

基于人类视觉系统的图像感知哈希算法

张 慧¹, 张海滨¹, 李 琼², 牛夏牧²

(1. 哈尔滨工业大学深圳研究生院信息安全研究中心, 广州深圳 518055;
2. 哈尔滨工业大学信息对抗研究所, 黑龙江哈尔滨 150001)

摘 要: 图像感知哈希 (Perceptual Hashing) 是一门新兴技术, 它通过对图像感知信息的简短摘要和基于摘要的匹配, 来支持图像的认证和识别, 具有广泛的应用前景。目前关于图像感知哈希的研究主要集中在图像特征的提取上, 但是特征的选择缺乏对人眼视觉特性的考虑。本文从不同的侧面提出几种基于人类视觉系统的图像感知哈希算法。通过这几种算法之间和已有传统算法之间的测试比较, 结果表明考虑了人眼视觉特性的图像感知哈希算法在鲁棒性和区分性上能够得到提高, 算法给出的感知距离度量更符合人的主观感受。

关键词: 人类视觉系统 (HVS); 感知哈希; 图像认证; 图像质量评价

中图分类号: TN911.73 **文献标识码:** A **文章编号:** 0372-2112 (2008) 12A-030-05

Image Perceptual Hashing Based on Human Visual System

ZHANG Hui¹, ZHANG Hai-bin¹, LI Qiong², NIU Xia-mu²

(1. Information Security Technology Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, Guangdong 518055, China;
2. Information Countermeasure Technology Institute, Harbin Institute of Technology, Harbin, Heilongjiang 150001, China)

Abstract: As an emerging technology, Image Perceptual Hashing (IPH) is becoming a new hotspot and have broad potential applications. Through extracting the digest of perceptual information of an image and matching based on the digest, IPH supports the identification and authentication of images. Currently, the IPH algorithms in the literature are mainly focused on the image feature extraction, but they do not introduce sufficient perceptual factors. In this paper, several new IPH algorithms based on HVS are proposed. Experiments test and compare the proposed algorithms. The results suggest our methods take more perception information into account and have better performance on robustness and discriminability.

Key words: human visual system (HVS); perceptual hashing; image authentication; image quality assessment

1 引言

感知哈希通过对媒体感知内容和信息的提取, 压缩形成一个简短的摘要, 用于记录或者标识媒体内容, 或者更进一步用于发现媒体内容遭受的篡改和/或篡改的位置。由于其广泛的应用前景, 感知哈希正吸引着越来越多的研究者。虽然它与传统密码学意义上的哈希函数 (Cryptographic Hash Function) 有很多相似之处, 比如单向性, 抗碰撞性等等, 但它们之间存在着本质的区别:

(1) 数据源: 传统密码学的哈希函数以信号的二进制数据作为输入; 而感知哈希关注的是人对信号的内容感知。

(2) 鲁棒性: 传统密码学的哈希函数的单向性和抗碰撞性要求它对输入的二进制数据比特变化敏感, 也就是说, 即使输入数据的一位的比特变化, 也会导致输出

产生明显的随机变化; 而对于感知哈希, 只要媒体的感知内容不变, 不论其存储格式如何, 或者曾经遭受什么内容保持操作, 其哈希序列都应该保持不变或者在一个指定的阈值之内变化。

图像感知哈希是感知哈希的一大分支, 其特征提取和编码阶段需要包含更多的视觉感知信息, 以满足其鲁棒性要求。但是当前的图像感知哈希算法研究缺乏对人类视觉特性的考虑, 关注的特征还停留在图像灰度, 特征点等一般的图像特征上。比如, Fridrich 利用图像 DCT 低频系数的不变性来获得图像信息的稳定编码^[1]; 而 Vishal Monga 等人提出的一组图像感知哈希算法分别利用了图像的特征点 (feature points) 不变性^[2], 图像线性操作的矩阵不变性 (matrix invariants)^[3] 以及图像小波分解得到的低通部分的灰度平均值^[4] 等。这些特征信息一方面反映了图像的感知特征, 另一方面也使得基于它们的

收稿日期: 2008-08-25; 修回日期: 2008-12-13

基金项目: 国家自然科学基金 (No. 60832010, No. 60671064, No. 60703011); 国家 863 计划 (No. 2007AA01Z458); 高等学校博士学科专项科研基金 (No. 20070213047)

