Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Supplementary materials for

Chen JIA, Fan SHI, Meng ZHAO, Shengyong CHEN, 2022. Light field imaging for computer vision: a survey. *Front Inform Technol Electron Eng*, 23(7):1077-1097. https://doi.org/10.1631/FITEE.2100180

1 Light field imaging processing

1.1 Light field super-resolution

1.1.1 Light field data structure based methods

The light field (LF) data structure based methods focus mainly on the use and exploration of spatial and angular information.

Considering the trade-off between spatial and angular resolution, Georgiev et al. (2006) proposed that sparse sampling is beneficial to obtain better spatial resolution for integral photography. Bishop et al. (2009) first proposed LF super-resolution (LFSR) reconstruction. They performed LFSR reconstruction of down-sampled images collected in different directions from the microlens arrays of the traditional plenoptic camera. They applied the blind deconvolution method (BDM) to restore the LF. Bishop and Favaro (2012) introduced a novel approach that uses the Bayesian framework to reconstruct more LF information in super-resolution. Wanner and Goldluecke (2014) proposed a fully continuous model for variational super-resolution view synthesis.They also introduced a variational model for super-resolution view synthesis to further improve the mathematical framework for variational LF analysis. Mitra and Veeraraghavan (2012) presented a Gaussian mixture model (GMM) to address the denoising and spatial and angular super-resolution of LFs. Zhang ZT et al. (2015) presented a novel phase-based framework that can derive high-quality 4D LF from baseline stereo pairs. Zhang FL et al. (2017) introduced a highly autonomous system for LF editing relying on patch-based image synthesis methods. They also demonstrated a layered approach for the central view and a way for viewpoint consistency. Rossi and Frossard (2018) presented a novel LFSR algorithm to increase the resolution of the total views, requiring only the disparity estimation process. Based on the epipolar plane image (EPI) geometry, Zhang S et al. (2021) proposed to learn the LF sub-pixel details end-to-end at a high spatial resolution for LF spatial super-resolution and achieved superior results.

1.1.2 Learning-based methods

Since deep learning has demonstrated its superiority in data representation, a growing number of convolution neural network (CNN) based super-resolution methods have been introduced. These methods primarily concentrate on learning the relationship between high-resolution and low-resolution images.

Specifically, Dong et al. (2014) first applied CNN to the image super-resolution field and proposed a super-resolution CNN reconstruction algorithm. The algorithm takes the pre-processed low-resolution (LR) image as the input of the network and directly learns the relationship between the input and output. A convolution operation realizes the main processes of feature extraction, nonlinear mapping, and reconstruction. Wang ZW et al. (2015) combined sparse coding with CNN and proposed a cascade sparse coding based network (CSCN), which

was effectively trained in the end-to-end cascade structure, showing advantages in accuracy and visual comfort of super-resolution. Kim JW et al. (2015a, 2015b) combined residual learning with CNN to reduce the number of network parameters, to increase the speed of super-resolution, and to solve the gradient dispersion problem with the deepening of network layers. Based on the traditional deep convolutional neural network (DCNN), Yoon et al. (2015) proposed to upsample the spatial and angular resolution of the LF image via the data-driven learning method. It can generate high-resolution subaperture images efficiently. Similarly, Yoon et al. (2017) proposed a spatial super-resolution network and an angular super-resolution network based on 4D LF data analysis. These networks were trained end-to-end. In addition, to reduce the complexity of super-resolution operations in the high-resolution (HR) space, Shi et al. (2016) presented a novel CNN framework with the capability of displaying 1080p super-resolution videos in real time without a high-performance processor. Kalantari et al. (2016) proposed a method using two sequential CNNs to achieve balance between spatial and angular resolution in a LF super-resolution task. Similarly, Gul and Gunturk (2018) presented a supervised learning LFSR enhancement method with two independent subnetworks. Compared with other subnetwork methods, the proposed networks have low computational complexity, and the network structure has only three layers. Wu GC et al. (2017) converted the LF reconstruction problem into a CNN-based angle to solve the problem, as the EPI in LF data has a clear texture structure. Furthermore, Wang YL et al. (2018) applied CNN to model the correlation between adjacent subaperture images and proposed a bidirectional recurrent CNN for super-resolution reconstruction. Lim et al. (2017) proposed a novel multiscale network that can share the parameters for different scales in an LFSR. Yuan et al. (2018) introduced a combined super-resolution DCNN, including a single-image super-resolution CNN and an EPI enhancement CNN. The core of these two networks for super-resolution is to improve the spatial resolution of subaperture images and restore the LF structure. In addition, Yeung et al. (2019) introduced novel end-to-end CNN hourglass shape models for super-resolution reconstruction of LF images. Zhang S et al. (2019) proposed a residual CNN to enhance the spatial resolution of LF images. The network can inherit the LF structure information and infer subpixel information at a high resolution. Farrugia and Guillemot (2020) presented a DCNN LFSR approach that can restore the LF across all angular views. Wu GC et al. (2019) trained the CNN with a set of sheared EPIs to increase the super-resolution performance. Wang YQ et al. (2020) applied an interactive strategy to simultaneously process spatial and angular features for an LFSR. Ko et al. (2021) proposed a spatial–angular learning super-resolution algorithm that can efficiently improve low-resolution LF images. Wang YQ et al. (2021) proposed a deformable convolution super-resolution network, which overcame the challenge of incorporating angular information due to LF disparity.

1.1.3 Multi-sensor-based methods

Multi-sensor-based methods usually combine a high-resolution digital single lens reflex (SLR) camera and light field camera (LFC) and constitute a new hybrid imaging system for super-resolution.

