https://doi.org/10.2352/issn.2169-2629.2019.27.71

Future Directions in Image Quality

Seyed Ali Amirshahi[*†], *Marius Pedersen*[*],
[*] **The Norwegian Colour and Visual Computing Laboratory, Norwegian University of Science and Technology, Gjøvik, Norway**
[†] **L2TI, Universite Paris 13, Sorbonne Paris Cite, Villetaneuse France**
s.ali.amirshahi@ntnu.no, marius.pedersen@ntnu.no

Abstract

With the advancements made in the field of image processing and computer vision, the last few decades have seen an increase in studies focused on image quality assessment. While this has resulted in the introduction of different new metrics which some show high correlation with the perceptual judgement of the human observers there still exists a huge room for improvement. In this short paper which is prepared as a complement to the workshop on Future Directions in Image Quality at CIC 27 in Paris, France we aim to introduce future directions in the field and challenges facing ahead.

Introduction

Image quality has been an active field of research for decades [1]. Traditionally, subjective quality assessment has been the preferred method for estimating quality. However, this is resource demanding and expensive which has resulted in objective quality assessment methods becoming the preferred method of choice. Objective methods, commonly referred to as quality metrics, have the goal of predicting the quality of images without asking human observers. These metrics are classified based on their use of the reference image; Full Reference metrics (FR) use the reference and the test images, Reduced Reference (RR) metrics use the test and partial information about the reference images, and No Reference (NR) metrics are only based on the use of the test images [2]. The number of image quality metrics proposed in the literature is very high [3], and a number of studies have been carried out to evaluate the performance of such metrics against subjective data [4, 5, 6, 7, 3, 8]. In recent years a number of datasets which provide standard subjective scores for the evaluation of image quality metrics have been introduced [9, 10, 11, 12].

Despite the extensive research on image quality assessment, there is still many unsolved challenges. In this paper we will address open challenges and future direction of research on image quality assessment. These challenges will be addressed in the following order:

- Image enhancement and issues on quality assessment
- Recent trends in machine learning for image quality
- Image quality in biometrics
- Quality assessment of XR applications
- Color image quality assessment in smartphones
- Future of subjective evaluation and crowdsourcing
- Challenges in medical image quality
- The future of video quality metrics

Image enhancement and issues on quality assessment

The rapid advancements in the imaging industry has resulted in the introduction of different imaging systems. Image enhancement is a crucial part of any current and new imaging systems where the goal is to increase the quality of the image through different image processing techniques. Over the years different techniques for image enhancement, including contrast enhancement [13], denoising [14], sharpening [15], color enhancement [16, 17], and so on have been proposed. Techniques for enhancement can be general or for specific applications, such as medical imaging [18, 19, 20], biometrics [21, 22], printing [23], displays [24], image acquisition [25], video conferencing [26], video [27] and, etc.

A challenge in image enhancement is to evaluate the performance of the enhancement techniques. The number of techniques proposed are huge, making the selection of the methods one uses in the evaluation difficult. To evaluate the performance of image enhancement techniques a limited number of different options are available.

1. Compare such techniques against standard techniques.
2. Subjective evaluation.
3. Objective evaluation using image quality metrics. This approach is mainly used for this task which different challenges.

Since most standard techniques are outdated, such standards are not necessarily the go to methods for image enhancement. Using subjective evaluations are time and resource demanding, which results in some cases to include few images and few observers, making it difficult to draw strong conclusions. Keeping in mind the drawbacks of the first two approaches, in recent years image quality metrics have been used to evaluate different image enhancement techniques.

Keeping in mind that most image quality metrics are designed on the assumption that the reference image has always a better quality than the test image, there is still a challenge when it comes to having image quality metrics performing well for image enhancement [28, 29]. Simply said, when it comes to image enhancement and its relationship with image quality assessment,

- there is a need for new image quality metrics specifically designed for image enhancement techniques.
- there is a need for new datasets containing enhanced images in order to evaluate the performance of image quality metrics. Keeping in mind this need, in recent years a number of datasets have tried to answer this issue [30, 31, 32].

