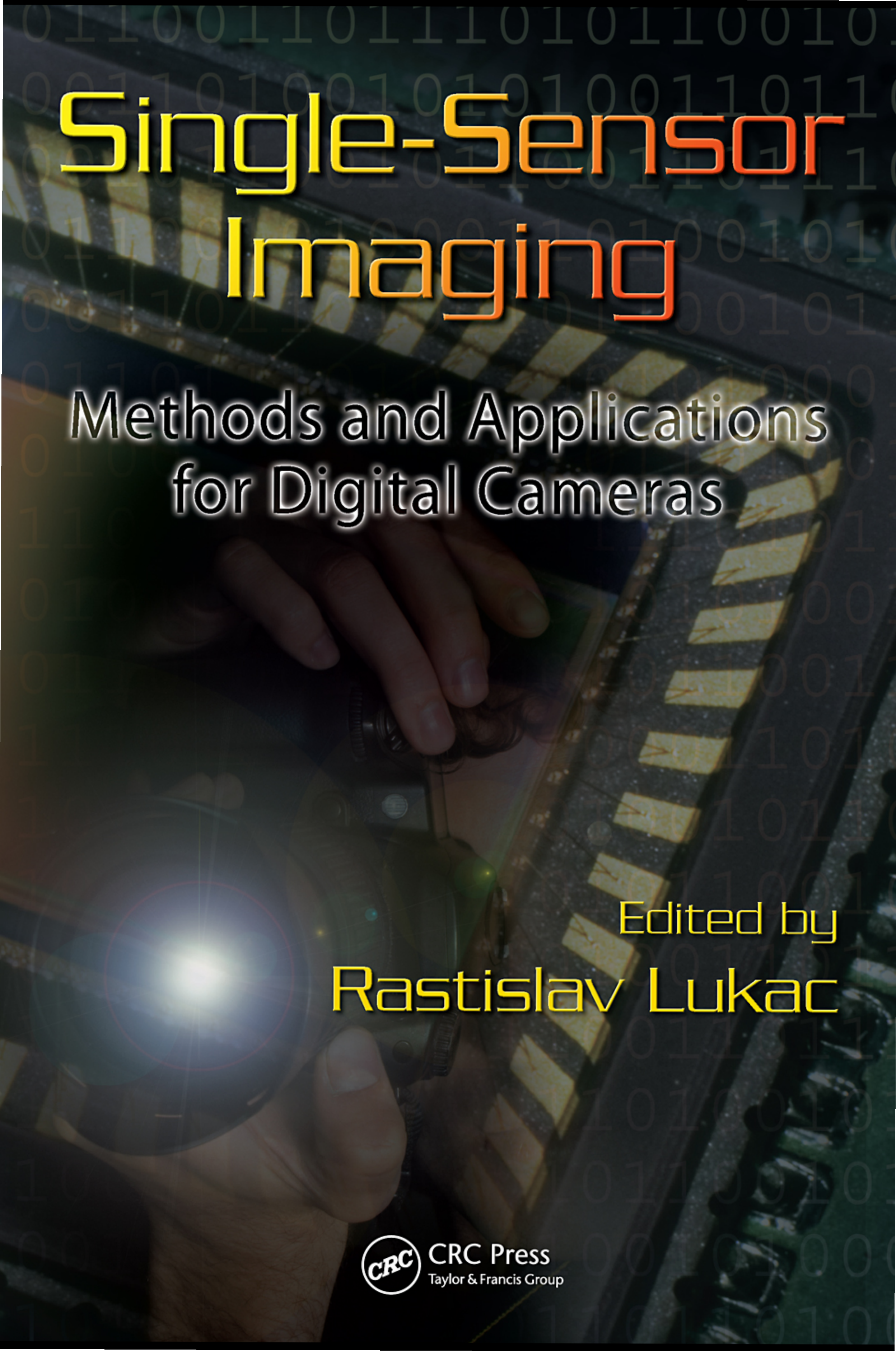


Single-Sensor Imaging

Methods and Applications for Digital Cameras

Edited by
Rastislav Lukac

CRC Press
Taylor & Francis Group

19

Simultaneous Demosaicking and Resolution Enhancement from Under-Sampled Image Sequences

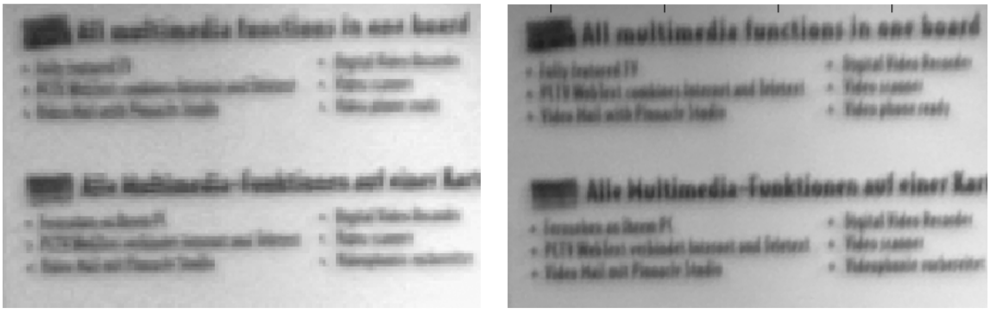
Sina Farsiu, Dirk Robinson, Michael Elad, and Peyman Milanfar

19.1 Introduction

Single-sensor imaging provides a convenient, yet low-cost solution for the problem of capturing color images at multiple wavelengths. To capture color, single-sensor cameras add a color filter array (CFA) [1], [2] on top of the photo-detector array. The CFA provides a simple mechanism for sampling different color channels in a multiplexed fashion.[1] Using a CFA trades-off the spatial sampling resolution of the three-sensor system, to achieve multispectral sampling. Refer to Chapters 1 and 5 for details.

Visual inspection of a real image shown in Figure 19.1a reveals various image quality issues encountered in single-sensor imaging. These shortcomings motivated the development of image processing algorithms to improve image quality for single-sensor imaging

[1]Design, properties, and performance characteristics of many CFA patterns are discussed in Reference [2].



(a) (b)

FIGURE 19.1 (See color insert.)

(a) Example showing the typical image quality problems associated with inexpensive single-sensor imaging devices. The image is noisy with evident color artifacts and has limited overall resolution as evidenced by the difficulty in reading the text numbers. (b) Image enhanced using a multiframe color supperresolution approach.

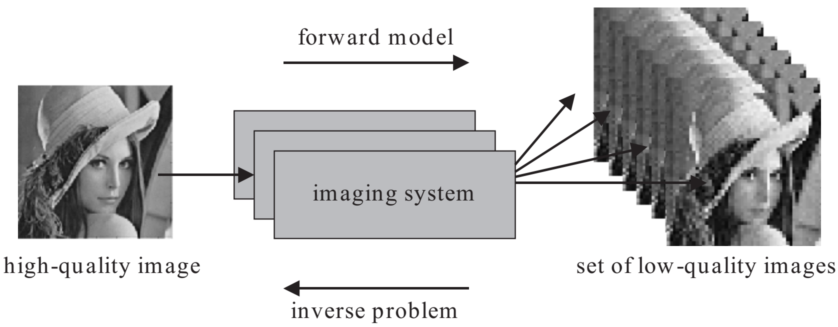


FIGURE 19.2

A block diagram representation of the multiframe image processing. The forward model is a mathematical description of the image degradation process. The multiframe inverse problem addresses the issue of retrieving (or estimating) the original scene from a *set* of low-quality images.

systems. Arguably, the most powerful image processing methods are the multiframe image processing approaches. The multiframe approaches combine the information from multiple images to produce higher quality images (Figure 19.1b). Figure 19.2 provides a general picture of the multiframe imaging approach to mitigate the shortcomings of imaging systems.

The need to interpolate the missing values in CFA images calls for a model that ties the full-color images in the sequence to the measurements obtained. The mathematical model of the forward imaging model for single-sensor imaging is

$$\mathbf{y}(k) = \mathbf{ADHF}(k)\mathbf{x} + \mathbf{v}(k) \tag{19.1}$$

where $\mathbf{y}$ is a vector containing all of the captured pixel values for a single image. The vector $\mathbf{x}$ contains the pixel values of the unknown high-resolution, noise-free image. Conceptually,

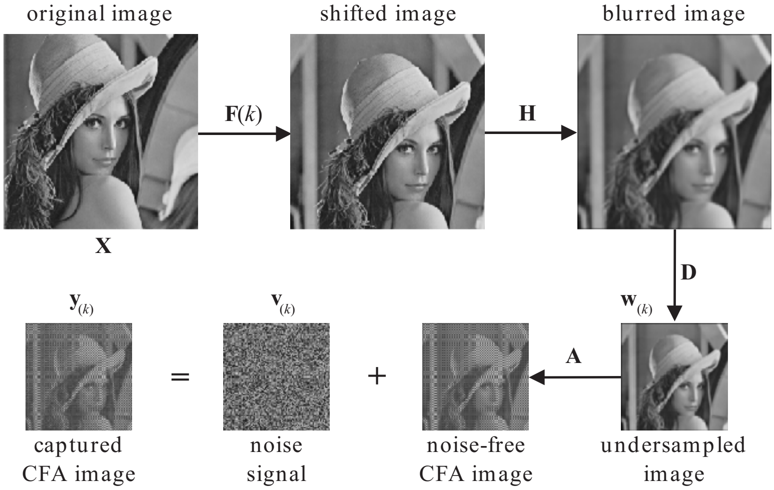


FIGURE 19.3

Block diagram representing the image formation model considered in this chapter, where $\mathbf{x}$ is the perfect intensity image of the scene, $\mathbf{v}$ is the additive noise, and $\mathbf{y}$ is the resulting color-filtered low-quality image. The operators $\mathbf{F}$, $\mathbf{H}$, $\mathbf{D}$, and $\mathbf{A}$ are representatives of the warping, blurring, downsampling, and color-filtering processes, respectively.

we can consider this vector to be comprised of the three unknown color channel images

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_R \\ \mathbf{x}_G \\ \mathbf{x}_B \end{pmatrix}. \tag{19.2}$$

The matrix $\mathbf{F}$ represents a motion operator capturing the image motion induced by the relative motion between the imaging system and the camera. The matrix $\mathbf{H}$ captures the blurring effects due to the camera optics, and $\mathbf{D}$ represents the downsampling effect due to limited resolution of the sensor array. These three operators ($\mathbf{F}$, $\mathbf{H}$, and $\mathbf{D}$) are applied to each of the color channels, either separately or jointly. In this work we assume that their effect is separated to each color channel.

The matrix $\mathbf{A}$ represents the sampling effect due to the CFA. This again stands for an operator that operates on the three color channels, sampling each differently, based on the Bayer pattern. The vectors $\mathbf{v}$ represent the additive random noise inherent in the imaging process. For the multiframe imaging case, the variable k indexes the N captured frames. Notice that in our proposed model, only the motion operators are assumed to vary as a function of k, implying that the blur, downsampling, and CFA sampling are all independent of time. An extension of this model could be proposed, where the various operators vary in time, and operate on the RGB in a mixed way.