Based on traditional and LF imaging, Wu JD et al. (2015) introduced an innovative LFSR hybrid imaging system that relies on the improved match patch algorithm and the dictionary learning approach. Boominathan et al. (2014) proposed a hybrid imaging system and a simple patch-based algorithm for super-resolution without any camera calibration information. Nevertheless, the patch-based algorithm easily blurs the high-frequency details. Therefore, by analyzing the LF structure, Zhao MD et al. (2018) proposed a high-frequency compensation super-resolution scheme that better avoids the blur of high-frequency details. Alam and Gunturk (2018) proposed a hybrid stereo imaging system, including an LFC and an Allied Vision Technologies (AVT) Mako G095C camera. High-resolution images captured by the AVT camera were used to increase the spatial resolution of subaperture images. Similarly, Zheng et al. (2017) proposed a hybrid system for super-resolution, including an LFC and a high-resolution camera. The images captured by the high-resolution camera were considered as a reference to super-resolve the low-resolution LF images. Based on a little increase in weight and cost, Wang YW et al. (2017) proposed a concept for an LF lens attachment consisting of eight low-resolution low-quality cameras that can convert a digital single lens reflex (DSLR) camera and lens into an LFC. Wang X

et al. (2016) presented a hybrid imaging system in which the LFC and DSLR cameras can share the same optical path.

Finally, to compare the results of the algorithms on three datasets in a comprehensive way, we adopt two metrics, including the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM). The quantitative comparison results are shown in Tables S1 and S2 (7×7 angular resolution, ×2 LF spatial super-resolution).

Table S1 The peak signal-to-noise ratio comparision on three LF datasets

Dataset	PSNR (dB)						
	Bicubic	Lim et al. (2017)	Rossi and Frossard (2018)	Wang YL et al. (2018)	Yeung et al. (2019)	Zhang S et al. (2019)	Wang YQ et al. (2020)
HCI1	35.82	37.10	38.04	36.46	41.31	41.02	41.50
HCI2	31.88	32.72	34.61	33.63	36.38	36.40	36.36
EPFL	31.51	32.29	32.46	32.70	35.21	34.40	35.69

Table S2 The structural similarity comparision on three LF datasets

Dataset	SSIM						
	Bicubic	Lim et al. (2017)	Rossi and Frossard (2018)	Wang YL et al. (2018)	Yeung et al. (2019)	Zhang S et al. (2019)	Wang YQ et al. (2020)
HCI1	0.941	0.954	0.964	0.965	0.977	0.975	0.978
HCI2	0.901	0.918	0.942	0.932	0.953	0.957	0.957
EPFL	0.902	0.925	0.930	0.935	0.951	0.948	0.960

PSNR and SSIM metrics are widely used in LFSR. These two metrics determine whether the super-resolution image is good or bad. The definition of PSNR is given in Eqs. (S1) and (S2). The definition of SSIM is given in Eq. (S3).

$$\mathrm{MSE}(x,y)=\frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(x_{ij}-y_{ij}\right)^2, \tag{S1}$$

$$\mathrm{PSNR}(x,y)=10\lg\left(A^2/\mathrm{MSE}(x,y)\right), \tag{S2}$$

$$\mathrm{SSIM}(x,y)=\frac{\left(2\mu_x\mu_y+C_1\right)\left(2\sigma_{xy}+C_2\right)}{\left(\mu_x^2+\mu_y^2+C_1\right)\left(\sigma_x^2+\sigma_y^2+C_2\right)}, \tag{S3}$$

where x is the input image, y is the super-resolution image, M and N are the width and height of these two images respectively, A represents the gray level in images, μ_x and μ_y are the pixel density averages, σ_x and σ_y are the standard deviations, σ_{xy} is the covariance of the images, and C_1 and C_2 are constants.

1.2 Depth estimation

1.2.1 Data-based methods

1. EPI-based methods

The essence of EPI is a 2D slice image of 4D LF data. The slope of EPI is the different positions of the same object in different views. Thus, the depth map estimation method depends on the use of different optimization techniques to measure the slope of EPI.

Wanner and Goldluecke (2012b) introduced a data item with a new expression form and estimated the depth information based on the total variation. Meanwhile, they introduced continuous optimization to the process of 4D LF data, and the local depth information obtained in the dominant direction of the EPI was continuously mapped to obtain the global depth map.

However, this algorithm cannot estimate the accurate depth information in the presence of specular reflection. In addition, LFCs have a microlens array, and the captured LF data allow the modification of both

focus and perspective viewpoints. Tao et al. (2013) leveraged the microlens array to obtain defocus and corresponding depth cues, and proposed a controllable depth image acquisition method based on x-u 2D EPI analysis.

Furthermore, to evaluate the influence of occlusion on depth estimation, Tosic and Berkner (2014) analyzed the EPI structure and converted the depth estimation task into computing the angle of rays for each pixel in a given EPI. Wanner and Goldluecke (2014) presented a new concept of local data term tailored to the continuous structure of LFs, and analyzed the EPI for disparity estimation. Krolla et al. (2014) introduced the concept of spherical LFs to conveniently construct EPIs to quickly estimate the disparity. Li JQ et al. (2015) explored the relationship between the LF structure information and the reliability of EPI, and introduced a linear system with two matrix forms, which can considerably enhance the global consistency in depth map acquisition. Diebold et al. (2016) proposed a new structural tensor with a low estimate of orientation error for depth estimation, which can effectively deal with the problem of reduced reliability in depth estimation in heterogeneous LFs. Zhang S et al. (2016) adopted the method of integrating the spinning parallelogram operator (SPO) in the 2D EPI and the depth estimation framework to keep the correct depth information. Kim C et al. (2013) proposed a course-to-coarse depth estimation scheme based on continuous EPI calculations. Johannsen et al. (2016) applied the idea of sparse coding to learn the EPI orientation–depth relationship, which improved the performance of depth estimation in accuracy and robustness. Zhang YB et al. (2017) introduced a reliability measure to remove the interference of unreliable pixels and improve the detail and overall performance of the depth map. To accurately obtain the depth information of the occlusion edge, Sheng et al. (2018) constructed a specific depth estimation function and proposed a method to calculate EPI differences in multiple directions. Schilling et al. (2018) proposed a new hybrid framework for the depth information solution, which unifies the depth information solution and occlusion information processing, ensuring that all information is effectively used.

