https://doi.org/10.2352/EI.2023.35.8.IQSP-262
© 2023, Society for Imaging Science and Technology

Evaluation of Motion Blur Image Quality in Video Frame Interpolation

Hai Dinh, Qinyi Wang, Fangwen Tu, Brett Frymire, and Bo Mu; OmniVision Technologies, Santa Clara, California 95054, USA

Abstract

Slow motion video has become a standard feature in mainstream cell phones. Due to constraints in hardware, such as power consumption, limited memory or data transfer throughput, a high capture framerate is not always feasible. As a solution, video frame interpolation (VFI), a software-based approach, has been widely employed. Conventionally, researchers use relative image quality merits between ground-truth and reconstructed frames, such as mean square error (MSE), peak signal-to-noise ratio (PSNR) [6], or structural similarity index measure (SSIM) [7], to evaluate VFI algorithm implementations. When the ground-truth data with specified high frame rate is not available, these metrics cannot be applied. For video interpolation, especially for fast moving objects, motion blur as well as ghosting are more significant to the audience subjective judgment. We developed an objective method and an apparatus that can rapidly assess such video quality factors of an interpolated slow motion video without the dependence on ground-truth data.

This paper discusses the development of an apparatus, designed to evaluate motion blur, and the analysis of blurred edge width (BEW) metric in quality assessment for VFI.

Introduction

Slow motion videos are those captured at a high frame rate and played back at normal speed, e.g. 30 frame-per-second (fps). It has become a standard feature in mainstream cell phones. An ever increasing capture frame rate might be desirable. However, this comes at the expense of increased power consumption and bandwidth requirement of mobile phone camera sensors, application processors and memory. If a frame rate higher than 480 fps is expected, a software-based approach can be considered [2][3][4]. Video Frame Interpolation (VFI) is a software-based method to create and insert synthesized frames, referred as interpolated frames, between two consecutive captured frames, referred as key frames. VFI can be utilized for video compression [1], slow motion generation, etc. These approaches present reasonable performance at the cost of significant computation resource requirement, since a bi-directional optical flow or even image depth are calculated. [5] introduces a more efficient framework to directly estimate the intermediate flow and blending mask within one network instead of calculating an end-to-end flow followed by backward warping. Evaluation of these algorithms with reliable quantitative metrics is challenging since most of them are designed for normal speed videos. Fast motion videos with obvious motion blur bring additional challenges for frame interpolation performance evaluation using conventional global metrics. It is noteworthy that some smartphone manufacturers refer to videos produced using VFI as "super slow-motion". In this paper, we only assess the quality of VFI-based videos. Hence, the term super slow motion and slow motion are sometime used interchangeably but both will refer to slow motion videos that are created with VFI.

Motion blur [9] in a captured frame is caused by a combination of pixel exposure and object movement. During the shutter exposure time, a moving object is captured as the integration of all positions on its trajectory. As a result, this object will appear blurred, which makes it difficult to differentiate objects. Motion blur has different behavior than optical-related blur such as out-of-focus or lens aberration.

In this work, we propose a test fixture design to evaluate motion blur quality of camera modules and video interpolation algorithms. We aim to solve the problem of slow motion video quality metrics by proposing a general approach based on blurred edge evaluation. The goal of this method is to provide an objective measurement of slow motion video quality, so that it could be used as an efficient baseline score for further evaluation.

Related Work

Conventionally, the quality of video frames can be evaluated using pixel-based metrics like mean square error (MSE), peak signal-to-noise ratio (PSNR), or structural similarity index (SSIM). However, these metrics have two drawbacks: (i) they require reference frames and (ii) they don't correlate well with the motion blur level [8].

For camera system performance, the modulation transfer function (MTF) is a well-established measurement for sharpness. There are various methods for MTF evaluation. The most common method is the slanted-edge, which derives the MTF from a standardized test chart. Recent work [8] have examined MTF of VFI and show the sharpness loss in the interpolated frames.

