# IDENTIFYING A MULTIPLE PLANE PLENOPTIC FUNCTION FROM A SWIPED IMAGE

*Michael Lawson, Mike Brookes and Pier Luigi Dragotti*

Imperial College London

## ABSTRACT

Blur in images, caused by camera motion with an open shutter, is usually thought of as a problem. The algorithm described in this paper shows instead that it is possible to use the blur caused by the integration of light rays at different locations along a moving camera trajectory to extract information about the light rays that are present within the scene. Retrieving the light rays present within a scene from different viewpoints is equivalent to retrieving the plenoptic function of the scene. In this paper, we focus on a specific case in which the blurred image of a scene, containing fronto-parallel planes with uniform unknown textures, is analysed to recreate the plenoptic function. The image is captured by a digital single lens camera with shutter open, moving in a straight line between two points, resulting in a swiped image. We estimate the EPI from this blurred image, and the EPI can be used to generate unblurred images for a given camera location.

***Index Terms*—** Plenoptic function, Plenoptic camera, Layer based depth, Blurred images

## 1. INTRODUCTION

Blur in images caused by camera motion is usually thought of as a problem, but in fact the blurring gives information about the structure of the scene that is absent from a single unblurred image. Figure 1a) shows an image of a building lobby that is blurred due to camera motion perpendicular to the optical axis during the exposure. It can be seen that the blurring gives information about the scene geometry since the camera motion affects objects near to the camera more than it affects distant objects. The goal of the work described in this paper is to use the information encapsulated in the image blurring to allow the reconstruction of unblurred images from viewpoints within the range of camera motion during the exposure, as illustrated in Figure 1b). In this paper, we refer to these blurred images as swiped images.

A convenient framework for describing images of a scene from different camera positions is the plenoptic function, which was introduced by Adelson and Bergen in in [1]. The plenoptic function describes a scene in terms of light rays observed by a camera at an arbitrary location. The complete plenoptic function has 7 dimensions, with the light intensity being $I(x, y, z, v, w, \lambda, \tau)$, where $(x, y, z)$ is the camera location, $(v, w)$ specifies the direction of the light ray, $\lambda$ is the wavelength, and $\tau$ is time [1, 2, 3, 4, 5, 6].


**Fig. 1**. a) shows a swiped image created by opening the camera shutter as the camera moves. b) shows the images reconstructed using the plenoptic function recovered from a).

In this paper, we consider a simplified version of the plenoptic function, the Epipolar Plane Image (EPI) [7] in which we restrict ourselves to monochrome images of a static scene from a camera that is constrained to move along a straight line. This reduces the dimensionality of the plenoptic function from 7 to 3, resulting in image intensity $I(x, v, w)$. It is also assumed that the surfaces in the scene are Lambertian, and so the intensity of the light rays does not vary with viewing angle.

Obtaining the plenoptic function of a scene normally requires either an array of cameras or a specialised plenoptic camera with multiple lenses, such as, for example, the Lytro Illum [8, 9, 10, 11]. Both of these options cause an increase in the cost and complexity of the camera hardware, and so are not widespread in general use consumer electronics. In this paper we show that it is possible to extract a detailed characterisation of the plenoptic function from a single swiped image.

Previous work on the plenoptic function has characterised spectral properties [2, 3, 5] and established the sampling densities needed to reconstruct images from novel viewpoints [2]. In particular it has been shown that photorealistic images can be reconstructed using a layer-based model of the scene in which object depths are coarsely quantised [12, 13, 14, 15]. This layer based model can be used as justification for the

use of fronto-parallel planes in the scene. This work builds upon work previously undertaken by Lawson et al. [16] to reconstruct the plenoptic function of a single slanted or fronto-parallel plane from a sampled swiped image. The multiple plane case is, however, much harder since we show it is similar to the problem of unlabelled sensing [17]. Unlabelled sensing refers to the case where you have a set of linear measurements of a phenomenon, but you do not know the order of observations.

The paper is organised as follows: in Section 2.1 we describe the problem formulation and the process of capturing the swiped image, in Section 2.2 we describe the process of recovering the information encapsulated within the EPI from the swiped image, and in Section 3 we show the application of the algorithm from Section 2.2 on images acquired from an experimental setup. Finally, we conclude in Section 4.

