https://doi.org/10.2352/issn.2169-2629.2021.29.258

Colourlab Image Database: Geometric Distortions

Marius Pedersen and Seyed Ali Amirshahi
The Norwegian Colour and Visual Computing Laboratory, Norwegian University of Science and Technology, Gjøvik, Norway
marius.pedersen@ntnu.no, s.ali.amirshahi@ntnu.no

Abstract

Over the years, a high number of different objective image quality metrics have been proposed. While some image quality metrics show a high correlation with subjective scores provided in different datasets, there still exists room for improvement. Different studies have pointed to evaluating the quality of images affected by geometrical distortions as a challenge for current image quality metrics. In this work, we introduce the Colourlab Image Database: Geometric Distortions (CID:GD) with 49 different reference images made specifically to evaluate image quality metrics. CID:GD is one of the first datasets which include three different types of geometrical distortions; seam carving, lens distortion, and image rotation. 35 state-of-the-art image quality metrics are tested on this dataset, showing that apart from a handful of these objective metrics, most are not able to show a high performance. The dataset is available at `www.colourlab.no/cid`.

Introduction

Measuring perceived image quality is important in many imaging applications. For this task, both subjective and objective techniques have been proposed. A plethora of objective Image Quality Metrics (IQMs) have been proposed in the literature [1], with the goal to complement or replace subjective assessment. These have been suggested for a number of applications, such as printing [2, 3], gamut mapping [4], tone mapping [5], and 360 degree images [6]. To evaluate the performance of IQMs, datasets containing distorted images and their respective subjective quality scores have been introduced [7, 8, 9, 10]. IQMs have shown high accuracy [1, 11] in most datasets with common spatial distortions such as blur and compression artifacts.

It is well known that compared to spatial distortions, when it comes to predicting the quality of images affected by geometric distortions such as translation, scaling, and rotation current IQMs show a rather low accuracy. Being able to handle such distortions has been stated as one of the main challenges facing IQMs [12]. While geometric distortions, when small, usually have little influence on the perceived quality of images they can have a large impact on IQMs. This is more evident in the case of those metrics that are pixel based. Naturally, such difference between the subjective and objective evaluation is a good evidence of how IQMs show a low performance in the case of images affected by geometrical distortions. Works such as [13] have linked the performance of IQMs to the datasets they are trained and tested on. Since benchmark datasets such as the TID2013 [7], LIVE [8], CSIQ [9], and CID:IQ [10], do not include geometric distortions, making it difficult to evaluate and benchmark IQMs regarding this quality aspect. In this study, we address this issue by introducing the Colourlab Image Database: Geometric Distortions (CID:GD) which is one of the first subjective datasets in the field of image quality which is solely focused on geometric distortions. Our goal is to not only provide the research community with a unique and much needed dataset, but also to evaluate the performance of state-of-the-art IQMs on images which have been affected by geometrical distortions.

In the rest of the paper, we will first introduce the CID:GD dataset, the subjective experiment, and the IQMs evaluated in this study are introduced in the next sections followed by the results of the subjective experiment and the performance of the mentioned IQMs. Finally, a conclusion and the future direction of the work is presented.

Overview of the dataset

The CID:GD dataset is built on the CID:IQ dataset [10] and extends the number of images and distortions in that dataset. CID:GD consists of 49 reference images (Figure 1) of which 23 were from the CID:IQ dataset and 26 images were used for the first time in this dataset. Similar to the CID:IQ dataset, all images were cropped in a way that the images had a size of 800×800 pixels. To analyze and assess the characteristics of the reference images, we have followed the recommendations by Winkler [14]. That is, we have calculated the Spatial Information (SI),

$$SI = \sqrt{L/1080}\sqrt{\sum s_r^2/P}, \tag{1}$$

an indicator of edge energy and the colorfulness,

$$CF = \sqrt{\sigma_{rg}^2+\sigma_{by}^2}+0.3\sqrt{\mu_{rg}^2+\mu_{by}^2}, \tag{2}$$

an indicator for variety and intensity of colors in the reference images. In Eq. (1), L is the number of vertical lines (in our case the images have 800 lines), P is the number of pixels (in our case all images are of size 800×800 pixels), and s_r^2 is the grayscale image filtered horizontally and vertically with a Sobel kernel $s_r=\sqrt{s_v^2+s_h^2}$. In the case of colorfulness represented by Eq. (2), $rg=R-G$ and $by=0.5(R+G)-B$, being an opponent colorspace.

