

面向图像集的置信区间内一致性增强质量评价准则

刘浩* 孙晓帆 张鑫生 吴乐明 况奇刚

(东华大学信息科学与技术学院 上海 201620)

摘要: 在对整个图像集进行增强质量评价时, 现有的平均准则会随着不同图像集非一致性地变化, 从而导致较大的评价质量波动。为此, 该文提出一个面向图像集的置信区间内一致性增强质量评价准则, 通过设置应用参数并使用置信区间筛选数据, 再比较各图像增强前后的质量分数差值, 由此评估图像质量增强的一致性, 最终计算出一致性增强质量分数有效值。在众多图像增强算法中, 所提准则能够挑选出具体应用所需要的稳定性强、可靠性高的增强算法。实验结果表明, 所提准则具有良好的主客观评价一致性, 性能优于当前的平均准则, 为各种图像增强算法提供了一个可用于任意图像集的质量评价准则。

关键词: 图像增强; 图像集; 质量评价; 一致性增强; 置信区间

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2019)01-0219-07

DOI: 10.11999/JEIT180272

Consistency Enhancement Quality Assessment Criterion in Confidence Interval for Image Set

LIU Hao SUN Xiaofan ZHANG Xinsheng WU Leming KUANG Qigang

(College of Information Science and Technology, Donghua University, Shanghai 201620, China)

Abstract: When evaluating the enhancement quality of a whole image set, the existing average score criterion varies inconsistently with different image sets and produces a large evaluation quality fluctuation. Therefore, this paper proposes a consistency enhancement quality assessment criterion in confidence interval for any image set. By setting application parameters and using confidence interval to screen data, the proposed criterion compares the quality score difference before and after enhancing each image, and evaluates the consistency of image quality enhancement, and then calculates the effective value of consistency enhancement quality scores. Among many image enhancement algorithms, the proposed criterion can select the high-reliability enhancement algorithm for a specific application. The experimental results show that the proposed criterion has good subjective and objective consistency and outperforms the existing average score criterion, which provides an evaluation criterion for those image enhancement algorithms applied to any image set.

Key words: Image enhancement; Image set; Quality assessment; Consistency enhancement; Confidence interval

1 引言

图像增强算法的优劣需要有效的图像质量评价方法来衡量, 现有的图像质量评价方法均是对单幅图像进行图像质量评价的有效方法^[1]。随着大数据时代的到来, 图像数量日益庞大, 通常需要以图像集为单位来增强图像, 因此需要一个面向图像集的增强质量评价准则。目前对于图像集的增强质量评价, 通常采用单幅图像质量分数的平均值作为评价

准则, 以平均值最好来说明某一图像增强算法最优^[2-5]。然而从大量的实验数据可知, 平均准则所获结果随着图像集的变化存在较大的波动, 其可靠性较低, 如果将由此选出的图像增强算法投入到实际应用中, 会产生很大的误差。因此, 如何为人脸识别、安全监控、生物医学等不同应用领域分别找到有效的图像增强算法, 需要一个可靠性更高的增强质量评价准则。本文提出一种面向图像集的置信区间内一致性增强质量评价准则, 可以在应用参数约束下找出适当的图像增强算法, 并为所有的图像增强算法提供一个鲁棒性高的质量评价体系。

本文以水下图像集为例展开论述, 其余部分组织如下: 第2节介绍本文使用的经典图像增强算法与适用于单幅图像的典型质量评价方法, 作为文中

收稿日期: 2018-03-23; 改回日期: 2018-09-05; 网络出版: 2018-09-19

*通信作者: 刘浩 liuhao@dhu.edu.cn

基金项目: 上海市自然科学基金(18ZR1400300)

Foundation Item: The Natural Science Foundation of Shanghai (18ZR1400300)

实验操作的基础；第3节引出问题，通过数据展示了平均值评价准则的缺陷；第4节介绍本文提出的面向图像集的置信区间内一致性增强质量评价准则；第5节对所提评价准则进行比较与实验应用；第6节得出结论并分析未来可能的工作。

2 图像增强算法与评价方法

本文用到的一些经典水下图像增强算法包括：逆滤波算法^[6]、均值滤波算法^[7]、动态阈值白平衡算法^[8]、直方图均衡算法^[9]、对比度受限的自适应直方图均衡化(Contrast Limited Adaptive Histogram Equalization RGB, CLAHE-RGB)算法^[10]、对比度拉伸算法^[11]、无监督颜色模型算法^[12]。

