

基于相关性的 JPEG 高压压缩图像峰值信噪比盲估计

陆敏俊¹, 王 慈²

(1. 上海交通大学 电子工程系, 上海 200240; 2. 华东师范大学 计算机科学技术系, 上海 200241)

摘 要: 量化噪声是 JPEG 图像压缩过程中产生的主要失真, 在图像质量评估时需要对其大小进行估计。针对峰值信噪比 (PSNR) 估计方法在实际应用时终端用户无法获取原始图像的限制, 基于图像空域相关性, 引入 JPEG 压缩图像的斜率均方差和最大非零离散余弦变换系数的位置作为可测特征量, 利用训练图像进行机器学习得到训练模型, 将测试图像进行解析得到图像 PSNR 盲估计结果。实验结果表明, 该 PSNR 估计结果优于目前流行的峰值信噪比盲估计方法。

关键词: 量化噪声; JPEG 图像; 峰值信噪比; 图像相关性; 斜率均方差; 机器学习

中文引用格式: 陆敏俊, 王 慈. 基于相关性的 JPEG 高压压缩图像峰值信噪比盲估计[J]. 计算机工程, 2017, 43(8): 253-257.

英文引用格式: Lu Minjun, Wang Ci. Blind Estimation of Peak Signal to Noise Ratio for JPEG Highly Compressed Images Based on Correlation[J]. Computer Engineering, 2017, 43(8): 253-257.

Blind Estimation of Peak Signal to Noise Ratio for JPEG Highly Compressed Images Based on Correlation

LU Minjun¹, WANG Ci²

(1. Department of Electrical Engineering, Shanghai Jiaotong University, Shanghai 200240, China;

2. Department of Computer Science and Technology, East China Normal University, Shanghai 200241, China)

[Abstract] Quantization noise is one of the dominant distortions of JPEG image compression, so its amplitude usually needs to be estimated for image quality assessment. Meanwhile, Peak Signal to Noise Ratio (PSNR) estimation method cannot get the original image in the practical applications, so Mean Squared Difference Slope (MSDS) and the position of maximum nonzero Discrete Cosine Transform (DCT) coefficient are introduced as the measurable characteristics based on image spatial correlation. Some images as well as their PSNR are trained to produce the PSNR estimation model, and the test image is fed into the model to estimate its PSNR. Experimental results show that the proposed algorithm is superior to the existing blind PSNR estimation algorithms.

[Key words] quantization noise; JPEG image; Peak Signal to Noise Ratio (PSNR); image correlation; Mean Squared Difference Slope (MSDS); machine learning

DOI: 10.3969/j.issn.1000-3428.2017.08.043

0 概述

JPEG 图像压缩是一种基于分块的离散余弦变换 (Discrete Cosine Transform, DCT) 压缩方法, 具有较好的重建质量, 且格式灵活, 支持多种压缩级别, 同时对色彩的信息保留较好, 被广泛应用于图像视频处理领域。因此, 对 JPEG 图像进行准确的客观质量评估是十分必要的。

为了定量地评估压缩图像质量进而评估压缩系统的性能, 本文引入峰值信噪比 (Peak Signal to

Noise Ratio, PSNR) 作为图像降质的测度。图像质量评价分为无参考评价、全参考评价及半参考评价 3 种。有参考、半参考评价需要原始无失真图像信息^[1-2]作为参考, 但在实际应用中由于接收端并不能得到压缩图像的原始样本, 因此本文采用盲估计的方法对 JPEG 图像进行 PSNR 评估。

由于 PSNR 估计的难点在于高压压缩图像, 本文建模以高压压缩图像为实验对象, 相比于峰值信噪比盲估计算法^[3-4]通过估计 DCT 系数的拉普拉斯分布参数 λ 计算压缩图像 PSNR 值, 本文针对 JPEG 图像

基金项目: 国家自然科学基金面上项目“基于内容的视频事件检测与描述研究”(61375016)。

作者简介: 陆敏俊 (1992—), 男, 硕士研究生, 主研方向为图像处理; 王 慈, 副教授、博士后。

收稿日期: 2016-06-23 **修回日期:** 2016-07-29 **E-mail:** lmjxqdlp@sjtu.edu.cn

的特点,利用图像在空域的相关性建立与压缩图像 PSNR 的关系,同时引入斜率均方差 (MSDS)、最大非零 DCT 系数位置 2 个特征量,通过机器学习得到训练模型,实现对 JPEG 图像的峰值信噪比盲估计,以取得较好的估计效果。

