

Visual Quality Assessment for Projected Content

Hoang Le¹, Carl Marshall², Thong Doan¹, Long Mai¹, Feng Liu¹

¹Portland State University ²Intel Corporation

Portland, OR USA Hillsboro, OR USA

{hoanl, thong, mtlong, fliu}@pdx.edu carl.s.marshall@intel.com

Abstract—Today’s projectors are widely used for information and media display in a stationary setup. There is also a growing effort to deploy projectors creatively, such as using a mobile projector to display visual content on an arbitrary surface. However, the quality of projected content is often limited by the quality of projection surface, environment lighting, and non-optimal projector settings. This paper presents a visual quality assessment method for projected content. Our method assesses the quality of the projected image by analyzing the projected image captured by a camera. *The key challenge is that the quality of the captured image is often different from the perceived quality by a viewer as she “sees” the projected image differently than the camera.* To address this problem, our method employs a data-driven approach that learns from the labeled data to bridge this gap. Our method integrates both manually crafted features and deep learning features and formulates projection quality assessment as a regression problem. Our experiments on a wide range of projection content, projection surfaces, and environment lighting show that our method can reliably score the quality of projected visual content in a way that is consistent with the human perception.

Keywords—Image quality assessment, projector-camera system, human visual experience

I. INTRODUCTION

Projectors provide a convenient way to display information and are now widely used in a wide variety of applications [1]. Besides standard deployments, such as those in classrooms and conference rooms, people are now developing more creative ways to design and deploy projector systems. For example, projectors are used in augmented reality applications to display information in various environments to enrich the visual experience [2], [3], [4]. Projectors are also deployed on mobile platforms for information display. In these applications, visual content needs to be displayed on arbitrary surfaces instead of dedicated projector screens.

The projection quality, however, is often compromised when projecting visual content onto the less ideal surfaces such as those with noticeable textures and highly specular reflections. Furthermore, strong ambient lighting and directional lighting can also lead to low-quality projection. Therefore it is important to understand the projection quality to enable automatic compensation [5] or automatic selection of a projection surface to support high-quality projection.

This paper aims to develop a quality assessment method for projected visual content. Like existing work [6], our method couples a camera with a projector to capture the

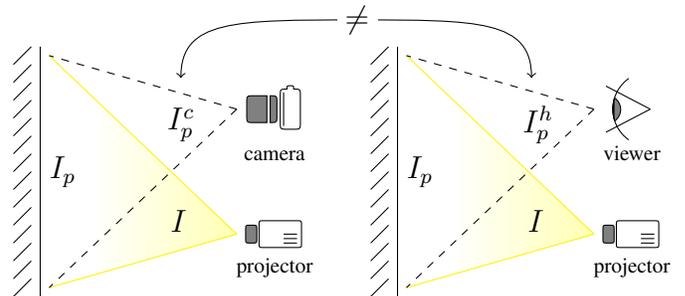


Figure 1. Challenge of projection quality assessment. The camera and the viewer perceive the projected image on the surface differently.

projected visual content. For simplicity, in the rest of this paper, we use image to refer to general visual content, such as photos, illustrations, tables, graphs, and documents. Our method uses the captured image and the original input image to evaluate the quality of the projected image. While there are many image quality assessment [7], [8], [9], [10], [11] and related aesthetics assessment methods available [12], [13], [14], [15], [16], [17], [18], [19], these methods cannot be directly applied due to a unique challenge in projection quality assessment: *the captured image I_p^c is often not what a viewer perceives when she looks at the projected image I_p* , as illustrated in Figure 1. The different sensor specifics of the camera and the human visual system make the viewer and the camera perceive the same projected image differently.

This paper presents a data-driven approach to bridge the gap between the quality of the captured projected image I_p^c and that of the perceived projected image I_p^h by viewers when they look at the projected image directly. Ideally, if the perceived projected image I_p^h is available, we can use the input image I as a reference to score the quality of the projected image. However, we only have the captured projected image I_p^c that is different from the perceived projected image I_p^h . Our solution is to use machine learning algorithms to learn to calibrate the difference between the captured projected image and the perceived projected image. To this end, we design a wide range of image features and then formulate projection quality assessment as a regression problem that learns from the labelled data to score the quality of the projected image.