本文用到的单幅图像的典型质量评价方法包

括：源自香农信息论的熵值(entropy)^[13]方法、Yang与Sowmya提出的UCIQE(Underwater Color Image Quality Evaluation)^[5]方法。

3 图像集的平均准则

3.1 平均准则的定性分析

本节运用不同的水下图像增强算法对图1(a)^[14]所示原图进行图像增强定性实验，得到不同的质量增强效果，如图1(b)—图1(h)所示。

针对图1(b)—图1(h)中不同图像增强算法的处理结果，本文采用UCIQE和熵值两种图像质量评价方法进行客观评测，图2分别给出了相应的质量分数曲线。从图2可以看出，一些经典的图像增强算法运用于图1(a)后，存在着降低原图质量分数的情况。

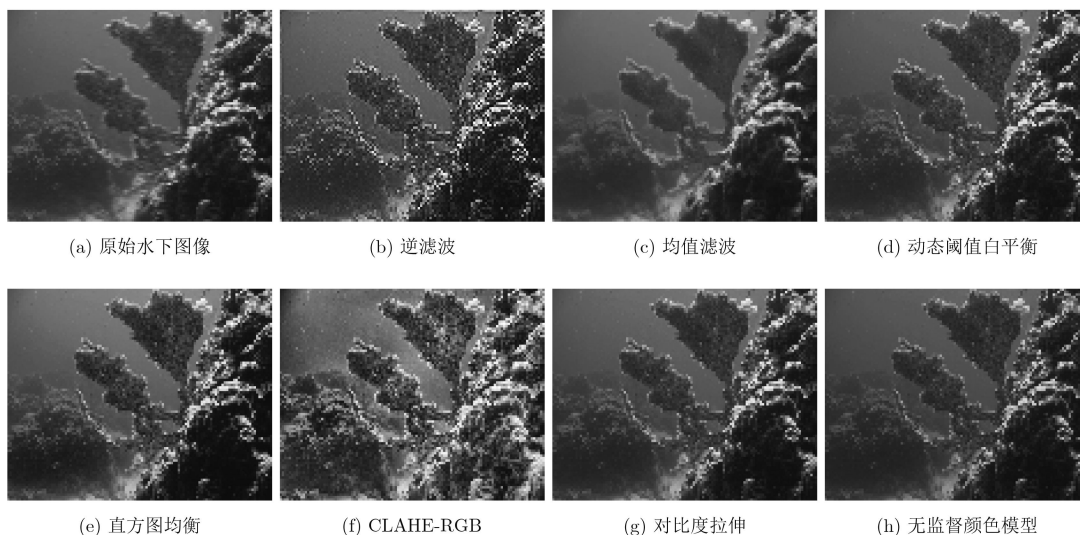


图1 不同图像增强算法的定性比较

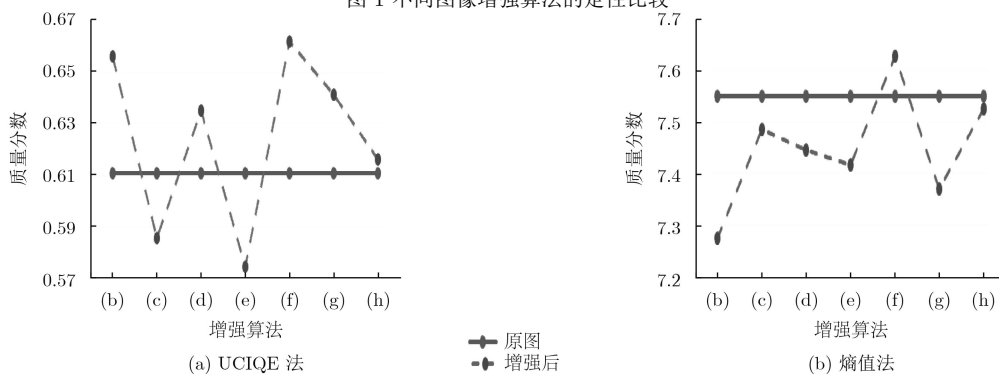


图2 各增强算法下的质量分数曲线

3.2 平均准则的定量分析

为了避免定性实验存在特殊性，进一步采用定量实验对10幅典型图像组成的图像集P进行图像增强实验，图3给出了典型图像集P中的6幅图像，它们在相关文献中被广泛使用。

选取第2节所述的各种图像增强算法对图像集

P逐幅图像进行增强，求出原图与增强后图像的熵值质量分数，将其中6幅图像的实验结果展示在图4中，图标号1~6依次代表图像Ancuti1, Ancuti2, Eustice, Fish, Reef1, Reef2。随后分别为每一增强算法计算出所有图像质量分数的平均值，如表1所示。