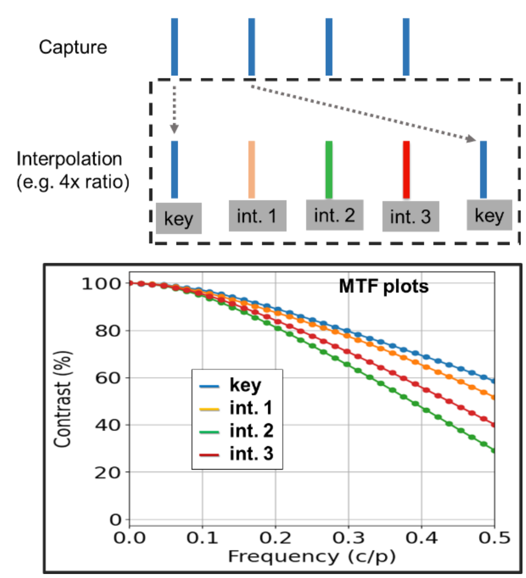


Figure 1: *Sharpness loss in interpolated frames [8].*

This is seen in Figure 1, where the MTF was measured from interpolated frames in comparison to key frames. It shows that interpolated frames with larger distance to their key frames exhibit lower MTF. MTF evaluation based on slanted edge method is appropriate for translational motion, but exhibits challenges for rotational motion.

Apparatus

We designed a test apparatus to study the motion blur in slow motion videos capturing moving objects. The requirements are:

1. Compact for repeatable experimentation with various targets
2. The sensor and phone holder can be easily swapped.
3. The target speed can be controlled
4. The luminance is adjustable

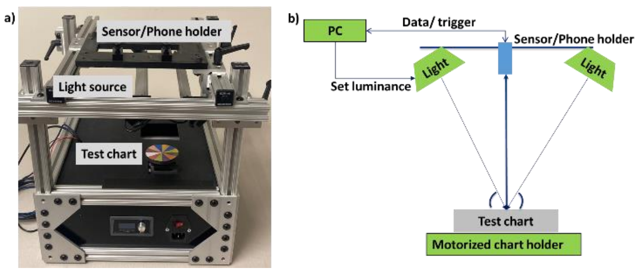


***Figure 2:** (a) front-view of the apparatus. (b) schematic of the apparatus setup.*

Figure 2 depicts the apparatus designed for capturing videos for motion blur analysis. A test chart depicting a colored Siemens star mounted on a motorized plate is used for the evaluation. The rotation speed can be adjusted from 0 to 300 rounds-per-minute (rpm). Two daylight light-emitting diode (LED) floodlights with 5600 K color temperature are placed at 45° angle with respect to the test chart normal. These LEDs are controlled by a constant current LED driver which allows adjustable luminance without flickering. The illuminance, at the test chart could be adjusted from 0 to 12,000 lux. The tested mobile phone or camera is held by a holder that can be moved in left-right and up-down direction. The holder is placed 11 cm away from the test chart. Videos are captured by rotating the disk while keeping the holder still.

Calculate Blurred Edge Width by Curve Fitting

The intensity profile across the edge is evaluated using the Blurred Edge Width (BEW), which is defined as the width (in pixels) along which the intensity changes from 10 to 90 percent of its local range. Frame interpolation and sharpening algorithms may introduce extra noise or over/under-shooting along these edges. As shown in Figure 3(a), different curves sampled from the same edge from different frames may have different minimum and maximum references, which causes inconsistency when calculating 10 to 90 percent values (blue and red lines). Moreover, noise or over/under-shooting also leads to multiple 10 to 90 percent values (red line). As a result, it is difficult to obtain valid BEW for a fair comparison. To solve the above mentioned issues and obtain an accurate estimation of BEW, we introduce a curve-fitting method. The target function selected for curve fitting is a sigmoid function. A few parameters are added to tune the shape of the sigmoid function for better fitting:

$$S(x) = |Y1 - Y2| \frac{1}{1 + e^{-(ax+b)}} + \min(Y1, Y2).$$

Here, Y1 and Y2 are the minimum and maximum luma values, |Y1-Y2| is the edge contrast and min(Y1, Y2) controls the up/down shift of the fitting curve. The fitting parameters a and b are introduced to control the slope and position of the edge obtained by regression among sampled edge points. As shown in Figure 3(b), the proposed curve fitting method is capable to determine the edge width correctly even in noisy condition.