By plotting SI against CF values (Figure 2) we can see that the reference images in CID:GD span a wide area with a good distribution in the space. This indicates that the reference images have a broad distribution of content. Compared with other state-of-the-art datasets such as CID:IQ [10] (Figure 2(b)), Ajagamelle [15] (Figure 2(c)), TID [7] (Figure 2(d)), Dugay [16] (Figure 2(e)), CSIQ [9] (Figure 2(f)), VLCFER [17] (Figure 2(g)), IVC [18] (Figure 2(h)), LIVE [8] (Figure 2(i)), and LIVE MD [19] (Figure 2(j)) the CID:GD dataset has a good distribution.

For the 49 reference images in the CID:GD dataset, we introduce the following three different types of distortions; namely seam carving, lens distortion, and image rotation.

Seam carving [20] was developed for content-aware image resizing, where seams (or paths of pixels connected from top to bottom or left to right) through the image are found and used to extend or reduce the size of the image. Using the seam carving method, 7.5% and 15% of the pixels in both horizontal and vertical directions were first removed and then inserted, making the resulting image to have the same pixel size. Such a distortion can result in changes in the aspect ratio of objects in the

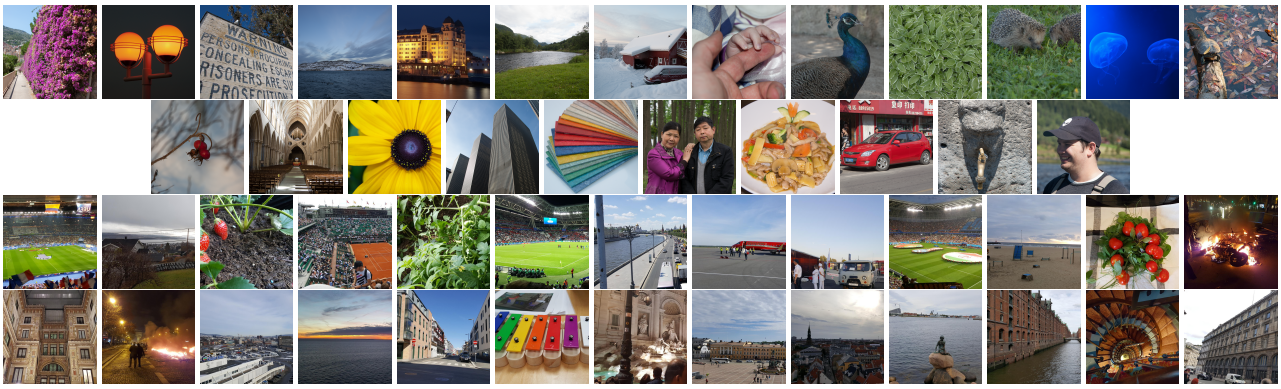
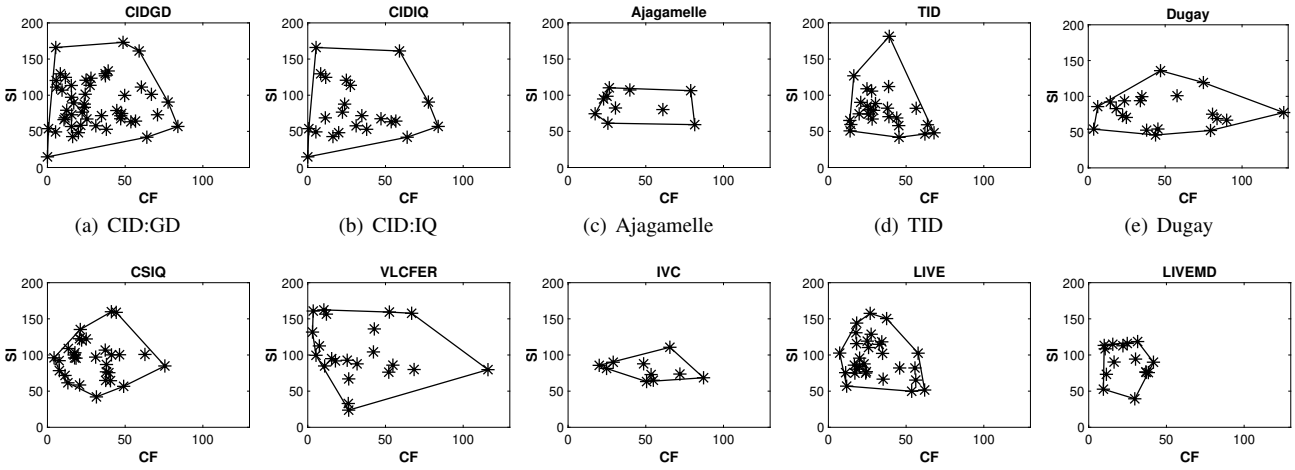


