LightNet: Deep Learning based Illumination Estimation from Virtual Images

Sabari Nathan[1],Parisa Beham M[2]

[1]Couger Inc.,Tokyo,Japan
[2]Sethu Institute of Technology, India

Abstract. In the era of virtual reality (VR), estimating illumination with lighting direction and lighting virtual objects has been a challenging problem. In VR, poor estimation of illumination and lighting direction makes any virtual objects into unrealistic. The inaccurate estimation of lighting can also cause strong artifacts in relighting of the virtual images. Inspired by these issues, the main objective of this paper is to enrich visual rationality of single image by providing accurate assessments of real illumination and lighting direction. We proposed a LightNet architecture by modelling Denseset121 network to estimate the light direction and color temperature level in any virtual reality images. We present quantitative results on VIDIT dataset to evaluate the performance and achieved good results in all the performance metrics. The experimental results proved that the proposed model is robust and provides a good level of accuracy in estimating illumination and lighting direction.

Keywords: Augmented reality, Illumination estimation, lighting direction, virtual images, Densenet, deep learning

1 Introduction

Fig. 1. Challenges in estimating the illumination level and direction in virtual reality images

Augmented reality incorporates digital content and images onto the real world. Interpreting virtual objects into real scenes has been widely used in smart city development and planning, art design, animation and film production [16]. The appearance of things in a scene depends on their illumination and lighting direction. This illumination and their direction is not often taken into account in augmented reality, which makes any virtual object look unrealistic. The perfection of virtual objects and their consistency against actual scene are determined by the lighting effects. Predicting light sources offers a way of automatically locating the precise positions of light sources in a photograph. It can be used to render virtual objects and insert artificial or real objects in the image by illuminating them under the same lighting conditions. Lighting virtual objects with proper illumination and correct orientation is a major focus in computer graphics domain. However, the estimation of real illumination from one image of the scene is a challenging problem, especially if the light sources are not directly visible in the image. Fig.1 shows the challenges in estimating the illumination level and light direction in virtual reality images. From the virtual images shown, one can able to estimate the light direction properly but difficult to estimate the lighting temperature.

Previous research showed that if no priors are used in light source estimation from a single image, it is an ill-conditioned problem [22]. In this paper, our approach is based on an assumption that prior data about illumination and its direction can be learned from a large set of virtual images with known light sources. This learned information can be trained in a deep learning network which is used to estimate lighting levels in a virtual reality scene which was not previously trained. The proposed network can succeed sufficient generality to estimate illumination in various scenes. By maintaining the convergence of training with incrementing network layers and to avoid a vanishing gradients problem, our network has been customized using dense blocks with 121 convolutional layers [8]. In a virtual reality scenario, changing camera orientations cause problems for illumination estimation by a convolutional network. This is because of high dimensionality and complexity of input if a network should handle separate camera orientations in real world.

This work presents the abilities of deep learning network for illumination and light direction estimation in virtual reality by integrating the presented method into a real-time AR rendering system. We also evaluate the results of our method and compare them with the results of a state-of-the-art networks for illumination estimation. Our results indicate that a deep neural network can be used to estimate light sources on scenes which have not been previously seen in the training process.

The main contributions of this paper can be summarized as follows:

1. A novel LightNet architecture for the estimation of illumination and lighting direction in virtual images.
2. Evaluation of the proposed architecture on challenging VIDIT dataset.
3. Comprehensive experiments are conducted and shown the excellence of the proposed method to the baseline methods.
4. Achieve good experimental results in terms of accuracy, F1 score and loss.