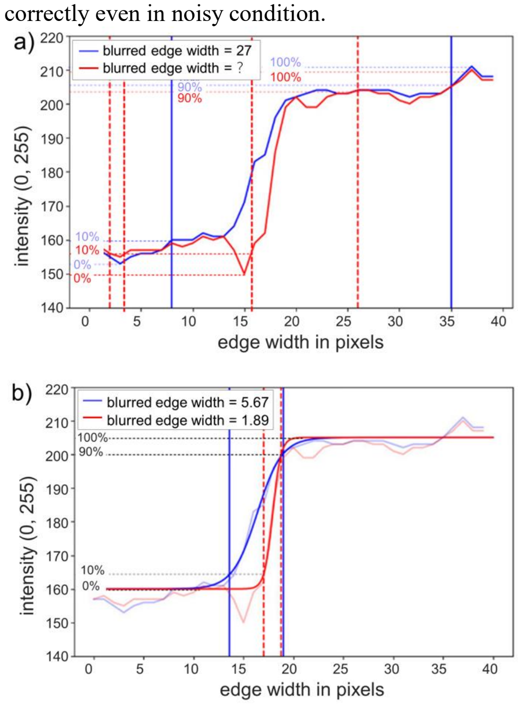


***Figure 3**: Blue and red curves are sampled intensity profiles from the same edge but different frames. (a)BEW without curve fitting. (b)BEW with proposed curve fitting*

Experimental Setup

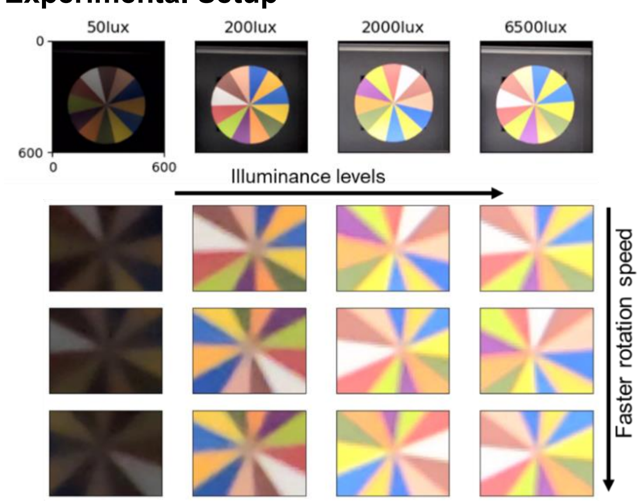


***Figure 4:** Sample of data collected from a phone with multiple illuminance levels and rotation speed.*

A colored Siemens star is imprinted on a motorized disk surface. The disk can be controlled to rotate at different speeds. Two flag-ship mobile phones – phone A (7680 fps, 720p, 4x frame rate interpolation) and phone B (960 fps, 720p, 2x frame rate interpolation) are used in the evaluation. Different ambient light conditions are also evaluated. A color star test chart is chosen instead of a gray-scale target in order to study color artifacts commonly observed in interpolated videos. Figure 4 depicts slow motion video frames captured under varying ambient light conditions and rotation speeds.

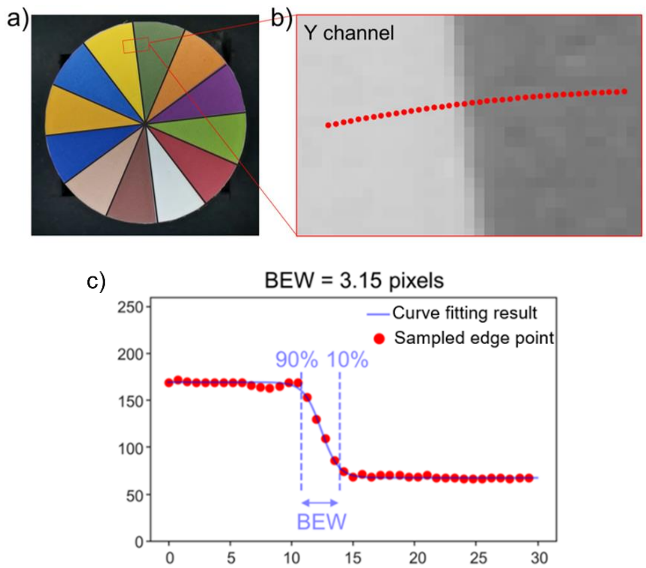


Figure 5: *An automated pipeline to calculate BEW with given input test frame. (a) auto-edge detection, (b) automatically generated samples perpendicular to the edges, and (c) apply the proposed curve fitting method to calculate BEW.*

Figure 5 illustrates key aspects of the algorithm calculating the BEW. Firstly, an edge detection module is introduced to identify the edges of the color disk. Secondly, a series of sample points with a total length of 30 pixels perpendicular to the target edge at any given radius is generated. Thirdly, the proposed curve fitting method is applied to calculate BEW based on the sampled intensity profile.

