

用于图像 Hash 的视觉相似度客观评价测度

唐振军 王朔中 魏为民 苏胜君

(上海大学通信与信息工程学院, 上海 200072)

摘要 由于评价图像 Hash 性能时,要求对两幅图像是否在感知上相似做出判断,因此针对这一需求,提出了一种衡量感知相似程度的评价测度。该测度的确定是先对图像进行低通滤波,再进行图像重叠分块;然后运用相关系数检测法计算每一对分块的相似程度,并对相似系数归一化,再分别计算若干个最小和最大的归一化相似系数的乘积;最后用最小相似系数乘积与最大相似系数乘积的比值作为感知相似性的测度。实验结果表明,该测度不仅可有效反映图像视觉质量的变化,而且能较好地地区分两幅图像是否存在重要的视觉差异,其对感知相似进行评价的性能优于峰值信噪比。

关键词 感知相似 客观评价 图像 Hash 图像摘要 图像索引

中图分类号:TP391.4 文献标识码:A 文章编号:1006-8961(2008)10-2039-04

Perceptual Similarity Metric for Application to Robust Image Hashing

TANG Zhen-jun, WANG Shuo-zhong, WEI Wei-min, SU Sheng-jun

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

Abstract To measure perceptual similarity between an original image and its modified version, we propose an objective measure that is relatively stable to normal image processing but quite sensitive to significant changes of the image content in local areas. This is achieved by low-pass filtering the two images, dividing them into overlapping blocks and determining similarity between the corresponding blocks in terms of correlation coefficient that is mapped to the interval $[0, 1]$. Based on previous calculated correlation coefficients, a ratio is calculated between the smallest and largest correlation coefficients and defined as the perceptual similarity. Experimental results show that the proposed metric is not substantially affected by normal image processing. It provides indication of changes in the image contents when its value becomes less than a given threshold. The proposed metric is useful in applications such as image hashing and CBIR.

Keywords perceptual similarity, objective measure, image hashing, image digest, image indexing

1 引言

图像 Hash 是根据图像的内容特征进行单向映射而提取出来的短小数字序列,可用来标识该图像,已广泛应用于内容认证、图像标注、基于内容的图像检索(CBIR)等领域。与数字水印不同,由于提取 Hash 并不对图像数据进行修改,因此能完好地保持原图像的品质。由于密码学中常用的 Hash 函数,如 SHA-1 和 MD5 对输入数据的变化非常敏感,输入数据任何

1bit 的改变都会完全改变输出的 Hash 值,因此不适用于图像认证。实际应用中,往往要对图像进行正常的数字处理,如增强、JPEG 压缩等,由于其内容并未发生实质性改变,因此希望图像 Hash 保持不变。图像 Hash 通常应该满足以下 3 个条件:(1)感知鲁棒性,即对于感知相似的两幅图像,不管其内部数据是否一致,它们的 Hash 应以很大的概率相同或十分接近;(2)对内容修改敏感,即当图像内容发生重要改变时,其对应的 Hash 应出现重大变化;(3)对密钥依赖性,即图像 Hash 应在密钥控制下生成,在无法获取密

基金项目:国家自然科学基金项目(60502039,60773079);国家高技术研究发展计划(863)项目(2007AA01Z477)

收稿日期:2008-06-20; **改回日期:**2008-07-08

第一作者简介:唐振军(1979~),男。2006年获广西师范大学计算机软件与理论专业硕士学位。现为上海大学博士研究生。研究方向为图像处理、多媒体信息安全。E-mail:tangzj230@163.com

钥的前提下,成功伪造图像 Hash 的概率接近于 0。后两种性质可统称为图像 Hash 的安全性。

近年来图像 Hash 引起了研究者的广泛关注,并提出了一系列有效的技术方案。其中代表性的方法包括:利用小波系数统计量^[1]、选择 DCT 系数矩阵^[2]、寻找不变关系^[3]、Radon 变换分解图像^[4,5]、基于特征点和规范化的三角形网格^[6]、提取 Fourier-Mellin 变换系数^[7]、运用矩阵分解技术^[8,9]等。在图像 Hash 研究中,需要对不同图像的感知相似性进行评定,例如一幅图像曾被篡改,它与原始图像之间就会存在不可接受的内容差异。对图像内容差异进行评价的方法通常分为主观和客观两类。其中主观评价需要人的参与,但局限于定性分析,因此判断结果可能因人而异;而客观评价则不需要人的参与,并可得到定量结果,而且性能优良的客观评价测度更适合于图像 Hash 研究。