1 图像相关性分析

自然图像的像素值在空域是高度相关的, JPEG 图像压缩过程中的 DCT 变换,是利用这个特性对图像的空域能量进行集中去相关性,从而达到压缩的效果。本文从图像相关性出发建立峰值信噪比、JPEG 压缩图像可测特征量和相关性的联系,最后通过机器学习达到盲估计的效果。

文献[5]的研究表明,图像像素的相关性可以用像素之间的协方差来表达。

$$r(m, n) = \sigma_f^2 \rho^{|m|} \rho^{|n|} \quad (1)$$

其中, m, n 为图像像素之间的水平与垂直距离; σ_f^2 为像素值的方差; ρ 为图像相关性系数, 并且 $|\rho| \leq 1$ 。

式(11)表明,对于一幅图像进行 JPEG 压缩时的每个 8×8 图像块,像素之间的相关性只和块像素方差、像素之间的水平与垂直距离以及图像块的相关性系数有关。通过以上结论可以进一步地推导 PSNR 以及特征量与图像相关性的关系。

2 PSNR 及特征量与图像相关性的关系

如果能建立 PSNR 与图像块像素方差 σ_f^2 及图像相关性系数 ρ 的表达式,就可以通过这 2 个变量来直接估计 JPEG 压缩图像的 PSNR。但是由于这 2 个量只能从原始图像中获得而无法从压缩后的图像中直接获取,因此还需要建立与在压缩图像中这 2 个变量有关的可测特征量来间接估计出 PSNR 值。

2.1 PSNR 与图像相关性的关系推导

文献[6]提出一种基于拉式参数 λ 的 PSNR 估计方法,图像经过基于块的离散余弦变换后,其 DCT 域系数的概率分布更加接近于一个以 λ 为参数的拉普拉斯分布。文献[7]对于一个未经量化的 DCT 系数块,系数 $x(i, j)$ 服从以 $\lambda(i, j)$ 为参数的拉普拉斯分布,即:

$$f(x) = \frac{\lambda(i, j)}{2} e^{-\lambda(i, j)|x|} \quad (2)$$

量化后得到 X 的概率密度为:

$$P(X) = \int_{X-q/2}^{X+q/2} f(x) dx \quad (3)$$

其中, (i, j) 为 8×8 DCT 系数块的频率位置; X 为 DCT 系数量化后的值; q 为量化步长。

最终得到该点的平方误差为:

$$e^2 = \frac{1}{P(X)} \int_{X-q/2}^{X+q/2} f(x) (X-x)^2 dx \quad (4)$$

$$P_{\text{PSNR}} = 20 \cdot \lg \left(\frac{255}{\sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} e(i, j)^2}} \right) \quad (5)$$

文献[8]对于一个服从拉普拉斯分布的 DCT 系数块的方差,满足以下关系:

$$\sigma_F^2(i, j) = \sigma_f^2 [\mathbf{A} \mathbf{R} \mathbf{A}^T]_{i, i} [\mathbf{A} \mathbf{R} \mathbf{A}^T]_{j, j}$$

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^7 \\ \rho & 1 & \rho & \cdots & \cdots \\ \rho^2 & \rho & 1 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^7 & \cdots & \cdots & \cdots & 1 \end{bmatrix}_{8 \times 8} \quad (6)$$

其中, $\sigma_F^2(i, j)$ 为 DCT 系数 (i, j) 位置的方差; σ_f^2 为像素值的方差; \mathbf{A} 为 DCT 变换矩阵。

DCT 系数块的方差和拉式分布参数满足以下关系:

$$\lambda(i, j) = \frac{\sqrt{2}}{\sigma_F(i, j)} \quad (7)$$

其中, $\lambda(i, j)$ 为 8×8 DCT 系数块第 (i, j) 位置系数的拉普拉斯分布参数。通过式(6)、式(7)可以得到由 σ_f^2, ρ 表示的 DCT 系数拉式分布参数 $\lambda(i, j)$, 再将它们代入式(4)、式(5)中,最终可以得到 $\text{PSNR} = F(\sigma_f^2, \rho)$; F 为 σ_f^2, ρ 的函数。这证明可以通过估计式子中的 2 个变量 σ_f^2, ρ 直接估计 PSNR。