Figure 1. *Images in the CID:GD dataset. The two top rows show the images in the CID:IQ dataset while the bottom two rows correspond to images which are used for the first time in the CID:GD dataset.*



(a) CID:GD (b) CID:IQ (c) Ajagamelle (d) TID (e) Dugay

(f) CSIQ (g) VLCFER (h) IVC (i) LIVE (j) LIVE MD

Figure 2. *Spatial information (SI) plotted against colorfulness (CF) for reference images in the (a) CID:GD , (b) CID:IQ [10], (c) Ajagamelle [15], (d) TID [7], (e) Dugay [16], (f) CSIQ [9], (g) VLCFER [17], (h) IVC [18], (i) LIVE [8] and (j) LIVE MD [19] datasets.*

scene, remove content, and introduce contours in the image (Figure 3(b)). In our experiments, we refer to the distorted images of seam carving as "Seam Carving 7.5%" and "Seam Carving 15%". For the production of the new images, we used the implementation provided by Afifi [21] in Matlab.

Lens distortion was applied to produce four different images, each with different levels of lens distortion using Matlab. This includes pincushion distortion with a distortion parameter of -0.1 and -0.5, referred to as "Lens Distortion -0.1" and "Lens Distortion -0.5" respectively (Figure 3(c)) and barrel distortion with a distortion parameter of $+0.1$ and $+0.25$ which we refer to as "Lens Distortion $+0.1$" and "Lens Distortion $+0.25$" respectively (Figure 3(d)). The distorted images are the same size as the reference image.

Image rotation was applied to the images by +1 and -3 degrees referred to as "Rotation +1 degree", "Rotation -3 degree" respectively (Figure 3(e)) using Matlab. The pixels in the image which need to be replaced due to the nature of the rotation are replaced by the median pixel value in the image calculated before rotation. The images after rotation are the same size as the reference image, meaning that some content has been cropped.

The CID:GD dataset provides eight different distorted images per reference image, resulting in a total of 392 test images. All images having the same size as the original image allows for the dataset to also be used to evaluate full-reference IQMs.

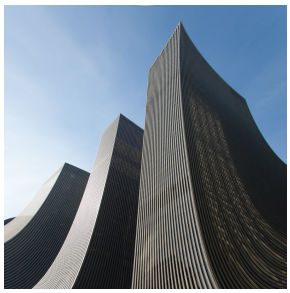
Subjective experiment

The subjective experiment was carried out as a single stimulus category judgement experiment using the web platform QuickEval [22]. During the experiment the observers rated the reference image along with the eight distorted images. Prior to the experiment, the observers were not informed about the type of distortions in the dataset. During the experiment the observers were asked to "Rate the quality of the image using a 5 point scale". The scale was "1-bad", "2-poor", "3-fair", "4-good", "5-excellent", following the ITU recommendations [23]. Due to the number of images in the dataset, the experiment was split into four different image sets three containing 110 images and one 110 to include all 441 images in the dataset (49 reference and 392 distorted images) selected randomly from the dataset. We should point out that the mentioned image sets do not share any images among each other. Observers were then given the chance to evaluate one or more sets in each observation session. Apart from randomly selected images for each image set, the images which were shown on a neutral (gray) background in each set were also presented in a random order to the observers so that the order the images are evaluated would not have an effect on the observer's evaluation.

33 different observers (20 males and 13 females) participated in the experiment. A minimum of 15 observers participated in each session, following the recommendation by CIE [24]. Five observers evaluated all four image sets. The judgements from the

(a) Reference image

(b) Seam Carving 15%



(c) Lens Distortion -0.5

(d) Lens Distortion +0.25

(e) Rotation -3 degrees

Figure 3. *Example of distorted images (a) Reference image, (b) Seam Carving 15%, (c) Lens Distortion -0.5, (d) Lens Distortion +0.25, and (e) Rotation -3 degrees.*

observers were processed into z-scores [25].

