

Optimization of ISP Parameters for Low Light Conditions Using a non-linear Reference-based Approach

Shubham Alai, Radhesh Bhat; KPIT Technology Pvt. Ltd., Bangalore, India

Abstract

Image signal processors (ISPs) play a significant role in camera systems by processing the RAW image from the image sensor to the final image. These final images are then used for human vision or computer vision tasks. A camera tuning engineer tunes the ISP parameters iteratively for different lighting conditions and scenarios to improve image quality. Due to the large number of ISP parameters that need to be tuned and the number of iterations required to achieve the final desired image quality which has a better trade-off, the tuning process often takes several weeks to months. For low lighting conditions, the complexity increases exponentially since details in the images are not even visible to human eyes. In low-light scenarios, extracting the maximum information from images typically requires an increase in ISO gain. However, this often results in image artifacts and adverse effects on noise levels and color accuracy. To address these issues, the number of ISP parameters needed for tuning substantially increases. With manual tuning, it becomes impractical to achieve optimal settings by considering all of them address the complexities of manual tuning, particularly in low-light situations, we have proposed a novel approach to tune ISP parameters automatically with the help of a reference image, using covariance matrix adaption evolution strategy (CMA-ES) as a non-linear optimizer. In this approach, we have used MSSIM as a loss metric to optimize the ISP parameters for low-lux images by considering the high-lux, well-tuned image as the reference. OV13880 sensor with Qualcomm's Spectra 380 ISP simulator and AR0233 sensor with Texas Instruments' TDA4X ISP is used in our experiment to validate the approach. The performance of the automatically tuned image is compared with an image obtained with default as well as manually tuned ISP parameters. We have demonstrated that in comparison to the manual tuning approach, the automatic ISP tuning approach significantly outperforms in terms of subjective as well as objective IQ metrics while also reducing the overall tuning process time.

Introduction

Nowadays, camera systems are used in various electronic appliances having a variety of use cases. The application areas of these systems include communication, entertainment, personal security, surveillance, autonomous driving, etc. Depending on its application, these systems are categorized into human vision (HV) and computer vision (CV). The image signal processor (ISP) is an important and complex component of the camera system.

Role of the ISP is to convert the RAW image captured by image sensors into a final processed RGB image for a desired HV or CV applications. ISP consists of various black blocks, such as defective pixel correction, demosaicing, noise reduction, edge enhancement,

color correction, tone mapping, etc. Fig.1 shows the typical pipeline structure of an ISP. There are tens to hundreds of parameters in each of these processing blocks, all of which act in various ways to influence the image quality for the intended use case. After performing the sensor and optical calibration procedure, an engineer must tune these parameters to improve the image quality. The tuning procedure is an iterative process where the parameters from each block has been adjusted based on the objective as well as subjective image quality (IQ) criteria. Generally, IQ experts tunes the ISP parameters for various scenes captured with controlled lighting conditions in a lab. The objective image quality is evaluated on standard ISO charts with some well-defined key performance indicators (KPIs). Some of the commonly evaluated objective metrics include Signal-to-Noise Ratio (SNR), Modular Transfer Function (MTF), Texture Acutance, Dynamic Range, etc. Whereas subjective IQ evaluation requires numerous reviews from professionals in the field of camera image quality. Although each of the ISP blocks is independent in nature, a small change in one ISP parameter from one block will affect the entire image and, therefore the end KPIs. It requires a lot of effort to tune an Image Signal Processor (ISP) due to the involvement of the large number of parameters to be tuned. Generally, the process of manual tuning takes several weeks to months to achieve the desired image quality. Fig. 2 provides an overview of the manual tuning process for different ISPs.

