# 라이트필드 영상의 Perspective 및 재초점 화질측정방법 비교

*Vinh Van Duong, Thuc Huu Nguyen, 전병우
성균관대학교 전자전기컴퓨터공학과
e-mail: *duongvinh@skku.edu*, *thuckechsu@skku.edu*, bjeon@skku.edu

# Comparison of Quality Metrics of Perspective and Refocused Images in Light Field Images

* Vinh Van Duong, Thuc Huu Nguyen, and Byeungwoo Jeon
Department of Electrical and Computer Engineering, Sungkyunkwan University

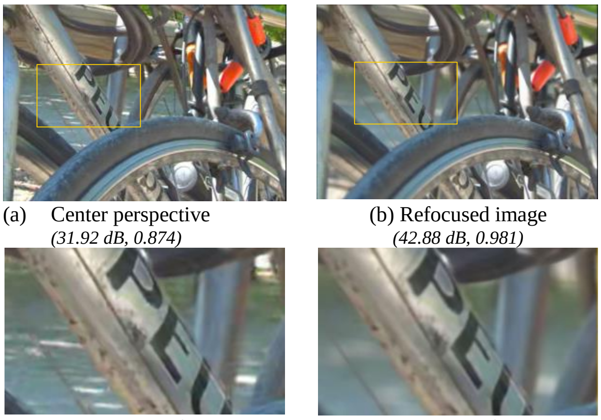
## Abstract

Digital refocusing and perspective change are the most expected applications of light field (LF) images. As LF image has a large amount of data, its compression is very essential. The fidelity of LF image after compression needs to be evaluated differently depending on a specific application such as perspective change or refocusing. In this paper, we investigate the fidelity of images after perspective change and refocusing. Several state-of-the-art objective quality metrics are compared. Our experiment shows that IWPSNR is the most reliable metric for both perspective and focus changes, but it does not outperform the popular metrics such as PSNR and SSIM.

## 1. Introduction

Recently, commercial LF cameras (e.g., Lytro or Raytrix) have come out in the market, and gained much attention from researcher to develop a new processing framework for LF images. Due to its huge amount of data, an efficient compression method is strongly required. In order to evaluate the most efficient coding method, reliable subjective and objective metrics are needed in assessing its quality. Unlike the conventional 2D image where the quality of the decoded image is not much different from that of displayed image to users, the LF image has an additional rendering process which users can directly interact with. Therefore, one needs to evaluate the quality of LF image after rendering. For example, Fig. 1 shows the quality of LF Bikes image. Fig. 1(a) shows a perspective image after compression while Fig. 1(b) shows an image after refocussing. There is a large difference between two images in terms of PSNR or SSIM, which motivates us to think more about evaluating effects of compression differently according to specific rendering process.

As raw lenslet LF data can be converted to multi-view images (i.e., sub-aperture images (SAIs)), the quality of LF image after compression can be measured using conventional 2D quality metrics by averaging the PSNR or SSIM values of all SAIs. It deserves noting that while the quality of SAI is only affected by compression, that of the refocused image (RFI) is affected not only by compression but also by rendering process. It explains the large difference of qualities between the two images in Fig. 1. The process to do refusing involves shift-and-add which acts like an averaging filter removing the artifacts in the image [1]. The blurs are mixed due to compression and refocusing, so it is bit challenging to identify the sole impact of compression on the quality of RFIs. In this paper, we are using several state-of-the-art objective quality metrics, i.e., PSNR, SSIM [2], MSSSIM [3], FSIM [4], and IWPSNR/IWSSIM [5], to evaluate the quality of perspective and RFIs rendered from compressed image by HEVC and VP9 encoders. The subjective score in [6] is also used for metric validation.



(a) Center perspective *(31.92 dB, 0.874)*
(b) Refocused image *(42.88 dB, 0.981)*
Zoom out of (a)
Zoom out of (b)

Fig. 1. Quality of compressed Bikes LF image (0.02 bpp by HEVC). (a). Perspective image, and (b). Refocused background image. The numbers in bracket are values of average PSNR-Y and SSIM-Y.

## 2. Image Quality Assessment

In this paper, we use the VALID dataset in [6] which is compressed by using two encoders, HEVC and VP9, at different bitrate. The dataset also provides subjective scores for the entire perspective and RFI based on the Double Stimulus Impairment Scale (DSIS) methodology.

### A. *Objective quality evaluation of the perspective images*

PSNR of the perspective image is computed for $Y$ channel as:

$$pPSNR_Y(k,l) = 10log_{10}\frac{255^2}{MSE(k,l)} \quad (1)$$

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017R1A2B2006518).

where $k$ and $l$ are the indices of the SAI, and $MSE(k,l)$ of the SAI image corresponding to $(k,l)$ is computed as:

$$MSE(k,l) = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - R(i,j)]^2, \qquad (2)$$

where $m = 625$ and $n = 434$ indicating image size of SAI. $I(i,j)$ and $R(i,j)$ are respectively the $Y$ channel values of compressed and referenced SAI. The average perspective PSNR for Y channel of all SAIs is computed as:

$$pPSNR_{Y_{mean}} = \frac{1}{(K-2)(L-2)}\sum_{k=2}^{K-1}\sum_{l=2}^{L-1} pPSNR_Y(k,l), \qquad (3)$$

where $K = 15$ and $L = 15$ are respectively the number of SAIs in horizontal and vertical directions. In a similar fashion, pSSIM-Y, pMSSIM-Y, pFSM-Y, pIWPSNR-Y, and pIWSSIM-Y are computed on the Y channel the same as pPSNR-Y described above.

