DOI:10.11721/cqnuj20130320

一种基于混沌理论的图像 Hash 算法*

于海鹏，文政颖
(河南工程学院 计算机科学与工程系，郑州 451191)

摘要:为了对数字图像进行快速有效的认证,提出了基于混沌理论的对数字图像进行认证的 Hash 算法。首先提出了按块对图像进行量化的思想,构造出图像矩阵,采用了混沌理论的 Logistic 映射方法,对图像矩阵进行置乱得到置乱矩阵,然后构造出差值矩阵,使用 N 次混浊调制生成调制矩阵,对调制矩阵进行二值化量化得到 1 bit 的 Hash 序列,通过多次调制和量化从而得到图像的 Hash 序列。仿真实验结果表明,算法能够有效认证数字图像,同时对于图像的缩放认证,算法也表现出较好的性能。

关键词:图像认证;混沌理论;Logistic 映射;图像 Hash
中图分类号:TP751;O415.5 **文献标志码**:A **文章编号**:1672-6693(2013)03-0099-04

在信息技术与互联网高速发展的过程中,数字图像、多媒体技术等技术已成为人们生活中不可缺少的一部分。数字图像应用领域相当广泛,除了在个人日常生活应用外,在军事领域、医疗、新闻出版、科学研究各个行业均广泛应用。图形图像技术随着计算机技术与网络通信的不断发展而迅速普及,与此同时,数字图像的真实性与完整性用肉眼几乎无法辨别。因此,对数字图像进行安全认证和完整性检测,如何确保数字图像的真实性便成为一个亟待解决的问题[1]。

数字图像的完整性与真实性认证与传统的消息完整性验证技术有所不同,数字图像认证技术具有鲁棒性、易碎性、安全性和篡改可定位等特性。近年来,针对图像内容的鉴别与认证问题,不少学者提出了各种各样的相关算法,现有的算法大致可以分为 4 类:1)基于图像统计特性的方法,如 Venkatesa 等人提出的利用图像小波变换统计各子波特性的图像 Hash 算法[2],虽然小波系数的统计量鲁棒性较好,但该算法并不能很好地发现图像内容的改变,特别是被恶意篡改后的特性;2)基于图像变换系数对关系的方法,如 Lin 等人提出的一种基于图像分块离散余弦变换的数字签名方法[3],该方法利用不同的块图像同一位置的 DCT 系数对之间具有不变性关系,使用此方法形成的图像有较好的鲁棒性,但它对图像中感知不明显区域的鉴别较为脆弱;3)基于粗略图像特征描述的方法,如 Fridrich 等人提出的基于 DCT 的图像可视化 Hash 函数构造方法[4],该算法对数字图像的滤波操作具有较好的稳定性,但对于几何扭曲效果不佳;4)基于低层图像特征提取的方法,如 Monga 等人提出的基于人类视觉系统特征对特征点感知的图像 Hash 框架[5],该算法具有很好的鲁棒性,但复杂度较高。

图像 Hash 算法是从图像中提取出能够表示图像特征的短序列标志图像,同时,在信息安全领域与密码学领域,混沌理论具有优良的密码学特性,本身具有离散性、随机性与不规则的非线性特性,对初始状态的敏感性以及对值域良好的遍历性使得它非常符合现代密码学中所要求的混淆原则和扩散原则。因此,将混沌理论与图像 Hash 算法有效结合,可以大幅提高图像的安全性与抗攻击性能。

1 图像 Hash 算法认证

图像 Hash 函数[6]主要用于数字签名,假定 $H(\cdot)$是一个图像的 Hash 函数,如果输入参数为 I,则可以得到输出的 Hash 值 $y=H(I)$。采用加密的 Hash 函数可以用 $H(I,K)$表示,其中 I 表示输入图像参数,K 表示加密密钥。因此,对图像的认证过程可以表示为:$R=H(I_1,K)\&H(I_2,K)$;如果图像 I_1 和 I_2 相似时,则用密钥 K 加密后的 Hash 函数相同或在很大程度上相似,R 为 1 或真值;如果图像 I_1 和 I_2 有