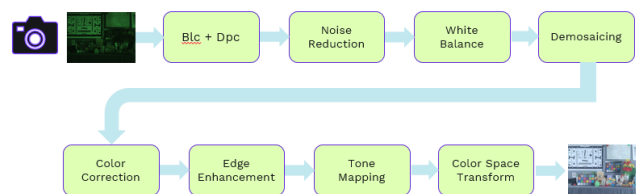


Figure 1. Camera Image Signal Processor (ISP) pipeline

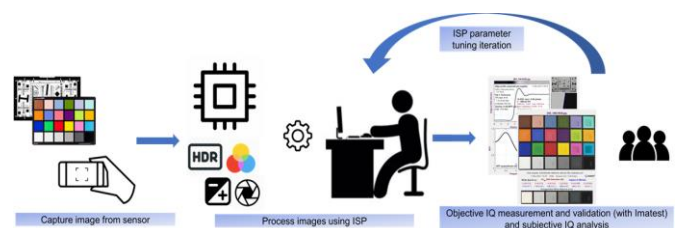


Figure 2. Camera ISP tuning process

To decrease the number of manual iterations required (thereby reducing time complexity) and achieve the optimal set of ISP parameters, this paper introduces an automatic ISP parameter optimization method. This approach aims to attain optimal image quality without manual intervention by utilizing a high-lux reference image with the Covariance Matrix Adaptation Evolution Strategy (CMA-ES). In this work, we work on automatically optimizing important ISP blocks especially for low to very-low-light scenarios. We are focusing on parameters from noise reduction, edge enhancement, and tone mapping blocks that contribute majorly to improving IQ in human-vision camera systems. The aim is to find the optimized ISP parameter settings that can bring a low-lux image (test image) close to a high-lux (reference image) with the help of similarity measurement. So, here MS-SSIM will be used as a loss metric, and the goal is to maximize its value using a non-linear optimizer. We have used CMA-ES as a non-linear optimizer to find the best suitable ISP parameter by maximizing the loss metric. CMA-ES is a stochastic optimizer that belongs to the family of evolutionary algorithms mainly used in black-box optimization. It is also known for its ability to handle high-dimensional optimization problems along with noisy and ill-conditioned objective functions. All these qualities make it highly suitable to tackle the challenges involved in ISP optimization. We have also included modulation transfer function (MTF) as another loss metric to mitigate the overdenoising of low-lux image biased due to high-lux image and enhance the details in the low-lux images along with the contrast. The proposed approach is divided into two main stages - the reference generation stage and the parameter optimization stage.

The experiments are carried out on two different ISPs and sensor combinations, Qualcomm's spectra 380 with OV13880 and Texas instrument's TDA4x ISP with AR0233. In experimental results, we have compared objective IQ for images 1) obtained with default ISP parameters, 2) with manual tuning, and 3) with optimized parameters. We have shown that the automatically tuned ISP parameters perform better in terms of image quality than the conventional manual tuning method for low to very low lighting conditions.

Related Work

In previous research on automatic parameter tuning, the proposed approach focused on optimizing individual blocks separately rather than the entire ISP box [1]. However, it's worth noting that the majority of hardware ISPs function as complete black box, which poses challenges in generating intermediate reference images due to the limited access to detailed knowledge about each block. Additionally, since it works with each block independently, this approach is limited in its ability to manage trade-offs between different blocks, such as denoising and sharpness. Nishimura et al. used an artificial bee colony (ABC) [7] as a global optimizer, serving as a tuner, and the Nelder-Mead simplex as a local optimizer for fine-tuning. However, this approach resulted in increased time and computational complexity. Moreover, it faces limitations regarding the number of loss matrices. Most of the are also limited by the number of available loss metrics. While [2] focuses on end-to-end optimization by mimicking the hardware ISP, it has the drawback of having to work within limited number of tuning parameters. Additionally, it requires extra effort to replicate the behavior of a hardware ISP.

Proposed Approach

In the proposed algorithm, we are focusing on the ISP blocks in which ISP parameters are highly dependent on lighting conditions and the blocks that cause a trade-off between different KPIs. Basic calibrations such as black level corrections, lens shading/distortion correction, white balancing, colour corrections, etc., required to be performed even before the reference generation stage.

Reference generation stage

Reference generation stage aims generation of the ISP parameters for high-lux images using a manual or automatic tuning approach. Tuning can be done using objective KPIs such as Signal to noise ratio (SNR), MTF, and Dynamic range (DR) with the help of ISO charts. By using parameters generated from the above stage, the high-lux RAW image is converted into a high-lux reference RGB image post-ISP. Low-lux optimized image generated from parameter optimization stage will be highly dependent on the reference image and hence its ISP parameters. For example, if the reference image has some blurred edges, same will be reflected in the low-lux optimized image. So, it is important to tune the high-lux reference image according to our requirements from the low-lux image. Since we are using MSSSIM as a loss-metric reference image should satisfy the pixel-to-pixel matching criteria to avoid false scores. To obtain pixel-to-pixel matching both images should be captured just by varying light intensity with a stable setup, as shown in Fig. 3.