Objective metrics

In our experiments the performances of a wide range of different state of the art IQMs were evaluated on the CID:GD dataset. The IQMs selected were Structure Similarity Index Metric (SSIM) [26], Structural Content [27], PSNRHVSM [28], Multiscale SSIM (MSSIM) [29], Color Image Difference (CID) [30], Feature SIMilarity (FSIM) and the FSIMc (colour) [31], Edge Strength SIMilarity (ESSIM) [32], Convolutional Neural Networks Quality (CNNQ) [33], Spatial Difference of Gaussians (SDOG) [15], Spatial Hue Angle Metric (SHAMEII) [34], Saliency comparison (SAL) [35], Visual Saliency based Index (VSI) [36], Colorfulness-based Patch-based Contrast Quality Index (CPCQI) [37], High Order Statistics Aggregation (HOSA) [38], Blur [39], Blur Metric [40], Spectral and Spatial Sharpness (S3) [41], Cumulative Probability of Blur Detection (CPBD) [42], FRIQUEE [43], Blind Image Quality Assessment through Anisotropy (BIQAA) [44], Just Noticeable Blur Metric (JNBM) [45], and variations of the CNNQ metric (CNNQ-SSIM, CNNQ-PSNR, CNNQ-NAE, CNNQ-MSE, CNNQ-MAE and CNNQ-SC) [46]. This selection covers a wide range of IQMs, containing full-reference and no-reference IQMs, pixel-based and spatial IQMs, single-scale and multi-scale IQMs, etc. For a description of the IQMs we refer the reader to the respective publications.

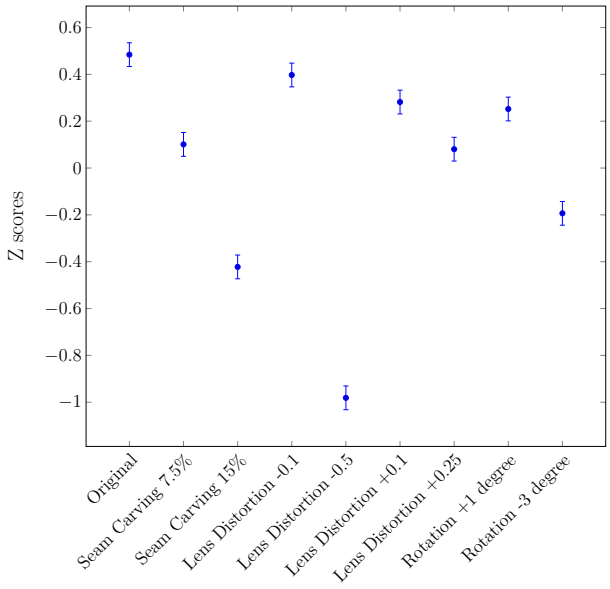


Figure 4. *Z-scores for different types of image distortions.*

To analyse the performance of the different IQMs we will calculate the non-linear Pearson correlation based on a 5-parameter logistic function [8]. The predicted scores were mapped to the subjective scores

$$Q = \beta_1 \left(\frac{1}{2} - \frac{1}{e^{-\beta_2(Q_p-\beta_3)}}\right) + \beta_4 Q_p + \beta_5 \tag{3}$$

where Q_p and Q are the predicted and the mapped scores and β_1–β_5 are fitting parameters. We will also calculate the correlation coefficients for each of the 49 different images, as this could reveal differences between the metrics for images with different content.

Results

In this section we provide an analysis of the subjective scores collected in our observation test followed by an in-depth study on the performance of different IQMs on the dataset.

Analysis of the subjective scores

By analysing the subjective scores, it is clear that the reference image has the highest z-score, but its 95% confidence interval overlaps with the test images with a -0.1 lens distortion (Figure 4). It is not surprising that the reference image has the highest z-score as it only contains natural distortions from the acquisition process. In the case of test images, the lowest z-score is for -0.5 lens distortion followed by the 15% seam carving distortion. In the case of -0.5 lens distortion, significant artifacts are introduced to the images, this is especially observed at the boundary were perspective changes are introduced. When it comes to 15% seam carving, some images have significant artifacts as well, this can be seen in images where lines that were straight became crooked. An example is the markings on the soccer field, which is known by the observers to be straight. With the -0.1 lens distortion most images have a higher positive z-score, but there are images where one can expect to see straight lines, and it is in these images that observers more easily notice the pincushion. In the images with +1 degree rotation, and for some with -3 degrees rotation, the content of the image could make detecting the rotation rather dif-

ficult. This is likely the reason this distortion has a higher z-score in certain images.

Analysis of the z-scores for each of the 49 images shows that in some images the 95% confidence interval overlaps between most distortions and the reference, indicating that the observers cannot differentiate between them. As an example, in images with larger uniform areas, such as the blue jellyfish image, the paths that are removed and added by seam carving are in the background and not introducing artifacts in the main objects, resulting in the observers not perceiving or finding the artifacts severe enough to differentiate. This is also an image where change in the shape of the jellyfish is difficult for observers to notice, which could be linked to the fact that jellyfish could have different shapes. Other distortions therefore introduce likely realistic content. In the case of other images, there is a much larger spread in the z-scores, indicating that the observers are able to differentiate between the distortions. These images have lines that the observers know should be straight, and contain known objects (such as a bus, people, houses). Having a known reference makes it easier for the observers to judge the severity of the geometrical distortion. It is clear from the subjective results that the content of the images plays an important role in the assessment carried out by the observers. The average standard deviation for the 49 different images per distortion is similar. However, we see some differences between the images. This for example can be seen in portrait images where some of the observers seem to be more sensitive to certain distortions.

