14 Fusion of Blurred Images

Filip Šroubek and Jan Flusser

CONTENTS

I. INTRODUCTION

In general, the term *fusion* means an approach to extraction of information spontaneously adopted in several domains. The goal of image fusion is to integrate combinations of complementary multisensor, multitemporal, and multiview information into one new image containing information of which the quality could not be achieved otherwise. The term *quality* depends on the application requirements.

Image fusion has been used in many application areas. In remote sensing and in astronomy,[1,2] multisensor fusion is used to achieve high spatial and spectral resolutions by combining images from two sensors, one of which has high spatial resolution and the other, high spectral resolution. Numerous fusion applications have appeared in medical imaging (see Ref. 3 or 4 for instance) such as simultaneous evaluation of a combination of computer tomography (CT), nuclear magnetic resonance (NMR), and positron emission tomography (PET) images. In the case of multiview fusion, a set of images of the same scene taken by the same sensor but from different viewpoints is fused to obtain an image with higher resolution than the sensor normally provides, or to recover the three-dimensional representation of the scene (shape from stereo). The multitemporal approach recognizes two different aims. Images of the same scene are acquired at different time instances either to find and evaluate changes in the scene or to obtain a less degraded image of the scene. The former aim is common in medical imaging, especially in change detection of organs and tumors, and in remote sensing for monitoring land or forest exploitation. The acquisition period is usually months or years. The latter aim requires the different measurements to be much closer to each other, typically in the scale of seconds, and possibly under different conditions. Our motivation for this work came from this area.

We assume that several images of the same scene called *channels* are available. We further assume all channels were acquired by the same sensor (or by different sensors of the same type) but under different conditions and acquisition parameters. Thus, all channels are of the same modality and represent similar physical properties of the scene.

Since imaging sensors and other devices have their physical limits and imperfections, the acquired image represents only a degraded version of the original scene. Two main categories of degradations are recognized: color (or brightness) degradations and geometric ones. The former degradations are caused by such factors as incorrect focus, motion of the scene, media turbulence, noise, and limited spatial and spectral resolution of the sensor; they usually result in a blurring of the image. The latter degradations originate from the fact that each image is a two-dimensional projection of a three-dimensional world. They cause deformations of object shapes and other spatial distortions of the image.

Individual channels are supposed to be degraded in different ways because of differences in acquisition parameters and imaging conditions (see Figure 14.1 for multichannel acquisition scheme). There are many sources of corruption and distortion that we have to cope with. Light rays (or other types of electromagnetic waves) reflected by objects on the scene travel to measuring sensors through a transport medium, for example, the atmosphere. Inevitably, each transport medium modifies the signal in some way. The imaging system is thus subject to blurring due to the medium's rapidly changing index of refraction, the finite broadcast bandwidth and the object motion. The source of corruption and its characteristics are often difficult to predict. In addition, the signal is corrupted inside a focusing set after reaching the sensor. This degradation is inherent to the system and cannot be bypassed, but it can often be measured and accounted for; typical examples are a range of lens imperfections. Finally, the signal must be stored on photographic material or first digitized with Charge-Coupled Devices (CCDs) and then stored. In both cases the recording exhibits a number

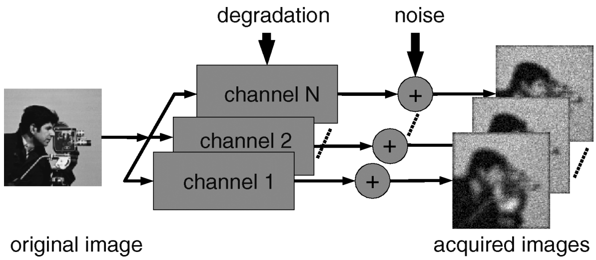


FIGURE 14.1 Multichannel acquisition model: the original scene is captured by N different channels which are subject to various degradations.

of degradations. Digital imaging systems suffer from low resolution and low sensitivity to the input signal, which are imposed by a finite number of intensity levels and a finite storage capacity. In analog systems, resolution artifacts are caused by the limited size of photographic material grain. Random noise is another crucial factor that severely affects the quality of image acquisition. In all real applications, measurements are degraded by noise. By utilizing suitable measuring techniques and appropriate devices, it can be considerably diminished, but unfortunately never cancelled.

Analysis and interpretation of degraded images is the key problem in real applications, because the degradations are, in principle, inevitable. A very promising approach to image quality enhancement is to fuse several channels with different degradations together in order to extract as much useful information as possible.

II. MULTICHANNEL IMAGE ACQUISITION MODELS

Regardless of its particular type, image degradations can be mathematically described by an operator based on an ideal representation of the scene. More formally, let $u(x, y)$ be an ideal image of the scene and let $z_1(x, y), \ldots, z_N(x, y)$ be acquired channels. The relation between each z_i and u is expressed as

$$z_i(x, y) = D_i(u(x, y)) + n_i(x, y) \tag{14.1}$$

where D_i is an unknown operator describing the image degradations of the ith channel and n_i denotes additive random noise. In the ideal situation, D_i would equal identity and n_i would be zero for each i. The major goal of the fusion is to obtain an image $\hat{u}$ as a “good estimate” of u; that means $\hat{u}$, in some sense, should be a better representation of the original scene than each individual channel z_i.

The fusion methodology depends significantly on the type of degradation operators D_i. In this work we focus on the cases where each operator D_i is a composition of image blurring and of geometric deformations caused by imaging geometry.

Under these assumptions, Equation 14.1 becomes

$$z_i(\tau_i(x,y)) = \int h_i(x,y,s,t)u(s,t)\mathrm{d}s\,\mathrm{d}t + n_i(x,y) \tag{14.2}$$

where $h_i(x,y,s,t)$ is called the *point spread function* (PSF) of the ith imaging system at location (x,y) and τ_i stands for the co-ordinate transform, describing geometric differences between the original scene and the ith channel (in simple cases, τ_i is limited to rotation and translation, but in general complex nonlinear deformations may be present too). Having N channels, Equation 14.2 can be viewed as a system of N integral equations of the first kind. Even if all h_is were known and neither geometric deformations nor noise were present, this system would not be generally solvable. In the sequel we simplify the model (2) by additional constraints and we show how to fuse the channels (that is, how to estimate the original scene) in these particular cases. The constraints are expressed as some restrictive assumptions on the PSFs and on the geometric deformations. We review five basic cases covering most situations occurring in practice.

III. PIECEWISE IDEAL IMAGING

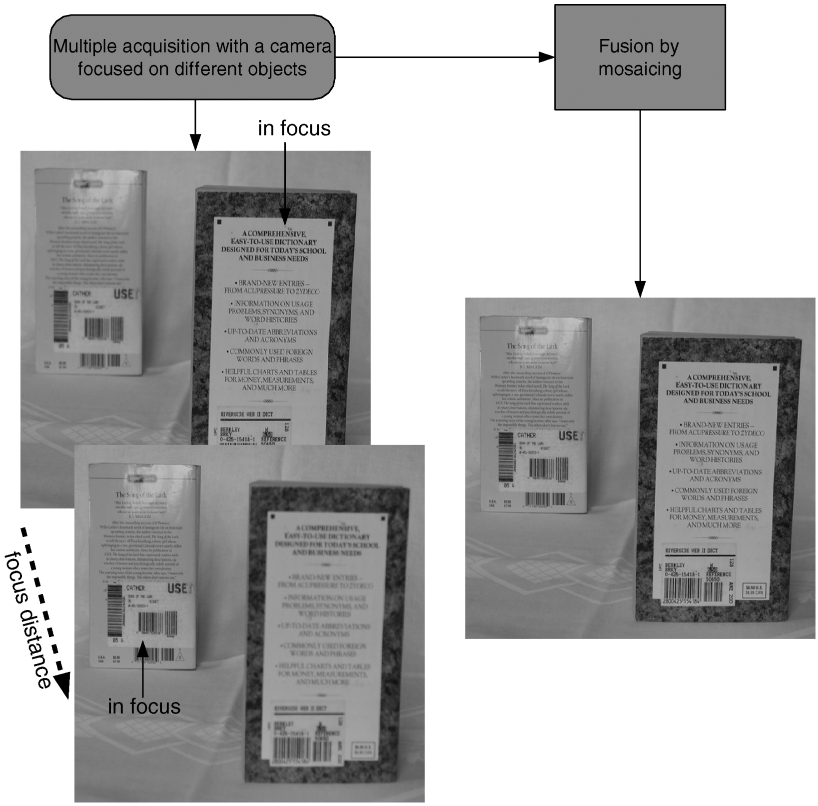
In this simplest model, the PSF of each channel is supposed to be piecewise space-invariant and every point (x,y) of the scene is assumed to be acquired undistorted in (at least) one channel. No geometric deformations are assumed.

