Sub-Pixel Layout for Super-Resolution with Images in the Octic Group

Boxin Shi[*†], Hang Zhao[*], Moshe Ben-Ezra[*], Sai-Kit Yeung[†],
Christy Fernandez-Cull[‡], R. Hamilton Shepard[‡],
Christopher Barsi[*], and Ramesh Raskar[*]

[*]MIT Media Lab, [†]Singapore University of Technology and Design, [‡]MIT Lincoln Lab

Abstract. This paper presents a novel super-resolution framework by exploring the properties of non-conventional pixel layouts and shapes. We show that recording multiple images, transformed in the octic group, with a sensor of asymmetric sub-pixel layout increases the spatial sampling compared to a conventional sensor with a rectilinear grid of pixels and hence increases the image resolution. We further prove a theoretical bound for achieving well-posed super-resolution with a designated magnification factor w.r.t. the number and distribution of sub-pixels. We also propose strategies for selecting good sub-pixel layouts and effective super-resolution algorithms for our setup. The experimental results validate the proposed theory and solution, which have the potential to guide the future CCD layout design with super-resolution functionality.

Keywords: Super-resolution, CCD sensor, Sub-pixel layout, Octic group

1 Introduction

High-resolution imaging is a goal commonly desired for many applications in computer vision. To overcome the upper limit of spatial frequency, determined by the interval between sampling points of the sensor, super-resolution (SR) can be performed by taking multiple frames with sub-pixel displacements of the same scene. However, even with a large number of images under sufficiently small-step displacements, the performance of SR algorithms can hardly extend beyond small magnification factors of 2 to 4 [1, 12], partially due to the challenges associated with proper alignment of local patches with finer translations [19].

We note here that the geometry of the pixels usually has been assumed to lie on a rectangular grid. This geometric restriction limits the information captured for each sub-pixel shift, because repeating the observations at integer pixel intervals are redundant. An aperiodic pixel layout [2] or a random disturbance to pixel shapes [17] could break the theoretical bottleneck of conventional SR by effectively avoiding the redundancy due to translational symmetry. These structures provide greater variation with sub-pixel displacements to result in more independent equations for recovering high-frequency spatial information.

This paper explores the properties of non-conventional pixel layouts and shapes. Some existing CCDs contain sub-pixels of different shapes and spatial locations within one pixel. We show two examples in the bottom row of Fig. 1. The
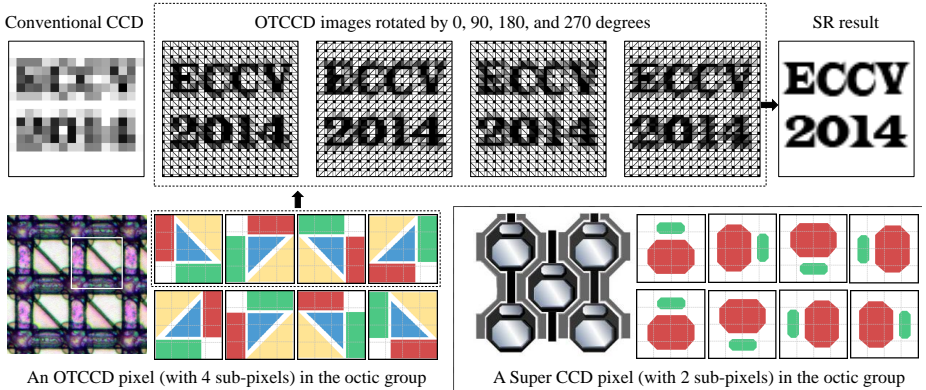
Fig. 1. Existing CCD sensors with non-conventional sub-pixel layouts (bottom row). By transforming the image plane, their images in the octic group show different sub-pixel layouts, which can be combined for significant resolution enhancement (top row). The OTCCD images under 4 rotations can perform 4× SR. Comparing the OTCCD with the super CCD, it can be seen that sub-pixels with an asymmetric layout produce more variation in their images in the octic group.

Orthogonal-Transfer Charge-Coupled Device (OTCCD) sensor [4] has four sub-pixels[1], and the super CCD [10] has two[2]. These sub-pixels naturally increase the spatial sampling rate. Instead of relying on sub-pixel displacements, however, we focus on forming multiple images via transformations in the octic group, *i.e.*, all symmetries of a square. We assume the pixel shape is square, so that each element in the octic group corresponds to one pose of a pixel. The sub-pixel layout varies with different poses of a pixel, and depends on the layout's symmetry. For example, the OTCCD can form 8 different sub-pixel layouts (through four 90° rotations and their reflections), but the super CCD has left-right symmetry and therefore shows only 4 different layouts. By combining multiple images recorded with different poses, a super-resolved image with higher resolution can be obtained. The intuition here is that more sub-pixels with asymmetric layouts can construct a higher resolution image. We discuss here the exact relationship between sub-pixel layout (including the number and distribution of sub-pixels) in the octic group and the magnification factor.

1.1 Contributions

Our key contributions are summarized as below:

---

[1] The OTCCD actually consists four phases in one pixel. Photon charge can integrate separately in each phase and shift between the phases. Here we interpret the four "phases" as four "sub-pixels".

[2] If we treat the gap among sub-pixels as another sub-pixel that does not record photon charges, the number of sub-pixels could be five for the OTCCD and three for the super CCD.

– Our new framework provides a novel view to the SR problem by using an asymmetric sub-pixel layout to form multiple mages in the octic group. Instead of focusing on a particular layout, we investigate the theoretical bound of SR performance w.r.t. the number and distribution of sub-pixels (Sec. 2.2).
– Based on the theoretical analysis, we propose a sub-pixel layout selection algorithm to choose good layouts for well-posed and effective SR (Sec. 2.3).
– We propose a simple yet effective SR reconstruction algorithm (Sec. 3) and validate our theory and algorithm using both synthetic and real-world data (Sec. 4).