* **收稿日期**:2012-07-20 **网络出版时间**:2013-05-20 18:04
资助项目:河南省科技计划项目(No. 122300410174)
作者简介:于海鹏,男,讲师,硕士,研究方向为计算机应用与网络技术;E-mail:nethk@qq.com,**通讯作者**:文政颖,E-mail:wenzy2009@126.com
网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20130520.1804.201303.99_019.html

较大程度上的不同,或者本身就不是一幅图像,则用密钥 K 加密后生成的 Hash 函数会截然不同,则 R 为 0 或假值。一般来说,对 Hash 函数的性能评价指标有鲁棒性、易碎性、安全性等。目前,图像 Hash 算法的执行过程大致有 4 类,其中较为安全的一种为图 1 所示[7]。

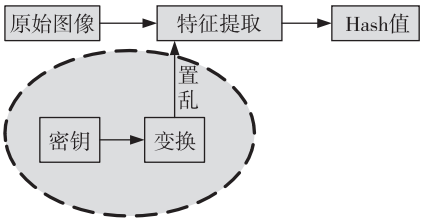

图 1 图像 Hash 算法产生流程图

针对任意图像,首先根据密钥值随机生成一个变换基函数,然后,将图像进行投影变换得到加密图像,对基再进行特征提取,最后得到图像的 Hash 值。这种方法的优点是基函数的随机性较大,不同的基函数会生成不同的 Hash 值,如果基函数未知,则得不到投影变换的图像结果,因此,这种变换过程生成的 Hash 值安全性很好。本文根据这种 Hash 值生成过程,对图像进行认证。

2 基于混沌理论 Hash 算法

2.1 混沌理论

Logistic 映射是一种典型的混沌系统的一种表示方式,由于其运算较为简单又易于实现,并且具有良好的混沌特性,成为目前研究最为广泛的一种动力系统[8]。Logistic 映射可以用一种非线性迭代方程表示

$$x_{k+1}=\mu x_k(1-x_k),x_k\in(0,1) \qquad (1)$$

其中,$k=0,1,2,\mu$ 为分岔参数,并且有当 $0<\mu\leqslant 4$,分岔图像如图 2 所示。从图中可以明显看出,图中右部分函数点的分布处于混沌状态,具体为当 $\mu>3.569\ 945\ 672$时进入了混沌区;当 $\mu=4$ 时,混沌映射为满映射,且混沌序列分布在 0～1 之间;映射生成的混沌序列$\{x_1,x_2,\cdots,x_i,\cdots\}$,从序列取值上观察,混沌序列显示出随机性、不确定性、收敛性与非周期性等特性,而且混沌序列对初始值是敏感的;从统计理论上分析,混沌序列显示出自相关性和互相关性为零的性质。

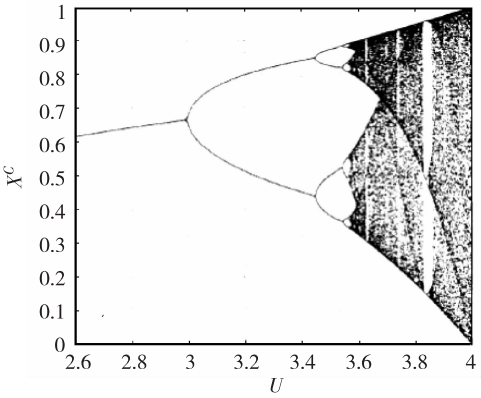

图 2 Logistic 映射分岔图像

2.2 图像置乱

文献[9]提出了一种按照图像 bit 位生成 Hash 值的算法,具有较好的认证性能,但由于图像本身具有信息冗余量大、像素数据量大等特点,基于 bit 位的 Hash 算法效率较低。因此,本文提出了基于块运算的图像 Hash 算法,块大小可根据具体的需求相应调整,从而大幅度提高算法执行效率。