More precisely, let $\Omega = \bigcup_{k=1}^{K} \Omega_k$ be a support of image function $u(x,y)$, where Ω_k are its disjunct subsets. Let h_i^k be a local PSF acting on the region Ω_k in the ith channel. Since every h_i^k is supposed to be space-invariant (that is, $h_i^k(x,y,s,t) = h_i^k(x-s,y-t)$), the imaging model is defined as

$$z_i(x,y) = (u * h_i^k)(x,y) \Leftrightarrow (x,y) \in \Omega_k \tag{14.3}$$

where $*$ stands for convolution and for each region Ω_k there exists channel j such that $h_j^k(x-s,y-t) = \delta(x-s,y-t)$.

This model is applicable in so-called *multifocus imaging*, when we photograph a static scene with a known piecewise-constant depth and focus channel-by-channel on each depth level. Image fusion then consists of comparing the channels in the image domain[5,6] or in the wavelet domain,[7,8] identifying the channel in which the pixel (or the region) is depicted undistorted and, finally, mosaicing the undistorted parts (no deconvolution is performed in this case, see Figure 14.2). To find the undistorted channel for the given pixel, a local focus measure is calculated over the pixel neighborhood and the channel which maximizes the focus measure is chosen. In most cases, the focus measures used are based on the idea of measuring the quantity of high frequencies of the image. It corresponds with an intuitive expectation that the blurring suppresses high frequencies regardless of the particular PSF. Image variance,[9] energy of a Fourier spectrum,[10] norm of image gradient,[9] norm of image Laplacian,[9] image



Q3 **FIGURE 14.2** Two-channel piecewise ideal imaging: in each channel, one book is in focus while the other one is out of focus. Image fusion is performed by mosaicing the channel regions which are in focus.

moments,[11] and energy of high-pass bands of a wavelet transform[7,8,12] belong to the most popular focus measures. Fusion in the image domain is seriously affected by the size of the neighborhood on which the focus measure is calculated. On the other hand, fusion in the wavelet domain is very sensitive to translation changes in the source images.

A. Application in Confocal Microscopy

A typical application area where piecewise ideal imaging appears is confocal microscopy of three-dimensional samples. Since the microscope has a very narrow depth of field, several images of the sample differing from each other by the focus distance are taken. Each of them shows in focus only the parts of the sample that are a certain distance from the lens, while other parts are blurred by an out-of-focus blur of various extents. These image layers form the so-called *stack image* of the sample. To obtain a focused image of the whole sample is beyond the scope of the microscope; the only possibility to get it employs a fusion of the stack image.

If the focal step used in the acquisition process is less than or equal to the depth of field of the microscope, then the assumptions of piecewise ideal imaging are fulfilled and fusion by mosaicing the undistorted parts of the individual layers can be applied.

A crucial question is how to find, for each pixel, the layer in which the given pixel (together with its neighborhood) depicted is least distorted or undistorted. Among the focus measures mentioned above, wavelet-based methods gave the best results. Their common idea is to maximize the energy of high-pass bands. Most wavelet-based focus measures ignore low-pass band(s), but Kautsky[12] pointed out that the energy of low-pass bands also reflects the degree of image blurring and, that considering both low-pass and high-pass bands increases the discrimination power of the focus measurement. We adopted and modified the idea from Ref. 12, and proposed to use a product of energies contained in low-pass and high-pass bands as a local focus measure

$$\varrho_i(x,y) = \|w_L(z_i)\|^2 \|w_H(z_i)\|^2 \tag{14.4}$$

Both energies are calculated from a certain neighborhood of the point (x,y); high-pass band energy $\|w_H(z_i)\|^2$ is calculated as the mean from three high-pass bands (in this version, we used decomposition to depth one only).

The fusion of the multifocus stack $\{z_1, \ldots, z_N\}$ is conducted in the wavelet domain as follows. First, we calculate the wavelet decomposition of each image z_i. Then a decision map $M(x,y)$ is created in accordance with a max-rule $M(x,y) = \arg\max\rangle_i(x,y)$. The decision map is the size of the subband, that is, a quarter of the original image, and it tells us from which image the wavelet coefficients should be used. The decision map is applied to all four bands and, finally, the fused image is obtained by inverse wavelet transform.

The performance of this method is shown here in fusion of microscopic images of a unicellular water organism (see Figure 14.3). The total number of the stack layers was 20; three of them are depicted. The fused image is shown on the bottom right.

In several experiments similar to this one we tested various modifications of the method. The definition 14.4 can be extended for deeper decompositions but it does not lead to any improvement. We compared the performance of various wavelets and studied the influence of the wavelet length. Short wavelets are too sensitive to noise, while long wavelets do not provide enough discrimination power. Best results on this kind of data were obtained by biorthogonal wavelets. Another possible modification is to calculate the decision map separately for each band but this also did not result in noticeable refinement. We also tested the performance of other fusion techniques. The proposed method always produced a visually sharp image, assessed by the observers as the best or one of the best.

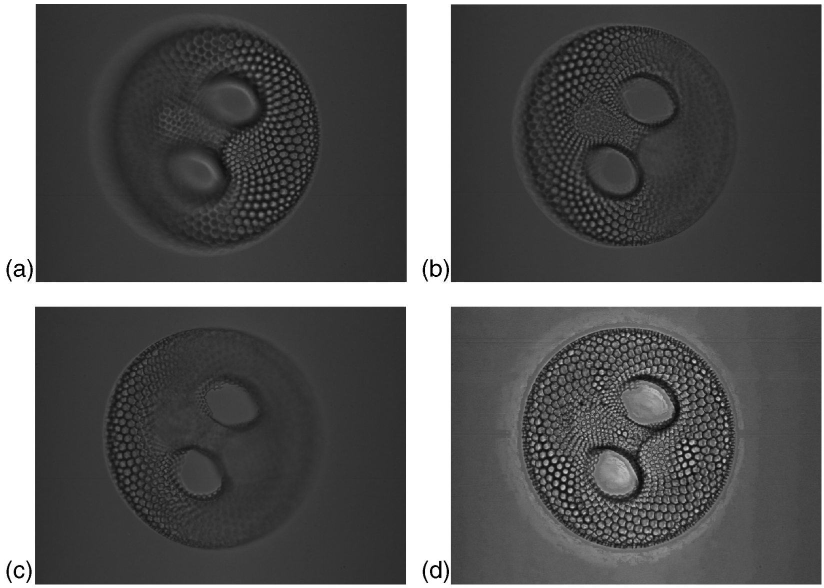


FIGURE 14.3. Fusion of a multifocus microscope image: three out of 20 layers in the multifocus stack and the result of the fusion in the wavelet domain (bottom right).

We can thus conclude that this method is very suitable for fusion of multifocus microscope images.

IV. UNIFORMLY BLURRED CHANNELS

An acquisition model with uniformly blurred channels assumes that every PSF h_i is space-invariant within the channel, that is, $h_i(x, y, s, t) = h_i(x - s, y - t)$. Equation 14.2 then turns into the form of "traditional" convolution in each channel with no geometric deformations:

$$z_i(x, y) = (u * h_i)(x, y) + n_i(x, y) \tag{14.5}$$

This model describes, for instance, photographing a flat static scene with different (but always wrong) focuses, or repetitively photographing a scene through a turbulent medium whose optical properties change between the frames (see Figure 14.4). The image fusion is performed via multichannel blind deconvolution (MBD). It should be noted that if the PSFs were known, then this task would turn into the classical problem of *image restoration* which has been considered in numerous publications, see Ref. 13 for a survey.

Blind deconvolution in its most general form is an unsolvable problem. All methods proposed in the literature inevitably make some assumptions about the PSFs h_i and/or the original image $u(x, y)$. Different assumptions give rise to

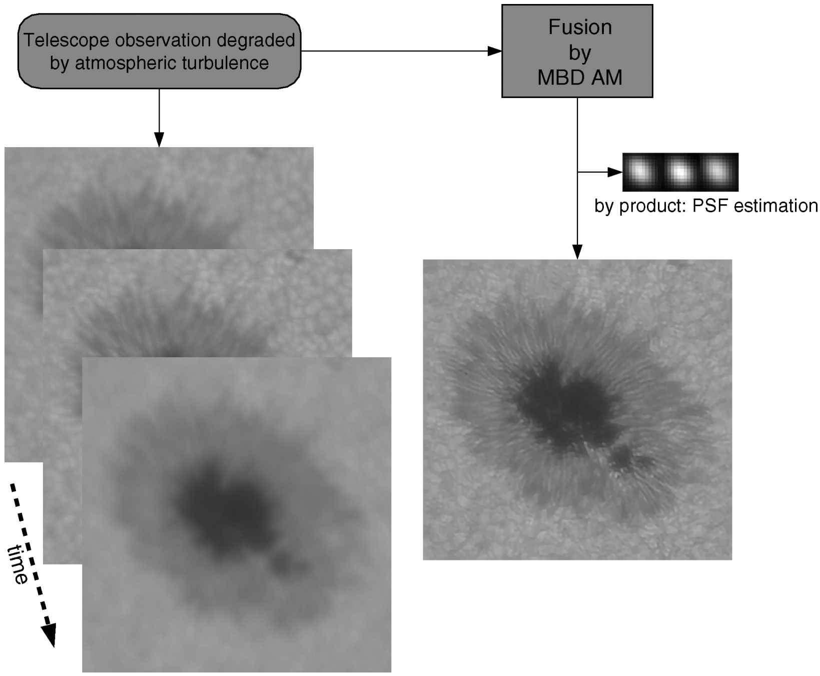


FIGURE 14.4 Images of a sunspot taken by a ground-based telescope and blurred due to perturbations of wavefronts in the Earth atmosphere. The perturbations vary in time which leads to different blurring of the individual frames. The resulting image was fused by the MBD–AM algorithm described in Section IV.A.

various deconvolution methods. There are two basic approaches to solving the MBD problem. The first one is to separately treat each channel by any single-channel deconvolution method and then to combine the results; the other is to employ deconvolution methods that are multichannel in their nature.