Future of subjective evaluation and crowdsourcing

Collecting subjective quality scores in controlled settings in a lab is time and resource demanding. Such limits results in datasets that are usually small in size and so have limited variability in image content, lack a divers set of distortions, are limited to few observers, and naturally have an insufficient number of subjective scores. Due to such issues collecting online crowdsourced datasets have always been an attractive option. In recent years datasets such as the LIVE In the Wild Image Quality Challenge Database [33], KonIQ-10k [34], and KADID-10k [35] datasets which compared to previous subjective datasets are large in numbers have been introduced. It is important to point out that through such large datasets it is now possible to train different machine learning techniques for image quality assessment.

Keeping in mind the progress made in the field, in the future, different issues such as the following need to be answered:

- reliability of the crowd and how to deal with unwanted behavior of the observers.
- how to cover a wide range of content and ensure the diversity of the dataset, while at the same time keeping its uniform.
- taking into account different viewing conditions, display devices, etc. between observers.
- introduction of recommendations and standards for crowdsourcing in the field of image quality assessment which could lead to the development of online platforms that are specifically aimed for such tasks.

Recent trends in machine learning for image quality assessment

With the introduction of Convolutional Neural Networks (CNNs) and various machine learning techniques different computer vision and image processing tasks have shown a dramatically better performance. While few works in the field of image enhancement [36] and assessment [37, 8] have taken advantage of CNNs, most image and video quality metrics are still focused on using a limited number of handcrafted features [38, 39]. Due to the lack of any large subjective dataset in the field, most image quality metrics based on machine learning are mostly focused on the use of transfer-learning. With the increase in crowdsourcing for image quality [40, 33], it is very likely that in the near future image quality metrics will be trained from scratch which based on the use of CNNs in other fields it is plausible that it would result in better performing image quality metrics. While the use of such machine learning techniques in the field are unavoidable, careful attention should be paid to better understand how such networks work and find links to the human visual system.

Challenges in medical image quality

With the increase in the use of different medical imaging devices, evaluating the quality of medical images has shown to be a critical issues. For example, evaluating the quality of images from different locations in the brain and cardiac MRI images [41] or to compare different CT reconstruction techniques [42]. In such specific medical applications, the quality assessment needs to be tuned specifically for each application, which requires a number of methods to be developed. However, some studies have tried to evaluate the performance of traditional image quality metrics on medical images. For example, in a study, the BRISQUE image quality metric [43] was used to evaluate and optimize the quality of capsule video endoscopy [44]. It is important to point out that compared to other image quality evaluation tasks, when it comes to evaluating the quality of medical images, due to the lack of a reference image we only focus on NR metrics [45] which makes such quality assessments more challenging.

Keeping in mind the progress made in evaluating the quality of medical images we could mention the following issues as the future challenges faced in the field:

- a real time image quality assessment for medical images is needed in order to optimize image acquisition and post-processing.
- due to the difference in nature and content of various medical imaging modalities introducing a universal image quality metric would be a huge step in the field.

Color image quality assessment in smartphones

With the huge advancements made in designing mobile phones and the high quality photographs they can now capture the sales of traditional photographic cameras have seen a decline. Add how social media and online social networks have taken over our daily life and you could see a shift from highly aesthetic images to images that are visually pleasing. Such images are mostly produced using different image processing techniques. In fact, a study by Araújo et al. indicated that a majority of images shared on Instagram have been gone through some kind of post processing [46]. It is clear that introducing an objective image quality metric to assess these very subjective aspects are difficult. Add the fact that such images are normally displayed on smartphone devices and not traditional large displays or printed media and you can see how these factors could affect the overall subjective judgment of the image quality and how challenging it would be to propose an objective metric for such cases.

Image quality in biometrics

Biometric systems which are used widely in our daily life are influenced by the quality of acquired images [47, 48]. Different studies have pointed out that low quality images are one of the main reasons for why such systems fail [49]. To prevent such issues different standards have been proposed [50] and various image quality issues in biometrics different studies have been produced.