To the best of our knowledge, this paper presents the first quality assessment method for projected visual content

that addresses the quality gap between the projected image captured by the camera and perceived by the viewer. Our experiments show that our data-driven approach can effectively learn to bridge this gap and assess the projected visual content in a way that is consistent with the human perception.

II. PROJECTED IMAGE QUALITY METRIC

We develop a projector-camera system to automatically assess the quality of the projected visual content. As shown in Figure 1, the projector projects an input image I onto the surface and produces a projected image I_p . Ideally, given the input image, we can assess the quality of the projected image using a state-of-the-art full reference quality metric as follows

$$Q(I_p) = Q_r(I_p^h, I), \quad (1)$$

where Q_r is a full-reference quality metric like [8], [10] and I_p^h is the perceived projected image by a human. In practice, a camera is adopted to capture the projected image and produces the captured projected image I_p^c , which is then used to approximate the perceived projected image I_p^h in Equation 1. However, it is difficult to obtain such a captured projected image that is a good approximation of I_p^h , as the camera and the viewer often see the projected image I_p differently, as discussed in Section II-A. To address this problem, we develop a data-driven approach to obtain a quality metric \hat{Q} that directly operates on the captured projected image and the input image and learns to bridge the gap between the perceived projected image and the captured projected image.

$$Q(I_p) = \hat{Q}(I_p^c, I), \quad (2)$$

Below we first briefly introduce our projector-camera system and then describe technical steps in our projection quality assessment method, including image rectification, feature extraction, and data-driven quality metric.

A. Projector-Camera System

We use an Epson 3-LCD projector to project visual content onto a planar surface in a perpendicular direction and a Nikon D7200 camera to capture the projected image. As shown in Figure 2, the camera is positioned next to the projector on a horizontal platform following a relevant design [20]. The baseline between them is 3.5 inches in our system. The projector-camera system is located about 45 inches away from the projection surface. The projection area is of size 22×35 inches. Note, the viewing experience when looking at a projected image depends on the viewing direction, especially on a surface with strong specular reflections. We position the camera close to the projector not only to make the system portable but also to simulate a typical scenario where the viewing direction for most of the viewers in a room does not significantly deviate from the projection direction.

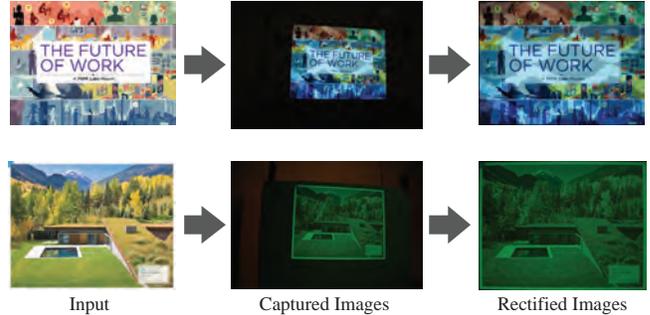


Figure 3. Image rectification. Note, the captured and rectified image in the bottom row appear green because the input image is projected onto a greenish surface.

We follow Zhao *et al.* to disable all advanced settings of the projector about brightness, contrast, and color enhancement [6]. We use the camera in the manual mode and adjust all its settings iteratively so that the captured images are not either underexposed or overexposed [6] in order to make the captured projected image as similar to what a viewer would see about the projected image as possible. However, we find that when the environmental lighting changes, it is very difficult to find a fixed camera setting that matches the captured projected image to what the viewer would perceive when she looks at the projected image. We address this problem by learning a quality metric to bridge this gap.



Figure 2. Projector-camera system.

