

## Repositório ISCTE-IUL

---

Deposited in *Repositório ISCTE-IUL*:

2022-08-01

Deposited version:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Hamad, M., Conti, C., Nunes, P. & Soares, L. D. (2022). View-consistent 4D Light Field style transfer using neural networks and over-segmentation. In 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). Nafplio: IEEE.

Further information on publisher's website:

10.1109/IVMSP54334.2022.9816312

Publisher's copyright statement:

This is the peer reviewed version of the following article: Hamad, M., Conti, C., Nunes, P. & Soares, L. D. (2022). View-consistent 4D Light Field style transfer using neural networks and over-segmentation. In 2022 IEEE 14th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP). Nafplio: IEEE., which has been published in final form at <https://dx.doi.org/10.1109/IVMSP54334.2022.9816312>. This article may be used for non-commercial purposes in accordance with the Publisher's Terms and Conditions for self-archiving.

---

### Use policy

Creative Commons CC BY 4.0

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in the Repository
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

---

# View-consistent 4D Light Field Style Transfer using Neural Networks and Over-segmentation

Maryam Hamad, Caroline Conti, Paulo Nunes, Luís Ducla Soares  
Instituto de Telecomunicações, Instituto Universitário de Lisboa (ISCTE-IUL)  
Lisbon, Portugal  
{maryam.hamad, caroline.conti, paulo.nunes, lds}@lx.it.pt

**Abstract**—Deep learning has shown promising results in several computer vision applications, such as style transfer applications. Style transfer aims at generating a new image by combining the content of one image with the style and color palette of another image. When applying style transfer to a 4D Light Field (LF) that represents the same scene from different angular perspectives, new challenges and requirements are involved. While the visually appealing quality of the stylized image is an important criterion in 2D images, cross-view consistency is essential in 4D LFs. Moreover, the need for large datasets to train new robust models arises as another challenge due to the limited LF datasets that are currently available. In this paper, a neural style transfer approach is used, along with a robust propagation based on over-segmentation, to stylize 4D LFs. Experimental results show that the proposed solution outperforms the state-of-the-art without any need for training or fine-tuning existing ones while maintaining consistency across LF views.

**Keywords**—light field; angular consistency; deep learning; neural style transfer; superpixels

## I. INTRODUCTION

Appealing paintings and artwork have attracted people for thousands of years. In the past, a skilled artist was always required to create a painting with a specific style, brush strokes and color palette, which typically took a long time. With the recent advances in learning-based techniques and the advent of style transfer, such creation is now possible to be performed by computers. Style transfer is an image editing application in which a new image is generated by combining the content of one image with the style of another one (e.g., a famous painting). Style transfer is a long-standing research area in the broader area of texture synthesis [1], [2]. Recently, with the rapid development of deep learning, neural networks are being used to solve the style transfer task. Gatys et al. [3] were the first to apply Convolutional Neural Networks (CNN) to stylize an image. In their work, CNNs are used to extract the feature maps of the content image (i.e., the image from which the content will be transferred) and style image (i.e., the image from which the style will be transferred). Afterwards, a target image (i.e., the stylized image that combines the content image with the style image) is iteratively optimized by minimizing a loss function. Johnson et al. [4] improved the performance of [3] by training a feed-forward network for each style image and generating a stylized image with only one forward pass in the testing stage. Although it is 3 times faster than [3], the solution in [4] is not flexible in terms of the number of used styles since it requires training for each style. Additionally, other neural networks have

also been exploited to achieve style transfer, such as generative adversarial networks that require paired training data to learn a specific style, which is not always available and may limit their applications [1]. Moreover, Neural Style Transfer (NST) has been extended to consider videos [5] and different imaging modalities, such as stereo imaging [6] and 4D Light Fields (LF) [7], [8]; interested readers are encouraged to read the recent comprehensive reviews of the existing NST solutions in [1], [2].

4D LFs involve rich information since not only the light intensity is captured but also ray directions [9]. LFs capture the same scene from different perspectives, thus allowing interesting applications such as depth or disparity estimation (i.e., the displacement of a point between different views, which is inversely proportional to the depth), view synthesis and post-capture refocusing [9], [10]. 4D LFs can be represented as an array of views  $I(x, y, u, v)$ , where  $(x, y)$  are the spatial coordinates, and  $(u, v)$  are the angular coordinates of each view. When fixing one angular and one spatial coordinates, an Epipolar Plane Image (EPIs) (i.e., the unique 2D spatio-angular LF slice typically containing a regular structure with several oriented lines [11]) can be obtained as illustrated in Figure 1.

