Anamorphic Pixels for Multi-Channel Superresolution

Alexander Oberdörster
Fraunhofer IOF
D-07749 Jena, Germany
alexander.oberdoerster@iof.fraunhofer.de

Paolo Favaro
University of Bern
CH-3012 Bern, Switzerland

Hendrik P. A. Lensch
Eberhard Karls University Tübingen
D-72076 Tübingen, Germany

Abstract

Superresolution from plenoptic cameras or camera arrays is usually treated similarly to superresolution from video streams. However, the transformation between the low-resolution views can be determined precisely from camera geometry and parallax. Furthermore, as each low-resolution image originates from a unique physical camera, its sampling properties can also be unique. We exploit this option with a custom design of either the optics or the sensor pixels. This design makes sure that the sampling matrix of the complete system is always well-formed, enabling robust and high-resolution image reconstruction. We show that simply changing the pixel aspect ratio from square to anamorphic is sufficient to achieve that goal, as long as each camera has a unique aspect ratio. We support this claim with theoretical analysis and image reconstruction of real images. We derive the optimal aspect ratios for sets of 2 or 4 cameras. Finally, we verify our solution with a camera system using an anamorphic lens.

1. Introduction

In multi-view camera systems, which record the same subject from several viewpoints, multiple images of a scene are assembled into novel views. Multi-view systems include plenoptic cameras, camera arrays or microscopic multi-aperture systems. Depending on the architecture, the source images are either recorded by single microlenses, cameras or optical channels. For simplicity, we use the term 'cameras' in this paper.

To improve the resolution of the computed views, superresolution (SR) methods such as reconstruction-based algorithms are often used. This is usually done as an afterthought, working on images from an existing camera system. Multi-view systems, however, have unique challenges and possibilities.

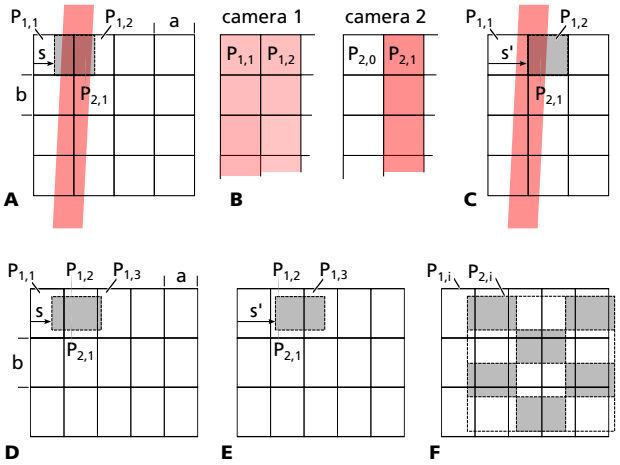

Figure 1. A prerequisite of superresolution is that a second measurement provides additional information over the first measurement. This is the case in **A**, where a pixel $P_{2,1}$ of camera 2 samples an area distinct from any pixels of camera 1 (like $P_{1,1}$ and $P_{1,2}$). It can resolve the red diagonal line that camera 1 missed (**B**). However, with multi-view systems, the shift depends on parallax. For another distance, $P_{2,1}$ does not provide useful additional information (**C**). We propose to change the sampling area to non-square, or anamorphic (**D**, $a \neq b$). Regardless of shift, sampling 2 now always provides additional information over sampling 1 (**E**). This also the case for all other pixels of camera 2 (**F**). There is a direct effect on the numerical stability of the superresolution problem (see Figure 5).

On one hand, the relation between cameras cannot be controlled completely: it is subject to parallax, which depends on the distance from objects in the scene to the cameras. For certain distances, some cameras may sample the scene at the same points in space, not adding any new spa-

tial information. Therefore, superresolution is ineffective at these distances.

On the other hand, performance can be improved by designing the camera system with superresolution in mind. In contrast to the traditional SR setup, where one camera sucessively captures multiple frames, each camera can now be tuned individually. As we will show, specifically the parallax problem can be adressed in this way.

Our proposed solution is to provide each camera with a unique pixel shape (Figure 1). We show that rectangular, non-square – or anamorphic – pixels with unique aspect ratios provide the necessary distinction between cameras. In contrast to more exotic pixel shapes, anamorphic pixels are easier to design and lay out on an image sensor. In fact, anamorphic pixels have already been used in commercial products in the past.

The effect of anamorphic pixels is two-fold: First, each camera now has a unique sampling frequency in X and Y, reducing the chance that two cameras will sample objects at exactly the same points. Even when they do, the unique pixel shape ensures that the information captured is still slightly different.

The same effect can also be achieved optically: Instead of modifying the image sensor itself, anamorphic lenses can be used. We discuss this option further in Section 8.

We first review previous work in Section 2. Then, we explain the challenges to SR caused by parallax (Section 3). In Section 4, the camera model is defined, along with the system sampling. We then show that the sampling matrix is better-formed for anamorphic pixels (Section 5). We also derive the optimal aspect ratios for systems of 2 and 4 cameras. Next, we describe our reconstruction algorithm in Section 6. Then, we simulate systems with square and anamorphic pixels and reconstruct images (Section 7). Finally, we experimentally verify our method with an anamorphic lens (Section 8).

2. Related Work

This work is an application of superresolution techniques to camera arrays. Superresolution techniques, where multiple low-resolution images (LRI) are combined to compute a high-resolution image (HRI), have a long history in computer vision; a comprehensive overview was provided by Park [14].

Before reconstruction, the LRIs have to be registered. Registration is an important and challenging process in itself. Often, no information about the relative positioning of the LRIs is known. Then, this part of the camera model has to be estimated along with the high-resolution image [10]. Good estimates of the relative transformations are crucial to high-quality reconstruction [11]. Sub-pixel precision is necessary for superresolution to be effective. In our case, the geometric relations between the cameras can be measured and fixed before image acquisition and reconstruction. A remaining parameter in the camera transformation, however, is parallax shift, which is different depending on object distance. During HRI reconstruction, parallax shift has to be estimated for all parts of the image. In this work, we limit our considerations to flat scenes with a known distance. However, Bishop and Favaro describe a technique where initial estimates of parallax and HRI are refined in an alternating fashion [3], which is also applicable to non-square pixels.