算法思想可以描述为:多一个数字图像以一个像素点为最小单位,构成一个二维矩阵,使用 Logistic 映射将每一个像素点通过混沌序列进行离散化,离散规则即序列混沌序列矩阵用 D 表示,此外用等大的矩阵 R 来描述横向点离散化过程中每个点的离散距离,矩阵 C 描述纵向点离散化过程中每个点的离散距离。通过 1 次使用 Logistic 混沌序列将图像矩阵处理后,图像便实现了置乱效果。算法具体过程如下。

对任意数字图像 I,用矩阵表示其所占存储空间为 $m*n*unit$,其中 $unit$ 因数字图像的格式不同而异,如在 bitmap 位图图像使用 RGB 色彩编码时,则 $unit$ 代表 3 byte,使用灰度色彩编码时,$unit$ 代表 1 byte。本文中 $unit$ 代表 1 byte,即数字图像是灰度色彩编码图像。在矩阵 $S(m\times n)$中,图像中的任意一个像素点位置可以用(i,j)表示,其中 $i\in\{0,1,\cdots,m-1\},j\in\{0,1,\cdots,n-1\}$。采用块长 $l=8$ bit 位,则图像宽度位长可表示为 $w=\lceil\log_2(m*unit/l)\rceil$,图像长度位长可表示为 $s=\lceil\log_2(n*unit/l)\rceil$。加密算法步骤如下。

1)读取数字图像构造数字矩阵 $S(m\times n)$,确定混沌序列初始值 x_0 与分岔参数 μ,并确定混沌序列的长度 $K=m*n$。

2)以 x_0 为混沌序列的起始值,通过(1)式的 Logistic 映射迭代获得 K 个混沌序列值 $X=\{x_0,x_1,\cdots,x_{mn-1}\}$。

将 X 中的混沌值进行整数化,生成图像置乱参考矩阵 D,本文所选取的块长 $unit=8$ bit 位,因此,取整算法可选为:$X_i=\lfloor(x_k+1)*255/2\rfloor,k=(1,2,\cdots,m*n)$。

对图像 S 进行混沌序列值变换操作得到最终置乱图像:将置乱参考矩阵 D 中的每一个像素点$D(i,j)$到 S 矩阵相对应的位(i,j)进行异或运算,生成置乱图像为 S'。

通过使用 Logisit 映射生成混沌序列值,将数字图像按块进行分割转换成像素矩阵,与置乱参考矩阵进

行变换操作，最终生成置乱图像，增加了图像的隐蔽性。流程图如图 3 所示。同时由于混沌系统对初始值具有强敏感性，在攻击者不知道 x_0 与 μ 的情况下，Hash 序列值难以构造，因此，算法具有较好的安全性能。

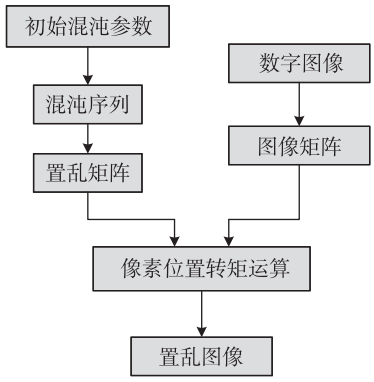

图 3　图像置乱算法流程图

2.3 混沌调制与判别量化

将置乱图像 $S'(i,j)$ 与原图像 $S(i,j)$ 的差值记为 $E(i,j)$，则有 $E(i,j)=S(i,j)-S'(i,j)$。用 Logistic 映射公式生成 $N*N$ 个混沌序列 $\{M|M_i\in(0,1),i=1,2,\ldots,N*N\}$，将其随机分为 N 组，其中第 j 组记为 $G_j=\{M_{i_1},M_{i_2},\cdots,M_{i_N}\}$，$j\in\{1,2,\cdots,N\}$，$i_1,i_2,\cdots,i_N\in\{0,1,2,\cdots,N*N\}$，则调制矩阵可以表示为 $G=[G_1,G_2,\cdots,G_N]^{\mathrm{T}}$，可以得到调制后的矩阵为

