

# INTERPOLATION ERROR AS A QUALITY METRIC FOR STEREO: ROBUST, OR NOT?

Jiangbo Lu<sup>\*,†</sup>, Qiong Yang<sup>\*</sup>, and Gauthier Lafruit<sup>†</sup>

<sup>\*</sup> Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium

<sup>†</sup> Multimedia Group, IMEC, Kapeldreef 75, B-3001, Leuven, Belgium

## ABSTRACT

To properly benchmark and stimulate current stereo algorithms specifically in the application context of view interpolation, a robust quantitative evaluation approach is important. As a prevailing quality assessment method, interpolation error has been widely used. It measures the distortions between an interpolated image and a real camera image for a desired virtual viewpoint. However, is it a robust quality metric, especially when state-of-the-art stereo technology is developing so fast? This paper hence focuses on revealing several rarely attended weaknesses that make the interpolation error evaluation paradigm vulnerable. In addition, we propose an alternative evaluation method as an early attempt at addressing these challenges, from a perspective of communication system. Evaluation of representative stereo methods from the Middlebury website shows that the new approach yields consistent quality assessment outcomes.

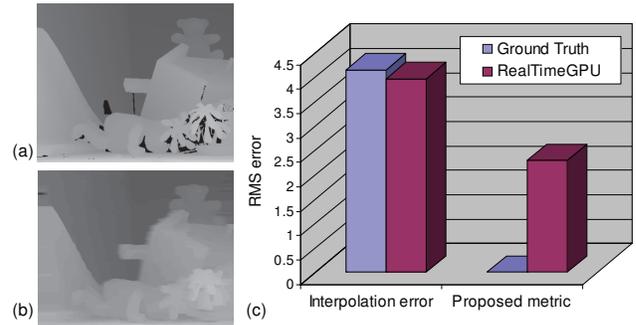
**Index Terms**— Stereo, view interpolation, interpolation error, evaluation methodology, communication system

## 1. INTRODUCTION

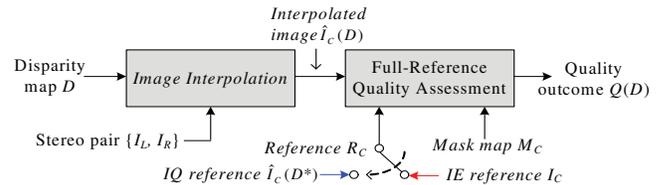
Quantitatively evaluating stereo algorithms for view interpolation is significant for many applications [1, 2]. For instance, in image-based rendering, researchers are not as directly concerned with error in disparity [3] as they are in the error in color values for the interpolated image. Hence, an automatic objective evaluation approach for stereo estimation, oriented toward interpolated image quality assessment, is necessary. Another example is free viewpoint video and 3D television [4]. Such an evaluation enables one to study the effects of various depth-image compression methods on rendered virtual views.

Interpolation error [2] was proposed as such a quality metric for stereo benchmarking. Given a disparity map  $\mathcal{D}$  estimated from a stereo image pair  $\{\mathcal{I}_L, \mathcal{I}_R\}$ , it first interpolates a user-specified virtual view image  $\hat{\mathcal{I}}_C(\mathcal{D})$ . Then,  $\hat{\mathcal{I}}_C(\mathcal{D})$  is compared against a real camera image  $\mathcal{I}_C$  captured at the same viewpoint. Without loss of generality, this paper focuses on the center viewpoint hereinafter.

As a prevailing quality metric that has been widely adopted to evaluate stereo and motion [1, 2], interpolation error has a few advantages. First, it more closely matches the requirement of the ultimate task (i.e., the quality of the interpolated image), compared with the disparity-based evaluation [3]. Second, it is a full-reference based image quality assessment approach [5], yielding quantitative evaluation outcomes on a well-understood ground. Third, quality assessment is easy to perform, without involving any sophisticated computation nor human intervention. However, despite of these advantages, we observed that it is *not always* robust to produce satisfactory quality outcomes. As a motivating example, Fig. 1 shows that an inaccurate disparity map of RealTimeGPU algorithm leads to a less root mean square (RMS) error than the ground-truth map. This result violates the common intuition, and it remains hard to explain.