Our results are relevant to a multitude of multi-view camera systems. These include plenoptic cameras with macroscopic relay lenses and microlens arrays, first described by Adelson and Wang [1] and in practical form by Ng [13]. It also applies to microscopic cameras that have optical channels with overlapping fields of view (FOV). This includes Lippman's original integral camera [12] and TOMBO [18], where each channel has the same FOV, but also architectures with fanned-out FOVs, such as the eCLEY [4]. It is in principle also applicable to macroscopic camera arrays such as the Stanford large camera array [20]. Our assumption that camera geometry is known does not apply in this case, however, and therefore it has to be established first [19]. Our considerations require a recording system with at least two physically distinct cameras or optical channels; they are therefore not relevant to video-based SR methods.

Non-square pixels have been routinely used in the camera industry. These pixels were part of some early HD broadcast cameras. Here, they were used because of bandwidth restrictions and for standards compatibility. Fujifilm has also produced several variants of image sensors with mixtures of hexagonal and square pixels of different sizes. These pixels have different sensitivity, increasing dynamic range.

More complex pixel shapes have been proposed by Ben-Ezra *et al*. [2]. They also identify pixel overlap as a main limitation to SR effectiveness. As a remedy, they propose an irregular Penrose tiling, which eliminates any chance of overlap. The rhombic pixel shape, however, poses a significant hurdle for implementation, as new pixel designs are expensive and image sensor companies prefer to iteratively tune pixels from existing designs.

Sasao *et al*. [16] achieve random pixel shapes by spraying an image sensor with toner particles. Schöberl *et al*. [17] use a deterministic, pseudo-random mask for the same end. Both methods reduce the active area of all pixels, however, and therefore reduce the light sensitivity of the system. This is not the case for our method, as fill factor and pixel area are unaffected.

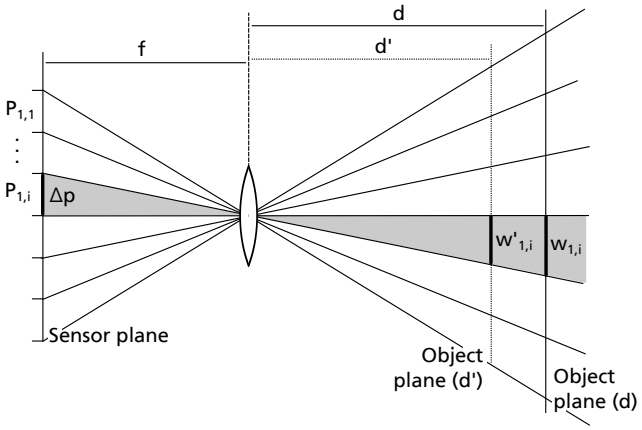

Figure 2. Through the lens with focal length f, pixel $P_{1,i}$ integrates light from the gray cone. On the object plane at distance d, it integrates over line $w_{1,i}$.

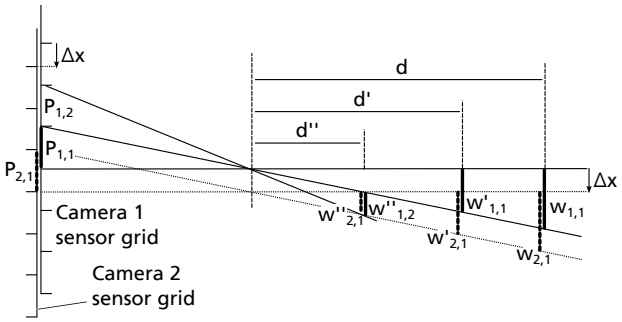

Figure 3. Camera 2 is shifted by Δx in relation to camera 1. The sampling areas of the pixels are shifted as well. At distance d, sampling width $w_{2,1}$ of pixel $P_{2,1}$ has a large overlap with sampling width $w_{1,1}$ of pixel $P_{1,1}$. This overlap decreases for a shorter distance d'. However, the overlap with $w'_{1,2}$ of pixel $P_{1,2}$ increases. At an even shorter distance d'', pixel 1 of camera 2 samples exactly the same area as pixel 2 of camera 1.

3. Parallax and sampling in multi-view systems

In video superresolution, the set of LRIs is comprised of subsequent frames from a video sequence. The relation between these frames is given by the movement of the camera between frames (and in general, also by object movement), which can be very complex and is generally unknown.

In multi-view superresolution, the relation between LRIs is given by the geometric relation between the cameras. In principle, each camera can be different, having a unique field of view, geometric distortion and viewing direction. These parameters can be established prior to recording, either from design data or from calibration [19] [7]. However, even if all cameras are identical, they are necessarily located at different places in space. We assume that they are situated on a plane perpendicular to their viewing direction.

To see how the cameras sample object space, let us first look at a single camera in a one-dimensional view (Figure 2). The camera has focal length f, pixel pitch Δp and a fill factor of 1. From the intercept theorem, we see that pixel $P_{1,i}$ of this camera 1 integrates over w_i on an object at distance d:

$$w_i = \frac{d}{f}\Delta p. \tag{1}$$

With fixed focal length, w_i increases proportionally with distance.

Adding a second camera at offset Δx (Figure 3), we see that the integration area of pixel $P_{2,1}$ of camera 2 is always offset by Δx, independent of distance. However, the relative shift $s = \frac{\Delta x}{w_i}$, measured in pixel integration widths, is

$$s = \frac{\Delta x}{\Delta p}\,\frac{f}{d}. \tag{2}$$

Let us now assume that the scene lies on a plane parallel to the camera plane. Then, this relative offset is the same for all pixels. Therefore, parallax causes a shift of the complete camera 2 sampling grid in relation to camera 1. This shift is inversely proportional to d and also proportional to the displacement of the cameras Δx.

Further, recall that s is measured in pixel widths. Therefore, whenever s has integer value, the integration area of $P_{2,1}$ in camera 2 is identical to that of another pixel $P_{1,2}$ in camera 1. In other words, the sampling grids of camera 1 and 2 align exactly. Figure 1 shows this phenomenon for a two-dimensional sampling grid[1]. Camera 2 now contributes a second measurement over the same integration areas. This second measurement can be used to reduce noise (by $\sqrt{2}$), but it adds no spatial information. Reconstruction from LRI 1 and LRI 2 is equivalent to averaging the LRIs and deblurring. Note that even for integer s, the value of s still has to be recovered prior to reconstruction.

For distances with a fractional s, however, LRI 2 does add information, and the reconstruction will be both more robust and of higher quality. For a given camera system, the SR reconstruction quality – or sharpness – therefore depends on object distance. As a real scene consists of objects at different depths, some regions of the HRI are easier to reconstruct; others are bound to be less sharp. This is true even with perfect knowledge of scene geometry.