Numerous single-channel blind deconvolution methods have been published extensively in the literature in the last two decades (see Ref. 14 or 15 for a basic survey). However, their adaptation to the MBD problem cannot reach the power of intrinsic multichannel methods and this approach seems to be a "dead-end".

The development of intrinsic multichannel methods has begun just recently. One of the earliest methods[16] was designed particularly for images blurred by atmospheric turbulence. Harikumar and Bresler[17,18] proposed indirect algorithms (EVAM), which first estimate the PSFs and then recover the original image by standard nonblind methods. Giannakis and Heath[19,20] (and at the same time Harikumar and Bresler[21]) developed another indirect algorithm based on Bezout's identity of coprime polynomials which finds inverse filters and, by convolving the filters with the observed images, recovers the original image. Pillai and Liang[22] have proposed another intrinsically multichannel method based on the greatest common divisor which is, unfortunately, even less numerically stable than the previous methods. Pai and Bovik[23,24] came with two direct multichannel restoration algorithms that directly estimate the original image from the null space or from the range of a special matrix. To reach higher robustness, Šroubek[25] proposed an iterative deconvolution method which employs anisotropic regularization of the image and between-channel regularization of the PSFs.

A. Alternating Minimization Algorithm

In this section, we present an alternating minimization algorithm for multichannel blind deconvolution (MBD–AM) and we demonstrate that it is a powerful tool for image fusion in the case of uniformly blurred channels.

Since the blind deconvolution problem is ill posed with respect to both u and h_i, a constrained minimization technique is required to find the solution of Equation 14.5. Constraints are built on prior knowledge that we have about the system. Typical assumptions valid for the majority of real acquisition processes are the following: n_i is supposed to have zero mean and the same variance σ^2 in each channel, and the PSFs are supposed to preserve the overall brightness (mean intensity) of the image. The imposed constraints then take the forms

$$\frac{1}{|\Omega|}\int_{\Omega}(h_i * u - z_i)^2 \mathrm{d}x = \sigma^2 \tag{14.6}$$

$$\int_{\Omega} h(x)\mathrm{d}x = 1 \tag{14.7}$$

(To simplify the notation, we drop the two-dimensional co-ordinates (x, y) or, if required, we only write x.)

Let $Q(u)$ and $R(h)$ denote some regularization functionals of the estimated original image u and blurs $h \equiv \{h_1, \ldots, h_N\}$, respectively. The constrained minimization problem is formulated as $\min_{u,h} Q(u) + \gamma R(h)$ subject to Equation 14.6 and Equation 14.7. The unconstrained optimization problem, obtained by means of Lagrange multipliers, is to find u and h which minimize the functional

$$E(u, h) = \frac{1}{2}\sum_{i=1}^{N}\|h_i * u - z_i\|^2 + \lambda Q(u) + \gamma R(h) \tag{14.8}$$

where λ and γ are positive parameters which penalize the regularity of the solutions u and h. The crucial questions are how to construct functionals Q and R and whether the global minimum can be reached. We propose an alternating minimization algorithm that iteratively searches for a minimum of Equation 14.8. Q1 Constraint Equation 14.7 was dropped since it will be automatically satisfied in the algorithm if the initial blurs satisfy the constraint. We now proceed the discussion with possible choices of $Q(u)$ and $R(h)$.

1. Regularization of the Image $Q(u)$

Regularization of Equation 14.5 with respect to the image function can adopt various forms. The classical approach of Tichonov chooses $Q(u) = \int_\Omega |\nabla u(x)|^2 \mathrm{d}x$, where ∇u denotes the gradient of u. Apart from easy implementation, this regularization is not suitable, since the L_2 norm of the image gradient penalizes too much the gradients corresponding to edges and an oversmoothing effect is observed. In real images, object edges create sharp steps that appear as discontinuities in the intensity function. It is the space of bounded variation (BV) functions that is widely accepted as a proper setting for real images. Rudin[26] first demonstrated very good anisotropic denoising properties of the total variation (TV) $Q_{\mathrm{TV}}(u) = \int |\nabla u(x)| \mathrm{d}x$. Existence and uniqueness of the minimum of TV is possible only in the BV-space, in which case ∇u denotes the gradient of u in the distributional sense. The same holds true for a more general case of convex functions of measures

$$Q_\phi(u) = \int \phi(|\nabla u(x)|)\mathrm{d}x$$

where ϕ is a strictly convex, nondecreasing function that grows at most linearly. Examples of $\phi(s)$ are s(TV), $\sqrt{1+s^2} - 1$ (hyper-surface minimal function) or $\log(\cosh(s))$. For nonconvex functions nothing can be said about the existence of the minimum. Nevertheless, nonconvex functions, such as $\log(1+s^2)$, $s^2/(1+s^2)$ or $\arctan(s^2)$ (Mumford–Shah functional[27]), are often used since they provide better results for segmentation problems.

2. Regularization of the Blurs $R(h)$

Regularization of the blurs h_i's directly follows from our model, Equation 14.5, and can be derived from the mutual relations of the channels. The blurs are assumed to have finite support S of the size (s_1, s_2) and certain channel disparity is necessary. The disparity is defined as weak coprimeness of the channel blurs, which states that the blurs have no common factor except a scalar constant. In other words, if the channel blurs can be expressed as a convolution of two subkernels then there is no subkernel that is common to all blurs. An exact definition of weakly coprime blurs can be found in Ref. 20. The channel coprimeness is satisfied for many practical cases, since the necessary channel disparity is mostly guaranteed by the nature of the acquisition scheme and random processes therein. We refer the reader to Ref. 18 for a relevant discussion.

Under the assumption of channel coprimeness, we can see that any two correct blurs h_i and h_j satisfy $\|z_i * h_j - z_j * h_i\|^2 = 0$ if the noise term in Equation 14.5 is omitted. We therefore propose to regularize the blurs by

$$R(h) = \frac{1}{2} \sum_{1 \le i < j \le N} \|z_i * h_j - z_j * h_i\|^2 \tag{14.9}$$

This regularization term does not penalize spurious factors, that is, $f * h_i$ for any factor f are all equivalent. We see that the functional $R(h)$ is convex but far from strictly convex. The dimensionality of the null space of $R(\hat{h})$ is proportional to the degree of size overestimation of $\hat{h}_i$ with respect to the size of the original blurs h_i's. Therefore to use the above regularization, we have to first estimate S of the original blurs and impose this support constraint in R. The size constraint is imposed automatically in the discretization of R, which is perfectly plausible since the calculations are done in the discrete domain anyway. An exact derivation of the size of the null space is given in Ref. 18.

3. Iterative Minimization Algorithm

We consider the following minimization problem

$$E(u,h) = \frac{1}{2}\sum_{i=1}^{N}\|h_i * u - z_i\|^2 + \frac{\gamma}{2}\sum_{1\le i<j\le N}\|z_i * h_j - z_j * h_i\|^2 + \lambda\int_{\Omega}\phi(|\nabla u(x)|)\mathrm{d}x \tag{14.10}$$

$E(u,h)$, as a function of variables u and h, is not convex due to the convolution in the first term. On the other hand, the energy function is convex with respect to u if h is fixed and it is convex with respect to h if u is fixed. The minimization sequence (u^n, h^n) can be thus built by alternating between two minimization subproblems

Q2

$$u^n = \arg\min_u E(u, h^{n-1}) \quad \text{and} \quad h^n = \arg\min_h E(u^n, h) \tag{14.11}$$

for some initial h^0 with the rectangular support S. The advantage of this scheme lies in its simplicity, since for each subproblem a unique minimum exists that can be easily calculated. However, we cannot guarantee that the global minimum is reached this way, but thorough testing indicates good convergence properties of the algorithm for many real problems.