Analysis of Experimental Results

First, the BEWs between key frames and interpolated frames are analyzed. The results shown in Figure 6(c) and 6(d) are the average BEWs among a series of frames from two slow motion videos recorded under different light conditions. The two plots both demonstrate that the key frames always have smaller BEWs than the adjacent interpolated frames. These results agree with [8] that the sharpness of interpolated frames is often reduced. The results shown in Figure 6(a) and 6(b) are BEWs of the same edge captured under different lighting conditions. The change in BEW can be explained by auto exposure forcing a longer exposure at reduced scene luminance.

Figure 7 shows the overall BEW distribution from videos recorded under different light conditions and motion speed. It appears that BEW is proportional to motion speed and exposure time of the key frames. It also appears that smaller interpolation ratios result in larger the BEW.

Motion Speed

To evaluate the effect of motion speed on BEW, we calculate the BEW at different radii as well as rotational speeds. In Figure 8, each dot represents the average BEW from every edge of every frame from slow motion videos recorded under the exact same condition. Within each subplot of Figure 8(a)-(d), the dots with the same color have different linear speed while keeping all other factors (exposure time, interpolation ratio and interpolation algorithm) the same. It can be observed that an increased linear speed leads to larger BEW. It is worth mentioning that the flat regions in Figure 8(a) (b) are limited by the spatial resolution of the edge in the frame. The change of edge width from 76rpm to 228rpm is only approximately 1 pixel due to the high initial key frame rate of phone A. One limitation of BEW is the incapability of differentiating sub-pixel level edge width accurately.