All multi-view systems that employ SR show this problem. It has been explicitly pointed out for the case of plenoptic cameras [3] [6]. For microscopic camera arrays, only the case of $s = 0$ has been noted, i.e. the case where object distance is much larger than channel separation, resulting in disappearing parallax shift [9].

[1] There is no parallax in Y, because the cameras themselves are placed at the same Y position.

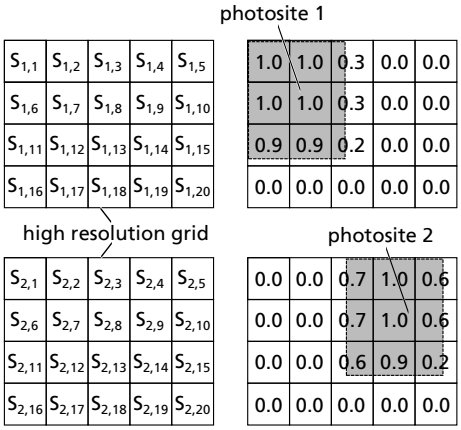

Figure 4. Each row of one camera's sampling matrix $\mathbf{S}_i$ is determined from the overlap of one photosite with the pixels of the high resolution grid. Note that while all sampling areas of one camera have the same shape, each camera has a unique sampling area shape.

4. Camera model

In order to analyze the performance of the camera system we model it as a linear system. First, we exclude lens blur from our model. The subject of this work is superresolution at the detector level. Sharpness of the LRIs should be ensured either by optics optimization or by deblurring before reconstruction.

The radiance of the scene $\mathbf{X}$ is then measured by the cameras into a set of measurements $\mathbf{Y}$. We consider a planar scene at distance d. Then, $\mathbf{X}$ is the HRI to be reconstructed and the measurements $\mathbf{Y}_i$ of a camera i – or one LRI – are given by

$$\mathbf{Y}_i = \mathbf{S}_i(d)\mathbf{X} + \mathbf{E}, \tag{3}$$

where $\mathbf{S}_i$ is the sampling matrix of camera i and $\mathbf{E}$ is the measurement noise. $\mathbf{X}$ is discretized on a high-resolution, square cartesian grid (the HRG) and sorted in lexicographical order. Each row of $\mathbf{S}_i$ describes one photosite of camera i. Row j of $\mathbf{S}_i$ is calculated as the overlap of photosite j with each location on the HRG (Figure 4).

Each photosite of camera i has the same rectangular shape, with width a_i and height b_i, measured in units of the HRG. Different cameras have photosites of distinct shapes, however, with a unique aspect ratio. To make sure all cameras have the same sensitivity, all photosites have the same area. When the HRG has n values in X and m values in Y direction, the LRG has $N_i = \frac{n}{a_i}$ values in X and $M_i = \frac{m}{b_i}$ in Y. $\mathbf{S}_i$ therefore has $N_i\, M_i \times n\, m$ elements.

$\mathbf{S}_i$ depends on object distance d. While the grid of photosites does not change with d, parallax causes a shift s_i (Equation 2) of the complete LRG, depending on the camera displacement Δx. Camera 1 is set as the reference with no parallax shift. The shift of the other cameras is calculated relative to this camera.

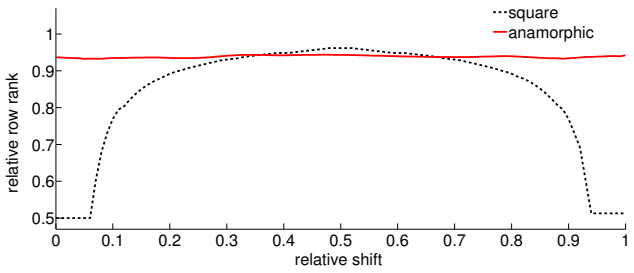

Figure 5. Relative row rank (RRR) of the sampling matrix of a two-camera system. Numerical rank with a threshold of 0.2 plotted over the shift between the sampling grids of the two cameras. With square pixels (black dotted line), RRR is highly shift-dependent. Anamorphic pixels avoid this problem.

We arrive at the complete model by concatenating all measurement matrices $\mathbf{S}_i$ vertically into one matrix $\mathbf{S}$. The resulting measurement vector

$$\mathbf{Y} = \mathbf{S}(d)\mathbf{X} + \mathbf{E} \tag{4}$$

contains $\sum_i N_i\, M_i$ measurements.

5. Numerical stability and error analysis

To reconstruct the HRI from the measurements, we have to solve the linear system (4). To develop an intuition for the complexity of the problem, we next analyse the sampling matrix $\mathbf{S}(d)$. We develop two metrics and use them to compare square and non-square pixel configurations.

First, recall that matrix $\mathbf{S}(d)$ has dimensions of $\sum_i N_i\, M_i \times n\, m$. It is obvious that the problem is underdetermined when $\sum_i N_i\, M_i < n\, m$. For square pixels, this is the case when the number of cameras is smaller than $\frac{n}{N} \cdot \frac{m}{M}$.

This implies that the row rank of $\mathbf{S}$ is equal to its number of rows. The system is still underdetermined, however, when d is so that $\Delta x = a_i$ for one of the cameras. Then, the sampling grids align exactly, some of the rows of $\mathbf{S}$ are now linearly dependent and the row rank of $\mathbf{S}$ is smaller than $\sum_i N_i\, M_i$.

We introduce the relative row rank (RRR) as the ratio between the row rank and the number of rows. It quantifies the actual information a camera system can capture in a certain configuration – particularly at a specific distance d – compared to its number of photosites.

We first compare systems of two cameras. Figure 5 shows the relative row rank (RRR) of a square configuration. RRR is plotted over the shift $\frac{s}{a_i}$ measured in LRG units. When shift is $\frac{1}{2}$, overlap between sampling units is minimal and RRR is close to 1. When shift is 0, the sampling grids align exactly and RRR drops to $\frac{1}{2}$. When we

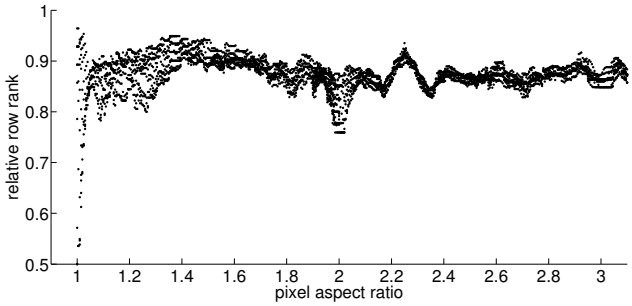

Figure 6. Dependency of relative row rank (RRR) on pixel aspect ratio (AR). For each aspect ratio, the RRR was calculated over a range of sampling grid shifts. The wider the scatter plot at a specific AR, the larger the depth dependence at this AR. The drop in rank is most dramatic for square pixels (AR of 1) and still significant for AR < 1.4. Aspect ratios with integer values (especially 2) also lead to a depth-dependent rank.

plot *numerical* rank to account for instabilities from inaccurate camera alignment and measurement noise, RRR is significantly smaller than 1 even for $s \neq 0$.

