Signal Processing: Image Communication 80 (2020) 115638

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

A simple framework to leverage state-of-the-art single-image super-resolution methods to restore light fields

Reuben A. Farrugia [a,*], Christine Guillemot [b]

[a] Department of Communications and Computer Engineering, University of Malta, Malta
[b] Institut National de Recherche en Informatique et en Automatique, Rennes 35042, France

ARTICLE INFO

Keywords:
Light fields
Super-resolution
Convolutional neural networks
Single image super resolution

ABSTRACT

This paper describes a simple framework allowing us to leverage state-of-the-art single image super-resolution (SISR) techniques into light fields, while taking into account specific light field geometrical constraints. The idea is to first compute a representation compacting most of the light field energy into as few components as possible. This is achieved by aligning the light field using optical flow and then by decomposing the aligned light field using singular value decomposition (SVD). The *principal basis* captures the information that is coherent across all the views, while the other basis contain the high angular frequencies. Super-resolving this *principal basis* using an SISR method allows us to super-resolve all the information that is coherent across the entire light field. In this paper, to demonstrate the effectiveness of the approach, we have used the very deep super resolution (VDSR) method, which is one of the leading SISR algorithms, to restore the *principal basis*. The information restored in the *principal basis* is then propagated to restore all the other views using the computed optical flow. This framework allows the proposed light field super-resolution method to inherit the benefits of the SISR method used. Experimental results show that the proposed method is competitive, and most of the time superior, to recent light field super-resolution methods in terms of both PSNR and SSIM quality metrics, with a lower complexity. Moreover, the subjective results demonstrate that our method manages to restore sharper light fields which enables to generate refocused images of higher quality.

1. Introduction

Light field imaging has recently emerged as a promising technology able to discriminate and capture light rays along different directions [1,2]. This rich visual description of the scene enables the creation of immersive experience in AR/VR applications and facilitates the integration of computer-generated graphics for post-production editing. Together with proper computational algorithms, this technology is expected to impact the field of digital photography, by enabling post-capture re-focusing, depth of field extension, or 3D scene model estimation.

However, light field imaging systems trade-off spatial resolution with angular information in the light field. Rigs of cameras capture views with a high spatial resolution but in general with limited angular sampling to reduce costs [2]. On the other hand, plenoptic cameras use an array of microlenses placed in front of the sensor to capture multiple low-resolution (LR) views in one 2D sensor image [1]. This is a way to cost-effectively capture multiple views with a high angular sampling, but at the expense of reducing the spatial resolution by orders of magnitude compared to the raw sensor image.

To tackle this problem, various methods have been developed, which are aimed to achieve better spatial and angular resolution trade-off from a plenoptic camera. These methods go from the use of coded aperture techniques, using e.g. a programmable non-refractive mask placed at the aperture as in [3], or optically coded projections as in [4], to light field super-resolution methods [5–12]. While research in light field super-resolution is at its infancy, research in the related field of single-image super-resolution (SISR) is quite mature with methods based on very deep convolutional neural networks achieving state-of-the-art performances [13–17].

This paper presents a framework which allows to leverage state-of-the-art 2D image super-resolution techniques to light field super-resolution. The energy of the light field is first compacted, to capture the coherent information, and to then apply 2D single-image super-resolution to restore the whole light field. To do so, we first align each view to the centre view using optical flow. The alignment plays an important role since it removes disparities between the views. These disparities can be easily recovered by inverting the alignment process. This aligned light field is then decomposed using singular value decomposition (SVD) where the eigenvectors of the SVD capture dominant

* Corresponding author.
E-mail address: reuben.farrugia@um.edu.mt (R.A. Farrugia).

https://doi.org/10.1016/j.image.2019.115638
Received 10 June 2019; Received in revised form 11 September 2019; Accepted 11 September 2019
Available online 25 September 2019
0923-5965/© 2019 Elsevier B.V. All rights reserved.

variations (or eigenimages) of the different views. It will be shown in Section 3 that, thanks to the alignment of the light field views, the *first principal basis* (*a.k.a.* dominant eigenvector) captures most of the angularly coherent spatial information, while the remaining basis contain only small angular variations, i.e. angular high frequencies. Based on the observation that the *first principal basis* is a natural image, its spatial details can be restored using any state-of-the-art SISR algorithm. In the experiments, we first consider three recent deep learning based SISR methods, namely SRCNN [18], VDSR [13] and Lab402 [17]. The VDSR method showing superior performances has then been retained to show the effectiveness of the proposed light field super-resolution framework. The spatial details restored in the *first principal basis* are then propagated to all the other views in a consistent manner, using inverse warping.

We have evaluated the efficiency of the proposed framework, considering input low resolution light fields in which each view is a downscaled version of the corresponding high resolution view, using a bicubic downscaling as most SISR methods and as recommended in track 1 of the NITRE 2017 challenge on SISR [19].

The results in Section 5 show that the proposed method achieves sharper light field images compared with existing light field super-resolution methods, namely the convolutional neural network based method in [11], the linear subspace projection based method (BM-PCARR) in [9] and the Graph-based light field super resolution method in [12]. Supplementary material attached to this paper also shows that the restored light fields are angularly coherent and that it is able to restore real-world plenoptic light fields. It is also shown that the method manages to restore light fields containing non-Lambertian surfaces.[1]

The main contributions of this paper are as follows:

- We present a framework that enables to leverage SISR methods to restore the *principal basis* capturing the coherent information across the entire light field.
- Based on this framework, we describe a light field super-resolution method that yield sharper light field images with results superior than existing methods for applications such as digital refocusing.
- The proposed framework allows to inherit the benefits of the SISR methods employed to restore the *principal basis* and we are therefore presenting the first light field super-resolution algorithm which uses only one model to cater for different magnification factors.

The remainder of this paper is organized as follows. Work related to the method described in this paper is provided in Section 2 while the light field energy compaction method is explained in Section 3. The proposed *principal basis* VDSR (PB-VDSR) is described in Section 4 while the experimental results are delivered in the following section. Section 6 concludes with the final remarks.

2. Related work

This section gives a brief overview of work related to the key concepts of the proposed spatial light field super-resolution approach and the light field super-resolution methods that are found in literature.

2.1. Single image super-resolution

Single-image super-resolution is an ill-posed inverse problem with infinite possible solutions. These methods use priors to derive a more plausible solution that satisfies a predefined assumption. These priors are either hand-crafted, such as total variation or Bayesian models, or data driven that are learned using machine learning methods. Pixel-based methods have been proposed in [20,21] where each pixel in the high-resolution (HR) image is inferred via statistical learning. To improve spatial coherency, patch-based approaches, referred to as example-based methods, have been proposed. Freeman et al. [22] presented the first single-image example-based super-resolution algorithm that used a coupled dictionary to learn a mapping between LR and HR patches. More advanced methods based on manifold learning [23–25] and sparse coding [26,27] were investigated to regularize the problem and were found to provide sharper images. Other approaches [28–31] utilized image self-similarities to avoid using dictionaries constructed using external images.

Deep neural networks have contributed to a drastic improvement in the field of single-image super-resolution. Dong et al. [18] were the first to use a rather shallow convolutional neural network (SRCNN). Residual learning was introduced in [13,14,17] for training deeper network architectures and achieved state-of-the-art performance. The authors in [15] pose the general image restoration problem with encoder–decoder networks and systematic skip connections. This architecture was later on extended in [16] where the authors expanded the model size and removed unnecessary modules in the convolutional residual networks.