Liu et al. [47] evaluated the performance of general purpose image quality metrics for face biometrics. The work found that such metrics could be suitable to measure quality in contactless-face biometrics. In a similar study Liu et al. [51] showed that the same image quality metrics are not able to predict system performance in an iris biometric system. This is most likely because iris images are different in nature to natural images which the metrics were initially designed for. With the introduction of multimodal biometric systems which commonly include both iris, face and/or fingerprints for example in smartphones and tablets, there has also been a push to find multi-modality assessment methods for biometrics [52]. However, due to the different characteristics of these images, this is a challenging task. In addition, since biometric matching is performed in real-time, any image quality metric used in biometric tasks should also perform in real-time, or close to real-time. While Jenadeleh et al. [53] introduced a real-time image quality metric for iris images, there is still room for improvement.

Quality assessment of XR applications

Studies have shown that the observers focus plays an important role in evaluating the quality of 2D and 3D images [54]. Compared to 2D images, the content of the image has a higher impact in evaluating the quality of 3D images [55]. Naturally, the image content will play an influential role when it comes to evaluating the quality of XR applications and so traditional image quality metrics designed for 2D images will not perform well for such a task.

An important challenge in this field of studies is the lack of reliable subjective assessment methodologies for 3D content [55], which in turn influences the evaluation of quality metrics. Add the complexities in XR applications which do not exist in

2D content and image quality assessment of XR applications are one of the most challenging task in the field of image processing.

The future of video quality metrics

Video systems are more complex than still image systems, and the video systems have different components including capture and display hardware, multiplexers, converters, codecs, routers, streamers, and switches which can potentially have an affect on the video quality [56]. To evaluate the quality of videos, traditionally MSE and PSNR have been used which are only loosely correlated with perceived video quality [56].

Different studies have pointed out that when it comes to evaluating the quality of videos different issues such as attention [57] and content [58] play an important role. While in recent years different subjective datasets have been introduced [59, 60], there still exist a huge room for improvement.

Conclusion

In this short paper which is prepared as a complement to the workshop on Future Directions in Image Quality at CIC 27 in Paris, France we provided a short overview of different challenges facing the field of image quality assessment. We specifically focused on eight different applications in the field. While traditional image quality metrics could be used in all the mentioned application, we have pointed out that in order to reach better results and improve the performance of the metrics each application should be treated separately. Simply said, in each case the specific conditions and characteristics of the images used in the application should be taken into account.

Acknowledgements

This work has received funding from The Research Council of Norway through a FRIPRO Mobility Grant, contract no 250653/F20. The FRIPRO Mobility grant scheme (FRICON) is co-funded by the European Unions Seventh Framework Programme for research, technological development and demonstration under Marie Curie grant agreement no 608695.

This work was also supported by the Research Council of Norway through project no. 247689. IQ-MED: Image Quality enhancement in MEDical diagnosis, monitoring and treatment.

References

[1] Farah Torkamani-Azar and Seyed Ali Amirshahi. A new approach for image quality assessment using svd. In *2007 9th International Symposium on Signal Processing and Its Applications*, pages 1–4. IEEE, 2007.

[2] Zhou Wang and Alan C Bovik. Modern image quality assessment. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 2(1):1–156, 2006.

[3] Marius Pedersen and Jon Yngve Hardeberg. Full-reference image quality metrics: Classification and evaluation. *Foundations and Trends® in Computer Graphics and Vision*, 7(1):1–80, 2012.

[4] Hamid R Sheikh, Muhammad F Sabir, and Alan C Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on image processing*, 15(11):3440–3451, 2006.

[5] Nikolay Ponomarenko, Federica Battisti, Karen Egiazarian, Jaakko Astola, and Vladimir Lukin. Metrics performance comparison for color image database. In *Fourth international workshop on video processing and quality metrics for consumer electronics*, volume 27, page 6, 2009.

[6] Atidel Lahoulou, Ahmed Bouridane, Emmanuel Viennet, and Mourad Haddadi. Full-reference image quality metrics performance evaluation over image quality databases. *Arabian Journal for Science and Engineering*, 38(9):2327–2356, 2013.