2. LF image based methods

Jiang et al. (2018) proposed a depth estimation algorithm that can restore the whole view from a sparse set of views under occlusion conditions, and the continuity of the depth information can be kept intact. Heber and Pock (2014) proposed a new type of matching term and transformed the depth estimation problem into a multiple view stereo matching one. The distortion of the image of the microlens array leads to inaccurate depth map acquisition. Jeon et al. (2015) proposed an energy minimization formula to eliminate the distortion effect. Lin et al. (2015) proposed a new data consistency measure for depth estimation by analyzing the LF focal stack. Furthermore, based on the separation of the foreground and background for different focus planes, Lee and Park (2017) proposed an LF depth estimation method with high computational efficiency by accumulating binary maps. Chen C et al. (2014) introduced a bilateral consistency metric for LF stereo matching to estimate the depth under occlusion. Tao et al. (2015) proposed a three-stage iterative algorithm based on LF data analysis to eliminate the inaccuracy of depth information acquisition caused by specular highlights. Wang TC et al. (2016a) proposed a robust depth estimation algorithm for occluded edges by taking advantage of LF pixels to maintain photo consistency after LF refocusing. The special structure of the microlens array makes it possible to obtain additional defocus, correspondence, and shadow information in single image acquisition. At the same time, it is better to obtain the depth information of continuous dense areas. As shown in Fig. S1, Tao et al. (2017) proposed a multistage consistency depth estimation framework, including photo consistency, depth consistency, and shadow consistency. Williem and Park (2016) extended the application of correspondence and defocus cues, and then proposed a depth estimation framework that can effectively reduce noise interference based on the angular entropy metric and the adaptive defocus response. Tao et al. (2016) introduced a novel photo consistent LF depth estimation framework, which was available for glossy surfaces. Similarly, relying on deriving a spatially varying bidirectional reflectance distribution function (SVBRDF) invariant theory, Wang TC et al. (2016b) proposed an LF depth estimation method for recovering 3D shapes from LFCs. Chen J et al. (2018) analyzed the influence of the initial label confidence map and edge strength weights, which increased the accuracy of depth estimation. Honauer et al. (2017) constructed a densely sampled 4D LF dataset, including highly accurate disparity ground truth, Heidelberg Collaboratory for image processing2 (HCI2). Williem et al.

(2018) presented two novel data costs, constrained adaptive defocusing cost (CAD) and constrained angle entropy cost (CAE), which were robust to occlusion and noise in depth estimation. Zhu et al. (2017) solved the problem about the occluder-consistency between the spatial and angle spaces. They also proposed a depth estimation algorithm that can reliably maintain the occlusion boundaries. Tian et al. (2017) proposed an algorithm to solve the problem of depth restoration in scattering media, such as water and fog. Mishiba (2020) introduced the offline viewpoint selection method and a fast weighted media filter into depth estimation, which greatly accelerated the depth estimation.

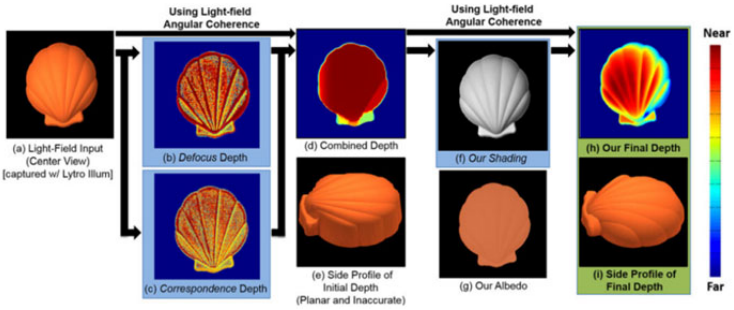

Fig. S1 The framework of depth estimation based on LF information (Tao et al., 2017)

1.2.2 Learning-based methods

In recent years, with the rapid development of deep learning technology, researchers have begun to apply deep learning technology to the LF depth estimation algorithm.

Heber and Pock (2016) trained a two-dimensional hyperplane CNN model with LF data, which can effectively predict the correspondence between the 4D LF data and the depth information it contains. Meanwhile, they set up a new dataset containing more synthetic LF information for experimental research. Based on some synthetic LF images and a CNN, Feng et al. (2018) used CNN to learn the correlation between a part of the pixels in the vertical and horizontal EPIs. Then, the correlation model between the learned pixels was used to estimate the depth of each image. Shin et al. (2018) considered the geometry of the LF polarities and used a combination of a multi-stream network and a merging network learning the spatial and angle information to obtain an estimated depth map. Jeon et al. (2019) proposed an auto pipeline, which can obtain the optimal depth label with high quality. At the pixel level, Zhou et al. (2019) used a CNN to learn the focal stack to obtain relevant depth cues for depth estimation. The final clues for depth estimation were mainly in two parts: depth semantic features and local structure information. Li Y et al. (2021) integrated the LF angular information and an attention module to propose a lightweight network, which can accurately complete the depth estimation of occluded edges.

Finally, to compare the results of the algorithms on the dataset (Honauer et al., 2017) in a comprehensive way, we provided two metrics, including mean square error (MSE) (×100) and BadPix (0.07). The quantitative comparison results are shown in Tables S3 and S4.

The MSE (×100) in depth estimation can be quantified as

$$\mathrm{MSE}_M = \frac{1}{|M|}\sum_{x\in M}\left(d(x)-\mathrm{gt}(x)\right)^2 \times 100. \tag{S4}$$

BadPix (0.07) is quantified as

$$\mathrm{BadPix}_M(t) = \frac{1}{|M|}\left|\left\{x \mid x\in M, \left|d(x)-\mathrm{gt}(x)\right| > t\right\}\right|. \tag{S5}$$

Herein d is the estimated disparity map, gt is the ground truth, and M is an evaluation mask.