## 2. EPI RECOVERY FROM A SWIPED IMAGE

### 2.1. Problem formulation

Consider a scene that is comprised of $P$ fronto-parallel planes, with a texture on each plane that is of constant value. The scene, shown in Figure 2, is viewed from above. In the scene a point in space has coordinates $(x, z)$, the camera centre has coordinates $(t, 0)$. The horizontal limits of plane $p$ are at $x_{p1}$ and $x_{p2}$, and the vertical height is at $z_p$, with the planes not having to be at different heights. The number of planes is unknown in advance of the swiped image acquisition. In this scene, the focal length defines the image plane at $z = f$. The surface intensity of each plane is defined as $P_{I_p}$. A 2D slice of the EPI is analysed by restricting $y = w = 0$.

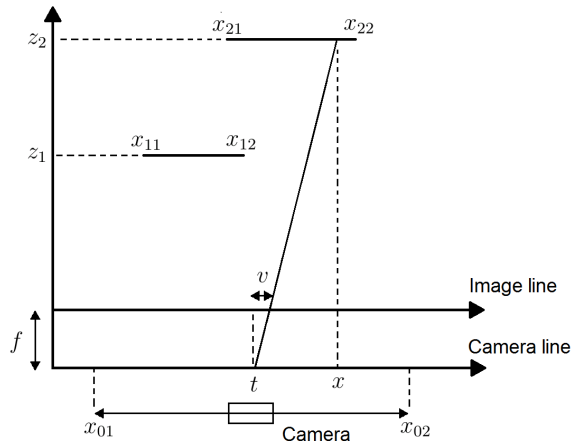


**Fig. 2**. A scene containing multiple fronto-parallel planes

Now, a swiped image of the scene can be created by swiping the camera from one position $(x_{01})$ to another $(x_{02})$, with the shutter open, and the camera moving at a constant velocity, as shown in Figure 3a). This can be thought of as being equivalent to having an additional plane $p = 0$ at $z = 0$. This plane can be thought of as a masking window, as the camera can only see through this plane, with the camera being occluded at all other locations as it is swiped from minus infinity to plus infinity, with this plane being defined by the limits $x_{01}$ and $x_{02}$. The swiped image can be created by integrating the EPI in the range $x_{01} < t < x_{02}$:

$$I(v) = \int_{t_1}^{t_2} E(v, t)dt \tag{1}$$

where $E(v, t)$ is the EPI of the scene, and $I(v)$ is the swiped image.

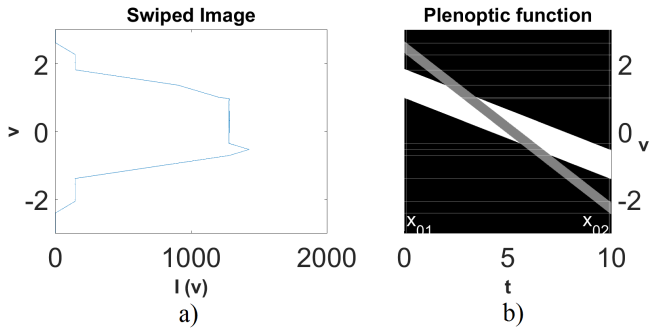


**Fig. 3**. A swiped image, a), created by integrating the plenoptic function shown in b).

A switchpoint on the swiped image is created every time a plane occludes or disoccludes another plane in the EPI, with the total number of switchpoints in the swiped image being $V$. This occurs when plane edges $(x_{pi}, z_p)$ and $(x_{qj}, z_q)$ are aligned at the camera position $t = t_{pi,qj}$, and the aligned points are at $v = v_{pi,qj}$ in the swiped image. The switchpoint locations are shown in the EPI in Figure 3b) by the horizontal white lines.

From similar triangles in Figure 2 we can write $\frac{v_{pi,qj}}{f} = \frac{x_{pi}-x_{qj}}{z_p-z_q}$ for each switchpoint. We can then rearrange this equation into the linear form:

$$x_{pi} - \bar{v}_{pi,qj} z_p - x_{qj} + \bar{v}_{pi,qj} z_q = 0 \tag{2}$$

where $\bar{v} = \frac{v}{f}$ is the normalised version of $v$. To avoid duplication, we can assume that $p > q$ and so for $0 \leq q < p \leq P$ and $1 \leq i, j \leq 2$, we have a total of $2P^2 + 2P$ equations in $3P$ unknowns. When $q = 0$, we have $4P$ simpler equations:

$$x_{pi} - \bar{v}_{pi,0j} z_p - x_{0j} = 0 \tag{3}$$

which can be rewritten as:

$$x_{pi} - \bar{v}_{pi,0j} z_p = x_{0j}. \tag{4}$$

If, for example, $P = 2$ then we get a total of 12 equations in 6 unknowns where $piqj$ take the sequential values $(1101, 1201, 1102, 2101, 2201, 2202, 2111, 2112, 2212)$. These equations can be arranged into matrix form:

$$\begin{pmatrix} 1 & 0 & -\bar{v}_{11,01} & 0 & 0 & 0 \\ 0 & 1 & -\bar{v}_{12,01} & 0 & 0 & 0 \\ 1 & 0 & -\bar{v}_{11,02} & 0 & 0 & 0 \\ 0 & 1 & -\bar{v}_{12,02} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -\bar{v}_{21,01} \\ 0 & 0 & 0 & 0 & 1 & -\bar{v}_{22,01} \\ 0 & 0 & 0 & 1 & 0 & -\bar{v}_{21,02} \\ 0 & 0 & 0 & 0 & 1 & -\bar{v}_{22,02} \\ -1 & 0 & \bar{v}_{21,11} & 1 & 0 & -\bar{v}_{21,11} \\ -1 & 0 & \bar{v}_{22,11} & 0 & 1 & -\bar{v}_{22,11} \\ 0 & -1 & \bar{v}_{21,12} & 1 & 0 & -\bar{v}_{21,12} \\ 0 & -1 & \bar{v}_{22,12} & 0 & 1 & -\bar{v}_{22,12} \end{pmatrix} \cdot \begin{pmatrix} x_{11} \\ x_{12} \\ z_1 \\ x_{21} \\ x_{22} \\ z_2 \end{pmatrix} = \begin{pmatrix} x_{01} \\ x_{01} \\ x_{02} \\ x_{02} \\ x_{01} \\ x_{01} \\ x_{02} \\ x_{02} \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Due to the geometry of a particular scene, it is possible that a switchpoint may not exist. Each switchpoint, $\bar{v}_{pi,qj}$ occurs in only one equation, so if the switchpoint doesn't exist, the corresponding equation is deleted. Therefore, given enough known switchpoints, the scene geometry can be determined. The difficulty in this problem is that switchpoints from the swiped image need to be labelled correctly, as the order of switchpoints in the swiped image is undetermined. This is similar to the problem of unlabelled sensing [17]. Once the switchpoints have been labelled correctly, the problem is reduced to that of solving a set of linear equations.

The swiped image value at each switchpoint can be calculated by the integral of the planes at each switchpoint, with the equation being:

$$\sum_{p=1}^{P} P_{w_p}(v_n) P_{I_p} = I(v_n) \tag{5}$$

where $P_{w_p}(v_n)$ is the size of plane in the $t$ axis in the EPI at switchpoint $n$ between $t_1$ and $t_2$, minus the occlusions from closer planes, $P_I$ is the intensity of the plane surfaces and $I(v_n)$ is the value of the swiped image at switchpoint $n$. There will be an Equation (5) for each switchpoint in the swiped image.

Given $I(v)$ and $x_{0i}$, the goal is to determine $P$, $x_{pi}$,$z_p$ and $P_{I_p}$ for $1 \leq p \leq P$ and $1 \leq i \leq 2$. From this information, it is possible to reconstruct the EPI of the scene.

### 2.2. EPI recovery algorithm

Within the swiped image, the switchpoints created by the alignment between the planes $p$ and plane $q = 0$ can be used to hypothesise planes. When a plane is seen unoccluded at both extremes of camera range, the distance between the switchpoints caused by the camera swipe will be equal, so switchpoint pairs with equal distances can be used to hypothesise planes. The hypothesised planes can be confirmed by checking if the switchpoints created by the occlusion or disocclusion of one hypothesised plane by another are present within the swiped image. Once these planes have been detected, remaining switchpoints can be labelled in a combinatorial approach. We can recover the EPI for a scene, given the condition that each plane is seen unoccluded at one end of the camera swipe range. Once the switchpoints have been labelled and verified, swiped image can be reconstructed to find the values of the planes and check the algorithm result correctness.