2 Related Works

Augmented reality (AR) become more popular in the past decade due to benchmark achievements in computer vision and computer graphics. Several methods were proposed and developed to estimate the illumination conditions in virtual objects in an AR system [10]. In the literature, image processing algorithms have been used to detect the illumination level, and light direction in virtual images [23] [4] [10]. Jachnik et al. [10] , developed an algorithm for real-time surface light-field extract from a single hand-held camera for capturing dense illumination information from specular surfaces. In their work, the light-field is divided into diffuse and specular components where the specular component can be used for environment map estimation. Since the intensity of the shadow is measured as a brute-force approach, computational cost of this method is more. Xing et al. [21] proposed an approach to render virtual objects into a sample image of an outdoor scene by simulating the illumination and the shadow casting between virtual objects and actual scenes. Arief et al. [9] proposed a method for real-time illumination direction estimation for mobile virtual reality systems using analysis of shadows created by a reference object. This method could predict the direction of a single light source in a controlled condition with better accuracy. In this method, the major drawback is that the estimation takes around 15 seconds. In spite of the good results, these image processing algorithms are not capable of running in a real-time devices due to their high computational cost. Special hardware approaches were also dealt in the literature to generate a 3D reconstruction of the scene. In this method, illumination can be estimated by knowing the position of the objects and the light sources. Gruber et al. [15] developed a method for real-time illumination estimation and picture realistic rendering in virtual reality. Rohmer et al. [19] proposed a differential illumination method to obtain a constant illumination of the inserted virtual objects on mobile devices. Multiple HDR video cameras have been used in a predetermined scenario. Recent work on lighting level estimation decomposed the RGB-D input into albedo and shading fields in order to elaborate the scene [11].

Boom et al. [1] developed a first hybrid CPU-GPU based method for estimating a light source position in a scene recorded by an RGB-D camera. The image and depth information from the Kinect is used to estimate a light position in a scene that appears realistic enough for augmented reality purposes. Chen et al. [20], proposed an illumination estimation method which estimate coarse scene geometry and intrinsic components including shading image and reflectance image. Then they used sparse radiance map of the scene to illuminate virtual objects by using the estimated sparse radiance map.

Success of deep learning network approach paves the way for better enhancement in the field of virtual reality. In the light source estimation by neural networks, the space of light directions is discretized into the set of N classes and the network classifies an image as one of these classes [2]. Previous research also proved that dominant light direction can be directly regressed from an input image by a neural network [6]. In [14], Kan and Kafumann proposed an approach based on a similar deep learning method aiming at higher complexity of a scene, temporal coherence and direct application of the network to an augmented reality scenario. Elizondo et al. [6] presents a novel neural network-based approach for recovering light source direction in relation to the viewpoint direction of a graphical image in noisy environments. The estimated light source direction can be used for the generation of 3D images from 2D ones. Frahm et al. [7] presents an approach which exploits a two camera system, the TV camera captures the video for the augmentation while the fish-eye camera observes the upper hemisphere to track the light sources. Thus the virtual objects are rendered by direct lighting. Soulier et al. [12] present a low-cost approach to detect the direction of the environment illumination, allowing the illumination of virtual objects according to the real light of the ambient, improving the integration of the scene.

Compared to the literature works, our proposed LightNet (Densenet121) architecture preserves information that is added to the network. Densenet layers are very narrow adding only a small set of feature-maps and keep the remaining feature maps unaffected, finally the classifier classify based on all feature-maps in the network. In contrast to ResNets [24], the features have been combined by concatenation. Besides better parameter efficiency, major advantage of densenet is that their enhanced flow of data and gradients throughout the network which makes them easy to train. Each layer has direct access to the gradients leading to an implicit deep supervision [24]. Further, it is also proved that dense connections have a regularizing effect, which reduces over fitting on tasks with smaller training set sizes. Thus our proposed architecture achieves excellent performance in estimating illumination and lighting direction.

3 LightNet : Illumination and Light direction Estimation

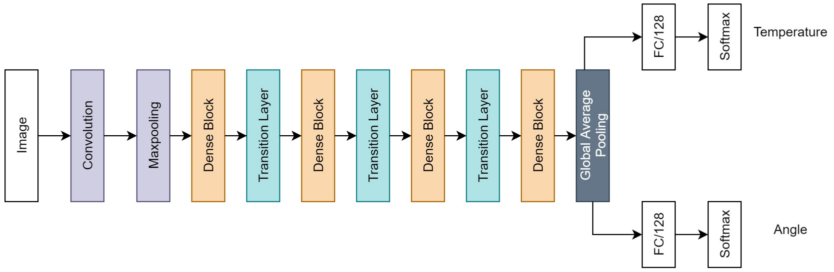

Fig. 2. Overall Block diagram of the proposed LightNet architecture

Our proposed method for estimating illumination and lighting direction uses a deep convolutional network to learn a connection between the input image and a dominant light direction. This network needs to be trained only once on a variety of augmented reality scenes, and then, it can be applied in a new scene.