The solution of the subproblem, Equation 14.10, formally satisfies the Euler–Lagrange equation

$$\frac{\partial E}{\partial u} = \sum_{i=1}^{N} h_i' * (h_i * u - z_i) - \lambda\,\mathrm{div}\left(\frac{\mathrm{d}\phi(|\nabla u|)}{\mathrm{d}|\nabla u|}\frac{\nabla u}{|\nabla u|}\right) = 0 \tag{14.12}$$

where the prime means mirror reflection of the function, that is, $h_i'(x,y) = h_i(-x,-y)$. One can prove (see for example Ref. 28) that a unique solution exists in the BV-space, where the image gradient is a measure. To circumvent the difficulty connected with implementing the measure and with the nonlinearity of the divergence term in Equation 14.12, the solution can be found by relaxing ϕ and following a half-quadratic algorithm originally proposed in Ref. 29 and generalized for convex functions of measures in Ref. 28.

The solution of the subproblem, Equation 14.11, formally satisfies the Euler–Lagrange equations

$$\frac{\partial E}{\partial h_k} = u' * (u * h_k - z_k) - \gamma \sum_{\substack{i=1 \\ i \neq k}}^{N} z_i' * (z_i * h_k - z_k * h_i) = 0, \quad k = 1, \ldots, N \tag{14.3}$$

This is a set of linear equations and thus finding h is a straightforward task.

It is important to note that the algorithm runs in the discrete domain and that a correct estimation of the weighting constants, λ, and γ, and mainly of the blur support S is crucial. In addition, the algorithm is iterative and the energy (Equation 14.9) as a function of the image and blurs does not have one minimum, so the initial guess g^0 plays an important role as well. The positive weighting constants λ and γ are proportional to the noise levels σ and can be calculated in theory from the set of Equation 14.6, Equation 14.12 and Equation 14.13 if the noise variance is known. This is, however, impossible to carry out directly and techniques such as generalized cross validation must be used instead. Such techniques are computationally very expensive and we suggested an alternative approach in Ref. 25 which uses bottom limits of λ and γ. Estimation of the size of the blur support S is even more vexatious. Methods proposed in Refs. 18,20 provide a reliable estimate of the blur size only under ideal noise-free conditions. In the noisy case they suggest a full search, that is, for each discrete rectangular support S estimate the blurs and compare the results.

4. Experiment with Artificial Data

First, we demonstrate the performance of the MBD–AM algorithm on images degraded by computer-generated blurring and noise and we compare the results with two recent methods — Harrikumar's EVAM and Pai's method.

For the evaluation, we use the percentage mean-square error of the fused image $\hat{u}$, defined as

$$\mathrm{PMSE}(u) \equiv 100 \frac{\|\hat{u} - u\|}{\|u\|} \tag{14.14}$$

Although the mean-square error does not always correspond to visual evaluation of the image quality, it has been commonly used for quantitative evaluation and comparison.

A test image of size 250×250 in Figure 14.5(a) was first convolved with four 7×7 PSFs in Figure 14.5(b) and then white Gaussian noise at five different levels (SNR $= 50$, 40, 30, 20, and 10 dB, respectively) was added. This way, we simulated four acquisition channels ($N = 4$) with a variable noise level. The size of the blurs and the noise level were assumed to be known. All three algorithms were therefore started with the correct blur size $S = (7, 7)$. In the case of MBD–AM, λ and γ were estimated as described in Ref. 25 and the starting position h^0 was set to the delta functions. The reconstructed images and blurs are

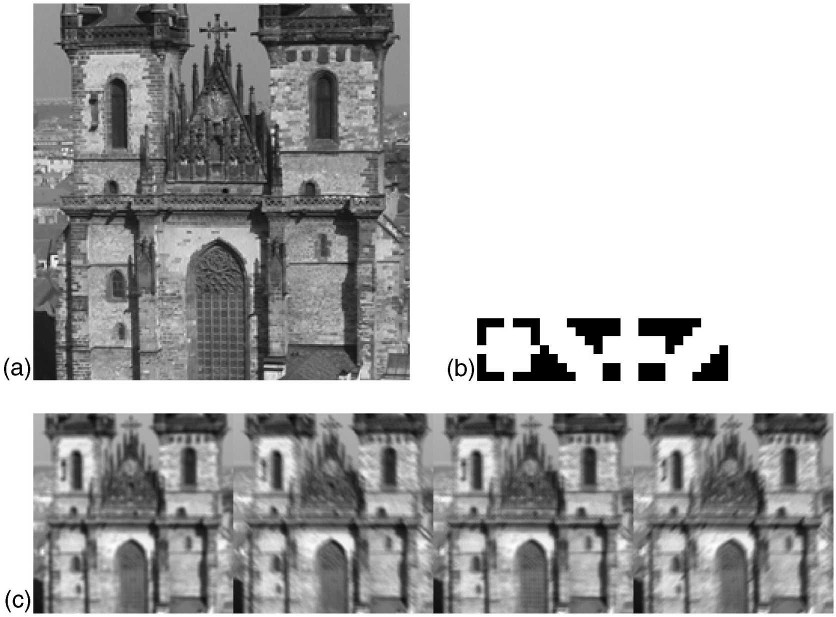


FIGURE 14.5 Synthetic data: (a) original 250×250 image of the Tyn church, (b) four 7×7 PSFs, and (c) four blurred channels. Q3

shown in Figure 14.6 with the percentage mean-square errors summarized in Table 14.1. Note that Pai's method reconstructed only the original image and not the blurs. The performance of the EVAM method quickly decreases as SNR decreases, since noise is not utilized in the derivation of this method. Pai's method shows superior stability but for lower SNR the reconstructed images are still considerably blurred. Contrary to the previous two methods, the MBD–AM algorithm is stable and performs well even for lower SNRs (20 dB, 10 dB). One slight drawback is that the output increasingly resembles a piecewise constant function which is due to the variational regularization $Q(u)$.

5. Experiment with Real Data

The following experiment was conducted to test the applicability of MBD–AM on real data. Four images of a bookcase were acquired with a standard digital camera focused to 80 cm (bookcase in focus), 40, 39, and 38 cm distance, respectively. The acquired data were stored as low resolution 640×480 24 bit color images and only the central rectangular part of the green band of size 250×200 was used for the fusion. The central part of the first image, which captures the scene in focus, is shown in Figure 14.7(a). Three remaining images, Figure 14.7(c), were used as the input for the MBD–AM algorithm. The parameter $\lambda = 1.6 \times 10^{-4}$ was estimated experimentally by running the

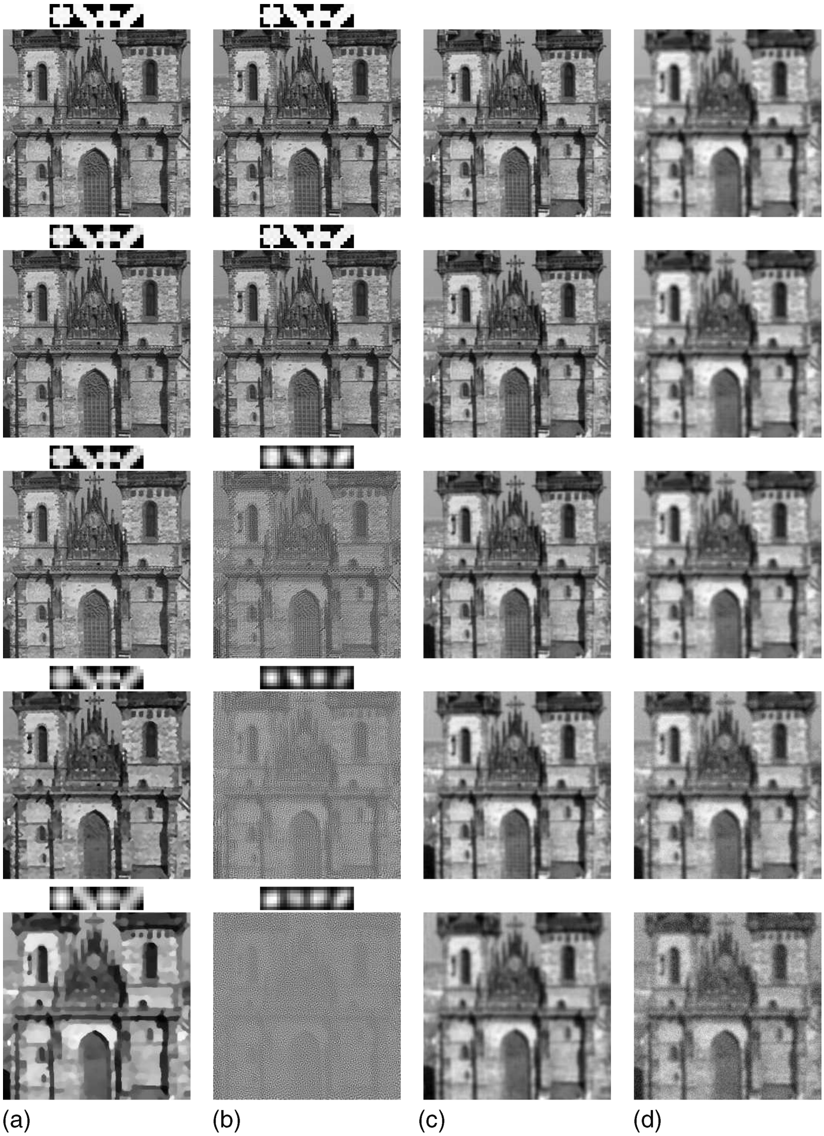


FIGURE 14.6 Reconstruction of the test image and blurs from four degraded images using (a) MBD–AM, (b) EVAM, and (c) Pai's method. The first of the four degraded channels is in column (d) for comparison. From top to bottom SNR = 50, 40, 30, 20, and 10 dB, respectively.