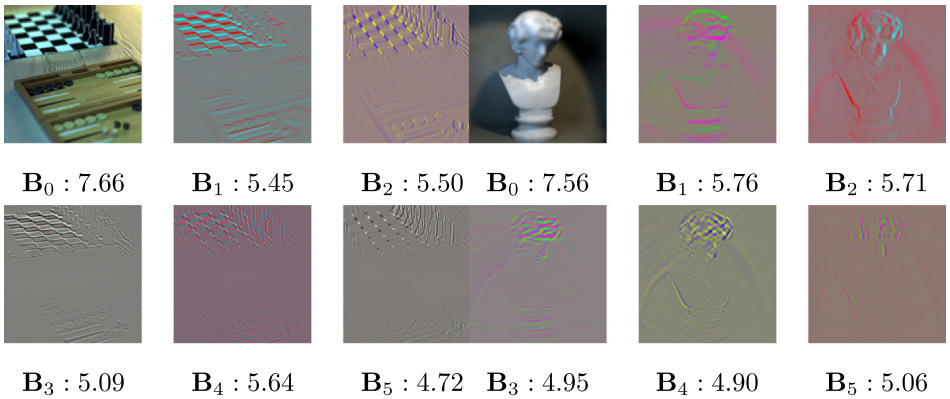
2.2. Light field super-resolution

Early light field super-resolution approaches pose the problem as one of recovering the high-resolution views from multiple low-resolution images with unknown non-integer translation misalignment. The authors in [5,6] proposed a two-step approach where they first estimate a depth map and then formulate the super-resolution problem either as a simple linear problem [5] or as a Bayesian inference problem [6] assuming an image formation model with Lambertian reflectance priors and depth-dependent blurring kernels. A patch-based technique was proposed in [7] where high-resolution 4D patches are estimated using a linear minimum mean square error (LMMSE) estimator assuming a disparity-dependent Gaussian Mixture Model (GMM) for the patch structure. A variational optimization framework was proposed in [8] to spatially super-resolve the light field given their estimated depth maps and to increase the angular resolution.

Example-based light field super-resolution methods have been recently proposed. These methods use machine learning to learn a mapping between low- and high-resolution light fields. In [9], the authors show that a 3D patch-volume resides on a low-dimensional subspace and propose to learn a projection between low- and high-resolution subspaces of patch-volumes using ridge-regression. Deep learning techniques for light field super-resolution have been first proposed in [10] where 4-tuples of neighbouring views are stacked into groups and restored using SRCNN [18]. The spatially restored light field is then fed into a second CNN that up-scales the angular resolution. The same authors have later proposed to restore each view independently using SRCNN in [11] showing superior performance over their original method. More recently, graph based light field super-resolution algorithm was presented in [12] that enforces the optimization to preserve the light field structure. A shallow neural network was proposed in [32] to restore light fields captured by a plenoptic camera. However, this method is only suitable to achieve a magnification factor of $\times 2$ and needs to train a CNN for every angular view. Very recently, a multi-scale fusion scheme was used to accumulate contextual information from multiple scales while Recurrent Convolutional Neural Networks (BRCNN) are used to model the spatial relation between adjacent views and restore the light field [33].

A hybrid light field super-resolution method is proposed in [34] and [35] where a high-resolution camera is coupled with a plenoptic camera. The authors in [36] describe an acquisition device formed by eight low-resolution side cameras arranged around a central high-quality camera. Iterative patch- and depth-based synthesis (iPADS) is then used to reconstruct a light field with the spatial resolution of the SLR camera and an increased number of views.

[1] Supplementary material can be viewed at https://youtu.be/HHmUZSP7HU4.



(a) First six-basis computed using SVD.

(b) First six-basis computed using SVD of the aligned light field.

Fig. 1. Comparing the information contained in each basis when computing (a) SVD decomposition and (b) Aligned SVD decomposition for the (left) Boardgames and (right) Antinous synthetic light fields. The entropy measure for each basis is given below the corresponding image.

While the methods in [10,11,32,33] use deep learning to super-resolve the light field, our method is considerably different. The novelty of our approach is that the proposed framework allows to use SISR techniques for light field super-resolution. The deep learning SR method used are not retrained on light fields and use models that are trained on natural images. Moreover, our framework inherits the benefits of the SISR algorithm used. In our study, we used VDSR [13] which allows us to use a very deep super-resolution method that adopts one single model to cater for different magnification factors.

2.3. Light field edit propagation

Light field edit propagation involves the restoration of the centre view followed by the propagation of the restored information to all the other views. The authors in [37] described an approach using a 3D voxel-based model of the scene with an associated radiance function to propagate pixel edits and illumination changes in a consistent manner from one view to the other views of the light field. The authors in [38] extend the 2D image stroke-based edit propagation method of [39] to light fields, where they reduce the complexity by propagating the edits in a downscaled version of the light field. In [40], a method based on a reparameterization of the light field is proposed to better preserve coherence of the edits in the angular domain. However, these methods deal with simple stroke-based editing and are not suitable to propagate complex edits, such as inpainting or super-resolution, to all the other views.

A patch-based depth-layer-aware image synthesis algorithm was adopted in [41] to propagate the edits from the centre view to all the other views. The authors in [42] use tensor driven diffusion to propagate information from the centre view along the Epipolar Plane Image (EPI) structure of the light field. These methods were used to propagate either simple edits, recolorization or inpainting from the centre view to all the other views. However, up to the knowledge of the authors, such approaches were never been considered for light field super-resolution.

3. Light field energy compaction

Lets consider an input light field $I(x, y, s, t)$ represented with the two plane parametrization proposed in [43,44], where (x, y) and (s, t) represent spatial and angular coordinates respectively. The light field can be seen as a 2D array of images, where each image $\mathbf{I}_{s,t}$ captures the scene from a viewpoint defined by angular coordinates (s, t). One can use single image super-resolution technique to restore every angular view independently. However, these methods do not exploit the geometrical structure of the light field [45] and are not guaranteed to provide angularly coherent solutions [9]. On the other hand, several light field super-resolution techniques have been proposed that either exploit the disparity/depth information [5–8] or else use learning based methods [9–11,32,33] to improve the quality of the light field. However, these algorithms do not benefit from the recent advances in single image super-resolution where very deep Convolutional Neural Networks are achieving outstanding performances [13–16,18].

A light field consists of a very large volume of high-dimensional data. Nevertheless, it exhibits redundancies in all four dimensions since every view captures the same scene from a slightly different viewpoint. Early work in the field of light field compression used 3D/4D wavelet transforms to decompose the light field into a number of sub-bands [51–53], where each sub-band gives information at different spatial and angular frequencies. Fig. 1(a) shows the first six orthogonal

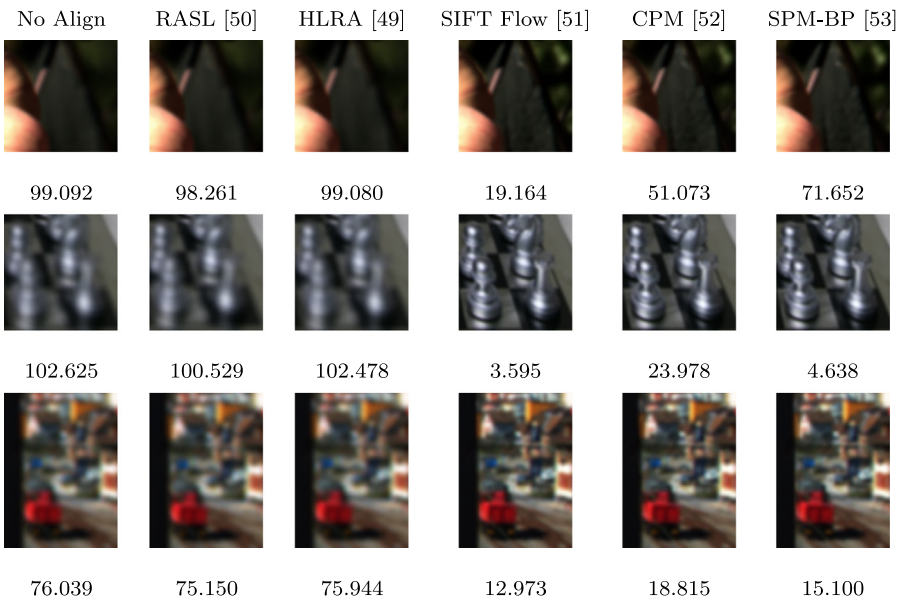


Fig. 2. Cropped regions of the mean view when using different disparity compensation methods. Underneath each image we provide the average variance across the n angular views which was used in [9] to characterize the performance of the alignment algorithm, where smaller values indicate better alignment (see Refs. [46–50]).

basis when decomposing the light field using SVD. It can be seen that while most of the energy resides in the *principal basis* $\mathbf{B}_0$, there is still a lot of high frequency detail in the other basis. Moreover, the *principal basis* $\mathbf{B}_0$, which captures the average energy in the scene is blurred. This is attributed to variations in disparities across the views which result in high-frequency angular details that are not captured by the *principal basis*.

