DEEP COLOR MISMATCH CORRECTION IN STEREOSCOPIC 3D IMAGES

Simone Croci[†], *Cagri Ozcinar*[‡], *Emin Zerman*[†], *Roman Dudek*[§], *Sebastian Knorr*[⋆], *Aljosa Smolic*[†]

[†] V-SENSE Project, School of Computer Science and Statistics, Trinity College Dublin
[‡] Samsung Research Institute UK
[§] Universidad de Las Palmas de Gran Canaria
[⋆] Ernst Abbe University of Applied Sciences Jena

ABSTRACT

Color mismatch in stereoscopic 3D (S3D) images can create visual discomfort and affect the performance of S3D image processing algorithms, *e.g.*, for depth estimation. In this paper, we propose a new deep learning-based solution for the problem of color mismatch correction. The proposed solution consists of a multi-task convolutional neural network, where color correction is the primary task and correspondence estimation is the secondary task. For the training and evaluation of the proposed network, a new S3D image dataset with color mismatch was created. Based on this dataset, experiments were conducted showing the effectiveness of our solution.

***Index Terms*—** Color mismatch, color correction, stereoscopic 3D, convolutional neural network

1. INTRODUCTION

Stereoscopic 3D images can contain color mismatch between the left and right image due to reasons like different camera and lens characteristics, different illumination and reflections resulting from different camera orientations, polarized light, etc. The presence of color mismatch can reduce the quality of experience and cause problems when processing the S3D images [1]. For color mismatch correction, there are already several traditional methods [2–9], but not so many based on deep learning, which has already been successfully applied in other computer vision problems. Therefore, here we propose an effective solution based on a convolutional network.

Color correction of S3D images consists of selecting either the left or right image as the reference image and correcting the other image, which is called the target image. The reference image contains the color information that must be mapped to the target image while preserving the structure information, *i.e.*, the edges, present in the target image.

For color correction, there are two categories of approaches, namely, global [2, 3] and local [4] methods. Global methods estimate a single color transformation, while local methods estimate different local transformations. Global methods usually analyze the color distributions of the reference and target images, but they fail in the presence of local color mismatch. Local methods can fix local color mismatch, but they are usually sensitive to the quality of correspondence estimation between reference and target images.

In this paper, we propose a multi-task deep learning-based solution that is essentially a local approach capable of fixing local color mismatch and that is also trained to estimate correspondences in the presence of color mismatch. In this way, this multi-task solution tries to improve correspondence estimation and reduce the negative effect of inaccuracies in this secondary task.

For the correspondence estimation, we use the so-called parallax-attention mechanism (PAM) [10]. PAM is a non-local network that estimates correspondences along horizontal epipolar lines, assuming that the input images are rectified. In our case, PAM is trained in an unsupervised manner due to the lack of ground truth correspondences.

The main contribution of this paper is a convolutional neural network (CNN) for color mismatch correction. We also created a dataset with undistorted and distorted S3D images necessary for the training of the network. Moreover, using this dataset, we conducted experiments for the optimization of the network and its evaluation. The code and dataset are publicly available[1].

2. RELATED WORK

In computer vision and multi-view video processing communities, the color mismatch problem is generally solved using either global [2, 3] or local [4] approaches. The global approaches include exposure compensation (or gain compensation) [2] or use 3D lookup tables [3]. However, they may fail in the case of local differences. Local approaches do not explicitly compensate for the image as a whole, and they can focus on correcting the image region by region using local feature correspondences [4].

New studies attempt to combine global and local approaches [5, 6]. They generally first focus on a global color correction to rectify the color values and bring them as close to the reference as possible. Afterward, the local color correction is handled in a region-based approach using optical flow estimation. More recently, a deep learning-based method [11] was proposed, which is based on a relatively simple CNN and uses per-pixel and perceptual losses.

The image processing and computer graphics communities have developed similar color manipulation methods, called color transfer techniques. These methods transfer the color feel from a palette image to a target image, and they assume that the contents of the two images are different. The earliest work in this area was by Reinhard *et al.* [12], who proposed transforming the mean and standard deviation of each color channel in the target image to match that of the palette image. Since then, more complex techniques have been used to model the color distributions of the images more accurately, including histograms and Gaussian mixture models [13, 14]. While global color transfer functions are often used, including affine, radial basis and optimal transport functions [15, 16], local techniques

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under the Grant Number 15/RP/2776.

