Quality Assessment of Images Undergoing Multiple Distortion Stages

UNIVERSITY OF WATERLOO

Shahrukh Athar, Abdul Rehman and Zhou Wang

{*shahrukh.athar, abdul.rehman, zhou.wang*}*@uwaterloo.ca, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1*

Introduction

- Objective Image Quality Assessment (IQA) methods are usually classified as *full-reference* (FR), *reduced-reference* (RR) and *no-reference* (NR) depending upon accessibility to pristine reference content (Figure 1).

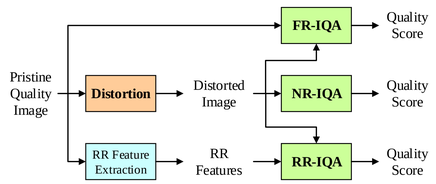


Figure 1: General framework of FR, RR and NR IQA.

- Typically IQA methods are tested and at times trained on image databases of different distortion types, but a single distortion stage.
- This is in clear contrast to real-world visual content distribution scenarios, where visual content may undergo multiple stages of distortions (Figure 2).

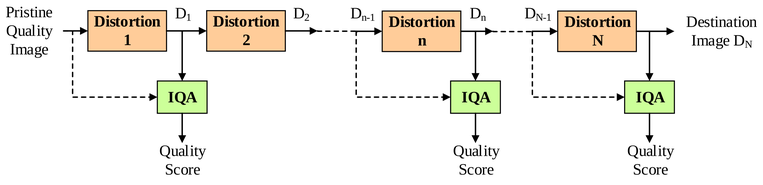


Figure 2: Quality tracking in visual content distribution with multiple stages of distortions.

- Objectives:
  - Understand the impact of multiple distortion stages on the performance of contemporary IQA models.
  - Investigate the potential of performing IQA at middle distortion stages.

References

1. A. Mittal, A. K. Moorthy and A. C. Bovik, "No reference image quality assessment in the spatial domain," *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, 2012.

2. P. Ye, J. Kumar, L. Kang and D. Doermann, "Unsupervised feature learning framework for no-reference image quality assessment," in *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2012, pp. 1098-1105.

3. Q. Wu, Z. Wang and H. Li, "A highly efficient method for blind image quality assessment," in *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2015, pp. 339-343.

4. A. Mittal, R. Soundararajan and A. C. Bovik, "Making a completely blind image quality analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209-212, 2013.

5. Z. Wang, H. R. Sheikh and A. C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images," in *2002 IEEE International Conference on Image Processing (ICIP)*, IEEE, 2002, vol. 1, pp. I477-I480.

6. H. R. Sheikh, Z. Wang, L. Cormack and A. C. Bovik, "LIVE image quality assessment database release 2," 2005.

7. E. C. Larson and D. M. Chandler, "Most apparent distortion: full-reference image quality assessment and the role of strategy," *Journal of Electronic Imaging*, vol. 19, no. 1, pp. 011006(1-21), 2010.

8 N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli and F. Battisti, "TID2008-a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, no. 4, pp. 30-45, 2009.

9. Z. Wang, E. P. Simoncelli and A. C Bovik, "Multiscale structural similarity for image quality assessment," in *Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers (ASILOMAR)*, IEEE, 2003, vol. 2, pp. 1398-1402.

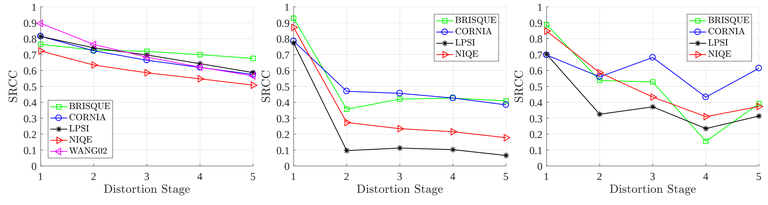
Performance Variations of IQA Models at Multiple Stages of Distortions

- We created the *IVC-MD5* database to analyze the performance of state-of-the-art NR IQA algorithms in multiple distortion scenarios (Table 1).

Table 1: Composition of the *IVC-MD5* Database.

Number of Pristine Images in Database				70
Distortion Stage	**Distortion Combination**			**Number of Images**
	1	**2**	**3**	
1	JPEG	Noise	Noise	210
2	JPEG	JPEG	JPEG2000	630
3	JPEG	JPEG	JPEG	1890
4	JPEG	JPEG	JPEG2000	5670
5	JPEG	JPEG	JPEG	17010

- Representative NR IQA algorithms that were tested include BRISQUE [1], CORNIA [2], LPSI [3], NIQE [4] and WANG02 [5].
- These NR IQA algorithms are tested and/or trained on image databases with a single distortion stage [6, 7, 8].
- FR IQA algorithm MS-SSIM [9] was used for benchmarking purposes.