Analysis of the objective scores

To analyse the performance of different IQMs non-linear Pearson correlation for the IQMs based on a 5-parameter logistic function was calculated (Figure 5). From the results we can observe that the CNNQ IQM [33], which although it is based on the use of CNNs which due to the use of pre-trained CNN models its computational time is fast and comparable with traditional IQMs, shows the best performance. It is interesting to observe that other variations of the CNNQ metric [46], which are more sensitive to changes in the geometrical properties of the image, have a slightly lower correlation value. We can also notice that in general, compared to no-reference IQMs, full-reference IQMs show a better performance. Keeping in mind that the reference images have been evaluated as having the highest quality overall, full-reference IQMs have a clear advantage. The no-reference IQM with the highest correlation rate of 0.10 is HOSA. Further analysis also reveals that some of the IQMs have problems with scale differences between images, which has been shown in the literature for other datasets [47].

We have also analyzed the correlation for each of the 49 images in the dataset (Figure 6). From the figure we can notice that the CNNQ metric in most images provides a high linear Pearson correlation. In 28 of the images the CNNQ-CONV5 has the highest linear Pearson correlation, while CNNQ-CONV4 has the highest in 9, and CNNQ-CONV1 in 4 images. We notice that in most cases the CNNQ IQM is able to correctly predict seam carving 15% or lens distortions -0.5 to be of the lowest quality. In comparison, in most cases (43 images) SSIM judges seam carving to have the lowest quality, the same can be seen in the case of iCID, SHAMEII, SDOG and other IQMs. This can be linked to the larger pixel displacement by seam carving compared to other distortions, but such a geometric displacement is not seen in the same way by the observers for most images. Other metrics such as the SAL IQM, which is based on saliency and not structural similarity or color difference, perform differently. In regard to ranking a specific distortion as having the lowest quality, no-reference IQMs are less consistent. Some IQMs have a larger interquartile range indicating that they are less consistent in their predictions. Our analysis of the Spearman correlation coefficients shows similar results.

The most difficult image, on average, for the IQMs to predict is reference image seven (winter scene with a barn and snow) and its distorted versions. This is one of the few images, where the reference image has not been rated as the best quality image resulting in a low performance in full-reference IQMs. The same can be said for reference images six (landscape) and eight (baby holding a hand). In reference image six, despite the fact that one has more changes than the other, the observers have judged the two seam carving images to be of equal quality. This is possibly due to the content of the image containing higher frequencies, making it challenging for the observers to detect the image artifacts. In general, observers have been able to detect changes in perspective or introduce artifacts resulting in perceived low quality. These geometric distortions are difficult to detect for IQMs as the metrics predict larger pixel displacements to have lower quality rather than being able to predict smaller pixel displacements, but introducing perspective or artifacts are perceptually more annoying.

Conclusions and Future Work

In this study we have introduced the Colourlab Image Database: Geometric Distortions (CID:GD) with 49 different reference images that include three different types of geometrical distortions; seam carving, lens distortion, and image rotation. The images have been evaluated by observers in a category judgement experiment. In addition, we have evaluated 35 state-of-the-art image quality metrics on the new dataset. Our evaluations show that most metrics have a low or average correlation, with the highest being around 0.67. The results indicate that the evaluation of geometric distortions is challenging for image quality metrics, and that this new dataset can be an important step in developing new image quality metrics for such types of distortions. The dataset is available at `www.colourlab.no/cid`.

Acknowledgment

In this work Seyed Ali Amirshahi has received funding from the Peder Sather Grant. Marius Pedersen has received funding from the Research Council of Norway (project number 324663).

References

[1] 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.

[2] Marius Pedersen, Nicolas Bonnier, Jon Y Hardeberg, and Fritz Albregtsen. Attributes of a new image quality model for color prints. In *Color and Imaging Conference*, volume 2009, pages 204–209. Society for Imaging Science and Technology, 2009.

[3] Altynay Kadyrova, Vlado Kitanovski, and Marius Pedersen. Quality assessment of 2.5d prints using 2d image quality metrics. *Applied Sciences*, 11(16), 2021. ISSN 2076-3417. doi: 10.3390/app11167470. URL `https://www.mdpi.com/2076-3417/11/16/7470`.