The proposed architecture for illumination estimation was unified into a virtual reality rendering context and evaluated on several real scenes which were not used during training.The overall block diagram of the proposed architecture is shown in Fig.2 which mainly consist of a Dense network (Densenet) architecture [8]. The Densenet has also a superior property that it alleviates the vanishing gradient and reuses the extracted features properly. The detailed architecture of the Densenet block is shown in Fig.3. It is provided with the simple connectivity pattern to confirm maximum data flow among layers in the network, where all the layers are connected directly with each other. Each layer in the network gets additional inputs from all the previous layers and directs the feature maps to all the subsequent layers thus preserved the feed-forward nature of the Densenet architecture.

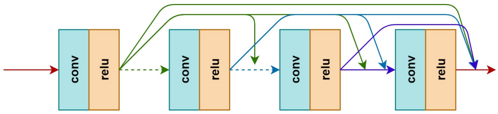

Fig. 3. Dense block

The proposed network includes N layers, each of which implements a non-linear transformation $T_n(*)$. Here n indicates the layer and denote the output of the n_{th} layer as x_n. Also different connectivity pattern is introduced in densenet that direct connections from any layer is directed to all subsequent layers. Accordingly, the nth layer obtains the feature-maps of all preceding layers, $x_0, x_1, x_2, \ldots..x_{n-1}$ as input:

$$x_n = T_n([x_0, x_1, x_2, ..., x_{n-1}]) \tag{1}$$

where $[x_0, x_1, x_2, \ldots..x_{n-1}]$ refers to the concatenation of the feature-maps formed in layers 0,1,2,... n-1. To simplify the implementation and to reduce the computational cost, the multiple inputs of $T_n(*)$ in eqn.(1) are concatenated into a single tensor.

Our network performs three consecutive operations such as Batch Normalization (BN) followed by a Rectified linear units (ReLU) and a 3×3 Convolution (Conv) function. Since the size of feature-maps changes concatenation is not possible. However, because of down-sampling operation, which changes the size of feature-maps, the proposed network is divided into multiple densely allied blocks called as dense blocks. The transition layers, the connecting lines between the dense blocks will perform convolution and pooling operation. In our experimentation, the transition layers contain a batch normalization and a 1×1 convolutional layer followed by a 2×2 average pooling layer. To decrease the number of input feature maps, bottleneck layers are involved in the network before each 3×3 convolution and thus reduce the computational cost. Our Densenet architecture consists of 58 dense blocks, followed by three transition blocks and three fully-connected layers. Totally our proposed LightNet architecture consist of 121 layers. We replaced the global average pooling and fully connected layers from the pre-trained network with new global average pooling and fully connected layers along with two output layers for lighting angle and temperature estimation. The model is trained with two softmax outputs. One output layer with 8×1 consider for lighting angle prediction and another output layer with 5×1 consider for temperature prediction. The model loss was updated by sum of temperature and lighting angle categorical cross-entropy function.

4 Evaluation and Results

4.1 Dataset

For experimentation and evaluation, we used novel Virtual Image Dataset for Illumination Transfer (VIDIT) dataset [5] for illumination estimation. VIDIT is used for the lighting estimation challenge in the AIM workshop, ECCV 2020. VIDIT comprises of 390 various unreal engine scenes, each captured with 40 illumination settings. The illumination settings are captured in all the combinations of 5 color temperatures (2500K, 3500K, 4500K, 5500K and 6500K) and 8 light directions (N, NE, E, SE, S, SW, W, NW). Resolution of each original image is 1024×1024. For evaluation all the images are normalized from 0 to 1 scale using mean-max normalization method. An example of virtual images from VIDIT dataset is shown in Fig.4. First column of the Fig.4 represents all the light directions for 2500k color temperature. Similarly, 2nd, 3rd, 4th and 5th columns of that figure represents 3500K, 4500K, 5500K and 6500K color temperature respectively.

