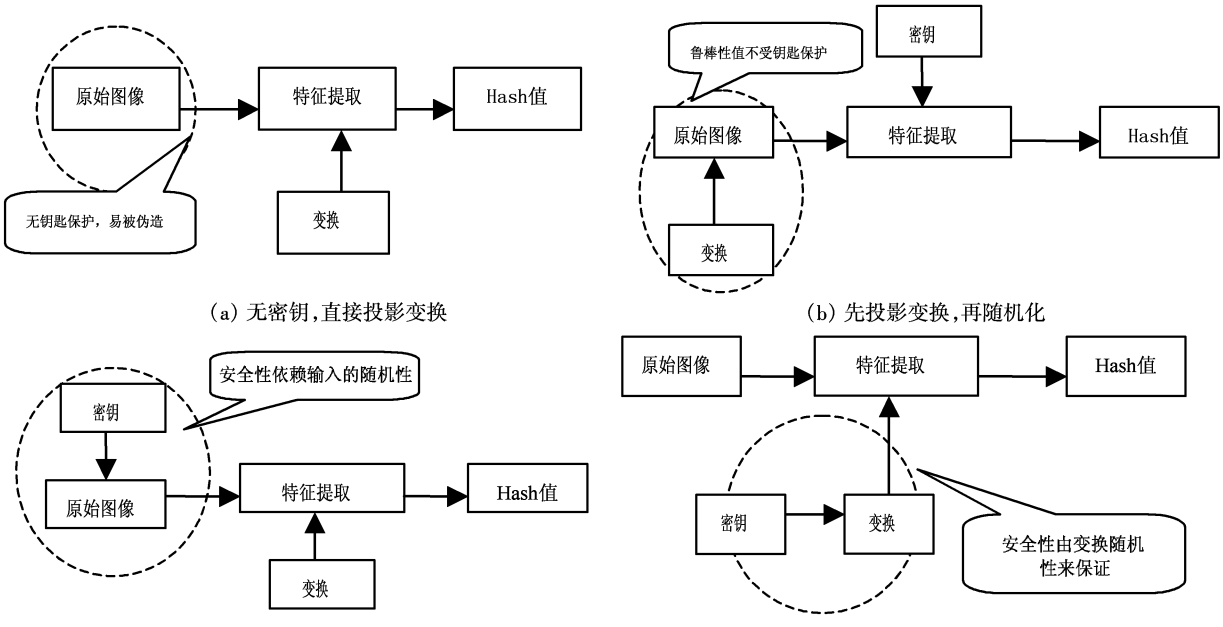

图 1 图像 hash算法的安全保护机制

使得攻击方很难发现真正的 hash取值. 虽然 Randlet变换的安全性很好, 但它的鲁棒性相对有限.

3.3 性能折中分析

本文对公开文献谈到的安全机制进行了解剖, 并就鲁棒性、易碎性、随机性和可定位性等方面总结得出其相对性能, 见表 2

表 2 图像 hash算法的性能评估

图像 Hash 算法		安全机制	鲁棒性	易碎性	随机性 (安全性)	可定位性
密码 Hash	MD5、SHA-1	D	低	高	高	无
基于统计的方法	文献 [2]	A	低	低	低	有
	文献 [3]	A	中	低	低	有
	文献 [4]	B	高	中	中	有
基于关系的方法	文献 [7]	A	低	低	低	有
	文献 [8]	A	中	低	低	有
基于粗略特征的方法	文献 [9]	C	中	低	中	无
	文献 [10]	C	高	低	中	无
	文献 [11]	C	中	中	中	无
	文献 [12]	C	中	中	中	无
	文献 [20]	D	中	中	高	无
基于低层特征的方法	文献 [13]	A	低	低	低	有
	文献 [14]	B	高	中	中	无

从表 2可以看出, 目前尚不存在能满足全部性能需求的完善算法. 虽然文献 [4, 10, 14]等提出的算法具有很高的鲁棒性, 但它们的脆弱性和随机性不够; 而文献 [20]提出的方法有很高的随机性, 但其鲁棒性和易碎性不够.

总的来说, 与 hash属性相关的 4种性能是互相冲突的. 第 1个属性要求在小的扰动下保持良好的鲁棒性, 而第 2个属性则要求尽可能降低感知不相似输入的碰撞可能性. 这里明显存在一种性能的折中问题, 例如, 如果采纳一些粗略的特征集, 虽然 hash值可以保持对变换的不变 (即鲁棒性), 但可能加大了感知不同图像碰撞性的可能性. 同样地, 就完全的随机性而言, 其要求输出 hash值分布的均匀性, 但这也可能阻止第 1种属性的获得. 从安全的角度看, 第 2种和第 3种属性是非常重要的, 它必须保证敌方很难通过操纵图像内容以获得相同的 hash值. 很显然, 一个 hash算法应通过某种程度的性能折中来实现上述互相冲突的目标.