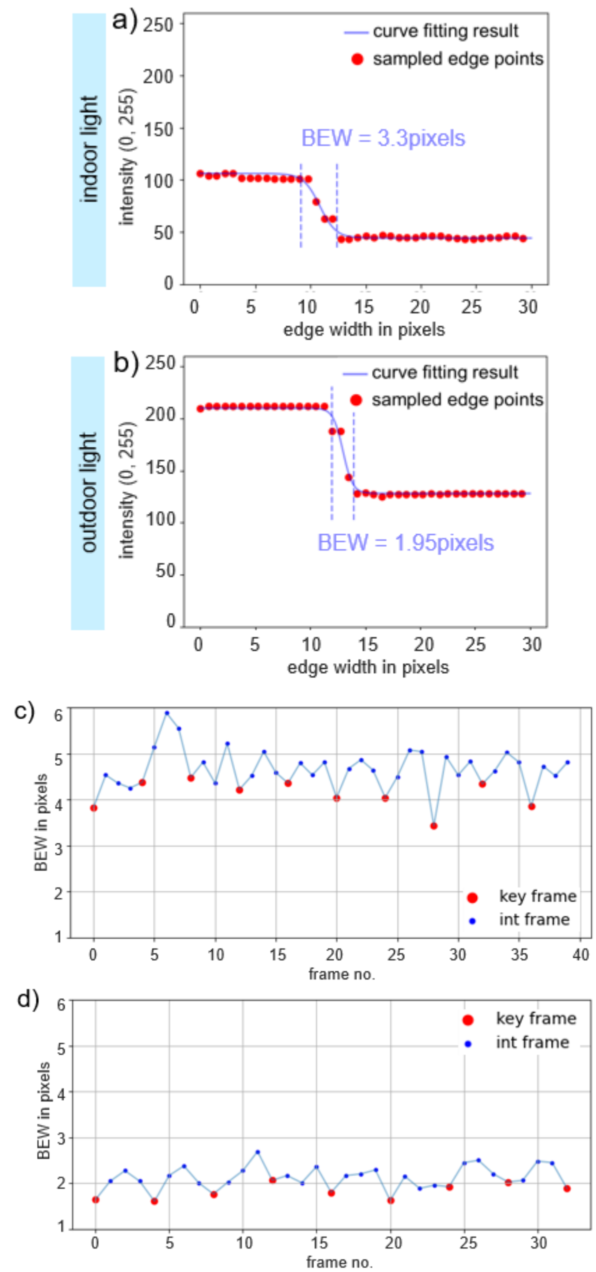


Figure 6: *(a) BEW of an edge under the indoor light condition with longer exposure time. (b) BEW of an edge under the outdoor light condition with shorter exposure time. (c) Average BEW of a series of frames from a slow motion video recorded under the indoor light condition. (d) Average BEW of a series of frames from a slow motion video recorded under the outdoor light condition.*

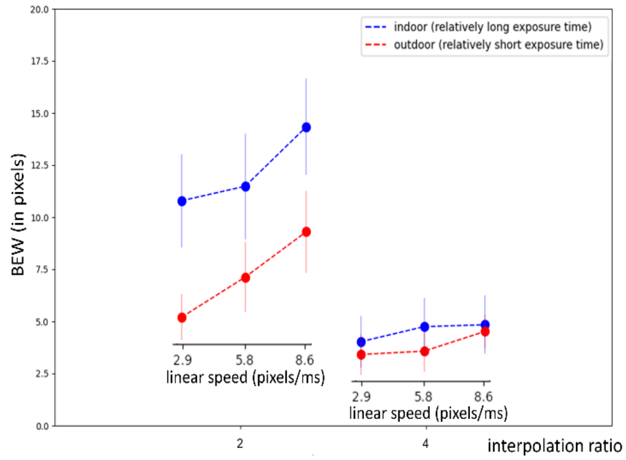
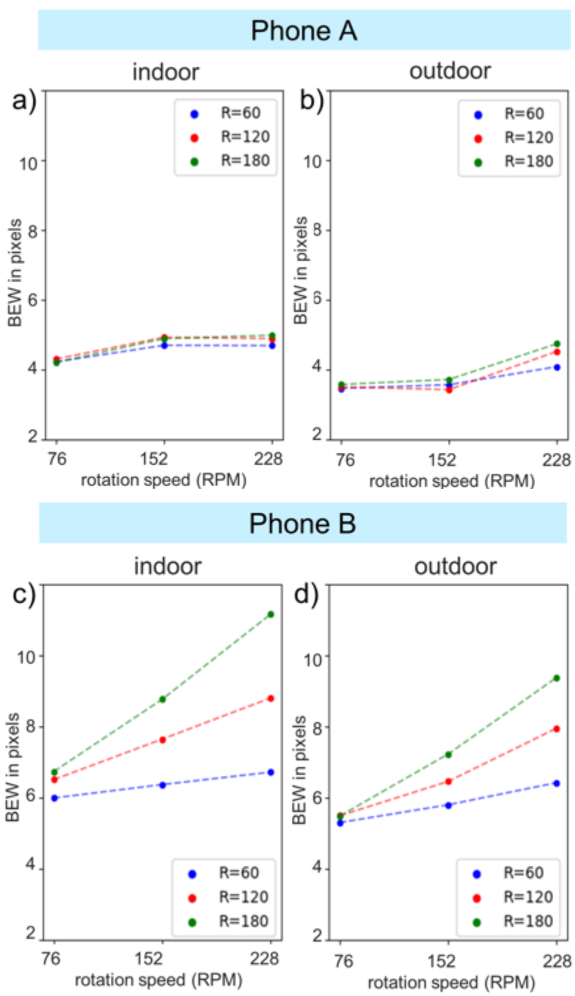
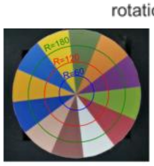


Figure 7: *The overall BEW distribution from videos recorded under different light conditions and motion speed of the two phones.*





	Linear speed (pixels / ms)		
	76 RPM	152 RPM	228 RPM
R = 60 pixels	0.5	0.9	1.4
R = 120 pixels	0.9	1.8	2.8
R = 180 pixels	1.4	2.8	4.3

Figure 8: *Effect of motion speed on BEW.*

Exposure Time

To evaluate the effect of exposure time on BEW, we calculate the BEW of slow motion videos recorded for both indoor and outdoor conditions.

