https://doi.org/10.2352/ISSN.2470-1173.2020.9.IQSP-315


# DNN-based ISP Parameter Inference Algorithm for Automatic Image Quality Optimization

***Younghoon Kim, Jungmin Lee, Sung-Su Kim, Cheoljong Yang, TaeHyung Kim and JoonSeo Yim***
***Samsung Electronics, Hwasung-City, Gyeonggi, 18448, Republic of Korea***

## Abstract

*In camera development, because the image quality is subjective and the tuning complexity is increasing, building a correlated model with image signal processor (ISP) pipeline is very demanding task. In order to overcome those problems, this paper proposes an automatic image quality tuning framework based on Deep Neural Network (DNN). The image quality metric (IQM) have been defined to quantifies subjective image quality, which effectively represents the actual user perception. In this way, fast reproduction of the desired image has been possible through the minimized computing resource. Proposed Optimization methodology consists of Phase 1, a ISP modeling, and Phase 2, parameter optimization. Phase 1 construct a model between the parameters of ISP and the image quality metric. At phase 2, we add partially connected layer at input layer in order to optimize the parameters of ISP. Using back-propagation approach, the network selectively updates only the weights of partial connections, which allow to automatically derive the optimal parameters for high quality image. This idea has been implemented and experimented through commercial 16 Mega pixel resolution CMOS image sensor (CIS) and the state-of-the art ISP.*

## Introduction

Recently, camera has become a killer application of major smart phones and also has derived the higher image quality. In order to response those demands, high resolution CMOS image sensors (CIS) are being developed. In addition, the evolution and the complexity of Image Signal Processors (ISP) has made the tuning process increasingly demanding task. Especially, the algorithm blocks have trade-off between such as de-noise and sharpening performance. The process of tuning the optimal parameter requires a lot of experience and takes lots of limitation, as therefore, there is a growing interest in auto-tuning systems based on machine learning.

Nishimura J. *et. al.* proposed a non-linear optimization and automatic generation method of reference image [1]. Ethan Tseng, E. et al. find the auto-tuned ISP parameters by leveraging differentiable mapping between the ISP parameter and image metrics [2]. Both approaches need to set up a target quality by reference images. Then, the similarity is evaluated with reference images by only changing the ISP parameters.

In this paper, we propose a framework for auto-tuning of the ISP parameters which includes Image Quality Metric (IQM) measurement and the ISP parameter optimization process. Unlike prior studies, we set the reference quality with desired IQM score rather than image. This methodology allows to easily reflect user preferences. The proposed methods translate a subjective image quality to objective score in IQM measurement part. Deep Neural Network (DNN) based optimization process has been proposed by using combining of architecture. The proposed architecture consists of two phases. The Phase 1 is a training process that learns the relations between the tuning parameter and image quality (IQM). Phase 2 derives the optimal ISP parameters in order to reach the desired IQM score set by the user. Section 2 presents the quantification of image quality with IQM score and pre-processing. In section 3, we address the optimization of image quality based on DNN using those pre-processed IQM scores. The proposed methods were experimentally validated through a set of TE42 chart images in Section 4. Concluding remarks are given in section 5.

## Quantification of Image Quality

We introduce the image quality metrics and IQM pre-process ahead of actual learning and optimization process. In order to apply neural network, quantifying the input/output variables is essential. We also consider the abstraction step of the input data set and feature for the robust learning performance.

### *Definition of Image Quality Metrics*

**Figure 1.** shows the conceptual ISP pipeline which consists of several image processing blocks. Out of many building blocks, we consider two image quality metric, "noise" and "detail", which are determined by the NR (noise reduction) and sharpening blocks.

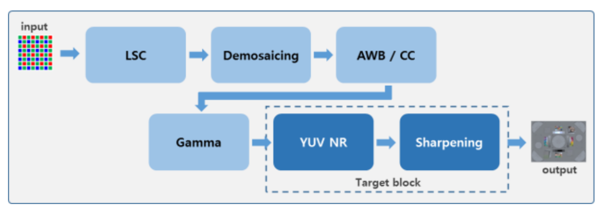



*Figure 1. Conceptual ISP pipeline*

When it comes to "noise" metric, in order to represents the actual user perception, we utilized Visual Noise (VN) rather than general SNR (signal to noise ratio). VN reflects the cognitive characteristics of human visual system. As for "detail" metric, we consider the concept of acutance and effective pixel count (EPC). Acutance is the weighted spatial frequency response of the image with human perception, while EPC is an effective resolution calculated from MTF10, minimally discernible signal difference. TE42 chart shown in **Figure 2** (a) is a multi-purpose standard chart for subjective/quantitative evaluation [10]. The mean and maximum of VN are calculated from gray patches in **Figure 2** (b). The detail metrics, acutance and EPC, are measured with Siemens star, slanted edge, and dead leaves pattern in **Figure 2** (c), (d), (e), respectively. [3-6].