1.2   Related work

Our approach belongs to the category of reconstruction-based SR with multiple images. SR algorithms using single images such as learning-based methods (*e.g.*, [9]) are beyond the scope of this paper. We refer the readers to survey papers (*e.g.*, [16]) for a discussion of various categories of SR algorithms.

For regular pixel layouts and shapes, there are various SR reconstruction methods for images with sub-pixel displacements. Popular approaches include iterative back projection (IBP) [11], maximum a posteriori and regularized maximum likelihood [6] and sparse representation [18]. These reconstruction techniques focus on solving the ill-posed inverse problem with a conventional sensor and setup.

In contrast, this paper studies asymmetric sub-pixel layouts and is therefore similar to previous techniques using non-conventional pixel layouts [2] and pixel shapes [17]. The former work used a Penrose pixel layout, which never repeats itself, on an infinite plane. The latter work implemented random pixel shapes by spraying fine-grained black powder on the CCD. Both methods focus on one type of layout or shape and use multiple images with sub-pixel displacements. Our work is different from them in two ways: 1) We transform the image plane to form multiple images in the octic group; 2) We propose a general theory and categorize good sub-pixel layouts for deeper understanding of SR performance with non-conventional pixels.

2   Good Sub-pixel Layout for Super-Resolution

2.1   Single image case

Similar to Penrose tiling [2], we ignore optical deblurring and assume that it can be applied after sub-pixel sampling. Thus, in the discrete domain, reconstruction-based SR can be represented as a linear system as

$$\mathbf{L} = \mathbf{PH} + \mathbf{E}, \tag{1}$$

where $\mathbf{H}$ includes all pixels of a high resolution (HR) image in a column vector, $\mathbf{L}$ concatenates column vectors formed by all low resolution (LR) images, $\mathbf{P}$ is the matrix that maps HR to LR images, and $\mathbf{E}$ is the per-pixel noise.
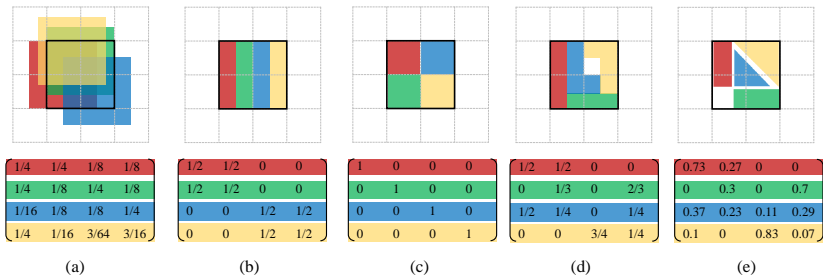
Fig. 2. Different sub-pixel layouts and their HR-to-LR mapping matrices **P**, with magnification factor $\mathcal{M} = 2$. (a) Conventional SR: LR pixels with double the size as HR pixels undergo a sub-pixel displacement; (b)-(e): examples of different sub-pixel layouts. Different color-shaded areas (RGBYW here) represent different sub-pixels, and the white area within one pixel is to simulate a gap that does not record photon charges.

In the ideal case when noise can be ignored, to double the resolution ($2\times$ SR), we need at least 4 LR images with exactly half-pixel shifts to produce a full reconstruction (the inverse problem is well-posed). In general, the displacements of LR images can be arbitrary, and they determine the values in **P**. An example of **P** is shown in Fig. 2(a). The HR grid is drawn with dashed lines, and the shaded squares with different colors represent LR pixels from different images. In this example, **P** is evaluated for the 2×2 area indicated by the bold black square (values out of this area are not shown). Each row of **P** corresponds to one displaced-LR pixel (shaded area with the same color), and the element in each row is calculated for all HR pixels (bold black square in Fig. 2) as the area ratio of overlapping regions to the LR pixel size. The analysis of **P** plays a key role in understanding the performance of SR.

Similar to sub-pixel displacement with multiples images, the increase in spatial sampling can also be implemented by splitting one LR pixel into smaller sub-pixels with a single image. The most straightforward example for the $2\times$ SR is splitting one LR pixel into 4 square regions, as shown in Fig. 2(c). In such a case, **P** is an identity matrix. By treating each sub-pixel as a displaced LR pixel, we can build **P** for sub-pixel layouts in Fig. 2(b), (d), and (e) in a similar way as Fig. 2(a). Note the layout in (b) has rank(**P**) = 2, so it cannot produce $2\times$ SR. The layouts in (d) and (e) have rank(**P**) = 4, so they can achieve $2\times$ SR. For easy analysis, we assume the sub-pixels completely cover one LR pixel, so the layouts in Fig. 2(d) and (e) actually have 5 sub-pixels. We treat the gap among sub-pixels as a dumb sub-pixel that does not record photon charges; therefore, strictly speaking, **P** for layouts (d) and (e) should have an all-zero row, which is omitted in the figure.

In general, the size of **P** equals to $r \times \mathcal{M}^2$, where r is the number of sub-pixels, and $\mathcal{M}$ as the magnification factor. It is easy to infer that for a single pixel with r sub-pixels, to achieve $\mathcal{M}\times$ SR, the sufficient condition for full reconstruction is
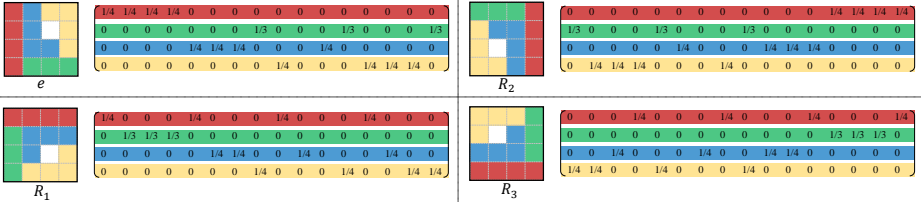
Fig. 3. Sub-pixel layouts with $r = 5$ (the gap among sub-pixels is a dumb sub-pixel) in a sub-group of the octic group $\hat{\mathcal{G}} = \{e, R_1, R_2, R_3\}$. These layouts could build a **P** with rank(**P**) = 16. Four images captured with such sub-pixel layouts can be used to perform $4\times$ SR.

when $r \geq \mathcal{M}^2$. Because the full reconstruction is achieved when rank(**P**) = $\mathcal{M}^2$ and **P** has the size of $r \times \mathcal{M}^2$, rank(**P**) < $\mathcal{M}^2$ holds if $r < \mathcal{M}^2$.