算法在不同程度上对各种内容保持操作和变换鲁棒。但是,这些图像感知哈希算法并没有对人类对图像的感知过程以及人类视觉的感知特性做专门的研究。

图像的认知过程包含两个阶段,首先是人的视觉系统(包括眼睛、视觉传导神经和大脑视皮层区域)对光信号输入的响应过程,然后是大脑对接收到的信号的识别、判断、联想和记忆等再处理的过程。而关于人类视觉系统(HVS)的研究成果一直以来主要应用在图像质量评价(Image Quality Assessment)上^[5,6]。研究者们已经可以将人眼观察事物的过程,抽象为若干视觉感知特性,并从对比度敏感函数(CSF)、多通道分解、亮度适应性、掩蔽效应和误差合并等多个方面对其进行数学建模^[7,8]。引入人类视觉系统(HVS)后的图像质量客观评价更符合人的主观感受。因此,借鉴人类视觉系统既有的研究成果,探讨如何在图像感知哈希中融入更多的感知因素,必然对感知哈希算法的性能和应用起到重要的促进作用。

2 基于人类视觉系统的图像感知哈希算法

本节在简述 Watson 模型的基础上,提出三种基于人类视觉系统的图像感知哈希算法,从不同的侧面说明人类视觉系统用于图像感知哈希的方法和作用。

2.1 Watson 模型用于感知距离度量

Watson 模型是一个经典的综合对比敏感度、掩蔽效应和误差合并的人类视觉模型。它定义了一个对比敏感表,表中的每一个数字代表在不存在掩蔽的情况下可被感知的最小 DCT 系数,并将这个系数视为一个 JND(Just Noticeable Difference)。对比敏感表中的数值越小,表示人眼对该数值位置的相应频率越敏感。基于对比敏感表,Watson 给出了计算亮度掩蔽阈值和对比度掩蔽阈值的方法。其中对比度掩蔽阈值表达了每块 DCT 系数中各项在一个 JND 范围内可调的大小^[7]。Watson 模型用于感知距离度量的方法可以简述为如下步骤:

(1) 计算原图与测试图的 DCT 系数差,求得 DCT 系数的错误图像 E 。

(2) 将 E 除以对比度掩蔽阈值,得到每一个像素的感知误差 d_{ij} 。 d_{ij} 是 JND 的倍数或者分数。

(3) 用明科夫斯基和对误差进行合并:

$$D(I_o, I_t) = \left| \left| d_{ij} \right| \right|_{ij}^{\nu}$$

基于 Watson 模型的各种应用表明, Watson 模型能够很好的概括人眼的视觉特性,图像的感知距离度量符合人类对感知距离的主观度量。

2.2 简化的 Watson 模型作为感知哈希算法

Watson 模型直接用于感知距离度量具有很好的效果,而且考虑了人类视觉的感知特性。但是把 Watson 模

型直接作为感知哈希算法需要保存原图的分块 DCT 系数和对比度掩蔽阈值矩阵作为哈希序列,这显然不符合感知哈希的摘要性要求。本文在保留 Watson 模型原有优势的同时,基于简化的 Watson 模型的感知哈希算法(为了便于叙述,标记为 Watson-Predigest)提出一种满足感知哈希摘要性的方法。具体方法简述如下:

(1) 将图像大小归一化为 512×512 ;

(2) 将归一化图像分成若干随机的分块, I_1, I_2, \dots, I_N , 其中, N 是最后的哈希序列长度,每个分块的尺寸应该为 8×8 的整数倍;

(3) 将图像 I 做分块 DCT 变换,求对比度掩蔽矩阵,得到 $DctI$ 和 $DctCT$;

(4) 按照第 2 步中的分块,求 I_1, I_2, \dots, I_N 中每一个分块中所包含的所有 DC 分量和 DC 分量对应的对比度掩蔽值的平均值,得到 DC 分量序列 $DC_1, DC_2 \dots DC_N$ 和对比度掩蔽序列 $CT_1, CT_2 \dots CT_N$ 。

Watson-Predigest 方法将 DC 分量序列和相应的对比度掩蔽序列作为最后的哈希序列保存。哈希序列为长度为 $2N$ 的实向量。在本文的实验中, $N = 256$ 。

为了获得应用于认证时的安全性, Watson-Predigest 方法采用 Venkatesan 在文献[4]中提出的方法来产生随机分块。具体过程可见文献[4]中的描述。Watson-Predigest 方法与 Watson 模型一样使用明科夫斯基和的误差合并来计算感知距离,只是合并的对象 d_{ij} 变成了经过简化的感知误差:

$$d_i = \left(\overline{DC_i(org)} - \overline{DC_i(dst)} \right) / \overline{DC_i(org)}$$