Figure 19.3 depicts the image capture model for multiframe image processing. The goal of multiframe processing is to estimate the high resolution image $\mathbf{x}$ from a collection of low-resolution images $\{\mathbf{y}(k)\}$. In general, the problem of multiframe processing for single-sensor imaging systems is complicated. Earlier approaches to this problem sequentially solved single frame demosaicking, followed by multiframe resolution enhancement or su-

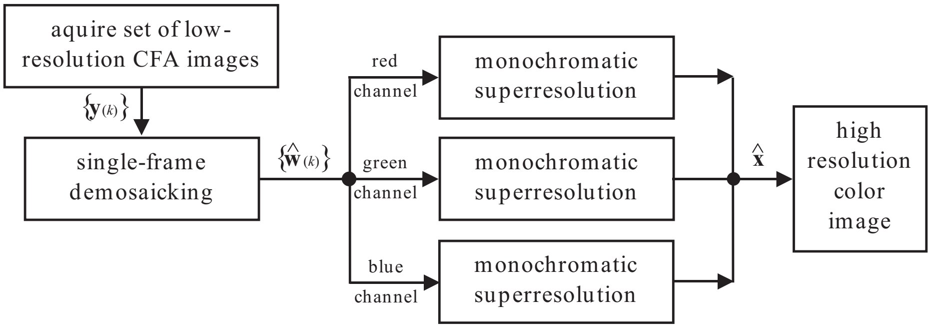


FIGURE 19.4

Block diagram representing the traditional two-step approach to the multiframe reconstruction of color images. First, a demosaicking algorithm is applied to each of the CFA images independently. Second, a superresolution algorithm is applied to each of the resulting color channel image sets independently.

perresolution. The proposed approach, a multiframe demosaicking scheme, solves both problems in a joint fashion offering improved performance. In Sections 19.2 and 19.3 we describe both of these approaches in more detail. In Section 19.4, we describe a fast implementation of the joint approach enabling multiframe enhancement of single-sensor video sequences. Concluding remarks of this chapter are given in Section 19.5.

19.2 Sequential Demosaicking and Superresolution

The classic approach for multiframe processing of single-sensor color images involves two stages of processing. First, single frame demosaicking interpolates the missing color pixel values in the mosaic data. This single frame demosaicking is applied to a collection of images independently, producing a set of full-color low-resolution images. Second, these restored low-resolution color images are processed using multiframe resolution enhancement or *superresolution* algorithms. The superresolution algorithms are either applied to each color channel independently [3], [4], or to the luminance (grayscale) components of the color images [5]. This traditional two-step approach is illustrated in Figure 19.4.

Producing a high resolution image requires a smart combination of data processing and prior knowledge on the spatial and spectral (color) properties of images. The prior information about the images characterizes both the color and the spatial properties of images. In the following sections, we review the various approaches to single-frame demosaicking and superresolution, including analysis of the different forms of prior information needed to solve these two problems.

19.2.1 Single-Frame Demosaicking

Single-frame demosaicking algorithms, or just demosaicking algorithms, attempt to restore the missing color components eliminated by using a color filter array. Essentially,

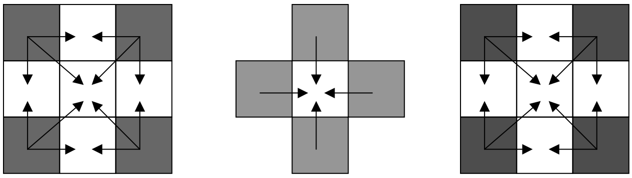


FIGURE 19.5

Illustration of the basic interpolation problem in single-frame demosaicking. The demosaicking algorithm must estimate the missing color pixel value using the observed neighboring values. Individual pixel arrangements correspond to: (left) blue, (middle) green, and (right) red channel of Bayer CFA captured images.

demosaicking algorithms try to estimate a low-resolution full-color image $\mathbf{w}(k)$ defined as

$$\mathbf{w}(k) = \mathbf{DHF}(k)\mathbf{x}. \tag{19.3}$$

The full-color image $\mathbf{w}(k)$ has three color values for every pixel location, whereas the captured image $\mathbf{y}(k) = \mathbf{Aw}(k) + \mathbf{v}(k)$ contains only one color value per pixel location. The single-frame demosaicking algorithm is applied to each captured image $\mathbf{y}(k)$ independently.

Numerous demosaicking methods have been proposed over the years to solve this underdetermined problem and in this section we review some of the more popular approaches. The simplest approach estimates the unknown pixel values using linear interpolation of the known color values in the neighboring pixel locations. Figure 19.5 illustrates the standard neighborhood interpolation problem of single-frame demosaicking.

The straightforward interpolation approach ignores some important information about the correlation between the color channel images and often produces serious color artifacts. It can be shown that the red and blue channels are related to the green channel via the approximate equality of the ratios $\frac{red}{green}$ and $\frac{blue}{green}$. As observed in Figure 19.5, the green channel has twice the number of pixels, compared to the red and blue channels. Thus, using the correlations between the color channels just described allows us to provide better interpolation of the red and blue channels using the more reliably interpolated green channel. This approach forms the basis of the smooth-hue transition method first discussed in Reference [6].

Note that this correlation between color channels does not hold across edges in an image. Consequently, although the smooth-hue transition algorithm works for smooth regions of the image, it does not work in the high-frequency (edge) areas. Considering this fact, gradient-based methods, first addressed in Reference [7], do not perform interpolation across the edges of an image. This noniterative method [7] uses the second derivative of the red and blue channels to estimate the edge direction in the green channel. Later, the green channel is used to compute the missing values in the red and blue channels.

A variation of this method was later proposed in Reference [8], where the second derivative of the green channel and the first derivative of the red (or blue) channels are used to estimate the edge direction in the green channel. Both the smooth hue and edge-based methods were later combined in Reference [9]. In this iterative method, the smooth hue interpolation is performed along the local image gradients computed in eight directions about a pixel of interest. A second stage using anisotropic inverse diffusion further enhances the quality of the reconstructed image. This two step approach of interpolation followed by an enhancement step has been used in many other settings. In Reference [10], spatial and spectral correlations among neighboring pixels are exploited to define the interpolation step, while adaptive median filtering is used as the enhancement step. A different iterative implementation of the median filter is used as the enhancement step of the method described in Reference [11], that takes advantage of a homogeneity assumption in the neighboring pixels.

Maximum a posteriori (MAP) methods form another important category of demosaicking methods. These MAP approaches apply global assumptions about the correlations between the color channels and the spatial correlation using a penalty function of the form

$$\Omega(\mathbf{w}) = J_0(\mathbf{w}, \mathbf{y}) + P(\mathbf{w}) \tag{19.4}$$

where $J_0(\cdot,\cdot)$ captures the correspondence between the estimated image $\mathbf{w}$ and the observed data $\mathbf{y}$ and $P(\cdot)$ captures the quality of the proposed solution $\mathbf{w}$ based on the prior knowledge about the spatial and color correlations in the destination image. For example, a MAP algorithm with a smooth chrominance prior is discussed in Reference [12]. The smooth chrominance prior is also used in Reference [13], where the original image is first transformed to YIQ representation.[2] The chrominance interpolation is performed using isotropic smoothing. The luminance interpolation is done using edge directions computed in a steerable wavelet pyramidal structure.

Other examples of popular demosaicking methods available in published literature are [15], [16], [17], [18], [19], [20], [21], and [22]. For additional information on demosaicking refer to Chapters 1, 3, and 6 to 9. We also note that a few postdemosaicking image enhancement techniques have been proposed (e.g., References [23] and [24]) to reduce artifacts introduced through the color-interpolation process. Almost all of the above demosaicking methods are based on one or more of these following assumptions:

1. In the measured image with the mosaic pattern, there are more green sensors with regular pattern of distribution than blue or red ones (in the case of Bayer CFA there are twice as many greens than Red or Blue pixels and each is surrounded by four green pixels).
2. The CFA pattern is most likely the Bayer pattern.
3. For each pixel, one and only one color band value is available.
4. The color pattern of available pixels does not change through the measured image.
5. The human eye is more sensitive to the details in the luminance component of the image than the details in chrominance component [13].
6. The human eye is more sensitive to chromatic changes in the low-spatial frequency regions than the luminance changes [18].
7. Interpolation should be performed along and not across the edges.
8. Different color bands are correlated with each other.
9. Edges should align between color channels.

[2] YIQ is the standard color representation used in broadcast television (NTSC systems) [14].

FIGURE 19.6

A high-resolution image (a) captured by a three-CCD sensor camera is (b) downsampled by a factor of four. In (c) the original three-sensor image is blurred by a Gaussian kernel before downsampling by a factor of four. The images in (a-c) are color-filtered and then demosaicked by the method of Reference [9] to produce the results shown in (d-f), respectively. © 2007 IEEE

The accuracy of these single-frame demosaicking algorithms depends to a large degree on the resolution of the single-sensor imaging system. For example, Figure 19.6a shows a high-resolution image captured by a three-sensor camera. If we had captured this image instead using a single-sensor camera and then applied the single frame demosaicking algorithm of Reference [9], we would obtain the image depicted in Figure 19.6d. This demosaicked image shows negligible color artifacts. Figure 19.6b shows the same scene from a simulated three-sensor camera which is undersampled by a factor of four. Figure 19.6e shows the single-frame demosaicking result, where the color artifacts in this image are much more evident than Figure 19.6d. In this case, inadequate sampling resolution and the subsequent aliasing produces severe color artifacts. Such severe aliasing happens in cheap commercial still or video digital cameras, with small number of sensor pixels [25].

Some imaging systems attempt to avoid these aliasing artifacts by adding optical low-pass filters to eliminate high spatial frequency image content prior to sampling. Refer to Chapter 4 for details. Such optical low-pass filters not only increase system cost but lower image contrast and sharpness. These optical low-pass filters can remove some, but not all color artifacts from the demosaicking process. Figure 19.6c shows a simulated low-resolution image which was presmoothed by a Gaussian low-pass digital filter simulating the effects of an optical low-pass filter prior to downsampling. The demosaicked version of this captured image is shown in Figure 19.6f. The demosaicked image contains fewer color artifacts than Figure 19.6e; however, it has lost some high-frequency details.

The residual aliasing artifacts amplified by the demosaicking algorithm motivated the application of multiframe resolution enhancement algorithms. We describe these algorithms in the next section.

19.2.2 Multiframe Superresolution

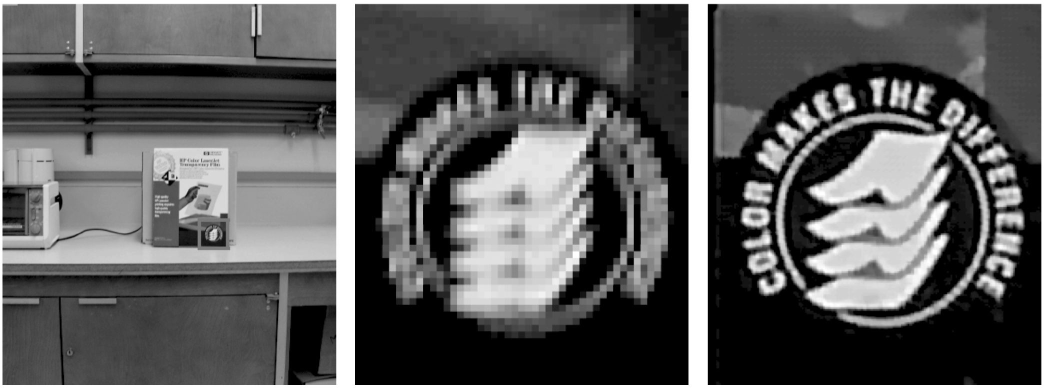
A relatively recent approach to dealing with aliasing in imaging systems is to combine multiple aliased images with some phase modulation between them to reconstruct a single high resolution image free of (or with significantly less) aliasing artifacts with improved SNR.[3] Such methods have been studied extensively for monochromatic imaging [26], [27], [28] and are commonly known as multiframe resolution enhancement (superresolution) algorithms, which produce sharp images with less noise and higher *spatial resolution* than the capture images.