2.2 Multiple images in the octic group

Enhancing the resolution by only using sub-pixels in one image has limited performance (requires $r \geq \mathcal{M}^2$). Further, in practice, increasing the sub-pixel number cannot continue indefinitely, due to manufacturing limitations and the proportionality between pixel size and light collection efficiency (*i.e.*, signal-to-noise ratio (SNR) decreases with pixel size). Combining different sub-pixel layouts for one pixel can further enhance the resolution, but physically modifying the layout in a fabricated sensor is cost prohibitive. Instead, we observe that simple operations on the image plane can serve to change the sub-pixel layouts, if we make multiple images to form the octic group.

Octic group In group theory, a square belongs to the octic group, which is the 4-th order dihedral group. This group contains 8 components that keep all symmetric properties of a square, denoted as

$$\mathcal{G} = \{e, R_1, R_2, R_3, S_e, S_{R_1}, S_{R_2}, S_{R_3}\}, \tag{2}$$

where e represents the original pose; R_1, R_2 and R_3 represent 90°, 180° and 270° rotations of the original pose; and S_e, S_{R_1}, S_{R_2}, and S_{R_3} represent the reflections (horizontal or vertical mirror flipping) to the first 4 elements, respectively. These 8 poses can transform into each other according to the multiplication table of the octic group.

An intuitive example We show an intuitive example in Fig. 3. Given a rotation-asymmetric sub-pixel layout with $r = 5$ (4 effective sub-pixels and 1 dumb sub-pixel), the maximum $\mathcal{M}$ allowed for such a structure is 2 (see Fig. 2(c)-(e)) for a single image. We denote $\hat{\mathcal{G}} = \{e, R_1, R_2, R_3\}$ as the sub-group of $\mathcal{G}$ with the first 4 elements. By rotating the image plane three times with a step of 90°,

we obtain 4 images in group $\hat{\mathcal{G}}$. Similar to Fig. 2, we reconstruct $r-1$ (exclude the gap) rows of $\mathbf{P}$ for each pose in $\hat{\mathcal{G}}$, and by stacking the layouts with all 4 poses we obtain $\mathbf{P}$. Here rank($\mathbf{P}$) = 16, so it is well-posed for full reconstruction of $4\times$ SR.

We assume the image plane is square and all pixels are congruent squares, then all images in the octic group will have their pixel contours exactly overlapped with different sub-pixel layouts inside. This makes the following analysis and SR reconstruction independent of pixel locations.

Full reconstruction conditions The full reconstruction of SR is determined by rank($\mathbf{P}$). The structure of $\mathbf{P}$ is determined by various factors: the size and distribution of sub-pixels, denoted as Γ; the number of sub-pixels r; the number of elements in $\mathcal{G}$ or its subgroup, denoted as t (it is equal to the number of different images used for SR); and the magnification factor $\mathcal{M}$. We denote P as the function constructing $\mathbf{P}$: $\mathbf{P} = P(\Gamma, r, t, \mathcal{M})$. Assuming we have found a Γ that satisfies rank($\mathbf{P}^\Gamma$) = argmax$_\Gamma$ rank($P(\Gamma, r, t, \mathcal{M})$) given a fixed combination of $(r, t, \mathcal{M})$, the exact value of rank($\mathbf{P}^\Gamma$) depends on $(r, t, \mathcal{M})$. According to Fig. 2 and Fig. 3, the intuition is that Γ should be a rotation/reflection-asymmetric sub-pixel layout. In this paper, we restrict the discussion to two different t values: $t = 4$ means 4 images in the group $\hat{\mathcal{G}}$ (only rotations), and $t = 8$ means 8 images that form the group $\mathcal{G}$ (rotations and reflections). With these constraints on Γ and t, we explore the relationship between r and $\mathcal{M}$.

1) For small $\mathcal{M}$: If $\mathcal{M}^2 << rt$, the upper bound U_1 of rank($\mathbf{P}^\Gamma$) is determined by $\mathcal{M}$ as $U_1 = \mathcal{M}^2$. This is understood by noting that $\mathbf{P}^\Gamma$ has a size of $rt \times \mathcal{M}^2$. But, this case is not very meaningful for practical applications, since people expect larger $\mathcal{M}$ with smaller r and t.

2) For large $\mathcal{M}$: If $\mathcal{M}^2 >> rt$, the upper bound U_2 of rank($\mathbf{P}^\Gamma$) is determined by the values of rt. Unfortunately, rank($\mathbf{P}^\Gamma$) might not reach the maximum number of rows of $\mathbf{P}^\Gamma$, which is rt, because of some linear dependence across the rows of $\mathbf{P}^\Gamma$. For example, the layout in Fig. 3 has $r = 5$ and $t = 4$, and Lemma 1 below explains that it is impossible to produce rank($\mathbf{P}^\Gamma$) = 16 with only $r = 4$.