The second graph in Figure 5 plots the RRR for two cameras with anamorphic pixels. In this case, the pixel's aspect ratio is 3.2. The RRR is now almost shift-independent, meaning that the complexity of the reconstruction problem is independent of object distance. The RRR is also never significantly lower than RRR for square pixels, indicating that non-square pixels offer comparable performance even for favorable shifts.

Next, we use RRR as a metric to find the optimal aspect ratio (AR) of the pixels. In the case of two cameras, only a single AR has to be chosen; the second pixel is derived from the first pixel by rotating it by $90°$. This is also beneficial for the implementation on silicon, as each particular aspect ratio requires a new pixel design.

In Figure 6, we can see that any AR larger than 1.4 is a good choice, as the rank is consistently large. However, integer values should be avoided, as they lead to a drop in RRR for certain shifts. For these values, a partial alignment of the sampling grid may still happen for every second or third pixel.

Next, we examine the robustness to inaccurate knowledge of depth. To this end, we derive the flawed system matrix $\tilde{\mathbf{S}}$ from the true system matrix $\mathbf{S}$ by moving each of the sampling areas by a random vector $\mathbf{r}$ with a variance of σ. We generated the low-resolution images $\mathbf{Y}$ by applying $\mathbf{S}$ to the high-resolution image $\mathbf{X}$. Then, we estimated $\mathbf{X}$ from $\mathbf{Y}$ with the pseudo-inverse of both sampling matrices $\mathbf{S}$ and $\tilde{\mathbf{S}}$. In this way, we obtained an optimal reconstruction $\hat{\mathbf{X}}_{true}$ and a perturbed reconstruction $\hat{\mathbf{X}}_{err}$. From their differences

$$\mathrm{err}_1 = ||\hat{\mathbf{X}}_{true} - \hat{\mathbf{X}}_{err}||_2^2 \qquad (5)$$

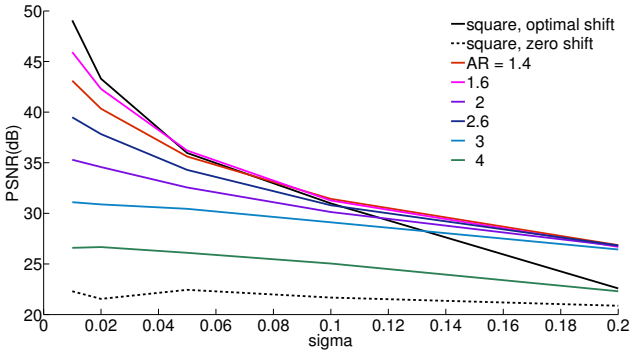

Figure 7. Robustness of reconstruction to inaccurate knowledge of depth. The sampling pixels are moved from their actual positions randomly, with variance σ. The figure compares the reconstruction from this flawed matrix to the one from the accurate matrix. Anamorphic pixels (colored lines) outperform square pixels (black lines) in most cases. For aspect ratios of 1.4 to 1.6, robustness is comparable or better than than square pixels in the optimal case. With zero shift (dotted line), square pixels perform worse than all tested anamorphic pixels.

we obtain the reconstruction error caused by a perturbed sampling matrix $\tilde{\mathbf{S}}$.

For $\mathbf{X}$, we used a random set of images randomly downloaded from the internet service Flickr [8]. From these images, we randomly selected small patches containing high spatial frequencies. In Figure 7, we see that anamorphic pixels lead to more robust reconstructions for low aspect ratios. Robustness is worse only for very oblong pixels.

We next used the same database of patches to explore how the reconstruction error depends on aspect ratio. Now, we compared the optimal reconstruction with ground truth:

$$\mathrm{err}_2 = ||\hat{\mathbf{X}}_{true} - \mathbf{X}||_2^2. \qquad (6)$$

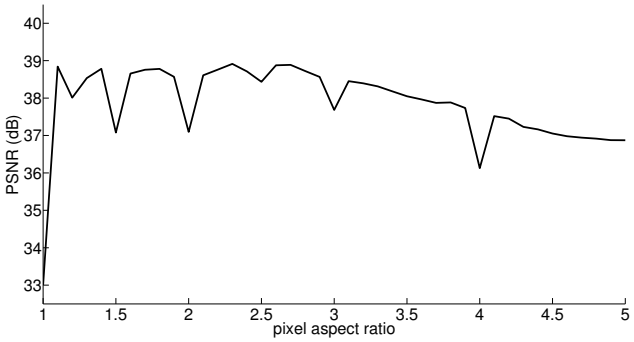

Figure 8. Mean reconstruction quality for 100 random 50×50 patches from 100 images, plotted over the aspect ratio. Low aspect ratios offer the best performance. Integer aspect ratios or simple ratios such as $\frac{3}{2}$ cause partial sampling grid alignment and reduced quality.

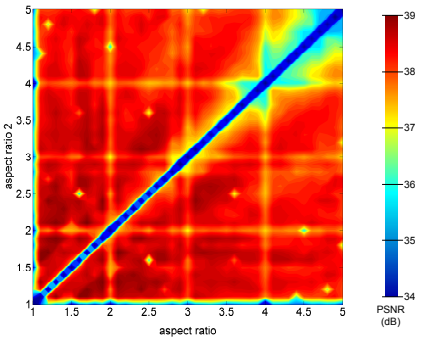

Figure 9. Mean reconstruction quality (PSNR) for the case of 4 cameras, with 2 pairs of aspect ratios shown on the X and Y axis. Shift is set to 0 to emphasize the differences between aspect ratios. As in Figure 8, lower aspect ratios lead to better results and integer ARs should be avoided. Furthermore, the aspect ratios of the two camera pairs should be different from each other.