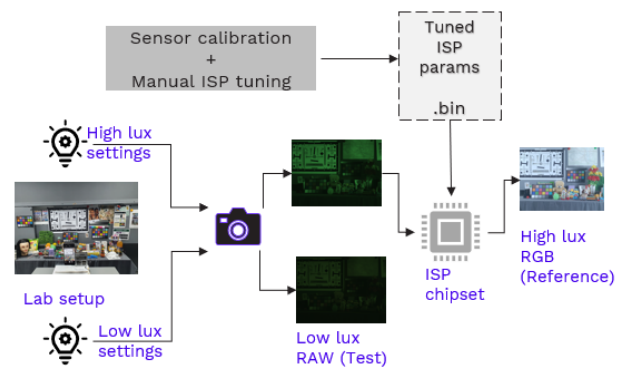


Figure 3. Reference generation stage

Parameter optimization stage

The optimization stage aims to get the best possible ISP hyperparameters that can improve the brightness, contrast, SNR, sharpness, and texture in the low-lux image and bring it closer to the high-lux reference image obtained from the first stage. An optimization algorithm with an appropriate loss metric plays a crucial role in this problem. We experimented with the optimization framework on several reference-based loss metrics and filtered out MS-SSIM as the best suitable for the problem statement. To make this selection, we have obtained a few low-lux images with different combinations of ISP parameters and ranked them based on subjective IQ with the help of some objective KPIs. MSSSIM score

follows this trend better as compared with other reference matrices. Three main attributes of the MSSSIM metric are luminance, contrast, and structure, which are directly correlated with the human visual system (HVS). Hence, the scores relatively follow the trend with subjective IQ ranking. For low-light imaging, it becomes difficult to decide the right combination of DR, SNR, MTF, Accutane, etc. Hence, MSSSIM can replace multiple objective KPIs with a single loss metric. The operation of the proposed framework is depicted in Fig. 4. Here, we calculate a loss metric by comparing the reference and test images. The CMA-ES optimizer then uses this loss metric to derive the subsequent set of ISP parameters. These newly obtained ISP parameters are subsequently applied to the ISP, which, in turn, generates the next image using the static metadata. The advantages of using CMA-ES are 1) it is a non-linear stochastic optimization process designed for black-box scenarios by [6] 2) it shows fast convergence compared with other evolutionary methods. 3) it can handle the curse of dimensionality (number of increasing parameters) and ill-conditioning problems. Due to the above reasons CMA-ES, and MSSSIM combination is the best suitable for the desired application.

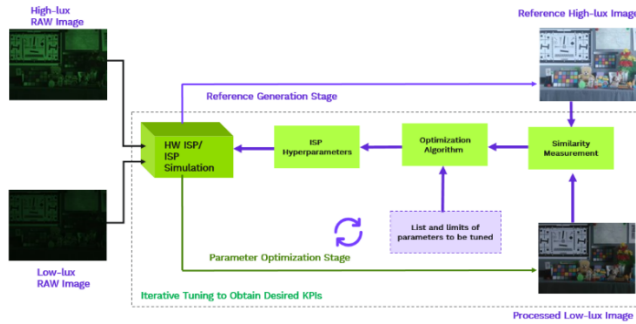


Figure 4. Proposed auto ISP optimization framework.

Loss function Definition

The Multi-Scale Structural Similarity Index (MSSSIM) is a mathematical measure used to evaluate the similarity between two images, typically a reference image (considered as the ground truth) and a test image (which is being assessed for quality or distortion). It is an extension of the more basic Structural Similarity Index (SSIM) that operates at multiple scales to provide a more comprehensive analysis of image quality. It is designed by considering all the various aspects of the human visual system.

MSSSIM is calculated by comparing the luminance (ℓ) (brightness) contrast (c) components, structural similarity (s) of the reference and test images at various scales. The mathematical expression for SSIM can be described as follows,

$$SSIM(x, y) = [\ell(x, y)]^\alpha \times [c(x, y)]^\beta \times [s(x, y)]^\gamma$$

where α , β and γ are constants that control the relative importance of luminance, contrast, and structure respectively.