[1] https://v-sense.scss.tcd.ie/research/color-mismatch-correction/

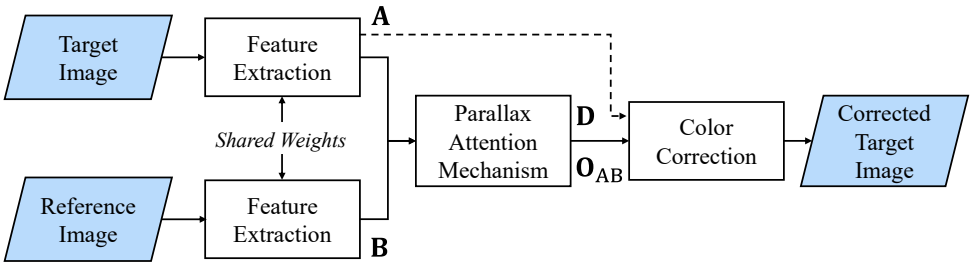


Fig. 1: Illustration of the proposed solution for color mismatch correction.

have also been proposed to allow for more flexibility in the recoloring [17, 18]. An efficient method was developed by Pitie *et al.* [7]. It first estimates a global color function that converts the color distribution of one image into another, and then it reduces possible grain artifacts generated by the color function. Recently, Grogan *et al.* [9] proposed a color transfer technique that could also be enhanced to take into account color correspondences between the target and palette images, ensuring the method could be used to color correct images of the same scene.

In our multi-task solution, we use the parallax-attention mechanism (PAM) for the correspondence estimation. This component has already been successfully used in other methods for disparity estimation [10], S3D image super-resolution [10, 19], binocular image dehazing [20], light field reconstruction [21], and object pose estimation [22]. Differently from the other methods, we apply PAM to S3D images with color mismatch and we show that it works also in this condition.

3. PROPOSED METHOD

The proposed color correction method takes as input an S3D image consisting of a reference image $\mathbf{I}_{ref}$ and a target image $\mathbf{I}_{targ}$ that are rectified. The target image is the image that needs to be color corrected based on the colors of the reference image. The proposed color mismatch correction method is based on a CNN that has three main components as illustrated in Figure 1: feature extraction, PAM, and color correction.

3.1. Feature Extraction

The feature extraction component extracts the features **A** and **B** necessary for the color correction from the target image $\mathbf{I}_{targ}$ and the reference image $\mathbf{I}_{ref}$, respectively. This component consists of a 3×3 convolution layer followed by a sequence of eighteen residual blocks (see Figure 3). The number of channels of the feature maps is kept constant through the entire CNN equal to 64.

3.2. Parallax-Attention Mechanism

The second component is the parallax-attention mechanism (PAM) [10] that estimates correspondences along horizontal epipolar lines. It also computes an occlusion map, and it warps the features of the reference image into the target image.

PAM is illustrated in Figure 2. The inputs of PAM are the feature maps $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{H \times W \times C}$ (H is height, W is width, and C are the channels) from the previous component, corresponding to the features extracted from the target image $\mathbf{I}_{targ}$ and the reference image $\mathbf{I}_{ref}$, respectively. In the beginning, there are two residual blocks with shared weights that adapt the input features for the estimation of the correspondences and that generate the feature maps $\mathbf{A}_0$ and

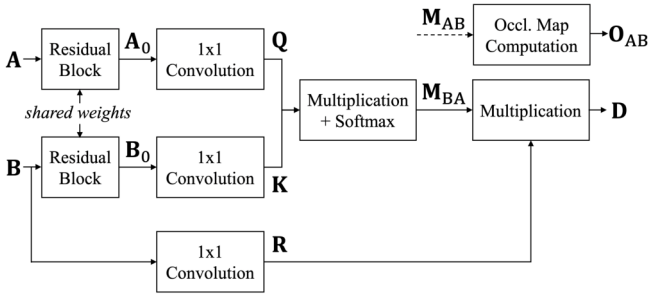


Fig. 2: Parallax-attention mechanism.