Lemma 1 *Given a group of pixels with t poses in $\mathcal{G}$, with each pixel containing r sub-pixels, for a sufficiently large $\mathcal{M}$, the upper bound of* rank($\mathbf{P}^\Gamma$)*, denoted as U_2, is $U_2 = t(r-1) + 1$.*

Proof. A sufficiently large $\mathcal{M}$ means the HR pixel is quite small comparing to the LR pixel. So we can assume that each sub-pixel covers several integer HR pixels (*e.g.*, the example in Fig. 3). Set the image plane as its original pose, and assume the i-th ($1 \leq i \leq r$) sub-pixel has an area of a_i by covering a_i unit-area HR pixels. Then, the i-th row of $\mathbf{P}^\Gamma$ denoted as $\mathbf{P}^\Gamma_{i*}$ contains a_i elements with value of $\frac{1}{a_i}$ and all other elements of 0. Given $r-1$ such rows, and a $1 \times \mathcal{M}^2$ row vector $\mathbf{I}$ with all values as 1, we can represent the r-th row as:

$$\mathbf{P}^\Gamma_{r*} = \frac{1}{\mathcal{M}^2 - \sum_{i=1}^{r-1} a_i} \left(\mathbf{I} - \sum_{i=1}^{r-1} a_i \mathbf{P}^\Gamma_{i*} \right). \tag{3}$$

According to the composition of $\mathbf{P}^\Gamma$, each sub-matrix of r rows corresponds to one image with a pose from $\mathcal{G}$. Therefore, the i-th row and the $(i+kr)$-th row ($1 \leq k \leq t-1, k \in \mathbb{Z}$) have the same values permutated to different columns (compare all rows with the same color in Fig. 3). Then, the r-th row in each sub-matrix can also be calculated by using $\mathbf{I}$ and Eq. (3).

Finally, $\mathbf{P}^\Gamma$ is concatenated by t sub-matrices of $r-1$ rows, plus another row vector $\mathbf{I}$. Thus, its maximum rank is equal to its number of rows $t(r-1)+1$. $\square$

Combining the inequality relationships above, we naturally come up with the following proposition about the upper bound of rank$(\mathbf{P}^\Gamma)$.

Proposition 1 *Given a group of pixels with t poses in $\mathcal{G}$ with each pixel containing r sub-pixels, for a designated magnification factor $\mathcal{M}$, the value of* rank$(\mathbf{P}^\Gamma)$ *is bounded as*

$$\text{rank}(\mathbf{P}^\Gamma) \leq \min(U_1, U_2) = \min(\mathcal{M}^2, t(r-1)+1). \tag{4}$$

Validation If $\mathcal{M}^2 \approx rt$, rank$(\mathbf{P}^\Gamma)$ might have a value below the upper bound of Proposition 1, but the exact value is very difficult to write as a closed-form solution, because rank maximization is a highly nonlinear problem. We use numerical simulation to plot these exact values and verify Proposition 1.

We randomly select r positions within one pixel area as centers and expand these centers in all 8 discrete directions. The expanding process is stopped when the whole pixel is filled. The pixel is transformed to different poses and forms a group $\mathcal{G}$. Then $\mathbf{P}$ is built and evaluated. This process is repeated 100 times to avoid symmetric sub-pixel layouts. We empirically observe that the possibility of generating a rank-deficient (partially or completely symmetric) layout is usually less than 1%, and almost all layouts have constant rank$(\mathbf{P}^\Gamma)$.

The rank$(\mathbf{P}^\Gamma)$ value distribution with varying r and $\mathcal{M}$ is shown in Fig. 4. The top row corresponds to $\hat{\mathcal{G}}$ ($t=4$) and the bottom row shows the case for $\mathcal{G}$ ($t=8$). The exact value distribution is illustrated in the first column, and the upper bound calculated from Proposition 1 is shown in the second column. The third column is the 2D planar view of the first column. It is interesting to note that the left side of the distribution shows a parabolic shape corresponding to U_1, while the right side of the distribution shows a planar shape corresponding to U_2. For $t=4$, the exact values perfectly match the upper bound. As the number of images in the group increases, the possibility that $\mathbf{P}^\Gamma$ has more linearly dependent rows increases, so when $t=8$ some values around $\mathcal{M} \approx rt$ cannot reach the upper bound. From the similarity of (d) to (e) and their small offsets indicated by numbers in (f), it can be seen that the upper bound is quite tight.

With the analysis above, it is easy to evaluate the SR performance for a specific sensor. For the two real sensors in Fig. 1, the OTCCD has an asymmetric layout with $r=5$, it could perform $4\times$ SR with $t=4$ images in $\hat{\mathcal{G}}$ and $5\times$ SR with $t=8$ images in $\mathcal{G}$; while the super CCD has $r=3$ sub-pixels with left-right symmetric, it only performs $2\times$ SR (rank$(\mathbf{P})=8$) with both $\hat{\mathcal{G}}$ and $\mathcal{G}$.
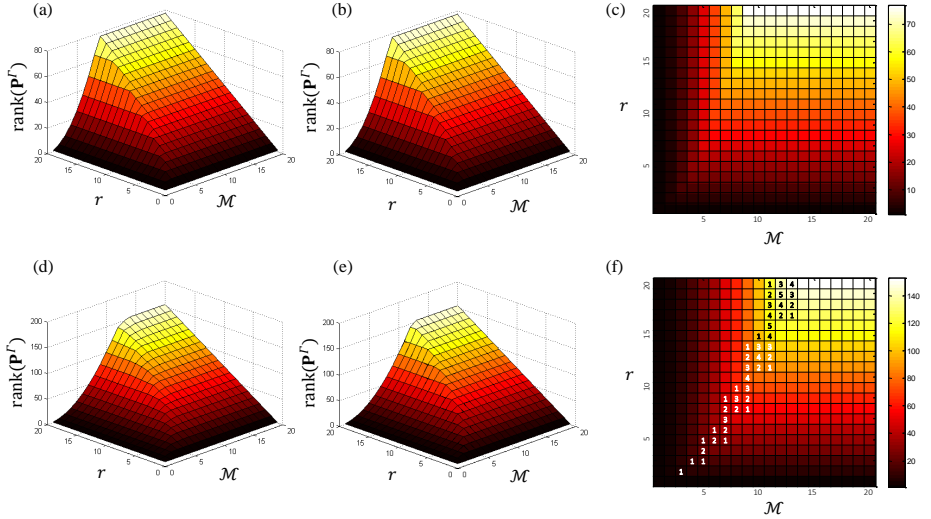
Fig. 4. Values of rank($\mathbf{P}^{\Gamma}$) varying with different r, t, and $\mathcal{M}$. Top row: $t = 4$; bottom row: $t = 8$. (a) and (d) are exact values from simulation; (b) and (e) are upper bound from Proposition 1; (c) and (f) are 2D planar views of (a) and (d). The numbers overlaid on the matrix area of (f) indicate difference from (d) to (e) (cells without numbers mean the upper bound is reached).