Table S3 MSE performance evaluation on the synthetic images

Category	Image	MSE (×100)							
		Johannsen et al. (2016)	Jeon et al. (2015)	Sheng et al. (2018)	Zhang S et al. (2016)	Wanner and Goldluecke (2012a)	Shin et al. (2018)	Lee and Park (2017)	Jeon et al. (2019)
Stratified	Backgammon	9.56	13.01	21.59	4.59	20.75	–	5.67	6.89
Training	Dots	5.73	5.68	3.30	5.24	6.66	–	2.09	8.34
	Pyramids	0.03	0.27	0.10	0.04	0.02	–	0.03	0.04
	Stripes	2.67	17.45	8.13	6.96	6.10	–	1.32	1.38
	Boxes	8.72	17.43	9.85	9.11	12.54	–	8.90	12.31
	Cotton	2.25	9.17	1.07	1.31	4.51	–	0.76	1.16
	Dino	1.23	1.16	1.14	0.31	2.10	–	0.66	0.75
	Sideboard	2.85	5.07	2.30	1.02	3.86	–	1.16	2.85
	Bedroom	0.57	0.47	0.63	0.21	–	–	0.48	–
	Bicycle	8.52	11.73	7.67	5.57	–	–	5.81	–
	Herbs	24.70	21.34	22.20	11.23	–	–	15.46	–
	Origami	5.01	6.76	2.30	2.03	–	–	3.32	–
Average		5.99	9.13	6.69	3.97	7.07	1.75	3.80	4.22

Table S4 BadPix performance evaluation on the synthetic images

Category	Image	BadPix (0.07)						
		Johannsen et al. (2016)	Jeon et al. (2015)	Sheng et al. (2018)	Zhang S et al. (2016)	Shin et al. (2018)	Lee and Park (2017)	Jeon et al. (2019)
Stratified	Backgammon	21.33	5.52	19.01	3.78	3.50	10.31	7.14
Training	Dots	62.00	2.90	5.82	16.27	2.49	12.84	7.96
	Pyramids	0.86	12.35	3.17	0.86	0.16	0.55	0.11
	Stripes	25.81	35.74	18.41	14.99	2.46	19.12	2.96
	Boxes	24.45	23.02	26.52	15.89	12.30	27.15	34.09
	Cotton	13.93	7.83	6.22	2.59	0.45	5.01	2.43
	Dino	10.35	19.03	14.91	2.18	1.21	7.55	4.39
	Sideboard	18.38	21.99	18.50	9.30	4.46	13.92	13.25
	Bedroom	13.59	13.86	17.57	4.86	2.30	5.77	–
	Bicycle	25.21	19.79	21.56	10.91	9.61	19.66	–
	Herbs	47.08	18.11	36.83	8.26	11.00	14.75	–
	Origami	28.90	14.18	22.43	11.70	5.81	22.86	–
Average		24.32	16.19	17.58	8.47	4.65	13.29	9.04

2 Tasks and applications

2.1 Face recognition, detection, and light field face dataset

2.1.1 Texture-based methods

Kim SY et al. (2014) proposed to distinguish real and deceptive face samples by computing edge features and ray difference features. The main deceptive samples in their performance evaluation included printed photos and an IPAD. By exploiting the relationship between color and angle in LF images, Sepas-Moghaddam et al. (2018) proposed an effective descriptor, LF angular local binary patterns (LFALBP), for face detection relying on disparity exploitation. In addition, based on the acquired LF data, visual geometry group (VGG)-16 model, and long short-term memory (LSTM) network, Sepas-Moghaddam et al. (2020) presented an advanced face recognition framework, as shown in Fig. S2. First, they analyzed the LF data and selected some subaperture images. Second, they used VGG-16 to train the feature descriptors corresponding to each subaperture image. Finally, the trained feature descriptors were input to LSTM for classification.

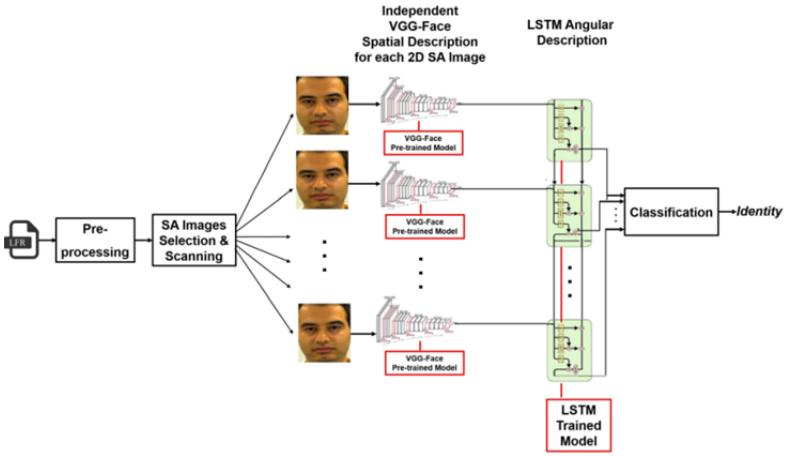

Fig. S2 Face recognition framework based on subaperture image analysis (Sepas-Moghaddam et al., 2020)

2.1.2 Focus and depth based methods

Raghavendra et al. (2013a) proposed a new scheme for face recognition by selecting the best focused face image from all focus images, and built a face recognition dataset using both an LFC and a traditional camera. Raghavendra et al. (2013b) effectively combined the super-resolution reconstruction algorithm with the discrete wavelet transform technology to improve the resolution of multiscale depth images, enhancing the face recognition ability in wide area scene monitoring. In addition, Zhao BQ et al. (2020) proposed an automatic face recognition framework algorithm that does not need extra manual management. At the same time, by analyzing the variation of focus between images rendered at different depths, they proposed a method based on posteriori refocusing to avoid an attack in face recognition. Similarly, to improve the overall performance, Raghavendra et al. (2016) explored the additional information available from different depth images in LFC using feature extraction and various super-resolution schemes, as shown in Fig. S3. Based on LF imaging technology and exploiting the available disparity information, Ji et al. (2016) presented a light field histogram of gradient (LFHoG) based method for face detection and achieved 99.75% accuracy. In addition, Xie et al. (2017) extracted three independent features from refocused images to correctly recognize face attacks. Similarly, Chiesa and Dugelay (2018) found useful information on the distance between the object and LFC to avoid face attacks, and thereby proposed a novel method for face detection by analyzing LF image properties. Sepas-Moghaddam et al. (2021) proposed a capsule network that can learn the angular part–whole relations of LF 2D subaperture images for face recognition in the wild.