While generating stylized images that are visually pleasant is an important criterion for 2D images, maintaining cross-view consistency is also essential for 4D LFs. More precisely, directly applying 2D image or video style transfer methods to the entire 4D LF views, without considering the correlation between them, may result in inconsistent stylized LFs with highly unnatural artifacts. Only a few solutions are available in the literature that consider 4D LF cues in the style transfer application. Hart et al. [7] proposed an extension to the work of Johnson et al. [4] by adding a disparity loss term to the loss function. The disparity loss is computed by finding the difference between each stylized LF view and the stylized central view warped into that view. The disparity loss is then backpropagated through the network. This repeats for each LF view until convergence is reached. While their work considers cross-view consistency, it requires optimizing each LF view iteratively (assuming dense LFs).

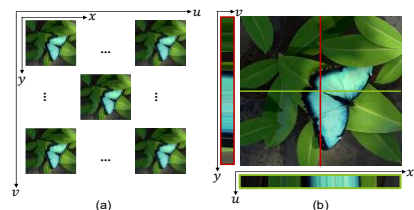


Fig. 1. Example of light field representations. a) 4D light field represented as an array of views; b) Horizontal and vertical EPIs.

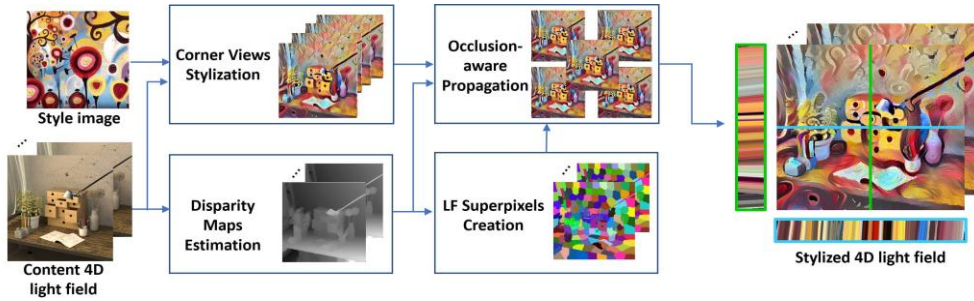


Fig. 2. Overview of the proposed method for view-consistent 4D LF neural style transfer. By combining the style of 2D image with the content of 4D LFs and applying an occlusion-aware propagation, a consistent 4D stylized LF is generated.

Moreover, although the feed-forward approach is fast, it needs to be trained for each style, hence, limiting style selection flexibility. Egan et al. [8] addressed these drawbacks and proposed a novel NST method that considers local angular consistency. Their work extended the Gatys et al. work [3] by adding the local angular consistency loss in the total loss function. Although their work ensures local angular consistency for LFs with larger disparity ranges, applying optimization using this technique for each view is very time-consuming.

The contribution of this paper is a novel 4D LF NST method that overcomes the limitations of the existing methods by:

- **Enabling NST flexibility (in terms of the number of styles that can be used) with less computational complexity:** to achieve that, the optimization-based NST [3] method (which does not require training a model for each style image) is used. To reduce the optimization-based NST complexity significantly, only a limited set of views (i.e., the four corner views) are initially stylized using the method in [3] (different from [7] and [8] that require optimizing each LF view).
- **Improving 4D LF view-consistency:** by exploiting LF over-segmentation (that adheres to object boundaries and maintains LF view-consistency), the edits from all corner views are propagated into each LF view using per-pixel disparity in an occlusion-aware manner. The proposed method outperforms the existing solutions without training or fine-tuning the existing NST models.

The remainder of this paper is organized as follows: Section II describes the proposed method in detail, and Section III evaluates its performance through a series of experimental results. Finally, Section IV concludes the paper with some final remarks and proposes directions for future work.

## II. PROPOSED METHOD

The proposed method contains four main steps as presented in Figure 2. Given a style image and a 4D LF, the four corner views are initially stylized using optimization-based NST [3]. After that, disparity maps for all input LF views are estimated using [12]; to ensure spatio-angular consistency during the propagation. Next, the 4D LF is over-segmented into spatio-angular coherent regions (a.k.a superpixels), as in [13] to facilitate the propagation and respect object boundaries and occlusions. Afterwards, the stylization is propagated into all LF views through occlusion-aware back-projection from each view into all corner views. Finally, remaining isolated non-stylized

pixels that emerged after back-projections due to occlusions, are filled robustly. Each step is detailed in the following subsections.