**Fig. 1.** An example of quality evaluation results using the (c) interpolation error metric and proposed metric. (a) Ground-truth disparity map of the *Teddy* image [6]. (b) The disparity map of RealTimeGPU algorithm, with a nonoccluded disparity error rate of 7.23% [6]. We measured the quality of interpolated center images in RMS error.

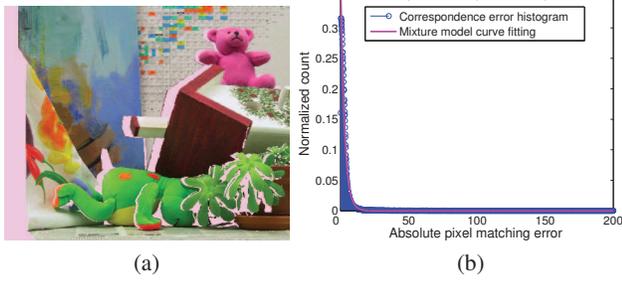


**Fig. 2.** System framework of the interpolation error metric ( $\mathcal{R}_C$  toggled to  $\mathcal{I}_C$ ) and proposed metric for stereo ( $\mathcal{R}_C$  toggled to  $\hat{\mathcal{I}}_C(\mathcal{D}^*)$ ).

Motivated by this perplexing result discovered in practice, we attempt at addressing the following questions in this paper. 1) What has caused the misleading quality assessment outcomes? Or, what are the weaknesses that have been overlooked in the interpolation error metric? On what conditions, they are no longer negligible? 2) Is there any pragmatic alternative to overcome the revealed limitations? Our exploratory answers to them are given in Sect. 2, 3, and 4.

## 2. INTERPOLATION ERROR METRIC REVISITED

To have a more structured revisit to the interpolation error (IE) metric, a schematic summary of this method is first shown in Fig. 2. For a given disparity map  $\mathcal{D}$ , its IE assessment outcome  $Q(\mathcal{D})$  is defined with three inputs, i.e., a reference image  $\mathcal{R}_C$ , the interpolated image  $\hat{\mathcal{I}}_C(\mathcal{D})$ , and a mask map  $\mathcal{M}_C$  that specifies the regions of evaluation interest. To generate the interpolated image  $\hat{\mathcal{I}}_C(\mathcal{D})$ , a baseline interpolation algorithm [2] is also presented in conjunction with the IE metric. Next, we discuss these three inputs in more details.



**Fig. 3.** The histogram of corresponding pixel errors in (b), measured between the registered right image in (a) and the left image  $\mathcal{I}_L$ . Holes are filled in light magenta, not counted in the histogram.

### 2.1. Weakness of the reference image $\mathcal{R}_C$

The IE metric uses the real captured camera image  $\mathcal{I}_C$  as the reference  $\mathcal{R}_C$ , and this image is assumed to have perfect quality. In fact, the interpolated image  $\hat{\mathcal{I}}_C(\mathcal{D})$  and the real camera image  $\mathcal{I}_C$  are formed from essentially different processes. The former is generated from the input stereo pair  $\{\mathcal{I}_L, \mathcal{I}_R\}$  and the estimated disparity map  $\mathcal{D}$ , however, the real camera image  $\mathcal{I}_C$  originates from the imaging of the scene under realistic environments. Because of the nonideal acquisition process, several inherent factors—such as imaging bias, sensor noise, lens blur, discrete image sampling, sub-pixel mis-registrations of cameras—make the observed color values of the corresponding pixels in  $\{\mathcal{I}_L, \mathcal{I}_R\}$  different from those observed by  $\mathcal{I}_C$ . This fact was also reported by Zhang and Seitz [7] and Strecha *et al.* [8]. Fig. 3(b) shows a histogram of pixel matching errors, measured between the registered right image (to the left viewpoint) and the left image  $\mathcal{I}_L$ . Consistent with the reported observation [7], the heavy tail and mixture model fitting in Fig. 3(b) indicate the violation of brightness constancy even between  $\mathcal{I}_L$  and  $\mathcal{I}_R$ . As a result, treating  $\mathcal{I}_C$  blindly as a perfect photometric reference brings nontrivial system bias to an evaluation method [1, 2].

Even worse, such a pre-captured reference at an intermediate viewpoint is mostly unavailable in 3DTV or free viewpoint video applications. Seeking a pragmatic approach that still enables ideal full-reference image quality assessment is clearly desired [4].