2.3   Good sub-pixel layout

The theoretical analysis in Sec. 2.2 explains the relationship of $(r, t, \mathcal{M})$ by assuming a good layout Γ has been found. We propose four merits to select good sub-pixel layout Γ, from randomly generated candidates. The first and most important one is to ensure the full SR reconstruction as Proposition 1: 1) With t images in a group, the pixel should contain at least r sub-pixels to ensure rank($\mathbf{P}^{\Gamma}$) = $\mathcal{M}^2$, which we call full-rank layouts (note that there can be infinite many solutions for full-rank layouts Γ).

Three additional constraints benefiting the sensor layout design and SR performance should be considered among candidates with full-rank layouts: 2) We set the sub-pixel with smallest area as a dumb sub-pixel (or gap), so we actually use only $r - 1$ effective sub-pixels to achieve the same performance of r sub-pixels. We do not use larger sub-pixels as the dumb one to maximize the size of effective sub-pixels for capturing more light. 3) The layouts with smaller sub-pixel area variance are preferred. Because our goal is to increase the spatial sampling rather than the dynamic range, sub-pixels with approximately equal areas will perform similarly in receiving light, thus too bright or too dark sub-pixels are easily avoidable. 4) We prefer $\mathbf{P}^{\Gamma}$ with smaller condition number, denoted as cond($\mathbf{P}^{\Gamma}$), which makes the inverse problem better-conditioned under noise. Considering the above four merits, we propose the good sub-pixel layout selection method in Algorithm 1.

Algorithm 1 Select good sub-pixel layout.

Input: t and $\mathcal{M}$.

Output: The good layout Γ.

1: Determine r according to Proposition 1 and Fig. 4;
2: Generate sufficiently many (> 10000) random sub-pixel layouts;
3: Label the sub-pixel with smallest area using 0 (dumb), and other sub-pixels using $\{1, 2, \cdots, r-1\}$; Calculate **P** based on the labels and areas of sub-pixels;
4: Remove rank-deficient (rank(**P**) $< \mathcal{M}^2$) layouts and keep the remaining layouts as Γ candidate set;
5: Sort current Γ candidates according to their sub-pixel area variance and remove layouts with larger variance (keep only smallest 10%);
6: Choose the layout with smallest cond($\mathbf{P}^\Gamma$) as Γ.

---

3 Reconstruction Algorithm

For r sub-pixels and t images in the octic group (or its sub-group), we have rt observations for each pixel location[3]. By concatenating these observations, we obtain the LR observations **L**. **P** is determined by the sub-pixel layout and image poses in the octic group as described in Sec. 2. For good layouts Γ with proper r and t, rank(**P**) is equal to $\mathcal{M}^2$. Therefore, the HR image **H** can be easily recovered by solving the linear system in Eq. (1).

The reconstruction is performed independently for each pixel by solving the linear least squares (ℓ^2) with a Tikhonov regularization term, denoted as

$$\underset{\mathbf{H}}{\arg\min} \|\mathbf{PH} - \mathbf{L}\|_2^2 + \lambda \|\mathbf{H}\|_2^2, \tag{5}$$

where λ is the weight of regularization term. This problem can be solved by using the LSQR method in [15].

When the noise is stronger, the problem can also be solved by minimizing the total variation (TV) with quadratic constraints as

$$\underset{\mathbf{H}}{\arg\min} \, \mathrm{TV}(\mathbf{H}) \;\; \text{subject to} \;\; \|\mathbf{PH} - \mathbf{L}\|_2^2 < \epsilon, \tag{6}$$

where ϵ is the constraint relaxation parameter. We solve the above problem using "ℓ^1-Magic" [5]. This approach needs more computation, but can better suppress the noise. We empirically find that under moderate noise, the ℓ^2-based solution is accurate with far less computation. We will verify this in Sec. 4.

The modified IBP algorithm in [2] dealing with non-conventional pixel layouts and shapes can also be naturally applied to solve our problem. Similar to [2], we can apply IBP in the HR domain. The average of all LR images is used as an initial HR image. Then the iterations are performed to update the residual between LR images upsampled to the HR domain and the images resampled using our sub-pixel layouts in the octic group. Please refer to [2] for details. We will also evaluate and compare this approach in Sec. 4.