Creating multiple scaled images using a Gaussian pyramid includes a process called image decimation. During this process, the original image is successively reduced in size to generate a sequence of

lower-resolution versions. This is usually achieved by applying a Gaussian filter and then down sampling the image (reducing its dimensions). The original image is labelled as Scale 1, and the highest scale is denoted as M . The loss metric is evaluated by consolidating measurements from various scales using,

$$MSSSIM(x, y) = [\ell_M(x, y)]^{\alpha_M} \prod_{j=1}^M [c_j(x, y)]^{\beta_j} [s_j(x, y)]^{\gamma_j}$$

Similar to SSIM, the exponents, α_M , β_M and γ_M are used to adjust the relative importance of different components.

Experimental Setup

In the proposed work, we are using Qualcomm's Spectra 380 ISP simulator, which will convert the raw image captured by the sensor into an RGB image using multiple blocks used in the ISP. With this simulator, we will receive the same output as the hardware ISP given all the required meta data provided matches with the hardware ISP. The RAW images are captured with Omnivision's 13880 smartphone 13MP sensor in a lab having controlled lighting conditions. The scene has a Macbeth color chart, ISO12233 charts to measure signal-to-noise ratio, and modulation transfer function (MTF50) respectively. Apart from ISO charts, it also has some real-world objects with various textures and details. For the experiment associated with Qualcomm's Spectra 380 ISP, a total of 28 isp parameters from 6 blocks have been used, as shown in Table 1. These parameters are responsible for controlling the noise, sharpness, and contrast in the image. Having some prior knowledge and experimental results we have restricted the maximum range of few parameters. The framework is set to optimize the 28-dimensional space for specified KPI that is, multi-scale structural similarity measurement (MSSSIM) and modulation transfer function (MTF). To evaluate MSSSIM high lux image obtained from the reference generation stage has been used as a reference. Whereas low light test images are captured by varying lux and color temperature to check the robustness of the proposed framework. The vertical slanted edge in the ISO12233 chart has been used as an ROI to evaluate MTF as shown in Fig. 5. The upper limit for MTF50 and MSSSIM is set as 0.6 Cy/Pxl and 1 respectively. A higher MSSSIM score suggests that the reference and test images are closely resemble each other, while MTF indicates the level of sharpness in the test image.

To check the scalability and robustness we have tested the proposed work on the Texas instrument's TDA4x isp with ON Semiconductor's 2MP AR0233 sensor. Similar to Qualcomm, we have used a total of 18 ISP parameters associated with the noise reduction, sharpness, and tone mapping blocks, as shown in Table 2. We then validate the image quality performance for the above-mentioned experiments in the next section. Apart from MS-SSIM and MTF, we have also evaluated the performance of the approach by evaluating SNR in the image.

Table 1: Spectra 380 ISP parameters

ISP Blocks	Parameter (length)	Value range
ABF, ANR (Noise Reduction)	Noise preservation (5), Denoise strength (4), Edge softness (4), Noise preserve anchor (5),	[0, 1], [0, 1], [0, 16], [0, 1],

	Noise preserve base (5), Frequency-based NR for luma channel (64), Frequency-based NR for chroma channel (64)	[0,1], [0,100], [0,100]
ASF (Sharpness)	Layer 1 gain positive (64), Layer 1 gain negative (64), Layer 2 gain positive (64), Layer 2 gain negative (64), smoothing strength, gain contrast positive (32), gain contrast negative (32), Layer1 activity normalization (64), Layer2 activity normalization (64), Layer1 clamp(1), Layer2 clamp(1)	[0.7,9], [0,7.9], [0,7.9], [0,7.9], [0,0.9], [0,1], [0,1], [0,0.9], [0,0.9], [0,255], [0,255]
GTM, LTM (Tone Mapping)	GTM curve param a_middletone(1), middletone_w (1), LTM strength (1), LCE strength (1), Dark boost(1), Bright suppress (1)	[0,1], [0,32], [0,4], [0,4], [0,4], [0,4]



Figure 5. Gray patch ROI (left) and Slanted edge ROI (right) used for visual noise and MTF metric evaluation respectively.