The authors in [47] tried to reduce the energy within the high-frequency basis by jointly aligning the angular views and estimating a low-rank approximation (LRA) of the light field. This approach has shown very promising results in the field of light field compression. In the same spirit, the RASL algorithm [46] was used to find the homographies that globally align a batch of linearly correlated images. Both methods find an optimal set of homographies such that the matrix of aligned images can be decomposed in a low–rank matrix of aligned images, with the latter constraining the error matrix to be sparse. However, as it can be seen in Fig. 2, while both RASL and HLRA methods manage to globally align the angular views, the resulting mean view, that is computed by averaging all the views, are still blurred indicating that the views are not well aligned.

In the sequel, we consider $\mathbf{I}_{s,t}$ to represent different views, where (s,t) define the angular coordinates. This notation will be further simplified as $\mathbf{I}_i$ with a bijection between (s,t) and i. The complete light field can hence be represented by a matrix $\mathbf{I} \in \mathbb{R}^{m,n}$:

$$\mathbf{I} = [vec(\mathbf{I}_1) \quad | \quad vec(\mathbf{I}_2) \quad | \quad \cdots \quad | \quad vec(\mathbf{I}_n)] \tag{1}$$

with $vec(\mathbf{I}_i)$ being the vectorized representation of the ith angular view, m represents the number of pixels in each view ($m = X \times Y$) and n is the number of views in the light field ($n = P \times Q$), where P and Q represent the number of vertical and horizontal angular views respectively. We then formulate the light field decomposition problem as that of finding a set of orthogonal basis $\mathbf{B}$ that is able to capture most of the information contained in the light field. This can be achieved by minimizing the following optimization problem

$$\min_{\mathbf{u},\mathbf{v},\mathbf{B},\mathbf{C}} \|\Gamma_{\mathbf{u},\mathbf{v}}(\mathbf{I}) - \mathbf{BC}\|_2^2 \tag{2}$$

where $\mathbf{u} \in \mathbb{R}^{m,n}$ and $\mathbf{v} \in \mathbb{R}^{m,n}$ are flow vectors that specify the displacement of each pixel needed to align each view with the centre view, $\mathbf{B} \in \mathbb{R}^{m,n}$ represents the basis matrix, $\mathbf{C} \in \mathbb{R}^{n,n}$ is the combination weight matrix and $\Gamma_{\mathbf{u},\mathbf{v}}(\cdot)$ is the forward warping process that projects pixels from the source image onto the target image.

This optimization problem is computationally intractable. Instead, we decompose this problem in two sub-problems: (i) use an optical flow estimation technique to find the flow vectors $\mathbf{u}$ and $\mathbf{v}$ that best align each view with the centre view and (ii) decompose the aligned light field into a set of basis $\mathbf{B}$ and coefficient matrix $\mathbf{C}$ using SVD. The results in Fig. 2 clearly show that the mean views are much sharper when aligning the light field using optical flow. Moreover, optical flow significantly reduce the variance across the angular views, with the SIFT flow method [48] achieving the best performance. It reduces the mean variance across views by a factor of nine, and thus we will use it to align the views. Reducing the total variance across the views (as shown in Fig. 2) allows to compact more information in the low-frequency basis.

The solution of the first sub-problem gives the flow-vectors $\mathbf{u}$ and $\mathbf{v}$ which are used to align the light field using forward warping *i.e.* $\tilde{\mathbf{I}} = \Gamma_{\mathbf{u},\mathbf{v}}(\mathbf{I})$. The aligned light field $\tilde{\mathbf{I}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$ is then decomposed using SVD, where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices and $\boldsymbol{\Sigma}$ is a diagonal matrix containing the singular values. The basis matrix is computed as $\mathbf{B} = \mathbf{U}\boldsymbol{\Sigma}$ while the coefficient matrix is given by $\mathbf{C} = \mathbf{V}^T$.

Fig. 1(b) shows the first six orthogonal basis using our proposed light field decomposition method. It can be seen that the *principal basis* $\mathbf{B}_0$ is much sharper indicating that it captures more information from the light field. Moreover, the energy in the higher-frequency basis is significantly reduced as indicated by the significant drop in entropy when using our proposed light field decomposition method. Thanks to the alignment, the first principal basis captures angularly coherent spatial information, while the remaining basis mainly capture higher angular frequencies. In addition, while the first principal basis of the low and high resolution light fields are highly correlated, the correlation drops for higher angular frequencies.

The light field can then be easily reconstructed using $\tilde{\mathbf{I}} = \mathbf{BC}$ without losing any information since $\mathbf{B}$ is orthogonal and full-rank. In the sequel, the decomposition of the aligned light field will be referred to as A-SVD.

4. Principal basis super-resolution

Let $\mathbf{I}^H$ and $\mathbf{I}^L$ denote the high- and low-resolution light fields. The super-resolution problem can be formulated in Banach space as

$$\mathbf{I}^L = \downarrow_\alpha \mathbf{B}\mathbf{I}^H + \boldsymbol{\eta} \tag{3}$$

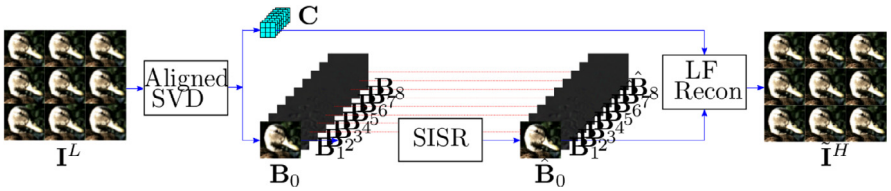


Fig. 3. The proposed light field super-resolution algorithm that takes a 3 × 3 matrix of low-resolution views as input, denoted by $\mathbf{I}^L$, to estimate the high-resolution light field $\mathbf{I}^H$.

where $\boldsymbol{\eta}$ is an additive noise matrix, $\mathbf{B}$ is the blurring kernel and $\downarrow_\alpha$ is a downsampling operator applied on each angular view with a scale-factor α. To show the effectiveness of the proposed light field super-resolution framework, we consider three recent deep learning based SISR methods, namely SRCNN [18], VDSR [13] and Lab402 [17], to restore the principal basis. These methods were trained to restore input low resolution light fields in which each view is a downscaled version of the corresponding high resolution view, using a bicubic downscaling as most SISR methods and as recommended in track 1 of the NITRE 2017 challenge on SISR [19]. The noise is assumed to be negligible. We therefore assume a similar degradation process as the one considered in these papers.

Fig. 3 shows the block diagram of the proposed light field super-resolution algorithm where, for simplicity, a 3 × 3 matrix of angular views is shown. The A-SVD algorithm, described in Section 3, is applied on the low-resolution light field $\mathbf{I}^L$ (that is upscaled using bicubic interpolation so that its resolution matches that of $\mathbf{I}^H$) to decompose the light field into a set of orthogonal basis $\mathbf{B} \in \mathbb{R}^{m,n}$ and coefficient matrix $\mathbf{C} \in \mathbb{R}^{n,n}$. As shown in more detail in Section 3, the A-SVD algorithm is able to capture more information in the *principal basis vector* $\mathbf{B}_0$.

Driven by the observation that the *principal basis vector* $\mathbf{B}_0$ is a natural image that captures most of the information in the light field, we pose the problem of light field super-resolution as that of restoring the resolution of the *principal basis*. The higher order basis $\mathbf{B}_j$, $j \in [1, n-1]$, that capture the discrepancies across the views in terms of occlusions and illumination, are not modified. Any single SISR method can be used to restore the low resolution *principal basis vector* $\mathbf{B}_0$ and to estimate the high resolution *principal basis vector* $\hat{\mathbf{B}}_0$. The high frequency basis are simply approximated using $\hat{\mathbf{B}}_j = \mathbf{B}_j$ for $j \in [1, n-1]$. The restored aligned light field is then reconstructed using a simple matrix multiplication which is then inverse-warped to restore the original disparities *i.e.* $\tilde{\mathbf{I}}^H = \Gamma^{-1}_{\mathbf{u},\mathbf{v}}(\hat{\mathbf{B}}\mathbf{C})$, where $\Gamma^{-1}_{\mathbf{u},\mathbf{v}}(\cdot)$ stands for the inverse warping process that maps pixels from the target image to the source image. While the theoretical and implementation details of A-SVD were provided in Section 3, the following sub-sections will deal with the implementation detail of the SISR and LF Reconstruction modules.