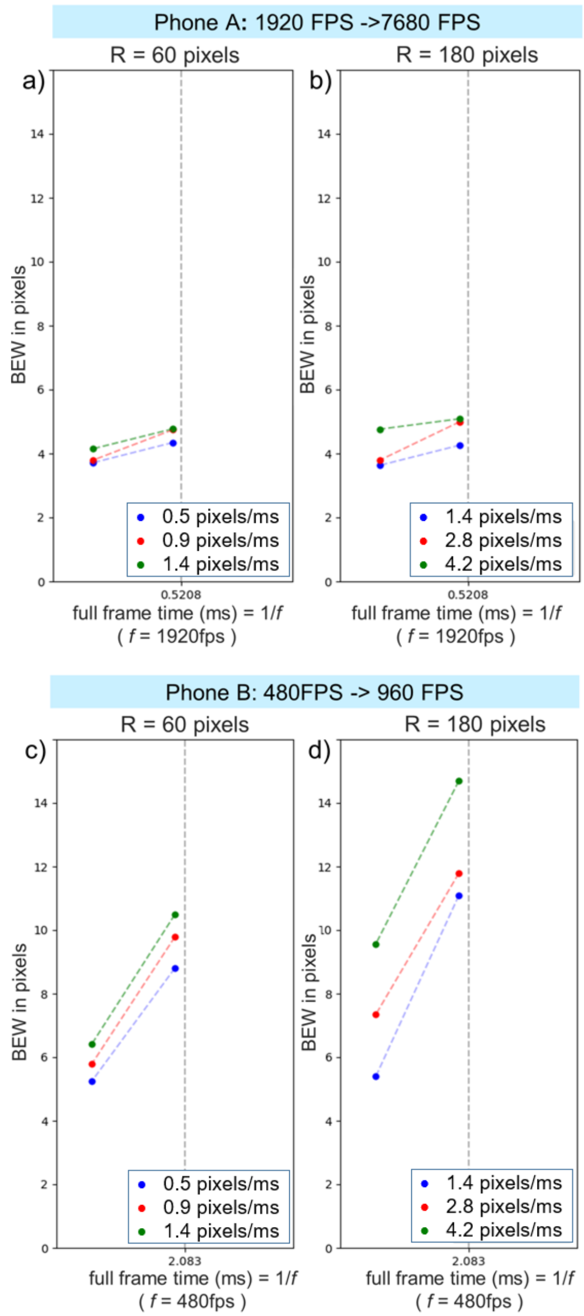


Figure 9: *Effect of exposure time of key frame on BEW.*

In each subplot of Figure 9, the vertical dash line indicates the full frame time $= 1/f$, where f is the initial frame rate before interpolation. As the accurate exposure time from each video is not contained in its meta information, the full frame time is used as a reference. The dots closer to the reference line are BEWs from videos recorded under outdoor light conditions while the dots further to the reference line are BEWs from videos recorded under indoor light conditions. When comparing each indoor and outdoor pairs within each subplot keeping all other factors (speed, interpolation ratio and interpolation algorithm) the same, it can be concluded that longer exposure time leads to larger BEW as is to be expected.

Interpolation ratio

The phones used in the test do not allow customized interpolation ratios. To achieve fair comparison, we interpolate key frames of videos captured by phone using the RIFE algorithm [5] under varying ratio of 2, 4, and 8. When comparing each dashed line connected by the dots within each subplot keeping all other factors (speed, exposure time and interpolation algorithm) the same, the interpolation ratio does not show very obvious influence on the BEW.

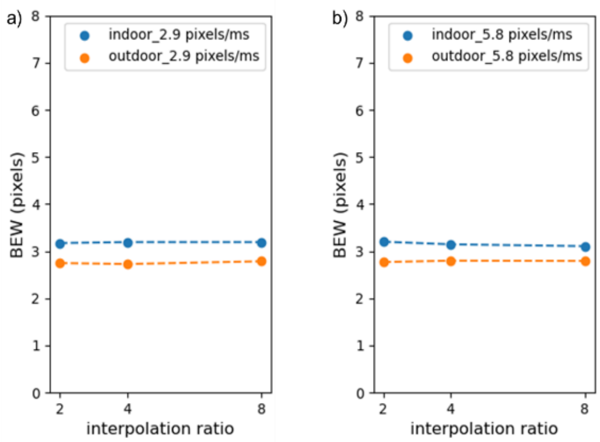


Figure 10: *Effect of interpolation ratio on BEW.*

Conclusion

We proposed an apparatus and a general objective approach based on BEW evaluation for video frame interpolation.