4.2 Training Data

VIDIT illumination estimation dataset consists of 11999 images. From the whole dataset, we randomly split 67 percentage of images for training and 33 percentage for validation. Adam optimizer is learned at the rate from 0.001 to 0.00001 with 500 epochs to train the proposed model. The proposed architecture has been evaluated for temperature and lighting direction estimation on the set of test images, which are not learned by the network in the training process. Performance evaluation in terms of accuracy has been computed for color temperature and lighting direction is based on the eqn. (2) as given below.

$$Accuracy(\%) = \frac{No.of\, correct\, predictions}{Total\, number\, of\, predictions} \times 100 \tag{2}$$

The model loss was updated by sum of temperature and lighting angle categorical cross-entropy function. The temperature and angle categorical loss is defined as follows:

$$Temperature_{CLoss} = -\sum_{i=1}^{N} T_i log \hat{T}_i \tag{3}$$

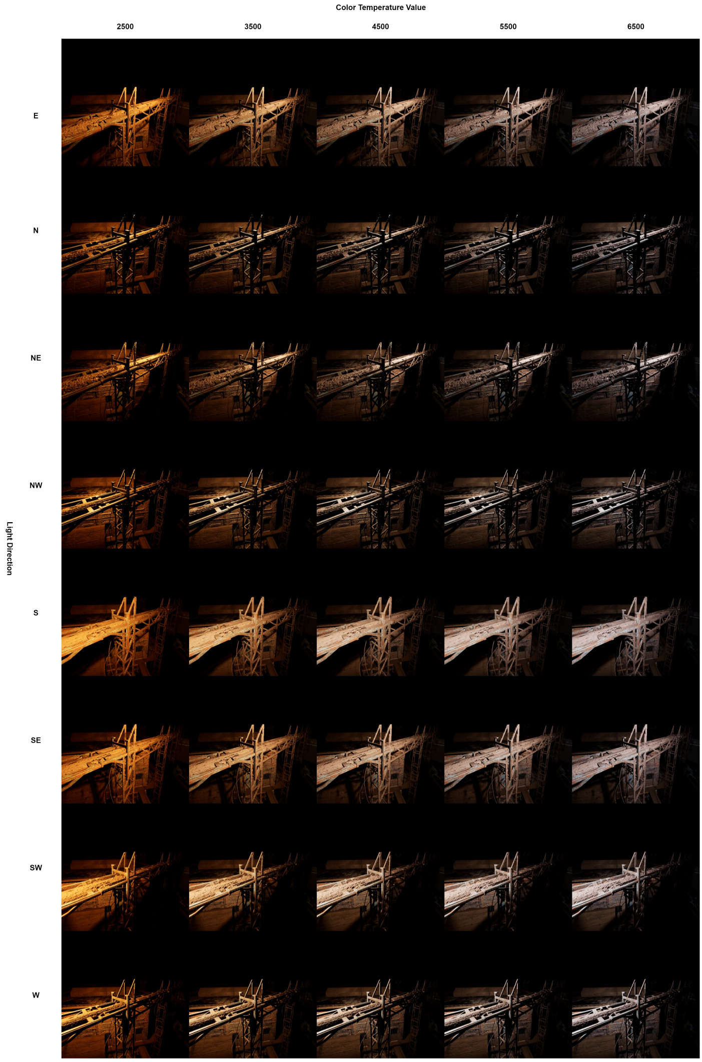

Fig. 4. An example of virtual images from VIDIT dataset. First Column: Images shown for 8 light directions (From top to bottom: E, N, NE, NW, S, SE, SW and W) for 2500k color temperature 2^{nd} column: 8 light directions for 3500k color temperature. 3^{d} column: 8 light directions for 4500k color temperature. 4^{th} column: 8 light directions for 5500k color temperature. 5^{th} column: 8 light directions for 6500k color temperature.

$$Angle_{CLoss} = -\sum_{i=1}^{N} \phi_i log\hat{\phi}_i \quad (4)$$

where i refers to all N test samples, ϕ_i are angle values [0,360] $\hat{\phi}_i$ indicates the predicted angle value, T_i is the color temperature value and $\hat{T}_i$ is the predicted color temperature value. The color temperature values 2500K, 3500K, 4500K, 5500K and 6500K, in short are takes the values of 0, 0.25, 0.5, 0.75 and 1 respectively.

The performance evaluation of the proposed LightNet model is based on the accuracy of the predictions following this formula for a loss metric:

$$Loss_M = \sqrt{\sum_{i=0}^{N-1} \left(\frac{|\hat{\phi}_i - \phi_i| mod 180}{180} \right)^2 + (\hat{T}_i - T_i)^2} \quad (5)$$

For higher accuracy the loss, $Loss_M$ should be lower.

4.3 Experimental Results

The proposed model is evaluated by estimating the illumination and light directions of VIDIT dataset. Table 1 shows the results of the proposed method on training, validation, development and testing set.The performance metrics such as Temperature loss, Angle loss and Loss metric have been computed as described in Eqn (3), Eqn (4) and Eqn (5) respectively and are tabulated in Table 1.