### A. Corner Views Stylization

Initially, only the extreme four corner views are stylized using the approach in [3]. The corner views are selected since they typically contain the maximum scene information including dis-occlusions. The approach in [3] aims at minimizing the distances of the feature representation between the content/style image and the target one in one or more layers of the CNN. The target image is initially generated using a white noise image and iteratively optimized using the loss function,  $\mathcal{L}_{total}$ , defined by (1), where  $\mathcal{L}_{content}$ , is the content loss and  $\mathcal{L}_{style}$ , is the style loss. To ensure view-consistent stylization, the initial white noise is set the same for all corner views. Finally, to control the output, two weighting factors (i.e., the content weight,  $\alpha$ , and the style weight,  $\beta$ ) are included:

$$\mathcal{L}_{total} = \alpha\mathcal{L}_{content} + \beta\mathcal{L}_{style}. \quad (1)$$

Notice that the proposed method is independent of the used 2D NST method. However, the approach in [3] is used due to its flexibility to transfer any style and it enables controlling the target images by adjusting the weights in (1). Moreover, any number of views or angular positions can be used but the results may be influenced accordingly.

### B. Disparity Maps Estimation

LF imaging provides rich information, which makes it possible to estimate a disparity map for each LF view. In this paper, the proposed method in [12] (that estimates disparities from each view to its right adjacent view) is used to estimate disparity maps for all LF views. Per-pixel disparity is used here to ensure consistent pixel projection during the propagation step.

### C. Light Field Superpixel Creation

LF over-segmentation is capable of adhering to object boundaries and creating a unique label for each homogenous region to facilitate subsequent editing tasks. In this paper, the recently proposed Adaptive LF Over-segmentation (ALFO) method [13] is used to guide the propagation in an occlusion-aware manner. The ALFO method exploits color, disparity and position features to apply adaptive K-means clustering. Additionally, it can robustly balance accuracy, shape regularity and view-consistency. In our experiments, the superpixel size is set to 20 as suggested in [13] as a reasonable size for robust adhesion to the borders.

#### D. Occlusion-aware Propagation

Given the LF disparity maps, LF superpixels and stylized corner views, the stylization now can be propagated into all other 4D LF views. Initially, each LF view is back-projected into all corner views using its disparity map (2):

$$\begin{aligned} x_i^{ref} &= x_i^{(u,v)} + d_{hor}^{(u,v) \rightarrow ref}, \\ y_i^{ref} &= y_i^{(u,v)} + d_{ver}^{(u,v) \rightarrow ref}, \end{aligned} \quad (2)$$

where  $x_i^{(u,v)}$ ,  $y_i^{(u,v)}$  are the spatial position coordinates of a pixel,  $i$ , which is located in a view of angular coordinates  $(u, v)$ ,  $x_i^{ref}$  and  $y_i^{ref}$  are the spatial position in a reference view (i.e., *ref* in this paper represents a single corner view, hence, the same equation is applied for all corner views independently), and  $d_{hor}^{(u,v) \rightarrow ref}$ ,  $d_{ver}^{(u,v) \rightarrow ref}$  are the horizontal and vertical disparity from view  $(u, v)$  to the reference view. The used disparity estimation method [12] estimates disparity for adjacent views, therefore, for regularly sampled 4D LF views back-projection is applied by multiplying the disparity value by  $(u_{ref} - u)$ ,  $(v_{ref} - v)$  when computing  $x_i^{ref}$  and  $y_i^{ref}$ , respectively [13]. These equations are applied in the case of parallel optical light field capturing assumption, as in [13]–[16]. Otherwise, intrinsic and extrinsic camera parameters should be considered.

Since the projected pixel coordinates may belong to  $\mathbb{R}^2$ , and to ensure integer indexing (since the visual information is only available for integer indices), the four neighboring pixels,  $\mathcal{N}_i \in \{I_a, I_b, I_c, I_d\}$ , of the back-projected pixel with integer positions ( $\in \mathbb{Z}^2$ ) are considered as presented in Figure 3. However, consistency is checked by comparing the label and disparity of the pixel in  $(x_i^{(u,v)}, y_i^{(u,v)})$  and all pixels in  $\mathcal{N}_i$  to choose which ones to be used for the interpolation. To overcome possible projection errors, due to disparity errors or discontinuities in superpixels, two conditions are checked before interpolation:

- At least one pixel in  $\mathcal{N}_i$  has the same label as the pixel in its original location  $(x_i^{(u,v)}, y_i^{(u,v)})$ .
- The absolute disparity difference between a pixel disparity in view  $(u, v)$  and at least one pixel disparity in  $\mathcal{N}_i$  is less than a threshold value,  $\varepsilon$ . We empirically set  $\varepsilon = 0.1$ ; since a superpixel with size (i.e., 20) is noticed to have, typically, similar disparity values.