$\mathbf{B}_0$. This is important since different tasks require different features, otherwise, the proposed multi-task solution would suffer from training conflicts [23]. Then, a 1×1 convolution layer converts $\mathbf{A}_0$ into a feature map $\mathbf{Q} \in \mathbb{R}^{H \times W \times C}$, and another 1×1 convolution layer converts $\mathbf{B}_0$ into a feature map $\mathbf{K} \in \mathbb{R}^{H \times W \times C}$ that is reshaped to $\mathbb{R}^{H \times C \times W}$. **Q** and **K** are multiplied and softmax is applied obtaining a parallax attention map $\mathbf{M}_{\mathbf{B}\to\mathbf{A}} \in \mathbb{R}^{H \times W \times W}$. $\mathbf{M}_{\mathbf{B}\to\mathbf{A}}$ can be seen as a cost matrix that encodes the correspondences along horizontal epipolar lines. In the next step, **B** is processed by a 1×1 convolution layer obtaining $\mathbf{R} \in \mathbb{R}^{H \times W \times C}$, which is multiplied by $\mathbf{M}_{\mathbf{B}\to\mathbf{A}}$ to generate $\mathbf{D} \in \mathbb{R}^{H \times W \times C}$. **D** can be interpreted as the warping of **B** into **A**. PAM also estimates the occlusion map $\mathbf{O}_{\mathbf{A}\to\mathbf{B}}$. For the occlusion map, a second parallax attention map $\mathbf{M}_{\mathbf{A}\to\mathbf{B}}$ is estimated by exchanging **A** and **B**. Refer to [10] for the details of the occlusion map computation.

3.3. Color Correction

The color correction component takes as input the features **A** of the target image, the warped features **D** of the reference image together with the occlusion map $\mathbf{O}_{\mathbf{A}\to\mathbf{B}}$, and it computes the color corrected target image. First, **A**, **D**, and $\mathbf{O}_{\mathbf{A}\to\mathbf{B}}$ are concatenated and the resulting features are fused by a 1×1 convolution layer. The fused features are then processed by a sequence of six residual blocks (see Figure 3) followed by two 3×3 convolution layers.

3.4. Losses

The multi-task CNN is trained for color correction in a supervised manner, and for correspondence estimation in an unsupervised way. For these tasks, we proposed to use two losses, namely, the color correction loss $\mathcal{L}_{CC}$ and the PAM correspondence estimation loss $\mathcal{L}_{PAM}$. The overall loss $\mathcal{L}$ consists of the following weighted sum:

$$\mathcal{L} = \mathcal{L}_{CC} + 0.005\, \mathcal{L}_{PAM}. \quad (1)$$

The color correction loss $\mathcal{L}_{CC}$ evaluates how different the color corrected target image $\mathbf{I}'_{targ}$ is from the ground truth target image $\mathbf{I}^{\text{GT}}_{targ}$. For this loss, we decided to use a combination of pixel-based and perceptual losses. More precisely, we propose the sum of the mean absolute error (MAE), the mean squared error (MSE), and the negative of the structural similarity index measure (SSIM) [24]:

$$\mathcal{L}_{CC} = \frac{1}{N}\sum_{p} \|\mathbf{I}'_{targ}(p) - \mathbf{I}^{\text{GT}}_{targ}(p)\|_1 + \frac{1}{N}\sum_{p} \|\mathbf{I}'_{targ}(p) - \mathbf{I}^{\text{GT}}_{targ}(p)\|_2^2 - \text{SSIM}(\mathbf{I}'_{targ}, \mathbf{I}^{\text{GT}}_{targ}), \quad (2)$$

where p are the pixels, and N is their number.