Q3

algorithm with different λ's and selecting the most visually acceptable results. The parameter γ was calculated as described in Ref. 25. A defocused camera causes image degradation approximately modeled by cylindrical blurs. A cepstrum analysis in Ref. 30 was used to estimate diameters of these blurs, which

TABLE 14.1
Performance of the MBD–AM, the EVAM and the Pai's Algorithms on the Data in Figure 14.5

SNR (dB)	MBD–AM	EVAM	Pai
50	0.93	0.99	5.09
40	2.61	2.99	7.87
30	5.17	24.1	10.9
20	10.2	35.7	13.9
10	15.3	38.3	16.4

The table shows percentage mean-square error of the fused image.

were determined to be around eight pixels. Obtained results after ten iterations are shown in Figure 14.7(b). Further iterations did not produce any visual enhancement. Simple visual comparison reveals that the letters printed on shelf backs are more legible in the restored image but still lack the clarity of the focused image, and that the reconstructed blurs resemble the cylindrical blurs as was expected. It is remarkable how successful the restoration was, since one would expect that the similarity of blurs would violate the coprimeness assumption. It is believed that the algorithm would perform even better if a wider disparity between blurs was assured.

V. SLIGHTLY MISREGISTERED BLURRED CHANNELS

This is a generalization of the previous model which allows between-channel shifts (misregistrations) of extent up to a few pixels.

$$z_i(x + a_i, y + b_i) = (u * h_i)(x, y) + n_i(x, y) \tag{14.15}$$

where a_i, b_i are unknown translation parameters of the ith channel. This model is applicable in numerous practical tasks when the scene or the camera moves slightly between consecutive channel acquisitions (see Figure 14.8). Such a situation typically occurs when the camera is subject to vibrations or in multitemporal imaging if the scene is not perfectly still. Sometimes a subpixel between-channel shift is even introduced intentionally in order to enhance spatial resolution of the fused image (this technique is called *superresolution imaging*, see Ref. 31 for a survey and other references).

Images degraded according to this model cannot be fused by the methods mentioned in Section IV.A. If they were applied, the channel misregistrations would lead to strong artifacts in the fused image. On the other hand, the misregistrations considered in this model are too small to be fully removed by image registration techniques (in case of blurred images, registration methods


(a)
(b)
(c)
Thinking in Complexity
2nd Edition
WYNDHAM · THE SEEDS OF TIME
ISAAC ASIMOV
THE NAKED SUN

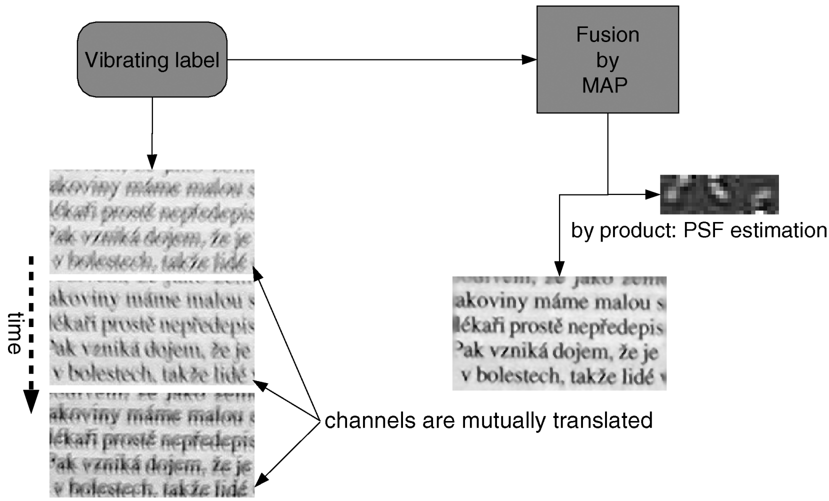


FIGURE 14.8 Multiple acquisition of a vibrating text label. Motion blur of various kind Q3 and small spatial misalignments of the individual frames can be observed. The fused image was achieved by the MAP algorithm described in Section I.

usually can suppress large spatial misalignments but seldom reach subpixel accuracy).

Fusion of images degraded according to this model requires special blind deconvolution methods, which can — in addition to the deconvolution itself — identify and compensate the between-channel misregistration. A successful method based on a stochastic approach is described below.

A. Maximum *a Posteriori* Probability Algorithm

Equation 14.15 can be expressed into equivalent form

$$z_i(x, y) = u(x, y) * h_i(x - a_i, y - b_i) + n_i(x, y) \tag{14.16}$$

which can be further rewritten as

$$z_i(x, y) = (u * g_i)(x, y) + n_i(x, y) \tag{14.17}$$

where g_i is a shifted version of the original PSF $h_i : g_i(x, y) = h_i(x - a_i, y - b_i)$. We can therefore work only with g_i and use the MBD–AM algorithm (Equation 14.10 and Equation 14.11). In this case, the estimate of the blur size has to include

---

Q4 **FIGURE 14.7** Real bookcase images: (a) 250×250 image acquired with the digital camera set to the correct focus distance of 80 cm, (b) MBD–AM fused image along with 10×10 estimated blurs obtained from three images, and (c) image of false focus distances 40 cm, 39 cm and 38 cm, after 10 iterations and $\lambda = 1.6 \times 10^{-4}$.

also the maximum shift between the channels. Since this is difficult to determine, standard MBD techniques including MBD–AM in its present form cannot be applied.

To overcome the difficulties connected with the parameter estimation, we adopt in Ref. 32 a stochastic approach to the minimization problem. The restoration can be formulated then as a maximum *a posteriori* (MAP) estimation. We assume that the matrices u, $g = \{g_1, \ldots, g_N\}$ and $z = \{z_1, \ldots, z_N\}$ are random vector fields with given probability density functions (PDFs) $p(u)$, $p(g)$ and $p(z)$, respectively, and we look for such realizations of u and g which maximize the *a posteriori* probability $p(u, g|z)$. The MAP estimation is equivalent to minimizing $-\log(p(u, g|z))$. The only two assumptions that we must make in addition to those in the energy minimization problem are: u and g are supposed to be statistically independent and n_i is white (that is, uncorrelated) Gaussian noise. Using the Bayes rule, the relation between *a priori* densities $p(u)$, $p(g)$ and the *a posteriori* density is $p(u, g|z) \propto p(z|u, g)p(u)p(g)$. The conditional PDF $p(z|u, g)$ follows from our model, Equation 14.5, and from our assumption of white noise. The blur PDF $p(g)$ can be derived from the regularization $R(g)$, which is also Gaussian noise with a covariance matrix that can be easily calculated. If the image PDF $p(u)$ is chosen in such a way that $-\log(p(u)) \propto Q(u)$ then the MAP estimation is almost identical to the minimization problem, Equation 14.9, for $\lambda = \sigma^2$ and $\gamma = |S|/2$ and we can use the alternating iterative algorithm. To improve stability of the algorithm against the overestimation of S and thus handle inaccurate registration, it suffices to add the constraint of positivity $h(x) > 0$ to Equation 14.7 and perform in Equation 14.11 the minimization subject to the new constraints.

Setting appropriately initial blurs can help our iterative algorithm to converge to the global minimum. This issue is especially critical for the case of overestimated blur size. One can readily see that translated versions of the correct blurs are all equivalent as long as they fit into our estimated blur size. We have seen that the regularization of the blurs R is unable to distinguish between the correct blurs and the correct blurs convolved with an arbitrary spurious factor. This has a negative impact on the convergence mainly if channel misalignment occurs, since new local minima appear for blurs that cope with the misalignment by convolving the correct blurs with an interpolating kernel. To get closer to the correct solution, we thus propose to set the initial blurs g^0 to delta functions positioned at the centers of gravity of blurs $\hat{g} = \arg\min R(g)$. This technique enables us to compensate for the channel shifts right from the start of the algorithm and get away from the incorrect interpolated solutions.

