Research Internship: Task 1

Sriram dhanasekaran

December 2025

Question 1

Why do we need to conduct Underwater Image Quality Assessment?
The Ocean covers 70% of the world, where a variety of living beings reside.[1] Researchers and Scientists are actively studying them with the help of divers or autonomous systems as they capture images of underwater flora and fauna, along with geographical features. But the images captured couldn't aid much in the research conducted, in their raw form at least, because of the physical phenomena called "absorption and scattering of light."As the sunlight gets absorbed (red light first), the images captured have a blue, blueish-green, or green tint and a hazy effect on them.[2, 3] Thus, the images have poor visibility, varied saturation, poor contrast, and lesser brightness in some cases.
This makes it challenging to compare and attain key information from the raw underwater images, requiring **Underwater Image Quality Assessment (UIQA)**. This produces a metric to measure the visibility, colour contrast, and saturation required to enhance the image for better gathering of information.[3, 4]

Uses of UIQA

- improves scientific analysis by helping in avoiding misinterpretation in the features extracted from the images
- allows Autonomous machines to analyse and interpret data from underwater images without the aid of human evaluation of the images

This concludes the UIQA's importance in underwater research and autonomous systems, and how it helps to analyse data of underwater images and gives us a metric to compare them.[5]

Question 2

What are the types of Quality Assessment techniques available?
There are five main types of Quality Assessment techniques available here

Full Reference Image Quality Assessment (FR-IQA)

In this method, the assessment is done by comparing the distorted image with the pristine counterpart. This generally uses MSE (Mean Squared Error), PSNR (Peak Signal to Noise Ratio), or SSIM (Structural Similarity Index Measure). Here, MSE produces raw pixel error, PSNR represents a measure of peak error, and SSIM measures structure, brightness, and contrast.[6, 7, 8]

$$MSE(x, y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

Where $i \in [1, 2, 3, ..., N]$ [9] represents the Mean Squared Error value for x, y where as

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

where MAX_I represents the maximum possible pixel value ($2^B - 1$ if B is the no. of bits) and MSE represents Mean Squared error value [9] represents the Peak Signal to Noise Ratio where as

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

where μ_x, μ_y are mean intensities (average pixel value), σ_x^2, σ_y^2 are variances, σ_{xy} represents covariance and C_1, C_2 are constants for stability [6] represents Structural Similarity Index Measure.

Reduced Reference Image Quality Assessment (RR-IQA)

In this method, the assessment is done by comparing the statistical outcomes of the pristine image (mean of pixels, standard deviation of pixels, or edge data of the image) with the distorted image. Example, Natural Scene Statistics (NSS) models.[10]

No Reference Image Quality Assessment (NR-IQA)

In this method, the assessment is done with no reference to the pristine image (as it is practically difficult to obtain pristine images of the underwater environment). There are three types of No Reference Image Quality Assessment.[11]

Statistical Model based NR-IQA

In this method, the assessment is done with no reference to the pristine image, but with expected statistical values of Natural Scenes from Natural Scene Statistics (NSS). There is a consistent pattern in the distorted image's pixel values and local contrast (contrast in a small localised area). First, the image is normalised using the Mean Subtracted Contrast Normalised (MSCN) coefficients and then modeled using the Generalised Gaussian Distribution (GGD) to obtain a result, which is then given to a regressor (Support Vector Regressor or Random Forest) to give a quality score.[11]

Feature-based NR-IQA

In this method, the assessment is done by extracting hand-crafted features that capture specific distortions such as blur, low contrast, noise, etc, without needing a reference image, which is then combined using a regressor to produce a quality score.[12]

Learning based NR-IQA

In this method, the assessment is done by learning a large dataset with underwater images and their Mean Opinion Score (MOS) using deep Neural Networks. It either uses Convolutional Neural Networks (CNN), Transformers, or any hybrid architecture. This method is limited by the need for large datasets, high computational power, and careful training to avoid overfitting.[13]

Subjective Quality Assessment (subjective QA)

In this method, the assessment is done manually, i.e, with a group of people evaluating the images for their quality and providing values like Mean Opinion Score (MOS). It is the most accurate way of assessment because it fully relies on human perception, but the drawback is that it is more expensive and time-consuming than other methods; however, it is necessary for aiding other methods.[13]

Task-specific Quality Assessment (Task-specific QA)

In this method, the assessment is task dependent, i.e, it evaluates an image by how much it aids the task present. This method is effective cause it is useful for practical application, but it is dependent on that particular task and doesn't work with general cases.[14]

Question 3

Mean Squared Error (MSE), Is it good enough to use it in Underwater Image Quality Assessment?

Mean Squared Error definition

$$MSE(x, y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

Where $i \in [1, 2, 3, ..., N]$ [9] represents the Mean Squared Error value for x, y

Significance of MSE in IQA

There are no parameters involved in MSE, making it easy to compute (one multiplication and two additions), and it doesn't require any memory. It has a physical meaning, like the energy of errors, which is easy to optimize as it is easy to calculate the gradient of MSE.