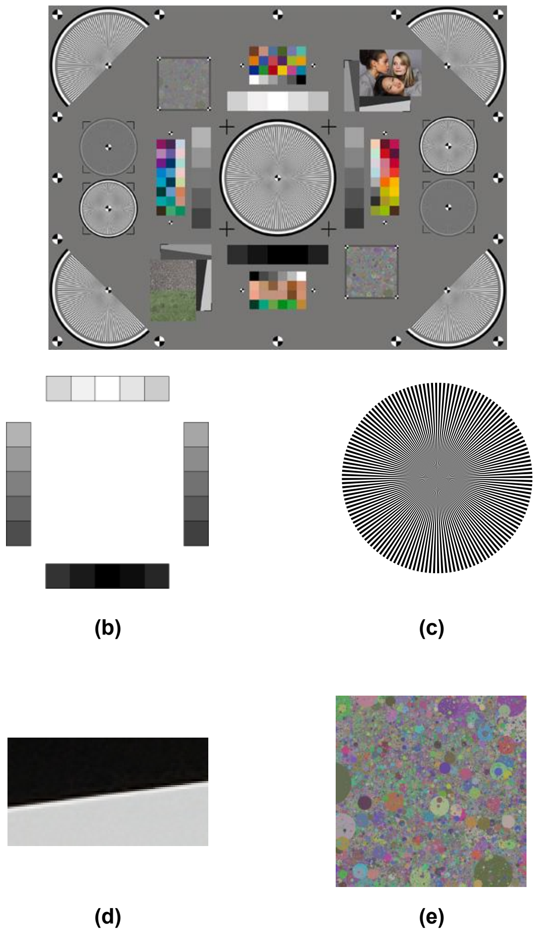



*Figure 2. TE42: (a) entire chart, (b) gray patch, (c) Siemens star, (d) slanted edge, (e) dead leaves pattern*

## IQM Pre-processing before learning phase

To build a training dataset, by taking a picture of TE42 chart, we generate 11,000 simulation images by applying uniformly distributed perturbation of 148 major ISP parameters in the de-noise and sharpening blocks. For each parameter, simulated images are scored as IQM.

By Feature Abstraction, generated 148 ISP parameters are compressed into 20 high level features. Feature Abstraction improves the learning performance and is very instrumental in reducing artifact.

In Feature Abstraction, we proposed two schemes, i.e., linearization, and dimension reduction. In linearization, parameters were reduced in the form of a quadratic equation using mathematical model. The dimension has been reduced by using well known Feature Extraction algorithm, with the principle of components analysis, that minimizes the information loss [11]. In this way, we can minimize the tuning parameters and can produce more robust experimental results.

## Deep Learning based ISP Modeling and Optimization

Proposed optimization method based on deep neural network (DNN), consist of two phases, a virtual ISP modeling and ISP parameter optimization:

1. **Phase 1. (Virtual ISP Learning)**: learning the ISP algorithms in the form of a black box to derive a virtual ISP model.

2. **Phase 2. (ISP Optimization)**: deriving a logical ISP parameter to achieve the optimized image quality for ISP.

As because, commercial ISP chain is not open to public, ISP behavior model is derived from phase 1, which build a virtual ISP model. At phase 2, the neural network tunes the optimal ISP parameter. This is connected to the virtual ISP model and finds the optimal ISP parameter.