Table 1. Results of the proposed method on VIDIT - Illumination estimation dataset

Type	Image Count	Loss$_M$	Angle Loss	Temp Loss
Training	8000	0.000315	0.00028906	2.6875e-05
Validation	3999	0.07378	0.071142	0.00264
Development	45	0.0974	0.0597	0.0377
Test	45	0.0984	0.0513	0.0471

In the testing phase, the proposed LightNet model achieved overall loss, angle loss and temperature loss as 0.0984, 0.0513 and 0.0471 respectively. To prove the superior performance of the proposed network, evaluation has also been done with 11 different benchmark baseline models which are listed in Table 2. Compared to the results of the baseline models, our proposed LightNet architecture was outperformed in estimating the illumination level and lighting direction. Performance of the light direction accuracy, temperature accuracy and loss value are considered for comparison with all the models. From the table it is inferred that our Densenet121 model provided better results compared to other benchmark models. Similarly, Table 3 shows the experimental results of all the existing network models for the development phase data.Table 3 also shows the performance comparison of the proposed Densenet121 model with other baseline models. From the table, it is inferred that our proposed model provided results on par with Resnet 50 and DenseNet-169. Even though the Resnet-50 and Densenet -169 models produced low loss value on training and validation data, these models failed to perform well on development data. It is also observed that, EfficientNetB5 model has a higher loss than DenseNet-121 despite having a larger accuracy in the angle and color temperature estimation. This is because the loss function computation is different from the accuracy calculation. Here we calculate the loss value based on root sum squared analysis as described in Eqn.(5). Based on that metrics, the loss value, temperature loss and angle loss of EfficientNetB5 was 0.1065, 0.02733 and 0.07916 respectively. Thus despite of the accuracy, the loss value of the EfficientNetB5 was high because of high temperature loss and angle loss.

Table 2. Performance comparison of the proposed Densenet121 model with other benchmark models. T_{AA} -Training Angle Accuracy ,T_{TA} -Training Temperature Accuracy,T_{Loss_M} -Training Loss Metric,V_{DA} -Validation Angle Accuracy,V_{TA} -Validation Temperature Accuracy,V_{Loss_M} -Validation Loss Metric

Model Type	T_{AA}	T_{TA}	T_{loss_M}	V_{DA}	V_{TA}	V_{Loss_M}
DenseNet-169	99.725	99.7625	0.000492812	79.769	82.14	**0.061512**
DenseNet-201	99.912	99.93	0.000324	78.46961	85.096274	0.0622824
Xception [3]	97.625	95.925	0.00826	75.5188	73.2933	0.08175
MobileNetV2 [18]	99.3125	97.675	0.0024743	68.54	77.779	0.102825
Resnet50 [13]	99.975	99.9375	**2.84E-05**	73.49337	83.14578	0.085773
EfficientNetB0 [17]	98.15	96.2	0.005799	67.866	78.769	0.0952894
EfficientNetB1	99.175	98.25	0.00342593	72.568	80.5701	0.08677
EfficientNetB2	94.975	87.05	0.021672	65.616	75.393	0.11652
EfficientNetB3	99.5625	97.05	0.002173	74.793	77.469	0.0838
EfficientNetB4	99.45	97.35	0.002149	74.89	80.6201	0.077486
EfficientNetB5	99.45	98.2625	0.0020628	77.844	83.795	0.06732
DenseNet-121	99.9375	99.7875	0.000315	76.069	84.2460	0.07378

To prove the efficacy of the proposed LightNet architecture, we also measured the performance metrics such as Precision, Recall and F1 score values. Table 4 lists the above performance metrics values on development data for the light directions 0, 45, 90, 135, 180, 225, 270 and 315. From the table it is observed that for 90 angle light direction,our method achieved maximum of one in all the three metrics. Likewise, Table 5 shows the performance on development data for various color temperature values. It is observed that, for 0.25 temperature, our method attained higher value of precision, recall and F1 score as 0.8, 0.5 and 0.62 respectively. The experimental results overall indicate that the proposed model achieves higher accuracy on light direction and illumination estimation than the compared baseline methods on challenging dataset.