The standard model for such algorithms assumes that a single monochromatic low-resolution aliased image in a set of N images is defined by

$$\mathbf{m}(k) = \mathbf{DHF}(k)\mathbf{g} + \mathbf{v}(k), k = 0 \ldots N-1 \tag{19.5}$$

where $\mathbf{m}(k)$ is the kth captured monochrome image. In this case, $\mathbf{g}$ represents the unknown high resolution monochromatic image. In the context of the sequential multiframe demosaicking approach, the captured images $\mathbf{m}(k)$ are obtained by extracting one of the color channels from the demosaicked images. In other words, $\mathbf{m}(k)$ is either $\hat{\mathbf{w}}_R(k)$ or $\hat{\mathbf{w}}_G(k)$ or $\hat{\mathbf{w}}_B(k)$ and $\mathbf{g}$ is either $\mathbf{x}_R$, or $\mathbf{x}_G$, or $\mathbf{x}_B$. As before, $\mathbf{F}(k)$ is a warping operator capturing the relative motion between the kth frame and the anchor frame $k = 0$ and $\mathbf{v}(k)$ represents the additive noise in the image system. The matrix $\mathbf{H}$ represents the blurring associated with the optical point spread function and $\mathbf{D}$ represents the downsampling operator reflecting the limited spatial resolution associated with the image sensor. In effect, the downsampling operator is the cause of the aliasing artifacts observed in the final image. While we do not address the motion estimation in full detail in this chapter, an appendix provides a brief discussion about this topic, with an emphasis on the need to estimate the cross-motion between the N frames jointly to avoid accumulated error.

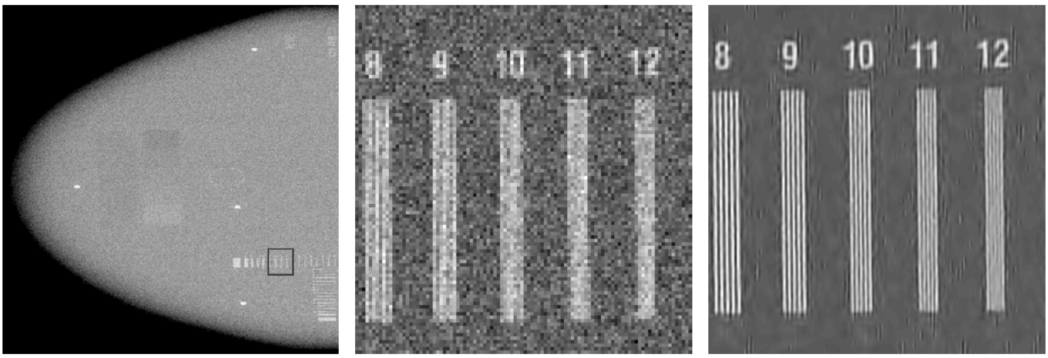
[3]Signal to noise ratio (SNR) is defined as $10\log_{10}\frac{\sigma^2}{\sigma_n^2}$, where σ^2, σ_n^2 are variance of a clean frame and noise, respectively.



(a) (b) (c)

FIGURE 19.7

Superresolution experiment on real image data. Twenty-six low quality images were combined to produce a higher quality image. One captured image is shown in (a). The red square section of (a) is zoomed in (b). Superresolved image in (c) is the high quality output image.



(a) (b) (c)

FIGURE 19.8

Superresolution experiment on real image data. Seventeen low quality images from a digital X-ray machine were combined to produce a higher quality image. One captured image is shown in (a). The red square section of (a) is zoomed in (b). Superresolved image in (c) is the high quality output image.

Early studies of superresolution showed that the aliasing artifacts in the low-resolution images enable the recovery of the high-resolution fused image, provided that a relative sub-pixel motion exists between the under-sampled input images [29].

Figure 19.7 shows an example of the resulting image after applying superresolution processing to a collection of images taken by a commercial digital still camera. The application of superresolution is not restricted to optical imaging. An example is illustrated in Figure 19.8, where seventeen images captured by a Siemens digital X-ray imaging system were fused to create the shown high resolution X-ray image [30]. Figure 19.8a shows one of the input images, a selected region of which is zoomed in Figure 19.8b for a closer examination. A factor of three superresolved image of Figure 19.8a is zoomed in Figure 19.8c, showing a significant reduction in aliasing artifacts.

The multiframe superresolution problem was first addressed in Reference [29], where the authors proposed a frequency domain approach, later extended by others [31]. Although the frequency domain methods are computationally cheap, they are very sensitive to noise and modelling errors [26]. Also, by design, these approaches handle only pure translational motion.

Another popular class of methods solves the problem of resolution enhancement in the spatial domain. The noniterative spatial domain data fusion approaches are studied in References [32], [33], and [34]. The iterative methods are based on construction of a cost function, which is minimized in an iterative function minimization approach. The cost function is of the form

$$\Omega(\mathbf{g}) = J_0(\mathbf{g}, \{\mathbf{m}(k)\}) + P(\mathbf{g}) \tag{19.6}$$

where $J_0(\mathbf{g}, \{\mathbf{m}(k)\})$ is the data penalty term that measures the closeness of the observed data $\{\mathbf{m}(k)\}$ to the predicted measurements using an estimate of the image $\mathbf{g}$. The $P(\mathbf{g})$ term represents the prior information about the unknown signal $\mathbf{g}$, as described before.

The most common data fidelity term is based on minimizing the L_2 difference between the observed data and the predicted data for an estimated image $\mathbf{g}$. The cost function looks like

$$J_0(\mathbf{g}, \{\mathbf{m}(k)\}) = \sum_{k=0}^{N-1} \|\mathbf{DHF}(k)\mathbf{g} - \mathbf{m}(k)\|_2^2. \tag{19.7}$$

The solution to this penalty function is equivalent to the least squares estimate of the image $\mathbf{g}$, and is the optimal maximum likelihood (ML) solution [35] when noise follows the white Gaussian model, and when the motion vectors implicit in $\mathbf{F}(k)$ are assumed to be known exactly, and a priori.

One distinct advantage of the spatial-domain approaches is the ability to incorporate complicated prior information into the estimation process. The most common form of prior information arises as a type of *regularization* to the commonly ill-conditioned or ill-posed estimation problem. This regularization or prior information takes the form of some assumptions about the spatial variability of the unknown high resolution image $\mathbf{g}$. Many of the early superresolution algorithms, however, applied rudimentary regularization such as Tikhonov regularization [36], which describes the smoothness of the image $\mathbf{g}$. A Tikhonov regularization has the form

$$P(\mathbf{g}) = \lambda \|\Lambda \mathbf{g}\|_2^2 \tag{19.8}$$

where Λ represents a spatial high-pass operator applied to the image $\mathbf{g}$. In other words, the estimated image will have a reduced high spatial frequency content, depending on the magnitude of the weighting parameter λ.

The early iterative algorithms used search techniques to minimize cost functions. The simplest approach to minimize a cost function of the form of Equation 19.6 is to use the steepest descent (SD) algorithm. In SD, the derivative of Equation 19.6 is computed with respect to the unknown parameter vector $\mathbf{g}$ to update the estimate. The estimate for the $n+1$th iteration is updated according to

$$\widehat{\mathbf{g}}^{n+1} = \widehat{\mathbf{g}}^{n} - \beta \nabla_x \Omega(\widehat{\mathbf{g}}^{n}) \tag{19.9}$$

where β is the SD step size. This approach is similar to the iterative back-projection method proposed in References [5] and [37].

These spatial domain methods discussed so far were often computationally expensive, especially when solved via explicit construction of the modelling matrices of Equation 19.5. The authors of Reference [38] introduced a block circulant preconditioner for solving the Tikhonov regularized superresolution problem formulated in Reference [36]. In addition, they addressed the calculation of regularization factor for the under-determined case[4] by generalized cross-validation presented in Reference [39].

Later, a very fast superresolution algorithm for pure translational motion and common space invariant blur was developed in Reference [33]. This fast approach was based on the observation that the deblurring can be separated from the merging process of the low-resolution images, and that merging of the images is obtained simply by shift-and-add operation. Using these, it was shown that the result is nevertheless ML optimal. Another interesting observation in this work is the fact that the matrices $\mathbf{H}$, Λ, $\mathbf{D}$, etc., can be applied directly in the image domain using standard operations such as convolution, masking, downsampling, and shifting. Implementing the effects of these matrices as a sequence of operators on the images eliminates the need to explicitly construct the matrices. As we shall see momentarily, this property helps us develop fast and memory efficient algorithms.

A different fast spatial domain method was recently suggested in Reference [40], where low-resolution images are registered with respect to a reference frame defining a nonuniformly spaced high-resolution grid. Then, an interpolation method called Delaunay triangulation is used for creating a noisy and blurry high-resolution image, which is subsequently deblurred. All of the above methods assumed the additive Gaussian noise model. Furthermore, regularization was either not implemented or it was limited to Tikhonov regularization. Another class of superresolution techniques uses the projection onto convex sets (POCS) formulation [41], [42], [43]. The POCS approach allows complicated forms of prior information such as set inclusions to be invoked.

Recently, a powerful class of superresolution algorithms leverage sophisticated *learned*-based prior information. In these algorithms, the prior penalty functions $P(\mathbf{g})$ are derived from collections of training image samples [44], [45]. For example, in Reference [45] an explicit relationship between low-resolution images of faces and their known high-resolution image is learned from a face database. This learned information is later used in reconstructing face images from low-resolution images. Due to the need for gathering a vast number of examples, often these methods are only effective when applied to very specific scenarios, such as faces or text.

Another class of superresolution algorithms focuses on being robust to modelling errors and different noise. These methods implicitly [46] or explicitly [47], [48] take advantage of L_1 distance metrics when constructing the cost functions. Compared to Equation 19.7, the methods in References [47] and [48] apply the L_1 metric as the data fidelity term

$$J_0(\mathbf{g}, \{\mathbf{m}(k)\}) = \sum_{k=0}^{N-1} \|\mathbf{DHF}(k)\mathbf{g} - \mathbf{m}(k)\|_1, \tag{19.10}$$

[4]In under-determined case, the number of nonredundant low-resolution frames is smaller than the square of resolution enhancement factor. A resolution enhancement factor of r means that low-resolution images of dimension $Q_1 \times Q_2$ produce a high-resolution output of dimension $rQ_1 \times rQ_2$. Scalars Q_1 and Q_2 are the number of pixels in the vertical and horizontal axes of the low-resolution images, respectively.

(a) (b) (c)

FIGURE 19.9

The image (a) shows an example low-resolution CFA image processed by the single frame demosaicking algorithm of Reference [9]. The image (b) shows the resulting color image after applying the robust superresolution algorithm [47] to each of the color channels independently completing the sequential multiframe processing algorithm, which shows improved resolution with significantly reduced artifacts. Some residual color artifacts, however, are apparent along the fence. Better reconstruction results are shown in (c), where the multiframe demosaicking method presented in Section 19.3 has fused the information of ten CFA frames. © 2007 IEEE

which is the optimal ML solution in the presence of additive white Laplacian noise.