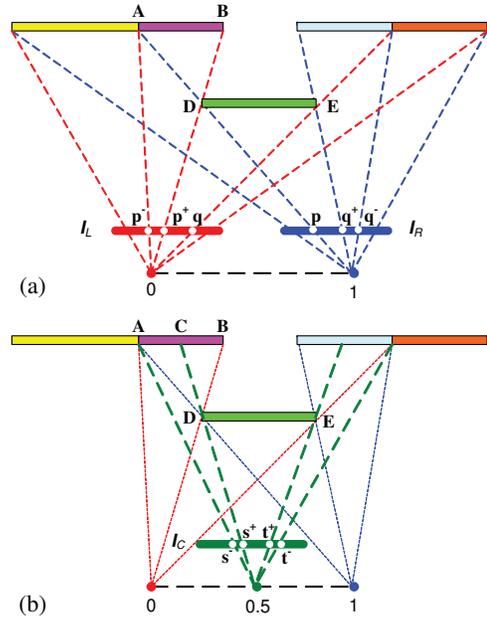
### 2.2. Weakness of the evaluation mask map $\mathcal{M}_C$

In the IE metric design, a region mask map  $\mathcal{M}_C$  is also used to exclude half-occluded pixels from evaluation. It effectively prevents the ill-posed pixels from contaminating performance evaluation, as the disparity for these pixels—visible in one input image—can only be inferred by enforcing some image priors. However, besides the half-occluded pixels, there are still many mixed pixels near depth discontinuities, which are marked as nonoccluded pixels in  $\mathcal{M}_C$ .

Fig. 4 gives a geometric explanation of these mixed pixels, e.g.,  $\mathbf{p}^+ \in \mathcal{I}_L$ ,  $\mathbf{p} \in \mathcal{I}_R$ , and  $\mathbf{s}^+ \in \mathcal{I}_C$ . They are located along depth discontinuities, and blend the contributions from both foreground and background colors. Based on the well-known compositing equation [9], the correspondence colors for  $\mathbf{p}^+$  and  $\mathbf{p}$  are modeled as

$$\begin{cases} \mathcal{I}_L(\mathbf{p}^+) = \alpha F_{\mathbf{D}} + (1 - \alpha) B_{\mathbf{B}} \\ \mathcal{I}_R(\mathbf{p}) = \alpha F_{\mathbf{D}} + (1 - \alpha) B_{\mathbf{A}} \end{cases} \quad (1)$$

where  $F$  and  $B$  represent the irradiance of foreground and back-



**Fig. 4.** Geometric view of the mixed pixel problem. The projections and boundary rays of  $\{\mathcal{I}_L, \mathcal{I}_R, \mathcal{I}_C\}$  are in {red, blue, green}. The depth discontinuity pixel  $\mathbf{p} \in \mathcal{I}_R$  has a one-to-two mapping, i.e.,  $\{\mathbf{p}^-, \mathbf{p}^+\} \in \mathcal{I}_L$  and  $\{\mathbf{s}^-, \mathbf{s}^+\} \in \mathcal{I}_C$ . Similar for the pixel  $\mathbf{q} \in \mathcal{I}_L$ .

ground scene points, respectively. The 3D scene points involved are  $\mathbf{A}, \mathbf{B}, \mathbf{D}$ , and  $\alpha$  is the contribution ratio used to linearly blend between foreground and background. Given only two stereo images, solving for unknowns in (1), i.e.,  $\alpha, F_{\mathbf{D}}, B_{\mathbf{B}}$ , is an under-constrained problem [9]. Subject to the surmised values of  $\alpha$  and  $F_{\mathbf{D}}$ , the composition color  $\hat{\mathcal{I}}_C(\mathbf{s}^+)$  for the center image, in principle, has an infinite number of reasonable interpretations. As an example, the locations of mixed pixels for the *Teddy* scene are marked in black in Fig. 6(b). Similar to half-occluded pixels, these ill-posed mixed pixels should not be included in quality assessment, otherwise it brings another kind of system bias to the evaluation system. Since the amount of local intensity variation is usually very large along depth discontinuities, such a system error actually results in pronounced interpolation errors, making stereo evaluation severely biased and unreliable.