100 random patches were sampled and reconstructed with a range of aspect ratios, yielding 100 curves. The mean of these curves is shown in Figure 8, giving the average reconstruction quality for each aspect ratio. Cameras with aspect ratios between 1.1 and 2.5 give the best quality, confirming matrix analysis (Figure 6). The results also confirm that integer ARs should be avoided.

Finally, we added two more cameras with a different aspect ratio and repeated the simulation. The new camera pair was shifted vertically, for a 2×2 array. Again, the results confirm the previous simulations (Figure 9). Additionally, we see that the two sets of cameras should have distinct aspect ratios.

6. Image reconstruction

In the last section, we reconstructed the scene by applying the pseudo-inverse of the sampling matrix to the vector of low-resolution samples. This approach is sufficient to calculate the optimal aspect ratio, but is also sensitive to noise. Furthermore, if the reconstructed HRI has dimensions exceeding the total number of input pixels, the system would be underdetermined and therefore allow infinite solutions.

A practical and theoretically-sound technique to deal with such ambiguities or uncertainty in the data is *regularization*. We therefore pose the image reconstruction task as the following minimization problem

$$\hat{\mathbf{X}} = \arg\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{Y} - \mathbf{S}\mathbf{X}\|^2 + \lambda J(\mathbf{X}) \qquad (7)$$

where $\lambda > 0$ is a constant that determines the amount of regularization and $J(\mathbf{X})$ is the *total variation* of $\mathbf{X}$ reshaped as an $m \times n$ image [15]. To minimize this cost, we employ an iterative algorithm based on local linearization of the Euler-Lagrange equation of problem (7), *i.e.*,

$$\mathbf{S}^T(\mathbf{Y} - \mathbf{S}\mathbf{X}) - \lambda \,\mathrm{vec}\left[\nabla \cdot \frac{\nabla \mathbf{X}_\square}{|\nabla \mathbf{X}_\square|}\right] = 0 \qquad (8)$$

where $\mathrm{vec}[\mathbf{A}]$ denotes the vectorized form of an image $\mathbf{A}$ with pixels arranged in lexicographical order and $\mathbf{A}_\square$ denotes the opposite transformation. At iteration t, the current estimated image $\mathbf{X}^t$ is then used to define the local linearization for the next estimate $\mathbf{X}^{t+1}$

$$\mathbf{S}^T(\mathbf{Y} - \mathbf{S}\mathbf{X}^{t+1}) - \lambda \,\mathrm{vec}\left(\nabla \cdot \frac{\nabla \mathbf{X}^{t+1}_\square}{|\nabla \mathbf{X}^t_\square|}\right) = 0. \qquad (9)$$

The linearized system is then solved via a standard conjugate gradient descent. More modern techniques can also be employed, such as [5]. However, the design of the image reconstruction algorithm is beyond the focus of this paper.

7. Simulations

We now apply the algorithm from the last section to reconstruct images. First, we simulate the imaging process for four cameras. For each of them, we generate the sampling matrices. The first camera's matrix is at the origin, the other camera's matrices are shifted vertically and/or horizontally to simulate parallax from cameras that are side by side. A high-resolution image is sampled with these matrices to generate the LRIs. To each LRI, we add Gaussian noise with variance σ. Then, we reconstructed the original image from the four LRIs. We compare the reconstruction with the original HRI to quantify reconstruction error.

The downsampling and superresolution factor was set to 2.5. For square pixels, this is also the pixel pitch. For anamorphic pixels, a_i and b_i were adjusted for constant area $a_i \cdot b_i$, ensuring that square and anamorphic LRIs have the same number of samples.

Figure 10 shows the results of the simulation. In the case of square pixels, we see that sharpness is is higher when cameras 2-4 are shifted by $\frac{a_i}{2}$. When the cameras are not shifted, the reconstructed image has significantly lower sharpness.

In the case of anamorphic pixels (AR = 1.3 and 1.56), image quality is independent of shift and better or comparable to square pixels in both cases.

Next, we examine the influence of the pixel aspect ratio. A reconstructed image of a resolution chart (Figure 11) suggests that systems with high aspect ratios have better resolution than those with low AR; however, this is only true for strictly horizontal or vertical frequencies. Diagonal resolution suffers with high aspect ratios, which leads to artifacts in natural images. Low aspect ratios yield more balanced results with less artifacts and are therefore preferable, consistent with Figure 8.

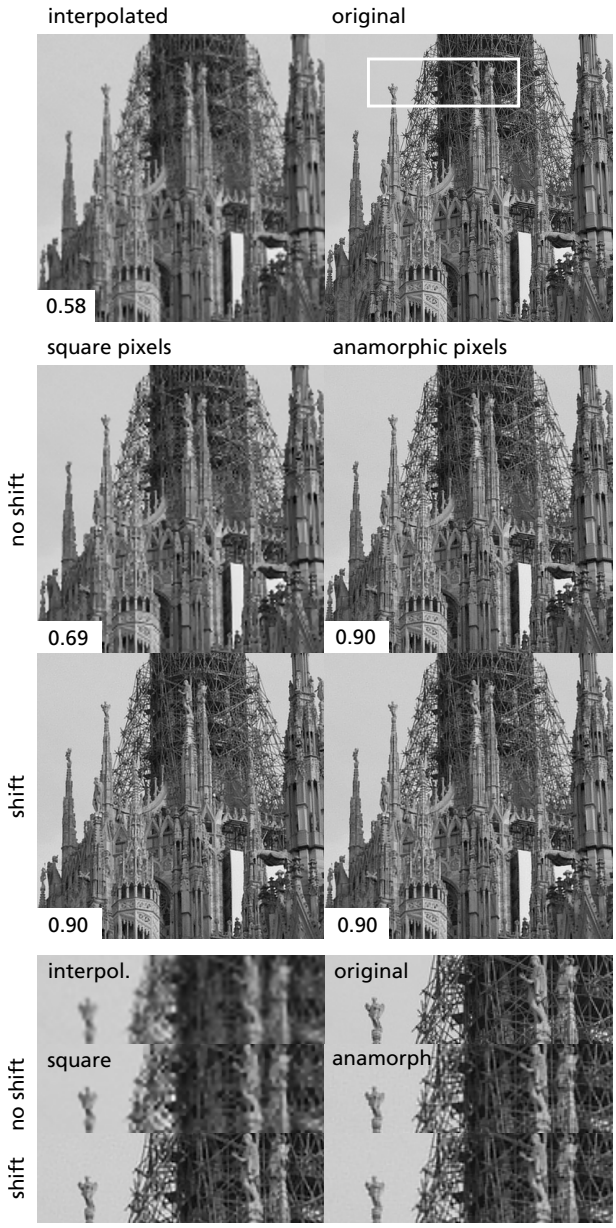

Figure 10. Reconstruction of a high-resolution image from downsampled data. The original (top right) was downsampled by a factor of 2.5 and degraded with Gaussian noise of $\sigma = 1$ gray level (top left, upsampled by interpolation). This process was performed four times, with the sampling grids either aligned or shifted by half a pixel pitch, yielding four low-resolution images (LRI). The different shifts simulate different parallax in a four-camera setup. Images were then reconstructed from all four LRIs with the algorithm from Section 6. Either square or anamorphic pixels (aspect ratios of 1.3 and 1.56) were used for the camera model. The inset numbers are mean SSIM metrics comparing the reconstructed image to the original. While quality depends on shift for square pixels, it is independent of shift for anamorphic pixels.