1. Experiment with Misregistered Images

Although the MAP fusion method can also be applied to registered channels, its main advantageous property, that discriminates it from other methods, is the ability to fuse channels which are not accurately registered. This property is illustrated by the following experiment.

The 230×260 test image in Figure 14.9(a) was degraded with two different 5×5 blurs and noise of SNR = 50 dB. One blurred image was shifted by 5×5 pixels and then both images were cropped to the same size; see Figure 14.9(c). The MAP algorithm was initialized with the overestimated blur size 12×12. The fused image and the estimated blur masks are shown in Figure 14.10. Recovered blurs contain negligible spurious factors and are properly shifted to compensate for the misregistration. The fused image is, by visual comparison, much sharper than the input channels and very similar to the original, which demonstrates excellent performance. This conclusion is supported also by the real experiment shown in Figure 14.8, where both blurring and shift were introduced by object vibrations. Unlike the input channels, the text on the fused image is clearly legible.

VI. HEAVILY MISREGISTERED BLURRED CHANNELS

This model is a further generalization of the previous model. Blurring of each channel is still uniform and is modeled by a convolution, but significant misregistrations between the channels are allowed.

$$z_i(\tau_i(x, y)) = (u * h_i)(x, y) + n_i(x, y) \tag{14.18}$$

In this model there are almost no restrictions on the extent and the type of τ_i; it may have a complex nonlinear form (the only constraint is that the individual frames must have sufficient overlap in the region of interest). This is a very realistic model of photographing a flat scene, where the camera moves in three-dimensional space in an arbitrary manner (see Figure 14.11).

Because of the complex nature of τ_i, it cannot be compensated for during the deconvolution step. Thus, fusion of images degraded according to this model is a two-stage process — it consists of image registration (spatial alignment) followed by MBD discussed in the previous section. Since all deconvolution methods require either perfectly aligned channels (which is not realistic) or allow, at most, small shift differences, the registration is a crucial step of the fusion.

Image registration in general is a process of transforming two or more images into a geometrically equivalent form. It eliminates the degradation effects caused by geometric distortion. From a mathematical point of view, it consists of approximating τ_i^{-1} and of resampling the image. For images which are not blurred, the registration has been extensively studied in the recent literature (see Ref. 33 for a survey). However, blurred images require special registration techniques. They can, as well as the general-purpose registration methods, be divided in two groups — global and landmark-based ones. Regardless of the particular technique, all feature extraction methods, similarity measures, and matching algorithms used in the registration process must be insensitive to image blurring.

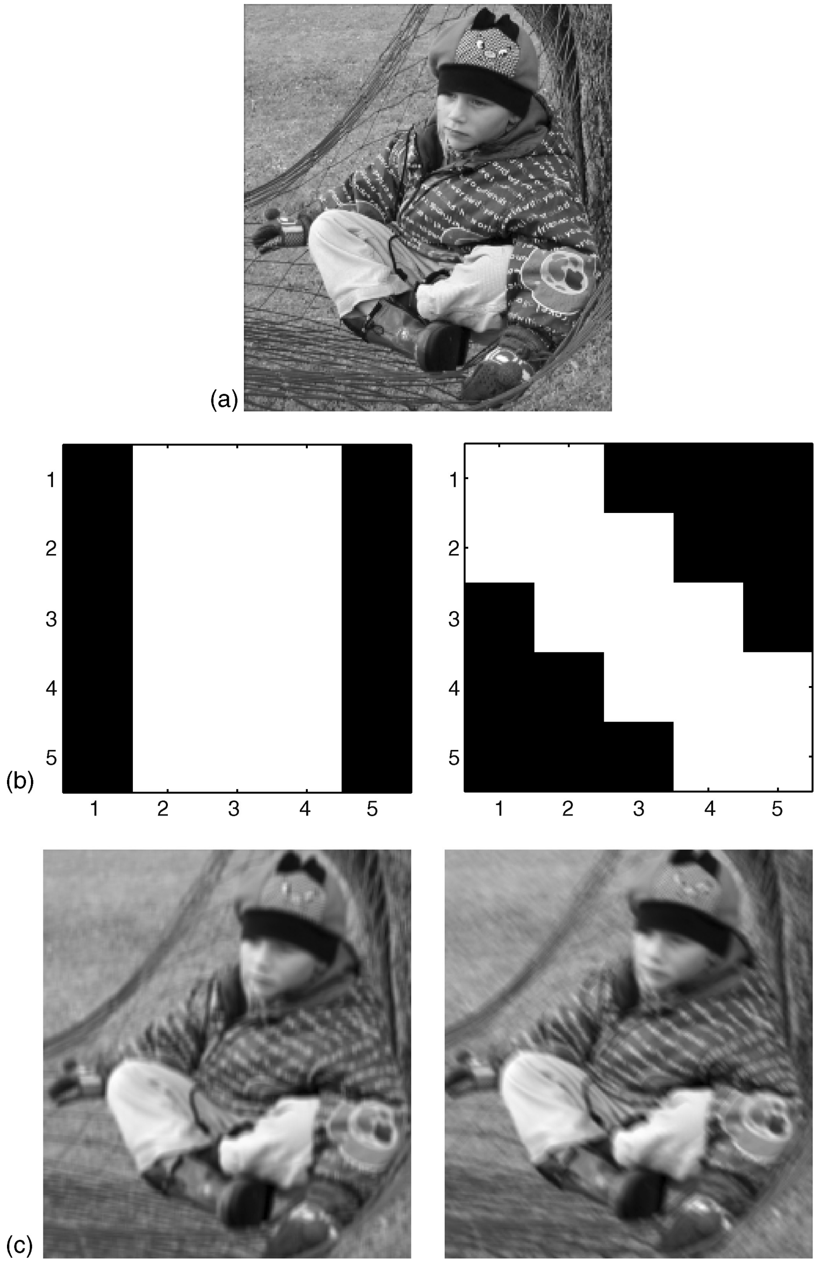
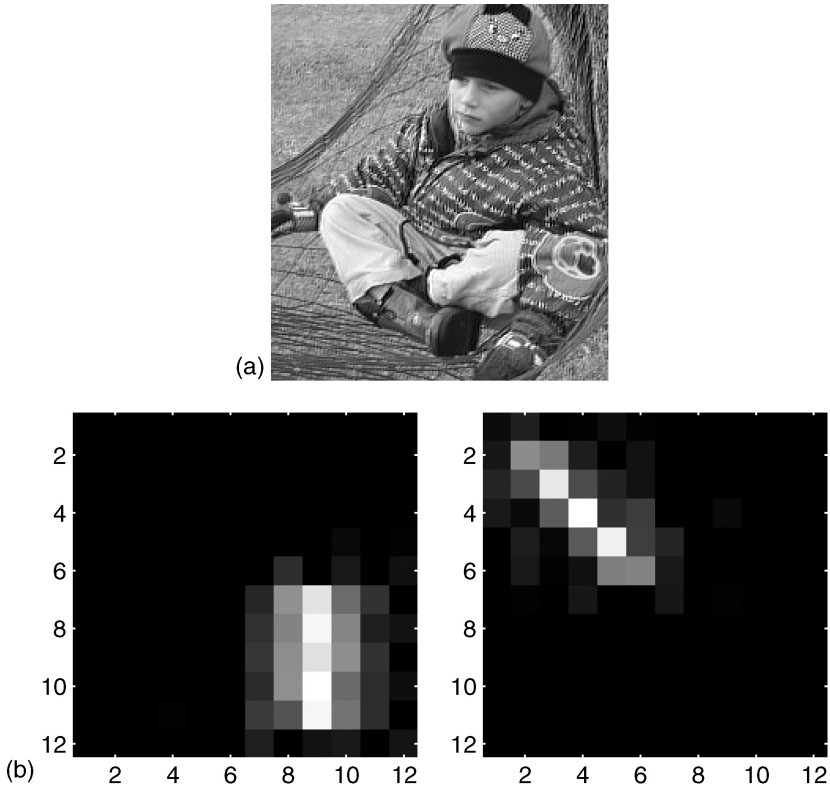


FIGURE 14.9 (a) original test image 230×260 pixels, (b) two 5×5 PSFs, and (c) blurred and shifted images.



Q3 **FIGURE 14.10** MAP image fusion: (a) fused image and (b) estimated blur masks with between-channel shift.

Global methods do not search for particular landmarks in the images. They try to estimate directly the between-channel translation and rotation. Myles and Lobo[34] proposed an iterative method working well if a good initial estimate of the transformation parameters is available. Zhang et al.,[35,36] proposed to estimate the registration parameters by bringing the channels into canonical form. Since blur-invariant moments were used to define the normalization constraints, neither the type nor the level of the blur influences the parameter estimation. Kubota et al.[37] proposed a two-stage registration method based on hierarchical matching, where the amount of blur is considered as another parameter of the search space. Zhang and Blum[38] proposed an iterative multiscale registration based on optical flow estimation in each scale, claiming that optical flow estimation is robust to image blurring. All global methods require considerable (or even complete) spatial overlap of the channels to yield reliable results, which is their major drawback.