其中, org 表示原图, dst 表示受攻击的测试图, d_i 是在每个分块的感知误差。

2.3 对比度掩蔽阈值作为加权的感知哈希算法

Watson 模型的对比敏感表描述了人眼对频率的敏感程度,基于此得到的对比度掩蔽阈值矩阵表示 DCT 系数在一个 JND 范围内可调整的大小。使用对比度掩蔽阈值来加权人眼敏感的频率系数,使得人眼敏感的视觉内容在感知哈希编码中占有更大的比重,得到的哈希序列必然能够反映更多的人眼视觉特性。基于这个思想,本文提出使用对比度掩蔽阈值作为加权的感知哈希算法(标记为 Watson-CTWeighted)。该算法使用 Fridrich 在文献[1]提出的随机序列构造方法,具体过程简述如下:

(1) 将图像大小归一化为 512×512 ;

(2) 将图像 I 做分块 DCT 变换,求对比度掩蔽矩阵,得到 $DctI$ 和 $DctCT$;

(3) 利用密钥 K 产生与原图大小相同的 N 个伪随机矩阵 $P_n(i, j)$ 。其中, $1 \leq n \leq N, 1 \leq i \leq 512, 1 \leq j \leq 512$, N 为最后的哈希序列长度,每个伪随机矩阵均服从标

准正态分布. 在本文的实验中, $N = 256$;

(4) 计算 $Y_n = \sum_{i=1}^{512} \sum_{j=1}^{512} [\text{DetI}(i, j) \times \text{DetCT}(i, j) \times P_n(i, j)]$;

(5) 以 0 为比较阈值二值化 Y_n 得到哈希序列 H_n . 即将 Y_n 中大于等于 0 的元素置 1, 其余置 0.

Watson-CTWeighted 算法使用汉明距计算感知距离.

2.4 基于多通道分解和误差合并的方法

经典的人类视觉系统通过多通道分解来从多个角度反映图像的视觉特征, 并且通过误差合并的方法来综合各个通道的评价结果. 而当前的感知哈希算法往往都基于对某一种图像特征的提取, 由于所选择的特征的局限性, 算法往往对不同的攻击表现出性能上的偏差. 比如文献[3]的方法只能抵抗 5 度的旋转攻击, 而文献[4]可以达到 20 度; 相反, 对于高斯噪声, 文献[3]的性能却更好一些.

表 1 各种算法在各种攻击类型的极端强度下的感知距离

攻击方式	攻击强度	Watson-Original	Venkatesan	Watson-Predigest	Watson-CTWeighted	MinkowskiSum
JPEG	质量因子:1	0.022	0.057	0.058	0.097	0.053
高斯噪声	均值:0.3	0.796	0.211	0.466	0.445	0.190
中值滤波	10 次	0.041	0.024	0.010	0.054	0.029
模糊	10 次	0.118	0.035	0.028	0.132	0.038
旋转和剪切	10 度	0.660	0.148	0.211	0.332	0.136
剪切	放大 1.2 倍	0.616	0.180	0.238	0.319	0.158
马赛克	10 像素窗口	0.200	0.022	0.011	0.113	0.023

3 实验结果及分析

为了验证人类视觉系统用于图像感知哈希时的效果, 本文使用文献[9]数据库提供的 82 张图像, 将原始的 Watson 模型也视作一个感知距离度量方式 (标识为 Watson-Original), 与本文提出的 Watson-Predigest、Watson-CTWeighted、MinkowskiSum 方法以及不考虑人类视觉特性的 Venkatesan 的方法作一个统一的比较.

3.1 不同算法结果的归一化

不同的感知哈希算法可能使用不同的哈希序列形式, 而不同的哈希序列数据格式又要求不同的距离度量方法. 最常见的有二进制向量和汉明距, 实数向量和欧氏距离等. 目前在感知哈希算法的测试和比较中, 还缺乏一个能够对采用不同编码方式的感知哈希算法进行比较的方法.

本文提出一个感知哈希感知距离的归一化方法. 首先, 执行组间测试, 其分布直方图的峰值位置对应的感知距离 d_{diff} 代表比较两幅感知内容完全不同的图像时所得到的感知距离的期望. 然后计算两幅完全相同的图像的感知距离 d_{same} . 最后, 将不同算法得到的区间

本文将基于不同图像特征的感知哈希算法视作不同的感知通道, 使用改进的明科斯基和公式来合并它们的结果, 提出基于多通道分解和误差合并的方法 (标识为 MinkowskiSum). 具体步骤如下:

(1) 归一化每个通道中不同的感知哈希算法得到的感知距离;

(2) 使用改进的明科斯基和合并结果:

$$PD = \frac{1}{\sum_i |1/d_i|}^{1/\nu}$$

其中, PD 为求和后的感知距离, d_i 为每个通道的感知距离, $\nu = 4$ 为求和指数.