### 2.3. Weakness of the baseline interpolation algorithm

The IE metric integrates a baseline interpolation algorithm, which performs a depth-based image interpolation as shown in Fig. 2. For a given disparity map  $\mathcal{D}$ , it generates an interpolated image  $\hat{\mathcal{I}}_C(\mathcal{D})$  by blending a pair of corresponding colors from  $\{\mathcal{I}_L, \mathcal{I}_R\}$ . Unfortunately, this baseline algorithm results in notorious color bleeding [10] along background occlusion boundaries in  $\hat{\mathcal{I}}_C(\mathcal{D})$ . Fig. 5 shows such artifacts, even when the ground-truth depth map  $\mathcal{D}^*$  is used.

Again, we rely on Fig. 4 to account for the weakness of the baseline interpolation algorithm. To generate the pixel  $\mathbf{s}^- \in \hat{\mathcal{I}}_C(\mathcal{D})$ , the baseline algorithm simply averages the colors of the pixel  $\mathbf{p}^- \in \mathcal{I}_L$  and  $\mathbf{p} \in \mathcal{I}_R$ . However, it is clear from Fig. 4(a) that the mixed pixel  $\mathbf{p}$  is of a contaminated color, while  $\mathcal{I}_L(\mathbf{p}^-) = B_{\mathbf{A}}$  alone gives a clean and correct background color. The locations of such problematic pixels are marked as red or blue pixels in Fig. 6(b). Without a prudent handling of color bleeding artifacts, the baseline interpola-



**Fig. 5.** Color bleeding artifacts in the interpolated center image due to the inappropriate baseline interpolation algorithm [2].

tion algorithm brings extra system errors to each individual stereo method under evaluation, making the IE metric more vulnerable.

In summary, the system bias caused by these three weaknesses tends to make the IE metric unreliable in benchmarking stereo, when their impact is comparable with the true performance difference.

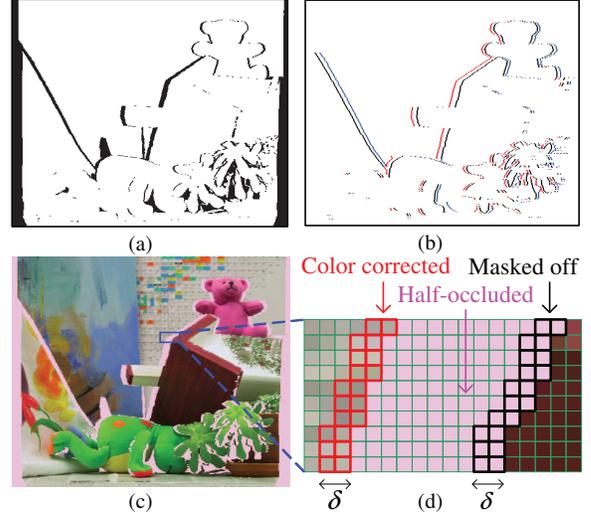
### 3. PROPOSED EVALUATION METHOD

To overcome the weaknesses of the IE metric, we have developed a new interpolation quality (IQ) evaluation metric, while preserving the primary advantages of IE discussed in Sect. 1. Correspondingly, our IQ metric makes the following three major contributions.

**Redefinition of the reference image.** Unlike the IE metric, which firmly assumes the real camera image  $\mathcal{I}_C$  is of photometrically perfect quality, our new evaluation philosophy is cast from a communication system perspective. Regarding image interpolation as a lossy processing channel, we define a pragmatic reference  $\hat{\mathcal{I}}_C(\mathcal{D}^*)$  given by the ideal lossless parameter, i.e., the ground-truth depth map  $\mathcal{D}^*$ , as shown in Fig. 2. Following the identical processing channel, for each depth map  $\mathcal{D}$  or  $\mathcal{D}^*$ , an interpolated image is generated, i.e.,  $\hat{\mathcal{I}}_C(\mathcal{D})$  or  $\hat{\mathcal{I}}_C(\mathcal{D}^*)$ . The quality of  $\hat{\mathcal{I}}_C(\mathcal{D})$  is hence quantified by the fidelity with the lossless reference  $\hat{\mathcal{I}}_C(\mathcal{D}^*)$ , namely the distortion caused by such a lossy channel. This new evaluation paradigm removes the negative impact of photometric deviations in Fig. 3(b), so the effects of disparity errors on the interpolated image can be more faithfully measured. Even if a ground-truth depth map  $\mathcal{D}^*$  is sometimes unavailable, the reference  $\hat{\mathcal{I}}_C(\mathcal{D}^*)$  can still be approximated via an accurate depth map estimated from  $\{\mathcal{I}_L, \mathcal{I}_R\}$ . In contrast, the IE metric does not work without a camera image  $\mathcal{I}_C$ .