Table 3. Comparison of experimental Results of Development phase data with benchmark models

Model Type	Angle Accuracy	Temperature Accuracy	Loss_M
DenseNet-169	66.67	42.223	0.159
DenseNet-201	66.67	42.224	0.16389
Xception	64.445	42.223	0.19827
MobileNetV2	51.112	46.667	0.2456
Resnet50	57.78	44.45	1.91E-01
EfficientNetB0	64.445	40	0.2123
EfficientNetB1	68.89	42.23	0.1899
EfficientNetB2	64.445	42.23	0.160778
EfficientNetB3	71.12	37.75	0.16988
EfficientNetB4	68.889	46.667	0.16678
EfficientNetB5	77.77	51.11	0.1065
DenseNet-121	75.556	42.223	**0.09744**

Table 4. Evaluation of performance metrics of light direction on development data. I_{count} represents the class wise image count.

Angle	Precision	Recall	F1-Score	I_{count}
0	0.62	1	0.77	5
45	1	0.8	0.89	5
90	1	1	1	4
135	1	0.5	0.67	4
180	0.71	0.83	0.77	6
225	0.71	0.71	0.71	7
270	0.6	0.5	0.55	6
315	0.75	0.75	0.75	8

From the Fig.5 it is understood that the light direction can be estimated accurately compared to color temperature estimation which is still an unsolvable issue in the field of virtual reality. As an extension, along with the training features, the depth information can also be added to estimate the color temperature accurately.It is also observed that for all kind of images our method estimates illumination and light direction which is visually acceptable and comparable. Moreover, the results show that the proposed LightNet model can estimate illumination and light direction in virtual reality images which were not learnt in a training set.

Table 5. Evaluation of performance metrics of color temperature on development data

Temperature	Precision	Recall	F1-Score	I_{count}
2500	0	0	0	13
3500	0.8	0.5	0.62	8
4500	0.33	0.6	0.43	5
5500	0	0	0	15
6500	0	0	0	4

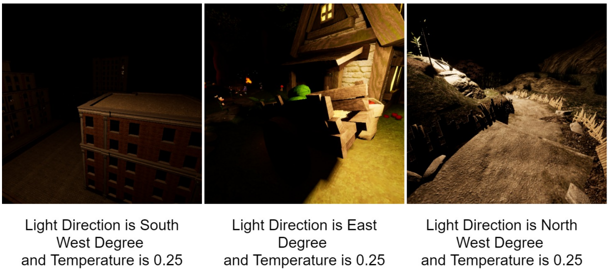

Fig. 5. Sample Images of development set

Computational complexity: In the proposed method, the training parameters of the model are 7,219,981 and the model size is 29.66MB. Finally, we measured the computation time of our method. During training and testing, the run time consumed per image with the size of $224 \times 224 \times 3$ is 0.019979 seconds. For training the other baseline models, same loss function and optimization was trained with 100 epochs. The proposed network was trained and tested with the Intel Core i7 processor, GTX 1080 GPU, 8GB RAM in Keras.

5 Conclusions

Factual illumination of virtual objects inserted in real scenes is one of the important challenges of a virtual reality system. This paper presented a novel LightNet architecture using Densenet121 for estimating illumination and lighting direction in virtual images. Our model is trained with two softmax outputs for color temperature and lighting direction prediction. The proposed architecture is evaluated on challenging VIDIT-illumination estimation dataset. Our experimental results, overall proved that our method achieves higher accuracy, F1 score and minimal loss on estimated light direction and illumination than the compared state-of-the-art methods. Our proposed model have also submitted to the AIM 2020 challenge and secured top position among the participants. In that we used the ImageNet pre-trained weights and obtained the result in terms of loss as 0.0984. Future direction of the present work should focus on the estimation of dynamic illumination as well as intensity of light sources to adjust the illumination even more realistic in virtual reality scenes. Also, an attentionbased classifier network can be created to train the model inclusive of depth information with the input to increase the overall accuracy.