As proposed in [10], for the correspondence estimation loss $\mathcal{L}_{PAM}$, we use the sum of three losses: the photometric loss $\mathcal{L}_{pm}$, the smoothness loss $\mathcal{L}_{smooth}$, and the cycle loss $\mathcal{L}_{cycle}$:

$$\mathcal{L}_{PAM} = \mathcal{L}_{pm} + \mathcal{L}_{smooth} + \mathcal{L}_{cycle}. \quad (3)$$

No ground-truth correspondences are used in these losses since PAM is trained in an unsupervised manner. The photometric loss $\mathcal{L}_{pm}$ computes the difference between the warped reference and target image and between the warped target and reference image. The smoothness loss $\mathcal{L}_{smooth}$ is applied for the correct handling of textureless regions when computing the correspondences. The cycle loss $\mathcal{L}_{cycle}$ is introduced to achieve cycle consistency.

4. EXPERIMENTS

4.1. Dataset

For the training and evaluation of the proposed CNN, we took S3D images with very low color mismatch and we introduced color distortions. The S3D images were taken from three datasets: Flickr1024 [25], InStereo2K [26], and the IVY LAB Stereoscopic 3D image database [27].

In order to exclude images with repetitive content and with large color mismatch, the S3D images were manually checked for repetition and analyzed automatically for color mismatch with the method proposed by Croci *et al.* [28]. In the end, we obtained 1035 undistorted S3D images.

In order to introduce color mismatch in the undistorted S3D images, we used the same approach of [11, 29]. More precisely, we modified the target images by applying the following six color modification operators found in Photoshop 2021 with different intensity levels: brightness (-90, -60, -30, 30, 60, 90), color balance (-90, -60, -30, 30, 60, 90), contrast (-60, -40, -20, 20, 40, 60), exposure (-3, -2, -1, 1, 2, 3), hue (-60, -40, -20, 20, 40, 60), saturation (-40, -20, 20, 40, 60). In the end, we obtained 36225 distorted S3D images. The final dataset consists of a total of 37260 undistorted and distorted S3D images. Both types of S3D images were used for the training and evaluation. 80% of the dataset was randomly selected for the training set, 10% for the validation set, and another 10% for the test set.

4.2. Training Procedure

During the training, random patches were extracted from the input images. As data augmentation strategy, we applied random vertical and horizontal flipping. Furthermore, we used the Adam optimizer with learning rate equal to 0.0001 and batch size equal to eight. The neural network was implemented with Pytorch.

4.3. Color Correction Quality Metrics

For the evaluation of the results, two different full-reference quality metrics were applied between the corrected and ground truth target images. The first quality metric $\Delta\hat{E}^*_{ab}$ [30] is the mean of the color differences between corresponding pixels defined as follows:

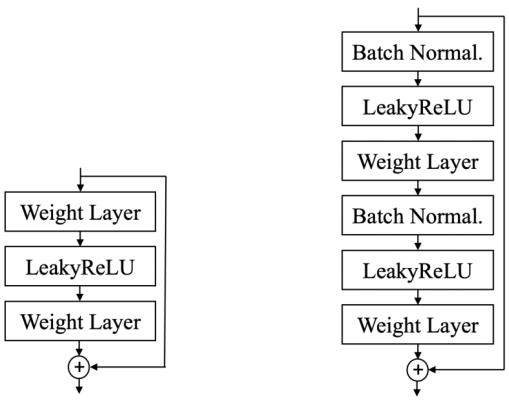
$$\Delta E^*_{ab} = \sqrt{(L^*_2 - L^*_1)^2 + (a^*_2 - a^*_1)^2 + (b^*_2 - b^*_1)^2}, \quad (4)$$

where (L^*_1, a^*_1, b^*_1) and (L^*_2, a^*_2, b^*_2) are pixel colors defined in the CIELAB color space. $\Delta\hat{E}^*_{ab}$ was chosen because it is based on a perceptually uniform color space. The second quality metric is the structural similarity index measure (SSIM) [24]. $\Delta\hat{E}^*_{ab}$ is an indicator for the correctness of the color information, while SSIM is an indicator for the structure information.