---
[3] If there is one dumb sub-pixel, the effective size of **P** could be $(r-1)t$ by removing rows with all elements as 0.
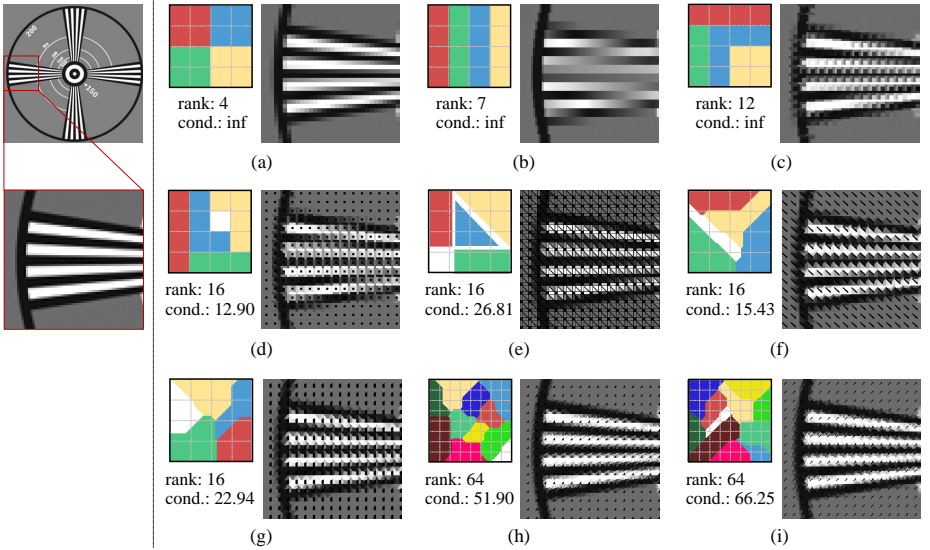
Fig. 5. Simulated images under various sub-pixel layouts for $4\times$ SR. The rank and condition number of their corresponding $\mathbf{P}$ are indicated below the sub-pixel layout. (a)-(c): $r = 4$; (d)-(g): $r = 5$; (h), (i): $r = 11$. For each layout, white means the dumb sub-pixel, and other colors indicate other effective sub-pixels.

4   Performance Evaluation

4.1   Synthetic test

Sub-pixel layouts We show image appearances under sensors with different sub-pixel layouts in Fig. 5. We model spatial integration of photon charges by using a box function by overlaying the sensor plane on the HR grid and taking average values within each sub-pixel region. The layouts in the first row have $r = 4$. They cannot reach full rank for $t = 4$, because for $r = 4$ the maximum rank is only 13 according to Lemma 1. Fig. 5(d)-(g) show some full-rank layouts with $r = 5$ and $t = 4$, and they could produce $4\times$ SR; (d) is a manually designed layout; (e) is from the real structure of an OTCCD sensor (Fig. 1); (f) and (g) are generated from Algorithm 1; (h) and (i) with $r = 11$ are also generated by Algorithm 1, they could produce $8\times$ SR with $t = 8$.

SR results with different layouts We then evaluate the SR performance by using the sub-pixel layouts in Fig. 5. In addition to the three reconstruction methods introduced in Sec. 3, we also compare SR using sub-pixel shift with our sub-pixel layouts, denoted as "IBP-shift" [2]. All test images contain 8-bit quantization noise. With only quantization noise, we found solving Eq. (5), denoted as "L2Reg", and Eq. (6) (TV) give almost the same results, so we omit the results from TV-based method here. For the IBP-based method, we run the
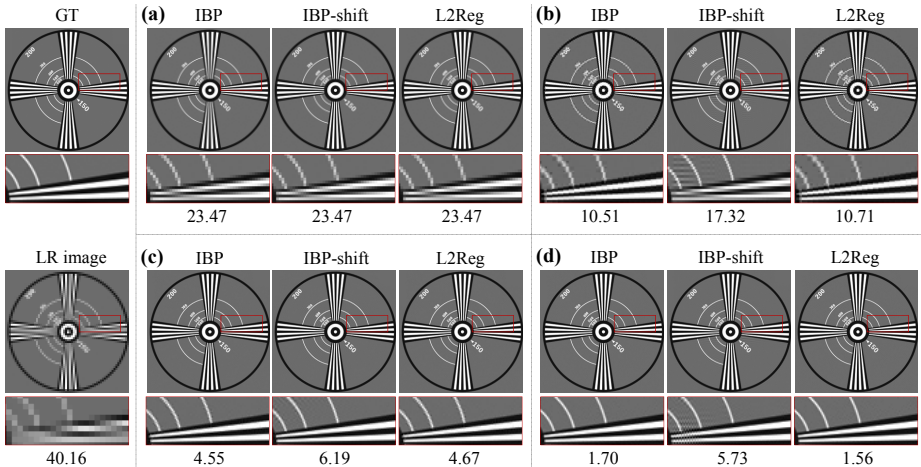
Fig. 6. SR results varying with sub-pixel layouts. The left most column shows the ground truth image and LR image observed by the conventional sensor with the same pixel size as our pixel. (a)-(d) here show 4× SR results under various sub-pixel layouts from Fig. 5(a)-(d). The number below each image is the RMSE value w.r.t. ground truth.

algorithm for 1000 iterations, and for L2Reg, we use $\lambda = 0.01$. These parameters are consistent for all of the following experiments, unless otherwise specified.

From the results in Fig. 6, we can tell that the conventional grid structure shows the worst accuracy, which actually performs 2× SR, because it keeps the layout unchanged for all images in the octic group. The layouts in (b) and (c) also show (partial) symmetric properties for different images, thus have limited enhancement in resolution. Generally, higher rank($\mathbf{P}$) produces higher resolution.

For the full-rank layouts, we show the reconstructed images using Fig. 5(d) as an example. All full-rank layouts are equivalently optimal in terms of full SR reconstruction. When there is no noise, all of them produce perfect reconstruction with RMSE = 0. Even if there is noise, these layouts produce SR images with similar appearances. There are some slight differences in RMSEs depending on the condition number of $\mathbf{P}$, *e.g.*, the SR result from layout in Fig. 5(e) has RMSE of 2.42, while Fig. 5(f) has 2.18. Fig. 5(d) has the smallest condition number, whose RMSE is also the smallest (1.56). However, this manually designed layout is not well-balanced in sub-pixel sizes.