4.1. SISR module

In this work we consider some of the most promising SISR methods found in literature to restore the *principal basis* and their performance is summarized in Table 1. In essence we consider the first deep-learning based super-resolution method SRCNN [18], the very deep convolutional neural network (VDSR) which uses residual learning with 20 convolutional layers [13] and the Lab402 method which was ranked third in the recent NTIRE workshop challenge. The network models of these methods were not retrained on light field data and therefore this experiment evaluates the generalization abilities of these methods. These results demonstrate that while both VDSR and Lab402 manage to outperform SRCNN, the VDSR method is able to achieve the best performance in terms of both PSNR and SSIM quality measures. This indicates that while other methods can be used to restore the *principal basis*, the VDSR algorithm achieves the best performance and will therefore be considered in the experimental results in Section 5. Given that our method uses VDSR to restore the *principal basis* we named our method PB-VDSR. It is important to mention here that unlike SRCNN, VDSR uses a single network model to cater for different magnification factors and PB-VDSR inherits this property. Moreover, it must be mentioned that while VDSR will be used in the remaining part of the paper, this approach can use deep feature based methods [54,55] that optimize perceptual metric instead of PSNR.

Table 1
Quality analysis (PSNR with SSIM in parenthesis) using different single-image super-resolution algorithms to restore the *principal basis* $\mathbf{B}_0$ at a magnification factor ×3.

Light Field	SRCNN [18]	VDSR [13]	Lab402 [17]
Antinous	33.32 (0.954)	**35.74** (**0.978**)	33.81 (0.977)
Boardgames	23.68 (0.835)	**24.65** (**0.865**)	23.92 (0.859)
Greek	30.78 (0.935)	**33.55** (**0.966**)	31.70 (0.961)
Medieval 2	30.32 (0.952)	**32.10** (**0.962**)	31.74 (**0.962**)
Origami	25.32 (0.951)	28.89 (**0.973**)	**28.97** (**0.973**)
Books	29.73 (0.966)	**30.78** (0.974)	29.86 (0.970))
Friends 2	29.31 (0.935)	**31.13** (**0.944**)	30.79 (**0.944**)
Game Board	31.75 (0.972)	**32.12** (**0.976**)	31.54 (0.974)
Graffiti	28.56 (0.870)	29.90 (0.880)	**29.91** (**0.883**)
Parc du Luxembourg	28.49 (0.926)	**29.08** (**0.935**)	28.39 (0.928)

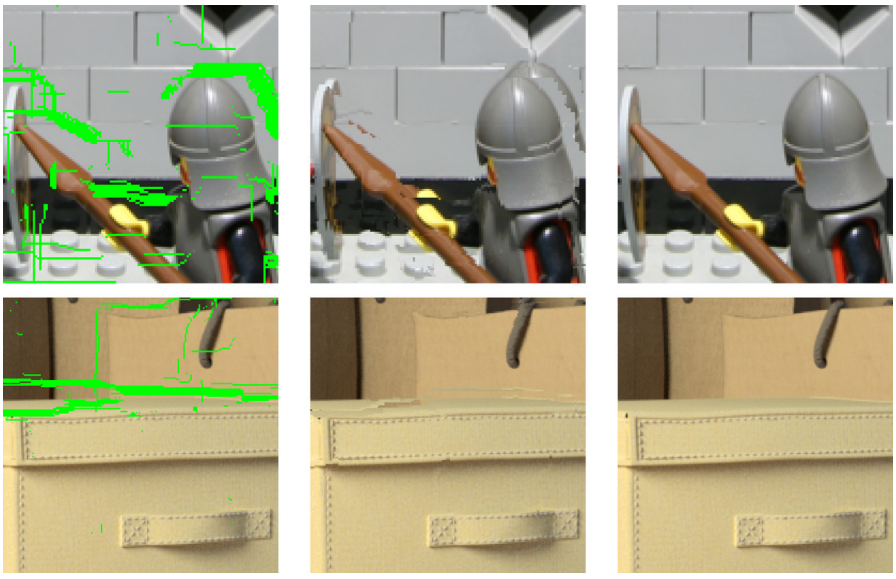
4.2. Light field reconstruction module

The aligned high resolution light field can be estimated by multiplying the restored basis $\hat{\mathbf{B}}$ and weight matrix $\mathbf{C}$ *i.e.* $\hat{\mathbf{I}}^H = \hat{\mathbf{B}}\mathbf{C}$. The views of $\hat{\mathbf{I}}^H$ are aligned with the centre view. Forward warping can be used to recover the original disparities of the restored views. However, as can be seen in the first column of Fig. 4, forward warping is not able to restore all pixels and results in a number of cracks and holes. Another approach is to use inverse warping and use neighbouring pixels to estimate the missing information. However, as can be seen in the second column of Fig. 4, missing pixels due to occlusion are not well correlated with the neighbouring pixels and result in inaccurate estimates.

In this work we observe that the pixels warped using forward warping are very accurate. Instead of interpolating the missing pixels, in this work we simply copy the collocated pixels (same spatial x and y coordinates) from the low-resolution light field to replace the missing pixels. The light field reconstructed in this work is depicted as the third column of Fig. 4 where it can be seen that the recovered pixels are more accurate than those estimated using inverse warping.

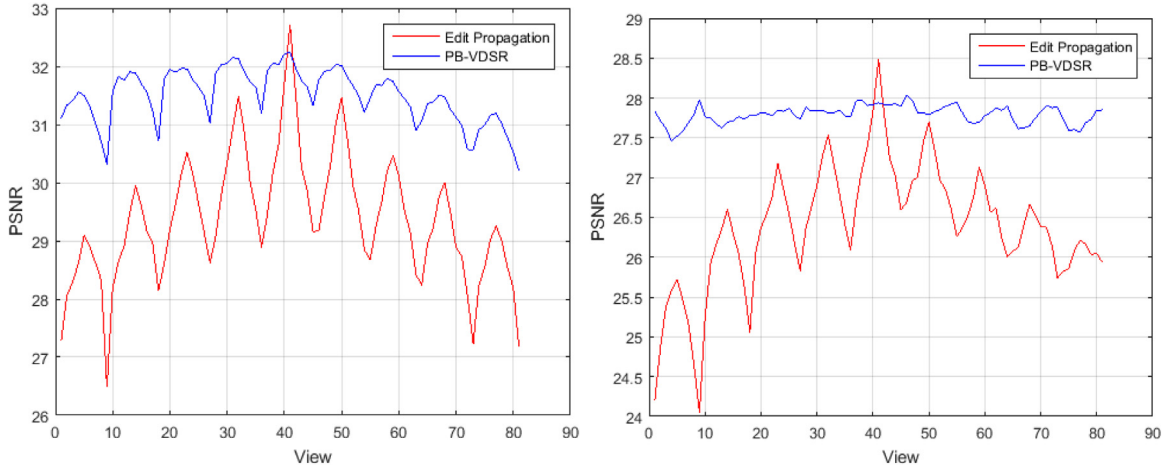
4.3. Edit propagation methods

This work is related to the Light Field Edit Propagation methods described in Section 2.3 which allow the user to edit the centre view and propagate the edits to all the other views. Fig. 5 illustrates the PSNR measure at each view and compares the proposed method, which we call PB-VDSR, against the edit propagation method that will be described next. The edit propagation considered applies VDSR to restore the centre view and then propagate the information to the other views using forward warping. The missing pixels due to occlusions are estimated using collocated pixels from the low-resolution light field as described in the previous subsection.



Forward Warping Inverse Warping Proposed Inpainting

Fig. 4. Inpainting the cracks marked in green.



Bee 2 (INRIA) Bikes (EPFL)

Fig. 5. PSNR analysis at each view comparing the proposed PB-VDSR to an edit propagation method. The view index indicates the index of the view when scanned using raster scan ordering.

It can be seen that edit propagation achieves larger PSNR for the centre view (view 41). However its performance degrades significantly when propagating the information to all the other views. This can be explained since the edit propagation ignores the variations across the views caused by illumination and occlusions. On the other hand, PB-VDSR restores the *principal basis* that captures the angular consistent information in the light field while the variations caused by illumination and occlusion are preserved in the higher frequency basis. This implies that PB-VDSR propagates the high angular frequency information in $\hat{\mathbf{I}}^H$. While we use the VDSR model trained on natural images, the performance might be improved by retraining the model on principal basis components of light fields. However, the goal here was to leverage models learned for SISR. However, retraining the SISR is not in scope of this paper since the objective here is to show that SISR can be extended using our framework to restore light fields.