峰值信噪比(PSNR)是进行图像客观评价常用而简单的评价指标,其反映了图像的峰值总能量与两幅图像对应像素的均方差之比。PSNR 值越大,表示两幅图像的总体差异越小。但它只代表两幅图像差异的全局平均,而不能充分反映小面积的重大差异。例如对一幅图像分别进行轻度滤波和局部篡改,其与原图像相比,所得到的两幅图像的 PSNR 值则可能完全相同。因此,PSNR 值并不适用于图像 Hash 研究中的感知相似度评价。

本文提出了一种基于分块计算的感知相似的评价方案,由于其顾及了图像之间的局部差异,因此得到的客观评价测度能较好地反映 JPEG 压缩、滤波、噪声干扰、水印嵌入等操作对图像视觉质量的影响,同时由于其对图像中小范围的内容变化敏感,因此能较好地判断图像是否被篡改。此外,这种感知差异性测度也可以直接应用于鲁棒 Hash 的生成算法。

2 基于分块的图像相似性测度

将图像 I 经过不同处理后得到的一组图像记为集合 A ,其中由与图像 I 感知相似的图像所组成的集合记为 A_{ident} ,与图像 I 在视觉上有重要差异的图像的集合记为 A_{diff} 。考虑两幅待评价的图像 I_1 和 I_2 ,如果用 $P(\cdot)$ 表示事件发生的概率,则对于 $\varepsilon_1, \varepsilon_2, \varepsilon_3$ 和 t ($0 < \varepsilon_1, \varepsilon_2, \varepsilon_3 < 1, t > 0$),一个理想的图像感知相似评价函数 $F(I_1, I_2)$ 应该满足如下条件:

(1) 感知鲁棒:若 $I_j \in A_{\text{ident}}$,则 $P[F(I, I_j) \geq t] \geq$

$1 - \varepsilon_1$;

(2) 篡改敏感:若 $I_j \in A_{\text{diff}}$,则 $P[F(I, I_j) < t] \geq 1 - \varepsilon_2$;

(3) 反映视觉差异:若 I_k 和 I 的相似程度低于 I_j 和 I 的相似程度,则 $P[F(I, I_k) < F(I, I_j)] \geq 1 - \varepsilon_3$;

(4) 对称性:交换 I_1 和 I_2 在函数中的输入位置,输出的相似值不变,即 $F(I_1, I_2) = F(I_2, I_1)$ 。

计算相似度时,要先对图像进行低通滤波,以减轻图像的轻微改变对感知相似度的影响,如噪声干扰和滤波。通常,由于被篡改的内容主要集中在图像的某一局部区域,因此图像感知相似的评价指标必须具有分辨局部内容差异的能力。为此,要先对图像 I_1 和 I_2 进行分块,并计算对应块的相似程度,在此基础上,即可达到分辨图像局部内容的目的,其具体步骤如下。

首先将图像 I_1 分割成大小为 $2s \times 2s$ 的重叠块,相邻图像块之间的重叠为 s 。设一共分为 n 块,记第 i 个分块为 $B_i^{(1)}$,于是 $I_1 = B_1^{(1)} \cup B_2^{(1)} \cup \dots \cup B_n^{(1)}$ 。同理,将图像 I_2 进行同样的重叠分块,即可得到 $I_2 = B_1^{(2)} \cup B_2^{(2)} \cup \dots \cup B_n^{(2)}$ 。若某一图像块包含的篡改内容越多,则与对应块之间的差异就越大。之所以采用重叠分块是为了保证当篡改区域比分块区域小时,存在一个分块能包含篡改区域的主要内容。判断分块 $B_i^{(1)}$ 与 $B_i^{(2)}$ 的相似程度时,可采用以下 Pearson 相关系数检测法:

$$\rho_i = \frac{\frac{1}{MN-1} \sum_{j=1}^M \sum_{k=1}^N [B_i^{(1)}(j,k) - \mu_{B_i^{(1)}}][B_i^{(2)}(j,k) - \mu_{B_i^{(2)}}]}{\sigma_{B_i^{(1)}} \sigma_{B_i^{(2)}}} \quad (1)$$