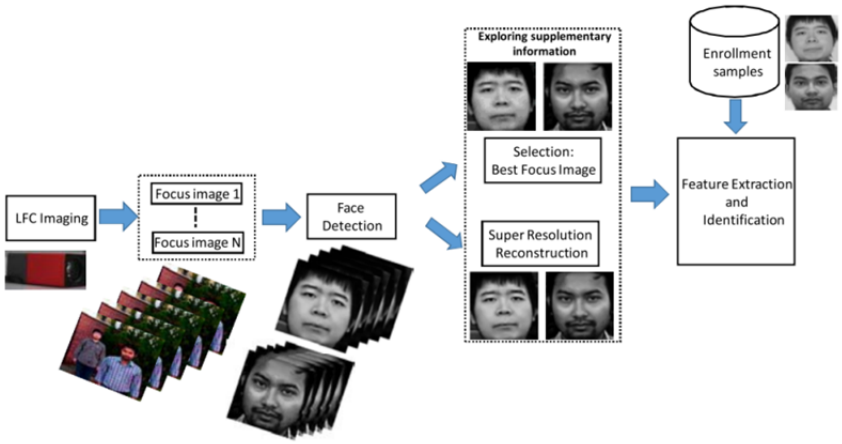

Fig. S3 The face recognition framework based on different depth images (Raghavendra et al., 2016)

In addition, several papers (Raghavendra et al., 2013a, 2015; Kim SY et al., 2014; Sepa-Moghaddam et al., 2018, 2020) reported a large number of LF face datasets, but the problem was that most of these data were not available or were used only for individual studies. Currently, the existing public benchmarks used mainly for face recognition include the Gjøvik University College (GUC) light field face artefact database (GUC-LiFFAD) (Raghavendra et al., 2015) and Instituto Superior Técnico (IST) lenslet light field face spoofing database (IST LLFFSD) (Sepas-Moghaddam et al., 2018).

GUC-LiFFAD was the first LF dataset for avoiding face attacks using Lytro. It consists mainly of two parts: normal face LF data and artefact face LF data. This dataset consists of a set of greyscale 2D images focused at different depths, notably exploiting a posteriori refocusing supported by LF imaging. Meanwhile, the samples in the GUC-LiFFAD dataset are widely representative, because there are mainly 80 relevant persons of different genders, ages, etc. However, the images in the dataset do not provide the original light field raw (LFR) information, and this dataset is limited for later research. IST LLFFSD is another widely used dataset for face recognition, including 50 subjects, 100 real images, and 600 face spoofing attack images. Specifically, it is composed mainly of three types of images: raw LF images, 2D rendered images, and depth maps. However, the difference between Raghavendra et al. (2015) and Sepas-Moghaddam et al. (2018) is that in the latter, LFC used a higher resolution, while instead of greyscale images, the work by Sepas-Moghaddam et al. (2018) is composed of 2D RGB rendered face images. In addition, the work of Sepas-Moghaddam et al. (2018) is rich in types (Raghavendra et al., 2015), and the comparison of the two datasets is shown in Table S5.

Table S5 Overview of publicly avalable LF-based face artefact datasets

Year	Number of subjects	Number of images	Type of content	Paper	Wrapped paper	Mobile	Tablet	Laptop	3D mask
2015	80	4826	2D rendered from LF	√	×	×	√	×	×
2017	50	700	LF+2D+depth	√	√	√	√	√	×

References

Alam MZ, Gunturk BK, 2018. Hybrid light field imaging for improved spatial resolution and depth range. *Mach Vis Appl*, 29(1):11-22. https://doi.org/10.1007/s00138-017-0862-2

Bishop TE, Favaro P, 2012. The light field camera: extended depth of field, aliasing, and superresolution. *IEEE Trans Patt Anal Mach Intell*, 34(5):972-986. https://doi.org/10.1109/TPAMI.2011.168

Bishop TE, Zanetti S, Favaro P, 2009. Light field superresolution. IEEE Int Conf on Computational Photography, p.1-9. https://doi.org/10.1109/ICCPHOT.2009.5559010

Boominathan V, Mitra K, Veeraraghavan A, 2014. Improving resolution and depth-of-field of light field cameras using a hybrid imaging system. IEEE Int Conf on Computational Photography, p.1-10. https://doi.org/10.1109/ICCPHOT.2014.6831814

Chen C, Lin HT, Zhan Y, et al., 2014. Light field stereo matching using bilateral statistics of surface cameras. IEEE Conf on Computer Vision and Pattern Recognition. https://doi.org/10.1109/CVPR.2014.197

Chen J, Hou J, Ni Y, Chau LP, 2018. Accurate light field depth estimation with superpixel regularization over partially occluded regions. *IEEE Trans Image Process*, 27(10):4889-4900. https://doi.org/10.1109/TIP.2018.2839524

Chiesa V, Dugelay JL, 2018. Advanced face presentation attack detection on light field database. IEEE Int Conf on Biometrics Special Interest Group, p.1-4. https://doi.org/10.23919/BIOSIG.2018.8553342

Diebold M, Jahne B, Gatto A, 2016. Heterogeneous light fields. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.1745-1753. https://doi.org/10.1109/CVPR.2016.193

Dong C, Loy CC, He KM, et al., 2014. Learning a deep convolutional network for image super-resolution. European Conf on Computer Vision, p.184-199. https://doi.org/10.1007/978-3-319-10593-2_13

Farrugia RA, Guillemot C, 2020. Light field super-resolution using a low-rank prior and deep convolutional neural network. *IEEE Trans Patt Anal Mach Intell*, 42(5):1162-1175. https://doi.org/10.1109/TPAMI.2019.2893666

Feng MT, Wang YN, Liu J, et al., 2018. Benchmark data set and method for depth estimation from light field images. *IEEE Trans Image Process*, 27(7):3586-3598. https://doi.org/10.1109/TIP.2018.2814217

Georgiev T, Zheng KC, Curless B, et al., 2006. Spatio-angular resolution tradeoffs in integral photography. Proc Eurographics Conf on Rendering Techniques, p.263-272.