B. Image Rectification

The first step of our method is to align the captured projected image I_p^c and the input image I . Since our method is targeted for applications such as automatic image compensation and automatic selection of projection surfaces where the projector-camera configuration may be changed to optimize the projection quality, we perform image rectification for each individual projected image instead of relying on pre-calibration for a fixed projector system. Specifically, we directly establish the correspondence between the captured projected image I_p^c and the input image I via image rectification. We first detect SIFT feature points in these two images and then use their feature descriptors to establish the feature correspondence [21]. Since the projection surface is a plane, we can robustly estimate a homography between these two images using a RANSAC approach [22]. This homography is finally used to warp the captured projected image to obtain pixel-wise alignment. We denote the rectified captured image as \tilde{I}_p^c . Figure 3 shows some rectification examples.

C. Data-driven Quality Assessment

Given the input image I and the rectified version of the captured projected image \tilde{I}_p^c , our method extracts a range of features that reflect the image quality and then uses a regression method to score the quality of the projected image I_p . We experimented with several off-the-shelf regression methods and find that the random forest regression method works best [23]. Below we detail the image features used in our method.

1) *Deep Learning Features*: Deep learning features have been shown very effective for image quality assessment [18], [19]. They capture the useful image characteristics that reflect the aesthetic quality of an image. Therefore, we include deep learning features into our method. Specifically, we generate the deep learning features using the deep convolutional neural network shared by Mai *et al.* [19]. We use this deep neural network to produce two 4096-dimensional feature vectors for an image, each extracted at a different image scale. At each scale, we concatenate the feature vector for the input image and the rectified captured image and then use PCA to reduce its dimension to 20. We finally concatenate the two resulting feature vectors into a 40-dimensional deep feature vector.

2) *Manually Crafted Photo Quality Features*: A variety of carefully crafted features have been used in image quality and aesthetics assessment. Our method includes some relevant ones and also designs some new features dedicated for projection quality assessment. These features can be grouped into four categories, including global image appearance, local image distortion, saliency-based features and superpixel-based features.

Global Image Appearances. Existing research on image aesthetics showed that global image statistics can serve as good indications of image quality [12], [13]. We, therefore, compute these image statistics as features, including the average hue, saturation, intensity, and brightness value. We compute these values for both the input image and the rectified captured image and concatenate them together. As shown by Marchesotti *et al.*, the GIST feature, which captures the global appearance of an image [24], is also helpful for quality assessment [15]. Thus, our method includes the GIST feature for the two images. Moreover, our method considers several other image characteristics that are particularly affected by projection and models them as features. We briefly describe them below.

Color transfer. Due to the different optics of the projector and the camera and the effect of environment lighting, the projected image as well as its captured and the perceived version undergo different global color transformations. To reflect this, we adopt the color transfer method that transforms the color distribution of one image to match the other and compute the color transfer feature between the input image and the rectified captured image. Specifically, we convert the color space of each image from RGB to Lab

and then compute the mean and standard deviation of each channel [25]. We include these measurements for the two images as features to account for the color transformations of the projector-camera system.

Hue count. Relevant to the color transfer feature, we also consider the impact of the projection on the hue count of the input image as it has been shown to be an important aspect of image quality. Briefly, we first convert an image into the HSV color space, then build a 20-bin hue histogram, and finally compute the hue count based on the histogram. Please refer to Ke *et al.* for more details [12].

Contrast and blurriness. Projection often compromises the contrast of an image, especially with strong ambient lighting. We follow the method from Ke *et al.* to compute the contrast values for the two images as features. Specifically, we first compute the gray scale histogram and then measure the width of the middle 98% mass of the histogram as the contrast [13]. Our method also measures blurriness, a relevant feature to contrast, of the two images as features using the blur detection method from Tong *et al.* [26].

Local Image Distortion. Unlike a dedicated projection surface, an arbitrary surface in a workspace or living room is often textured. The texture can often significantly change the local content of the projected image and compromise the projection quality. To detect the local distortion caused by the textured surface, we detect salient feature points, SIFT in our implementation, in the input image and the rectified captured image. We then establish the feature correspondence between these two images and measure the local image distortion as follows.

$$f_{sift} = \frac{\sum_{k=1}^m e^{-\frac{d_k^2}{2\sigma^2}}}{m_i + m_c} \quad (3)$$

where m_i and m_c are the numbers of SIFT feature points in the input image and the rectified captured image, respectively, and m is the number of the pairs of matched feature points between these two images. d_k is the Euclidean distance between two feature descriptors of each pair of matched features. σ is a parameter with default value 10^3 .