Limitations of MSE in IQA

The MSE method couldn't predict the visual quality perceived by humans properly. It gives a high error value for slightly transformed images, even though they aren't much different from the original image. Also, the MSE can be the same for two completely different images. It treats all pixels equally, i.e, it can't differentiate spatial features. [6]

Why we can't use MSE in UIQA

Even though MSE is a venerable method and is the first method people go to when it comes to signal or image processing, it cannot differentiate between different types of distortion available underwater (noise, blur, low visibility, haze, etc); therefore, we can't use MSE in UIQA.

Question 4

There are two inputs x, y, where x corresponds to the subjective opinion score of an image, while y corresponds to the predicted score of an image by a Machine Learning model.How do we infer the differences between x and y?

We compare x, y using statistical error and correlation coefficients, statistical error contains methods like Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), where if the statistical error is less, the predicted score y is more similar to the subjective opinion score x. However, in the correlation coefficients method, which contains the Pearson Linear Correlation Coefficient (PLCC) and Spearman Rank-order Correlation Coefficient (SRCC), if the correlation coefficient is high, the predicted score y is more similar to the subjective opinion score x. The equation for Mean Average Error is

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x_i - y_i|$$

where N is the total number of images in the dataset, x_i, y_i are the subjective opinion score and predicted score for the i^{th} image respectively. The Root Mean Square Error equation is

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}$$

where N is the total number of images in the dataset, x_i, y_i are the subjective opinion score and predicted score for the i^{th} image respectively. The Pearson Linear Correlation Coefficient equation is

$$PLCC = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}}$$

where $\bar{x}, \bar{y}$ are mean subjective scores and mean predicted scores respectively. The numerator is the covariance (σ_{xy}), and the denominator contains the standard deviation of both values separately $(\sigma_x.\sigma_y)$.The Spearman Rank Correlation Coefficient equation is

$$SRCC = 1 - \frac{6\sum_{i=1}^{N} d_i^2}{N(N^2 - 1)}$$

where d_i =Rank(x_i)−Rank(y_i).[15, 16]

Question 5

What is the Statistical significance of F-tests?

F-test's definition

The F-test is calculated by comparing the F-statistic to the critical value (from the F-distribution table for the corresponding significance α value). If the F-statistic value ¿ critical value, then we can conclude that there is a significant difference; otherwise, there is not much difference.[17]

Methodology

F-statistic is defined as the ratio between the variance between groups and the variance within the groups; i.e, the ratio between the variance explained by the groups and the variance due to random error. The ratio is

$$F = \frac{MSR}{MSE}$$

where MSR is Mean Squared due to Regression and MSE is Mean Squared Error.[17]

Statistical significance of F-tests

F-tests are statistically significant when $F > F_\alpha$, where F is the F-statistic and F_α is the critical value. Meaning the effect is real and not accidental; i.e, in UIQA terms, the features predict the Mean Opinion Score (MOS) significantly.[17]

References

[1] N. O. Service, "How much of the ocean have we explored?," 2020. `https://oceanservice.noaa.gov/facts/exploration.html`.

[2] C. J. Mobley, *Light and Water: Radiative Transfer in Natural Waters*. Academic Press, 1994.

[3] J. Y. Chiang and Y.-C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1756–1769, 2012.

[4] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE Journal of Oceanic Engineering*, vol. 40, no. 3, pp. 676–676, 2015.

[5] S. Yang, Y. Chen, *et al.*, "Underwater image quality assessment: A benchmark dataset and evaluation," *JOSA A*, vol. 40, no. 7, pp. 1276–1289, 2023.

[6] 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*, vol. 13, no. 4, pp. 600–612, 2004.

[7] Q. Huynh-Thu and M. Ghanbari, "Scope of validity of psnr in image/video quality assessment," *Electronics Letters*, vol. 44, no. 13, pp. 800–801, 2008.

[8] A. Hore and D. Ziou, "Image quality metrics: Psnr vs. ssim," *2010 20th International Conference on Pattern Recognition*, pp. 2366–2369, 2010.

[9] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Prentice Hall, 4th ed., 2018.

[10] H. Sheikh and A. Bovik, "A visual information fidelity approach to video quality assessment," *IEEE Transactions on Image Processing*, vol. 15, no. 2, pp. 430–444, 2005.

[11] A. K. Moorthy and A. C. Bovik, "A two-step framework for constructing blind image quality indices," *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 513–516, 2010.

[12] S. Golestaneh and D. Chandler, "No-reference quality assessment of contrast-distorted images using feature maps," *Signal Processing: Image Communication*, vol. 29, no. 8, pp. 877–888, 2014.

[13] F. Dondi, S. Croci, and A. Giachetti, "Mos-based quality assessment and datasets: A survey," *Journal of Imaging*, vol. 9, no. 2, p. 45, 2023.

[14] A. C. Bovik, "Automatic prediction of perceptual image and video quality," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2008–2024, 2013.

[15] H. R. Sheikh, A. C. Bovik, and L. K. Cormack, "The live image quality assessment database," *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 99–104, 2006.

[16] P. Le Callet, S. Möller, and A. Perkis, "Qualinet white paper on perceptual quality assessment," in *Qualinet*, European Network on Quality of Experience in Multimedia Systems, 2012.

[17] D. C. Montgomery and G. C. Runger, *Applied Statistics and Probability for Engineers.* John Wiley & Sons, 6th ed., 2014.