Gul MSK, Gunturk BK, 2018. Spatial and angular resolution enhancement of light fields using convolutional neural networks. *IEEE Trans Image Process*, 27(5):2146-2159. https://doi.org/10.1109/TIP.2018.2794181

Heber S, Pock T, 2014. Shape from light field meets robust PCA. European Conf on Computer Vision, p.751-767. https://doi.org/10.1007/978-3-319-10599-4_48

Heber S, Pock T, 2016. Convolutional networks for shape from light field. IEEE Conf on Computer Vision and Pattern Recognition, p.3746-3754. https://doi.org/10.1109/CVPR.2016.407

Honauer K, Johannsen O, Kondermann D, et al., 2017. A dataset and evaluation methodology for depth estimation on 4D light fields. Proc Asian Conf on Computer Vision, 10113:19-34. https://doi.org/10.1007/978-3-319-54187-7_2

Jeon HG, Park J, Choe G, et al., 2015. Accurate depth map estimation from a lenslet light field camera. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.1547-1555. https://doi.org/10.1109/CVPR.2015.7298762

Jeon HG, Park J, Choe G, et al., 2019. Depth from a light field image with learning-based matching costs. *IEEE Trans Patt Anal Mach Intell*, 41(2):297-310. https://doi.org/10.1109/TPAMI.2018.2794979

Ji Z, Zhu H, Wang Q, et al., 2016. LFHOG: a discriminative descriptor for live face detection from light field image. IEEE Int Conf on Image Processing, p.1474-1478. https://doi.org/10.1109/ICIP.2016.7532603

Jiang XR, Pendu ML, Guillemot C, et al., 2018. Depth estimation with occlusion handling from a sparse set of light field views. 25th IEEE Int Conf on Image Processing, p.634-638. https://doi.org/10.1109/ICIP.2018.8451466

Johannsen O, Sulc A, Goldluecke B, 2016. What sparse light field coding reveals about scene structure. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.3262-3270. https://doi.org/10.1109/CVPR.2016.355

Kalantari NK, Wang TC, Ramamoorthi R, 2016. Learning-based view synthesis for light field cameras. *ACM Trans Graph*, 35(6):193. https://doi.org/10.1145/2980179.2980251

Kim C, Zimmer H, Pritch Y, et al., 2013. Scene reconstruction from high spatio-angular resolution light fields. *ACM Trans Graph*, 32(4):1-12. https://doi.org/10.1145/2461912.2461926

Kim JW, Lee JK, Lee KM, 2015a. Deeply-recursive convolutional network for image super-resolution. Proc IEEE Int Conf on Computer Vision and Pattern Recognition, p.1637-1645. https://doi.org/10.1109/CVPR.2016.181

Kim JW, Lee JK, Lee KM, 2015b. Accurate image super-resolution using very deep convolutional networks. Proc IEEE Int Conf on Computer Vision and Pattern Recognition, p.1646-1654. https://doi.org/10.1109/CVPR.2016.182

Kim SY, Ban YS, Lee SY, 2014. Face liveness detection using a light field camera. *Sensors*, 14(12):22471-22499. https://doi.org/10.3390/s141222471

Ko KS, Koh YJ, Chang SK, et al., 2021. Light field super-resolution via adaptive feature remixing. *IEEE Trans Image Process*, 30:4114-4128. https://doi.org/10.1109/TIP.2021.3069291

Krolla B, Diebold M, Goldlucke B, et al., 2014. Spherical light fields. Proc British Machine Vision Conf, p.1-12. https://doi.org/10.5244/C.28.67

Lee JY, Park RH, 2017. Depth estimation from light field by accumulating binary maps based on foreground-background separation. *IEEE J Sel Top Signal Process*, 11(7):955-964. https://doi.org/10.1109/JSTSP.2017.2747154

Li JQ, Lu ML, Li ZN, 2015. Continuous depth map reconstruction from light fields. *IEEE Trans Image Process*, 24(11): 3257-3265. https://doi.org/10.1109/TIP.2015.2440760

Li Y, Wang Q, Zhang L, et al., 2021. A lightweight depth estimation network for wide-baseline light fields. *IEEE Trans Image Process*, 30:2288-2300. https://doi.org/10.1109/TIP.2021.3051761

Lim B, Son SH, Kim HW, et al., 2017. Enhanced deep residual networks for single image super-resolution. IEEE Conf on Computer Vision and Pattern Recognition Workshops, p.136-144. https://doi.org/10.1109/CVPRW.2017.151

Lin H, Chen C, Kang SB, et al., 2015. Depth recovery from light field using focal stack symmetry. IEEE Int Conf on Computer Vision, p.3451-3459. https://doi.org/10.1109/ICCV.2015.394

Mishiba K, 2020. Fast depth estimation for light field cameras. *IEEE Trans Image Process*, 29:4232-4242. https://doi.org/10.1109/TIP.2020.2970814

Mitra K, Veeraraghavan A, 2012. Light field denoising, light field superresolution and stereo camera based refocussing using a GMM light field patch prior. IEEE Computer Society Conf on Computer Vision and Pattern Recognition Workshops, p.22-28. https://doi.org/10.1109/CVPRW.2012.6239346