2.2 分块特征值提取

由于 σ_f^2, ρ 只能在原始图像中获得,而无法在压缩后的 JPEG 图像中直接获取,因此必须从压缩图像中寻找包含以上变量的可测特征量来间接获取相关性的信息从而估计 PSNR。下文通过推导证明这些可测特征量与图像相关性之间的关系。

特征 1 斜率均方差 (MSDS)

MSDS 是通过度量图像边界的不连续程度来反映块效应的^[9-10],计算模型如图 1 所示。

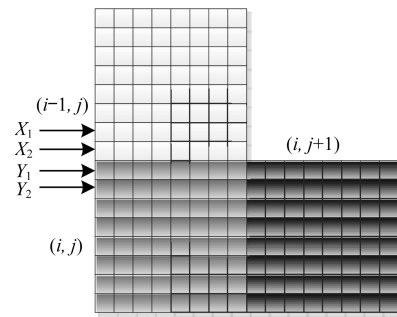


图 1 MSDS 计算模型

斜率均方差定义如下:

$$\text{MSDS}_{i, j} = \text{MSDS}_{L, i, j} + \text{MSDS}_{R, i, j} + \text{MSDS}_{U, i, j} + \text{MSDS}_{D, i, j} \quad (8)$$

$$\text{MSDS}_{L, i, j} = \sum_{k=0}^{N-1} \left[\frac{1}{2} (3p_{i, j}(k, 0) - p_{i, j}(k, 1)) - \frac{1}{2} (3p_{i, j-1}(k, N-1) - p_{i, j-1}(k, N-2)) \right]^2 \quad (9)$$

以图 1 中心块 (i,j) 第 1 行、第 2 行与上边界块 $(i-1,j)$ 最后 2 行最左侧的 4 个像素为例,假设它们的原始像素的大小从上到下分别为 X_1, X_2, Y_1, Y_2 ,其中 X_1, X_2, Y_1, Y_2 服从均值为 0、方差为 σ_f^2 的拉普拉斯分布。假设 4 个像素压缩后所带的量化噪声大小分别为 n_1, n_2, m_1, m_2 服从均值为 0 的拉普拉斯分布^[11],那么压缩后的这 4 个像素大小为 $X_1 + n_1, X_2 + n_2, Y_1 + m_1, Y_2 + m_2$;假设的 8 个量均为随机变量,不妨设:

$$M = \frac{1}{2}(3Y_1 - Y_2 - 3X_1 + X_2 + 3m_1 - m_2 - 3n_1 + n_2) \quad (10)$$

其中, M 为中心块 (i,j) 像素的计算;其余 31 组像素计算过程相同,因此求出 M 后,加上另外 31 组像素就得到 8×8 像素块的 MSDS 期望:

$$E(MSDS) = \sum_{k=1}^{32} E(M_k^2) \quad (11)$$

通过上文提到的相关性公式 $r(i,j)$ 与 σ_f^2, ρ 建立关系,从而得到 (i,j) 中心块的 MSDS 期望 $E(MSDS) = G(\sigma_f^2, \rho)$, G 为 σ_f^2, ρ 的函数。

特征 2 最大非零 DCT 系数位置(MAXD)

原始图像 JPEG 压缩过程中各个 8×8 DCT 系数块会产生量化噪声,在假设 DCT 系数服从拉普拉斯分布时推导得到包含 σ_f^2, ρ 的量化噪声函数^[12]。另一方面,图像 DCT 系数假设服从均匀分布时可求出各个 AC 系数的量化噪声大小。基于图像相关性,在固定相关性下,最大非 0 系数与噪声之间的大小是确定的关系,因此以理论量化噪声模型和实际噪声估计值对最大非零 DCT 系数建立等式,得到包含 σ_f^2, ρ 的函数,最后结合特征 MSDS 的等式联立求出 σ_f^2, ρ 。

由于在实际操作中难以直接得到该方程组的解析解,因此本文通过机器学习的方法间接求取 PSNR 与 σ_f^2, ρ 的关系。均匀分布假设 AC 系数的量化噪声只和它所在的频率位置有关,且不同 AC 系数的量化误差间存在确定的相关性,因此本文用最大非零 DCT 系数位置(MAXD)来间接表示最大非零 DCT 系数的量化噪声作为可测特征量。具体的提取步骤是:对于每个 8×8 的 DCT 系数块,对其 DCT 系数进行 Zigzag 扫描编号,再排序得到最大非零 DCT 系数的位置信息作为特征向量。对于直流块,其最大非零 DCT 系数位置为直流系数的位置。