5. Experimental results

The experiments conducted in this paper use real-world light fields from the EPFL [56], INRIA,[2] and Stanford[3] datasets. Both EPFL and INRIA are light fields that are captured by a plenoptic camera and therefore have low angular disparities while the Stanford dataset is captured using a Gantry which have larger angular disparities. While the angular views of the EPFL and Stanford datasets are available, the light fields in the INRIA dataset were decoded using the method in [57]. In all our experiments we consider a 9×9 matrix of angular views. For computational purposes, the high-resolution views of the Stanford dataset were down-sampled such that the lowest dimension is set to 400 pixels. The high-resolution images of the other datasets were kept unchanged *i.e.* 625×434.

[2] INRIA dataset: https://www.irisa.fr/temics/demos/IllumDatasetLF/index.html.

[3] Stanford dataset: http://lightfield.stanford.edu/.

Table 2
Quality analysis (PSNR with SSIM in parenthesis) using different light field super-resolution algorithms when considering a magnification factor of ×3.

Light Field	Bicubic	BM-PCARR	LF-SRCNN	GRAPH	PB-VDSR
Bikes	27.55 (0.87)	28.78 (0.89)	28.73 (0.88)	29.22 (0.90)	29.87 (0.90)
Bench in Paris	22.43 (0.79)	23.33 (0.83)	23.07 (0.82)	23.25 (0.83)	23.48 (0.83)
Friends 1	31.17 (0.90)	32.17 (0.92)	32.17 (0.92)	32.17 (0.92)	33.16 (0.92)
Sphynx	27.65 (0.77)	28.73 (0.81)	28.45 (0.80)	28.88 (0.81)	28.76 (0.80)
Bee 2	31.02 (0.91)	32.03 (0.91)	32.24 (0.92)	32.74 (0.93)	32.62 (0.92)
Duck	23.35 (0.84)	24.22 (0.86)	24.19 (0.87)	24.43 (0.88)	24.50 (0.88)
Fruits	28.74 (0.85)	30.21 (0.89)	29.87 (0.88)	30.91 (0.91)	30.20 (0.89)
Rose	34.05 (0.90)	35.30 (0.92)	35.00 (0.91)	36.19 (0.94)	34.98 (0.91)
Mini	27.30 (0.77)	28.23 (0.79)	28.03 (0.79)	28.31 (0.81)	28.55 (0.80)
Chess	30.04 (0.92)	31.02 (0.93)	30.88 (0.93)	31.69 (0.94)	31.61 (0.94)
Bunny	32.91 (0.94)	34.31 (0.94)	34.14 (0.94)	35.31 (0.96)	35.64 (0.95)
Lego Bulldozer	26.21 (0.86)	27.05 (0.87)	27.10 (0.87)	28.27 (0.90)	28.15 (0.89)
Lego Truck	30.26 (0.89)	31.18 (0.91)	30.99 (0.91)	31.62 (0.92)	31.39 (0.92)
Lego Knights	27.28 (0.86)	28.15 (0.88)	28.24 (0.87)	28.62 (0.90)	29.01 (0.90)
Overall	**28.57(0.86)**	**29.62 (0.88)**	**29.51 (0.88)**	**30.12 (0.90)**	**30.14 (0.89)**

Table 3
Quality analysis (PSNR with SSIM in parenthesis) using different light field super-resolution algorithms when considering a magnification factor of ×4.

Light Field	Bicubic	BM-PCARR	LF-SRCNN	GRAPH	PB-VDSR
Bikes	25.33 (0.80)	26.42 (0.82)	26.28 (0.82)	26.62 (0.84)	27.85 (0.82)
Bench in Paris	21.00 (0.72)	21.74 (0.75)	21.50 (0.75)	21.57 (0.75)	21.73 (0.75)
Friends 1	29.15 (0.86)	30.14 (0.88)	30.10 (0.88)	30.08 (0.88)	30.89 (0.88)
Sphynx	25.88 (0.70)	26.89 (0.74)	26.62 (0.72)	26.86 (0.74)	26.79 (0.72)
Bee 2	28.72 (0.86)	29.85 (0.87)	29.85 (0.88)	30.25 (0.89)	30.27 (0.88)
Duck	21.62 (0.76)	22.29 (0.79)	22.25 (0.79)	22.44 (0.81)	22.50 (0.80)
Fruits	26.60 (0.78)	27.82 (0.82)	27.53 (0.80)	28.28 (0.84)	27.28 (0.80)
Rose	31.86 (0.84)	33.05 (0.87)	32.57 (0.85)	33.42 (0.88)	32.12 (0.84)
Mini	25.71 (0.70)	26.40 (0.72)	26.30 (0.71)	26.45 (0.73)	26.73 (0.72)
Chess	28.03 (0.87)	28.90 (0.88)	28.77 (0.88)	29.31 (0.90)	28.65 (0.88)
Bunny	30.47 (0.90)	31.80 (0.91)	31.57 (0.91)	32.30 (0.92)	32.33 (0.91)
Lego Bulldozer	24.29 (0.79)	25.00 (0.80)	25.02 (0.81)	25.85 (0.84)	25.13 (0.81)
Lego Truck	28.55 (0.85)	29.35 (0.87)	29.15 (0.86)	29.56 (0.87)	29.09 (0.86)
Lego Knights	25.20 (0.79)	26.13 (0.81)	26.84 (0.81)	26.67 (0.84)	25.89 (0.81)
Overall	**26.60 (0.80)**	**27.60 (0.82)**	**27.45 (0.82)**	**27.83 (0.84)**	**27.66 (0.82)**

We compare the performance of our proposed PB-VDSR method against some of the best performing methods in the field of light field super-resolution, namely the CNN based light field super-resolution algorithm (LF-SRCNN) [11], the linear subspace projection based method (BM-PCARR) [9] and the Graph-based light field super resolution (GRAPH) [12]. It must be mentioned that while the BM+PCARR and LF-SRCNN were retrained on 98 light fields that were not considered in the evaluation phase, the network model adopted by VDSR was not retrained on light fields and we used the original model adopted for single image super-resolution. Moreover, PB-VDSR inherits the benefits of VDSR and adopts one single model to cater for different magnification factors. The other light field super-resolution methods described in the related work section were not considered since they either were reported to achieve performance inferior to the methods considered here [9,11,12] or the code was not made publicly available by the authors at the time of writing the paper. The MATLAB code of the proposed method will be made available online upon publication.[4] All methods were configured using the configuration specified in the original papers, where a disparity range of $[-6, 6]$ was adopted for the GRAPH method.

The results in Tables 2 and 3 compare these light field super-resolution methods in terms of both PSNR and SSIM for magnification factors of ×3 and ×4 respectively. It can be seen that our proposed method outperforms both BM-PCARR and LF-SRCNN and it is competitive to the GRAPH light field super-resolution method when considering both PSNR and SSIM objective quality metrics. Moreover, it can be seen in Fig. 6 that our method is able to restore central views that are much sharper and of higher quality (see bee in first row, text on the bicycle in second row, eyes of the duck in fourth row and edges of the chess board in the sixth row of Fig. 6) compared to the three leading light field super-resolution methods found in literature. One can also notice that the other methods provide aliasing (see bicycle rim in second row and teeth of the second female on the left in the fifth row of Fig. 6) and ghosting artifacts (see rabbit ears in third row in Fig. 6). It can also be seen that our proposed method yields sharper results on non-Lambertian surfaces as can be seen on the Mini light field (bottom row in Fig. 6).

One important feature of a light field is that it enables to digitally refocus the image after production. The quality of the refocused image depends on the quality of the light field and of its coherence across all the views. The results in Fig. 7 shows a number of refocused images obtained from light fields restored using the GRAPH [12] and PB-VDSR where the images were refocused using the Light Field Toolbox [58]. These results clearly show that the refocused images computed on light fields restored using PB-VDSR are sharper and of better quality. Moreover, the supplementary multimedia files show a pseudovideo of different light fields reconstructed using PB-VDSR where it is evident that the proposed method manages to restore light fields of higher quality and better angularly coherence compared to those obtained using the GRAPH method, even when considering non-Lambertian surfaces as the Tarot Cards and Crystal Ball light field where the latter fails. Moreover, these result also show the restoration of real-world applications when super-resolving the plenoptic image from 625×434 to 1875×1302.