If any of the above conditions holds for all pixels in  $\mathcal{N}_i$  or part of them, then only these pixels are valid for interpolation. Interpolation is applied by computing the bilinear interpolation of valid pixels in  $\mathcal{N}_i$ , otherwise, no interpolation is computed.

After computing the interpolated value from each corner view, the pixel in its original angular location  $(u, v)$  is set to the mean color value of all valid back-projections from the four corner views. The mean is used after extensive experiments since it shows the best visual and numerical results when compared to using the median or weighted sum and maintains

consistency across LF views. By doing this, only very few sparse and isolated pixels that have no projection, or invisible regions due to the angle of view, remain unstylized. To fully stylize all LF views, these remaining isolated pixels are filled by applying inward interpolation using the widely used region filling based on the Laplace equation as in [17].

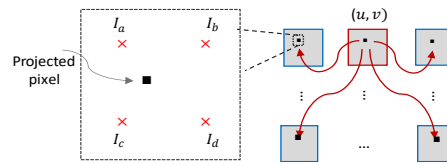


Fig. 3. Example of back-projection: a pixel in  $(u, v)$  view that needs to be stylized is back-projected into each corner view (in blue squares).

### III. EXPERIMENTAL RESULTS

In this section, several methods are used as benchmarks to evaluate the performance of the proposed method. Firstly, two different baseline methods are considered, as in [8]: i) by applying Independent View Stylization (IVS) using existing 2D NST [3] to all LF views independently; and ii) by applying Pseudo Video Stylization (PVS) as proposed (for videos) in [5] for styling a pseudo video sequence of 4D LF views. To the best of the authors' knowledge, only two recently proposed methods are specifically focused on tackling 4D LF challenges. The first one focuses on Global Angular Consistency Stylization (GACS) [7], and the second one focuses on Local Angular Consistency Stylization (LACS) [8]. Moreover, different synthetic and real-world LF datasets and style images are used, as shown in TABLE I. For quantitative evaluations, two different metrics are used to evaluate the view-consistency namely: i) the LF Epipolar Consistency (LFEC) metric defined in [18]; and ii) the LF Angular Consistency (LFAC) metric<sup>1</sup> defined in [8]. The LFEC and LFAC metrics evaluate the angular consistency by back-warping LF views into a reference view and finding the color variance. Different than the LFEC metric that back-warps all LF views into the central view, the LFAC metric back-warps into the center view of a local window of views; to robustly consider large occluded regions. Both metrics require estimating disparity to apply back-warping, therefore, we estimated per-pixel disparity maps, for our results and all benchmark methods by using [12]. We noticed that, by using [12], the metric results of the benchmark methods are improved. Moreover, the disparity loss (which is the amount of disparity changes) is evaluated by using the disparity Mean Square Error (MSE) metric defined in [7]. This metric computes the  $MSE \times 100$  between the central view disparity map estimated from the original LFs and the stylized ones. As in [8], the disparity estimation method in [19] is used. Results of all metrics are presented in TABLE II. Due to the limitation in the paper size, only the central view with horizontal EPIs are presented in TABLE III. However, we encourage the reader to see our and dynamic results<sup>2</sup> for all LF views for clear view-consistency evaluation. For the used NST implementation, we used standard GPU-based MATLAB implementation [20] and we set  $\alpha = 50$ ,  $\beta = 10^3$ , the same values as used in the benchmark methods.

<sup>1</sup>Software implementation of all the used metrics can be found at: <https://github.com/doegan32/Light-Field-Style-Transfer>

<sup>2</sup>Dynamic results for all LF views can be found at: <https://github.com/MaryamHamad/LFStyleTransfer>

The proposed method generates outperforming angular consistency in both LFEC and LFAC metrics, as can be seen in TABLE II. For the MSE metric, the GACS method achieves the best average results and preserves better object boundaries; hence, generates the central disparity maps that are similar to the original LF ones. However, it requires a pre-trained NST model as input for each style image. In this paper, corner views are used to minimize the occlusions, hence, there are no large holes left after propagation in densely sampled LFs. However, our method can be extended to consider sparse LFs that may have largely occluded regions by simply adding more reference views to consider all objects in LF views. The used technique for filling the holes in dis-occluded regions after propagation may generate some artifacts (which also occur in the benchmark methods) and thus requires further investigation. For time complexity, the proposed method reduces the needed time to stylize the entire LF significantly, i.e., for a LF with 81 views instead of taking  $81 \times T_s$ , where  $T_s$  is the average time needed to stylize a single view, it takes less than  $10 \times T_s$  including LF disparity estimation and superpixel generation. Finally, it can be observed that neither applying 2D methods for each view independently nor existing methods for video are adequate solutions for 4D LFs.