Visual Saliency-based Feature. Visual saliency is designed to capture low-level visual stimulation from the images and has been shown correlated to image quality assessment [10]. Projection should preserve the saliency of the input image as much as possible. For the input image and the rectified captured image, we first detect a saliency mask for each image [27] and then build a color histogram with 20 bins for each channel using the salient pixels. We include the two resulting color histogram as features.

Superpixel-wise Image Similarity. As shown in existing full-reference quality assessment methods, pixel-wise relationship between the input image and the distorted image is indicative for the quality of the distorted image. However, although our rectification algorithm can align the captured



Figure 4. Sample input images.

projected image and the input image well, it sometimes still cannot achieve pixel-wise alignment due to various reasons. For example, the projection surface is not a perfect plane and thus the transformation between two images cannot be perfectly modeled by a homography, which makes it difficult to measure the pixel-wise relationship between these two images. Based on the observation that the misalignment between the input and the rectified captured image is often small, we extract superpixels in the images and measure superpixel-based similarity between these two images. We first extend the SLIC method [28] to extract superpixels from the two images simultaneously. Specifically, we superimpose the input image onto the rectified captured image and create a 6-channel image, with the first three channels coming from the input image and the next three channels from the rectified captured image. We then apply SLIC to this 6-channel image and obtain a set of superpixels. We measure the similarity between the two images by computing the correlation score of the intensity histograms of superpixels in the two images. We finally build a 20-bin histogram of these correlation scores as a feature for our algorithm.

III. EXPERIMENTS

Since there is currently no large-scale benchmark for projected image quality assessment, we developed a benchmark to evaluate our method. We first collected a set of 40 images from SlideShare¹. These images cover a variety of projection content, such as photos, graphics, text, and data visualization, as shown in Figure 4. We also collected 13 different projection surfaces that are common in the office or living room environment, as shown in Figure 5. These surfaces cover a wide range of color, texture, and material properties. We further simulate five different environment lighting conditions, including full ambient light, half ambient

¹<http://www.slideshare.net/>



Figure 5. Sample projection surfaces.

light, no ambient light, and two configurations with directional light pointing toward the projection surface. Overall, each image was projected under 65 different configurations. In total we have 2600 captured projected images.

We conducted a study to obtain the subjective quality score of each projected image. Specifically, we projected each image onto a surface and concurrently showed the same image on a 27-inch Apple monitor. The image shown on the monitor serves as a reference of a high-quality display of the input image. We recruited three participants and asked them to score the quality of the projected image that they saw on the surface by comparing it to the input image displayed on the monitor. *Note, the observers did not see the projected images captured by the camera. They only looked at the projection results.* The score ranges from 1 to 9 (from worst to best). We average the subjective scores for the projected image as its quality score. In our experiments, we uniformly scale the user score to the range of [0, 1]. Figure 6 reports the score distributions of our benchmark. In our work, all the participants need to score images by looking at the projected images on the surfaces during the projecting and capturing process, which makes it difficult for us to recruit a large number of them to stay in the lab at the same time to score for 2600 times. We examined the quality of the labelling data. As shown in Figure 7, the standard deviations of scores for most images are low, which shows that they were rated by three participants consistently.

A. Evaluation

We experimented with our method and compared it to the state of the art full-reference quality metrics on our benchmark. All these methods take a rectified captured image and its corresponding input image as input and output a quality score. The rectification algorithm described in Section II-B was used to align each captured projected image and its corresponding input image. Our test showed that this rectification method failed on 2.2%, namely 57 out of 2600 captured projected images. We removed these 57 captured