Landmark-based blur-invariant registration methods have appeared very recently, just after the first paper on the moment-based blur-invariant features.[39] Originally, these features could only be used for registration of mutually

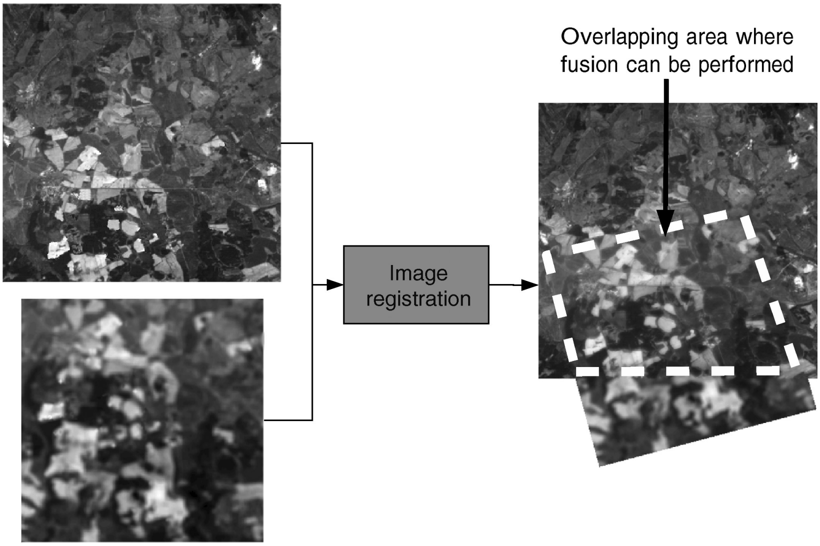


FIGURE 14.11. Two satellite images differing from one another by amount of blur due to different spatial resolution and by shift, rotation, and scaling (left). After the registration (right), the MAP fusion algorithm from the previous section can be applied on the overlapping area. The registration of these images was performed in IMARE Toolbox[50] by means of invariant-based method,[42] courtesy of Barbara Zitová.

shifted images.[40,41] The proposal of their rotational-invariant version[42] in combination with a robust detector of salient points[43] led to registration methods that are able to handle blurred, shifted and rotated images.[44,45]

Although the above-cited registration methods are very sophisticated and can be applied to almost all types of images, the result rarely tends to be perfect. The registration error usually varies from subpixel values to a few pixels, so only fusion methods sufficiently robust to between-channel misregistration can be applied to channel fusion.

VII. CHANNELS WITH SPACE-VARIANT BLURRING

This model comprises space-variant blurring of the channels as well as nonrigid geometric differences between the channels.

$$z_i(\tau_i(x, y)) = \int h_i(x, y, s, t)u(s, t)\mathrm{d}s\,\mathrm{d}t + n_i(x, y) \tag{14.19}$$

The substantial difference from the previous models is that image blurring is no longer uniform in each frame and thus it cannot be modeled as a convolution. Here, the PSF is a function of spatial co-ordinates (x, y) which makes the channel degradation variable depending on the location. This situation typically arises when photographing a three-dimensional scene by a camera with a narrow depth of field. Differently blurred channels are obtained by changing the focus distance of the camera (see Figure 14.12). Unlike piecewise ideal imaging, the depth of the scene can vary in a continuous manner and the existence of at least one "ideal" picture for each location is not guaranteed. Another example is photographing a dynamic scene where different parts move by different velocity and/or in different directions (see Figure 14.13).

Space-variant blurring is not a simple extension of the previous models. It requires qualitatively new approaches and methods. As for the previous model, the image fusion consists of image registration and multichannel blind deblurring but there is a significant difference. While the registration methods can be in principle the same, the techniques used here in the second step must be able to handle space-variant blurring.

Up until now, no papers have been published on multichannel space-variant deblurring. There are, however, a few papers on single-channel space-variant image deblurring, usually originating from space-invariant deconvolution methods. Guo et al.[46] proposed to divide the image into uniformly blurred regions (if possible) and then to apply a modified expectation-maximization algorithm in each region. You and Kaveh[47] considered parameterized PSF and used anisotropic regularization for image deblurring. Cristobal and Navarro[48]

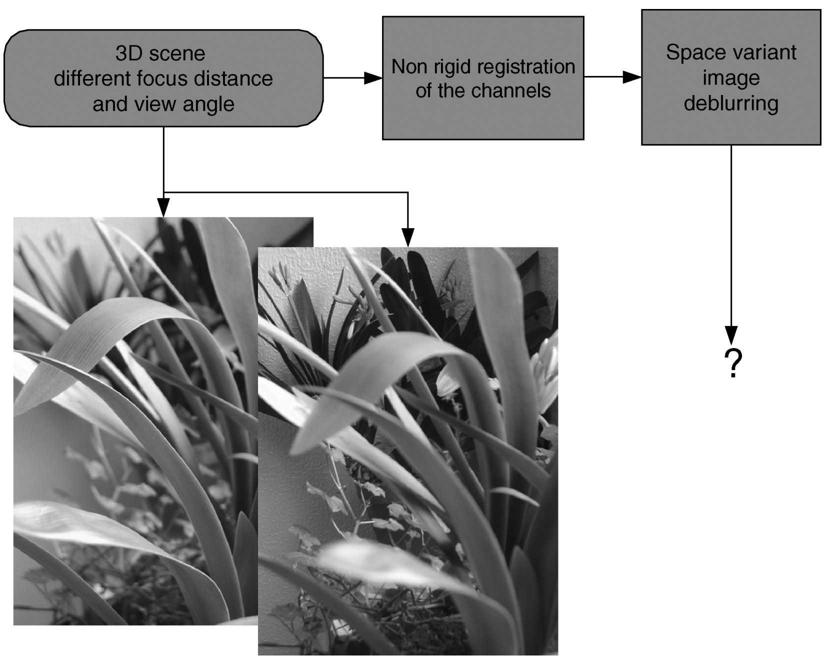


FIGURE 14.12 Space-variant blurring of the channels. Two pictures of a complex three-dimensional scene taken with a variably focused camera. The camera also changed its position and viewing angle between the acquisitions which lead to projective geometric deformation between the channels.

FIGURE 14.13 Space-variant degradation of the channels due to motion blur.

applied multiscale Gabor filters to the restoration. However, the extension of the above methods to the multichannel framework is questionable.

No doubt the solution to this problem is a big challenge for image fusion for the near future. Prospective methods should employ all available *a priori* information, such as a depth map or relief model. They may comprise depth-based, defocus-based or depth- and defocus-based segmentation of the input channels in order to find regions of the same type of blur. Nevertheless, a general solution probably does not exist.

VIII. CONCLUSION

In this chapter, we presented an overview of image fusion methods for the case where the input channels are blurred, noisy, and geometrically different. One has to face this problem in various application areas where the picture of the scene is taken under nonideal conditions. Mathematically, this task is ill posed and cannot be resolved by inverting all degradation factors. The only solution is multiple acquisition of the scene and consequent fusion of all acquired channels. It is believed that if the channel degradations are different, the channels can be fused together in such a way that the information missing in one channel can be supplemented by the others. The fusion approaches and methods differ from each other according to the type of assumed degradations of the channels. Here we classified the possible degradations into five major groups: piecewise ideal imaging, uniform blurring, slight and heavy channel misregistration, and space-variant blurring of the channels. For each category, except the last one, we presented reliable fusion methods whose performance was experimentally verified.

ACKNOWLEDGMENTS

This work has been partially supported by grant No. 102/04/0155 from the Grant Agency of the Czech Republic. Figure 14.4, Figure 14.7, Figure 14.8, and Figure 14.11 are modifications of illustrations which appeared in the authors' previous publications[25,32,49] and are used here with permission of the coauthors.