**Redefinition of the evaluation mask map.** As discussed in Sect. 2.2, ill-posed mixed pixels should also be excluded for unbiased interpolation quality assessment, similar to half-occluded pixels recorded in the IE mask map  $\mathcal{M}_C$  (see Fig. 6(a)). To this end, mixed pixels need to be located for the requested center viewpoint, as exemplified by the pixel  $s^+$  and  $t^+$  in Fig. 4(b). Based on the warped ground-truth depth map  $\mathcal{D}^*$  and  $\mathcal{M}_C$ , we detect these mixed pixels as any visible pixels on depth discontinuities from the center viewpoint. Depth discontinuities are defined as any disparity jump greater than  $\lambda$  ( $= 2$ ) pixels [3]. In reality, a neighborhood of pixels along depth discontinuities are affected, so we consider mixed pixels as a strip with the width of  $\delta$  (with  $\delta$  empirically set to 2 for *Teddy*).

**An improved image interpolation algorithm.** As shown in Fig. 4(b), the color bleeding artifacts (e.g.,  $s^-$ ) are located on the opposite side of occlusion boundaries to the ill-posed mixed pixels (e.g.,  $s^+$ ). Therefore, the detection of color bleeding artifacts can be easily combined with the mixed pixels detection in a single step. Fig. 6(b) marks pixels of problematic colors—due to the baseline



**Fig. 6.** The proposed evaluation mask map is defined by (a) the IE mask map  $\mathcal{M}_C$ , plus the ill-posed mixed pixels in black in (b). Color bleeding artifacts are detected as red (and blue) pixels in (b), where a single left (or right) correspondence is used to generate (c) the interpolated image. (d) Close-up view of an artifact-free region.

interpolation algorithm—in red and blue for *Teddy*. Once these erroneous pixels are detected, our improved algorithm corrects them by filling the clean background colors from a single source image, i.e., pixels in red (blue) are filled with the left (right) pixel colors. Thanks to this simple yet effective scheme, the interpolated image based on the ground-truth depth map is now artifact free (see Fig. 6(c)).

### 4. EXPERIMENTAL RESULTS

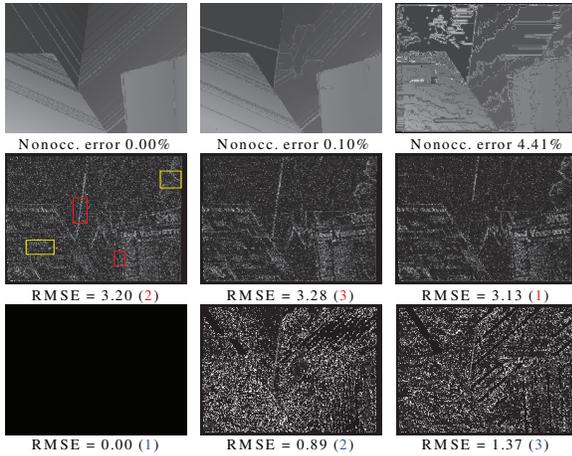
As the first comparative experiment, Fig. 1 shows quality assessment results on *Teddy* based on the IE and IQ metric. Addressing the weaknesses of the IE metric, the IQ metric yields consistent evaluation outcomes for the ground-truth depth map and that of Real-TimeGPU. We used the same RMS error measure in the IE metric.

Fig. 7 shows another comparison on *Venus*. It is clear that when the IE metric is applied, the ground-truth disparity map even results in non-trivial errors. Among them, the errors inside smooth depth surfaces (e.g., yellow rectangles in Fig. 7) are mainly caused by inappropriate use of the reference image. The errors along depth discontinuities are due to improper handling of ill-posed mixed pixels and color bleeding artifacts (e.g., red rectangles in Fig. 7). All these system errors together lead to the incorrect ranking of the three algorithms, making the worst algorithm to be ranked as best. The true defect of the *Infection* algorithm is not exposed. Compared with the IE metric, the IQ metric yields far more faithful evaluation results.