The algorithm can be described as follows:

1. Assume number of planes within the scene is the minimum possible given the number of switchpoints: $2P^2 + 2P < V$.
2. Create a histogram of switchpoint differences to find repeated differences.
3. For each repeated switchpoint difference, hypothesise that the switchpoints are $v_{p1,01}$,$v_{p2,01}$,$v_{p1,02}$,$v_{p2,02}$. These define the scene geometry ( $x_{p1}$, $x_{p2}$ and $z_p$) of a hypothetical plane $p$.
4. Create a list of hypothesised plane pair combinations, where one plane is plane $p$ and the other plane is plane $q$.
5. For each pair combination of hypothesised planes, find the $v$ location of the switchpoints that should occur by the planes occluding each other ($v_{pi,qj}$ for $1 \leq i, j \leq 2$), which are present within the swiped image if $t_{pi,qj}$ is within the range $x_{01} < t < x_{02}$. If these switchpoints exist then evidence for these planes exist.
6. If all switchpoint have been accounted for, recreate swiped image $I(v)$ and find plane surface intensities by solving Equation (5).
7. If the number of switchpoints from hypothesised planes with evidence for existence is less than $V$, hypothesise additional planes using sets of three remaining switchpoints.Two of these switchpoints are assumed to be either $v_{p1,01}$ and $v_{p2,01}$ or $v_{p1,02}$ and $v_{p2,02}$. If the switchpoints are $v_{p1,01}$ and $v_{p2,01}$ then the additional switchpoint is assumed to be $v_{p2,q1}$. If the switchpoints are $v_{p1,02}$ and $v_{p2,02}$ then the additional switchpoint is assumed to be $v_{p1,q2}$.
8. For each pair combination of hypothesised planes, find the $v$ location of the switchpoints that should occur by the planes occluding each other ($v_{pi,qj}$ for $1 \leq i, j \leq 2$), which are present within the swiped image if $t_{pi,qj}$ is within the range $x_{01} < t < x_{02}$. If these switchpoints exist then evidence for these planes exist.
9. Recreate swiped image $I(v)$ and find plane surface intensities $P_{I_p}$ by solving Equation (5).

10. If swiped image cannot be recreated, increase the number of hypothesised total planes in the scene.

In real images of scenes, it could not be assumed that the switchpoint locations were at the exact values of $v$ that would be expected from the scene geometry. This is because of small errors in camera alignment, and nearest pixel rounding errors for the switchpoint locations. To enable pairs of differences to be detected correctly, the position of the switchpoints was modelled as having a normal distribution, with $\mu = \bar{v}_{qi,pj}$ and the variance chosen based on the detected level of switchpoint uncertainty. The switchpoint difference distribution can be written as:

$$\mu_{1-2} = \mu_1 - \mu_2$$
$$\sigma^2_{1-2} = \sigma^2_1 - \sigma^2_2$$

where $\mu$ is the mean value of a switchpoint and $\sigma^2$ is the variance of a switchpoint. To find repeated differences, the Hellinger distance between the switchpoint distribution was calculated:

$$H^2 = 1 - \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2}} e^{-\frac{1}{4}\frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}}$$

If the Hellinger distance for two distances are below a threshold, the differences are considered to be the matched for the purposes of this algorithm. The threshold was chosen based on experimental data, with the chosen value being high enough to exclude most false positives, but low enough to not exclude correct repeated differences.

The complexity of this algorithm can be determined by the number of switchpoints, which is a product of the number of the planes within the scene. As the algorithm could potentially match each switchpoint to each other switchpoint to form a hypothetical plane the number of matches results in a worst case order of complexity of:

$$O\begin{pmatrix}(2P^2+2P)^2\\2\end{pmatrix} = O(P^8)$$

However, if the number of switchpoints is lower than the maximum, the complexity will be reduced. In our experiments the typical observed order of complexity was that of $O(P^4)$.

The technique outlined in this section can determine the plenoptic function given a number of assumptions. It is assumed that the swiped image acquired is unique to a scene geometry, and that each plane is seen unoccluded at either extreme of the camera range of movement. It is also assumed that each plane in the image will occlude or be occluded by another plane in the image, as the algorithm relies on the prediction of switchpoints from occlusions of planes. A single plane, which is not occluded at any point by another plane, will not have any switchpoints to predict with the algorithm. There is an ambiguity as to the location of the $\bar{v}_{11,01}$ and $\bar{v}_{12,02}$, which cannot be resolved by predicting switchpoints caused by occlusions of other planes in the scene.

## 3. NUMERICAL RESULTS

The algorithm proposed in Section 2 was verified by using real world experimentation. A DSLR camera was mounted on a motorised CineMoco camera slider. Two printed sheets of paper, with a constant surface texture on each, were setup as fronto-parallel planes. A black coloured board was used as the scene background. A large number of still images of the scene were taken as the camera slid between two points, which were linearly combined to form a swiped image (shown in Figure 4a). To reduce the impact of the black background, pixel values below a certain threshold were set to 0. The impact of noise was reduced by the application of median filtering, which preserved the edges in the swiped image for switchpoint detection. A row of pixels from the centre of the swiped image were used for analysis, as the image was assumed not to vary in the vertical direction.