The reported test fixture has adjustable illumination and motion speed control. The ability to swap targets expands the range of scenarios that can be captured, and the camera module holder has been designed to accommodate a wide variety of modules and phones. This versatility makes it an ideal tool for evaluating the quality of super slow-motion video in a variety of settings. We reported results obtained and analyzed using the proposed approach on two flagship phones. The result shows that BEW of slow motion videos is impacted by the motion speed of the objects and the exposure time of key frames. No obvious change of BEW was observed when changing the interpolation ratio of the RIFE algorithm between 2 and 8 for the given captured key frames.

Our approach could be further integrated to use human visual system-dependent metrics such as contrast sensitivity function in spatial and temporal domain. Various VFI algorithms could be also investigated. Furthermore, a method to improve BEW accuracy for edges at subpixel level could be studied in future work.

References

[1] C.Y. Wu, N. Singhal and P. Krahenbuhl, "Video compression through image interpolation," in *Proceedings of the European conference on computer vision (ECCV). 2018: 416-431.*

[2] H. Jiang, D. Sun, and V. Jampani,"Super slomo: High quality estimation of multiple intermediate frames for video interpolation", in *Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 9000-9008.*

[3] W. Bao, W.S. Lai, C. Ma, "Depth-aware video frame interpolation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 3703-3712.*

[4] F. Reda, J. Kontkanen, E. Tabellion, "FILM: Frame Interpolation for Large Motion," in *arXiv:2202.04901, 2022.*

[5] Z. Huang, T. Zhang, W. Heng, "Rife: Real-time intermediate flow estimation for video frame interpolation," in *arXiv:2011.06294, 2020.*

[6] Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM[C]//2010 20th international conference on pattern recognition. IEEE, 2010: 2366-2369.

[7] Sara U, Akter M, Uddin M S. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study[J]. Journal of Computer and Communications, 2019, 7(3): 8-18.

[8] L. Luo, C. Yurdakul, K. Feng, D.E. Seo, F. Tu, B. Mu, " Temporal MTF evaluation of slow motion mode in mobile phones," in Proc. IS&T Int'l. Symp. on Electronic Imaging: Image Quality and System Performance

[9] Tiwari S, Shukla V P, Singh A K, et al. Review of motion blur estimation techniques[J]. Journal of Image and Graphics, 2013, 1(4): 176-184.

Author Biography

Hai Dinh *received his M.S. degree in Electrical and Computer Engineering from University of Minnesota, Duluth, USA in 2011. Currently he is a Sr. Imaging System Engineer at OmniVision Technologies in Santa Clara, California, USA. His research interests include digital color imaging and image/video quality evaluation.*

Qinyi Wang *received her B.E. degree from Nanyang Technological University, Singapore, 2015 and Ph.D. degree in Electrical and Electronic Engineering from Nanyang Technological University, Singapore, 2021. She is currently working as a Staff Algorithm Engineer in OmniVision Technologies Singapore Pte. Ltd. Her research interests include event-driven computer vision and sensor fusion.*

Fangwen Tu *received his B.E. degree from Dalian University of Technology, Dalian, China, 2012 and Ph.D. degree in Electrical Engineering from National University of Singapore, Singapore, 2017. He is currently working as a Staff Algorithm Engineer in OmniVision Technologies Singapore Pte. Ltd. His research interests include machine learning, image processing, and sensor fusion.*

Brett Frymire *graduated from DeVry University, Phoenix, AZ, USA in 1978. He is currently a Senior Hardware Test Engineer, OmniVision Technologies in Santa Clara, California, USA. His research interests include imaging sensor and image quality metric.*

Bo Mu *received his Ph.D. degree in Imaging Science from Rochester Institute of Technology, Rochester, NY, USA in 2007. He is currently a Director of Algorithm Development, OmniVision Technologies in Santa Clara, California, USA. His research interests include color image and video processing, computer vision, computational imaging and image quality metric.*