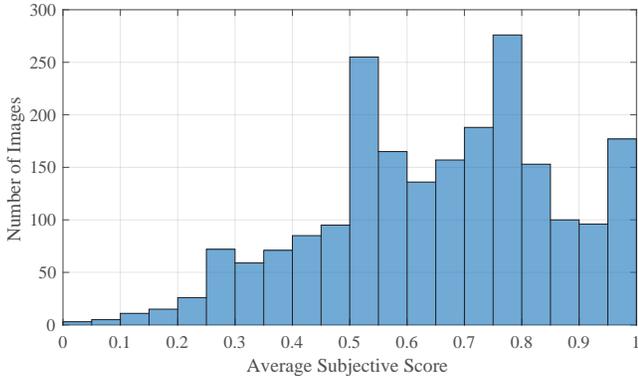


Figure 6. Mean opinion score distribution of the benchmark.

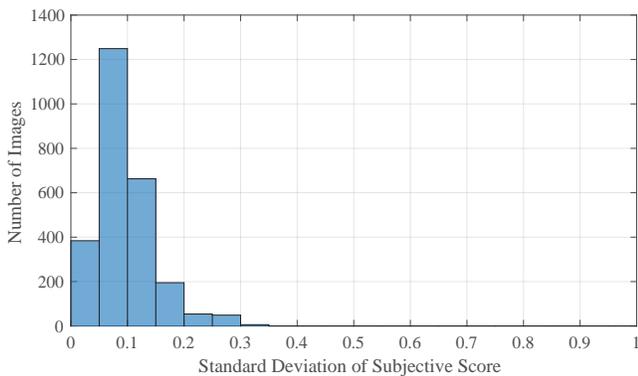


Figure 7. Standard deviation of opinion scores of the benchmark.

projected images during our evaluation.

We randomly partitioned our dataset into ten folds. Each fold contains four groups of captured projected images, in which, each group contains captured images of the same input image. For each fold, we trained our algorithm using the rest nine folds and tested its prediction performance on that fold. As shown in Figure 8, our predicted scores are well-correlated with the ground-truth subjective scores. Figure 9 shows an example where our method accurately scores the quality of projected images on different projection surfaces.

It is interesting to examine how the deep learning features (DL) and manually crafted features (MC) work in our method. As reported in Table I, MC features contribute most to the performance of our method. DL features alone are not sufficient, as these are optimized for assessing the quality of photos instead of projected visual content. We did not fine tune the deep neural network due to the lack of a sufficiently large training set. In future, we will develop a larger benchmark to fine tune the deep neural network. Nevertheless, DL features can still slightly improve our method when combined with MC features, as shown in Table I.

Since this is the first work on predicting the quality

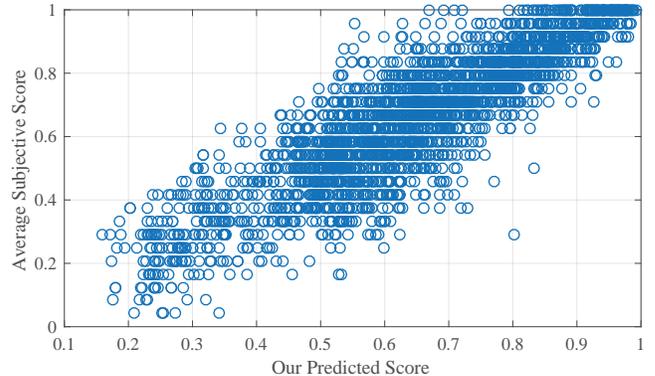


Figure 8. Our predicted scores.

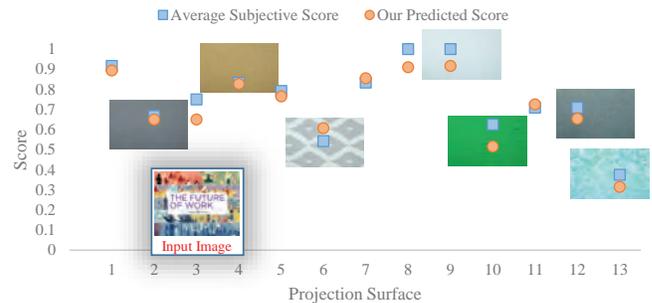


Figure 9. Our results for an image on various surfaces.