由表1看出，CLAHE-RGB，无监督颜色模型

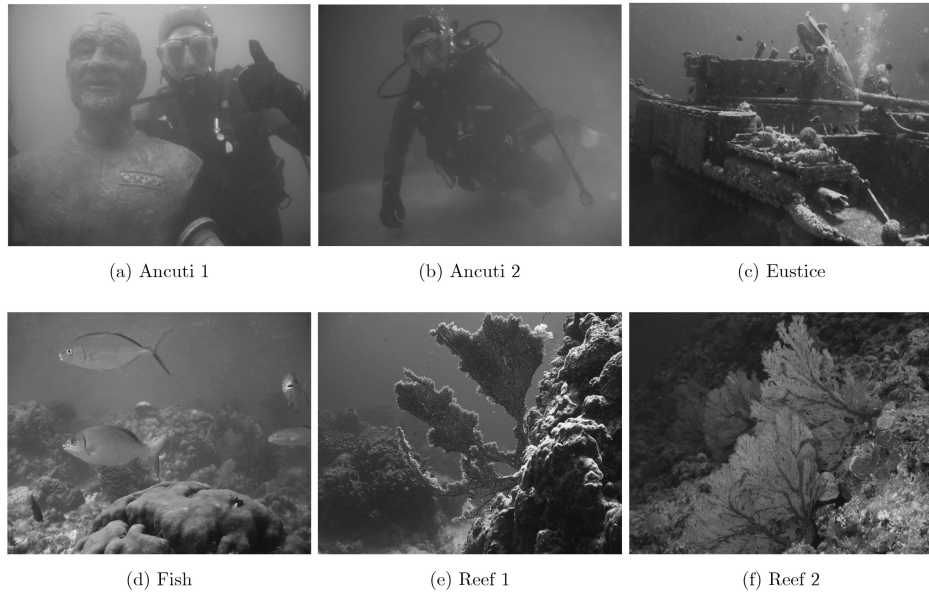


图3 典型水下图像

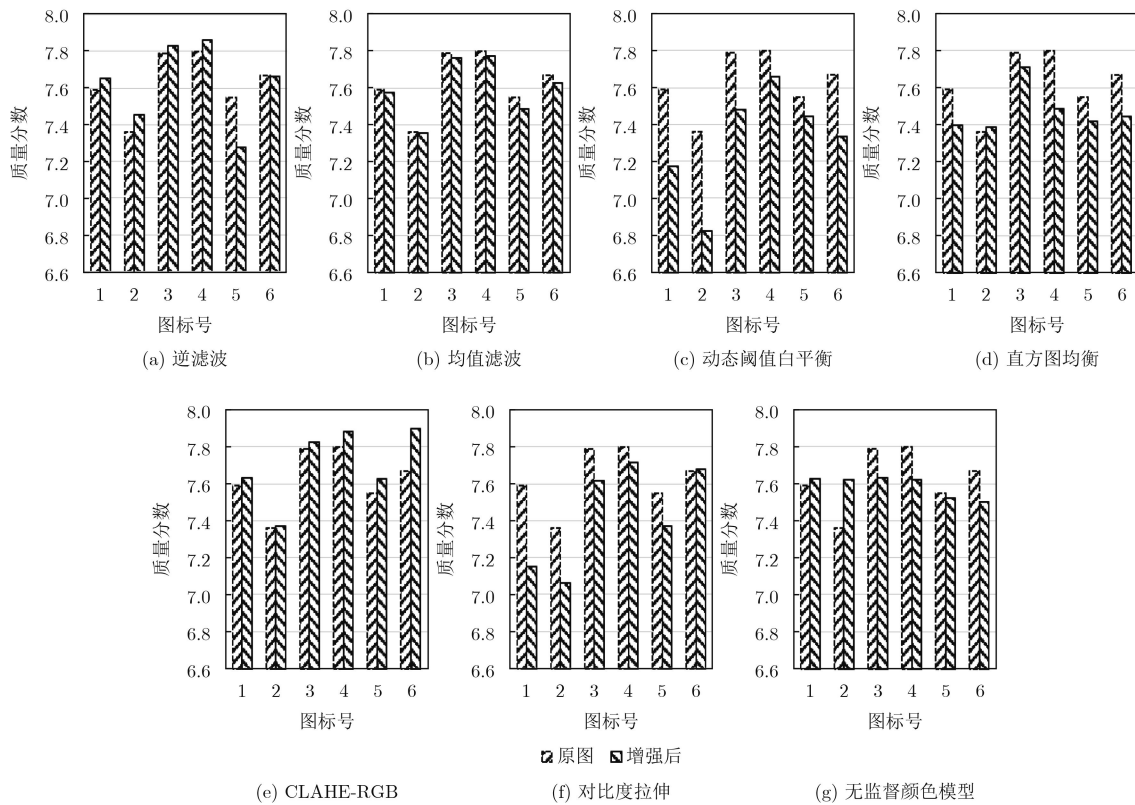


图4 图像增强前后的熵值质量分数对比

与直方图均衡算法对该图像集增强后所得到的熵值平均值较高，但在图4中可以发现，在这3种算法中，无监督颜色模型与直方图均衡算法的增强能力稳定性差，与原图的质量分数契合度低，其熵值的平均值较高是由于存在个别图像的增强效果超好，以及增强效果较差的图像本身分数较高等原因，由此可以得出：质量分数平均值的高低并不能全面地

说明所使用图像增强算法的好坏。从图4和表1可以看出：CLAHE-RGB算法的整体增强能力与稳定性最好；逆滤波算法的熵值评价分数虽然不是最优，但其增强能力稳定性好，增强后的质量分数与原图的质量分数契合度高。总体而言，CLAHE-RGB与逆滤波算法更适合应用于图像集较大的场合。因此在大数据时代，对于面向特定应用的图像增强算法

表1 熵值的平均分数对比

算法	原图	逆滤波 ^[6]	均值滤波 ^[7]	动态阈值白平衡 ^[8]	直方图均衡 ^[9]	CLAHE-RGB ^[10]	对比度拉伸 ^[11]	无监督颜色模型 ^[12]
熵值	7.41	7.47	7.38	7.18	7.49	7.72	7.21	7.55

的选取, 需要提出一个稳定性更强、可靠性更高的增强质量评价准则。

4 一致性增强质量评价准则