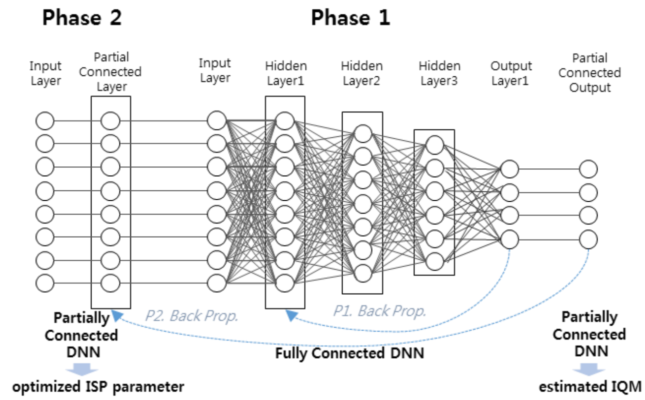



*Figure 3. The proposed two-combining neural architecture for automatic ISP optimization system*

## Phase1: Virtual ISP modeling

In Phase 1, the accuracy of the model is highly related to the image quality of the final image. Also there are potential risks to generate outlier with low quality image at Phase 2. For the robust model, we use multi-output DNN that consists of fully connected layers with Rectifier Linear Unit (ReLU) activation function as shown in Phase 1 of **Figure 3**. During the learning process, the objective function of multi-output regression is used for modeling the relation between ISP parameter $\boldsymbol{x}$ and IQM score $\boldsymbol{y}$. Additionally, Adam (adaptive moment estimation) [7] optimizer method is used for stochastic optimization with a learning rate of 0.001.

The output $\boldsymbol{y}_i$ is the predicted IQM score normalized to [0,1]. And we use the mean square error for the cost function shown as in Eq. (1).

$$J(\theta_h) = \frac{1}{2\mathrm{m}} \sum_{i=0}^{m} (\hat{y}_i - y_i)^2 \tag{1}$$

where, $\hat{y}_i$ and $y_i$are the $i$-th estimated score and ground truth score of training data, $\theta_h$ is parameters of DNN and $m$ is the number of score categories (*i.e.,* Visiaul noise, mean, max, accutance and effective pixel counts).

## Phase 2: ISP parameters optimization

The purpose of phase 2 is to find optimal ISP parameter corresponding to the desired IQM score. In order to achieve this goal, we proposed a Partial Connected Network (PCN) shown as

Phase 2 of **Figure 3.**, which has single connection between input layer and the first hidden layer of adaptive DNN in phase 1.

Through the inference process, PCN has to come up with the best $w_p$ value, the weight of partial connection. This weight determines the optimal ISP parameters from initial values of ISP parameter, $x_{init}$. Generally, in conventional DNN, output value is inferenced by feed-forward methods. However, in our case, the goal of framework is to inference the partial weight $w_p$, at input layer from the desired IQM score at output layer, instead. Henceforth, PCN uses back-propagation method in order to inference the optimal ISP parameters [8-9].

During the phase 2 process, the weights of the hidden layer $w_h$ computed from phase 1 are already fixed. As therefore, only the partial weight, $w_p$, will be updated during the optimization process from the desired IQM score. We use the weighted mean square error for the cost function as Eq. (2).

$$J(\theta_p|\theta_h) = \frac{1}{2\mathrm{m}}\sum_{i=0}^{m} w_u^i \left(\hat{y}^i - y_u^i\right)^2 \tag{2}$$

where, $\theta_p$ is the weight of partial connection layer, $w_u$ is the user defined weight for score and $y_u$ is the desired IQM score.

In the optimization of $w_p$, same as phase 1, we use gradient descent methods with Adam. In order to satisfy guideline of parameter range recommended by ISP manufacturer, we set the constraint during the optimization process, shown as Eq. (3).

$$w_{p,new}^j = \begin{cases} w_{p,old}^j - \alpha\nabla w_{p,old}^j, & \text{if } c_{min}^j < z < c_{max}^j \\ do\ not\ update, & otherwise \end{cases} \tag{3}$$

where, $c$ is the constraint of the ISP parameters region, $j$ is an order of input dimension, $\alpha$ is a learning rate in training, $z$ is the result of $w_p^j - \alpha\nabla w_p^j$ and a subscript of $w_p$ means the updated status.

Finally, PCN can inference optimal ISP parameter by the combination of weight of partial connection and initial ISP parameters, $\mathrm{w}_p \cdot \mathrm{x}_{init}$, which generates the desired image as a result.

### *Consideration of output reliability*

When it comes to optimization process, the training iteration usually finishes once the training loss reaches a certain thresholds or specific epoch. However, we adapt novel termination strategy in order to secure robust performance.

In phase 2 optimization, we estimate the IQM score ($\hat{y}$) based on DNN hidden layer weight $\theta_h$ and optimal ISP parameter $\mathrm{w}_p\mathrm{x}_{init}$. Next, this score is compared with desired IQM score defined by user. The phase 2 optimization process finishes when the ratio of two scores reaches the certain threshold as shown in Eq. (4).