#### IV. FINAL REMARKS

In this paper, a novel view-consistent 4D LF NST method is proposed. Without any further training for new deep learning models or fine-tuning existing ones, we exploited an existing optimization-based NST method to initially stylize only four corner views. Afterwards, the stylized views are propagated into all other LF views in an occlusion-aware manner by using LF superpixels. Experimental results have been shown to outperform the considered benchmark methods and produce visually appealing and consistent results across all LF views.

For future work, we will extend style transfer to sparse LFs that include wide occlusions. Additionally, we will study other applications of the proposed propagation technique, such as semantic segmentation and object removal, where the edits are applied in reference views and propagated into other LF views.

#### ACKNOWLEDGMENT

The authors would like to thank Mr. Dónal Egan for publishing the software of their method [8] including the evaluation metrics and results that facilitated their comparison.

TABLE I. TEST IMAGES USED IN OUR EXPERIMENTS

























	content 4D LFs	Disparity range	Style image	Thumbnails: (content, style)
a	Swan [21]	[-1, 1]	Candy	 
b	Lego knights [22]	[-3, 3]	Rain princess	 
c	Bikes [23]	[-1, 1]	Rain princess	 
d	Herbs [24]	[-3, 1.8]	Starry night	 
e	Table [24]	[-2, 1.6]	Candy	 

TABLE II. ANGULAR CONSISTENCY (LFEC, LFAC) AND DISPARITY LOSS (MSE $\times$ 100) METRICS

Metric		IVS (baseline)	PVS (baseline)	GACS [7]	LACS [8]	Ours
LFEC (↑)	a	19.29	25.24	30.31	28.43	<b>40.83</b>
	b	19.50	21.18	23.16	24.51	<b>29.21</b>
	c	22.14	23.10	33.77	27.48	<b>42.92</b>
	d	22.68	24.80	22.53	27.96	<b>31.29</b>
	e	19.02	22.03	28.19	25.72	<b>32.38</b>
Avg.		20.53	23.27	27.59	26.82	<b>35.33</b>
LFAC (↑)	a	29.74	41.27	46.48	42.72	<b>53.54</b>
	b	30.51	35.29	37.82	38.09	<b>44.00</b>
	c	33.66	37.70	48.45	42.13	<b>54.93</b>
	d	33.37	38.92	34.92	41.58	<b>45.49</b>
	e	29.65	36.69	42.54	39.81	<b>46.85</b>
Avg.		31.39	37.97	42.04	40.87	<b>48.96</b>
MSE $\times$ 100 (↓)	a	209.19	2.56	1.01	2.11	<b>0.54</b>
	b	256.36	22.15	<b>13.14</b>	19.50	18.85
	c	29.28	2.09	<b>0.91</b>	1.94	2.40
	d	96.08	11.68	8.10	8.31	<b>5.00</b>
	e	284.98	8.45	<b>1.39</b>	5.20	3.43
Avg.		175.18	9.39	<b>4.91</b>	7.41	6.04

TABLE III. VISUAL COMPARISON WITH BENCHMARK METHODS

	Content LF	Style image	IVS (baseline)	PVS (baseline)	GACS [7]	LACS [8]	Ours
a							
b							