The complexity of PB-VDSR is mainly affected by the computation of the optical flow used to align the light field (SIFT Flow in our case), of the SVD decomposition used to decompose the aligned light field, of the single image super resolution method used to restore the

[4] LF-Editing Repository: https://github.com/rrfarr/LF-Editing.

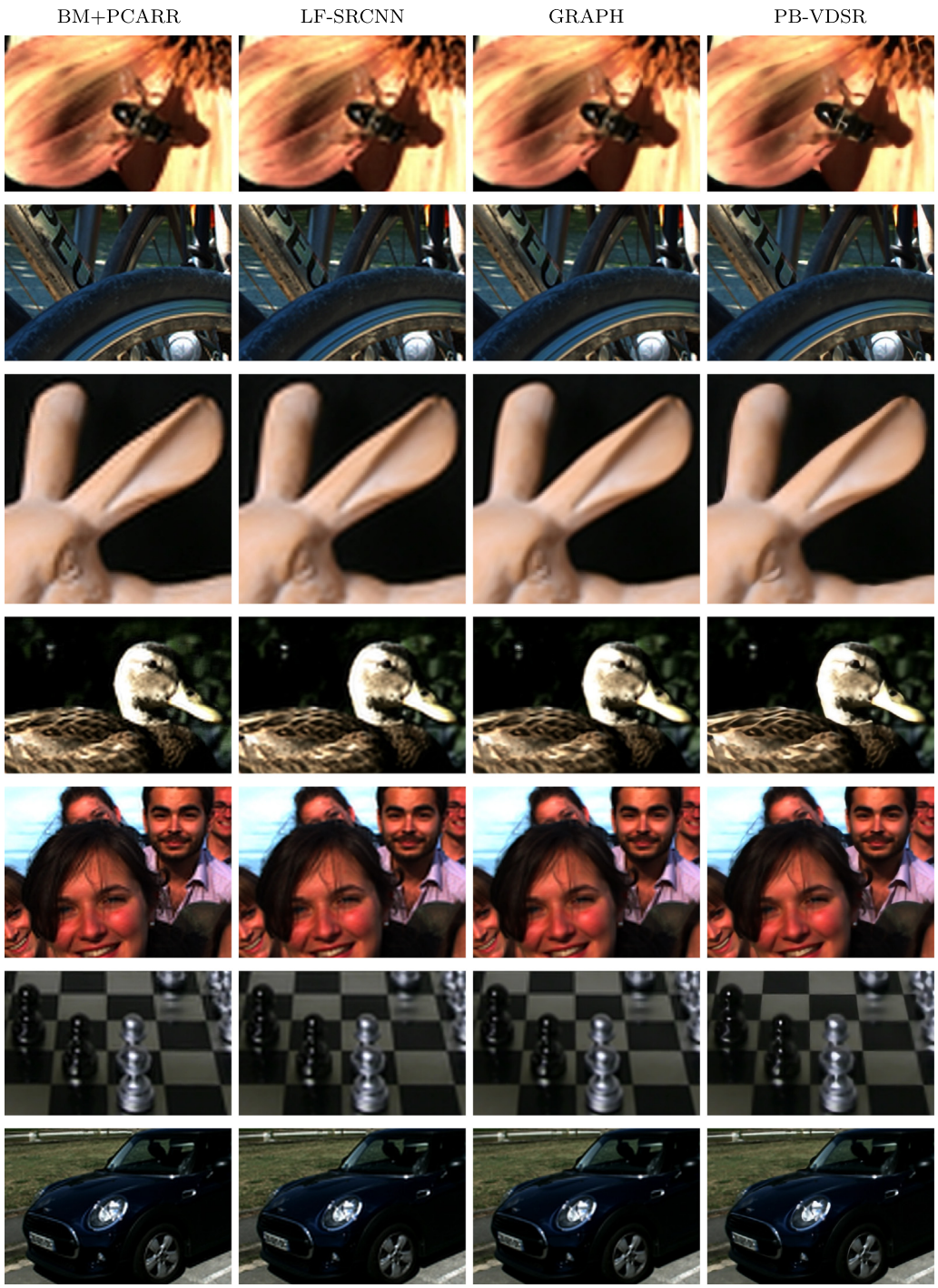
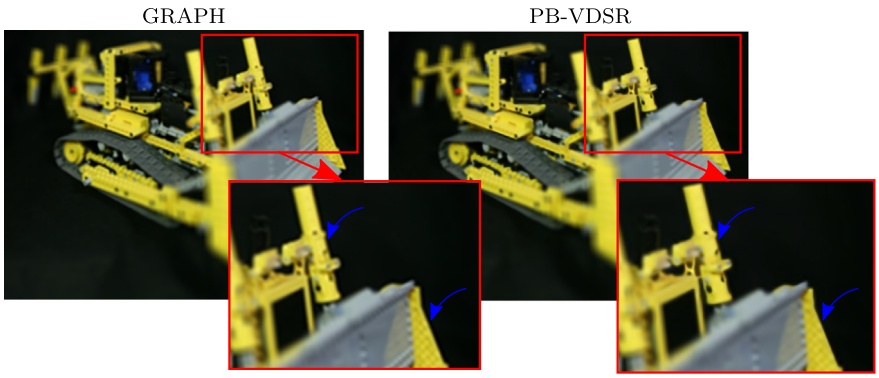
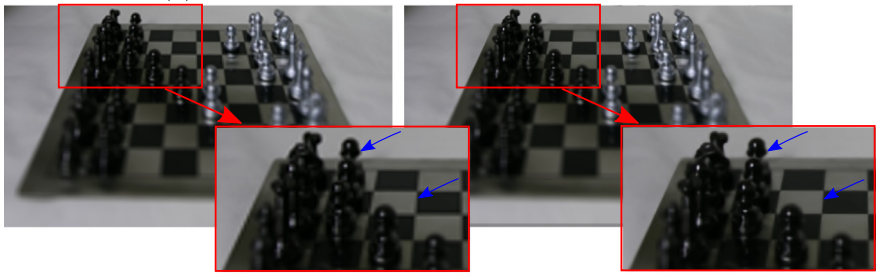


Fig. 6. Restored centre view of light fields using different light field super-resolution algorithms. These are best viewed in colour and by zooming on the views.

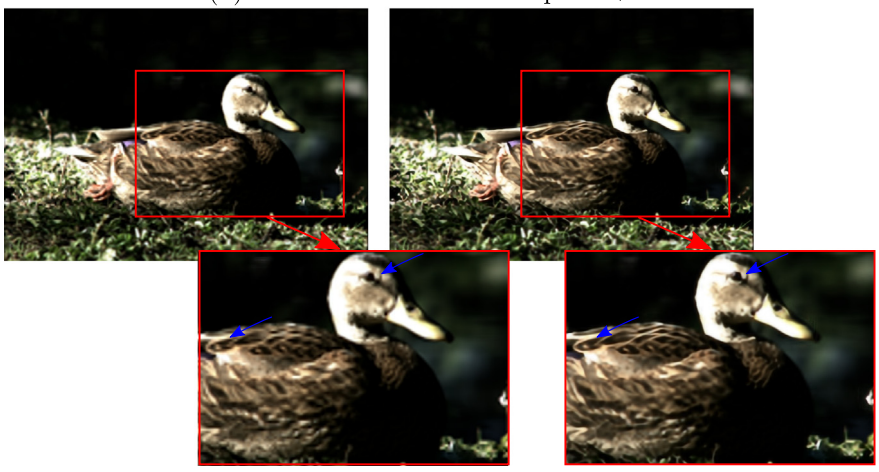
principal basis (VDSR in our case) and of the matrix multiplication that is used to propagate the restored information in the principle basis to all the other views. The Sift Flow is used to align all the n views to the centre view and is reported in [48] to have a time complexity of the order $O(nm\log(\sqrt{m}))$, where m represents the number of pixels in each view. The SVD decomposition and the matrix multiplication incur a time complexity of the order $O(n^2m)$ each. Moreover, the feed-forward part of VDSR which is used during evaluation has a fixed depth and width and its complexity is mainly dependent on the resolution of the *principal basis*. This implies that the VDSR algorithm has a time complexity of the order $O(m)$. This complexity analysis concludes that the proposed method has a time complexity that is mainly dependent on the resolution and number of views in the light field. This contrasts with the GRAPH method presented in [12] whose time complexity is proportional to α^4 where α is the magnification factor. A quantitative assessment of the complexity of different light field super-resolution methods considered in this work is summarized in Table 4, where it shows the time taken to process one light field. These methods were implemented using MATLAB with code provided by the authors and tested on an Intel Core(TM)i7 with a Windows 10 64-bit Operating System, 32-GByte RAM and a Titan GTX1080Ti GPU. The LF-SRCNN has the smallest time complexity. However, it registered the worst performance in terms of quality (see Tables 2 and 3). Our proposed method achieved the second lowest complexity which is clearly independent on the target magnification factor. On the other hand, the complexity



(a) Refocus of Lego Bulldozer at a slope of -0.6.