其中, $\mu_{B_i^{(1)}}$ 和 $\mu_{B_i^{(2)}}$ 分别代表 $B_i^{(1)}$ 和 $B_i^{(2)}$ 的均值,分子是 $B_i^{(1)}$ 和 $B_i^{(2)}$ 的协方差,分母为 $B_i^{(1)}$ 和 $B_i^{(2)}$ 的标准差。考虑到分母不能为 0,如果两个标准差均等于 0,则令 ρ_i 为 1;如果只有一个为 0,则令 ρ_i 为 0。为了方便后续处理,可将 ρ_i 归一化为

$$r_i = \frac{\rho_i - \min(\rho_i)}{\max(\rho_i) - \min(\rho_i)} \quad (2)$$

其中, $\max(\rho_i)$ 和 $\min(\rho_i)$ 分别表示 ρ_i 的最大和最小取值,由相关系数的取值范围知, $\max(\rho_i) = 1$, $\min(\rho_i) = -1$ 。若 $B_i^{(1)}$ 和 $B_i^{(2)}$ 越相似,则 r_i 值将越大。于是可将图像感知相似度(perceptual similarity)定义为

$$S(I_1, I_2) = \frac{\prod_{r \in I_{\text{small}}} r}{\prod_{r \in I_{\text{large}}} r} \quad (3)$$

其中, Γ_{small} 是由 m 个值最小的 r 组成的集合, 而 Γ_{large} 则是由 m 个值最大的 r 组成的集合。分块间的差异越小, 分块就越相似, 两者的相关系数也就越大, 式(3)反映了待评价图像间最不相似的 m 对图像块与最相似的 m 对图像块间的关系。由于 Pearson 相关系数与图像块的输入顺序无关, 因此式(3)满足对称性条件。由于 r 的取值范围为 $0 \sim 1$, 因此可知, S 的取值范围亦为 $[0, 1]$ 。由于待评价图像越相似, 其 r 值间的差异将越小, 因此 S 取值越大; 当图像 I_1 和 I_2 完全一致时, $S = 1$ 。如果图像 I_2 是图像 I_1 的篡改版本, 那么至少会存在 1 个较小的 r , 其值与最大的 r 差异较大, 这样即可使得 S 值明显变小。

3 实验结果

为验证本文评价测度的效果, 用 StirMark 4.0^[10] 对大小为 512×512 的彩色标准测试图像 Baboon、House、Peppers 和 Lena 进行 JPEG 压缩、嵌入水印及噪声干扰攻击, 详细参数见表 1。同时还考虑了滤波处理, 例如对上述图像进行模板为 3×3 的高斯低通滤波, 标准差取值为 $0.3 \sim 1.0$ 。为了检验评价指标的篡改敏感性, 对 Baboon 的同一个区域进行 3 次篡改操作, 且篡改的面积逐步增大(如图 1 所示)。

表 1 图 2 和图 3 横坐标对应的图像处理

Tab. 1 Processing corresponding to indices in Figures 2 and 3

序号	代码	操作	说明	参数
1 ~ 10	I	JPEG 压缩	质量因子	10, 20, ..., 100
11 ~ 20	II	水印嵌入	强度	10, 20, ..., 100
21 ~ 28	III	3×3 高斯低通滤波	标准差	0.3, 0.4, ..., 1.0
29 ~ 31	IV	噪声干扰	级别	1, 2, 3
32 ~ 34	V	局部篡改	—	—

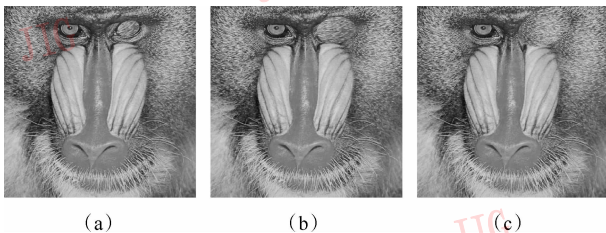


图 1 篡改面积不同的图像 Baboon
Fig. 1 Differently tampered Baboon

实验中采用 3×3 大小、标准差为 1 的高斯滤波器对输入图像进行低通滤波, 取相邻图像块的重叠大小 $s = 32, m = 2$ 。为了验证评价指标对图像感知