In Reference [47], in the spirit of the total variation criterion [49], [50] and a related method called the bilateral filter [51], [52], the L_1 distance measure was also incorporated into a prior term called the bilateral total variation (BTV) regularization penalty function, defined as

$$P(\mathbf{g}) = \lambda \sum_{l=-L_{max}}^{L_{max}} \sum_{m=-M_{max}}^{M_{max}} \alpha^{|m|+|l|} \|\mathbf{g} - \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{g}\|_1. \tag{19.11}$$

where S_x^l and S_y^m are the operators corresponding to shifting the image represented by $\mathbf{g}$ by l pixels in the horizontal and m pixels in the vertical directions, respectively. The parameters M_{max} and L_{max} define the size of the corresponding bilateral filter kernel. The scalar weight α, $0 < \alpha \leq 1$, is applied to give a spatially decaying effect to the summation of the regularization terms.

Alternative robust superresolution approaches were presented in References [53] and [54] based on normalized convolution and nonparametric kernel regression techniques. To achieve robustness with respect to errors in motion estimation, Reference [55] proposed a novel solution based on modifying camera hardware. Finally, in References [42], [56], and [57] the superresolution framework has been applied to reduce the quantization noise resulting from video compression. More comprehensive surveys of monochromatic multiframe superresolution methods can be found in References [26], [58], [59], and [60].

19.2.3 Sequential Processing Example

Figure 19.9 shows an example of the sequential multiframe processing. In this example, ten CFA images were synthetically generated using the high resolution image shown in Figure 19.6a. To produce the simulated CFA images, each color channel of the high resolution image was blurred by a 5×5 pixel Gaussian blurring kernel with unit standard deviation. Then, the ten images were taken from this blurry image by subsampling by a factor of four starting at randomly offset pixels. White Gaussian noise was added to each of the simulated low-resolution images such that the image had an SNR of 30dB. Each of the simulated CFA images was processed by the single frame demosaicking algorithm described in Reference [9].

Figure 19.9a shows an example low-resolution image after single frame demosaicking. Then, the robust monochromatic superresolution algorithm of Reference [47] was applied to each of the three color channels independently. The resulting high resolution image is shown in Figure 19.9b, which shows improvement in resolution and noise reduction. The high resolution image still, however, contains some color artifacts and has over-smoothing in some regions. The weakness of these results motivated the use of multiframe demosaicking algorithms, an example of which is shown in Figure 19.9c. These techniques are described in detail in the next section.

19.3 Multiframe Demosaicking

The sequential multiframe processing approach works reasonably well when the amount of aliasing artifacts is small. With significant color aliasing artifacts, however, the sequential approach fails to adequately correct all of the color artifacts. Basically, the failure is traced to the fact that all of the color prior information is applied at the individual (single) frame processing level during demosaicking. A more complicated, yet robust method performs the demosaicking and superresolution *simultaneously*. This approach is known as *multiframe demosaicking* and has been developed relatively recently [61], [62], [63], [64], [65], [66], [67]. Note that some of these methods (e.g., References [63] and [66]), not motivated by superresolution enhancement, were designed primarily for video-demosaicking, which is described in detail in Chapter 18.

19.3.1 Modelling the Problem

The multiframe demosaicking approach tries to estimate the full-color high resolution image $\mathbf{x}$ using the the *set* of unprocessed CFA images $\mathbf{y}(k)$. In this case, the complete forward data model given by Equation 19.1 is used to reconstruct the full-color high resolution image $\mathbf{x}$.

The generic MAP multiframe demosaicking approach estimates $\mathbf{x}$ by minimizing a cost function of the form

$$\Omega(\mathbf{x}) = J_1(\mathbf{x}, \{\mathbf{y}(k)\}) + P_1(\mathbf{x}) + P_2(\mathbf{x}) + P_3(\mathbf{x}). \tag{19.12}$$

This cost function is similar to the superresolution cost function of Equation 19.6 with the exception that now we have three prior information terms. These multiple terms include both the spatial and color prior information constraints. In Reference [65], the MAP-based penalty terms are described as:

1. $J_1(\cdot,\cdot)$: A penalty term to enforce similarities between the raw data and the high-resolution estimate (data fidelity penalty term).
2. $P_1(\cdot)$: A penalty term to encourage sharp edges in the luminance component of the high-resolution image (spatial luminance penalty term).
3. $P_2(\cdot)$: A penalty term to encourage smoothness in the chrominance component of the high-resolution image (spatial chrominance penalty term).
4. $P_3(\cdot)$: A penalty term to encourage homogeneity of the edge location and orientation in different color bands (spectral dependencies penalty term).

These terms combine several forms of prior information, forming a powerful constraint when performing multiframe demosaicking. Details about these terms are provided below:

• *Data Fidelity Penalty Term*
In the multiframe demosaicking case, the data fidelity term must include all three color channels. Considering the general motion and blur model of Equation 19.1, a reasonable L_2-based multiframe data fidelity penalty term is given by:

$$J_1(\mathbf{x},\{\mathbf{y}(k)\}) = \sum_{i=R,G,B} \sum_{k=0}^{N-1} \|\mathbf{ADHF}(k)\mathbf{x}_i - \mathbf{y}_i(k)\|_2^2. \tag{19.13}$$

Essentially, this functional penalizes estimates of the high resolution color image that, when passed through the forward model of Equation 19.1, differs from the observed image. The data penalty function, Equation 19.13, works well when the model is accurate (e.g., the motion estimation produces accurate estimates of the forward imaging model, and the blur is the correct one). When these are not true, the forward model contains errors, and a more robust form of Equation 19.13 could be proposed, based on the L_1 norm:

$$J_1(\mathbf{x},\{\mathbf{y}(k)\}) = \sum_{i=R,G,B} \sum_{k=0}^{N-1} \|\mathbf{ADHF}(k)\mathbf{x}_i - \mathbf{y}_i(k)\|_1. \tag{19.14}$$

This is the full-color generalization of the robust monochromatic superresolution penalty term of Equation 19.10. As before, this penalty term provides robustness at the expense of slight increase in computational complexity, and reduced denoising efficiency when applied to Gaussian noise.

• *Spatial Luminance Penalty Term*
The human eye is more sensitive to the details in the luminance component of an image than the details in the chrominance components [13]. It is important, therefore, that the edges in the luminance component of the reconstructed high-resolution image look sharp. As explained in Reference [47], the BTV functional of Equation 19.11 produces images with sharp edges. For the multiframe demosaicking application, we apply the BTV cost

function to the reconstruction of the luminance component. The luminance image can be calculated as the weighted sum $\mathbf{x}_L = 0.299\mathbf{x}_R + 0.597\mathbf{x}_G + 0.114\mathbf{x}_B$, [14]. This luminance image is the Y component of the YIQ image format commonly found in video coding. Using λ_1 to denote the strength of the image prior, the BTV luminance regularization term is then defined as

$$P_1(\mathbf{x}) = \lambda_1 \sum_{l=-L_{max}}^{L_{max}} \sum_{m=-M_{max}}^{M_{max}} \alpha^{|m|+|l|} \|\mathbf{x}_L - \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{x}_L\|_1. \tag{19.15}$$

• *Spatial Chrominance Penalty Term*

Spatial regularization is required also for the chrominance channels. However, since the human visual system is less sensitive to the resolution of these bands, we can apply a simpler regularization functional, based on the L_2 norm, in the form of

$$P_2(\mathbf{x}) = \lambda_2 \left(\|\Lambda \mathbf{x}_I\|_2^2 + \|\Lambda \mathbf{x}_Q\|_2^2 \right), \tag{19.16}$$

where the images $\mathbf{x}_I$ and $\mathbf{x}_Q$ are the I and Q portions of the YIQ color representation and λ_2 defines the strength of the image prior. As before, Λ is a spatial high-pass operator such as derivative, Laplacian, or even identity matrix. Such a penalty term encourages the two chrominance components to vary smoothly over the high resolution image.

• *Spectral Dependency Penalty Term*

This penalty term characterizes the spectral or interchannel correlation between the three channels of the original color images. This penalty function minimizes the mismatch between locations or orientations of edges across the color bands. Following Reference [12], minimizing the vector product norm of any two adjacent color pixels forces different bands to have similar edge location and orientation. Based on the theoretical justifications in Reference [68], the authors of Reference [12] suggest a pixelwise spectral dependencies cost function to be minimized. This term has the vector outer product norm of all pairs of neighboring color pixels. With some modifications to what was proposed in Reference [12], our spectral dependencies penalty term is a differentiable cost function

$$\begin{aligned} P_3(\mathbf{x}) = \lambda_3 \sum_{l,m=-1}^{1} \Big[&\|\mathbf{x}_G \odot \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{x}_B - \mathbf{x}_B \odot \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{x}_G\|_2^2 + \\ &\|\mathbf{x}_B \odot \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{x}_R - \mathbf{x}_R \odot \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{x}_B\|_2^2 + \|\mathbf{x}_R \odot \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{x}_G - \mathbf{x}_G \odot \mathbf{S}_x^l \mathbf{S}_y^m \mathbf{x}_R\|_2^2 \Big], \end{aligned} \tag{19.17}$$

where $\odot$ is the element by element multiplication operator. The λ_3 defines the strength of the image prior.

As with the monochromatic superresolution case, we minimize the generic full-color cost function from Equation 19.12 using a type of SD. In each step, the derivative will be computed with respect to one color channel assuming the other color channels are fixed. Thus, for the $n+1$ iteration, the algorithm updates the ith color channel according to

$$\widehat{\mathbf{x}}_i^{n+1} = \widehat{\mathbf{x}}_i^n - \beta \nabla_{x_i} \Omega(\mathbf{x}^n) \qquad i = R, G, B \tag{19.18}$$

where the scalar β is the steepest descent step size. This minimization approach converges relatively quickly for most image sequences, providing satisfactory results after ten to fifteen such iterations.

(a) (b)

FIGURE 19.10 (See color insert.)

Multiframe color superresolution applied to a real data sequence. (a) One of thirty-one low-resolution input images of size $141 \times 147 \times 3$ demosaicked by the single-frame method of Reference [7]. (b) Resulting image after multiframe demosaicking using the robust data penalty term of Equation 19.14. © 2007 IEEE

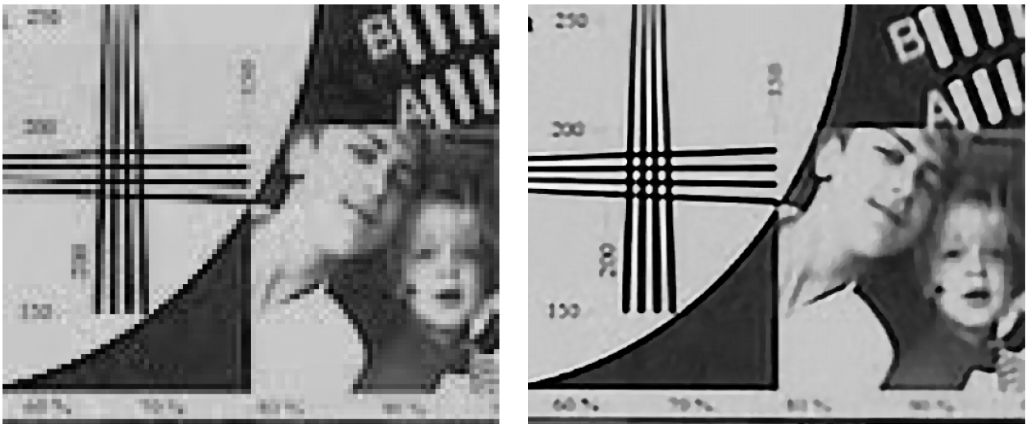
19.3.2 Multiframe Demosaicking Examples

In this section, we present some visual results of the multiframe demosaicking algorithm described in the previous section. The first example demonstrates the advantages of the multiframe demosaicking algorithm when dealing with raw CFA images taken directly from an image sensor. In this example, we acquired thirty-one uncompressed raw CFA images from a two megapixel CMOS sensor. The (unknown) camera point spread function (PSF) was modelled as a tapered 5×5 disk PSF. In all examples, to compute the unknown motion matrices, the raw CFA images were first demosaicked by the single-frame method from Reference [7]. Then, the luminance component of these images was registered in a pairwise fashion using the motion estimation algorithm[5] described in Reference [70].