Raghavendra R, Yang B, Raja KB, et al., 2013a. A new perspective—face recognition with light-field camera. IEEE Int Conf on Biometrics, p.1-8. https://doi.org/10.1109/ICB.2013.6612980

Raghavendra R, Raja KB, Yang B, et al., 2013b. Comparative evaluation of super-resolution techniques for multi-face recognition using light-field camera. 18th IEEE Int Conf on Digital Signal Processing, p.1-6. https://doi.org/10.1109/ICDSP.2013.6622829

Raghavendra R, Raja KB, Busch C, et al., 2015. Presentation attack detection for face recognition using light field camera. *IEEE Trans Image Process*, 24(3):1060-1075. https://doi.org/10.1109/TIP.2015.2395951

Raghavendra R, Raja KB, Busch C, et al., 2016. Exploring the usefulness of light field cameras for biometrics: an empirical study on face and iris recognition. *IEEE Trans Inform Forens Secur*, 11(5):922-936. https://doi.org/10.1109/TIFS.2015.2512559

Rossi M, Frossard P, 2018. Geometry-consistent light field super-resolution via graph-based regularization. *IEEE Trans Image*

Process, 27(9):4207-4218. https://doi.org/10.1109/TIP.2018.2828983

Sepas-Moghaddam A, Malhadas L, Correia PL, et al., 2018. Face spoofing detection using a light field imaging framework. *IET Biometr,* 7(1):39-48. https://doi.org/10.1049/iet-bmt.2017.0095

Sepas-Moghaddam A, Haque MA, Correia PL, et al., 2020. A double-deep spatio-angular learning framework for light field based face recognition. *IEEE Trans Circ Syst Video Technol*, 30(12):4496-4512. https://doi.org/10.1109/TCSVT.2019.2916669

Sepas-Moghaddam A, Etemad A, Pereira F, et al., 2021. CapsField: light field-based face and expression recognition in the wild using capsule routing. *IEEE Trans Image Process*, 30:2627-2642. https://doi.org/10.1109/TIP.2021.3054476

Schilling H, Diebold M, Rother C, et al., 2018. Trust your model: light field depth estimation with inline occlusion handling. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.4530-4538. https://doi.org/10.1109/CVPR.2018.00476

Sheng H, Zhao P, Zhang S, et al., 2018. Occlusion-aware depth estimation for light field using multi-orientation EPIs. *Patt Recogn*, 74:587-599. https://doi.org/10.1016/j.patcog.2017.09.010

Shi WZ, Caballero J, Huszár F, et al., 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.1874-1883. https://doi.org/10.1109/CVPR.2016.207

Shin CH, Jeon HG, Yoon YJ, et al., 2018. EPINET: a fully-convolutional neural network using epipolar geometry for depth from light field images. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.4748-4757. https://doi.org/10.1109/CVPR.2018.00499

Tao MW, Hadap S, Malik J, et al., 2013. Depth from combining defocus and correspondence using light-field cameras. Proc IEEE Int Conf on Computer Vision, p.673-680. https://doi.org/10.1109/ICCV.2013.89

Tao MW, Wang TC, Malik J, et al., 2015. Depth estimation for glossy surfaces with light-field cameras. European Conf on Computer Vision, 8926:533-547. https://doi.org/10.1007/978-3-319-16181-5_41

Tao MW, Su JC, Wang TC, et al., 2016. Depth estimation and specular removal for glossy surfaces using point and line consistency with light-field cameras. *IEEE Trans Patt Anal Mach Intell*, 38(6):1155-1169. https://doi.org/10.1109/TPAMI.2015.2477811

Tao MW, Srinivasan PP, Hadap S, et al., 2017. Shape estimation from shading, defocus, and correspondence using light-field angular coherence. *IEEE Trans Patt Anal Mach Intell*, 39(3):546-560. https://doi.org/10.1109/TPAMI.2016.2554121

Tian JD, Murez Z, Tong C, et al., 2017. Depth and image restoration from light field in a scattering medium. IEEE Int Conf on Computer Vision, p.2401-2410. https://doi.org/10.1109/ICCV.2017.263

Tosic I, Berkner K, 2014. Light field scale-depth space transform for dense depth estimation. IEEE Conf on Computer Vision and Pattern Recognition Workshops, p.435-442. https://doi.org/10.1109/CVPRW.2014.71

Wang TC, Efros AA, Ramamoorthi R, 2016a. Occlusion-aware depth estimation using light-field cameras. IEEE Int Conf on Computer Vision, p.3487-3495. https://doi.org/10.1109/ICCV.2015.398

Wang TC, Chandraker M, Efros AA, et al., 2016b. SVBRDF-invariant shape and reflectance estimation from light-field cameras. IEEE Conf on Computer Vision and Pattern Recognition, p.5451-5459. https://doi.org/10.1109/CVPR.2016.588

Wang X, Li L, Hou GQ, 2016. High-resolution light field reconstruction using a hybrid imaging system. *Appl Opt*, 55(10):2580-2593. https://doi.org/10.1364/AO.55.002580

Wang YL, Liu F, Zhang KB, et al., 2018. LFNet: a novel bidirectional recurrent convolutional neural network for light-field image super-resolution. *IEEE Trans Image Process*, 27(9):4274-4286. https://doi.org/10.1109/TIP.2018.2834819

Wang YQ, Wang LG, Yang JG, et al., 2020. Spatial-angular interaction for light field image super-resolution. Proc European Conf on Computer Vision, p.290-308. https://doi.org/10.1007/978-3-030-58592-1_18

Wang YQ, Yang JG, Wang LG, et al., 2021. Light field image super-resolution using deformable convolution. *IEEE Trans Image Process*, 30:1057-1071. https://doi.org/10.1109/TIP.2020.3042059

Wang YW, Liu YB, Heidrich W, et al., 2017. The light field attachment: turning a DSLR into a light field camera using a low budget camera ring. *IEEE Trans Visual Comput Graph*, 23(10):2357-2364. https://doi.org/10.1109/TVCG.2016.2628743