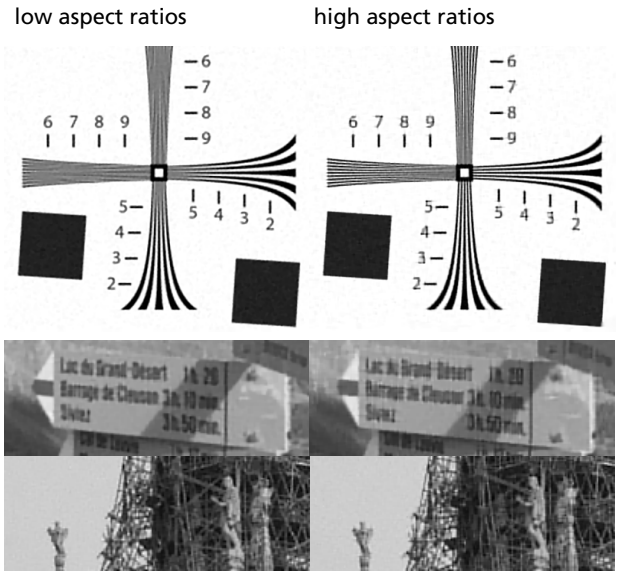

Figure 11. A simulation with a resolution chart (top two images) shows that 4 cameras with high aspect ratio pixels (4.0 and 4.8, in this case) have a higher horizontal and vertical resolution than pairs of low-AR cameras (1.3 and 1.56). Diagonal resolution is reduced, however, leading to artifacts in details that are neither horizontal nor vertical, which occur in most natural images (examples in the two rows below the test charts).

Additionally, we explored the performance of anamorphic pixels with different magnification factors (Figure 15) and numbers of cameras (Figure 16). Performance scales similarly to square pixels, as long as the cameras are shifted carefully. When the cameras are co-located, sharpness does not increase for systems with square pixels; only noise is reduced slightly.

Finally, we found that when reconstructing from a low number of LRIs and noise level is low, results with anamorphic pixels are superior to those with square pixels, even when the latter are shifted optimally (Figure 14). With increasing noise, the performance gap closes, but even at high noise levels, sharpness remains comparable.

8. Experiments with anamorphic lenses

Next, we reconstruct images taken with a digital camera. The proposed effect of avoiding sample area overlap is now accomplished by modifying the imaging lens. When the objective lens has different reproduction scale in X and Y direction, the sampling area of a square pixel becomes elongated. Such a lens is typically called an anamorphic lens[2]. With this type of lens, individual cameras can be given distinct shapes and pitches of their sampling areas,

[2]Anamorphic lenses are commonly used in cinematography to widen the horizontal field of view. A 1:1.3 anamorphic lens, for example, captures a 16:9 FOV on an image sensor with a 4:3 aspect ratio. Other optical

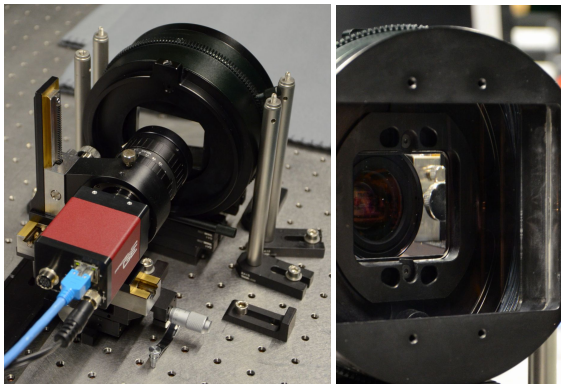

Figure 12. Experimental setup with camera, lens and anamorphic adapter (left). Looking through the adapter, the round lens of the camera is distorted into an oval (right).

even when these cameras have standard image sensors with square pixels.

We recorded a flat lab scene with an industrial camera, the Allied Vision GE1050 with 5 μm pixel pitch (Figure 12). Instead of an anamorphic lens, we used a conventional spherical lens and an anamorphic adapter. The lens was a Fujinon HF16SA-1, set to F4 for maximum sharpness. The anamorphic adapter consists of two cylindrical lens elements. It can be attached to the front of a standard objective lens, widening its horizontal FOV. Commercially avaliable models have nominal aspect ratios of either 1:1.33 or 1:2. From Figure 8, we chose the former. The actual AR at the scene distance was measured as 1.22.

We took exposures from different viewpoints, spaced 1 mm apart. From each position, we shot with and without the anamorphic adapter. At all positions, the adapter was also rotated 90° to achieve two distinct sampling areas. The recorded images were first downsampled by a factor of 4 to obtain the LRIs. Then, we applied superresolution with a factor of 2.

In horizontal position, an anamorphic adapter increases the FOV horizontally, also increasing the width of the sampling area on the object. The height is not affected. In total, the integration area of each sample increases, leading to a lower number of samples on the same object area. To make sure that LRIs taken with and without anamorphic adapter have the same number of samples and show the same part of the scene, we moved the camera backwards for the exposures without the adapter. The distance to the object has to be increased by a factor of $\sqrt{\text{AR}}$.

Figure 13 shows details from the full images. First, we can see that reconstruction is always sharper than simple interpolation from one image. In the worst case, however, a superresolved image from two LRIs without adapter is identical to the reconstruction from just one image. When the anamorphic adapter was used, the version calculated from two images is always superior.

effects, such as different resolution and depth of field in X and Y direction, are unintended byproducts.