Figure 19.10a shows a single low-resolution image after applying the single-frame demosaicking method of Reference [7]. The single-frame demosaicking method exhibits the typical problems of low-resolution, color aliasing, and sensor noise. Figure 19.10b shows the resulting image after applying the multiframe demosaicking algorithm directly to the set of low-resolution CFA images to increase the spatial resolution by a factor of three.

In the second example, we repeated the above experiment using a different camera, capturing twenty-six color filtered images. Figure 19.1a shows a single low-resolution image after applying more complex single-frame demosaicking from Reference [9] and Figure 19.1b shows the resulting image after applying the multiframe demosaicking algorithm with the data penalty term of Equation 19.14 directly to the set of low-resolution CFA images to increase the spatial resolution by a factor of three. Both examples depicted in Figure 19.1b and Figure 19.10b show improved resolution and almost no color artifacts.

[5]Application of an accurate multiframe registration technique (see Appendix) for directly registering the CFA data is described in Reference [69].



(a) (b)

FIGURE 19.11 (See color insert.)

(a) One of forty captured images which are demosaicked and compressed in an unknown fashion by the imaging system. These images were used to reconstruct the higher resolution image in (b) by exploiting the multiframe demosaicking algorithm. © 2007 IEEE

In the third example, we apply the multiframe demosaicking algorithm to a set of color images which have already undergone single-frame demosaicking and compression. Figure 19.11a shows an example low-resolution image which has undergone an unknown demosaicking algorithm followed by compression. Forty such images were combined using the multiframe demosaicking algorithm to produce the image on the right. This resulting image demonstrates the capability of the multiframe demosaicking approach to handle images later in the imaging pipeline after standard compression. While the results would presumably be better if the multiframe algorithm had direct access to the raw CFA images, the final image shows significant improvement over the input image.

19.4 Fast and Dynamic Multiframe Demosaicking

In this section, first we describe the conditions under which the multiframe demosaicking algorithm described in the previous section may be implemented very efficiently (Section 19.4.1). Using this fast formulation, in Section 19.4.2 we extend the multiframe demosaicking algorithm to *dynamic* multiframe processing which differs from basic multiframe demosaicking in that the final result is a high resolution image sequence or video. This is also known as video-to-video superresolution, which has been addressed in the literature for the grayscale [71], [72], [73] and color (demosaicked) [74] sequences.

19.4.1 Fast Multiframe Demosaicking

While the iterative method described in the previous section is fast enough for many practical applications, speeding up the implementation is desirable when handling large amounts of data. In this section, we describe a fast, two-stage approach to multiframe demosaicking, in the spirit of the method developed in Reference [75]. The two stage approach works only when two conditions are satisfied. First, the common system PSF is spatially invariant. Second, the motion between the low-resolution frames (at least locally) is translational in nature. When these two conditions are satisfied, the shifting operator $\mathbf{F}(k)$ and the blurring operator $\mathbf{H}$ commute. We then define $\mathbf{z} = \mathbf{Hx}$ as the unknown high-resolution blurry image. In the fast approach, multiframe demosaicking is divided into the following two steps:

1. Noniterative data fusion (shift-and-add – see below) provides an estimate of $\mathbf{z}$ from the captured data. The model for this step is

$$\mathbf{y}(k) = \mathbf{ADF}(k)\mathbf{z} + \mathbf{v}(k) \tag{19.19}$$

2. Iterative deblurring and interpolation provides an estimate of $\mathbf{x}$ from the estimate $\hat{\mathbf{z}}$.

The two-step approach is fast since the first stage is noniterative. After this stage, we no longer need to store the set of capture images $\{\mathbf{y}(k)\}$. In essence, the estimate $\hat{\mathbf{z}}$ is a sufficient statistic with which we can estimate the high resolution image $\mathbf{x}$.

19.4.1.1 Data Fusion Step

The data fusion step solves the estimation problem for the model in Equation 19.19 in a noniterative fashion for each color channel independently by using an analytic minima of a simple data penalty cost function [47]. If the penalty function is based on the L_2 norm penalty function

$$\sum_k \|\mathbf{y}_i(k) - \mathbf{DF}(k)\mathbf{z}_i\|_2^2, \;\; i = R, G, B \tag{19.20}$$

the fusion step is called the shift-and-add algorithm. In this algorithm, the input images are upsampled by zero-padding, shifted by the inverse of the translation amount, and averaged over the set of N images. Figure 19.12 shows a visual example of the shift-and-add process for one color channel. This is equivalent to the ML estimate of the image $\mathbf{z}$. As shown in Reference [47], if the robust L_1 data fidelity term

$$\sum_k \|\mathbf{y}_i(k) - \mathbf{DF}(k)\mathbf{z}_i\|_1 \;\; i = R, G, B \tag{19.21}$$

is used, then a pixelwise median operation over the set of frames implements the robust penalty function of Equation 19.14. This constitutes the robust shift-and-add algorithm, the result of which is a single image containing estimates of the blurry image $\mathbf{z}$.

As apparent in the fused image in Figure 19.12, the color sampling pattern after shift-and-add can be quite arbitrary depending on the relative motion of the low-resolution images. For some pixels we may not have any value, while for some we might have multiple measurements of one, two, or even all three color bands. Because this sampling pattern is arbitrary, we rely on the iterative deblurring and interpolation algorithm to restore the missing information. As an aside, Figure 19.12 partly explains the inferiority of the sequential strategy of first demosaicking the low-resolution images followed by superresolution. Image motion between captured frames will often provide data in some of the pixel

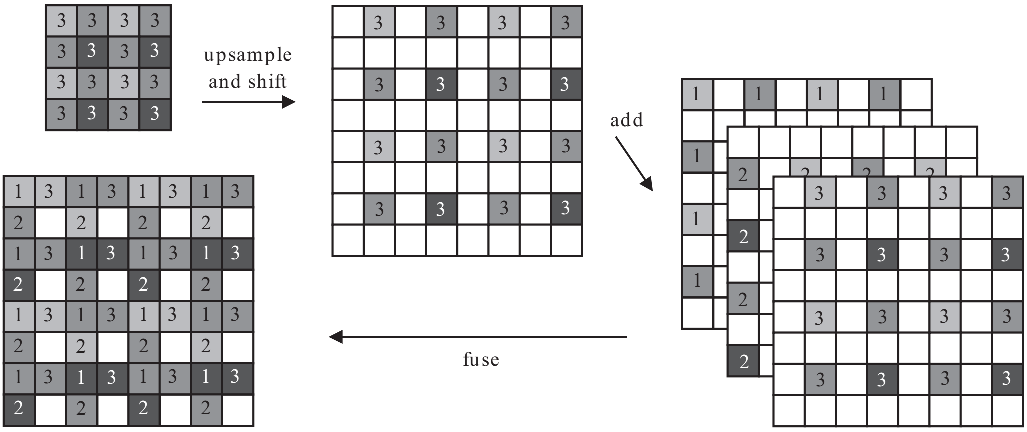


FIGURE 19.12 (See color insert.)

Diagram showing an example of the shift-and-add process. The input color channel image is upsampled by the resolution enhancement factor $r = 2$, and shifted according to the inverse of the image translation. This shifted image is added and averaged with the other $N-1$ (in this case 2) upsampled and shifted low-resolution images.

locations lacking color data in a single CFA image. Interpolating the missing pixels for each individual frame prematurely estimates the missing pixel data thereby limiting the final performance of multiframe resolution enhancement.

19.4.1.2 Deblurring and Interpolation Step

The second step uses the estimate of the blurry high resolution image $\hat{\mathbf{z}}$ to estimate the high resolution image $\mathbf{x}$. This step relies on a cost function very similar to that of Equation 19.12 with the exception that the data penalty term is replaced by

$$\Omega(\mathbf{x}, \hat{\mathbf{z}}) = \sum_{i=R,G,B} \|\Phi_i (\mathbf{H}\mathbf{x}_i - \hat{\mathbf{z}}_i)\|_2^2 + P_1(\mathbf{x}) + P_2(\mathbf{x}) + P_3(\mathbf{x}) \tag{19.22}$$

The matrix Φ_i (for $i = R, G, B$) is a diagonal matrix with diagonal values equal to the square root of the number of measurements that contributed to make each element of $\hat{\mathbf{z}}_i$. In this way, the data fidelity term applies no weight to pixel locations which have no observed data. On the other hand, those pixels which represent numerous measurements have a stronger influence on the data fidelity penalty function. Here, we observe that the data fidelity term no longer requires summation over the set of captured frames, saving a considerable amount of processing time as well as memory.

19.4.2 Dynamic Multiframe Demosaicking

Armed with the fast implementation of the multiframe demosaicking algorithm, we turn to apply this algorithm to video sequences. This is called *dynamic* multiframe demosaicking that produces an image *sequence* or video with higher resolution. A naive approach to this problem is to apply the static algorithm on a set of images while changing the reference frame. However, the memory and computational requirements for the static process are so taxing as to preclude its direct application to the dynamic case.

In contrast, we present a dynamic algorithm which takes advantage of the fact that the resulting high resolution image for the previous time frame $t-1$ helps predict the solution for the current time frame t. In this section, we replace the generic frame index k with t to indicate the temporal ordering of the low-resolution frames. A simple forward model capturing the temporal relationship of the low-resolution images is

$$\mathbf{x}(t) = \mathbf{F}(t)\mathbf{x}(t-1) + \mathbf{u}(t), \tag{19.23}$$

and

$$\mathbf{y}(t) = \mathbf{ADH}(t)\mathbf{x}(t) + \mathbf{v}(t). \tag{19.24}$$

In other words, the current frame is a shifted version of the previous frame with some additive noise with covariance $\mathbf{C}_u(t)$.

The equations given above describe a system in its *state-space* form, where the state is the desired ideal image. Thus, a Kalman-filter (KF) [76] formulation can be employed to recursively compute the optimal estimates $(\mathbf{x}(t), t \in \{1,...,N\})$ from the measurements $(\mathbf{y}(t), t \in \{1,...,N\})$, assuming that $\mathbf{D}, \mathbf{H}$, $\mathbf{F}(t)$, $\mathbf{C}_u(t)$, and $\mathbf{C}_v(t)$ (covariance of $\mathbf{v}(t)$) are known [35], [71], [72]. In the following, we study the application of the causal Kalman filter, where the estimates are obtained through on-line processing of an incoming sequence.[6]

To implement the dynamic multiframe demosaicking algorithm in a fast and efficient manner, we rewrite the state space model Equations 19.23 and 19.24 in terms of $\mathbf{z}$ ($\mathbf{z}(t) = \mathbf{Hx}(t)$) as

$$\mathbf{z}(t) = \mathbf{F}(t)\mathbf{z}(t-1) + \mathbf{e}(t), \tag{19.25}$$

and

$$\mathbf{y}(t) = \mathbf{ADz}(t) + \mathbf{v}(t). \tag{19.26}$$

Note that the first of the two equations is obtained by left multiplication of both sides of Equation 19.23 by $\mathbf{H}$ and using the fact that it commutes with $\mathbf{F}(t)$. Thus, the perturbation vector $\mathbf{e}(t)$ is a colored version of $\mathbf{u}(t)$, leading to $\mathbf{C}_e = \mathbf{HC}_u\mathbf{H}^T$ as its covariance matrix.