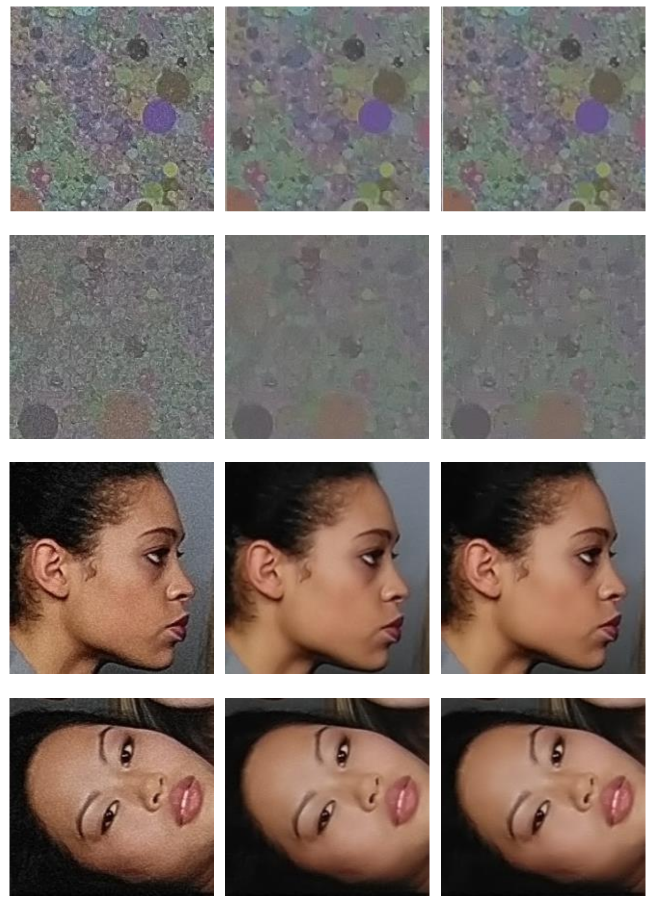
$$\mathrm{ABS}\left[\frac{w_u^i y_u^i - y^i\left(\mathrm{w}_p \cdot \mathrm{x}_{init}|\theta_h\right)}{w_u^i y_u^i}\right] < \varepsilon, for\ \forall\ i \tag{4}$$

where, ABS is absolute operation and $\boldsymbol{\varepsilon}$ is a constant value.

The final optimization phase includes conditions to validate the actual picture quality with respect to the numerical quality generated through deep learning process. These verification conditions ensure that the target quality is accurate and reliable.

## Experimental Results

In order to experimentally verify the proposed tuning framework, we applied the tuning process to camera system composed of 16 megapixel CIS and the state-of-the-art commercial ISP which has six algorithm blocks. We optimized the de-noise and sharpening blocks which are the most influential parts to image quality.



(a) (b) (c)

*Figure 4. Comparison of final image of TE42 chart: (a) initial, (b) hand-optimized, (c) generated from the proposed system*

We evaluate the performance of proposed methods based on the TE42 chart that includes various image scene such as gray patch, resolution chart, natural scene and etc. We captured one image of TE42 chart in a 1000 lux environment in order to obtain the original image for the data generation. Next, 11,000 and 2,500 images were randomly generated for the training and test data.

### *Virtual ISP modeling*

As because phase 2 results are very much influenced by the robustness of virtual ISP modeling, which is phase 1, so we pre-verified the performance of the virtual ISP modeling.

From the 148 tuning parameters, the ISP parameters are compressed into 20 variables as the input data $\boldsymbol{x}$ using the pre-processing described in section 2.2. For output $\boldsymbol{y}$, we evaluate four IQM score categories; VN mean, VN max, Siemens EPC and Siemens acutance. The performance are measured by Mean

Absolute Error (MAE) between the estimated IQM score $\hat{\mathbf{y}}$ and the ground truth score $\mathbf{y}$.

During the phase 1, all quality items (IQM score categories) showed less than 2% error as shown in Table 1, which is affordable level of image quality. This error affects the final image quality in the optimization during phase 2.