(b) Refocus of Chess at a slope of +0.1.



(c) Refocus of Duck at a slope of +0.0.

Fig. 7. Refocusing of different light field at different depths.

Table 4
Processing time of different light field super-resolution algorithms at different magnification factors.

Algorithm	×2	×3	×4
BM-PCARR	22 min	23 min	23 mi.
LF-SRCNN	33 s	33 s	33 s
GRAPH	4 h	7 h	1 day
PB-VDSR	9 min	9 min	9 min

of GRAPH is orders of magnitudes larger than our method and its complexity increases significantly at higher magnification factors.

6. Comments and conclusion

This paper has proposed a simple framework allowing to apply state-of-the-art SISR methods for light field super-resolution while preserving light field geometrical constraints. The problem is decomposed into two sub–problems where we first align each view to the centre view using optical flow and we then decompose the aligned light field using SVD. Experimental results show that the *principal basis* captures the coherent information in the light field and is a natural image that can be restored using state-of-the-art SISR methods. We also demonstrate that the information restored in the *principal basis* can be propagated in a consistent manner to all the other views. Experimental results show that the use of the VDSR SISR technique in the proposed framework manages to restore light fields that are sharper and coherent across the angular views, compared to existing light field super-resolution methods. Moreover, results in the supplementary material show that the restored light field is able to restore reflections on non-Lambertian surfaces. The proposed framework can be extended to other light field image processing applications such as inpainting and recolouring where one can edit the *principal basis* using state-of-the-art 2D image processing methods, and then propagate the restored information to all the other views as it was done here.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has been supported in part by the EU H2020 Research and Innovation Programme under grant agreement No 694122 (ERC advanced grant CLIM)

References

[1] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, P. Hanrahan, Light Field Photography with a Hand-Held Plenoptic Camera, Tech Report CSTR 2005-02, Stanford University Computer Science, Stanford University, 2005.

[2] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, M. Levoy, High performance imaging using large Camera arrays, ACM Trans. Graph. 24 (3) (2005) 765–776.

[3] C.-K. Liang, T.-H. Lin, B.-Y. Wong, C. Liu, H.H. Chen, Programmable aperture photography: multiplexed light field acquisition, in: Proc. of ACM SIGGRAPH, Vol. 27, 2008, pp. 1–10.

[4] Z. Xu, E. Lam, A high-resolution lightfield camera with dual-mask design, in: Proc. Image Reconstruction from Incomplete Data VII, Vol. 8500, SPIE Optical Engineering + Applications, 2012.

[5] A. Levin, W.T. Freeman, F. Durand, Understanding Camera trade-offs through a Bayesian analysis of light field projections, in: D. Forsyth, P. Torr, A. Zisserman (Eds.), Computer Vision – ECCV 2008, Springer, Berlin, Heidelberg, 2008, pp. 88–101.

[6] T.E. Bishop, P. Favaro, The light field Camera: Extended depth of field, aliasing, and superresolution, IEEE Trans. Pattern Anal. Mach. Intell. 34 (5) (2012) 972–986, http://dx.doi.org/10.1109/TPAMI.2011.168.

[7] K. Mitra, A. Veeraraghavan, Light field denoising, light field superresolution and stereo camera based refocussing using a GMM light field patch prior, in: 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2012, pp. 22–28. http://dx.doi.org/10.1109/CVPRW.2012.6239346.

[8] S. Wanner, B. Goldluecke, Variational light field analysis for disparity estimation and super-resolution, IEEE Trans. Pattern Anal. Mach. Intell. 36 (3) (2014) 606–619, http://dx.doi.org/10.1109/TPAMI.2013.147.

[9] R.A. Farrugia, C. Galea, C. Guillemot, Super resolution of light field images using linear subspace projection of patch-volumes, IEEE J. Sel. Top. Sign. Proces. 11 (7) (2017) 1058–1071, http://dx.doi.org/10.1109/JSTSP.2017.2747127.

[10] Y. Yoon, H.G. Jeon, D. Yoo, J.Y. Lee, I.S. Kweon, Learning a deep convolutional network for light-field image super-resolution, in: 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), 2015, pp. 57–65. http://dx.doi.org/10.1109/ICCVW.2015.17.

[11] Y. Yoon, H.G. Jeon, D. Yoo, J.Y. Lee, I.S. Kweon, Light-field image super-resolution using convolutional neural network, IEEE Signal Process. Lett. 24 (6) (2017) 848–852, http://dx.doi.org/10.1109/LSP.2017.2669333.

[12] M. Rossi, P. Frossard, Geometry-consistent light field super-resolution via graph-based regularization, IEEE Trans. Image Process. 27 (9) (2018) 4207–4218, http://dx.doi.org/10.1109/TIP.2018.2828983.

[13] J. Kim, J.K. Lee, K.M. Lee, Accurate image super-resolution using very deep convolutional networks, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1646–1654. http://dx.doi.org/10.1109/CVPR.2016.182.

[14] J. Kim, J.K. Lee, K.M. Lee, Deeply-recursive convolutional network for image super-resolution, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1637–1645. http://dx.doi.org/10.1109/CVPR.2016.181.

[15] X.-J. Mao, C. Shen, Y.-B. Yang, Image restoration using very deep fully convolutional encoder-decoder networks with symmetric skip connections, in: Advances in Neural Information Processing Systems (NIPS'16), 2016, arXiv:1603.09056.

[16] B. Lim, S. Son, H. Kim, S. Nah, K.M. Lee, Enhanced deep residual networks for single image super-resolution, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1132–1140. http://dx.doi.org/10.1109/CVPRW.2017.151.

[17] W. Bae, J. Yoo, J.C. Ye, Beyond deep residual learning for image restoration: persistent homology-guided manifold simplification, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1141–1149 http://dx.doi.org/10.1109/CVPRW.2017.152.

[18] C. Dong, C.C. Loy, K. He, X. Tang, Learning a deep convolutional network for image super-resolution, in: D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (Eds.), Computer Vision – ECCV 2014, Springer International Publishing, Cham, 2014, pp. 184–199.

[19] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, L. Zhang, et al., Ntire 2017 challenge on single image super-resolution: Methods and results, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Work-Shops, 2017.

[20] H. He, W. Siu, Single image super-resolution using Gaussian process regression, in: 2011 IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 449–456.

[21] K. Zhang, X. Gao, D. Tao, X. Li, Single image superresolution with non-local means and steering kernel regression, IEEE Trans. Image Process. 21 (11) (2012) 4544–4556.

[22] W.T. Freeman, T.R. Jones, E.C. Pasztor, Example-based super-resolution, IEEE Comput. Graph. Appl. 22 (2) (2002) 56–65, http://dx.doi.org/10.1109/38.988747.

[23] H. Chang, D.-Y. Yeung, Y. Xiong, Super-resolution through neighbor embedding, in: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004., Vol. 1, pp. I–I. http://dx.doi.org/10.1109/CVPR.2004.1315043.

[24] X. Gao, K. Zhang, D. Tao, X. Li, Image super-resolution with sparse neighbor embedding, IEEE Trans. Image Process. 21 (7) (2012) 3194–3205, http://dx.doi.org/10.1109/TIP.2012.2190080.