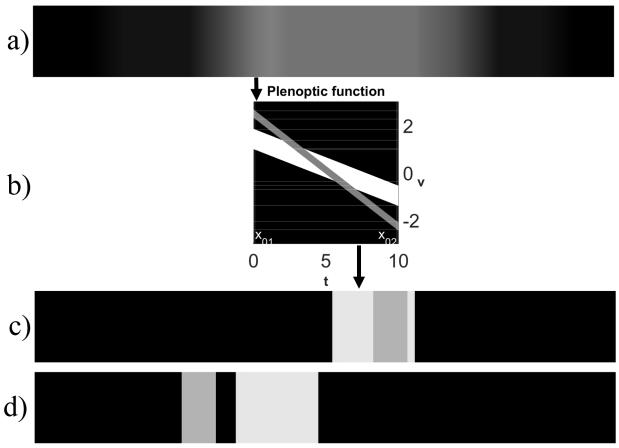


**Fig. 4**. a) shows a real swiped image created by opening the camera shutter as the camera moves. b) shows the plenoptic function recovered from a). c) and d) show recreated single images of the scene from two different camera positions.

The algorithm proposed in Section 2 was verified as recovering the location of the planes in Figure 4 to within an accuracy of 0.1%, and the texture surface intensity to within 1.5%, as well as other setups fulfilling the conditions from Section 2.2. Using the information recovered, the EPI of the scene could be successfully reconstructed, with the EPI being shown in Figure 4b), and new images of the scene were generated, as shown in Figure 4c) and Figure 4d).

## 4. CONCLUSION

In this paper, we demonstrate that it is possible to exactly reconstruct the plenoptic function of a scene with multiple fronto-parallel planes, given a swiped image of the scene. Extending this algorithm to a complex real-world scene is being investigated currently.

## 5. REFERENCES


[1] E. H. Adelson and J. R. Bergen, *The plenoptic function and the elements of early vision*, Vision and Modeling Group, Media Laboratory, Massachusetts Institute of Technology, 1991.

[2] C. Gilliam, P. L. Dragotti, and M. Brookes, "On the spectrum of the plenoptic function," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 502–516, 2014.

[3] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum, "Plenoptic sampling," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 307–318.

[4] A. L. Da Cunha, M. N. Do, and M. Vetterli, "On the information rates of the plenoptic function," *IEEE Transactions on Information Theory*, vol. 56, no. 3, pp. 1306–1321, 2010.

[5] M. N. Do, D. Marchand-Maillet, and M. Vetterli, "On the bandwidth of the plenoptic function," *IEEE Transactions on Image Processing*, vol. 21, no. 2, pp. 708–717, 2012.

[6] L. McMillan and G. Bishop, "Plenoptic modeling: An image-based rendering system," in *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*. ACM, 1995, pp. 39–46.

[7] R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," *International Journal of Computer Vision*, vol. 1, no. 1, pp. 7–55, 1987.

[8] E. H. Adelson and J. Y. A. Wang, "Single lens stereo with a plenoptic camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 99–106, Feb. 1992.

[9] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Computer Science Technical Report CSTR*, vol. 2, no. 11, pp. 1–11, 2005.

[10] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 924–934, 2006.

[11] R. Ng, "Fourier slice photography," in *ACM Transactions on Graphics (TOG)*. ACM, 2005, vol. 24, pp. 735–744.

[12] J. Berent and P. L. Dragotti, "Plenoptic manifolds," *IEEE Signal Processing Magazine*, vol. 24, no. 6, pp. 34–44, 2007.

[13] A. Gelman, P. L. Dragotti, and V. Velisavljevic, "Multiview image coding using depth layers and an optimized bit allocation," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4092–4105, 2012.

[14] A. Gelman, J. Berent, and P. L. Dragotti, "Layer-based sparse representation of multiview images," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1–15, 2012.

[15] J. Pearson, M. Brookes, and P. L. Dragotti, "Plenoptic layer-based modeling for image based rendering," *IEEE Transactions on Image Processing*, vol. 22, no. 9, pp. 3405–3419, 2013.

[16] M. Lawson, P. L. Dragotti, and M. Brookes, "Capturing the plenoptic function in a swipe," in *SPIE Optical Engineering + Applications*, 2016.

[17] J. Unnikrishnan, S. Haghighatshoar, and M. Vetterli, "Unlabeled sensing: Solving a linear system with unordered measurements," in *2015 53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2015, pp. 786–793.