4 结语

讨论了多媒体鉴别系统设计的一般需求和特征,并对目前用于内容鉴别的图像 hash技术进行深入的评述和性能分析. 由于图像内容操纵的方法很多,因篇幅限制,这里的比较做的还远远不够. 需要设计大规模的实验来完成相对性能的评估. 而且,作为一个新兴的研究领域,还存在很多值得深入研究和探索的地方, 如理论框架、鲁棒特征提取、hash算法的随机性以及图像 hash在认证水印中的应用等.

参考文献 (References)

[1] Rey C, Dugelay J L. A survey of watermarking algorithms for image authentication[J]. *EURASIP Journal on Applied Signal Processing*, 2002, 6(3): 613-621.

[2] Schneider M, Chang S F. A robust content based digital signature for image authentication[C]//*Proc IEEE Int Conf Image Processing*. Lausanne, Switzerland, 1996, 3: 227-230.

[3] Venkatesan R, Koon S M, Jakubowski M H, et al. Robust image hashing[C]//*Proc IEEE Conf on Image Processing*. Vancouver, Canada, 2000: 664-666.

[4] Alghoniemy M, Tewfik A H. Geometric invariance in image watermarking[J]. *IEEE Trans Image Process*, 2004, 13(2): 145-153.

[5] Simitopoulos D, Koutsonanos D E, Strintzis M G. Robust image watermarking based on generalized radon transformations[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2003, 13(8): 732-745.

[6] Kim H S, Lee H K. Invariant image watermark using Zernike moments[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2003, 13(8): 766-775.

[7] Lin C Y, Chang S F. A robust image authentication method distinguishing JPEG compression from malicious manipulation[J]. *IEEE Trans on Circuits and Systems for Video Technology*, 2001, 11(2): 153-168.

[8] Lu C S, Liao H Y M. Structural digital signature for image authentication: an incidental distortion resistant scheme[J]. *IEEE Trans on Multimedia*, 2003, 5(2): 161-173.

[9] Fridrich J, Goljan M. Robust hash functions for digital watermarking[C]//*Proc IEEE Int Conf Information Technology: Coding Computing*. Las Vegas, USA, 2000: 178-183.

[10] MihÇak M K, Venkatesan R. New iterative geometric methods for robust perceptual image hashing[C]//*Proc ACM Workshop Security and Privacy in Digital Rights Management*. Philadelphia, PA, 2001: 13-21.

[11] Kozat S S, Venkatesan R, MihÇak M K. Robust perceptual image hashing via matrix invariants[C]//*Proc IEEE Int Conf Image Processing*. Singapore, 2004, 5: 3443-3446.

[12] Swaminathan A, Mao Y, Wu M. Robust and secure image hashing[J]. *IEEE Transactions on Information Forensics and Security*, 2006, 1(2): 215-230.

[13] Dittmann J, Steinmetz A, Steinmetz R. Content-based digital signature for motion pictures authentication and content-fragile watermarking[C]//*Proc IEEE Int Conf on Multimedia Computing and System*. Florence, Italy, 1999, 2: 209-213.

[14] Monga V, Evans B L. Perceptual image hashing via feature points: performance evaluation and trade-offs[J]. *IEEE Transactions on Image Processing*, 2006, 15(11): 3452-3465.

[15] Fridrich J. Visual hash for oblivious watermarking[C]//*Proc SPIE: Security and Watermarking of Multimedia Contents*. San Jose, CA, USA, 2000, 2: 286-294.

[16] Kailasanathan C, Naini R C. Image authentication surviving acceptable modifications using statistical measures and *K*-mean segmentation[C]//*IEEE-EURASIP Work Nonlinear Sig and Image Proc*. Baltimore, USA, 2001.

[17] Holliman M, Memon N. Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemas[J]. *IEEE Trans on Image Processing*, 2000, 9(3): 432-441.

[18] Radhakrishnan R, Memon N. On the security of the SARI image authentication system[C]//*Proc IEEE Int Conf on Image Processing*. Thessaloniki, Greece, 2001, 3: 971-974.

[19] Radhakrishnan R, Xiong Z, Memon N. On the security of the visual hash function[J]. *Journal of Electronic Imaging*, 2005, 14(1): 1-10.

[20] Malkin M, Venkatesan R. The randlet transform: applications to universal perceptual hashing and image authentication[C]//*Proc Allerton Conf*. Monticello, IL, USA, 2004.