3 JPEG 图像峰值信噪比盲估计方法

本文采用机器学习的方法来间接估计峰值信噪

比。对图像按压缩系数 QS 分类压缩之后,通过计算压缩后训练图像每个 8×8 系数块的特征参数作为输入、块 MSE 作为输出进行 SVM 训练,得出块 MSDS、块最大非零 DCT 系数位置与块 MSE 之间的关系模型。同时,对测试图像按 QS 分类后,在模型中输入压缩后测试图像的特征向量,即可获得图像的 PSNR 估计结果。

3.1 图像数据库

将近年来 PSNR 盲估计算法中性能最优的方法^[13]作为本文的对比算法,本文使用与该文献相同的 29 幅标准图像库^[14]图像,采用其亮度成分进行仿真,同时将 29 幅图像中的每一幅图像作为测试图像时,另外 28 幅图像作为训练图像进行训练,从而保证了测试与训练间的独立性。

3.2 机器学习方法

本文选择了经典的支持向量回归模型,将同一 QS 下压缩图像每个 8×8 图像块的特征值:MSDS,最大非零 DCT 系数位置作为机器学习的输入训练特征参数^[15],将对应的 8×8 图像块 MSE 作为训练输出进行 SVM 训练得到训练模型。

具体流程如下:

1)按设定的压缩系数 QS 对训练图像和测试图像进行 JPEG 压缩。本文实验选取 QS 为 5, 10, 15, 20, 25, 30, 35, 40 的 8 种情况,分别对标准数据库图像进行压缩。在对每一幅图像作为测试图像时,将另外 28 幅图像作为训练图像。

2)对于训练图像,首先进行 8×8 分块,在同一 QS 下,分别求出每个图像块的特征参数:MSDS,最大非零 DCT 系数位置作为训练输入,图像块 MSE 作为训练输出,对其进行 SVM 训练得到同一 QS 下训练模型。

3)对于测试图像,将待测试图像解析,获取其 QS,再将其每个 8×8 图像块特征参数 MSDS、最大非零 DCT 系数位置输入到对应的 QS 训练模型,最后输出每个 8×8 图像块的 MSE 预测值作为估计值。

4)最后得到测试图像所有 8×8 图像块的 MSE 测试值,然后利用下式求解出每幅测试图像的 PSNR 估计结果:

$$P_{\text{PSNR}} = 20 \cdot \lg \left(\frac{255}{\sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N MSE_{i,j}}} \right) \quad (12)$$

其中, $MSE_{i,j}$ 是测试图像第 (i,j) 个 8×8 图像块的 MSE 预测值; M, N 为图像行方向与列方向 8×8 图像块数,255 为 8 bit 图像像素的最大值。