[4] Zofia Baranczuk, Peter Zolliker, and Joachim Giesen. Image quality measures for evaluating gamut mapping. In *Color and Imaging Conference*, volume 2009, pages 21–26. Society for Imaging Science and Technology, 2009.

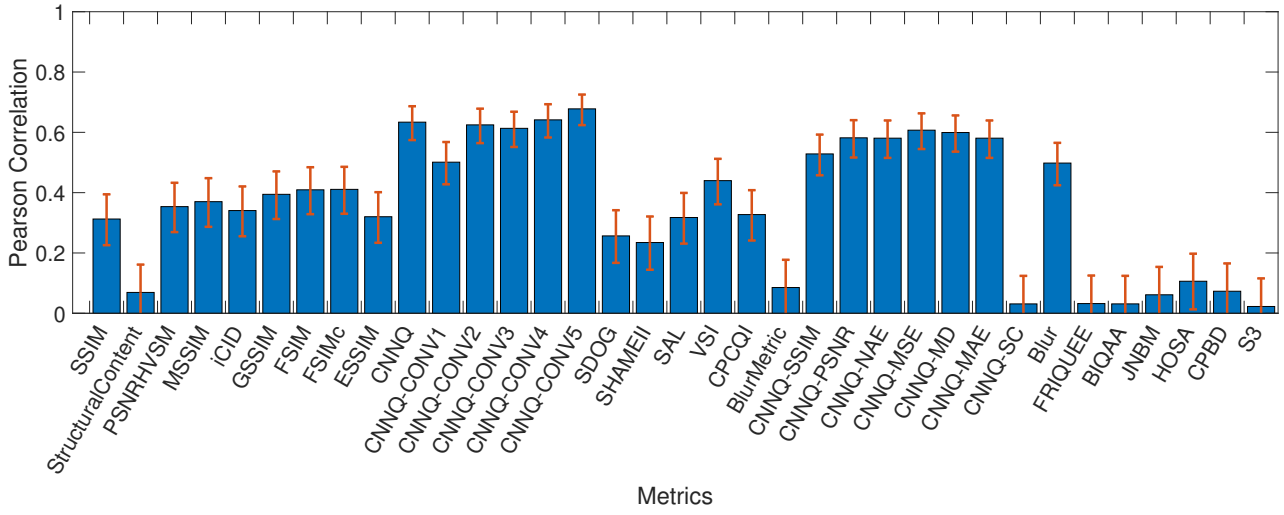


Figure 5. *Non-linear Pearson correlation values with a 95% confidence interval. The left most IQMs including CNNQ-SC are full-reference IQMS, while the right most IQMs from Blur are no-reference IQMs.*

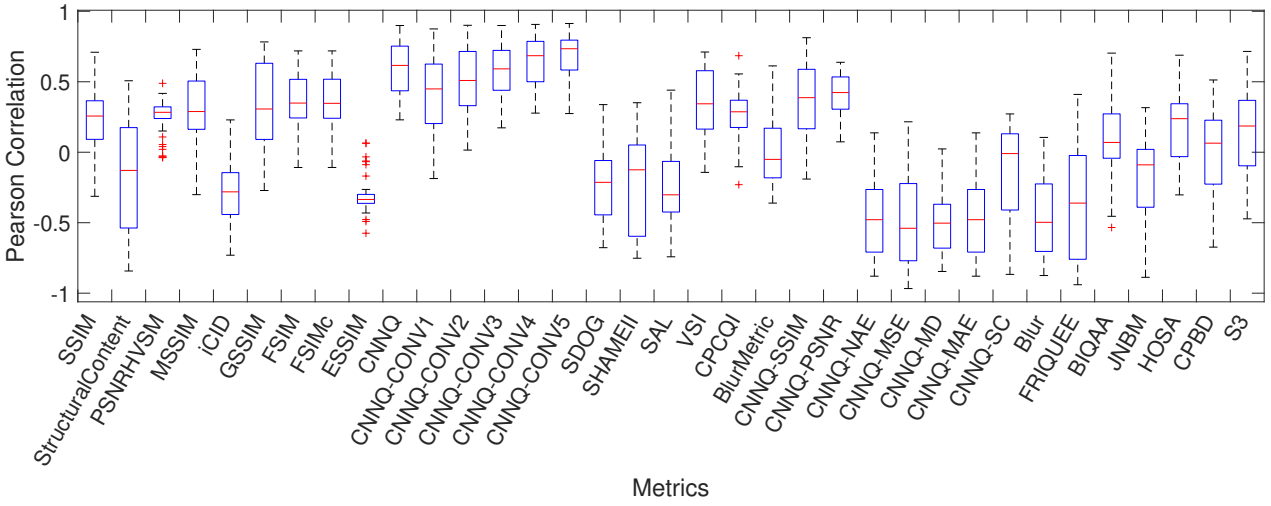


Figure 6. *Boxplot represents the Pearson correlation calculated for each image using the mentioned IQM. The left most IQMs including CNNQ-SC are full-reference IQMS, while the right most IQMs from Blur are no-reference IQMs.*