相似的鲁棒性能, 分别计算处理前后标准图像 Baboon、House、Peppers 和 Lena 的相似度值 S , 结果如图 2 所示, 纵坐标是 S 值, 横坐标代表不同的操作, 具体含义见表 1。横坐标 1 ~ 10 表明, 随着 JPEG 压缩质量因子的提高, 相似度不断增大; 与横坐标 11 ~ 20 对应的相似度 S 均接近于 1, 这说明 S 值对水印嵌入操作具有良好的鲁棒性; 横坐标 21 ~ 28 给出了 3×3 高斯低通滤波对相似度 S 的影响, 由于计算分块相似程度前, 图像已经过低通滤波预处理, 因此, 图像感知相似度对低通滤波操作较稳健, 即随着低通滤波器带宽的减小, S 值呈下降趋势, 其反映了图像的模糊程度。横坐标 29 ~ 31 对应噪声干扰时, 相似度 S 的变化情况, 即随着噪声级别的提高, 相似度 S 不断降低。实验中仅出现 3 个 S 值低于 0.9, 分别是 JPEG 压缩质量因子为 10 时, Baboon 和 House 图的 S 值, 以及噪声级别为 3 时, House 的 S 值, 其中, 最小的 S 值为 0.844, 其他情况下 S 值均大于 0.9。根据此实验结果, 可取感知相似阈值为 0.75, 如果两幅图像的 S 值大于 0.75, 则认为两者感知相似, 否则认为两者存在重要的视觉差异。

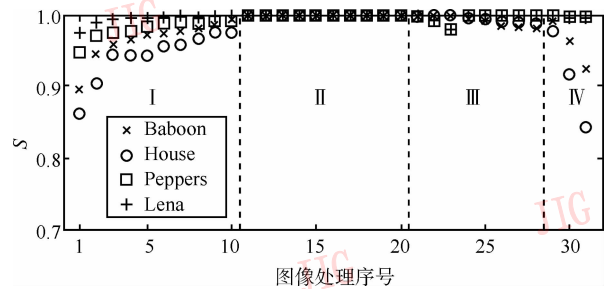


图 2 图像处理前后的感知相似度
Fig. 2 Perceptual similarity between original images and their processed or tampered versions. Indices in the abscissa represent different processing

第 2 组实验是分别计算 Baboon 图像处理前后的 PSNR 值和 S 值, 计算结果如图 3 所示, 其中图 3 (a) 是 PSNR 曲线, 图 3 (b) 是 S 曲线。对比发现, 在 S 曲线图中, 可找到一个阈值, 如 0.75, 可用其区分正常的数字处理与恶意篡改, 而 PSNR 值则无法找到这样的分割阈值。表 2 和表 3 分别列出了图 3 中几种代表性图像处理和局部篡改的例子。这一实验说明, 具有相同 PSNR 值的图像, 其内容可能相似, 也可能存在局部区域的重大差异, 因此 PSNR 并不适用于评价两幅图像是否感知相似。对于感知相似度 S 的大小, 当图像感知相似, 且没有局部重大

差异时,其取值较高;如果图像被恶意篡改,则取值较低,其感知评价性能明显优于 PSNR。

此外,计算得到的标准图像间的 S 值如表 4 所示。由表 4 可见,视觉内容完全不同的两幅图像的 S 值很低。总之,由于本文定义的感知相似度 S 对感知相似具有鲁棒性,不仅能区分内容不同的图像,而且能较好地分辨图像内容的局部篡改,因此可在图像 Hash 研究中用作图像感知相似的评价指标。