REFERENCES

1. Chavez, P., Sides, S., and Anderson, J., Comparison of three different methods to merge multiresolution and multispectral data: Landsat TM and SPOT panchromatic, *Photogrammetric Eng. Remote Sensing*, 57, 295–303, 1991.
2. Duport, B., Girel, J., Chassery, J., and Pautou, G., The use of multiresolution analysis and wavelets transform for merging SPOT panchromatic and multispectral image data, *Photogrammetric Eng. Remote Sensing*, 69(9), 1057–1066, 1996.
3. Oldmixon, E., and Carlsson, K., Methods for large data volumes from confocal scanning laser microscopy of lung, *J. Microsc. — Oxford*, 170, 221–228, 1993.
4. Li, H., Deklerck, R., Decuyper, B., Hermanus, A., Nyssen, E., and Cornelis, J., Object recognition in brain CT-scans: knowledge-based fusion of data from multiple feature extractors, *IEEE Trans. Med. Imaging*, 14, 212–229, 1995.
5. Li, S., Kwok, J., and Wang, Y., Combination of images with diverse focuses using the spatial frequency, *Inf. Fusion*, 2(3), 169–176, 2001.
6. Li, S., Kwok, J., and Wang, Y., Multifocus image fusion using artificial neural networks, *Pattern Recognit. Lett.*, 23(8), 985–997, 2002.
7. Li, H., Manjunath, B., and Mitra, S., Multisensor image fusion using the wavelet transform, *Graphical Model Image Process.*, 57(3), 235–245, 1995.
8. Zhang, Z., and Blum, R., A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. In *Proceedings of the IEEE*, Vol. 87, 1999, August.
9. Subbarao, M., Choi, T., and Nikzad, A., Focusing techniques, *J. Opt. Eng.*, 32, 2824–2836, 1993.
10. Subbarao, M., and Tyan, J. K., Selecting the optimal focus measure for autofocusing and depth-from-focus, *IEEE Trans. Pattern Anal. Machine Intell.*, 20, 864–870, 1998.
11. Zhang, Y., Zhang, Y., and Wen, C., A new focus measure method using moments, *Image Vis. Comput.*, 18, 959–965, 2000.
12. Kautsky, J., Flusser, J., Zitová, B., and Šimberová, S., A new wavelet-based measure of image focus, *Pattern Recognit. Lett.*, 23, 1785–1794, 2002.
13. Banham, M., and Katsaggelos, A., Digital image restoration, *IEEE Signal Process. Mag.*, 14(2), 24–41, 1997.
14. Sezan, M. I., and Tekalp, A. M., Survey of recent developments in digital image restoration, *Opt. Eng.*, 29, 393–404, 1990.
15. Kundur, D., and Hatzinakos, D., Blind image deconvolution, *IEEE Signal Process. Mag.*, 13(3), 43–64, 1996.
16. Schulz, T. J., Multiframe blind deconvolution of astronomical images, *J. Opt. Soc. Am. A*, 10, 1064–1073, 1993.

17. Harikumar, G., and Bresler, Y., Efficient algorithms for the blind recovery of images blurred by multiple filters. In *Proceedings IEEE International Conference on Image Processing*, Vol. 3, Lausanne, Switzerland, 1996.
18. Harikumar, G., and Bresler, Y., Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms, *IEEE Trans. Image Process.*, 8(2), 202–219, 1999.
19. Giannakis, G. B., and Heath, R. W., Blind identification of multichannel FIR blurs and perfect image restoration. In *Proceedings 13th International Conference on Pattern Recognition*, 1996.
20. Giannakis, G., and Heath, R., Blind identification of multichannel FIR blurs and perfect image restoration, *IEEE Trans. Image Process.*, 9(11), 1877–1896, 2000.
21. Harikumar, G., and Bresler, Y., Exact image deconvolution from multiple FIR blurs, *IEEE Trans. Image Process.*, 8, 846–862, 1999.
22. Pillai, S., and Liang, B., Blind image deconvolution using a robust GCD approach, *IEEE Trans. Image Process.*, 8(2), 295–301, 1999.
23. Pai, H.-T., and Bovik, A., Exact multichannel blind image restoration, *IEEE Signal Process. Lett.*, 4(8), 217–220, 1997.
24. Pai, H.-T., and Bovik, A., On eigenstructure-based direct multichannel blind image restoration, *IEEE Trans. Image Process.*, 10(10), 1434–1446, 2001.
25. Šroubek, F., and Flusser, J., Multichannel blind iterative image restoration, *IEEE Trans. Image Process.*, 12(9), 1094–1106, 2003.
26. Rudin, L., Osher, S., and Fatemi, E., Nonlinear total variation based noise removal algorithms, *Physica D*, 60, 259–268, 1992.
27. Mumford, D., and Shah, J., Optimal approximation by piecewise smooth functions and associated variational problems, *Comm. Pure Appl. Math.*, 42, 577–685, 1989.
28. Aubert, G., and Kornprobst, P., *Mathematical Problems in Image Processing*, Springer, New York, 2002.
29. Chambolle, A., and Lions, P., Image recovery via total variation minimization and related problems, *Numer. Math.*, 76(2), 167–188, 1997.
30. Chang, M., Tekalp, A., and Erdem, A., Blur identification using the bispectrum, *IEEE Trans. Signal Process.*, 39(10), 2323–2325, 1991.
31. Park, S., Park, M., and Kang, M., Super-resolution image reconstruction: a technical overview, *IEEE Signal Proc. Mag.*, 20(3), 21–36, 2003.
32. Šroubek, F., and Flusser, J., Shift-invariant multichannel blind restoration. In *Proceedings of the third International Symposium on Image and Signal Processing and Analysis, ISPA'03*, Rome, 2003, September.
33. Zitová, B., and Flusser, J., Image registration methods: a survey, *Image Vis. Comput.*, 21, 977–1000, 2003.
34. Myles, Z., and Lobo, N. V., Recovering affine motion and defocus blur simultaneously, *IEEE Trans. Pattern Anal. Machine Intell.*, 20, 652–658, 1998.
35. Zhang, Y., Wen, C., and Zhang, Y., Estimation of motion parameters from blurred images, *Pattern Recognit. Lett.*, 21, 425–433, 2000.
36. Zhang, Y., Wen, C., Zhang, Y., and Soh, Y. C., Determination of blur and affine combined invariants by normalization, *Pattern Recognit.*, 35, 211–221, 2002.
37. Kubota, A. Kodama, K., and Aizawa, K., Registration and blur estimation methods for multiple differently focused images. In *Proceedings International Conference on Image Processing*, Vol. II, 1999.

38. Zhang, Z., and Blum, R., A hybrid image registration technique for a digital camera image fusion application, *Inf. Fusion*, 2, 135–149, 2001.
39. Flusser, J., Suk, T., and Saic, S., Recognition of blurred images by the method of moments, *IEEE Trans. Image Process.*, 5, 533–538, 1996.
40. Flusser, J., and Suk, T., Degraded image analysis: an invariant approach, *IEEE Trans. Pattern Anal. Machine Intell.*, 20(6), 590–603, 1998.
41. Bentoutou, Y., Taleb, N., Mezouar, M., Taleb, M., and Jetto, L., An invariant approach for image registration in digital subtraction angiography, *Pattern Recognit.*, 35, 2853–2865, 2002.
42. Flusser, J., and Zitová, B., Combined invariants to linear filtering and rotation, *Int. J. Pattern Recognit. Art. Intell.*, 13(8), 1123–1136, 1999.
43. Zitová, B., Kautsky, J., Peters, G., and Flusser, J., Robust detection of significant points in multiframe images, *Pattern Recognit. Lett.*, 20, 199–206, 1999.
44. Flusser, J., Zitová, B., and Suk, T., Invariant-based registration of rotated and blurred images. In *Proceedings IEEE 1999 International Geoscience and Remote Sensing Symposium*. Los Alamitos, 1999, June.
45. Flusser, J., Boldyš, J., and Zitová, B., Moment forms invariant to rotation and blur in arbitrary number of dimensions, *IEEE Trans. Pattern Anal. Machine Intell.*, 25(2), 234–246, 2003.
46. Guo, Y., Lee, H., and Teo, C., Blind restoration of images degraded by space-variant blurs using iterative algorithms for both blur identification and image restoration, *Image Vis. Comput.*, 15, 399–410, 1997.
47. You, Y.-L., and Kaveh, M., Blind image restoration by anisotropic regularization, *IEEE Trans. Image Process.*, 8(3), 396–407, 1999.
48. Cristobal, G., and Navarro, R., Blind and adaptive image restoration in the framework of a multiscale Gabor representation. In *Proceedings IEEE Time-Frequency Time-scale analysis*, 1994.
49. Flusser, J., Boldyš, J., and Zitová, B., Invariants to convolution in arbitrary dimensions, *J. Math. Imaging Vis.*, 13, 101–113, 2000.
50. Zitová, B., and Flusser, J. IMARE — image registration toolbox for MATLAB. Freeware, http://www.utia.cas.cz/user_data/zitova/IMARE.htm.

Author Queries

JOB NUMBER: 9295

CHAPTER: Fusion of Blurred Images

Q1 "Constraint (7) was dropped…" I have assumed (7) refers to Equation (7) and changed to Equation 14.7 to conform to CRC style. Please confirm this is correct. If so, should the wording of this phrase be changed to "The constraint of Equation 14.7 was dropped…"?

Q2 Since there are two equations with the same number so the equations are reordered. Kindly check.

Q3 Please check the figure quality.

Q4 Please check whether the permission line is needed or not for this figure. If so, please provide.