For different reconstruction methods, L2Reg provides results similar to those of IBP. With full-rank layouts, L2Reg shows even higher accuracy. The asymmetric pixel structure also benefits the SR using sub-pixel shift (IBP-shift), but its accuracy is not as good as using images in the octic group (IBP). For a fair comparison, we evaluate only half-pixel displacement compared to our $t = 4$ rotations here. Using finer steps in shifting and more images further increases the resolution [2], but it is equivalent to using more sub-pixels with smaller sizes.
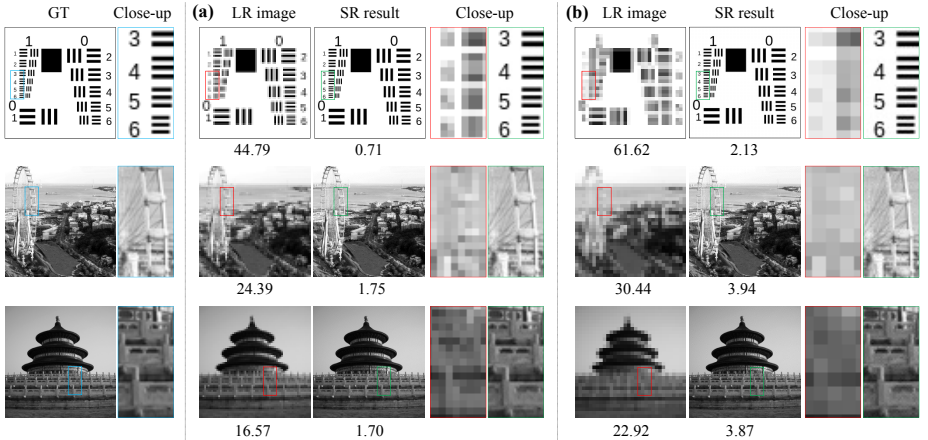
Fig. 7. (a) 4× SR results and (b) 8× SR results with full-rank sub-pixel layouts from Fig. 5(d) and Fig. 5(h). The left most column shows the ground truth image. The LR image refers to the results from a conventional sensor. Close-up views are shown in the rightmost of each column.

We show more SR results with full-rank layouts in Fig. 7 solved by L2Reg. Column (a) shows 4× SR with the layout in Fig. 5(d) and $t = 4$; column (b) shows the 8× SR results with the layout in Fig. 5(h) and $t = 8$.

Results varying with noise We show the influence of noise on the results in Fig. 8. 4× SR with full-rank layouts are evaluated by using three different reconstruction methods. We use a Matlab built-in function "imnoise" to add signal-dependent Poisson noise, which more closely models shot noise than does zero-mean Gaussian noise. We use a scaling factor η to adjust the strength of the noise[4] before quantizing the data to 8 bits.

In the presence of Poisson noise, IBP does not show good convergence, and the errors accumulate after a local minimum has been reached. To show the best results that IBP can obtain, we manually stop the iterations at 150 and 50 for the test in Fig. 8(a) and (b) (larger noise makes IBP worse in convergence), respectively. Even with manual interference, IBP still shows worse performance than ℓ^2- and TV-based methods. TV could produce reconstructions with smaller errors with noisy images. We use $\epsilon = 2$ in Eq. (6) for this experiment.

4.2   Real data test

We use a Canon EOS Rebel T3i camera to capture images with real noise. For each $\mathcal{M} \times \mathcal{M}$ area of the captured image, we create one pixel according to a

---

[4] For double-precision data, "imnoise" interprets pixel values as means of Poisson distributions scaled up by 10^{12}. To adjust the noise level, we scale the data by $\frac{1}{\eta}$ before applying "imnoise," and then scale it back by η.

GT | (a) IBP | L2Reg | TV | (b) IBP | L2Reg | TV
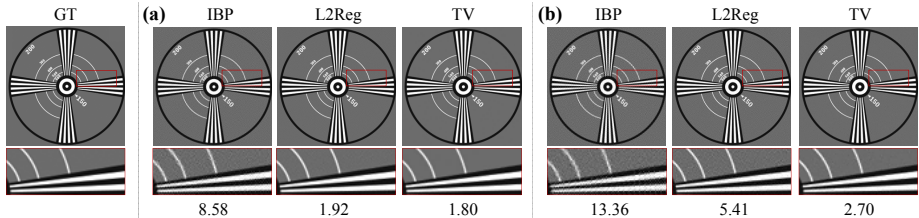
8.58 | 1.92 | 1.80 | 13.36 | 5.41 | 2.70

Fig. 8. SR results with noise. (a) Poisson noise with $\eta = 10^7$ plus 8-bit quantization; (b) Poisson noise with $\eta = 10^8$ plus 8-bit quantization. Three different reconstruction methods are compared: IBP, ℓ^2-, and TV-based methods.
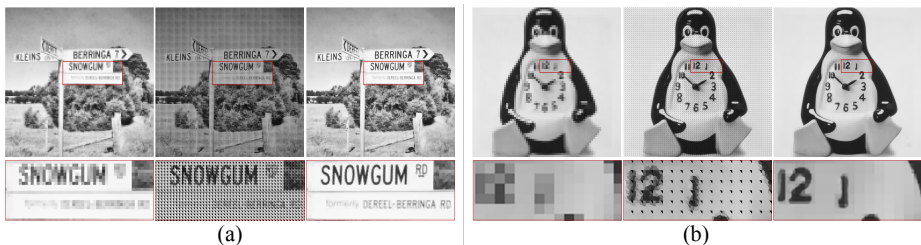

(a)   (b)

Fig. 9. SR results using real data. (a) $4\times$ SR with OTCCD sub-pixel layout in Fig. 1, (b) $8\times$ SR with our good sub-pixel layout in Fig. 5(h). From left to right: images using conventional sensor, image views from an sensor with sub-pixel layouts, and SR result.

sub-pixel layout. Here we evaluate the OTCCD layout in Fig. 1 for $4\times$ SR and the layout in Fig. 5(h) for $8\times$ SR. We manually rotate and flip the image plane in a controlled manner to obtain images in the octic group. The captured images are further registered using the method in [7].