**Table 1. IQM loss (Mean Absolute Error) of Phase 1**

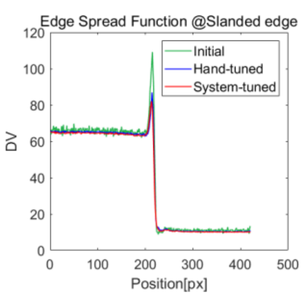
| VN mean | VN max | Siemens EPC | Siemens acutance |
|---|---|---|---|
| 0.0120 | 0.0158 | 0.0124 | 0.0125 |

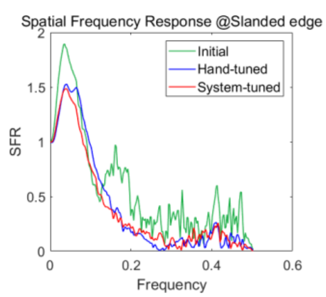
### *ISP parameter optimization*

In order to prove the performance of ISP parameter optimization process, we compare the result with the initial image and the hand-optimized image.

The randomly generated initial ISP parameters, $x_{init}$, and image are shown in **Figure 4** (a). The hand-optimized images by tuning expert are shown in (b). The resultant images shown in (c), are superior to (a) and comparable to (b).

The stability of image also should be considered during the optimization. In this regard, **Figure 5** (a) and (b) show the edge spread function and spatial frequency response of the slanted edge in TE42, respectively. The image quality of proposed system shows less noise in the flat region and the acceptable quality in the edge region compared to the initial image and comparable to the hand-tuned one.




**(a)** **(b)**

***Figure 5.*** *Comparison result of slanted edge of Figure. 3 (TE42): (a) edge Spread Function, (b) spatial frequency response*

The result has less fluctuation in the flat region and overshoot is well suppressed in the boundary region, compared to the initial and hand-optimized result shown in **Figure 5** (a). Excessive boosting in the high frequency range is also suppressed so that an equivalent quality can be obtained compared to the hand- optimized one shown in **Figure 5** (b).

**Figure 6** shows the converging behavior of IQM scores during optimization process from the initial score. Trade-off relation between the acutance score (of the Siemens Star chart) and visual noise (VN mean, VN max) has also been shown in this figure. This allows acutance score to move in the opposite direction of the visual noise. Later on, the similar behaviors are shown for final score.

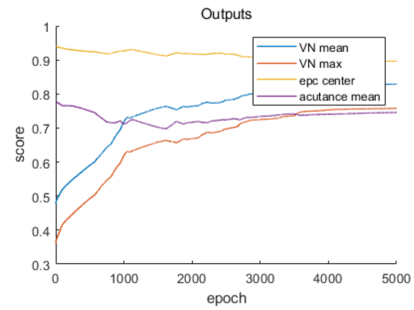



***Figure 6.*** *The converging behavior of IQM.*

From a quantitative point of view, the IQM score of the proposed result significantly improves as in Table 2. The EPC Siemens and the Siemens acutance decreased slightly -0.026 and -0.019, but the VN mean and the VN max were improved by 0.397 and 0.439, respectively.

**Table 2. Comparison of IQM score by tuning method**

| Method | VN mean | VN max | Siemens EPC | Siemens acutance |
|---|---|---|---|---|
| Initial | 0.452 | 0.338 | 0.943 | 0.776 |
| Hand-tuned | 0.789 | 0.677 | 0.927 | 0.781 |
| System-tuned | 0.849 | 0.777 | 0.917 | 0.757 |

As shown in the result of the experiment, we have demonstrated that the quality of detail and noise in result image are automatically improved by the proposed method.

### *Experiments on high dimensional features*

Previous experiments are based on 16 mega pixel size CIS and commercial ISP. In order to generalize, we extended the cases with different input and output dimensions by two trials.

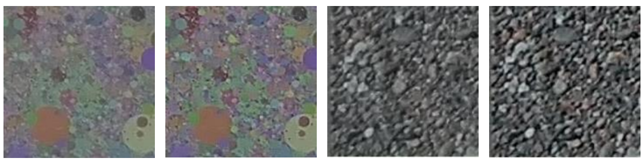

**(a)** **(b)** **(c)**

***Figure 7.*** *Variation of input dimension (the number of ISP parameter): (a) initial image, (b) 148 parameters are used, (c) 220 parameters are used*

<u>First trial:</u> The number of the input learning parameter has been increased. **Figure 7** shows the effects of the number of ISP parameters. If we increase the number of parameters, from 148 (b) to 220 (c), the image quality has been better optimized because we can control the de-noise and sharpening levels in fine manner.

By expanding a larger number of parameters, the resultant image is more natural.

Second trial: The number of output categories (IQM) has been increased. **Figure 8** shows the different texture image quality by increasing the number of IQM, *m*, from 4 to 6, where the additional extra items are Dead Leaves' acutance and Slanted Edge's acutance. By increasing the output categories (IQM), we can easily see the detail is also dramatically improved.