MinkowskiSum 方法可以集合不同算法的在鲁棒性上的优势的同时, 不降低原有算法的区分性. 但是, 该方法在空间和时间上显然要比原有算法开销更大.

作为实例, 本文合并了前文提到的文献[3]和文献[4]的两个算法.

$[d_{same}, d_{diff}]$ 归一化到相同的区间 $[0, 0.5]$ 上.

采用上述方法, 本文实验中包含的几个不同算法的测试结果得以在一个统一的标准下进行比较.

3.2 算法性能的比较和分析

首先通过组间测试的感知距离的直方图分布粗略观察算法的区分性. 图 1 例示了 Watson-Original 的组间测试直方图分布. 该分布的特点一方面说明算法具备潜在的好的区分性, 另一方面给出了组内测试结果所应该具备的极端情况. 上述被测试的算法在组间测试

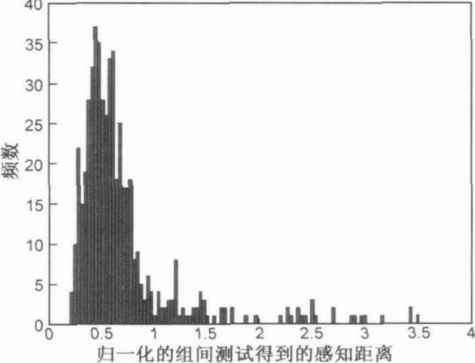


图 1 Watson模型的组间测试直方图分布

的感知距离的直方图分布上具有相似的特征,因此,上述算法性能的区别主要反映在组内测试的结果上。

针对每种攻击,当攻击强度增大时,算法给出的感知距离应该相应单调增大,当攻击强度增大到使图像感知质量已经超出人的接受程度时,相应的感知距离应该逼近或超过 0.5。表 1 给出了各种算法在各种攻击的攻击强度最大或图像质量不可接受的情况下的感知距离取值。可以看出,对于 JPEG、中值滤波、马赛克、模糊等攻击,因为即使它们的攻击强度相当大时,攻击得到的图像质量还是很好的,所以各种算法都给出的距离度量也都很小;旋转和剪切是两种几何攻击,因为所有算法都基于对图像分块的处理,所以它们都在攻击强度不大时就失去了合理性。

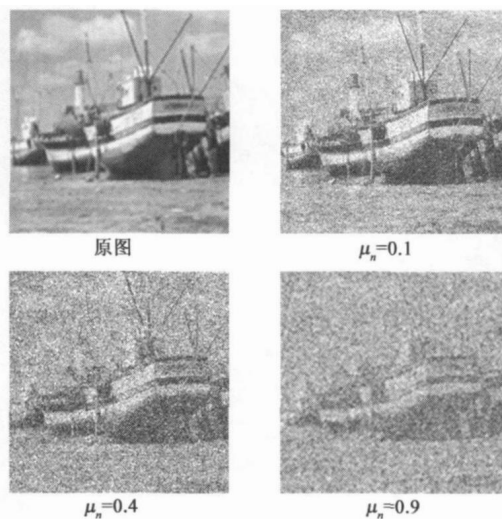


图2 不同噪声攻击强度下的图像质量

图 2 给出一个样本图像在受到不同高斯噪声攻击后的实例。通过观察它们原始尺寸的图像可知,在噪声均值大于 0.4 时,图像原始信息被噪声淹没,变得不可辨认。图 3 使用感知距离与攻击强度的关系曲线更加直观的比较了各种算法在高斯噪声下的整体性能。其中, Watson-Predigest 和 Watson-CTWeighted 在高斯噪声的均

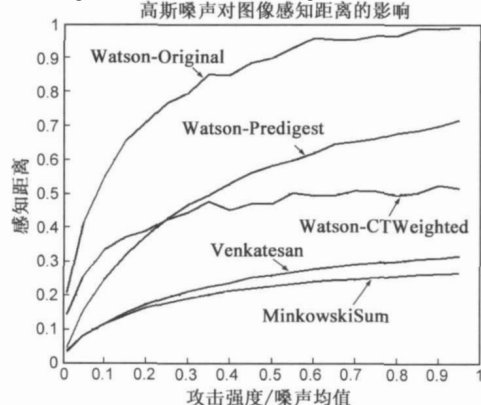


图3 各方法在高斯噪声攻击下感知距离的比较

值在 0.4 左右时的归一化感知距离接近于 0.5,这比较符合人眼的视觉特性;而 Watson-Original 在均值为 0.1 时即超过 0.5;Venkatesan 甚至在攻击强度最大时感知距离仍小于 0.5。可见, Watson-Predigest 和 Watson-CTWeighted 方法更符合人眼的主观判断。而 Watson-Original 相对于人的感受而言显得太严格;其它几种方法则显得太宽松。 Watson-Predigest 和 Watson-CTWeighted 相对于 Watson-Original 的性能提升均得益于算法对低通信息的保留。而 MinkowskiSum 则由于是由 Venkatesan 的方法改进而来,受限于原始算法的局限。