Wang ZW, Liu D, Yang JC, et al., 2015. Deep networks for image super-resolution with sparse prior. Proc IEEE Int Conf on Computer Vision, p.370-378. https://doi.org/10.1109/ICCV.2015.50

Wanner S, Goldluecke B, 2012a. Globally consistent depth labeling of 4D light fields. IEEE Conf on Computer Vision and Pattern Recognition, p.41-48. https://doi.org/10.1109/CVPR.2012.6247656

Wanner S, Goldluecke B, 2012b. Spatial and angular variational super-resolution of 4D light fields. European Conf on Computer Vision, p.608-621. https://doi.org/10.1007/978-3-642-33715-4_44

Wanner S, Goldluecke B, 2014. Variational light field analysis for disparity estimation and super-resolution. *IEEE Trans Patt Anal Mach Intell*, 36(3):606-619. https://doi.org/10.1109/TPAMI.2013.147

Williem W, Park IK, 2016. Robust light field depth estimation for noisy scene with occlusion. IEEE Conf on Computer Vision and Pattern Recognition, p.4396-4404. https://doi.org/10.1109/CVPR.2016.476

Williem W, Park IK, Lee KM, 2018. Robust light field depth estimation using occlusion-noise aware data costs. *IEEE Trans Patt Anal Mach Intell*, 40(10):2484-2497. https://doi.org/10.1109/TPAMI.2017.2746858

Wu GC, Zhao MD, Wang LY, et al., 2017. Light field reconstruction using deep convolutional network on EPI. Proc IEEE Conf on

Computer Vision and Pattern Recognition, p.6317-6327. https://doi.org/10.1109/CVPR.2017.178

Wu GC, Liu YB, Dai QH, et al., 2019. Learning sheared EPI structure for light field reconstruction. *IEEE Trans Image Process*, 28(7):3261-3273. https://doi.org/10.1109/TIP.2019.2895463

Wu JD, Wang HQ, Wang XZ, et al., 2015. A novel light field super-resolution framework based on hybrid imaging system. *Vis Commun Image Process*, p.1-4. https://doi.org/10.1109/VCIP.2015.7457904

Xie XH, Gao Y, Zheng WS, et al., 2017. One-snapshot face anti-spoofing using a light field camera. Chinese Conf on Biometric Recognition, 10568:108-117. https://doi.org/10.1007/978-3-319-69923-3_12

Yeung HWF, Ho JH, Chen XM, et al., 2019. Light field spatial super-resolution using deep efficient spatial-angular separable convolution. *IEEE Trans Image Process*, 28(5):2319-2330. https://doi.org/10.1109/TIP.2018.2885236

Yoon YJ, Jeon HG, Yoo DG, et al., 2015. Learning a deep convolutional network for light-field image super-resolution. Proc IEEE Int Conf on Computer Vision Workshop, p.24-32. https://doi.org/10.1109/ICCVW.2015.17

Yoon YJ, Jeon HG, Yoo DG, et al., 2017. Light-field image super-resolution using convolutional neural network. *IEEE Signal Process Lett*, 24(6):848-852. https://doi.org/10.1109/LSP.2017.2669333

Yuan Y, Cao ZQ, Su LJ, 2018. Light-field image superresolution using a combined deep CNN based on EPI. *IEEE Signal Process Lett*, 25(9):1359-1363. https://doi.org/10.1109/LSP.2018.2856619

Zhang FL, Wang J, Shechtman E, et al., 2017. PlenoPatch: patch-based plenoptic image manipulation. *IEEE Trans Vis Comput Graph*, 23(5):1561-1573. https://doi.org/10.1109/TVCG.2016.2532329

Zhang S, Sheng H, Li C, et al., 2016. Robust depth estimation for light field via spinning parallelogram operator. *Comput Vis Image Underst*, 145:148-159. https://doi.org/10.1016/j.cviu.2015.12.007

Zhang S, Lin YF, Sheng H, 2019. Residual networks for light field image super-resolution. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.11046-11055. https://doi.org/10.1109/CVPR.2019.01130

Zhang S, Chang S, Lin YF, 2021. End-to-end light field spatial super-resolution network using multiple epipolar geometry. *IEEE Trans Image Process*, 30:5956-5968. https://doi.org/10.1109/TIP.2021.3079805

Zhang YB, Lv HJ, Liu YB, et al., 2017. Light field depth estimation via epipolar plane image analysis and locally linear embedding. *IEEE Trans Circ Syst Video Technol*, 27(4):739-747. https://doi.org/10.1109/TCSVT.2016.2555778

Zhang ZT, Liu YB, Dai QH, 2015. Light field from micro-baseline image pair. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.3800-3809. https://doi.org/10.1109/CVPR.2015.7299004

Zhao BQ, Sun JM, Zheng Y, et al., 2020. Light field camera based automatic algorithm for optimal face search. 11th Int Conf on Graphics Image Processing, 11373:113730D. https://doi.org/10.1117/12.2557249

Zhao MD, Wu GC, Li YP, et al., 2018. Cross-scale reference-based light field super-resolution. *IEEE Trans Comput Imag*, 4(3):406-418. https://doi.org/10.1109/TCI.2018.2838457

Zheng HT, Guo MH, Wang HQ, et al., 2017. Combining exemplar-based approach and learning-based approach for light field super-resolution using a hybrid imaging system. Proc IEEE Int Conf on Computer Vision Workshop, p.2481-2486. https://doi.org/10.1109/ICCVW.2017.292

Zhou WH, Zhou EC, Yan YX, et al., 2019. Learning depth cues from focal stack for light field depth estimation. IEEE Int Conf on Image Processing, p.1074-1078. https://doi.org/10.1109/ICIP.2019.8804270

Zhu H, Wang Q, Yu J, 2017. Occlusion-model guided anti-occlusion depth estimation in light field. *IEEE J Sel Top Signal Process*, 11(7):965-978. https://doi.org/10.1109/JSTSP.2017.2730818