Underwater Image Restoration Through a Prior Guided Hybrid Sense Approach and Extensive Benchmark Analysis
August 3, 2025

---

1. Introduction / Abstract

Underwater images often suffer from color distortions and blurriness due to the way water absorbs and scatters light, especially the rapid absorption of red light compared to green and blue. This degradation hampers tasks like marine life segmentation and object recognition underwater. The paper addresses the challenge of restoring underwater images to a quality comparable to in-air images, a task termed Underwater Image Restoration (UIR).

The authors propose a novel framework called **GuidedHybSensUIR**, which combines multi-scale processing with a new **Color Balance Prior** (a simple average of the RGB channels used as a guide to correct color imbalance). The framework uses a hybrid of convolutional neural networks (CNNs) for fine detail restoration and Transformers (a type of deep learning model that captures long-range dependencies) for global color correction. Extensive benchmarking against 37 state-of-the-art methods on multiple real-world datasets shows that their method outperforms others overall, providing a new standard for underwater image restoration.

---

2. Methodology

The researchers designed a U-shaped neural network architecture that processes images at multiple scales to restore both local details and global color balance:

- **Detail Restorer**: A CNN-based module that focuses on recovering fine textures and sharpness at different scales. It uses **quaternion convolution** (a mathematical operation that treats RGB channels as components of a quaternion to better capture color interdependencies) combined with two specialized blocks: the Residual Context Block (RCB) and the Nonlinear Activation-Free Block (NAFB). These blocks enhance feature extraction and preserve important image details.
- **Feature Contextualizer**: A Transformer-based module placed at the network's bottleneck to capture long-range relationships and global color dependencies. It uses three types of attention mechanisms:
  - **Adjust Color Transformer (ACT)**: Focuses on aligning image features with the color balance prior.
  - **Keep Feature Transformer (KFT)**: Retains relevant prior features while suppressing inaccurate ones.

– **Self-Attention Transformer (SAT)**: Captures internal relationships within the image features themselves.

These attention mechanisms operate on feature channels rather than image patches, reducing computational complexity.

- **Scale Harmonizer**: A module in the decoder that fuses features from different scales, using learnable scaling and shifting parameters conditioned on the input features to harmonize multi-scale information for high-quality reconstruction.

The **Color Balance Prior** is introduced as a guiding feature, computed as the average of the red, green, and blue channels per pixel, reflecting the Gray World Assumption (the idea that average color in a natural scene should be neutral gray). This prior is embedded into the Feature Contextualizer and the final decoding stage to steer the restoration process toward balanced colors.

The model is trained using a composite loss function combining: - Pixel-level fidelity loss (Smooth L1 loss), - Structural similarity loss (SSIM), - Perceptual quality loss (LPIPS, which measures similarity based on deep features aligned with human perception).

---

3. Theory / Mathematics

A key theoretical concept is the **Color Balance Prior**, derived from the Gray World Assumption, which states that in a well-lit scene, the average intensities of the red, green, and blue channels should be equal. The paper formalizes this as:

$$\mathrm{Prior}_i(x, y) = \frac{R(x, y) + G(x, y) + B(x, y)}{3}$$

where (R, G, B) are the color channel intensities at pixel ((x,y)), and the prior assigns the same average value to each channel, guiding the model to correct color casts caused by underwater light absorption.

Another important formula is the inter-channel attention used in the Adjust Color Transformer:

$$\mathrm{InterCAttn}(Q, K, V) = \mathrm{Softmax}\left(\frac{\hat{Q} \cdot \hat{K}^\top}{\tau}\right) \cdot \hat{V}$$

Here, (Q, K, V) are the query, key, and value tensors derived from image and prior features, ( ) is a learnable temperature parameter controlling attention sharpness, and the operation computes attention across feature channels to adjust color information effectively.

4. Key Results & Visuals

- **Figure 1**: Shows 3D scatter plots of color distributions for an input underwater image, the color balance prior, and the restored output. The prior lies centrally in the restored image's color distribution, indicating its effectiveness in guiding color correction.
- **Table I**: Quantitative comparison on three paired test sets (UIEB, EUVP, LSUI) using metrics like PSNR (pixel accuracy), SSIM (structural similarity), and LPIPS (perceptual similarity). The proposed method achieves the highest or near-highest scores across most metrics, outperforming 37 other methods.
- **Figure 9 & 10**: Visual comparisons with traditional and deep learning methods show that the proposed method better corrects color distortions and restores fine details, producing more natural and clearer underwater images.
- **Table IV**: Ablation study results demonstrate the contribution of each module (Detail Restorer, Feature Contextualizer, Scale Harmonizer, and Color Balance Prior). The full model combining all components achieves the best performance, confirming the synergy of the hybrid architecture and prior guidance.

---

5. Conclusion & Real-World Impact

The paper presents a novel underwater image restoration framework that effectively combines CNNs and Transformers with a simple yet powerful Color Balance Prior to address both local detail recovery and global color correction. The extensive benchmark and retraining of existing methods provide a standardized evaluation platform, showing that the proposed method surpasses current state-of-the-art techniques in restoring underwater images.

Limitations include the approximate nature of the color balance prior and the computational cost, which, while reasonable, is not the lowest among compared methods. Future work could explore more precise priors and further efficiency improvements.

This research significantly advances underwater imaging quality, enabling better visual analysis and automated understanding of underwater scenes, which is crucial for marine biology, archaeology, and environmental monitoring.