为了解决上述问题, 本文提出了一种全新的增强质量评价准则: 一致性增强质量评价(Consistency Enhancement Quality Assessment, CEQA)准则。面向图像集的置信区间内CEQA准则的质量评价体系如图5所示, 首先对原始图像集进行逐幅图像的质量增强, 并对增强前后的各个图像进行质量评价, 获得质量分数差值, 接下来, 根据不同的需求设置参数, 并使用置信区间筛选数据, 再通过剩余的有效测试数据判断算法的一致性, 最后, 计算出一致性增强质量分数有效值(CEQA_{eff})。

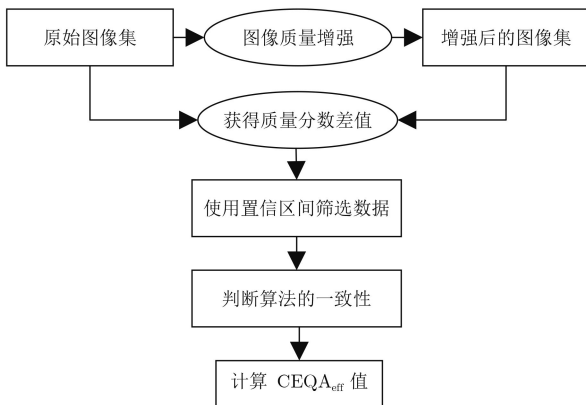


图5 CEQA准则的质量评价体系

CEQA准则的流程图如图6所示, 其具体步骤如下:

步骤1 首先使用某一质量评价方法M对获取的图像集中的全部原始图像 $I_1, I_2, \dots, I_n$ 进行质量评价, 得出各幅原始图像 I_i 的质量分数基准值 α_i , 作为CEQA公式中的参数之一, 也作为评价图像增强算法好坏的基准, 其中, $i(i=1, 2, \dots, n)$ 为图像标号, n 为图像集中所包含的图像数量;

步骤2 使用某一图像增强算法A对图像集中的全部原始图像进行图像质量增强, 得到增强后的图像集 $I'_1, I'_2, \dots, I'_n$;

步骤3 使用步骤1的质量评价方法M对增强后的全部图像 I'_i 进行质量评价, 得到增强后的质量分数 β_i 。若 β_i 大于与之相对应的 α_i , 则说明在质量

