8

Multidimensional fusion by image mosaics

Yoav Y. Schechner [a] and Shree K. Nayar [b]

[a] Department of Electrical Engineering, Technion – Israel Institute of Technology, Haifa, Israel

[b] Department of Computer Science, Columbia University, New York, USA

Image mosaicing creates a wide field of view image of a scene by fusing data from narrow field images. As a camera moves, each scene point is typically sensed multiple times during frame acquisition. Here we describe *generalised mosaicing*, which is an approach that enhances this process. An optical component with spatially varying properties is rigidly attached to the camera. This way, the multiple measurements corresponding to any scene point are made under different optical settings. Fusing the data captured by the multiple frames yields an image mosaic that includes additional information about the scene. This information can come in the form of extended dynamic range, high spectral quality, polarisation sensitivity or extended depth of field (focus). For instance, suppose the state of best focus in the camera is spatially varying. This can be achieved by placing a transparent dielectric on the detector array. As the camera rigidly moves to enlarge the field of view, it senses each scene point multiple times, each time in a different focus setting. This yields a wide depth of field, wide field of view image, and a rough depth map of the scene.

8.1 Introduction

Image mosaicing[1] is a common method to obtain a wide field of view (FOV) image of a scene [8–10]. The basic idea is to capture frames of different parts of a scene, and then fuse the data from these frames, to obtain a larger image. The data is acquired by a relative motion between the camera and the scene: this way, each frame captures different

[1] In different communities the terms *mosaicing* [1,2] and *mosaicking* [3–7] are used.

scene parts. Image mosaicing has long been used in a variety of fields, such as optical observational astronomy [11,12], radio astronomy [13], and remote sensing [3,7,14–18], optically or by synthetic aperture radar (SAR). It is also used in underwater research [5, 6,19–21]. Moreover, image mosaicing has found applications in consumer photography [1,4,9,22–29].

As depicted in Figure 8.1, image mosaicing mainly addressed the extension of the FOV. However, there are other imaging dimensions that require enhanced information by fusing multiple measurements. In the following, we show how this can be done, within a unified framework that includes mosaicing. The framework is termed *generalised mosaicing*. It extracts significantly more information about the scene, given an amount of acquired data similar to that acquired in traditional mosaicing.

A typical video sequence acquired during mosaicing has great redundancy in terms of the data it contains. The reason is that there is typically a significant overlap between frames acquired for the mosaic, thus each point is observed multiple times. Now, let us rigidly attach to the camera a fixed filter with spatially varying properties, as in the setup shown in Figure 8.2. As the camera moves (or simply rotates), each scene point is measured under different optical settings. This significantly reduces the redundancy in the captured video stream. In return, the filtering embeds in the acquired data more information about each point in the mosaic FOV. Except for mounting the fixed filter, the image acquisition in generalised mosaicing is identical to traditional mosaicing.

In the following sections we describe several realisations of this principle. Specifically, when a filter with spatially varying transmittance is attached to the camera, each scene point is effectively measured with different exposures as the camera moves. These measurements are then fused to a high dynamic range (HDR) mosaic. Similarly, if the filter transmits a spatially varying spectral band, multispectral information is obtained for each scene point. In another implementation, a spatially varying polarisation filter is used, yielding wide FOV polarimetric imaging. Such systems were described in [30–34].

A particular realisation, which we describe in more detail is one having spatially varying focus settings. Fusion of image data acquired by such a sensor can yield an all-focused image in a wide FOV, as well as a rough depth map of the scene.

8.2 Panoramic focus

8.2.1 Background on focus

Focusing is required in most cameras. Let the camera view an object at a distance s_{object} from the first (front) principal plane of the lens. A focused image of this object is formed at a distance s_{image} behind the second (back) principal plane of the lens, as illustrated in Figure 8.3. For simplicity, consider first an aberration-free flat-field camera, having an effective focal length f. Then,

$$\frac{1}{s_{\text{image}}} = \frac{1}{f} - \frac{1}{s_{\text{object}}} \tag{8.1}$$