[7] Marius Pedersen. Evaluation of 60 full-reference image quality metrics on the CID:IQ. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 1588–1592. IEEE, 2015.

[8] Seyed Ali Amirshahi, Marius Pedersen, and Azeddine Beghdadi. Reviving traditional image quality metrics using CNNs. In *Color and Imaging Conference*, volume 2018, pages 241–246. Society for Imaging Science and Technology, 2018.

[9] Nikolay Ponomarenko, Vladimir Lukin, Alexander Zelensky, Karen Egiazarian, Marco Carli, and Federica Battisti. Tid2008-a database for evaluation of full-reference visual quality assessment metrics. *Advances of Modern Radioelectronics*, 10(4):30–45, 2009.

[10] Xinwei Liu, Marius Pedersen, and Jon Yngve Hardeberg. CID:IQ–a new image quality database. In *International Conference on Image and Signal Processing*, pages 193–202. Springer, 2014.

[11] Patrick Le Callet and Florent Autrusseau. Subjective quality assessment irccyn/ivc database. 2005.

[12] HR Sheikh. Live image quality assessment database release 2. *http://live. ece. utexas. edu/research/quality*, 2005.

[13] Azeddine Beghdadi and Alain Le Negrate. Contrast enhancement technique based on local detection of edges. *Computer Vision, Graphics, and Image Processing*, 46(2):162–174, 1989.

[14] Guy Gilboa, Nir Sochen, and Yehoshua Y Zeevi. Image enhancement and denoising by complex diffusion processes. *IEEE transactions on pattern analysis and machine intelligence*, 26(8):1020–1036, 2004.

[15] Andrea Polesel, Giovanni Ramponi, and V John Mathews. Image enhancement via adaptive unsharp masking. *IEEE transactions on image processing*, 9(3):505–510, 2000.

[16] Shaohua Chen and Azeddine Beghdadi. Natural enhancement of color image. *EURASIP Journal on Image and Video Processing*, 2010(1):175203, 2010.

[17] Gabriele Simone, Giuseppe Audino, Ivar Farup, Fritz Albregtsen, and Alessandro Rizzi. Termite retinex: a new implementation based on a colony of intelligent agents. *Journal of electronic imaging*, 23(1):013006, 2014.

[18] Ahmed Mohammed, Ivar Farup, Marius Pedersen, Øistein Hovde, and Sule Yildirim Yayilgan. Stochastic capsule endoscopy image enhancement. *Journal of Imaging*, 4(6):75, 2018.

[19] Bilel Sdiri, Azeddine Beghdadi, Faouzi Alaya Cheikh, Marius Pedersen, and Ole Jakob Elle. An adaptive contrast enhancement method for stereo endoscopic images combining binocular just noticeable difference model and depth information. *Electronic Imaging*, 2016(13):1–7, 2016.

[20] Olga Cherepkova and Jon Yngve Hardeberg. Enhancing dermoscopy images to improve melanoma detection. In *2018 Colour and Visual Computing Symposium (CVCS)*, pages 1–6. IEEE, 2018.

[21] Lin Hong, Yifei Wan, and Anil Jain. Fingerprint image enhancement: algorithm and performance evaluation. *IEEE transactions on pattern analysis and machine intelligence*, 20(8):777–789, 1998.

[22] Xinwei Liu, Marius Pedersen, Christophe Charrier, and Patrick Bours. Can image quality enhancement meth-

ods improve the performance of biometric systems for degraded face images? In *2018 Colour and Visual Computing Symposium (CVCS)*, pages 1–5. IEEE, 2018.

[23] Wen-Chung Kao, Jia-An Ye, Ming-I Chu, and Chung-Yen Su. Image quality improvement for electrophoretic displays by combining contrast enhancement and halftoning techniques. *IEEE Transactions on Consumer Electronics*, 55(1):15–19, 2009.

[24] Pei-Shan Tsai, Chia-Kai Liang, Tai-Hsiang Huang, and Homer H Chen. Image enhancement for backlight-scaled tft-lcd displays. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(4):574–583, 2009.