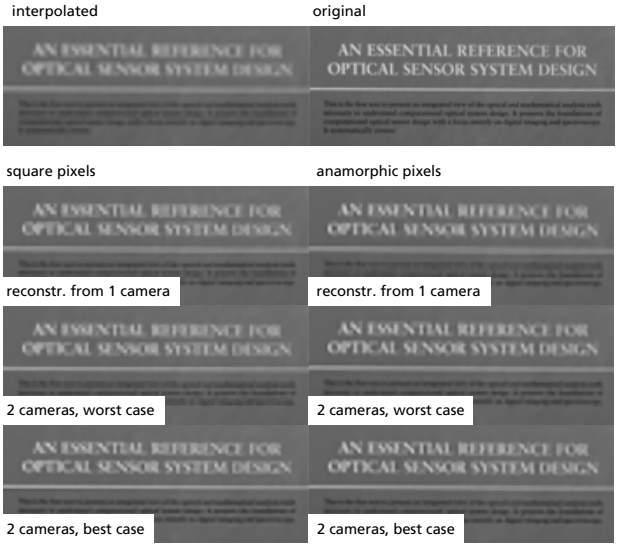

Figure 13. Superresolution from photographs of a natural scene. Photos taken either with or without an anamorphic adapter, yielding square or anamorphic sampling areas. Reconstruction from image pairs taken from shifted camera positions. For square pixels, reconstruction quality depends on shift; anamorphic pixels reduce this dependency. Please refer to the supplemental material for the full images.

9. Conclusion

We have shown that a custom design of the image sensor pixels simplifies the superresolution problem with multi-view systems. Cases where resolution drops because of sampling grid alignment are avoided. When each channel or camera has a unique, non-square pixel shape, the linear system is better conditioned and the reconstruction is more robust, leading to better resolution and fewer artifacts.

Our findings apply to camera arrays and microscopic multi-aperture systems as well as plenoptic cameras. Our solution can be implemented with technology that is readily available, either with a modified image sensor or with anamorphic optics. By making superresolution more practical for multi-view systems, achieving adequate resolutions is now feasible and these systems themselves become more useful for traditional photography.

In the future, we hope to implement our solution with microoptical anamorphic lenses, exploring the relation between geometric distortion and image reconstruction in multi-view systems. We also plan to extend our reconstruction techniques to arbitrary scenes and reconstruct depth-maps along with high-resolution images.

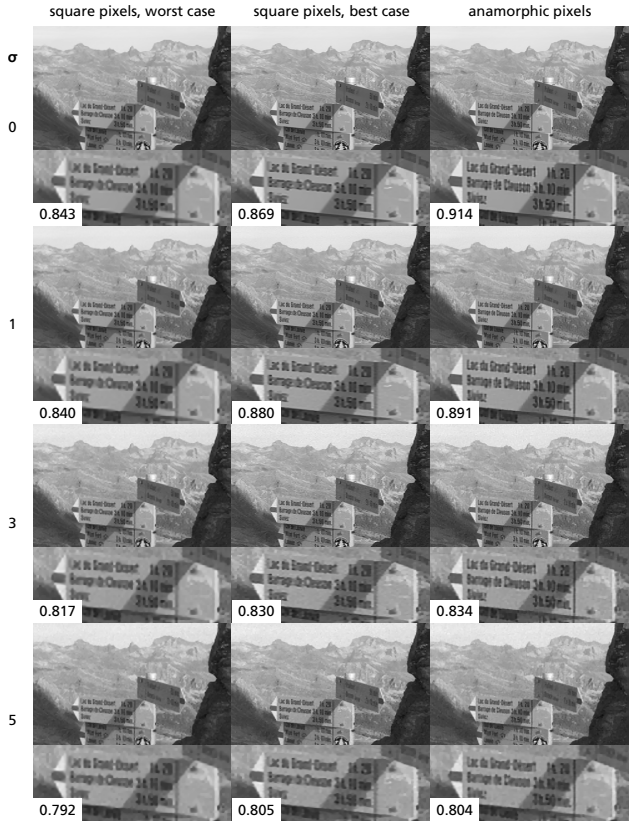

Figure 14. Reconstruction for 4 different noise levels ($\sigma = 0$ to 5 gray levels). Magnification of 2.5, two cameras, anamorphic pixels with an aspect ratio of 1.3. Anamorphic pixels outperform square pixels significantly for low noise levels. While noise compromises reconstruction quality in general and therefore closes the gap between different sampling schemes, results with anamorphic pixels remain comparable to square pixels even in the optimal case.

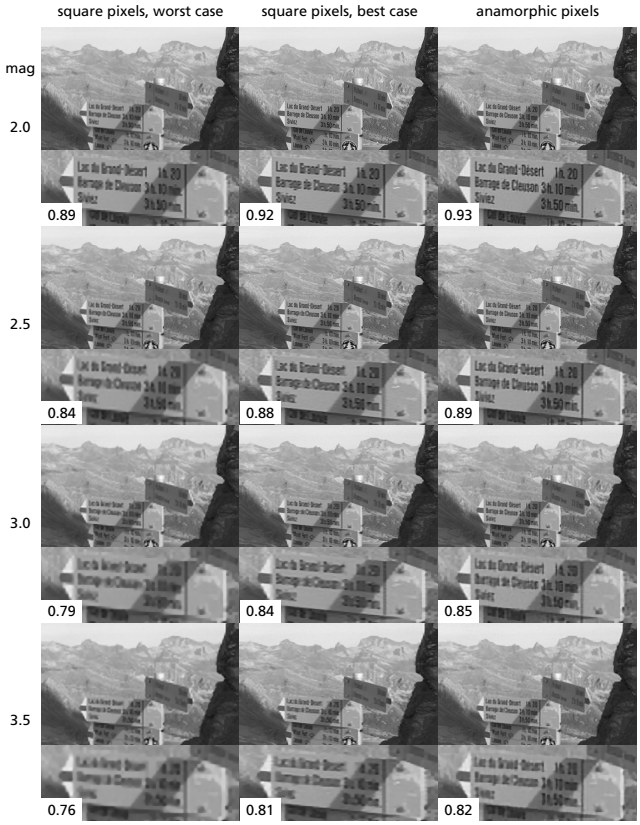

Figure 15. Reconstruction for 4 different magnification factors. Noise of $\sigma = 1$ gray level, two cameras (anamorphic pixels with aspect ratio of 1.3). Anamorphic pixels show a scaling behaviour that is similiar to square pixels (best case).

Acknowledgements

I would like to thank Prof. Andreas Tünnermann and Dr. Andreas Bräuer for the opportunity to stay at the University of Bern for researching non-regular pixel layouts. During my time there, Dr. Anita Sellent was always available for helpful discussions.