To have a more comprehensive understanding of two quality metrics for stereo, we have randomly selected nine stereo algorithms from the algorithms submitted to the Middlebury website [6]. These nine algorithms include the best performing and worst performing stereo algorithms, i.e., *AdaptingBP* and *Infection* as of Sep. 2008, as well as some representative methods ranked in between. In Table 1, we have reported the evaluation results of IE and IQ, along with the disparity-based metric. Based on the side by side comparison between the IE and IQ evaluation results, one can observe that these two metrics have significantly different quality assessment out-

**Table 1.** Evaluation of nine stereo algorithms based on the disparity error rate (Disp) for non-occluded areas [6], previous IE metric [2] measured in RMS error, and our IQ metric in RMS error. The superscript denotes the relative ranking among them. The IE results for *Tsukuba* are not applicable, since there is no real camera image available for the given stereo pair, but the proposed IQ metric works well.

Algorithm	<i>Tsukuba</i>		<i>Venus</i>			<i>Teddy</i>			<i>Cones</i>		
	Disp	IQ	Disp	IE	IQ	Disp	IE	IQ	Disp	IE	IQ
AdaptingBP	<b>1.11</b> <sup>1</sup>	1.72 <sup>2</sup>	<b>0.10</b> <sup>1</sup>	3.28 <sup>4</sup>	<b>0.89</b> <sup>1</sup>	<b>4.22</b> <sup>1</sup>	4.16 <sup>2</sup>	2.31 <sup>2</sup>	<b>2.48</b> <sup>1</sup>	5.98 <sup>2</sup>	3.14 <sup>2</sup>
OverSegmBP	1.69 <sup>4</sup>	1.92 <sup>4</sup>	0.50 <sup>2</sup>	3.45 <sup>8</sup>	1.24 <sup>2</sup>	6.74 <sup>2</sup>	4.68 <sup>5</sup>	3.03 <sup>4</sup>	3.19 <sup>2</sup>	6.26 <sup>3</sup>	3.30 <sup>3</sup>
AdaptWeight	1.38 <sup>3</sup>	1.80 <sup>3</sup>	0.71 <sup>3</sup>	3.36 <sup>7</sup>	1.53 <sup>6</sup>	7.88 <sup>4</sup>	4.58 <sup>3</sup>	3.01 <sup>3</sup>	3.97 <sup>3</sup>	<b>5.93</b> <sup>1</sup>	<b>2.97</b> <sup>1</sup>
GC + occ	1.19 <sup>2</sup>	<b>1.66</b> <sup>1</sup>	1.64 <sup>4</sup>	3.34 <sup>6</sup>	1.37 <sup>3</sup>	11.2 <sup>5</sup>	4.60 <sup>4</sup>	3.19 <sup>5</sup>	5.36 <sup>4</sup>	6.31 <sup>4</sup>	3.85 <sup>4</sup>
RealTimeGPU	2.05 <sup>6</sup>	2.14 <sup>5</sup>	1.92 <sup>6</sup>	3.24 <sup>3</sup>	1.50 <sup>5</sup>	7.23 <sup>3</sup>	<b>3.96</b> <sup>1</sup>	<b>2.29</b> <sup>1</sup>	6.41 <sup>5</sup>	7.56 <sup>9</sup>	4.96 <sup>9</sup>
GC	1.94 <sup>5</sup>	2.82 <sup>8</sup>	1.79 <sup>5</sup>	3.15 <sup>2</sup>	1.55 <sup>7</sup>	16.5 <sup>6</sup>	4.86 <sup>7</sup>	3.55 <sup>7</sup>	7.70 <sup>6</sup>	6.52 <sup>5</sup>	4.29 <sup>5</sup>
SSD + MF	5.23 <sup>8</sup>	3.37 <sup>9</sup>	3.74 <sup>7</sup>	3.30 <sup>5</sup>	1.56 <sup>8</sup>	16.5 <sup>6</sup>	5.31 <sup>9</sup>	3.79 <sup>8</sup>	10.6 <sup>7</sup>	6.95 <sup>6</sup>	4.41 <sup>6</sup>
SO	5.08 <sup>7</sup>	2.80 <sup>7</sup>	9.44 <sup>9</sup>	4.16 <sup>9</sup>	2.98 <sup>9</sup>	19.9 <sup>9</sup>	5.23 <sup>8</sup>	4.00 <sup>9</sup>	13.0 <sup>8</sup>	6.98 <sup>7</sup>	4.86 <sup>8</sup>
Infection	7.95 <sup>9</sup>	2.70 <sup>6</sup>	4.41 <sup>8</sup>	<b>3.13</b> <sup>1</sup>	1.37 <sup>3</sup>	17.7 <sup>8</sup>	4.77 <sup>6</sup>	3.32 <sup>6</sup>	14.3 <sup>9</sup>	7.40 <sup>8</sup>	4.83 <sup>7</sup>