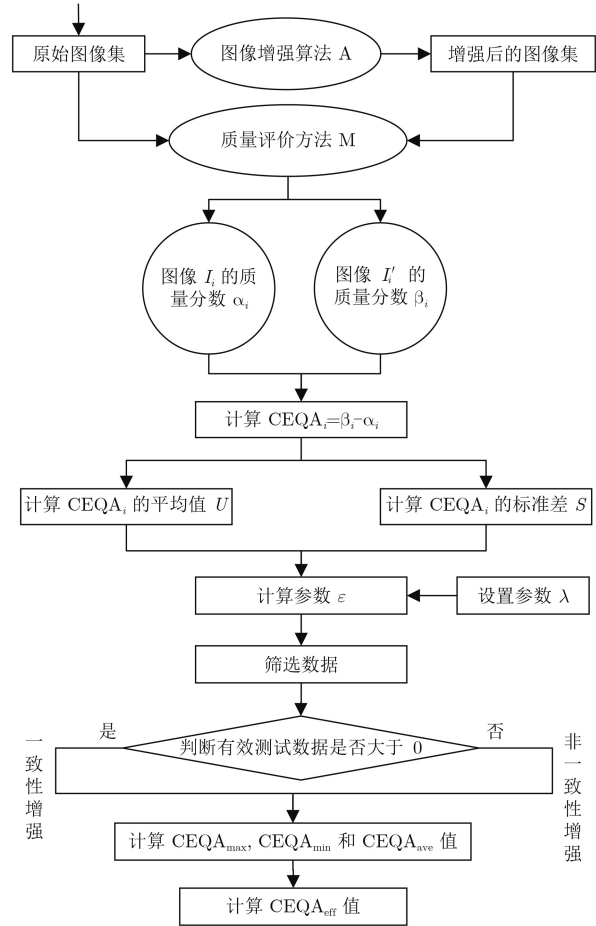


图6 CEQA准则的流程图

评价方法M的度量下, 图像增强算法A增强了这幅原始图像 I_i 的图像质量; 反之, 表明这一M与A的组合对于图像 I_i 造成了质量降低;

步骤4 将得到的质量分数 α_i 和 β_i 代入求质量分数差值CEQA_i的公式:

$$\text{CEQA}_i = \beta_i - \alpha_i \quad (1)$$

得到CEQA_i值, 若该值为正, 则增强后的图像质量分数 β_i 高于基准值 α_i , 即在质量评价方法M的度量下, 图像增强算法A增强了这一幅原始图像 I_i 的图像质量;

步骤5 求出各幅图像的质量分数差值CEQA_i的平均值 U :

$$U = \frac{1}{n} \sum_{i=1}^n \text{CEQA}_i \quad (2)$$

并求出CEQA_i的标准差 S :

$$S = \sqrt{\frac{1}{n} \sum_{i=1}^n (\text{CEQA}_i - U)^2} \quad (3)$$

步骤 6 根据具体应用的需求设置参数 λ ，将该参数代入公式：

$$\varepsilon = \lambda \cdot S \quad (4)$$

求出参数 ε ，得到置信区间 $[U - \varepsilon, U + \varepsilon]$ ，使用该置信区间筛选全部 CEQA_i 值，在此置信区间之外的 CEQA_i 值被视为可忽略误差，将其舍弃，保留在置信区间内的 CEQA_i 值作为有效测试数据，继续执行步骤7操作。典型地，当 $\lambda = 1.96$ 时， $[U - \varepsilon, U + \varepsilon]$ 为95%的置信区间；

步骤 7 将有效测试数据重新按照 CEQA_i 的下标编号升序排序，整理为 $\text{CEQA}_1, \text{CEQA}_2, \dots, \text{CEQA}_m$ ，其中， $j(j = 1, 2, \dots, m)$ 为新的图像标号， $m(m \leq n)$ 为有效测试数据的个数；

步骤 8 将有效测试数据与0值进行比较，若全部有效测试数据都大于0，则说明在该质量评价方法M的度量下，图像增强算法A为一致性增强算法；否则，图像增强算法A为非一致性增强算法；

步骤 9 基于有效测试数据，首先求出 CEQA_j 的最大值：

$$\text{CEQA}_{\max} = \max \{\text{CEQA}_1, \text{CEQA}_2, \dots, \text{CEQA}_m\} \quad (5)$$

然后求出 CEQA_j 的最小值：

$$\text{CEQA}_{\min} = \min \{\text{CEQA}_1, \text{CEQA}_2, \dots, \text{CEQA}_m\} \quad (6)$$