References

[1] E. Adelson and J. Wang. Single lens stereo with a plenoptic camera. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 14(2):99–106, 1992.

[2] M. Ben-Ezra, Z. Lin, B. Wilburn, and W. Zhang. Penrose pixels for super-resolution. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(7):1370–1383, 2011.

[3] T. Bishop and P. Favaro. The light field camera: Extended depth of field, aliasing, and superresolution. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(5):972–986, 2012.

[4] A. Brückner, J. Duparré, R. Leitel, P. Dannberg, A. Bräuer, and A. Tünnermann. Thin wafer-level camera lenses inspired by insect compound eyes. *Opt. Express*, 18(24):24379–24394, Nov 2010.

[5] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imaging Vis.*, 40(1):120–145, May 2011.

[6] M. Damghanian, R. Olsson, M. Sjöström, H. Navarro Fructuoso, and M. Martinez-Corral. Investigating the lateral resolution in a plenoptic capturing system using the SPC model. In *Proc. SPIE*, volume 8660, pages 86600T–86600T–8, 2013.

[7] C. B. Duane. Close-range camera calibration. *Photogrammetric engineering*, 37(8):855–866, 1971.

[8] J. Hays and A. A. Efros. im2gps: estimating geographic information from a single image. In *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[9] R. Horisaki, K. Kagawa, Y. Nakao, T. Toyoda, Y. Masaki, and J. Tanida. Irregular lens arrangement design to im-

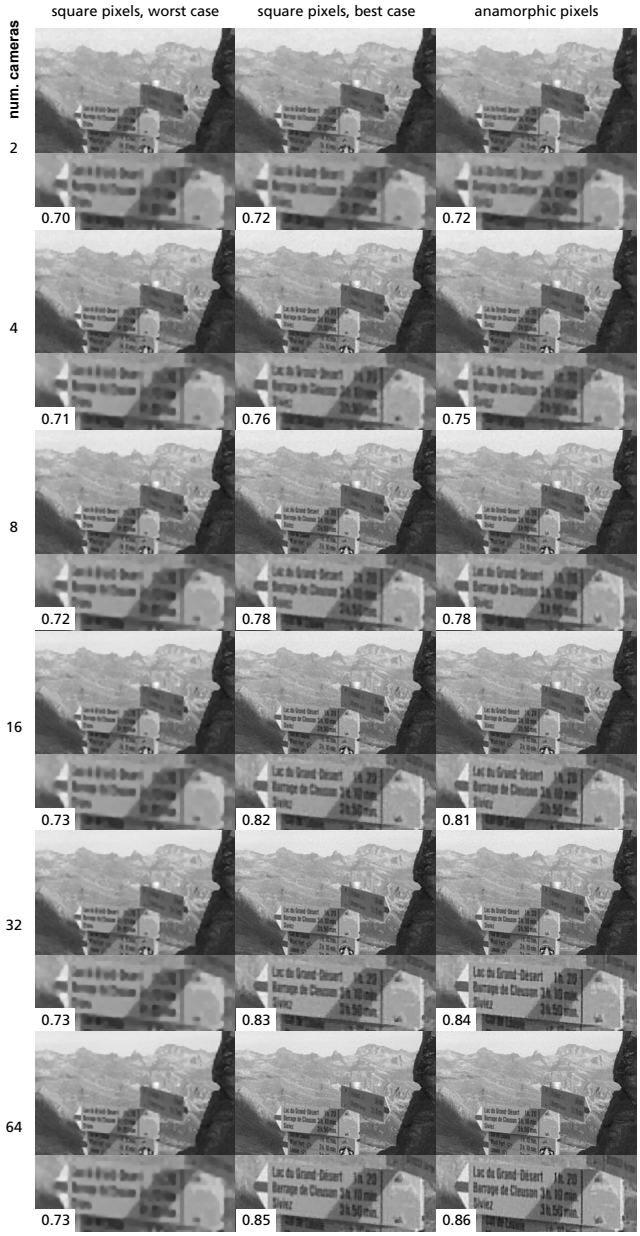

Figure 16. Reconstructing from a large number of low-resolution images to achieve a large magnification factor of 4.0 despite a high noise level of $\sigma = 5.0$. When cameras use square pixels and are co-located, adding cameras only decreases noise, but does not improve resolution. The result with 64 cameras is therefore similar to the 2-camera result (left column). For performance comparable to anamorphic pixels (right column), cameras with square pixels have to be carefully shifted so that sampling areas overlap as little as possible (middle column). For the anamorphic results, the first camera pair had an aspect ratio of 1.3; for each additional pair, the AR was multiplied by 1.04.

prove imaging performance of compound-eye imaging systems. *Applied Physics Express*, 3(2):022501, 2010.

[10] A. V. Kanaev, J. R. Ackerman, E. F. Fleet, and D. A. Scribner. TOMBO sensor with scene-independent superresolution processing. *Opt. Lett.*, 32(19):2855–2857, Oct 2007.

[11] Z. Lin and H.-Y. Shum. Fundamental limits of reconstruction-based superresolution algorithms under local translation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(1):83–97, 2004.

[12] G. Lippmann. Épreuves réversibles donnant la sensation du relief. *Journal de Physique*, 7(4):821–825, 1908.

[13] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. *Stanford Tech Report CTSR 2005-02*, 2005.

[14] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, May 2003.

[15] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268, Nov 1992.

[16] T. Sasao, S. Hiura, and K. Sato. Super-resolution with randomly shaped pixels and sparse regularization. In *Computational Photography (ICCP), 2013 IEEE International Conference on*, pages 1–11. IEEE, 2013.

[17] M. Schöberl, J. Seller, S. Foessel, and A. Kaup. Increasing imaging resolution by covering your sensor. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 1897–1900. IEEE, 2011.

[18] J. Tanida, T. Kumagai, K. Yamada, S. Miyatake, K. Ishida, T. Morimoto, N. Kondou, D. Miyazaki, and Y. Ichioka. Thin observation module by bound optics (TOMBO): Concept and experimental verification. *Appl. Opt.*, 40(11):1806–1813, Apr 2001.

[19] V. Vaish, B. Wilburn, N. Joshi, and M. Levoy. Using plane + parallax for calibrating dense camera arrays. *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:I2–I9, Jun 2004.

[20] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy. High performance imaging using large camera arrays. *ACM Trans. Graph.*, 24:765–776, Jul 2005.