[5] Muhammad Usman Khan, Imran Mehmood, Ming Ronnier Luo, and Muhammad Farhan Mughal. No-reference image quality metric for tone-mapped images. In *Color and Imaging Conference*, volume 2019, pages 252–255. Society for Imaging Science and Technology, 2019.

[6] Mohamed-Chaker Larabi, Audrey Girard, Sami Jaballah, and Fan Yu. Benchmark of 2d quality metrics for the assessment of 360-deg images. In *Color and Imaging Conference*, volume 2019, pages 262–267. Society for Imaging Science and Technology, 2019.

[7] Nikolay Ponomarenko, Lina Jin, Oleg Ieremeiev, Vladimir Lukin, Karen Egiazarian, Jaakko Astola, Benoit Vozel, Kacem Chehdi, Marco Carli, Federica Battisti, et al. Image database tid2013: Peculiarities, results and perspectives. *Signal processing: Image communication*, 30:57–77, 2015.

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

[9] Eric Cooper Larson and Damon Michael Chandler. Most apparent distortion: full-reference image quality assessment and the role of strategy. *Journal of electronic imaging*, 19(1):011006, 2010.

[10] X. Liu, M. Pedersen, and J.Y. Hardeberg. CID:IQ–a new image quality database. In *International Conference on Image and Signal Processing*, pages 193–202. Springer, 2014.

[11] 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.

[12] D.M. Chandler, M.M. Alam, and T.D. Phan. Seven challenges for image quality research. In *Human Vision and Electronic Imaging XIX*, volume 9014, page 901402. International Society for Optics and Photonics, 2014.

[13] S.A. Amirshahi and M. Pedersen. Future directions in image quality. In *Color and Imaging Conference*, volume 2019, pages 399–403, 2019.

[14] Stefan Winkler. Analysis of public image and video databases for quality assessment. *IEEE Journal of Selected Topics in Signal Processing*, 6(6):616–625, 2012.

[15] Sebastien A Ajagamelle, Marius Pedersen, and Gabriele Si-

mone. Analysis of the difference of gaussians model in image difference metrics. In *Conference on Colour in Graphics, Imaging, and Vision*, volume 2010, pages 489–496. Society for Imaging Science and Technology, 2010.

[16] Fabienne Dugay, Ivar Farup, and Jon Y Hardeberg. Perceptual evaluation of color gamut mapping algorithms. *Color Research & Application*, 33(6):470–476, 2008.

[17] Andela Zarić, Nenad Tatalović, Nikolina Brajković, Hrvoje Hlevnjak, Matej Lončarić, Emil Dumić, and Sonja Grgić. VCL@FER image quality assessment database. *AUTOMATIKA: časopis za automatiku, mjerenje, elektroniku, računarstvo i komunikacije*, 53(4):344–354, 2012.

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

[19] Dinesh Jayaraman, Anish Mittal, Anush K Moorthy, and Alan C Bovik. Objective quality assessment of multiply distorted images. In *2012 Conference record of the forty sixth asilomar conference on signals, systems and computers (ASILOMAR)*, pages 1693–1697. IEEE, 2012.

[20] S. Avidan and A. Shamir. Seam carving for content-aware image resizing. In *ACM SIGGRAPH 2007 papers*, pages 10–es. 2007.

[21] M. Afifi - York University, Canada. Implementation of seam carving for content-aware image resizing, 2017.

[22] K. Van Ngo, J. J. Storvik, C. A. Dokkeberg, I. Farup, and M. Pedersen. Quickeval: a web application for psychometric scaling experiments. In *Image Quality and System Performance XII*, volume 9396, page 93960O, 2015.

[23] Telecommunication Standardization sector of ITU. Methods for the subjective assessment of video quality, audio quality and audiovisual quality of internet video and distribution quality television in any environment, 03 2013. Series P: Terminals and Subjective and Objective Assessment Methods. Recommendation ITU-T P.913.

[24] J. Morovic. Guidelines for the evaluation of gamut mapping algorithms. *Commision Internationale De L'Eclaire (CIE)*, 153:D8–6, 2003.