4.4. Hyperparameter Optimization and Ablation Study

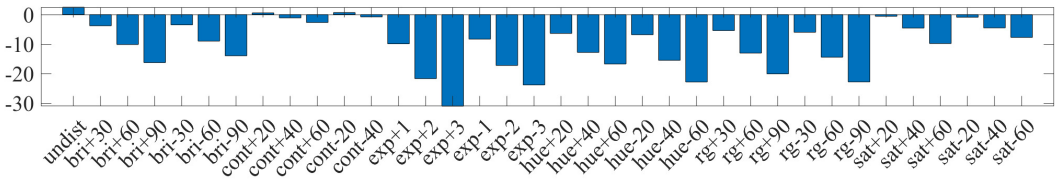
For the hyperparameter optimization, we selected a baseline model (BM) that is an instance of the proposed solution, and we modified its components in order to find the optimal solution. BM corresponds to our model with six residual blocks in the feature extraction component, and two residual blocks in the color correction component. The residual block type used in BM is ResBOne shown in Figure 3a. BM and its variations were trained for 50 epochs. First, we tried a different residual block type with batch normalization called ResBTwo shown in Figure 3b. Then, we tried different numbers of residual blocks in the feature extraction and correction components. The best overall loss $\mathcal{L}$ and color correction loss $\mathcal{L}_{CC}$ computed on the validation set are shown in Table 1. According to the table, we selected the optimal model with the residual block type ResBOne, eighteen residual blocks in the feature extraction component, and six residual blocks in the color correction component. We then trained this model for 100 epochs.

In the ablation study, we analyzed the contribution of PAM. For this study, we replaced PAM with a component that warps the reference image features based on correspondences obtained by SIFT-Flow [31], which is an optical flow estimation method robust to color mismatch. We trained this CNN without PAM for 50 epochs and we obtained the best $\mathcal{L}_{CC}$ computed on the validation set equal to -0.9109. This is larger than the value obtained with the CNN with PAM, and it shows the important contribution of PAM in our solution.

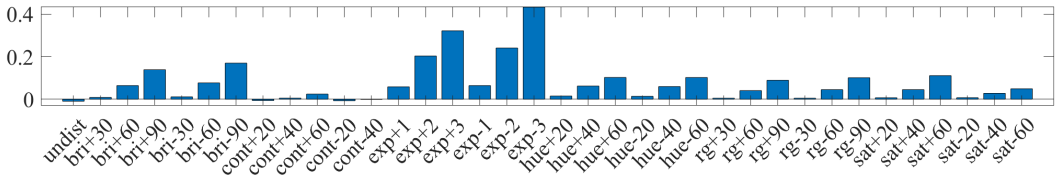


(a) ResBOne (b) ResBTwo

Fig. 3: Residual blocks.



(a) $\Delta\hat{E}^*_{ab}$ difference: $\Delta\hat{E}^*_{ab,\text{corrected}} - \Delta\hat{E}^*_{ab,\text{distorted}}$ (Lower values are better)



(b) SSIM difference: $\text{SSIM}_{\text{corrected}} - \text{SSIM}_{\text{distorted}}$ (Higher values are better)

Fig. 4: Difference of the metric scores of the corrected and distorted target images for each distortion. The labels of the horizontal axis refer to the different distortions with their intensity levels. undist corresponds to the undistorted images.

Table 1: Hyperparameter optimization.

Feat. Extr.	Col. Corr.	$\mathcal{L}$	$\mathcal{L}_{CC}$
6 ResBOne	2 ResBOne	-0.9577	-0.9584
6 ResBTwo	2 ResBTwo	-0.9460	-0.9468
12 ResBOne	2 ResBOne	-0.9594	-0.9601
6 ResBOne	4 ResBOne	-0.9626	-0.9633
12 ResBOne	4 ResBOne	-0.9645	-0.9652
18 ResBOne	4 ResBOne	-0.9647	-0.9654
12 ResBOne	6 ResBOne	-0.9659	-0.9666
18 ResBOne	6 ResBOne	-0.9665	-0.9672

4.5. Distortion-based Evaluation

The proposed method was applied to each distortion separately, and for each of them, Figure 4 shows the difference between the metric scores of the corrected and distorted target images. Our solution is able to improve the S3D images for almost all distortions, except for a few of them where small changes are introduced. This can be explained by the inaccuracy of the correspondence estimation.