Table 2: TDA4x ISP parameters

ISP Blocks	Parameter (length)	Value range
NSF4 (Noise Reduction)	Strength (1)	[0.1,10]
EE (Sharpness)	Merge select (1), shift amount (1), Edge sharpening gain (1), Edge Sharpening threshold(1)	[0, 1], [0, 31], [0, 255],
	Edge sharpening offset (1), Edge intensity (4096), Threshold before lut (1)	[0, 65520], [0, 1008], [0.512]
GLBCE (Tone Mapping)	Asymmetry (1), Second pole (1), Glbce strength (1), Intensity variance (1)	[0.30], [0,40], [0,255], [0,15]

Experimental Results

In this section, we show the experimental results of the proposed framework on different ISP chipsets. The performance of the proposed work has been evaluated using image quality evaluation matrices mentioned in the section Loss function definition and some objective IQ matrices. We have also asked the golden eye experts to evaluate the quality of the images subjectively. The results contain the comparison of objective KPIs between the manually tuned images, images with default isp parameters, and images with optimized isp parameters (proposed framework). MTF is measured using the slanted edge in the ISO12233 chart, and SNR is measured on patch number-22 of the Macbeth chart. The objective IQ metrics above are evaluated with Imatest software.

The Experimental results from Fig.6 to Fig.13 show the comparison of high-lux image with the images obtained with different isp-parameter settings. The proposed framework can improve the subjective image quality for low-lux images ranging from 100Lux to 1 lux. From Fig.8 and Fig.9, we can observe that the isp parameters given by the proposed framework improve the produced image with better contrast compared with manually tuned. Fig.10 and Fig.11 prove that the proposed work surpasses the image tuned by manually tuned approach and improves the brightness of the extremely low-lux image by containing sharpness as well as a signal-to-noise ratio. Fig. 12 and Fig.13 show the scalability of the proposed framework on TI's platform on the 1 lux image.



Figure 6 Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters for Qualcomm ISP. a) High lux reference image b) Low Lux image tuned with default parameters c) Low Lux manually tuned image d) Low Lux image with optimized parameters (proposed approach)

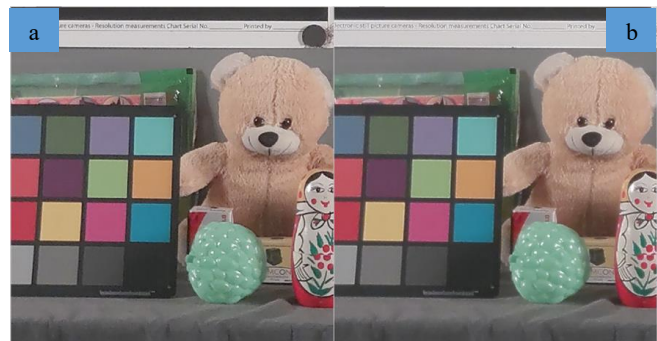


Figure 7: Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters for 100Lux image with Qualcomm ISP. a) Manually tuned image b) Image with optimized parameters (proposed approach).



Figure 8: Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters for Qualcomm ISP. a) 1000 lux reference image b) 10 Lux image tuned with default parameters c) 10 Lux manually tuned image d) 10 Lux image with optimized parameters (proposed approach)



Figure 11 Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters for 100Lux image with Qualcomm ISP. a) Manually tuned image b) Image with optimized parameters (proposed approach).

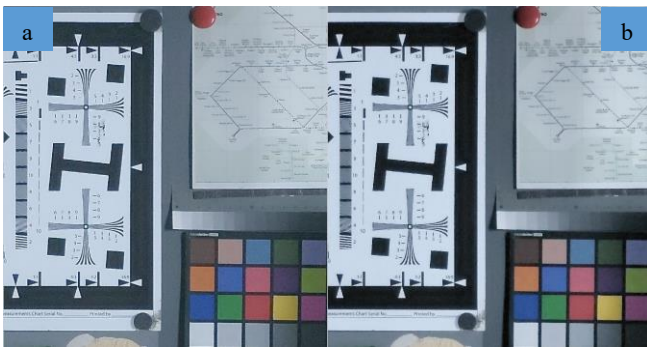


Figure 9: Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters for 10Lux image with Qualcomm ISP. a) Manually tuned image b) Image with optimized parameters (proposed approach).