并求出 CEQA_j 的平均值：

$$\text{CEQA}_{\text{ave}} = \text{ave} \{\text{CEQA}_1, \text{CEQA}_2, \dots, \text{CEQA}_m\} \quad (7)$$

步骤 10 获得在质量评价方法M下，图像增强算法A对该图像集的一致性增强质量分数有效值 CEQA_{eff} ：

$$\text{CEQA}_{\text{eff}} = \mu \cdot \frac{\text{CEQA}_{\text{ave}}}{\text{CEQA}_{\max} - \text{CEQA}_{\min}} + (1 - \mu) \cdot \frac{\text{CEQA}_{\min}}{\text{CEQA}_{\max} - \text{CEQA}_{\min}} \quad (8)$$

其中， $\mu (0 < \mu < 1)$ 为权重系数。对于这一图像集，该图像增强算法A在质量评价方法M的评价条件下，若算法A为一致性增强的图像增强算法， CEQA_{ave} 值不同时， CEQA_{eff} 值越高，说明其质量提升的强度越大，增强能力越好；而当 CEQA_{ave} 值相同时， CEQA_{eff} 值越高，说明其稳定性越好。同理，若算法A为非一致性增强的图像增强算法也可得到相同的结论，但当增强一致性不同时，不同增强算法的增强能力与稳定性，不可以通过 CEQA_{eff} 值的高低来比较。

本文提出的一致性增强质量评价准则可以为所有的图像增强算法提供一个评价准则，根据具体的应用设置不同的参数，得到的一致性增强算法即为适合该应用的有效算法。对于图像集应用，一致性增强质量评价准则比平均准则的可靠性更高。

5 实验结果

5.1 CEQA准则与平均准则的实验对比

为了说明CEQA准则可以比平均准则更好地评价图像增强算法的优劣，本文选取200幅水下图像^[15, 16]作为水下图像集c，随机选取其中的50幅图像作为图像集b，再随机选图像集b中的5幅图像作为图像集a，对其分别使用动态阈值白平衡算法与对比度拉伸算法来增强图像质量，得到增强后的图像，再使用质量评价方法UCIQE对以上增强前后的图像进行图像质量评价，所得图像质量分数如图7所示。由图7可以看出，当图像集a为5幅图像时，经过动态阈值白平衡增强后的图像集存在图像质量降低的情况，属于非一致性增强，而对比度拉伸算法则一致性地增强了各幅图像的图像质量，稳定性更好、可靠性更高，属于一致性增强的图像增强算法。若将一致性增强的图像增强算法应用到更大的图像集b时，对比度拉伸算法能够更可靠、稳定地增强图像质量，而在小图像集中非一致性增强的图像增强算法，当应用在大图像集时，进一步扩大了

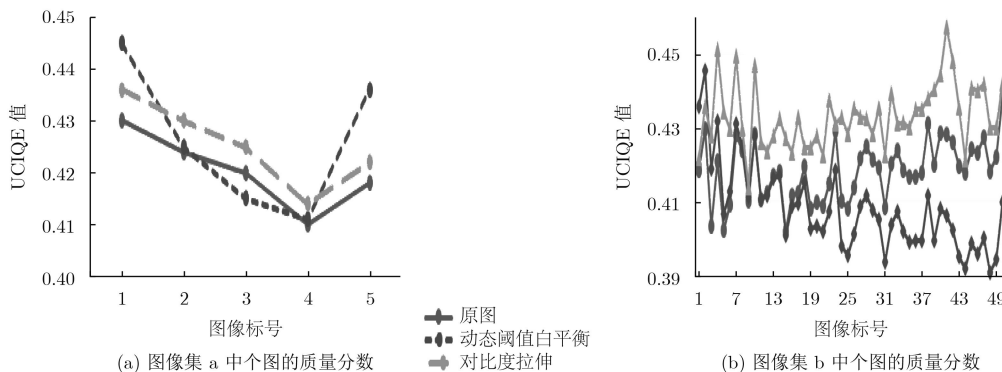


图 7 各幅图像的图像质量分数

误差,出现了大量降低图像质量的情况。

接下来,分别使用平均准则与CEQA准则对以上实验数据进行整理,其中,取 $\lambda = 2$, $\mu = 0.5$,得到对比结果如表2所示。从表2中发现,当图像集变化时,使用CEQA准则选择出的一致性增强算法的CEQA_{eff}值一直高于非一致性增强算法的CEQA_{eff}值,与图7所得的理论相符,具有良好的主客观评价一致性;而平均准则所得到的平均值分数随图像集的变化产生了较大波动,将图像集较小时所获得的高平均值方法应用到大图像集时,所得到的平均值分数甚至低于原图的平均值分数。由此表明,在大数据应用挑选有效的图像增强算法时,平均准则的可靠性低于CEQA准则。