$$EG=E\times G \tag{2}$$

对调制后的矩阵 EG 进行按元素求和，计算可由(3)式表示为

$$\mathrm{sum}(EG)=\sum_{i=0}^{N-1}\sum_{j=0}^{N-1}EG(i,j) \tag{3}$$

Hash 序列 h 可以根据矩阵 EG 与参考量化阈值 t 进行对比得到，关系式为

$$h=\begin{cases}0,\mathrm{sum}(EG)\leqslant t\\1,\mathrm{sum}(EG)>t\end{cases} \tag{4}$$

经过多次的重复运算，可以得到最终的 Hash 序列。Hash 序列的长度只与算法的运算次数有关，长度越长，Hash 精度越高，与不同的数字图像产生的 Hash 发生碰撞概率就越小。

3 仿真实验

仿真实验采用 MATLAB 工具实现，以标准灰度图像作为输入仿真图像，将基于混沌理论的图像 Hash 算法与基于 DCT 的图像 Hash 算法进行对比；将算法重复运算迭代 64 次，产生 64 位的长度的 Hash 值，混沌系统参数初始值定为 $x_0=0.72$，$\mu=3.76$。图像认证时需要比较认证图像的 Hash 序列值与原始图像的 Hash 值，如果之间的差值小于认证阈值 t_s，则通过认证，否则拒绝。本仿真实验当中，将认证阈值定为 $t_s=0.1$。在信息论中，将认证图像的 Hash 值与原始图像的 Hash 值之间的差值用标准汉明距离(Normalized Hamming distance)来表示，标准汉明距离定义为 $h(x,y)=\dfrac{\sum x_i\oplus y_i}{n}$，式中 x、y 表示输入，$\oplus$ 为异或操作。图 4 是使用两种算法对两张不同数字图像基于混沌理论的 Hash 算法对图像进行认证的仿真实验图。图 5 是两种算法对于同一张数字图像采用算法的仿真图。

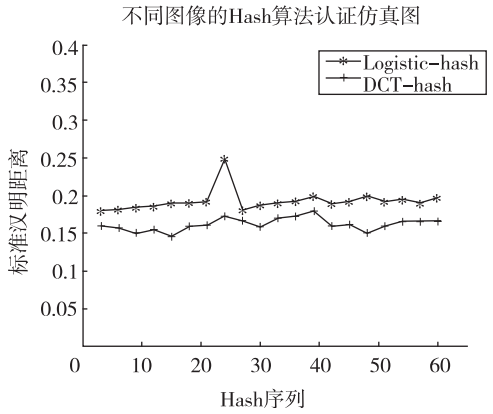

图 4　非同一张图像的 Hash 算法仿真结果图

由图 4 可以明显观察得出，两张图像 Hash 值的差值的浮动范围较大，即两张图像的标准汉明距离差别明显，图像认证失败。图 5 是对经过放大后的一张图像与原图像进行认证的仿真图，可以看出，两图的标准汉明差别均维持在较低水平，基本相同，即认证成功。