of projected images, we select several state-of-the-art full-reference image quality metrics as references to examine the performance of our method. These methods include SSIM [8], GSM [29], FSIM [30], VSI [10], MS-SSIM [9], IW-SSIM [31], VIF [11], and SRSIM [32]. We adopted four popular performance metrics, including Spearman rank-order correlation coefficient (SROC), Kendall rank-order correlation coefficient (KROC), Pearson linear correlation coefficient (PLCC), and root mean squared error (RMSE). SROC and KROC measure the correlation between the predicted ranking order and the ground truth quality ranking order for all the captured images of the same input image. For PLCC and RMSE, previous work recommended that we first apply a non-linear regression function to map the objective scores predicted by the reference methods in our experiments to the subjective scores before computing the PLCC and RMSE values. We followed [10] and used the non-linear mapping function from [11] in our experiments. Figure 10 shows the scatter plots of the ground truth subjective quality scores and objective scores predicted by the reference methods, as well as the corresponding non-linear mapping functions. These mapping functions are used to map the predicted objective quality scores to the subjective scores, which are used to compute the PLCC and RMSE values reported in Table I. Since our method directly predicts the quality scores of the projected images, no mapping is needed. As reported in Table I, our method consistently

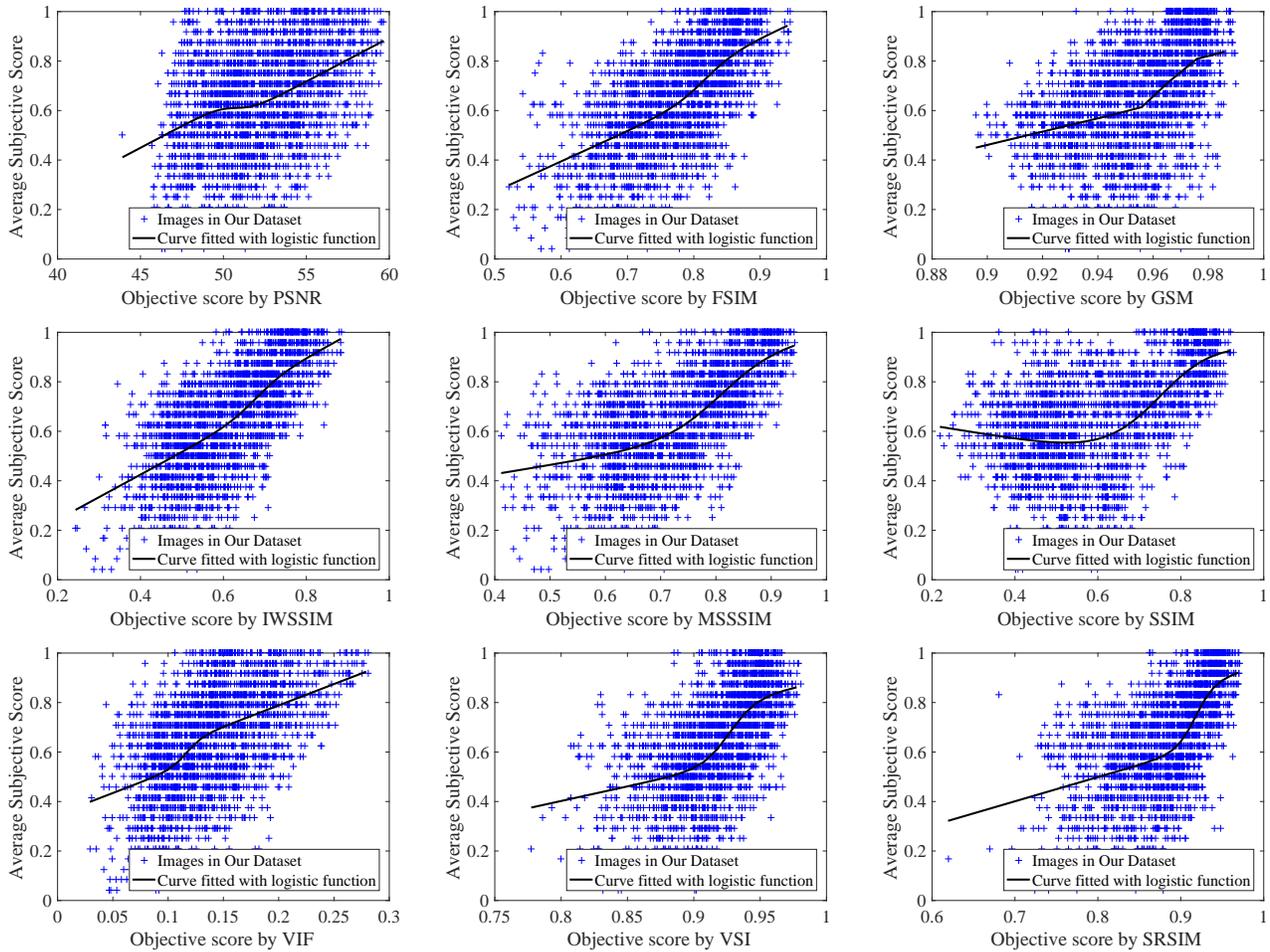


Figure 10. Scatter plots of the predicted objective scores using existing full-reference quality metrics.

Table I
COMPARISON AMONG DIFFERENT QUALITY METRICS.

Method	SROC	KROC	PLCC	RMSE
PSNR	0.6141	0.4472	0.5938	0.1896