The following defines the forward Kalman propagation and update equations [35] that account for a causal (on-line) process. We assume that at time $t-1$ we already have the mean-covariance pair, $(\hat{\mathbf{z}}(t-1), \hat{\Pi}(t-1))$, and those should be updated to account for the information obtained at time t. We start with the covariance matrix update based on Equation 19.25:

$$\tilde{\Pi}(t) = \mathbf{F}(t)\hat{\Pi}(t-1)\mathbf{F}^T(t) + \mathbf{C}_{e(t)}, \tag{19.27}$$

where $\tilde{\Pi}(t)$ is the propagated covariance matrix (initial estimate of the covariance matrix at time t). The KF gain matrix is given by

$$\mathbf{K}(t) = \tilde{\Pi}(t)(\mathbf{AD})^T[\mathbf{C}_v(t) + \mathbf{AD}\tilde{\Pi}(t)\mathbf{D}^T]^{-1}. \tag{19.28}$$

Based on $\mathbf{K}(t)$, the updated state vector mean is computed by

$$\hat{\mathbf{z}}(t) = \mathbf{F}(t)\hat{\mathbf{z}}(t-1) + \mathbf{K}(t)[\mathbf{y}(t) - \mathbf{ADF}(t)\hat{\mathbf{z}}(t-1)]. \tag{19.29}$$

[6] A closely related noncausal processing mode, where every high-resolution reconstructed image is derived as an optimal estimate incorporating information from all the frames in the sequence, is studied in Reference [74].

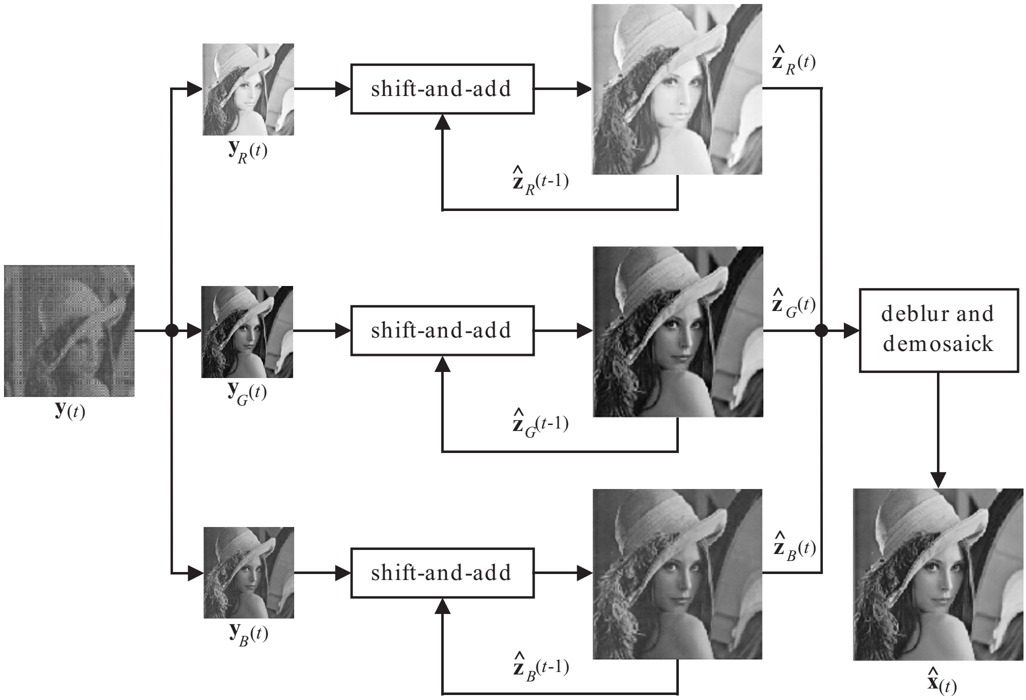


FIGURE 19.13

Block diagram representation of the fast dynamic multiframe demosaicking process.

The final stage requires the update of the covariance matrix, based on Equation 19.26 as follows:

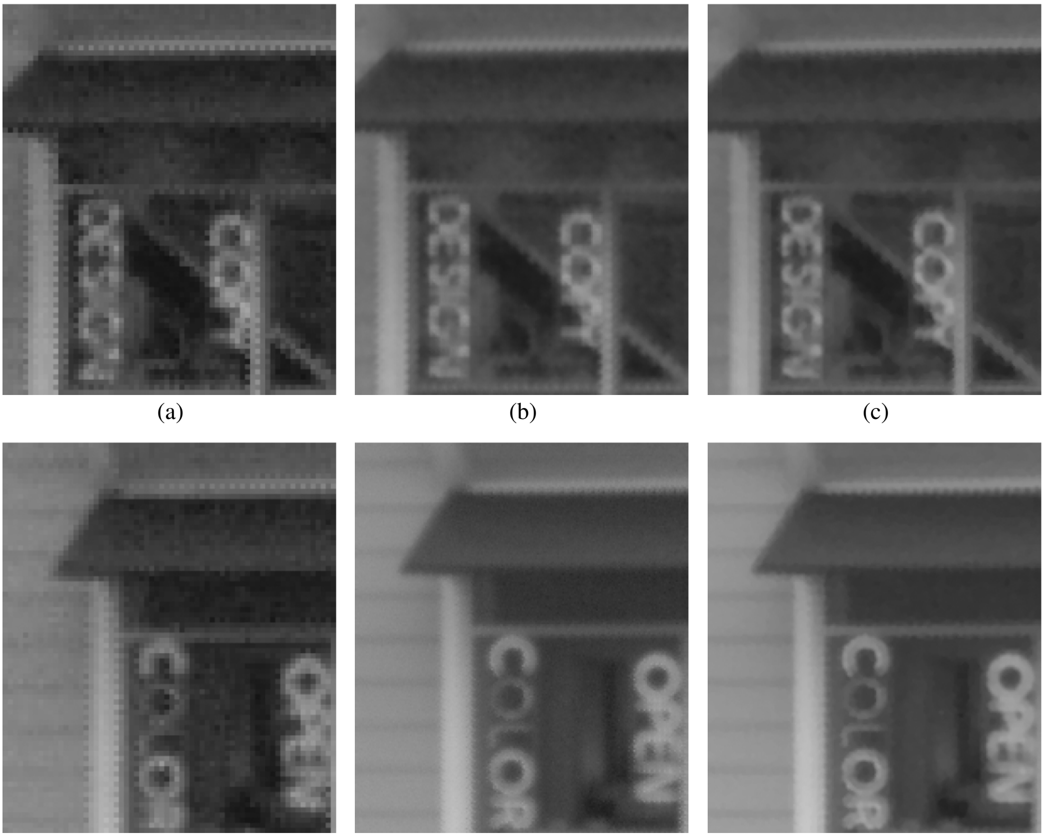
$$\hat{\Pi}(t) = \mathrm{Cov}\left(\hat{\mathbf{z}}(t)\right) = [\mathbf{I} - \mathbf{K}(t)\mathbf{A}\mathbf{D}]\tilde{\Pi}(t). \tag{19.30}$$

More on the meaning of these equations and how they are derived can be found in References [35] and [77].

While in general the above equations require the propagation of intolerably large matrices in time, if we refer to $\mathbf{C}_e(t)$ as a diagonal matrix, then $\tilde{\Pi}(t)$ and $\hat{\Pi}(t)$ are diagonal matrices. Following References [71] and [72], if we choose a matrix $\sigma_e^2\mathbf{I} \geq \mathbf{C}_e(t)$, it implies that $\sigma_e^2\mathbf{I} - \mathbf{C}_e(t)$ is a positive semi-definite matrix, and there is always a finite σ_e that satisfies this requirement. Replacing $\mathbf{C}_e(t)$ with this majorizing diagonal matrix, the new state-space system in Equations 19.25 and 19.26 simply assumes a stronger innovation process. The effect on the KF is to rely less on the temporal relation in Equation 19.25 and more on the measurements in Equation 19.26. Since $\mathbf{C}_e(t)$ is diagonal, Equations 19.27, 19.28, 19.29, and 19.30 can be implemented on an extremely fast pixel-by-pixel basis [74].

The output of the temporal Kalman filter equations (Equations 19.27, 19.28, 19.29, and 19.30) is an estimate of the blurred high resolution video sequence $\hat{\mathbf{z}}(t)$. We refer to this as the dynamic shift-and-add sequence. At this point, we apply the iterative deblurring and interpolation step described earlier. The iterative deblurring of the blurry image $\hat{\mathbf{z}}(t)$ for time t is accomplished by minimizing the cost function

$$\Omega(\mathbf{x}(t)) = J_1'(\mathbf{x}(t), \hat{\mathbf{z}}(t)) + P_1(\mathbf{x}(t)) + P_2(\mathbf{x}(t)) + P_3(\mathbf{x}(t)) \tag{19.31}$$



(a) (b) (c)

(d) (e) (f)

FIGURE 19.14

The images in the left column show frames #1 (top) and #69 (bottom) of the 74-frame low-resolution CFA image sequence. The second column shows the results after the recursive dynamic shift-and-add processing increasing their resolution by a factor of three in each direction. The right column shows the images after iteratively deblurring the dynamic shift-and-add images. We observe that the dynamic shift-and-add performs much of the work, while the subsequent deblurring restores the final details to the images.

using the forward shifted version of the previous estimate $\mathbf{F}(t)\hat{\mathbf{x}}(t-1)$ as the initial guess. By relying on the shifted high resolution estimate from the previous frame $\hat{\mathbf{x}}(t-1)$ as an initial guess, very few iterations are required for subsequent frames.

Figure 19.13 shows the block diagram describing fast dynamic multiframe demosaicking. The algorithm uses only the forward causal processing model; however, it can be extended to consider both forward filtering and smoothing, as described in References [65] and [78].

19.4.3 Example of Dynamic Multiframe Demosaicking

We present an example demonstrating the dynamic multiframe demosaicking algorithm. As before, we used seventy-four uncompressed, raw CFA images from a video camera (based on Zoran 2MP CMOS Sensors). The upper and lower images in the left column of Figure 19.14 show the low-resolution frames (frames number 1, and 69, respectively) demosaicked by the method in Reference [7]. The central images show the resulting color images after the recursive dynamic shift-and-add processing [74]. Some of the resolution has been restored at this point. The images in the right column show the iteratively deblurred versions of the shift-and-add images. Overall, the final image sequences show significant improvement over the original sequence.

19.5 Conclusion

In this chapter, we discussed superresolution and demosaicking as two important inverse problems in imaging. More significantly, we presented a unified approach for simultaneously solving these two interrelated problems. We illustrated how using the robust L_1 norm for the data fidelity term makes the proposed methods robust to errors in both measurements and in modelling. Furthermore, we used the bilateral regularization of the luminance term to achieve sharp reconstruction of edges; meanwhile the chrominance and spectral dependencies cost functions were tuned to remove color artifacts from the final estimate.