(a) Distortion Combination-1 (b) Distortion Combination-2 (c) Distortion Combination-3

Figure 3: Performance of NR IQA algorithms for different distortion combinations of the *IVC-MD5* database.

- Conclusions:
  - NR IQA algorithm performance degrades consistently with distortion stages.
  - Performance degradation is most severe in case of mixed distortion types.

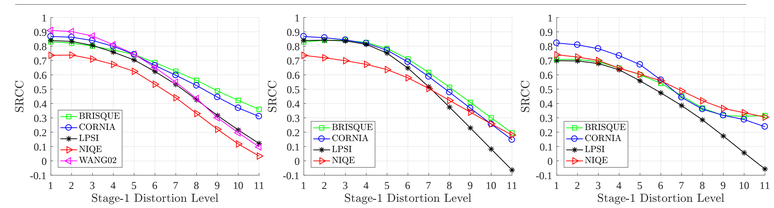
IQA at Mid-Stage

- We created a large-scale image dataset called the *IVC-MD-Te* database, that contains images which have undergone two distortion stages (Table 2).

Table 2: Composition of the *IVC-MD-Te* Database.

Number of Pristine Images in Database				3570
Distortion Stage	**Distortion Combination**			**Number of Images**
	1	**2**	**3**	
1	JPEG	Noise	Noise	39270
2	JPEG	JPEG	JPEG2000	667590

- Benchmarking was again done by using MS-SSIM [9].
- Three scenarios can be envisioned for mid-stage IQA.



(a) JPEG-JPEG (b) Noise-JPEG (c) Noise-JPEG2000

Figure 4: Performance of NR IQA algorithms at individual Stage-1 distortion levels of the *IVC-MD-Te* database.

- ***Scenario-1***: Access is available only to content after Distortion Stage n, i.e., to image D_n.
  - Only NR algorithms are applicable.
  - Performance of all NR IQA algorithms is unsatisfactory, especially for mixed distortion types (Table 3).
  - Performance of NR IQA algorithms is inconsistent across increasing Stage-1 distortion levels (Figure 4).
- ***Scenario-2***: Access is available to image D_{n-1} in addition to D_n (example: at transcoder input/output). We refer to image D_{n-1} as a *degraded-reference*.
  - FR algorithms can be used to evaluate the quality of D_n relative to D_{n-1}.
  - MS-SSIM [9] was used to implement this scenario.
  - *Scenario-2* outperforms *Scenario-1* (Table 4).
- ***Scenario-3***: In addition to images D_n and D_{n-1}, prior knowledge is also available about the quality of image D_{n-1} relative to the *pristine-reference* image.
  - Using the quality scores through the delivery chain, an SVR based model was used to predict the quality of image D_n.
  - *Scenario-3* far outperforms *Scenarios-1* and *2* (Table 5).

Table 3: Performance evaluation of *Scenario-1*.

Distortion Combination	**NR Method**	**PLCC**	**SRCC**
JPEG-JPEG	BRISQUE	0.7238	0.7163
	CORNIA	0.7131	0.7170
	LPSI	0.6800	0.6719
	NIQE	0.5795	0.5605
	WANG02	**0.7444**	**0.7399**
Noise-JPEG	BRISQUE	0.5770	0.5599
	CORNIA	**0.6263**	**0.6018**
	LPSI	0.4502	0.3920
	NIQE	0.4523	0.4161
Noise-JPEG2000	BRISQUE	0.5811	0.5717
	CORNIA	**0.6252**	**0.6358**
	LPSI	0.4187	0.3945
	NIQE	0.5735	0.5674

Table 4: Performance evaluation of *Scenario-2*.

Distortion Combination	**PLCC**	**SRCC**
JPEG-JPEG	0.8199	0.7443
Noise-JPEG	0.7080	0.6902
Noise-JPEG2000	0.6975	0.6780

Table 5: Performance evaluation of *Scenario-3*.

Distortion Combination	**PLCC**	**SRCC**
JPEG-JPEG	0.9931	0.9944
Noise-JPEG	0.9554	0.9523
Noise-JPEG2000	0.9408	0.9383

Conclusion

- Traditional FR and NR IQA frameworks and models fail to sustain their performance with multiple distortion stages.
- Relaying IQA results along the distortion chain and developing IQA models accordingly leads to substantially improved performance, showing great potential for *degraded-reference* IQA research.