某 QS 下的训练流程如图 2 所示。

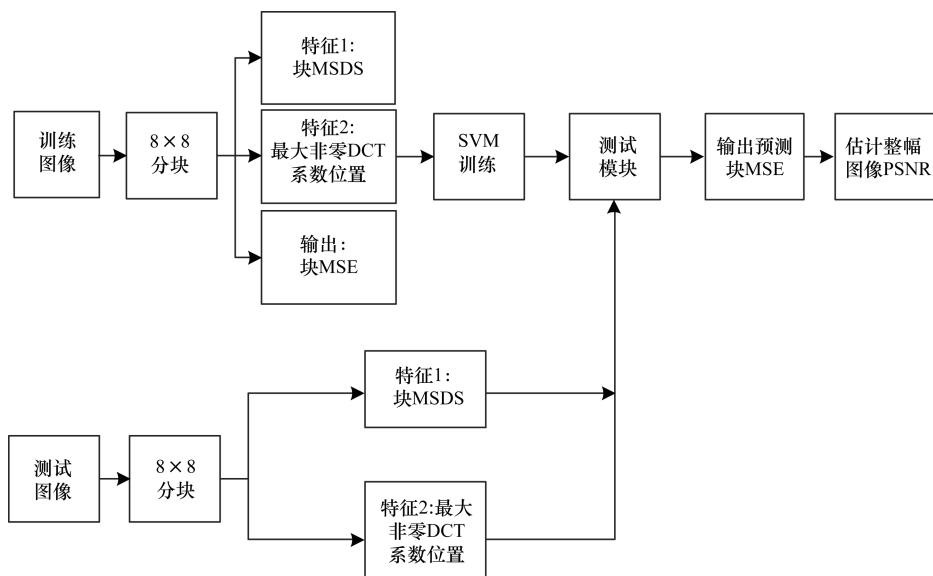


图2 QS算法流程

4 实验结果与分析

本文使用标准图像库,并引入交叉验证方法,当标准库中每一幅图像作为测试图像时,另外28幅图像作为训练图像,从而保证了PSNR估计结果的准确性,下文是对测试图像进行测试后得到的估计结果,同时与对比算法的PSNR盲估计方法进行对照,如图3、图4所示。

在图3、图4中, x 轴 y 轴分别为测试图像PSNR算法估计值与图像真实PSNR值,中间 $y=x$ 实线为等值线。散点离等值线越近表明估计误差越小。从图中可以看出,本文方法估计的散点图从整体来看比对比算法更加收拢于等值线。

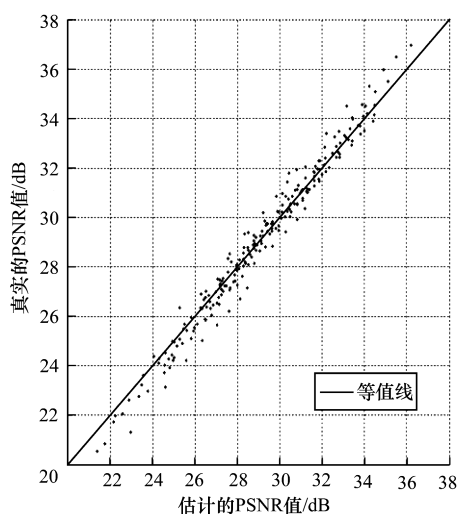


图3 本文算法测试结果散点图

从表1中可以看出,本文方法在平均绝对误差上比对比算法小了0.14 dB,误差减小了25.5%,而在均方根误差上也减小了0.11 dB,误差减小了17.2%。

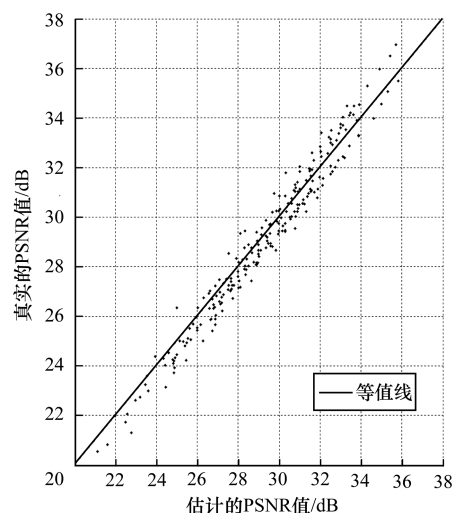


图4 对比算法测试结果散点图

表1 实验结果综合对比

方法	平均绝对误差/dB	均方根误差/dB	LCC	SROCC
本文方法	0.41	0.53	0.989	0.988
对比方法	0.55	0.64	0.981	0.979

因此,本文方法整体PSNR评估性能有了较大的提升,优于近年JPEG高压缩图像的PSNR盲估计方法。同时本文方法为量化拟合程度,使用线性相关系数(Liner Correlation Coefficient, LCC)、斯皮尔曼等级相关系数(The Spearman Rank-order Correlation Coefficient, SROCC)2个统计量作为评判标准, $LCC \in [0, 1]$ 体现了PSNR真实值与估计值的线性相关性, $SROCC \in [0, 1]$ 通过使用单调函数来描述PSNR真实值与估计值的相关性。从结果可以看出,本文方法在LCC与SROCC上均优于对比方法,更加接近于1,保证了本文方法在大样本下较好的拟合度。

5 结束语

本文基于图像相关性的特征参数,通过机器学习,得出 JPEG 图像峰值信噪比与相关性有关的图像特征参数的关系,从而实现一种适用于 JPEG 压缩图像的 PSNR 盲估计方法。实验结果表明,本文方法在线性相关系数、斯皮尔曼等级相关系数上均优于对比方法。