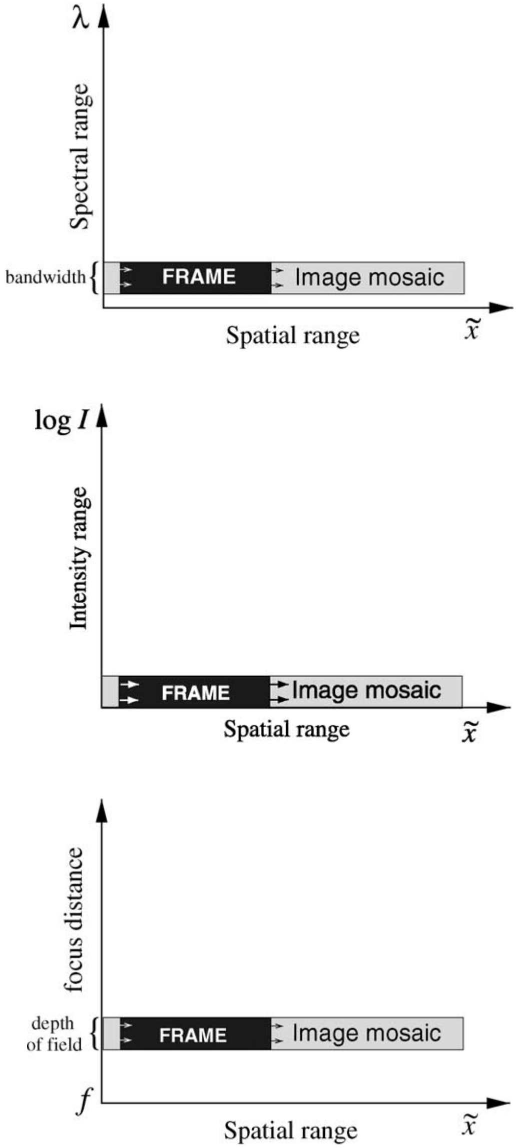


Figure 8.1 *An image frame has a limited FOV of the scene, i.e., it has a limited extent spatially ($\tilde{x}$ coordinates). By fusing partly overlapping frames, an image mosaic extends the FOV of any camera. However, there is a need to enhance additional imaging dimensions, such as the dynamic range [33] of intensity I, the hyper-spectral [32] quality (sensitivity to the wavelength λ), and depth of field. The latter refers to the need to view objects in-focus at a distance extending from f (front focal plane) to infinity.*

Hence, s_{image} is equivalent to s_{object}. The image is sensed by a detector array (e.g., a CCD), situated at a distance s_{detector} from the back principal plane. If $s_{\text{detector}} = s_{\text{image}}$, then the detector array senses the focused image. Generally, however, $s_{\text{detector}} \neq s_{\text{image}}$. If $|s_{\text{image}} - s_{\text{detector}}|$ is sufficiently large, then the detector senses an image which is *defocus blurred*.

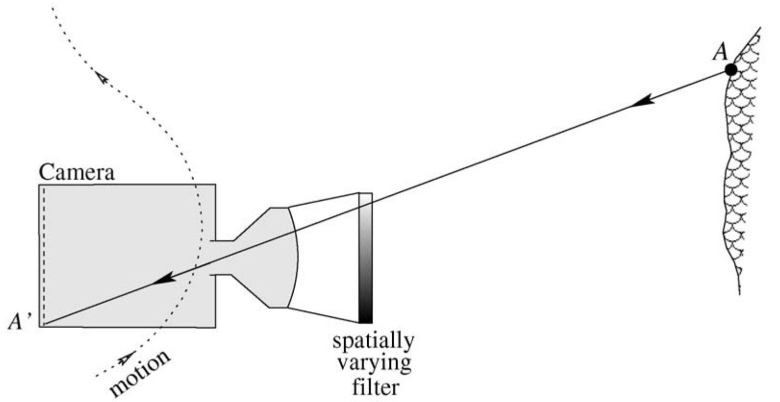


Figure 8.2 *Scene point A is imaged on the detector at A' through a spatially varying filter attached to the camera. As the imaging system moves [31], each scene point is sensed through different portions of the filter, thus multiple measurements are obtained under different optical settings.*