表2 平均准则的平均值与CEQA准则的CEQA_{eff}值的对比(在UCIQE方法下)

	图像集	原图	动态阈值白平衡 ^[6]	对比度拉伸 ^[11]
平均值	图像集a	0.422	0.427	0.426
	图像集b	0.418	0.408	0.433
	图像集c	0.419	0.389	0.446
CEQA _{eff} 值	图像集a		-0.101	0.698
	图像集b	无	-0.510	0.294
	图像集c		-0.792	0.199

5.2 CEQA准则的主客观评价一致性对比

接下来,对第3节所述图像集P进行CEQA准则的实验应用,以验证通过该准则计算得到的结论符

合图4中所得到的主观论证。取 $\lambda = 1.96$, $\mu = 0.5$,得到该图像集在熵值质量评价方法下,各算法的CEQA_{eff}值如表3所示:

从实验中发现,该图像集在熵值质量评价方法下,只有CLAHE-RGB算法成为一致性增强的图像增强算法,并且由表3可以看出,该水下图像集在熵值评价方法下,CLAHE-RGB算法为最优算法,增强效果最好,可靠性与稳定性最高,与图4中所得到的理论结果相符。此外,在非一致性增强的图像增强算法中,逆滤波算法得到的CEQA_{eff}值相对较高,其可靠性与稳定性相对最好,同样与图4中所得到的理论结果相符。因此,CEQA准则具有很好的主客观评价一致性。

6 结论

本文提出了一种面向图像集的置信区间内一致性增强质量评价准则,根据不同的应用需求设置特定的参数约束,基于置信区间对图像增强前后的质量分数差值进行筛选,保留可容忍误差内的有效测试数据,再通过有效测试数据计算出一致性增强质量分数有效值,进而选出可靠、稳定的图像增强算法,以期运用于大图像集等实际应用。所提准则克服了平均准则的缺陷,不易随着图像集所含图像数量的变化而产生波动,其稳定性高、可靠性强,实际操作符合理论结果。由于未来的应用可能需要更复杂的参数约束,仍然有待提出更具竞争力的增强质量评价准则。

表3 熵值的CEQA_{eff}值

增强算法	CLAHE-RGB ^[10]	逆滤波 ^[6]	无监督颜色模型 ^[12]	直方图均衡 ^[9]	动态阈值白平衡 ^[8]	对比度拉伸 ^[11]	均值滤波 ^[7]
CEQA _{eff} 值	0.202	0.193	-0.025	-0.205	-0.609	-0.713	-0.8057

参考文献

- [1] 王志明. 无参考图像质量评价综述[J]. 自动化学报, 2015, 41(6): 1062-1079. doi: 10.16383/j.aas.2015.c140404.
WANG Zhiming. Review of no-reference image quality assessment[J]. *Acta Automatica Sinica*, 2015, 41(6): 1062-1079. doi: 10.16383/j.aas.2015.c140404.
- [2] PENG Y T and COSMAN P C. Underwater image restoration based on image blurriness and light absorption[J]. *IEEE Transactions on Image Processing*, 2017, 26(4): 1579-1594. doi: 10.1109/TIP.2017.2663846.
- [3] LI Chongyi, GUO Jichang, CONG Runmin, et al. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior[J]. *IEEE Transactions on Image Processing*, 2016, 25(12): 5664-5677. doi: 10.1109/TIP.2016.2612882.
- [4] LI Chongyi, GUO Jichang, CHEN Shanji, et al. Underwater image restoration based on minimum information loss principle and optical properties of underwater imaging[C]. *IEEE International Conference on Image Processing*, Phoenix, USA, 2016: 1993-1997.
- [5] YANG M and SOWMYA A. An underwater color image quality evaluation metric[J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 6062-6071. doi: 10.1109/TIP.2015.2491020.
- [6] 吴雪根, 吴谨, 张鹤. 逆滤波法在图像复原中的应用[J]. 信息技术, 2011(10): 183-185. doi: 10.3969/j.issn.1009-2552.2011.10.050.
WU Xueyin, WU Jin, and ZHANG He. Research on image restoration techniques based on inverse filtering algorithm[J]. *Information Technology*, 2011(10): 183-185.