4.6. Comparison Study

The proposed method was compared with Dudek *et al.*'s local approach [5], and the global approaches of Grogan *et al.* [9], Pitie *et al.* [7], and Reinhard *et al.* [12]. We did not compare against the most recent deep learning-based method [11] as neither code nor data are available. As can be seen from Table 2 showing the method scores, our method has a better performance than the other four methods. Figure 5 shows a target image distorted with the hue operator with intensity $+60$, and its corrections with our and Pitie's method. Also in this figure, it is possible to see how our method generates better results.

Table 2: Comparison of color correction methods.

Method	$\Delta\hat{E}^*_{ab}$	SSIM
Dudek *et al.* [5]	6.4982	0.9366
Grogan *et al.* [9]	5.6269	0.9476
Pitie *et al.* [7]	5.8105	0.9350
Reinhard *et al.* [12]	13.4496	0.8247
Ours	3.3084	0.9791

(a) Distorted target image.

(b) Undistorted target image.

(c) Pitie *et al.* [7].

(d) Ours.

Fig. 5: Visual results.

5. CONCLUSION

We proposed a new deep learning-based color correction solution based on multi-task learning, where the primary and secondary tasks are color correction and correspondence estimation, respectively. We created a new dataset for the training and evaluation, and we conducted studies showing the effectiveness of the proposed solution.

In the future, we plan to improve the dataset in order to have undistorted S3D images with almost absent color mismatch by using photorealistic computer-generated images or 2D-to-3D converted images. Another goal is to investigate different feature extraction and color correction components in order to improve the color correction performance. We are also considering improving the correspondence estimation by training PAM with supervised learning based on ground-truth correspondences.

6. REFERENCES

[1] Sebastian Knorr, Kai Ide, Matthias Kunter, and Thomas Sikora, "The avoidance of visual discomfort and basic rules for producing "good 3D" pictures," *SMPTE Motion Imaging Journal*, pp. 72–79, oct 2012.

[2] Wei Xu and Jane Mulligan, "Performance evaluation of color correction approaches for automatic multi-view image and video stitching," in *IEEE Conf. on Comput. Vision and Pattern Recognition (CVPR)*, 2010, pp. 263–270.

[3] Clyde Mouffranc and Vincent Nozick, "Colorimetric correction for stereoscopic camera arrays," in *Comput. Vision - ACCV 2012 Workshops*, 2013, pp. 206–217.

[4] Qi Wang, Pingkun Yan, Yuan Yuan, and Xuelong Li, "Robust color correction in stereo vision," in *Int. Conf. Image Processing (ICIP)*, 2011, number 2, pp. 965–968.

[5] Roman Dudek, Simone Croci, Aljosa Smolic, and Sebastian Knorr, "Robust global and local color matching in stereoscopic omnidirectional content," *Signal Processing: Image Communication*, vol. 74, 2019.

[6] X. Zheng, Y. Niu, J. Chen, and Y. Chen, "Color correction for stereoscopic image based on matching and optimization," in *2017 International Conference on 3D Immersion (IC3D)*, 2017, pp. 1–8.

[7] F. Pitié and Rozenn Dahyot, "Automated colour grading using colour distribution transfer," *Computer Vision and Image Understanding*, vol. 107, 03 2007.

[8] Mairéad Grogan and Rozenn Dahyot, "Robust registration of Gaussian mixtures for colour transfer," *CoRR*, vol. abs/1705.06091, 2017.

[9] Mairéad Grogan and Rozenn Dahyot, "L2 divergence for robust colour transfer," *Computer Vision and Image Understanding*, vol. 181, pp. 39 – 49, 2019.

[10] Longguang Wang, Yulan Guo, Yingqian Wang, Zhengfa Liang, Zaiping Lin, Jungang Yang, and Wei An, "Parallax attention for unsupervised stereo correspondence learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2020.

[11] Yuanyuan Fan, Pengyu Liu, and Yuzhen Niu, "Deep residual optimization for stereoscopic image color correction," in *Parallel Architectures, Algorithms and Programming*, Hong Shen and Yingpeng Sang, Eds., Singapore, 2020, pp. 147–158, Springer Singapore.

[12] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley, "Color transfer between images," *IEEE Comput. Graph. Appl.*, vol. 21, no. 5, pp. 34–41, Sept. 2001.