图 4 说明 MinkowskiSum 方法可以综合两个感知特征通道的优势。其中 SVD 是对文献[3]的仿真, Wavelet 是对文献[4]的仿真,合并结果在各个阶段选择了性能较好的通道,从而更加符合人类的视觉感受。

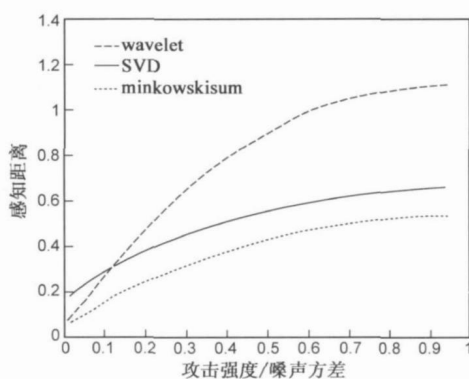


图4 MinkowskiSum方法对两个算法的合并

综合算法性能测试的结果可以看出, Watson-Predigest 和 Watson-CTWeighted 方法由于在编码中考虑了更多的人类视觉的因素,其结果比仅利用图像灰度特征的 Venkatesan 的方法更加符合人眼对图像质量的观感; MinkowskiSum 方法可以综合不同通道的优点,而这也是得益于对人类视觉系统多通道性的理解。

4 结论

图像感知哈希的研究从最初脱胎于数字水印和基于内容的图像检索 (CBIR) 的发展,到现在成为一个独立的方向,已经经历了一个逐渐深入的研究历程。但是直到目前,图像感知哈希的研究却很少在人类视觉系统的框架内得到讨论。本文从人类视觉特性的不同角度,提出三个基于人类视觉系统的感知哈希算法,并测试证明这些算法能够更好的概括人类视觉特性,其感知距离的度量结果更加符合人类对图像的主观感受。

参考文献:

- [1] J. Fridrich, M. Goljan. Robust hash functions for digital watermarking [A]. Proceedings. International Conference on Information Technology: Coding and Computing [C]. Las Vegas:

- IEEE, 2000. 178 - 183.
- [2] V Monga, B L Evans. Robust perceptual image hashing using feature points [A]. Proceedings of IEEE International Conference on Image Processing (ICIP) [C]. Singapore: IEEE, 2004. 1: 677 - 680.
- [3] S S Kozat, R Venkatesan, M K Mihcak. Robust perceptual image hashing via matrix invariants [A]. Proceedings of International Conference on Image Processing (ICIP) [C]. Singapore: IEEE, 2004. 5: 3443 - 3446.
- [4] R Venkatesan, S M Koon, et al. Robust image hashing. Proceedings of IEEE International Conference of Image Processing (ICIP) [C]. Vancouver: IEEE, 2000. 3: 664 - 666.
- [5] A De Angelis, A Moschitta, F Russo, P Carbone. Image quality assessment: an overview and some metrological considerations [A]. International Workshop on Advanced Methods for Uncertainty Estimation in Measurement [C]. Trento: IEEE, 2007. 47 - 52.
- [6] H R Sheikh, A C Bovik. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15(2): 430 - 444.
- [7] Andrew B Watson. DCT quantization matrices visually optimized for individual images [A]. The International Society for Optical Engineering [C]. California: SPIE, 1993. 1913: 202 - 216.
- [8] Wang Zhou, A C Bovik, H R Sheikh, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600 - 612.
- [9] Allan G Weber. The USC-SIPI Image Database: Version 5 [DB/OL]. <http://sipi.usc.edu/database>

作者简介:



张 慧 男, 1978 年生于湖南沅陵. 哈尔滨工业大学计算机科学与技术学院博士生. 研究方向为多媒体认证、感知哈希、文档安全.

E-mail: zhanghui1978@gmail.com



张海滨 男, 1983 年出生于黑龙江哈尔滨. 哈尔滨工业大学深圳研究生院, 硕士研究生, 研究方向为多媒体内容特征提取与安全、图像的质量评价.