[25] Angelo Bosco, Sebastiano Battiato, Arcangelo Bruna, and Rosetta Rizzo. Noise reduction for cfa image sensors exploiting hvs behaviour. *Sensors*, 9(3):1692–1713, 2009.

[26] Carlos Andres Arango Duque, Mekides Assefa Abebe, Muhammad Shahid, and Jon Yngve Hardeberg. Color and quality enhancement of videoconferencing whiteboards. *Electronic Imaging*, 2018(16):010504–1, 2018.

[27] Qing Wang and Rabab K Ward. Fast image/video contrast enhancement based on weighted thresholded histogram equalization. *IEEE transactions on Consumer Electronics*, 53(2):757–764, 2007.

[28] Damon M Chandler. Seven challenges in image quality assessment: past, present, and future research. *ISRN Signal Processing*, 2013, 2013.

[29] Yao Cheng, Marius Pedersen, and Guangxue Chen. Evaluation of image quality metrics for sharpness enhancement. In *Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis*, pages 115–120. IEEE, 2017.

[30] Azeddine Beghdadi, Muhammad Ali Qureshi, Bilel Sdiri, Mohamed Deriche, and Faouzi Alaya-Cheikh. Ceed-a database for image contrast enhancement evaluation. In *2018 Colour and Visual Computing Symposium (CVCS)*, pages 1–6. IEEE, 2018.

[31] Cuong T Vu, Thien D Phan, Punit S Banga, and Damon M Chandler. On the quality assessment of enhanced images: A database, analysis, and strategies for augmenting existing methods. In *2012 IEEE Southwest Symposium on Image Analysis and Interpretation*, pages 181–184, 2012.

[32] Muhammad Ali Qureshi, Azeddine Beghdadi, Bilel Sdiri, Mohamed Deriche, and Faouzi Alaya-Cheikh. A comprehensive performance evaluation of objective quality metrics for contrast enhancement techniques. In *EUVIP*, pages 1–5. IEEE, 2016.

[33] Deepti Ghadiyaram and Alan C Bovik. Massive online crowdsourced study of subjective and objective picture quality. *IEEE Transactions on Image Processing*, 25(1):372–387, 2015.

[34] Hanhe Lin, Vlad Hosu, and Dietmar Saupe. KonIQ-10K: Towards an ecologically valid and large-scale IQA database, 2018.

[35] Hanhe Lin, Vlad Hosu, and Dietmar Saupe. Kadid-10k: A large-scale artificially distorted iqa database. In *2019 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–3. IEEE, 2019.

[36] Xin Zhang and Ruiyuan Wu. Fast depth image denoising and enhancement using a deep convolutional network. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2499–2503. IEEE, 2016.

[37] Seyed Ali Amirshahi, Marius Pedersen, and Stella X Yu. Image quality assessment by comparing CNN features between images. *Journal of Imaging Science and Technology*, 60(6):60410–1, 2016.

[38] Seyed Ali Amirshahi and M-C Larabi. Spatial-temporal video quality metric based on an estimation of qoe. In *2011 Third International Workshop on Quality of Multimedia Experience*, pages 84–89. IEEE, 2011.

[39] Seyed Ali Amirshahi. Towards a perceptual metric for video quality assessment. Master's thesis, Norwegian University of Sciecne and Technology, 2010.

[40] Flávio Ribeiro, Dinei Florencio, and Vítor Nascimento. Crowdsourcing subjective image quality evaluation. In *2011 18th IEEE International Conference on Image Processing*, pages 3097–3100. IEEE, 2011.

[41] Michael Osadebey, Marius Pedersen, Douglas Arnold, and Katrina Wendel-Mitoraj. Image quality evaluation in clinical research: A case study on brain and cardiac mri images in multi-center clinical trials. *IEEE journal of translational engineering in health and medicine*, 6:1–15, 2018.

[42] Kristin Jensen, Anne Catrine T Martinsen, Anders Tingberg, Trond Mogens Aaløkken, and Erik Fosse. Comparing five different iterative reconstruction algorithms for computed tomography in an roc study. *European radiology*, 24(12):2989–3002, 2014.