[13] F. Pitie, A.C. Kokaram, and R. Dahyot, "N-dimensional probability density function transfer and its application to color transfer," in *IEEE Int. Conf. Comput. Vision (ICCV)*, Oct 2005, vol. 2, pp. 1434–1439 Vol. 2.

[14] Yu-Wing Tai, Jiaya Jia, and Chi-Keung Tang, "Local color transfer via probabilistic segmentation by expectation-maximization," in *IEEE Conf. on Comput. Vision and Pattern Recognition (CVPR)*, June 2005, vol. 1, pp. 747–754 vol. 1.

[15] F. Pitié and A. Kokaram, "The linear monge-kantorovitch linear colour mapping for example-based colour transfer," in *Visual Media Prod., 2007. IETCVMP. 4th European Conf. on*, Nov 2007, pp. 1–9.

[16] Mairéad Grogan, Rozenn Dahyot, and Aljosa Smolic, "User interaction for image recolouring using L2," in *Proc. of the 14th European Conf. on Visual Media Prod.*, New York, NY, USA, 2017, CVMP 2017, pp. 6:1–6:10, ACM.

[17] Baoyuan Wang, Yizhou Yu, Tien-Tsin Wong, Chun Chen, and Ying-Qing Xu, "Data-driven image color theme enhancement," *ACM Trans. Graph.*, vol. 29, no. 6, Dec. 2010.

[18] Yichang Shih, Sylvain Paris, Frédo Durand, and William T. Freeman, "Data-driven hallucination of different times of day from a single outdoor photo," *ACM Trans. Graph.*, vol. 32, no. 6, pp. 200:1–200:11, Nov. 2013.

[19] Longguang Wang, Yingqian Wang, Zhengfa Liang, Zaiping Lin, Jungang Yang, Wei An, and Yulan Guo, "Learning parallax attention for stereo image super-resolution," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[20] Y. Pang, J. Nie, J. Xie, J. Han, and X. Li, "BidNet: Binocular image dehazing without explicit disparity estimation," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 5930–5939.

[21] Gaochang Wu, Yebin Liu, Lu Fang, and Tianyou Chai, "Spatial-angular attention network for light field reconstruction," 2020.

[22] Yoshihiro Nakano, "Stereo vision based single-shot 6D object pose estimation for bin-picking by a robot manipulator," 2020.

[23] Ozan Sener and Vladlen Koltun, "Multi-task learning as multi-objective optimization," in *Advances in Neural Information Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., pp. 525–536. Curran Associates, Inc., 2018.

[24] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Trans. Img. Proc.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[25] Yingqian Wang, Longguang Wang, Jungang Yang, Wei An, and Yulan Guo, "Flickr1024: A large-scale dataset for stereo image super-resolution," in *International Conference on Computer Vision Workshops*, Oct 2019, pp. 3852–3857.

[26] Wei Bao, Wei Wang, Yuhua Xu, Yulan Guo, Siyu Hong, and Xiaohu Zhang, "InStereo2K: A large real dataset for stereo matching in indoor scenes," *Science China Information Sciences*, vol. 63, 11 2020.

[27] Y. J. Jung, H. Sohn, S. Lee, , and Y. M. Ro, "Ivy lab stereoscopic image database," [Online]: `http://ivylabprev.kaist.ac.kr/demo/3DVCA/3DVCA.htm`, 2012.

[28] Simone Croci, Sebastian Knorr, Lutz Goldmann, and Aljosa Smolic, "A framework for quality control in cinematic VR based on Voronoi patches and saliency," *2017 International Conference on 3D Immersion (IC3D)*, pp. 1–8, 2017.

[29] Y. Niu, H. Zhang, W. Guo, and R. Ji, "Image quality assessment for color correction based on color contrast similarity and color value difference," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 28, no. 4, pp. 849–862, 2018.

[30] Richard S. Hunter, "Photoelectric color difference meter∗," *J. Opt. Soc. Am.*, vol. 48, no. 12, pp. 985–995, Dec 1958.

[31] C. Liu, J. Yuen, and A. Torralba, "SIFT Flow: Dense correspondence across scenes and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 978–994, 2011.