**Fig. 7.** Quality assessment for the interpolated center *Venus* images using the disparity map of ground truth, *AdaptingBP*, and *Infection* (from left to right). Top row: Disparity maps and nonoccluded disparity error rates. Middle row: IE-based error maps and RMS errors (ranks). Bottom row: IQ-based error maps and RMS errors (ranks).

comes when disparity estimation errors are sufficiently small (e.g., *Venus*), while less difference in assessment outcomes when the test algorithms show a large gap in disparity performance (e.g., *Teddy* and *Cones*). As modern stereo algorithms evolve rapidly, the performance gap between different algorithms is steadily narrowed. For reliable benchmarking, quality metrics with strong performance discriminating power are hence increasingly necessary and desirable.

We also present the disparity-based evaluation results in Table 1. It is observed that view interpolation oriented quality metrics, either IE or IQ, do not necessarily have a high correlation with the disparity-based evaluation. For example, though *RealTimeGPU* has a moderate ranking among the test algorithms, it actually performs best on *Teddy* but also worst on *Cones* in terms of both IE and IQ assessment. The reason is that the “foreground fattened” depth maps of *RealTimeGPU* tend to be “healthy” to the scene of flat surfaces (e.g., *Teddy*), but they degrade the accuracy at depth discontinuities for scenes of subtle geometries, e.g., *Cones*. This fact pinpoints the importance of the interpolation-driven quality metrics for stereo.

## 5. CONCLUSION

In this paper, we have explored the inherent weaknesses of the interpolation error metric, which is commonly used for stereo evaluation in view interpolation applications. The revealed weaknesses suggest that a prudent design for each component in such a holistic evaluation system is not traditionally practiced. However, as stereo research advances, the systematic errors become increasingly distinct, making quality assessment outcomes fragile. As a first step towards addressing this challenge, we have developed a new, pragmatic quality metric that yields more consistent stereo evaluation results. We hope more creative exploration can be triggered to make quality assessment more robust and sensible in the future work. Besides that, we also plan to investigate the perceptual quality aspect.

## 6. REFERENCES

- [1] R. Szeliski, “Prediction error as a quality metric for motion and stereo,” in *Proc. of ICCV*, 1999, pp. 781–788.
- [2] S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. Black, and R. Szeliski, “A database and evaluation methodology for optical flow,” in *Proc. of ICCV*, 2007.
- [3] D. Scharstein and R. Szeliski, “A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,” *Int. J. of Computer Vision (IJCV)*, vol. 47(1), pp. 7–42, 2002.
- [4] A. Smolić and D. McCutchen, “3DAV exploration of video-based rendering technology in MPEG,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, pp. 348–356, Mar 2004.
- [5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Tran. IP*, vol. 13(4), pp. 600–612, 2004.
- [6] Middlebury stereo page, <http://vision.middlebury.edu/stereo/>.
- [7] L. Zhang and S. M. Seitz, “Estimating optimal parameters for MRF stereo from a single image pair,” *IEEE Trans. Patten Anal. Machine Intell.*, vol. 29(2), pp. 331–342, 2007.
- [8] C. Strecha, R. Fransens, and L. V. Gool, “Combined depth and outlier estimation in multi-view stereo,” in *Proc. of Computer Vision and Pattern Recognition*, 2006, pp. II: 2394–2401.
- [9] Y.-Y. Chuang, B. Curless, D. H. Salesin, and R. Szeliski, “A Bayesian approach to digital matting,” in *Proc. of Computer Vision and Pattern Recognition (CVPR)*, 2001, pp. II, 264–271.
- [10] Y. Taguchi, B. Wilburn, and C. L. Zitnick, “Stereo reconstruction with mixed pixels using adaptive over-segmentation,” in *Proc. of Computer Vision and Pattern Recognition*, 2008.