SSIM	0.5349	0.3816	0.6159	0.1666
GSM	0.6134	0.4410	0.6192	0.1762
FSIM	0.7532	0.5851	0.7538	0.1535
VSI	0.7395	0.5529	0.7198	0.1602
MS-SSIM	0.7360	0.5680	0.7543	0.1494
IW-SSIM	0.7468	0.5766	0.7572	0.1477
VIF	0.7128	0.5425	0.6929	0.1728
SRSIM	0.7226	0.5607	0.7329	0.1551
Ours MC	0.8871	0.7492	0.8979	0.1008
Ours DL	0.5494	0.4160	0.6126	0.1864
Ours	0.8935	0.7570	0.9010	0.0994

outperforms these reference methods.

IV. CONCLUSION

This paper presented a quality metric to evaluate the visual quality of projection content. A major challenge in evaluating the projection quality is that the camera and viewer “see” the same projected image differently, making it difficult to apply existing reference-based quality metrics. This paper addressed this challenging problem using a data-

driven approach that learns to bridge this gap and score the quality of the projected image. Our experiment showed that our method can assess the projection quality in a way that is consistent with the human perception.

ACKNOWLEDGMENTS

The images used in Figures 2, 3, 4, and 9 are used under a Creative Commons license from SlideShare users Pip Cleaves, PSFK, Dony Peter, Kleiner Perkins Caufield & Byers, Fabio Lalli, and R/GA. This research is supported by a gift from Intel Corporation.

REFERENCES

- [1] A. Majumder and M. S. Brown, *Practical multi-projector display design*. AK Peters USA, 2007.
- [2] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs, “The office of the future: a unified approach to image-based modeling and spatially immersive displays,” in *ACM SIGGRAPH*, 1998, pp. 179–188.
- [3] M. Brown, A. Majumder, and R. Yang, “Camera-based calibration techniques for seamless multiprojector displays,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 2, pp. 193–206, 2005.