(a) (b)

***Figure 8.*** *Variation of target categories (IQM m): (a) dead leaves (m=4 → m=6), (b) stones (m=4 → m=6)*

## Conclusion

This paper proposed an automatic image quality optimization methodology based on Neural Network. There are several novelties in our approaches. By hybridizing FCN (Fully connected network) for phase 1 ISP learning stage and PCN (Partially connected network) for phase 2 ISP tuning parameter optimization stages, the quantitative metric (IQM) makes the fast reproduction of the desired image quality from the limited computing resource. Other approaches utilize the image data themselves, which makes the iteration very time-consuming.

Through the experiments, our optimization methodology has provided accurate and affordable image quality along with complex ISP.

The proposed method would be advantageous in obtaining the initial image quality for new pixels or new sensors without long tuning period by camera experts. Another big advantage is that the fast-changing state-of-the-art ISP pipeline and new user-quality preferences can be reflected in the proposed frameworks in real-time basis.

## References


[1] Nishimura, J. et al. Automatic ISP Image Quality Tuning Using Nonlinear Optimization. *Proc. 25th IEEE International Conference on Image Processing*, 2471-2475 (2018)

[2] Tseng, E. et al. Hyperparameter Optimization in Black-box Image Processing using Differentiable Proxies. *ACM Trans. Graph.*, 1, No. 1 (2019)

[3] Wueller, D. & Kejser, U. B. Standardization of image quality analysis - ISO 19264. *Proc. IS&T Archiving Conference*. 1, 111-116 (2016)

[4] Cao, F. et al. Dead leaves model for measuring texture quality on a digital camera. *Proc. SPIE*, 7537, 75370E (2010)

[5] Baxter, D. et al. Development of the I3A CPIQ spatial metric. *Proc. SPIE*, 8293, Quality and System IX, 829302 (2012)

[6] Artmann, U. Image quality assessment using the dead leaves target: experience with the latest approach and further investigations. *Proc. SPIE*, 9404, 94040J (2015)

[7] Kingma, D. et al. Adam: A method for stochastic optimization. arXiv preprintarXiv, 1412.6980 (2014)

[8] LeCun, L. et al. Efficient backprop. *InNeural Networks: Tricks of the Trade*, 9–50 (1998)

[9] Hecht-Nielsen, R. Theory of the backpropagation neural network. *InInternational JointConference on Neural Networks* (IJCNN), 593–605 (1989)

[10] TE42. https://www.image-engineering.de/products/charts/all/425-te42

[11] Jolliffe, I. T. Principal component analysis. *New York: Springer-Verlag* (1986)


## Author Biography

*__Younghoon Kim__ is the engineer of Samsung Electronics Co. Ltd. He received B.S. (2011) in Computer Science and Engineering from Konkuk University, Korea. His research areas are machine learning, big data. Recently, his current research interests include deep learning based image quality assessment and calibration.*

*__Jungmin Lee__ is currently working as an engineer at Samsung Electronics. He received his B.S. in electronic engineering (2010) and M.S. in image processing and computer vision (2012) from Sogang University in Korea. His research interests include image processing, computer vision and objective image quality assessment.*

*__Sung-Su Kim__ received his B.S. in electronic engineering and his M.S. in human vision system from KyungPook National University. Since 2004, he has worked in Samsung Advanced Institute of Technology (SAIT) and Samsung Electronics Co. Ltd., Korea. His research interests include pattern recognition, image understanding, image quality metric, and machine learning.*

*__Cheoljong Yang__ received the B.S degree in electrical engineering (2010) and his PhD in vision information processing (2017) from Korea University, Seoul, Korea. Since then he has worked in the system LSI division at Samsung Electronics. His current research interests are in machine learning based image quality assessment for CMOS sensors.*

*__Taehyung Kim__ received his B.S degree in Control & Instrumentation Engineering from Seoul City University (1996). Since then he has worked in Samsung Electrics Co. Ltd, Hwasung, Korea. His work has focused on CMOS Image Sensor development and its image quality, SW and image enhancement algorithm.*

*__JoonSeo Yim__ received his B.S. and Ph.D. degree from Seoul National University (1991) and KAIST (1998) respectively, majored in Electrical and Electronics Engineering. He has been worked in Samsung Electronics. His research interests include camera sensor innovation, evolutionary computation and design optimization methodologies.*