*B. Objective quality evaluation of the refocused images*

The quality of the RFI is the average quality of the focal stack (i.e., a set of RFIs is called the focal stack). In this paper, we render eleven refocused images ($N = 11$) for each LF image, then the average PSNR of the focal stack is computed:

$$rPSNR_{Y_{mean}} = \frac{1}{N}\sum_{k=1}^{N} rPSNR_Y(k) \qquad (4)$$

where $rPSNR_Y(k)$ is the PSNR of the $k$th RFI of the compressed image. Similarly, rSSIM-Y, rMSSSIM-Y, rFSIM-Y, rIWPSNR-Y, and rIWSSIM-Y are computed the same process as rPSNR-Y.

## 3. Experimental Results

In order to compare the performance of objective metrics, we use four popular metrics: Person linear correlation coefficient (PLCC), Spearman rank-order coefficient (SROCC), Kendall rank-order correlation coefficient (KROCC), and root mean squared error (RMSE). The $\boldsymbol{MOS_p}$ values are predicted from the objective values using the logistic function, computed as [7]:

$$\boldsymbol{MOS_p(i) = b_1 + \frac{b_2}{1 + \exp(-b_3 \times (MR(i) - b_4)}} \qquad (5)$$

where $MOS_p(i)$ is a representation of the predicted MOS for the $i$th image. $b_j(j = 1,2,3,4)$ is the regression parameter and $MR$ is an objective metric result. $b_1$ to $b_4$ are the $MOS_{min}$, $MOS_{max}$, $MR_{min}$, $MR_{max}$. The predicted MOS ($MOS_p$) is compared to the ground truth MOS values provided by subjective quality assessment.

*A. Quality assessment of perspective images*

Table 1 shows the predicted MOS score of six state-of-the-art objective quality metrics in terms of PLCC, SROCC, KROCC, and RMSE. As can be seen, the performances of six objective metrics are quite similar. However, IWPSNR has the best performance compared to the other metrics. In [7], the author found that FSIM is the most reliable metric for the perspective images. However, in our work, we observe that IWPSNR is slightly better than FSIM in both perspective and RFI images. The popular PSNR and SSIM also have very similar performance compared to IWPSNR. This implies that they are also good quality metrics for quality assessment of the perspective images.

*B. Quality assessment of refocused image*

Table 1. Measures between $MOS_p$ & $MOS$ values.

| | | PLCC | SROCC | KROCC | RMSE |
|---|---|---|---|---|---|
| **Perspective images** | pPSNR-Y | 0.976 | 0.967 | 0.850 | 0.315 |
| | pSSIM-Y | 0.956 | 0.956 | 0.815 | 0.389 |
| | pMSSSIM-Y | 0.966 | 0.961 | 0.839 | 0.344 |
| | pFSIM_Y | 0.978 | 0.969 | 0.867 | 0.275 |
| | pIWPSNR-Y | **0.980** | **0.970** | **0.870** | **0.263** |
| | pIWSSIM-Y | 0.976 | 0.96 | 0.8623 | 0.289 |
| **Refocused images** | rPSNR-Y | 0.983 | **0.969** | **0.870** | 0.244 |
| | rSSIM-Y | 0.978 | 0.9652 | 0.8494 | 0.2786 |
| | rMSSSIM-Y | 0.975 | 0.955 | 0.828 | 0.294 |
| | rFSIM_Y | 0.981 | 0.956 | 0.839 | 0.253 |
| | rIWPSNR-Y | **0.987** | 0.969 | 0.864 | **0.210** |
| | rIWSSIM-Y | 0.978 | 0.959 | 0.846 | 0.274 |

**Best value shown in* ***bold***

Table 1 also shows which objective value can predict the MOS ($MOS_p$) values well enough to match the MOS of the ground truth. Different from the perspective image case, IWPSNR only shows the best performance in terms of PLCC and RMSE. PSNR has the best performance in terms of SROCC and KROCC. This means that the popular PSNR can be used to assess the quality of the refocused image.

## 4. Conclusions

In this paper, we evaluate the quality of the perspective image and refocused image using six different objective quality metrics to see the most suitable metric for quality assessment of these images corrupted by compression and rendering. The results show that IWPSNR has the best performance but not outperforms the other popular metrics such as PSNR or SSIM.

## References

[1] R. Ng, et al., "Light field photography with a hand-held plenoptic camera," in Computer Science Technical Report, vol. 2, no. 11, pp. 111, 2005.

[2] W. Zhou, et al., al.hou, has the best performance but nplenoptic camera. 11, psimilarity," in IEEE Trans. Image Process., vol. 13, no. 4, pp. 600camera. 11, psim

[3] Z. Wang, et al., al.ang, . Image Process., vol. 13, no. 4, pp. 600camera. 11, ps in Proc. of The thirty-seventh Asilomar Conf. on Signals Systems Computers, vol. 2, pp. 1398. 600camera. 11,

[4] L. Zhang, et al., "Fsim: A feature similarity index for image quality assessment,enin IEEE Trans. Image Process., vol. 20, no. 8, pp. 2378ex for image qua

[5] Z. Wang, et al., et al.g, no. content weighting for perceptual image quality assessment," in IEEE Trans. Image Process., vol. 20, no. 5, pp. 1185e quality assess

[6] I. Viola, et al., I. Viola, et al.Process., vol. 20, no. 5, pp. 1185e quality assessmin Proc. of the 10th Inter. Conf. on Quality of Multimedia Experience (QoMEX), Sardinia, Italy, 2018.

[7] Amirpour, et al., 2018.2018.rience (o. 5, pp. 1185e quality assessment," ,".11, 2005. SciencAssessment, et aProc. of the 2019 IEEE Inter. Conf. Acoustic, Speech, Signal Process (ICASSP), pp. 2402-2406, May 2019.