[43] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *IEEE Transactions on image processing*, 21(12):4695–4708, 2012.

[44] Marius Pedersen, Olga Cherepkova, and Ahmed Mohammed. Image quality metrics for the evaluation and optimization of capsule video endoscopy enhancement techniques. *Journal of Imaging Science and Technology*, 61(4):40402–1, 2017.

[45] Li Sze Chow and Raveendran Paramesran. Review of medical image quality assessment. *Biomedical signal processing and control*, 27:145–154, 2016.

[46] Camila Souza Araújo, Luiz Paulo Damilton Corrêa, Ana Paula Couto da Silva, Raquel Oliveira Prates, and Wagner Meira. It is not just a picture: revealing some user practices in instagram. In *2014 9th Latin American Web Congress*, pages 19–23. IEEE, 2014.

[47] Xinwei Liu, Marius Pedersen, Christophe M Charrier, and Patrick Bours. Performance evaluation of no-reference image quality metrics for face biometric images. *Journal of Electronic Imaging*, 27(2):023001, 2018.

[48] Fernando Alonso-Fernandez, Julian Fierrez, and Javier Ortega-Garcia. Quality measures in biometric systems. *IEEE Security & Privacy*, 10(6):52–62, 2011.

[49] Xiufeng Gao, Stan Z Li, Rong Liu, and Peiren Zhang. Standardization of face image sample quality. In *International Conference on Biometrics*, pages 242–251. Springer, 2007.

[50] International Organization for Standardization (ISO). Information technologybiometric sample qualitypart 5: face image data, ISO/IEC TR 29794-5, International Electrotechnical Commission (IEC), 2010.

[51] Xinwei Liu, Marius Pedersen, Christophe Charrier, and Patrick Bours. Can no-reference image quality metrics assess visible wavelength iris sample quality? In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3530–3534. IEEE, 2017.

[52] Xinwei Liu, Marius Pedersen, and Christophe Charrier. Image-based attributes of multi-modality image quality for contactless biometric samples. In *2016 3rd International*

Conference on Signal Processing and Integrated Networks (SPIN), pages 106–111. IEEE, 2016.

[53] Mohsen Jenadeleh, Marius Pedersen, and Dietmar Saupe. Realtime quality assessment of iris biometrics under visible light. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 443–452, 2018.

[54] Olivier Le Meur and Patrick Le Callet. What we see is most likely to be what matters: Visual attention and applications. In *2009 16th IEEE International Conference on Image Processing (ICIP)*, pages 3085–3088. IEEE, 2009.

[55] Quan Huynh-Thu, Patrick Le Callet, and Marcus Barkowsky. Video quality assessment: From 2d to 3dchallenges and future trends. In *2010 IEEE International Conference on Image Processing*, pages 4025–4028. IEEE, 2010.

[56] Stefan Winkler and Praveen Mohandas. The evolution of video quality measurement: From psnr to hybrid metrics. *IEEE Transactions on Broadcasting*, 54(3):660–668, 2008.

[57] Junyong You, Touradj Ebrahimi, and Andrew Perkis. Attention driven foveated video quality assessment. *IEEE Transactions on Image Processing*, 23(1):200–213, 2013.

[58] Shyamprasad Chikkerur, Vijay Sundaram, Martin Reisslein, and Lina J Karam. Objective video quality assessment methods: A classification, review, and performance comparison. *IEEE transactions on broadcasting*, 57(2):165–182, 2011.

[59] Christos George Bampis, Zhi Li, Anush Krishna Moorthy, Ioannis Katsavounidis, Anne Aaron, and Alan Conrad Bovik. Study of temporal effects on subjective video quality of experience. *IEEE Transactions on Image Processing*, 26(11):5217–5231, 2017.

[60] Vlad Hosu, Franz Hahn, Mohsen Jenadeleh, Hanhe Lin, Hui Men, Tamás Szirányi, Shujun Li, and Dietmar Saupe. The konstanz natural video database (konvid-1k). In *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–6. IEEE, 2017.