[25] P.G. Engeldrum. *Psychometric Scaling: A Toolkit for Imanging System Development*. Imcotek press, 2000.

[26] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4): 600–612, 2004.

[27] M Esckicioglu and P Fisher. Image quality measures and their performances. *IEEE Transactions on Communications*, 43(12):2959–2965, 1995.

[28] K. Egiazarian, J. Astola, N. Ponomarenko, V. Lukin, F. Battisti, and M. Carli. New full-reference quality metrics based on hvs. In *Proceedings of the Second International Workshop on Video Processing and Quality Metrics*, volume 4, 2006.

[29] Z. Wang, E.P. Simoncelli, and A.C. Bovik. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pages 1398–1402. Ieee, 2003.

[30] I. Lissner, J. Preiss, P. Urban, M.S. Lichtenauer, and P. Zolliker. Image-difference prediction: From grayscale to color. *IEEE Transactions on Image Processing*, 22(2):435–446, 2012.

[31] L. Zhang, L. Zhang, X. Mou, and D. Zhang. Fsim: A feature similarity index for image quality assessment. *IEEE transactions on Image Processing*, 20(8):2378–2386, 2011.

[32] X. Zhang, X. Feng, W. Wang, and W. Xue. Edge strength similarity for image quality assessment. *IEEE Signal processing letters*, 20(4):319–322, 2013.

[33] S.A. Amirshahi, M. Pedersen, and S.X. Yu. Image quality assessment by comparing cnn features between images. *Electronic Imaging*, (12):42–51, 2017.

[34] M. Pedersen and J.Y. Hardeberg. A new spatial filtering based image difference metric based on hue angle weighting. *Journal of Imaging Science and Technology*, 56(5): 50501–1, 2012.

[35] G. Cao, M. Pedersen, and Z. Barańczuk. Saliency models as gamut-mapping artifact detectors. In *Conference on Colour in Graphics, Imaging, and Vision*, volume 2010, pages 437–443, 2010.

[36] L. Zhang, Y. Shen, and H. Li. Vsi: A visual saliency-induced index for perceptual image quality assessment. *IEEE Transactions on Image processing*, 23(10):4270–4281, 2014.

[37] K. Gu, W. Lin, G. Zhai, X. Yang, W. Zhang, and C.W. Chen. No-reference quality metric of contrast-distorted images based on information maximization. *IEEE transactions on cybernetics*, 47(12):4559–4565, 2016.

[38] J. Xu, P. Ye, Q. Li, H. Du, Y. Liu, and D. Doermann. Blind image quality assessment based on high order statistics aggregation. *IEEE Transactions on Image Processing*, PP (99):1–1, 2016. ISSN 1057-7149. doi: 10.1109/TIP.2016.2585880.

[39] M. Elsayed, F. Sammani, A. Hamdi, A. Albaser, and H. Babalghoom. A new method for full reference image blur measure. *International Journal of Simulation: Systems, Science Technology*, 19:4, 2018.

[40] F. Crete, T. Dolmiere, P. Ladret, and M. Nicolas. The blur effect: perception and estimation with a new no-reference perceptual blur metric. In *Human vision and electronic imaging XII*, volume 6492, page 64920I, 2007.

[41] C.T. Vu, T.D. Phan, and D.M. Chandler. S3: A spectral and spatial measure of local perceived sharpness in natural images. *IEEE transactions on image processing*, 21(3): 934–945, 2011.

[42] N.D. Narvekar and L.J. Karam. A no-reference image blur metric based on the cumulative probability of blur detection (cpbd). *IEEE Transactions on Image Processing*, 20(9): 2678–2683, 2011.

[43] L. Liu, Y. Hua, Q. Zhao, H. Huang, and A.C. Bovik. Blind image quality assessment by relative gradient statistics and adaboosting neural network. *Signal Processing: Image Communication*, 40:1–15, 2016.

[44] S. Gabarda and G. Cristóbal. Blind image quality assessment through anisotropy. *JOSA A*, 24(12):B42–B51, 2007.

[45] R. Ferzli and L.J. Karam. A no-reference objective image sharpness metric based on the notion of just noticeable blur (jnb). *IEEE transactions on image processing*, 18(4):717–728, 2009.

[46] S.A. Amirshahi, M. Pedersen, and A. Beghdadi. Reviving traditional image quality metrics using CNNs. In *Color and imaging conference*, number 1, pages 241–246, 2018.

[47] M. Pedersen and I. Farup. Improving the robustness to image scale of the total variation of difference metric. In *3rd International Conference on Signal Processing and Integrated Networks*, pages 116–121. IEEE, 2016.