References

1. Bastiaan J.Boom, Sergio.Orts-Escolano, X.Ning, X., McDonagh.Steven, Peter.Sandilands, Robert B.Fisher: Interactive light source position estimation for augmented reality with an rgb-d camera. Computer Animation and Virtual Worlds (2017). https://doi.org/10.1002/cav.1686, e1686 cav.1686
2. Bruno Augusto.D.M, Rafael Rego.D, Cristina Nader.V, Esteban.Clua: Deep light source estimation for mixed reality. In: Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 1 GRAPP. pp. 303–311. INSTICC, SciTePress (2018). https://doi.org/10.5220/0006724303030311
3. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1800–1807 (2017)
4. Clements, M., Zakhor, A.: Interactive shadow analysis for camera heading in outdoor images. In: 2014 IEEE International Conference on Image Processing (ICIP). pp. 3367–3371 (2014)
5. El Helou, M., Zhou, R., Barthas, J., Sstrunk, S.: Vidit: Virtual image dataset for illumination transfer. arXiv preprint (2020). https://doi.org/arXiv:2005.05460
6. Elizondo, D.A.Zhou, S.M.Chrysostomou: Light source detection for digital images in noisy scenes: a neural network approach. In: Neural Comput. Appl. vol. 28, p. 899–909 (2017)
7. Frahm, Koeser, Grest, Koch: Markerless augmented reality with light source estimation for direct illumination. In: Visual Media Production, 2005. CVMP 2005. The 2nd IEE European Conference on. pp. 211–220 (2005)
8. G.Huang, Z.Liu, L.Van Der Maaten, K.Q.Weinberger: Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 2261–2269 (2017)
9. I.Arief, S.McCallum, J.Y.Hardeberg: Realtime estimation of illumination direction for augmented reality on mobile devices. In: Color and Imaging Conference 2012 . Society for Imaging Science and Technology. p. 111–116 (2012)
10. Jachnik, J., A.Newcombe, R., A.J.Davison: Real-time surface light-field capture for augmentation of planar specular surfaces. In: 2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 91–97 (2012)
11. J.T.Barron, J.Malik: Intrinsic scene properties from a single rgb-d image. In: 2013 IEEE Conference on Computer Vision and Pattern Recognition. pp. 17–24 (2013)
12. Kevin E. Soulier, M., Martín L. Larrea: Real-time estimation of illumination direction for augmented reality with low-cost sensors. In: Conference: XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) (2016)
13. K.He, X.Zhang, S.Ren, J.Sun: Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 770–778 (2016)
14. Kán, P., Kafumann, H.: Deeplight: light source estimation for augmented reality using deep learning. In: The Visual Computer (2019). vol. 35:873–883 (2019)
15. L.Gruber, T.Langlotz, P.Sen, T.Höherer, Schmalstieg, D.: Efficient and robust radiance transfer for probeless photorealistic augmented reality. In: 2014 IEEE Virtual Reality (VR). pp. 15–20 (2014)

16. Martînez, H., Laukkanen, S.: Towards an augmented reality guiding system for assisted indoor remote vehicle navigation. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems **2**(2) (2 2015). https://doi.org/10.4108/inis.2.2.e3
17. Mingxing Tan, Quoc V.Le: Efficientnet: Rethinking model scaling for convolutional neural networks. In: ICML) (2019). https://doi.org/arXiv:1905.11946v3
18. M.Sandler, A.Howard, M.Zhu, A.Zhmoginov, L.Chen: Mobilenetv2: Inverted residuals and linear bottlenecks. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4510–4520 (2018)
19. Rohmer, K., Büschel, W., Dachselt, R., Grosch, T.: Interactive near-field illumination for photorealistic augmented reality on mobile devices. In: 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 29–38 (2014)
20. X.Chen, K., X.Jin: Single image based illumination estimation for lighting virtual object in real scene. In: 12th International Conference on Computer-Aided Design and Computer Graphics. vol. 450-455 (2011)
21. Xing Guanyu, Zhou Xuehong, Peng Qunsheng, Liu Yanli, Qin Xueying: Lighting simulation of augmented outdoor scene based on a legacy photograph. Computer Graphics Forum pp. 101–110 (2013). https://doi.org/10.1111/cgf.12217, https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.12217
22. Y.Dong, G.Chen: Appearance from-motion: recovering spatially varying surface reflectance under unknown lighting. ACM Trans. Graph. (2014). https://doi.org/10.1145/2661229.2661283
23. Y.Liu, X.Granier: Online tracking of outdoor lighting variations for augmented reality with moving cameras. IEEE Transactions on Visualization and Computer Graphics pp. 573–580 (2012)
24. Zhou Hao, Hadap Sunil, Sunkavalli Kalyan, Jacobs David W.: Deep single-image portrait relighting. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (October 2019)