[25] J.C. Ferreira, E. Vural, C. Guillemot, Geometry-aware neighborhood search for learning local models for image superresolution, IEEE Trans. Image Process. 25 (3) (2016) 14.

[26] J. Yang, Z. Wang, Z. Lin, S. Cohen, T. Huang, Coupled dictionary training for image super-resolution, IEEE Trans. Image Process. 21 (8) (2012) 3467–3478, http://dx.doi.org/10.1109/TIP.2012.2192127.

[27] R. Timofte, V. De Smet, L. Van Gool, A+: Adjusted anchored neighborhood regression for fast super-resolution, in: D. Cremers, I. Reid, H. Saito, M.-H. Yang (Eds.), Computer Vision – ACCV 2014, Springer International Publishing, Cham, 2015, pp. 111–126.

[28] D. Glasner, S. Bagon, M. Irani, Super-resolution from a single image, in: 2009 IEEE 12th International Conference on Computer Vision, 2019, pp. 349–356. http://dx.doi.org/10.1109/ICCV.2009.5459271.

[29] G. Freedman, R. Fattal, Image and video upscaling from local self-examples, ACM Trans. Graph. 30 (2) (2011) 12:1–12:11, http://dx.doi.org/10.1145/1944846.1944852.

[30] J. Yang, Z. Lin, S. Cohen, Fast image super-resolution based on in-place example regression, in: 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 1059–1066. http://dx.doi.org/10.1109/CVPR.2013.137.

[31] M. Bevilacqua, A. Roumy, C. Guillemot, A. Morel, Single-image super-resolution via linear mapping of interpolated self-examples, IEEE Trans. Image Process. 23 (12) (2014) 5334–5347, http://dx.doi.org/10.1109/TIP.2014.2364116.

[32] M.S.K. Gul, B.K. Gunturk, Spatial and angular resolution enhancement of light fields using convolutional neural networks, IEEE Trans. Image Process. 27 (5) (2018) 2146–2159, http://dx.doi.org/10.1109/TIP.2018.2794181.

[33] Y. Wang, F. Liu, K. Zhang, G. Hou, Z. Sun, T. Tan, LFNet: A novel bidirectional recurrent convolutional neural network for light-field image super-resolution, IEEE Trans. Image Process. 27 (9) (2018) 4274–4286, http://dx.doi.org/10.1109/TIP.2018.2834819.

[34] T.-C. Wang, J.-Y. Zhu, N.K. Kalantari, A.A. Efros, R. Ramamoorthi, Light field video Capture using a learning-based hybrid imaging system, ACM Trans. Graph. 36 (4) (2017) 133:1–133:13.

[35] C.-H. Lu, S. Muenzel, J. Fleischer, High-resolution light-field imaging via phase space retrieval, Appl. Opt. 58 (2019) A142–A146.

[36] Y. Wang, Y. Liu, W. Heidrich, Q. Dai, The light field attachment: Turning a DSLR into a light field Camera using a low budget Camera ring, IEEE Trans. Vis. Comput. Graphics 23 (10) (2017) 2357–2364, http://dx.doi.org/10.1109/TVCG.2016.2628743.

[37] S.M. Seitz, K.N. Kutulakos, Plenoptic image editing, in: Sixth International Conference on Computer Vision, IEEE, 1998, pp. 17–24, http://dx.doi.org/10.1109/ICCV.1998.710696, Cat.No. 98CH36271.

[38] A. Jarabo, B. Masia, D. Gutierrez, Efficient propagation of light field edits, in: Proc. of SIACG'11, 2011, pp.75–80.

[39] X. An, F. Pellacini, AppProp: All-pairs appearance-space edit propagation, ACM Trans. Graph. 27 (3) (2008) 40:1–40:9, http://dx.doi.org/10.1145/1360612.1360639.

[40] H. Ao, Y. Zhang, A. Jarabo, B. Masia, Y. Liu, D. Gutierrez, Q. Dai, Light field editing based on reparameterization, in: Y.-S. Ho, J. Sang, Y.M. Ro, J. Kim, F. Wu (Eds.), Advances in Multimedia Information Processing – PCM 2015, Springer International Publishing, Cham, 2015, pp. 601–610.

[41] F.L. Zhang, J. Wang, E. Shechtman, Z.Y. Zhou, J.X. Shi, S.M. Hu, Plenopatch: Patch-based plenoptic image manipulation, IEEE Trans. Vis. Comput. Graphics 23 (5) (2017) 1561–1573, http://dx.doi.org/10.1109/TVCG.2016.2532329.

[42] O. Frigo, C. Guillemot, Epipolar Plane Diffusion: An Efficient Approach for Light Field Editing, in: British Machine Vision Conference (BMVC), London, France, 2017.

[43] M. Levoy, P. Hanrahan, Light field rendering, in: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, in: SIGGRAPH '96, ACM, New York, NY, USA, 1996, pp. 31–42, http://dx.doi.org/10.1145/237170.237199.

[44] S.J. Gortler, R. Grzeszczuk, R. Szeliski, M.F. Cohen, The lumigraph, in: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, in: SIGGRAPH '96, ACM, New York, NY, USA, 1996, pp. 43–54, http://dx.doi.org/10.1145/237170.237200.

[45] C.-K. Liang, R. Ramamoorthi, A light transport framework for lenslet light field Cameras, ACM Trans. Graph. 34 (2) (2015) 16:1–16:19, http://dx.doi.org/10.1145/2665075.

[46] Y. Peng, A. Ganesh, J. Wright, W. Xu, Y. Ma, Rasl: Robust alignment by sparse and low-rank decomposition for linearly correlated images, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010, pp. 763–770. http://dx.doi.org/10.1109/CVPR.2010.5540138.

[47] X. Jiang, M.L. Pendu, R.A. Farrugia, C. Guillemot, Light field compression with homography-based low-rank approximation, IEEE J. Sel. Top. Sign. Proces. 11 (7) (2017) 1132–1145, http://dx.doi.org/10.1109/JSTSP.2017.2747078.

[48] C. Liu, J. Yuen, A. Torralba, SIFT flow: Dense correspondence across scenes and its applications, IEEE Trans. Pattern Anal. Mach. Intell. 33 (5) (2011) 978–994, http://dx.doi.org/10.1109/TPAMI.2010.147.

[49] Y. Hu, R. Song, Y. Li, Efficient coarse-to-fine patch match for large displacement optical flow, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 5704–5712. http://dx.doi.org/10.1109/CVPR.2016.615.

[50] Y. Li, D. Min, M.S. Brown, M.N. Do, J. Lu, SPM-BP: Sped-up patchmatch belief propagation for continuous MRFs, in: 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 4006–4014. http://dx.doi.org/10.1109/ICCV.2015.456.

[51] I. Peter, W. Straßer, The wavelet stream - progressive transmission of compressed light field data, in: IEEE Visualization 1999 Late Breaking Hot Topics, 1999, pp. 69–72.

[52] P. Lalonde, A. Fournier, Interactive rendering of wavelet projected light fields, in: Proceedings of the 1999 Conference on Graphics Interface '99, San Francisco, CA, USA, pp. 107–114.

[53] C.-L. Chang, X. Zhu, P. Ramanathan, B. Girod, Light field compression using disparity-compensated lifting and shape adaptation, IEEE Trans. Image Process. 15 (4) (2006) 793–806, http://dx.doi.org/10.1109/TIP.2005.863954.

[54] J. Johnson, A. Alahi, L. Fei-Fei, Perceptual losses for real-time style transfer and super-resolution, in: European Conference on Computer Vision, 2016.

[55] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, Photo-realistic single image super-resolution using a generative adversarial network, in: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 105–114. http://dx.doi.org/10.1109/CVPR.2017.19.

[56] M. Rerabek, T. Ebrahimi, New light field image dataset, in: IEEE Internation Conference on Quality of Multimedia Experience, 2016.

[57] D.G. Dansereau, O. Pizarro, S.B. Williams, Decoding, calibration and rectification for lenselet-based plenoptic cameras, in: 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 1027–1034. http://dx.doi.org/10.1109/CVPR.2013.137.

[58] D.G. Dansereau, O. Pizarro, S.B. Williams, Linear volumetric focus for light field Cameras, ACM Trans. Graph. 34 (2) (2015) 15:1–15:20, http://dx.doi.org/10.1145/2665074.