Figure 12: Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters with TI ISP and 2MP sensor. a) 1000 lux reference image b) 10 Lux image tuned with default parameters c) 10 Lux manually tuned image d) 10 Lux image with optimized parameters (proposed approach)



Figure 10 Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters for Qualcomm ISP. a) 1000 lux reference image b) 1 Lux image tuned with default parameters c) 1 Lux manually tuned image d) 1 Lux image with optimized parameters (proposed approach)



Figure 13 Subjective comparison of noise reduction and sharpness tuning done with different ISP parameters for 1Lux image with TI ISP and 2MP sensor a) Manually tuned image b) Image with optimized parameters (proposed approach).

The Table 3. below shows the performance of the proposed work in different lighting conditions with 13 MP and 2MP sensors. From the table, it can observe that our framework outperformed in all tested

scenarios. Although SNR has not been given as an optimization metric, still the framework has shown significant improvement in the SNR for low-lux images. The reason behind it is high lux reference image has better SNR compared to the low lux image. So, to match the performance of the reference image our framework has improved the SNR. But we have to take care that denoising shouldn't be the dominant factor since, it may introduce haziness in the image and degrade the subjective image quality for low lux images.

We conducted a survey where experts ranked the images based on trade-offs and subjective image quality, while objective numbers remained concealed from them. Fig. 14 displays the rankings assigned by the experts for various images, illustrating that the proposed framework consistently yields superior image quality compared to manually tuned images, even in subjective evaluations. Fig. 15 shows the convergence of the framework for a given KPI. Each data point represents the image obtained with one combination of ISP parameters. One population in a non-linear optimizer gives the combination of ISP parameters.

Table 3: Comparison of performance of proposed method with manual tuned and default isp parameters for different lux and color temperature.

Lux level, CT	Tuning Approach	SNR (in dB)	MTF50 (in cy/pxl)	MSSIM Score with reference image
100L 2500K	Default ISP Parameters	25.3	0.28	0.54
	Manual Tuning	44.9	0.282	0.68
	Proposed automated tuning approach	46.2	0.26	0.69
10L 9900K	Default ISP Parameters	13.8	0.50	0.27
	Manual Tuning	34.3	0.34	0.70
	Proposed automated tuning approach	40.5	0.32	0.72
1L 9900K	Default ISP Parameters	7.5	.0.30	0.22
	Manual Tuning	22.7	0.257	0.61
	Proposed automated tuning approach	31.8	0.16	0.67
01 Lux 2MP 6500K	Default ISP Parameters	29.8	0.27	0.75
	Manual Tuning	23.8	0.27	0.84
	Proposed automated tuning approach	42.0	0.24	0.90

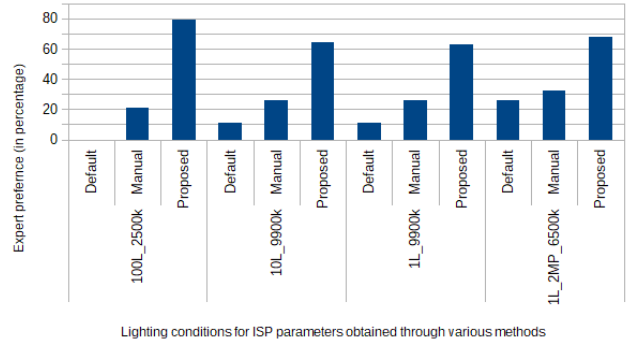


Figure 14: The proposed framework's performance in subjective analysis survey.

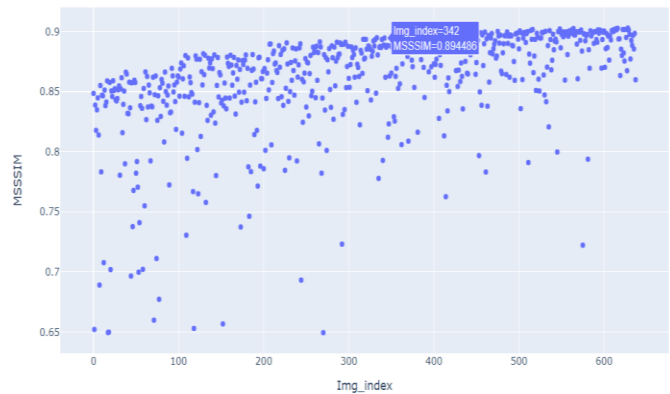


Figure 15: Visualizing data for images obtained during optimization, taking MSSSIM as the loss metric into account.