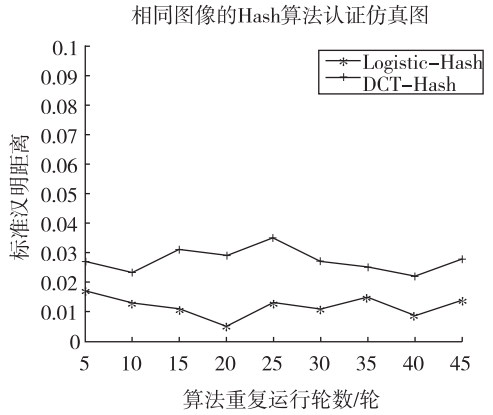

图 5　同一张图像的 Hash 算法仿真结果图

通过仿真实验可以得出，本算法对于图像认证具有较好的性能，同时支持认证经过缩放后的数字图像，在图像的缩放上展现出了良好的性能，同时具有一定的鲁棒性。

4 总结

本文结合混沌理论的 Logistic 映射方法，提出了基于混沌理论的对数字图像进行认证的 Hash 算法。引入了块模式思想将图像单元矩阵化，构造 Logistic 映射置乱参考矩阵，将图像矩阵进行置乱。该方法的最大优点是不仅提高了图像的安全性，而且提高了算法的执行效率，采用调制和量化最后生成图像 Hash 值。相对于基于 DCT 的图像可视化 Hash 函数构造方法而言，由于混沌系统的特点是具有初值敏感性和无法用后值反推回前面的值，因此，算法自身具有较高的安全性能。同时，由实验结果可得出，基于混沌理论的图像 Hash 算法具有更好的性能。算法对图像的缩放具有较好的性能，能够有效认证经过缩放操作的图像。

参考文献：

[1] 胡春强，邓绍江，张宇，等. 一种基于 Arnold 变换的图像 Hash 算法[J]. 世界科技研究与发展，2010，32(2)：138-140.
Hu Q Q，Deng S J，Zhang Y et al. Image Hashing algorithm Based on Arnold transformation[J]. World SCI-Tech R&D，2010，32(2)：138-140.

[2] Venkatesan R，Koon S M，Jakubowski M H，et al. Roubust image Hashing[C]. IEEE international conference on image processing. Vancouver，Cannda，2000(3)：664-666.

[3] 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.

[4] 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。

[5] 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.

[6] 秦川，王朔中，张新鹏. 一种基于视觉特性的图像摘要算法[J]. 中国图像图形学报，2006，11(11)：1678-1681.
Qin C，Wang S Z，Zhang X P. Image Hashing based on human visual system[J]，Journal of Image and Graphics，2006，11(11)：1678-1681.

[7] 叶卫国，韩水华. 基于内容的图像 Hash 算法及其性能评估[J]. 东南大学学报，2007，37(1)：109-113.
Ye W，Han S，Performance evaluation for content-based image authentication[J]. Journal of Southeast University，2007，37(1)：109-113.

[8] 颜世银，钱海峰，李志斌. 基于混沌系统的对称图像加密方案[J]. 计算机工程，2008，34(14)：155-160.
Yan S Y，Qian H F，Li Z B. Symmetric image Encryption scheme based on Chaotic system[J]. Computer Engineering，2008，34(14)：155-160.

[9] 邓绍江，张岱固，濮忠良. 一种基于混沌的图像置乱算法[J]. 计算机科学，2008，35(8)：238-240.
Deng S J，Zhang D G，Pu Z L. Digital image scrambling algorithm based on Chaotic system[J]. Computer Science，2008，35(8)：238- 240.

An Image Hash Algorithm Based on Chaos Theory

YU Hai-peng，WEN Zheng-ying

(Dept. of Computer Science and Engineering，Henan Institute of Engineering，Zhengzhou 451191，China)

Abstract：Hash algorithm for digital image authentication based on chaos theory. For beginning，the paper proposed the idea of constructing the image matrix from the quantifying image blocks，and used the chaos theory of logistic mapping method to make a scrambling matrix. Then，the differential matrix can gain from the scrambling matrix and image matrix. A modulated matrix generated by *N* turbidity modulator and one bit Hash sequence received from the binarization of the modulated matrix. Then an image sequence Hash obtained through multiple modulation and quantization. Finally，the algorithm is validated by simulation experiment，and the results show that the algorithm was effectively and powerful，it' s also certificated for the image scaling.

Key words：image authentication；chaos theory；logistic mapping；image Hash

(责任编辑　游中胜)