While the implementation of the resulting algorithms may appear complex at first blush, from a practical point of view, all matrix-vector operations in the proposed methods are implementable as simple image operators, making the methods practically feasible and possibly useful for implementation on specialized or general purpose hardware within commercially available cameras. Furthermore, as these operations are locally performed on pixel values on the HR grid, parallel processing may also be used to further increase the computational efficiency.

As we have pointed out, accurate subpixel motion estimation is an essential part of any image fusion process such as the proposed simultaneous multiframe superresolution and demosaicking. While multiscale methods based upon the optical flow or phase correlation principles are adequate, they are not specifically tailored or appropriate for estimating motions between color-filtered images. In fact, there is currently no significant body of literature that addresses the problem of estimating subpixel motion between Bayer filtered images. This gap in the literature needs to be filled, which could result in improved versions of the algorithms suggested in this chapter. More broadly, a new paradigm (Reference [25]) along the lines is needed in which accurate subpixel motion estimation is performed jointly inside and along with the overall superresolution/demosaicking problem, instead of what has traditionally been a preprocessing step.

Acknowledgments

We would like to thank Lior Zimet and Erez Galil from Zoran Corp. for providing the camera used to produce the raw CFA images. We would also like to thank Eyal Gordon from the Technion-Israel Institute of Technology for helping us capture the raw CFA images used in the Figure 19.1 experiment. We thank Prof. Ron Kimmel of the Technion for providing us with the code that implements the method from Reference [9]. We would like to thank Prof. Joseph Y. Lo and Prof. Cynthia A. Toth for helping us capturing and processing the low-resolution images in Figure 19.8. The image sequence used in the experiment of Figure 19.11 is courtesy of Adyoron Intelligent Systems Ltd., Tel Aviv, Israel.

This work was supported in part by the U.S. Air Force under Grant F49620-03-1-0387. Sina Farsiu was supported in part by the above and by the North Carolina Biotechnology Center's Collaborative Funding Grant (CFG).

Figure 19.7 and Figure 19.14 are reprinted from [79] with the permission of SPIE. Figure 19.6, Figure 19.9, Figure 19.10, and Figure 19.11 are reprinted from [65] with the permission of IEEE.

Appendix: Motion Estimation

The first step in solving the multiframe demosaicking problem is estimating the motion between the collection of CFA images. Numerous image registration techniques have been developed throughout the years [80]. Of these, optical flow [81], [82], and correlation-based methods [83] are among the most popular.

While the detailed description of these techniques is out of the scope of this chapter, we note that these methods are mainly developed to estimate the relative motion between *a pair* of frames. For cases where several images are to be registered with respect to each other (e.g., superresolution applications), two simple strategies are commonly used. The first is to register all frames with respect to a single reference frame [47]. This may be called the *anchoring* approach, as illustrated in Figure 19.15a. The choice of a reference or anchor frame is rather arbitrary, and can have a severe effect on the overall accuracy of the resulting estimates. This caveat aside, overall, this strategy is effective in cases where the camera motion is small and random (e.g., small vibrations of a gazing camera).

The other popular strategy is the *progressive* registration method [32], [78], where images

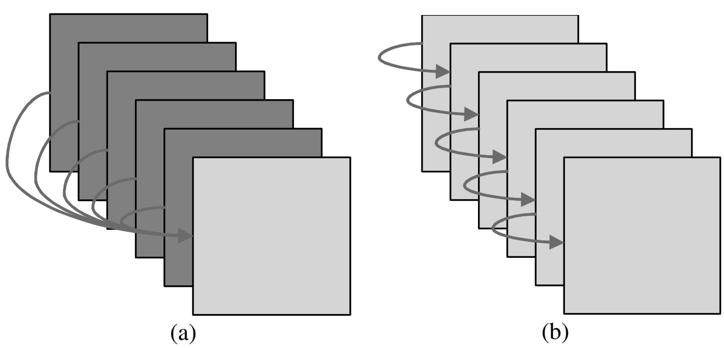


FIGURE 19.15

Common strategies used for registering frames of a video sequence. (a) Fixed reference ("anchored") estimation. (b) Pairwise ("progressive") estimation.

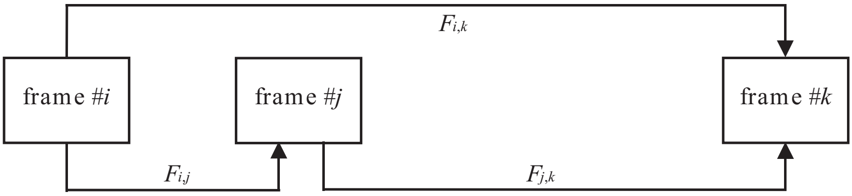


FIGURE 19.16

The consistent motion property dictates that the motion between any pair of frames must be the composition of the motion between two other pairs of frames.

in the sequence are registered in pairs, with one image in each pair acting as the reference frame. For instance, taking a causal view with increasing index denoting time, the ith frame of the sequence is registered with respect to the $(i+1)$th frame and the $(i+1)$th frame is registered with respect to the $(i+2)$th frame, and so on, as illustrated in Figure 19.15b. The motion between an arbitrary pair of frames is computed as the combined motion of the above incremental estimates. This method works best when the camera motion is smooth. However, in this method, the registration error between two "nearby" frames is accumulated and propagated when such values are used to compute motion between "far away" frames. Neither of the above approaches takes advantage of the important prior information available for the multiframe motion estimation problem. This prior information constrains the estimated motion vector fields as defined in the following.

To begin, let us define $\mathbf{F}_{i,j}$ as the operator which maps (registers) frames indexed i and j as follows:

$$\mathbf{y}_i = \mathbf{F}_{i,j}\{\mathbf{y}_j\},$$

where $\mathbf{y}_i$ and $\mathbf{y}_j$ are the lexicographic reordered vector representations of frames i and j. Now given a sequence of N frames, precisely $N(N-1)$ such operators can be considered. Regardless of considerations related to noise, sampling, and the finite dimensions of the data, there is an inherent intuitive relationship between these pair-wise registration operators. This condition dictates that the operator describing the motion between any pair of frames must be the composition of the operators between two other pairs of frames. More specifically, as illustrated in Figure 19.16, taking any triplet of frames i, j, and k, we have the motion consistency condition as:

$$\forall i,j,k \in \{1,...,N\}, \quad \mathbf{F}_{i,k} = \mathbf{F}_{i,j} \circ \mathbf{F}_{j,k}. \tag{19.32}$$

Using these types of constraints provides a means for reliable estimation of the sub-pixel motions and hence the warping operators. While a detailed discussion will be out of the scope of this paper, we shall note that the constrained motion property has been successfully implemented in different guises and for various applications such as mosaicking [84], motion estimation [85], [86], [87], and superresolution/demosaicking [25], [69], [82], [88], [89], [90], [91], [92].

References

[1] B.E. Bayer, "Color imaging array," U.S. Patent 3 971 065, July 1976.

[2] R. Lukac and K.N. Plataniotis, "Color filter arrays: Design and performance analysis," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 4, pp. 1260–1267, November 2005.

[3] N.R. Shah and A. Zakhor, "Resolution enhancement of color video sequences," *IEEE Transactions on Image Processing*, vol. 8, no. 6, pp. 879–885, June 1999.

[4] B.C. Tom and A. Katsaggelos, "Resolution enhancement of monochrome and color video using motion compensation," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 278–287, February 2001.

[5] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Model and Image Processing*, vol. 53, no. 3, pp. 231-239, May 1991.

[6] D.R. Cok, "Signal processing method and apparatus for sampled image signals," U.S. Patent 4 630 307, July 1987.

[7] C.A. Laroche and M.A. Prescott, "Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients," U.S. Patent 5 373 322, December 1994.

[8] J.F. Hamilton and J.E. Adams, "Adaptive color plan interpolation in single sensor color electronic camera," U.S. Patent 5 629 734, May 1997.

[9] R. Kimmel, "Demosaicing: Image reconstruction from color CCD samples," *IEEE Transactions on Image Processing*, vol. 8, no. 9, pp. 1221–1228, September 1999.

[10] L. Chang and Y.P. Tan, "Color filter array demosaicking: New method and performance measures," *IEEE Transactions on Image Processing*, vol. 12, no. 10, pp. 1194–1210, October 2002.

[11] K. Hirakawa and T.W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," in *Proceedings of the IEEE International Conference on Image Processing*, Barcelona, Spain, September 2003, vol. III, pp. 669–672.

[12] D. Keren and M. Osadchy, "Restoring subsampled color images," *Machine Vision and Applications*, vol. 11, no. 2, pp. 197–202, December 1999.

[13] Y. Hel-Or and D. Keren, "Demosaicing of color images using steerable wavelets," Technical report HPL-2002-206R1 20020830, HP Labs, Israel, citeseer.nj.nec.com/548392.html, 2002.

[14] W.K. Pratt, *Digital Image Processing*. New York: John Wiley & Sons, 3rd ed., 2001.

[15] D. Taubman, "Generalized Wiener reconstruction of images from colour sensor data using a scale invariant prior," in *Proceedings of the IEEE International Conference on Image Processing*, Vancouver, Canada, September 2000, vol. III, pp. 801–804.

[16] D.D. Muresan and T.W. Parks, "Demosaicing using optimal recovery," *IEEE Transactions on Image Processing*, vol. 14, no. 2, pp. 267–278, February 2005.

[17] B.K. Gunturk, Y.A. Altunbasak, and R.M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 997–1013, September 2002.

[18] S.C. Pei and I.K. Tam, "Effective color interpolation in CCD color filter arrays using signal correlation," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 503–513, June 2003.

[19] D. Alleysson, S. Süsstrunk, and J. Hérault, "Color demosaicing by estimating luminance and opponent chromatic signals in the Fourier domain," in *Proceedings of the IS&T/SID 10th Color Imaging Conference*, Scottsdale, AZ, USA, November 2002, pp. 331–336.

[20] R. Ramanath and W.E. Snyder, "Adaptive demosaicking," *Journal of Electronic Imaging*, vol. 12, no. 4, pp. 633–642, October 2003.

[21] L. Zhang and X. Wu, "Color demosaicing via directional linear minimum mean square-error estimation," *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2167–2178, December 2005.

[22] R. Lukac, K.N. Plataniotis, D. Hatzinakos, and M. Aleksic, "A new CFA interpolation framework," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 86, no. 7, pp. 1559–1579, July 2006.

[23] R. Lukac, K. Martin, and K.N. Plataniotis, "Colour-difference based demosaicked image postprocessing," *Electronics Letters*, vol. 39, no. 25, pp. 1805–1806, December 2003.

[24] R. Lukac, K. Martin, and K.N. Plataniotis, "Demosaicked image postprocessing using local color ratios," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 6, pp. 914–920, June 2004.

[25] D. Robinson, S. Farsiu, and P. Milanfar, "Optimal registration of aliased images using variable projection with applications to superresolution," *The Computer Journal, Special Issue on Super-Resolution in Imaging and Video*, doi: 10.1093/comjnl/bxm007, accepted for publication.

[26] S. Borman and R.L. Stevenson, "Super-resolution from image sequences — A review," in *Proceedings of the 1998 Midwest Symposium on Circuits and Systems*, Notre Dame, Indiana, April 1998, vol. V, p. 374.

[27] S.C. Park, M.K. Park, and M.G. Kang, "Super-resolution image reconstruction: A technical overview," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.

[28] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in superresolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, October 2004.

[29] T.S. Huang and R.Y. Tsai, Advances in Computer Vision and Image Processing, ch. Multiframe image restoration and registration, *Advances in Computer Vision and Image Processing*, Greenwich, Connecticut: JAI Press, 1984, vol. 1, pp. 317–339.