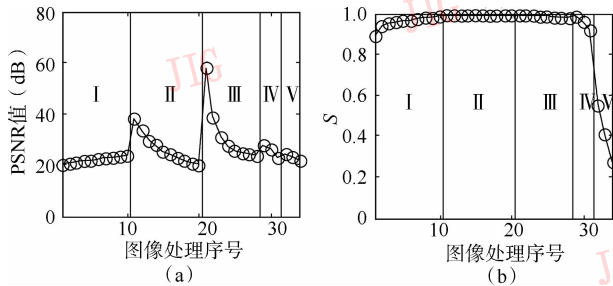


图 3 PSNR 值与 S 值的对比实验

Fig. 3 Comparison between PSNR and S . Indices in the abscissa represent different processing

表 2 处理后图像的 PSNR 值与 S 值的对比

Tab. 2 PSNR and S values before and after image processing

	高斯滤波 $\sigma = 0.9$	JPEG $Q = 50$	JPEG $Q = 90$
PSNR 值 (dB)	24.5	22.3	23.8
S	0.981	0.970	0.985

表 3 篡改后图像的 PSNR 值与 S 值的对比

Tab. 3 PSNR and S values before and after tampering

	图 1 (a)	图 1 (c)	图 1 (b)
PSNR 值 (dB)	24.5	22.3	23.6
S	0.558	0.273	0.415

表 4 不同图像的 S 值

Tab. 4 Perceptual similarity of different images

	Baboon	House	Lena	Peppers
Baboon	1.000	0.044	0.023	0.024
House	0.044	1.000	0.002	0.001
Lena	0.023	0.002	1.000	0.080
Peppers	0.024	0.001	0.080	1.000

4 结 论

本文提出了一个用于评价两幅图像感知相似度

的客观评价测度。该方法通过计算重叠分块的相似性不仅使评价指标能敏感地反映小面积图像块的变化,而且确保了对局部内容篡改的灵敏度。实验结果表明,本文提出的图像内容相似性测度能较好地区分两幅图像的视觉内容是否存在重要差异,因而可用于图像感知的相似性评价。该方法也可在图像 Hash 提取算法中得到应用。

本文方法由于尚未考虑几何变换的情况,因此该指标不适用于图像几何变换后的感知相似问题。今后在进行图像感知相似及相关图像 Hash 算法研究中,将进一步考虑几何变换下的图像感知相似的评价问题。

参考文献 (References)

- 1 Venkatesan R, Koon S-M, Jakubowski M H, *et al.* Robust image hashing [A]. In: Proceedings of IEEE International Conference on Image Processing [C], Vancouver, Canada, 2000: 664 ~ 666.
- 2 Fridrich J, Goljan M, Robust hash functions for digital watermarking [A]. In: Proceedings IEEE International Conference on Information Technology: Coding and Computing [C], Las Vegas, NA, USA, 2000: 178 ~ 183.
- 3 Lin C Y, Chang S F. A robust image authentication system distinguishing JPEG compression from malicious manipulation [J]. IEEE Transactions on Circuits System and Video Technology, 2001, 11(2): 153 ~ 168.
- 4 Lefebvre F, Macq B, Legat J D. RASH: Radon soft hash algorithm [A]. In: Proceedings of European Signal Processing Conference [C], Toulouse, France, 2002: 299 ~ 302.
- 5 Roover C De, Vleeschouwer C De, Lefebvre F, *et al.* Robust video hashing based on radial projections of key frames [J]. IEEE Transactions on Signal Processing, 2005, 53(10): 4020 ~ 4036.
- 6 Lu C-S, Hsu C-Y, Sun S-W, *et al.* Robust mesh-based hashing for copy detection and tracing of images [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo [C], Taipei, Taiwan, China, 2004: 731 ~ 734.
- 7 Swaminathan A, Mao Y, Wu M. Robust and secure image hashing [J]. IEEE Transactions on Information Forensics and Security, 2006, 1(2): 215 ~ 230.
- 8 Kozat S S, Mihcak K, Venkatesan R. Robust perceptual image hashing via matrix invariants [A]. In: Proceedings of IEEE Conference on Image Processing [C], Singapore, 2004: 3443 ~ 3446.
- 9 Monga V, Mihcak M K. Robust and secure image hashing via non-negative matrix factorizations [J]. IEEE Transactions on Information Forensics and Security, 2007, 2(3): 376 ~ 390.
- 10 Petitcolas F A P. Watermarking schemes evaluation [J]. IEEE Signal Processing Magazine, 2000, 17(5): 58 ~ 64.