Conclusion

In this paper, we have demonstrated the automatic ISP tuning process for low to very low-light images using ISP parameters derived from noise reduction, sharpness improvement, and tone mapping blocks. Objective numbers clearly indicate that, with the assistance of the high-lux reference image, we achieved superior image quality compared to manual tuning and default ISP parameters. Although noise reduction and dynamic range-related key performance indicators have not been included, the proposed framework still shows significant improvement in this regard. Based on the survey results, it is evident that the proposed framework outperforms manually tuned images in subjective evaluations

With the proposed approach, we can attain an image quality that is better or at least similar to manual tuning, all while reducing the time complexity by 90%. It reduces the time complexity of tuning low-light images from a matter of weeks to just a few hours.

Future Scope

The future prospects for this proposed work involve exploring alternative loss metrics to enhance color accuracy in images and expanding the framework's capabilities to optimize white balance

gains and color correction matrices. Additionally, we aim to develop machine learning-based loss metrics that better align with both subjective and objective image quality trends compared to MSSIM.

Acknowledgement

We would like to express our gratitude to Apeksha Chipade, Samrendra Singh, and Praveen Singh for their invaluable assistance in helping us understand the manual tuning process flow, aiding in the analysis of ISP parameters for Texas Instruments after optimization, and providing us with calibrated ISP parameters. Our sincere appreciation also extends to the Golden Eye Experts in the Imaging team at KPIT Technologies for generously dedicating their time to offer input on optimized images versus manually tuned images and images obtained with default parameters. Furthermore, we wish to acknowledge the efforts of Ajay Basrur for presenting this work on our behalf at the Electronics Imaging conference.

References

- [1] Nishimura, Jun, Timo Gerasimow, Rao Sushma, Aleksandar Sutic, Chyuan-Tyng Wu, and Gilad Michael. "Automatic ISP image quality tuning using nonlinear optimization." In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pp. 2471-2475. IEEE, 2018.
- [2] Pavithra, G., and Bhat Radhesh. "Automatic image quality tuning framework for optimization of ISP parameters based on multi-stage optimization approach." *Electronic Imaging* 2021, no. 9 (2021): 197-1.
- [3] Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik. "Multiscale structural similarity for image quality assessment." In *The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2, pp. 1398-1402. Ieee, 2003.
- [4] Mosleh, Ali, Avinash Sharma, Emmanuel Onzon, Fahim Mannan, Nicolas Robidoux, and Felix Heide. "Hardware-in-the-loop end-to-end optimization of camera image processing pipelines." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7529-7538. 2020.
- [5] Xi, Weijuan, Huan Zeng, and Jonathan B. Phillips. "An Automatic Tuning Method for Camera Denoising and Sharpening based on a Perception Model." *Electronic Imaging* 2018, no. 5 (2018): 442-1.
- [6] Hansen, Nikolaus. "The CMA evolution strategy: A tutorial." *arXiv preprint arXiv:1604.00772* (2016).
- [7] Karaboga, Dervis, and Bahriye Basturk. "Artificial bee colony (ABC) optimization algorithm for solving constrained optimization problems." In *International fuzzy systems association world congress*, pp. 789-798. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.

Author Biography

Shubham Alai received his M.Tech degree in Signal Processing from the National Institute of Technology, Calicut, in 2021. He currently works in the R&D department of the Digital Imaging team at KPIT Technologies Pvt. Ltd. His research interests include image processing, camera image signal processing algorithms, machine learning, deep learning, and computer vision.

Radhesh Bhat received his BE degree in Electronics and Communication from Visvesvaraya Technological University in 2004. Currently he is heading digital imaging division at KPIT Technology Pvt. Ltd. His research interests include camera image signal processing pipeline, image quality, computer ion and deep learning.