## REFERENCES

- [1] Y. Jing, Y. Yang, Z. Feng, J. Ye, Y. Yu, and M. Song, "Neural Style Transfer: A Review.," *IEEE Trans. Vis. Comput. Graph.*, vol. 26, no. 11, pp. 3365–3385, Nov. 2020.
- [2] G. Sohaliya and K. Sharma, "An Evolution of Style Transfer from Artistic to Photorealistic: A Review.," *2021 Asian Conf. Innov. Technol. (ASIANCON)*, Pune, India, Aug. 27-29, 2021, pp. 1-7.
- [3] L. A. Gatys, A. S. Ecker, and M. Bethge, "Image Style Transfer Using Convolutional Neural Networks.," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, June 27-30, 2016, pp. 2414–2423.
- [4] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution.," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9906 LNCS, pp. 694–711, 2016.
- [5] M. Ruder, A. Dosovitskiy, and T. Brox, "Artistic Style Transfer for Videos.," *German Conf. Pattern Recog.*, Springer Verlag, 2016, pp. 26–36.
- [6] D. Chen, L. Yuan, J. Liao, N. Yu, and G. Hua, "Stereoscopic Neural Style Transfer.," in *2018 IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, June 18-23, 2018, pp. 6654–6663.
- [7] D. Hart, J. Greenland, and B. Morse, "Style transfer for light field photography.," *IEEE Winter Conf. Appl. Comput. Vision (WACV)*, Snowmass, CO, USA, 1-5 Mar., 2020, pp. 99–108.
- [8] D. Egan, M. Alain, and A. Smolic, "Light Field Style Transfer with Local Angular Consistency.," *IEEE Int. Conf. Acoust., Speech, Signal Process (ICASSP)*, Toronto, ON, Canada, June 6-11, 2021, pp. 2300–2304.
- [9] G. Wu *et al.*, "Light Field Image Processing: An Overview.," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 7, pp. 926–954, Oct. 2017.
- [10] S. Zhou, T. Zhu, K. Shi, Y. Li, W. Zheng, and J. Yong, "Review of light field technologies.," *Vis. Comput. Ind. Biomed. Art*, vol. 4, no. 29, Dec. 2021.
- [11] R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion.," *Int. J. Comput. Vis.*, vol. 1, no. 1, pp. 7–55, 1987.
- [12] N. Khan, M. H. Kim, and J. Tompkin, "View-consistent 4D Light Field Depth Estimation.," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, 2020.
- [13] M. Hamad, C. Conti, P. Nunes, and L. D. Soares, "ALFO: Adaptive Light Field Over-Segmentation.," *IEEE Access*, vol. 9, pp. 131147–131165, 2021.
- [14] M. Hog, N. Sabater, and C. Guillemot, "Superrays for Efficient Light Field Processing.," *IEEE J. Sel. Top. Signal Process.*, vol. 11, no. 7, pp. 1187–1199, Oct. 2017.
- [15] N. Khan, Q. Zhang, L. Kasser, H. Stone, M. H. Kim, and J. Tompkin, "View-Consistent 4D Light Field Superpixel Segmentation.," in *Int. Conf. Comput. Vis. (ICCV)*, Seoul, Korea, 27 Oct.–2 Nov., 2019, pp. 7810–7818.
- [16] H. Zhu, Q. Zhang, Q. Wang, and H. Li, "4D Light Field Superpixel and Segmentation.," *IEEE Trans. Image Process.*, vol. 29, pp. 85–99, July. 2019.
- [17] "Fill in specified regions in image using inward interpolation - MATLAB regionfill - MathWorks.," [Online]. Available: <https://www.mathworks.com/help/images/ref/regionfill.html>. [Accessed: 28-Mar-2022].
- [18] P. David, M. Le Pendu, and C. Guillemot, "Angularly Consistent Light Field Video Interpolation.," *IEEE Int. Conf. Multimedia Expo (ICME)*, London, UK, July 6-10, 2020, pp. 1–6.
- [19] C. Shin, H. G. Jeon, Y. Yoon, I. S. Kweon, and S. J. Kim, "EPINET: A Fully-Convolutional Neural Network Using Epipolar Geometry for Depth from Light Field Images.," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 4748–4757, Dec. 2018.
- [20] "Neural Style Transfer Using Deep Learning - MATLAB & Simulink.," [Online]. Available: <https://www.mathworks.com/help/images/neural-style-transfer-using-deep-learning.html>. [Accessed: 8-Apr-2022].
- [21] M. Rerabek and T. Ebrahimi, "New Light Field Image Dataset.," in *8th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, Lisbon, Portugal, 2016.
- [22] V. Vaish and A. Adams, "The (New) Stanford Light Field Archive.," *Stanford*, 2008. [Online]. Available: <http://lightfield.stanford.edu/acq.html>. [Accessed: 8-Apr-2022].
- [23] "Stanford Lytro Light Field Archive.," [Online]. Available: <http://lightfields.stanford.edu/LF2016.html>. [Accessed: 8-Apr-2022].
- [24] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, "A Dataset and Evaluation Methodology for Depth Estimation on 4D Light Fields.," in *13th Asian Conf. Comput. Vis., Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part III*, S.-H. Lai, Ed. Cham: Springer, 2017, pp. 19–34.