Optimal Feature Selection for Blind Super-Resolution Image Quality Evaluation

Juan Berón* Hernán Darío Benítez Restrepo* Alan C. Bovik†

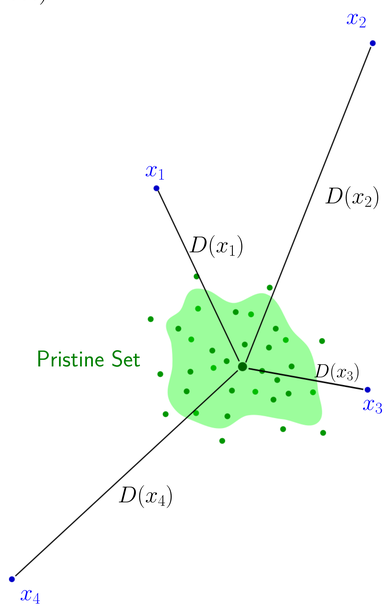
*Pontificia Universidad Javeriana †The University of Texas at Austin

Summary

- Definition of an Opinion Unaware Non Reference Image Quality Assessment (NR IQA) metric through the "bag of features" approach.
- Feature reduction through the implementation of the sequential forward floating search.
- Design of three Opinion Aware NR IQA metrics.
- Cross-Dataset validation of the performance of the resulting metrics

Metric Definition

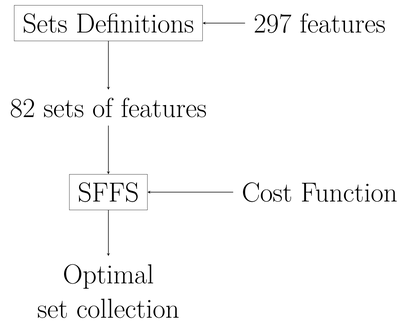
Each image is represented as a point in $\mathbb{R}^n$ where n is the number of features that are extracted from each image. The quality of an image is determined as the distance from a set of high quality images (pristine set).



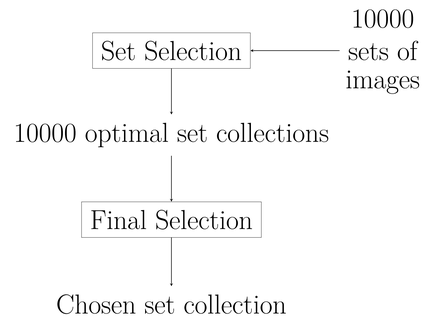
$$D(x) = \sqrt{\sum_{i=1}^{n} \frac{(x_i - \mu_i)^2}{c_i^2}}, \qquad (1)$$

Feature Selection

A total of 297 features are calculated per image. However, not all the features are the best for the metric. To reduce the number of features they are grouped in unbreakable sets and the SFFS will select the best collection of sets.



The cost function is the Pearson correlation coefficient between the scores of the metric and the human scores to a given set of images. Then, the cost function depends on the set of images used.



The final selection step is choosing the features that appeared 99% of the times. Two Opinion Unaware NR IQA metrics were created depending on the source of the sets of images chosen.

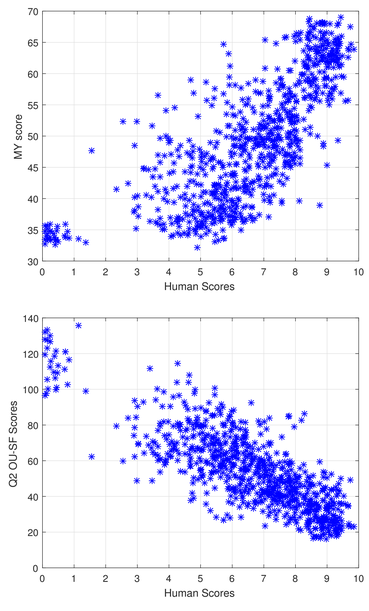
- $Q1_{\text{OU-SF}}$: 167 features (from Ma and Yang dataset).
- $Q2_{\text{OU-SF}}$: 134 features (from SR-IQA Javeriana).

Cross-Dataset Validation

Metrics were trained in one dataset and the performance was assessed in the other dataset.

	Testing Dataset	
Metric	Ma and Yang Dataset	SR-IQA Javeriana
$Q1_{\text{OU-SF}}$	**0.8434**	0.7856
$Q2_{\text{OU-SF}}$	0.7756	**0.8251**
NIQE	0.6413	0.4914
IL-NIQE	0.7671	0.6494
$Q_{\text{OA-ALL}}$	0.8158	0.7564
$Q1_{\text{OA-SF}}$	-	0.7609
$Q2_{\text{OA-SF}}$	0.8108	-
MY	0.7848	0.6787
BRISQUE	0.4755	0.5531
PI	0.7863	0.6289
MS-SSIM	0.6936	0.5575
FSIM	0.6857	0.5525
SSIM	0.5606	0.5795
VIF	0.7415	0.6135
IFC	0.7478	0.5598

SRCC of the scores of the metrics with human scores



Metric scores vs human scores in Ma and Yang dataset

Image Datasets

Two images datasets of super resolved images were used: Ma and Yang and SR IQA Javeriana.

Dataset	Num Images	SR algorithms
Ma and Yang [1]	810	9
SR-IQA Javeriana	608	7

Conclusions

- The feature selection yielded opinion unaware metrics with a performance comparable to opinion aware metrics.
- Selecting features for opinion unaware NR IQA metrics allowed the design of opinion aware NR IQA metrics
- The selection procedure is not bonded to Super Resolution and could be applied to other types of images.

References

[1] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang, "Learning a no-reference quality metric for single-image super-resolution," *Computer Vision and Image Understanding*, vol. 158, pp. 1 – 16, 2017.

Acknowledgements

The authors would like to thank NVIDIA Corporation for the donation of a TITAN XP GPU used in these experiments, Pontificia Universidad Javeriana and the authors would also like to acknowledge the grant provided by Comision Fulbright Colombia

Pontificia Universidad JAVERIANA Cali