- doi: [10.3969/j.issn.1009-2552.2011.10.050](https://doi.org/10.3969/j.issn.1009-2552.2011.10.050).
- [7] 何石, 潘晓璐, 李一民. 一种均值滤波的优化算法[J]. 信息技术, 2012(3): 133–137. doi: [10.3969/j.issn.1009-2552.2012.03.038](https://doi.org/10.3969/j.issn.1009-2552.2012.03.038).
- HE Shi, PAN Xiaolu, and LI Yimin. Optimization algorithm for average filtering[J]. *Information Technology*, 2012(3): 133–137. doi: [10.3969/j.issn.1009-2552.2012.03.038](https://doi.org/10.3969/j.issn.1009-2552.2012.03.038).
- [8] 苏志锋. 基于FPGA的图像预处理研究与实现[D]. [博士学位文], 华南理工大学, 2015.
- SU Zhifeng. Studying and implementation of image signal preprocessing based on FPGA[D]. [Ph.D. dissertation], South China University of Technology, 2015.
- [9] 李耀辉, 刘保军. 基于直方图均衡的图像增强[J]. 华北科技学院学报, 2003, 5(2): 65–67.
- LI Yaohui and LIU Baojun. The image enhancement based on histogram equalization[J]. *Journal of North China Institute of Science and Technology*, 2003, 5(2): 65–67.
- [10] HITAM M S, AWALLUDIN E A, and YUSSOF W. Mixture contrast limited adaptive histogram equalization for underwater image enhancement[C]. International Conference on Computer Application Technology, Sousse, Tunisia, 2013: 1–5.
- [11] 陈宇, 霍富荣, 苗华. 对比度拉伸在目标探测与识别中的应用研究[J]. 仪器仪表学报, 2008, 29(4): 795–798.
- CHEN Yu, HUO Furong, and MIAO Hua. Application of contrast stretching in optical correlation detection and recognition[J]. *Chinese Journal of Scientific Instrument*, 2008, 29(4): 795–798.
- [12] 杨勇, 郭玲, 王天江. 基于多尺度结构张量的多类无监督彩色纹理图像分割方法[J]. 计算机辅助设计与图形学学报, 2014, 26(5): 812–825.
- YANG Yong, GUO Ling, and WANG Tianjiang. Multi-scale structure tensor based unsupervised color-texture image segmentation approach in multiclass[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2014, 26(5): 812–825.
- [13] 蒋刚毅, 黄大江, 王旭, 等. 图像质量评价方法研究进展[J]. 电子与信息学报, 2010, 32(1): 219–226. doi: [10.3724/SP.J.1146.2009.00091](https://doi.org/10.3724/SP.J.1146.2009.00091).
- JIANG Gangyi, HUANG Dajiang, WANG Xu, et al. Overview on image quality assessment methods[J]. *Journal of Electronics & Information Technology*, 2010, 32(1): 219–226. doi: [10.3724/SP.J.1146.2009.00091](https://doi.org/10.3724/SP.J.1146.2009.00091).
- [14] EMBERTON S, CHITTKA L, and CAVALLARO A. Hierarchical rank-based veiling light estimation for underwater dehazing[C]. British Machine Vision Conference, Swansea, UK, 2015: 125.1–125.12.
- [15] JIAN Muwei, QI Qiang, DONG Junyu, et al. The OUC-Vision large-scale underwater image database[C]. IEEE International Conference on Multimedia & Expo, Hong Kong, China, 2017: 1297–1302.
- [16] JIAN Muwei, QI Qiang, DONG Junyu, et al. Saliency detection using quaternionic distance based weber local descriptor and level priors[J]. *Multimedia Tools and Applications*, 2018, 77(11): 14343–14360. doi: [10.1007/s11042-017-5032-z](https://doi.org/10.1007/s11042-017-5032-z).
- 刘 浩: 男, 1977年生, 副教授, 研究方向为图像处理与通信.
- 孙晓帆: 女, 1993年生, 硕士生, 研究方向为水下图像增强及质量评价.
- 张鑫生: 男, 1992年生, 硕士生, 研究方向为压缩域图像分析.
- 吴乐明: 男, 1993年生, 硕士生, 研究方向为图像集压缩及评价.
- 况奇刚: 男, 1994年生, 硕士生, 研究方向为图像超分辨率.