[30] D. Robinson, S. Farsiu, J.Y. Lo, P. Milanfar, and C.A. Toth, "Efficient multiframe registration of aliased X-ray images," in *Proceedings of the 41th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, USA, November 2007, pp. 215–219.

[31] N.K. Bose, H.C. Kim, and H.M. Valenzuela, "Recursive implementation of total least squares algorithm for image reconstruction from noisy, undersampled multiframes," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Minneapolis, MN, USA, April 1993, vol. V, pp. 269–272.

[32] L. Teodosio and W. Bender, "Salient video stills: Content and context preserved," in *Proceedings of the First ACM International Conference on Multimedia*, Anaheim, CA, USA, August 1993, vol. X, pp. 39–46.

[33] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1186–1193, August 2001.

[34] M.C. Chiang and T.E. Boulte, "Efficient super-resolution via image warping," *Image and Vision Computing*, vol. 18, no. 10, pp. 761–771, July 2000.

[35] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[36] M. Elad and A. Feuer, "Restoration of single super-resolution image from several blurred, noisy and down-sampled measured images," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1646–1658, December 1997.

[37] S. Peleg, D. Keren, and L. Schweitzer, "Improving image resolution using subpixel motion," *Pattern Recognition Letters*, vol. 5, no. 3, pp. 223–226, March 1987.

[38] N. Nguyen, P. Milanfar, and G.H. Golub, "A computationally efficient image superresolution algorithm," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 573–583, April 2001.

[39] N. Nguyen and P. Milanfar, "Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement," *IEEE Transactions on Image Processing*, vol. 10, no. 9, pp. 1299-1308, September 2001.

[40] S. Lertrattanapanich and N.K. Bose, "High resolution image formation from low resolution frames using Delaunay triangulation," *IEEE Transactions on Image Processing*, vol. 11, no. 12, pp. 1427–1441, December 2002.

[41] A.J. Patti, M.I. Sezan, and A.M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Transactions on Image Processing*, vol. 6, no. 8, pp. 1326–1333, March 1997.

[42] Y. Altunbasak, A. Patti, and R. Mersereau, "Super-resolution still and video reconstruction from MPEG-coded video," *IEEE Transactions on Circuits and System Video Technology*, vol. 12, no. 4, pp. 217–226, April 2002.

[43] O. Haik and Y. Yitzhaky, "Super-resolution reconstruction of a video captured by a vibrated TDI camera," *Journal of Electronic Imaging*, vol. 15, no. 2, pp. 023006-1-12, April-June 2006.

[44] C.B. Atkins, C.A. Bouman, and J.P. Allebach, "Tree-based resolution synthesis," in *Proceedings of the IS&T Conf. on Image Processing, Image Quality, Image Capture Systems*, Savannah, GA, USA, April 1999, pp. 405–410.

[45] S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 9, pp. 1167–1183, September 2002.

[46] A. Zomet, A. Rav-Acha, and S. Peleg, "Robust super resolution," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA, December 2001, vol. I, pp. 645–650.

[47] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Fast and robust multi-frame super-resolution," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1327–1344, October 2004.

[48] Z.A. Ivanovski, L. Panovski, and L.J. Karam, "Robust super-resolution based on pixel-level selectivity," in *Proceedings of the 2006 SPIE Conference on Visual Communications and Image Processing*, San Jose, CA, USA, January 2006, vol. 6077, pp. 607707-1–8.

[49] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, November 1992.

[50] T.F. Chan, S. Osher, and J. Shen, "The digital TV filter and nonlinear denoising," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 231–241, February 2001.

[51] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proceedings of the IEEE International Conference on Computer Vision*, New Delhi, India, January 1998, pp. 836–846.

[52] M. Elad, "On the bilateral filter and ways to improve it," *IEEE Transactions on Image Processing*, vol. 11, no. 10, pp. 1141–1151, October 2002.

[53] T.Q. Pham, L.J. van Vliet, and K. Schutte, "Robust fusion of irregularly sampled data using adaptive normalized convolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, id. 83268, ps. 12, 2006.

[54] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 349–366, February 2007.

[55] M.B. Ezra, A. Zomet, and S.K. Nayar, "Video super-resolution using controlled subpixel detector shifts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 977–987, June 2004.

[56] C.A. Segall, R. Molina, A. Katsaggelos, and J. Mateos, "Bayesian high-resolution reconstruction of low-resolution compressed video," in *Proceedings of the IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 2001, vol. II, pp. 25–28.

[57] C.A. Segall, A.K. Katsaggelos, R. Molina, and J. Mateos, "Bayesian resolution enhancement of compressed video," *IEEE Transactions on Image Processing*, vol. 13, no. 7, pp. 898–911, July 2004.

[58] S. Borman, "Topics in multiframe superresolution restoration," University of Notre Dame, Ph.D. Thesis, Notre Dame, IN, USA, May 2004.

[59] M.G. Kang and S. Chaudhuri, "Super-resolution image reconstruction," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–36, May 2003.

[60] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, August 2004.

[61] S. Farsiu, M. Elad, and P. Milanfar, "Multi-frame demosaicing and super-resolution from under-sampled color images," *Proceedings of the 2004 IS&T/SPIE 16th Annual Symposium on Electronic Imaging*, vol. 5299, pp. 222–233, January 2004.

[62] T. Gotoh and M. Okutomi, "Direct super-resolution and registration using raw CFA images," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Washington, D.C., USA, July 2004, vol. II, pp. 600–607.

[63] R. Lukac and K.N. Plataniotis, "Fast video demosaicking solution for mobile phone imaging applications," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 2, pp. 675–681, May 2005.

[64] R. Sasahara, H. Hasegawa, I. Yamada, and K. Sakaniwa, "A color super-resolution with multiple nonsmooth constraints by hybrid steepest descent method," in *Proceedings of the International Conference on Image Processing*, Genoa, Italy, September 2005, vol. I, pp. 857–860.

[65] S. Farsiu, M. Elad, and P. Milanfar, "Multiframe demosaicing and super-resolution of color images," *IEEE Transactions on Image Processing*, vol. 15, no. 1, pp. 141-159, January 2006.

[66] R. Lukac and K.N. Plataniotis, "Adaptive spatiotemporal video demosaicking using bidirectional multistage spectral filters," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 2, pp. 651–654, May 2006.

[67] P. Vandewalle, K. Krichane, D. Alleysson, and S. Süsstrunk, "Joint demosaicing and super-resolution imaging from a set of unregistered aliased images," in *IS&T SPIE Electronic Imaging: Digital Photography III*, San Jose, CA, USA, January 2007, vol. 6502, ps. 12.

[68] D. Keren and A. Gotlib, "Denoising color images using regularization and correlation terms," *Journal of Visual Communication and Image Representation*, vol. 9, no. 4, pp. 352–365, December 1998.

[69] S. Farsiu, M. Elad, and P. Milanfar, "Constrained, globally optimal, multi-frame motion estimation," in *Proceedings of the 2005 IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, July 2005, pp. 1396–1401.

[70] J.R. Bergen, P. Anandan, K.J. Hanna, and R. Hingorani, "Hierarchical model-based motion estimation," *Proceedings of the European Conference on Computer Vision*, vol. 20, pp. 237–252, May 1992.

[71] M. Elad and A. Feuer, "Super-resolution reconstruction of image sequences," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 817–834, September 1999.

[72] M. Elad and A. Feuer, "Superresolution restoration of an image sequence: Adaptive filtering approach," *IEEE Transactions on Image Processing*, vol. 8, no. 3, pp. 387–395, March 1999.

[73] C. Newland, D. Gray, and D. Gibbins, "Modified Kalman filtering for image super-resolution," in *Proceedings of the 2006 Conference on Image and Vision Computing*, Great Barrier Island, New Zealand, November 2006, pp. 79–84.

[74] S. Farsiu, M. Elad, and P. Milanfar, "Video-to-video dynamic super-resolution for grayscale and color sequences," *EURASIP Journal on Applied Signal Processing*, vol. 2006, id. 61859, 2006.

[75] M. Elad and Y. Hel-Or, "A fast super-resolution reconstruction algorithm for pure translational motion and common space invariant blur," *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1186–1193, August 2001.

[76] B. Anderson and J. Moore, *Optimal Filtering*. Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

[77] A.H. Jazwinski, *Stochastic Processes and Filtering Theory: The Visual Neurosciences*. New York: Academic Press, 1970.

[78] S. Farsiu, D. Robinson, M. Elad and P. Milanfar, "Dynamic demosaicing and color superresolution of video sequences," in *Proceedings of the SPIE Conference on Image Reconstruction from Incomplete Data III*, Denver, CO, USA, October 2004, vol. 5562, pp. 169–178.

[79] S. Farsiu, M. Elad, and P. Milanfar, "A practical approach to super-resolution," in *Proceedings of the SPIE Conference on Visual Communications and Image Processing*, San Jose, CA, USA, January 2006, vol. 6077, pp. 24–38.

[80] L.G. Brown, "A survey of image registration techniques," *ACM Computing Surveys*, vol. 24, no. 8, pp. 325–376, December 1992.

[81] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proceedings of the DARPA Image Understanding Workshop*, 1981, pp. 121–130.

[82] W.Y. Zhao and H.S. Sawhney, "Is super-resolution with optical flow feasible?" in *Proceedings of the European Conference on Computer Vision*, Copenhagen, Denmark, May 2002, vol. I, pp. 599–613.

[83] M. Alkhanhal, D. Turaga, and T. Chen, "Correlation based search algorithms for motion estimation," in *Proceedings of the Picture Coding Symposium*, Portland, OR, USA, April 1999, pp. 99–102.

[84] H.S. Sawhney, S.C. Hsu, and R. Kumar, "Robust video mosaicing through topology inference and local to global alignment," in *Proceedings of the European Conference on Computer Vision*, Freiburg, Germany, June 1998, vol. II, pp. 103–119.

[85] V.M. Govindu, "Combining two-view constraints for motion estimation," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Kauai, HI, USA, July 2001, vol. II, pp. 218–225.

[86] V.M. Govindu, "Lie-algebraic averaging for globally consistent motion estimation," in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Washington, D.C., USA, July 2004, vol. I, pp. 684–691.

[87] Y. Sheikh, Y. Zhai, and M. Shah, "An accumulative framework for the alignment of an image sequence," in *Proceedings of the Asian Conference on Computer Vision*, Jeju, Korea, January 2004.

[88] P. Vandewalle, S. Süsstrunk, and M. Vetterli, "A frequency domain approach to registration of aliased images with application to super-sesolution," *EURASIP Journal on Applied Signal Processing*, vol. 2006, id. 71459, 2006.

[89] H. Foroosh, J. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Transactions on Image Processing*, vol. 11, no. 3, pp. 188–200, March 2002.

[90] D. Robinson and P. Milanfar, "Statistical performance analysis of super-resolution," *IEEE Transactions on Image Processing*, vol. 15, no. 6, pp. 1413–1428, June 2006.

[91] N.A. Woods, N.P. Galatsanos, and A.K. Katsaggelos, "Stochastic methods for joint registration, restoration, and interpolation of multiple undersampled images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 1, pp. 201–213, January 2006.

[92] J. Chung, E. Haber, and J. Nagy, "Numerical methods for coupled super-resolution," *Inverse Problems*, vol. 22, no. 4, pp. 1261–1272, August 2006.