Various noise are included in the captured images, such as the registration error, image blur, sensor noise, and JPEG compression noise, so we apply the TV-based method with $\epsilon = 30$ to reconstruct the HR images. We show the results in Fig. 9. Note that the images from sensors with several sub-pixels already have some resolution enhancement, but with multiple images in the octic group the resolution could be further increased. Even with various types of real noise, our SR results could clearly recover delicate details in the original scene.

5   Discussion

A potential hardware implementation We suggest a potential implementation for building a prototype camera to realize our SR framework. As shown in Fig. 1, there are existing CCD sensors with asymmetric sub-pixel layouts. The image rotation can be implemented by placing a Dove prism in front of the main lens, similarly as done in rotational-shearing interferometry [14, 13]. The Dove prism has the property of rotating the image plane 2θ for its own rotation of θ.

It can be controlled with great precision using a rotary engine (such as a step-per motor or an ultrasound motor used to focus lenses), which is mechanically simpler and more accurate than XY translation stages used in previous SR work [3]. This will be sufficient for realizing the $t = 4$ group. Note that all the images after a single Dove prism will be mirror-flipped, so another, cascaded Dove prism would be required to obtain all images in the octic group.

Other considerations in SR system design As compared to conventional SR that involves inter-pixel overlapping, the proposed method based on octic groups can work independently and equivalently on each pixel, which has advantages in supporting parallel computation and saving memory in encoding $\mathbf{P}$ (do not need to consider neighboring pixels) for real-time functionality.

We do not directly compare our approach with SR algorithms that use conventional sensors, because the goal of this paper is to show the condition for full reconstruction rather than developing an advanced method for solving the inverse problem. As validated in the experiments, even with simple solutions in Sec. 4 the accuracy could be very high. We believe that by using more complicated regularization terms (*e.g.*, [8]) and modern robust methods (*e.g.*, [18]), the reconstruction accuracy could be further improved under severe noise.

6   Conclusion

The key observation of this paper is that when one pixel is split into several asymmetrically distributed sub-pixels, the images in the octic group could further increase the spatial sampling. This group of images can be combined to perform super-resolution. We analyzed the theoretical bound for this setup. With proper sub-pixel layouts, SR with desired magnification factor could be accurately achieved with simple computation. We verify our theory and algorithm with both synthetic and real-world data.

Acknowledgement

References

1. Baker, S., Kanade, T.: Limits on super-resolution and how to break them. IEEE Transactions on Pattern Analysis and Machine Intelligence 24(9), 1167–1183 (2002)
2. Ben-Ezra, M., Lin, Z., Wilburn, B., Zhang, W.: Penrose pixels for super-resolution. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(7), 1370–1383 (2011)
3. Ben-Ezra, M., Zomet, A., Nayar, S.K.: Video super-resolution using controlled subpixel detector shifts. IEEE Transactions on Pattern Analysis and Machine Intelligence 27(6), 977–987 (2005)
4. Burke, B.E., Tonry, J., Cooper, M., Luppino, G., Jacoby, G., Bredthauer, R., Boggs, K., Lesser, M., Onaka, P., Young, D., Doherty, P., Craig, D.: The orthogonal-transfer array: A new CCD architecture for astronomy. Proceedings of the SPIE, Optical and Infrared Detectors for Astronomy 5499, 185–192 (2004)
5. Candes, E., Romberg, J.: ℓ^1-magic: Recovery of sparse signals via convex programming. http://users.ece.gatech.edu/∼justin/l1magic/ (2005)
6. Elad, M., Feuer, A.: Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images. IEEE Transactions on Image Processing 6(12), 1646–1658 (1997)
7. Evangelidis, G.D., Psarakis, E.Z.: Parametric image alignment using enhanced correlation coefficient maximization. IEEE Transactions on Pattern Analysis and Machine Intelligence 30(10), 1858–1865 (2008)
8. Farsiu, S., Elad, M., Milanfar, P.: Multiframe demosaicing and super-resolution of color images. IEEE Transactions on Image Processing 15(1), 141–159 (2006)
9. Freeman, W.T., Pasztor, E.C.: Learning low-level vision. In: Proc. of International Conference on Computer Vision (ICCV). pp. 1182–1189 (1999)
10. Fuji film Super CCD: http://www.dcviews.com/press/fuji_superccd_4.htm (2003)
11. Irani, M., Peleg, S.: Improving resolution by image restoration. Computer Vision, Graphics, and Image Processing 53, 231–239 (1991)
12. Lin, Z., Shum, H.Y.: Fundamental limits of reconstruction-based superresolution algorithms under local translation. IEEE Transactions on Pattern Analysis and Machine Intelligence 26(1), 83–97 (2004)
13. Moreno, I., Paez, G., Strojnik, M.: Dove prism with increased throughput for implementation in a rotational-shearing interferometer. Applied Optics 42(22), 4514–4521 (2003)
14. Murty, M.V.R.K., Hagerott, E.C.: Rotationalshearing interferometry. Applied Optics 5(4), 615–619 (1966)
15. Paige, C.C., Saunders, M.A.: LSQR: An algorithm for sparse linear equations and sparse least squares. ACM Transactions on Mathematical Software 8(1), 43–71 (1982)
16. Park, S.C., Park, M.K., Kang, M.G.: Super-resolution image reconstruction: A technical overview. IEEE Signal Processing Magazine 20(3), 21–36 (2003)
17. Sasao, T., Hiura, S., Sato, K.: Super-resolution with randomly shaped pixels and sparse regularization. In: Proc. of International Conference on Computational Photography (ICCP). pp. 1–11 (2013)
18. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. IEEE Transactions on Image Processing 19(11), 2861–2873 (2008)
19. Zhao, W.Y., Sawhney, H.S.: Is super-resolution with optical flow feasible. In: Proc. of European Conference on Computer Vision (ECCV). pp. 599–613 (2002)