参考文献

- [1] Anush M K, Kalpana S, Rajiv S, et al. Wireless Video Quality Assessment: A Study of Subjective Scores and Objective Algorithms [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20 (4): 587-599.
- [2] Chen Xiaolin, Zhang Rui, Zheng Shibao. Image Quality Assessment Based on Local Edge Direction Histogram [C]//Proceedings of IEEE International Conference on Image Analysis and Signal Processing. Washington D. C., USA, IEEE Press, 2011: 108-112.
- [3] Ichigaya A, Kurozumi M, Hara N, et al. A Method of Estimating Coding PSNR Using Quantized DCT Coefficients [J]. IEEE Transactions on Circuits Systems Video Technology, 2006, 16(2): 251-259.
- [4] Brandao T, Queluz M P. No-reference Image Quality Assessment Based on DCT-domain Statistics [J]. Signal Processing, 2008, 88(4): 822-833.
- [5] 景 麟,王宏远,马 泳. 基于拉普拉斯分布模型的快速 DCT 算法 [J]. 电视技术, 2008, 32(11): 16-19.
- [6] Lam E Y, Goodman J W. A Mathematical Analysis of the DCT Coefficient Distributions for Images [J]. IEEE Transactions on Image Processing, 2000, 9 (10): 1661-1666.
- [7] 董皓远,方向忠. 基于图像先验知识的量化噪声盲估计算法 [J]. 计算机工程, 2010, 36(11): 195-197.
- [8] Pao I M, Sun Ming Ting. Modeling DCT Coefficients for fast Video Encodings [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(4): 608-616.
- [9] Minami S, Zakhori A. An Optimization Approach for Removing Blocking Effects in Transform Coding [J]. IEEE Transactions on Circuits System, 1995, 5: 74-82.
- [10] 沈诗文. 量化噪声建模及其在去块效应算法中的应用 [D]. 上海: 上海交通大学, 2011.
- [11] 程 燕. 图像超分辨率重建关键技术的研究 [D]. 上海: 上海交通大学, 2007.
- [12] Robertson M A, Stevenson R L. DCT Quantization Noise in Compressed Images [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(1): 27-38.
- [13] Yang Gao, Tan Yap Peng. Blind PSNR Estimation Using Shifted Blocks for JPEG Images [C]//Proceedings of ISCAS' 11. Washington D. C., USA: IEEE Press, 2011: 216-225.
- [14] Sheikh H R, Wang Z, Cormack L, et al. IVE Image Quality Assessment Database Release 2 [EB/OL]. [2015-12-21]. <http://live.ece.utexas.edu/research/quality>.
- [15] Anish M, Moorthy A K, Bovik A C. No-reference Image Quality Assessment in the Spatial Domain [J]. IEEE Transactions on Image Processing, 2012, 21 (12): 4695-4708.

编辑 索书志

(上接第 252 页)

- [9] Rainer L, Alexander K, Vadim P. Empirical Analysis of Detection Cascades of Boosted Classifiers for Rapid Object Detection [C]//Proceedings of Joint Pattern Recognition Symposium. Berlin, Germany: Springer, 2003: 297-304.
- [10] Proenca H, Alexandre L A. The NICE. I: Noisy Iris Challenge Evaluation, Part I [C]//Proceedings of the 1st IEEE International Conference on Biometrics: Theory, Applications, and Systems. Washington D. C., USA: IEEE Press, 2007: 1-4.
- [11] CASIA Iris Image Database [EB/OL]. [2016-04-11]. <http://biometrics.idealtest.org>.
- [12] Tan C W, Kumar A. Towards Online Iris and Periocular Recognition Under Relaxed Imaging Constraints [J]. IEEE Transactions on Image Processing, 2013, 22 (10): 3751-3765.
- [13] Farmanullah J, Imran U. Iris Localization in Frontal Eye Images for Less Constrained Iris Recognition Systems [J]. Digital Signal Processing, 2012, 22 (6): 9818-9865.
- [14] Zhao Z J, Kumar A. An Accurate Iris Segmentation Framework Under Relaxed Imaging Constraints Using Total Variation Model [C]//Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2015: 3828-3836.
- [15] Xu Li, Yan Qiong, Xia Yang. Structure Extraction from Texture via Relative Total Variation [J]. ACM Transactions on Graphics, 2012, 31(6): 139-149.

编辑 陆燕菲