- [4] B. R. Jones, H. Benko, E. Ofek, and A. D. Wilson, "Illumiroom: Peripheral projected illusions for interactive experiences," in *ACM CHI*, 2013, pp. 869–878.
- [5] O. Bimber, D. Iwai, G. Wetzstein, and A. Grundhfer, "The visual computing of projector-camera systems," *Computer Graphics Forum*, vol. 27, no. 8, pp. 2219–2245, 2008.
- [6] P. Zhao, M. Pedersen, J.-B. Thomas, and J. Y. Hardeberg, "Perceptual Spatial Uniformity Assessment of Projection Displays with a Calibrated Camera," *The 22nd Color and Imaging Conference*, pp. 159–164, 2014.
- [7] N. Damera-Venkata, T. Kite, W. Geisler, B. Evans, and A. Bovik, "Image quality assessment based on a degradation model," *IEEE Trans. on Image Processing*, vol. 9, no. 4, pp. 636–650, 2000.
- [8] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [9] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *37th Asilomar Conf. on Signals, Systems and Computers*, vol. 2. IEEE, 2003, pp. 1398–1402.
- [10] L. Zhang, Y. Shen, and H. Li, "Vsi: A visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. on Image Processing*, vol. 23, no. 10, pp. 4270–4281, 2014.
- [11] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. on Image Processing*, vol. 15, no. 2, pp. 430–444, 2006.
- [12] Y. Ke, X. Tang, and F. Jing, "The design of high-level features for photo quality assessment," in *IEEE CVPR*, 2006, pp. 419–426.
- [13] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Studying aesthetics in photographic images using a computational approach," in *ECCV*, 2006, pp. 288–301.
- [14] Y. Luo and X. Tang, "Photo and video quality evaluation: Focusing on the subject," in *ECCV*, 2008, pp. 386–399.
- [15] L. Marchesotti, F. Perronnin, D. Larlus, and G. Csurka, "Assessing the aesthetic quality of photographs using generic image descriptors," in *ICCV*, 2011, pp. 1784–1791.
- [16] H.-H. Su, T.-W. Chen, C.-C. Kao, W. H. Hsu, and S.-Y. Chien, "Scenic photo quality assessment with bag of aesthetics-preserving features," in *ACM Int. Conf. on Multimedia*, 2011, pp. 1213–1216.
- [17] X. Tang, W. Luo, and X. Wang, "Content-based photo quality assessment," *IEEE Trans. on Multimedia*, vol. 15, no. 8, pp. 1930–1943, Dec 2013.
- [18] X. Lu, Z. Lin, H. Jin, J. Yang, and J. Z. Wang, "Rating image aesthetics using deep learning," *IEEE Trans. on Multimedia*, vol. 17, no. 11, pp. 2021–2034, 2015.
- [19] L. Mai, H. Jin, and F. Liu, "Composition-preserving deep photo aesthetics assessment," in *IEEE CVPR*, 2016, pp. 497–506.
- [20] D. Moreno and G. Taubin, "Simple, accurate, and robust projector-camera calibration," in *Int. Conf. on 3D Imaging, Modeling, Processing, Visualization & Transmission*. IEEE, 2012, pp. 464–471.
- [21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [22] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [23] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
- [24] A. Oliva and A. Torralba, "Modeling the shape of the scene: A holistic representation of the spatial envelope," *International journal of computer vision*, vol. 42, no. 3, pp. 145–175, 2001.
- [25] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Comput. Graph. Appl.*, vol. 21, no. 5, pp. 34–41, Sep. 2001. [Online]. Available: <http://dx.doi.org/10.1109/38.946629>
- [26] H. Tong, M. Li, H. Zhang, and C. Zhang, "Blur detection for digital images using wavelet transform," in *IEEE ICME*, vol. 1, 2004, pp. 17–20.
- [27] W.-C. Tu, S. He, Q. Yang, and S.-Y. Chien, "Real-time salient object detection with a minimum spanning tree," in *IEEE CVPR*, June 2016.
- [28] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [29] A. Liu, W. Lin, and M. Narwaria, "Image quality assessment based on gradient similarity," *IEEE Trans. on Image Processing*, vol. 21, no. 4, pp. 1500–1512, 2012.
- [30] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: a feature similarity index for image quality assessment," *IEEE Trans. on Image Processing*, vol. 20, no. 8, pp. 2378–2386, 2011.
- [31] Z. Wang and Q. Li, "Information content weighting for perceptual image quality assessment," *IEEE Trans. on Image Processing*, vol. 20, no. 5, pp. 1185–1198, 2011.
- [32] L. Zhang and H. Li, "Sr-sim: A fast and high performance iqa index based on spectral residual," in *IEEE ICIP*, 2012, pp. 1473–1476.