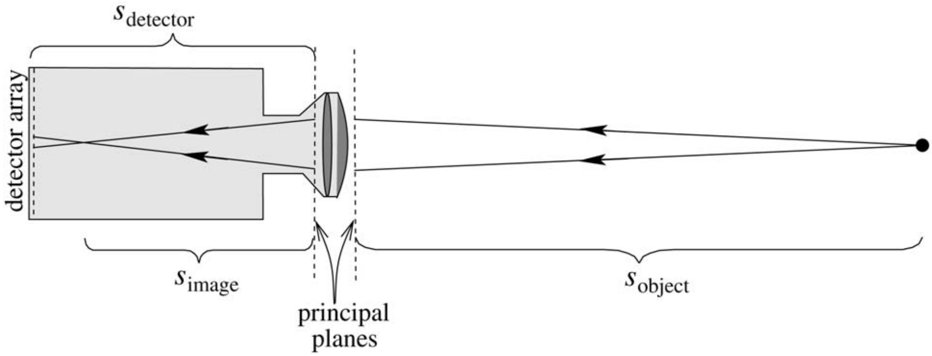


Figure 8.3 *Geometry of a simple camera system.*

For a given s_{detector}, there is a small range of s_{image} values for which the defocus blur is insignificant. This range is the *depth of focus*. Due to the equivalence of s_{image} to s_{object}, this corresponds to a range of object distances which are imaged sharply on the detector array. This range is the *depth of field* (DOF). Hence, a single frame can generally capture in focus objects that are in this limited span. However, typically, different objects or points in the FOV have different distances s_{object}, extending beyond the DOF. Hence, while some objects in a frame are in focus, others are defocus blurred.

There is a common method to capture each object point in focus, using a stationary camera. In this method, the FOV is fixed, while K frames of the scene are acquired. In each frame, indexed $k \in [1, K]$, the *focus settings* of the system change relative to the previous frame. Change of the settings can be achieved by varying s_{detector}, or f, or s_{object}, or any combination of them. This way, for any specific object point (x, y), there is a frame

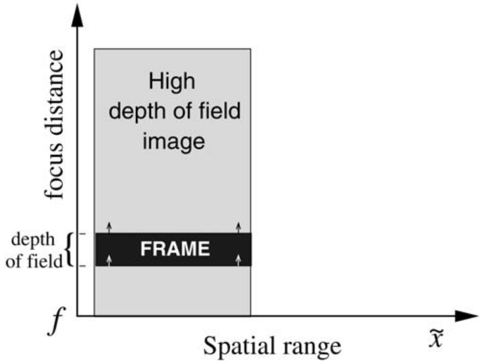


Figure 8.4 *An image frame has a limited FOV of the scene (marked by $\tilde{x}$) and a limited DOF. By fusing differently focused images, the DOF can be extended by image post processing, but the FOV remains limited.*

$k(x, y)$ for which Equation (8.1) is approximated as

$$\frac{1}{s_{\text{detector}}^{(k)}} \approx \frac{1}{f^{(k)}} - \frac{1}{s_{\text{object}}^{(k)}(x, y)} \tag{8.2}$$

i.e. $s_{\text{detector}}^{(k)} \approx s_{\text{image}}^{(k)}$, bringing the image of this object point into focus. This is the *focusing* process. Since each point (x, y) is acquired in focus at some frame k, then fusing the information from all K frames yields an image in which all points appear in focus. This principle is sketched in Figure 8.4. The result of this image fusion is effectively a high DOF image. However, the FOV remains limited, since the camera is static while the frames are acquired. In the subsequent sections, we will show that *focusing* and *extension of the FOV* can be obtained in a single, efficient scan.

The surface of least confusion

The object distance $s_{\text{object}}(x, y)$ is a function of the transversal coordinates. Following Equation (8.1), this function is equivalent to a surface $s_{\text{image}}(x, y)$ inside the camera chamber.[2] On this surface, the image is at best focus (least blur). This is the *surface of least confusion* (SLC) [35]. Apparently, for a flat object having a spatially invariant distance, the SLC is flat as well. In such a case, the entire object can be focused in a single frame, since the detector array is flat.

However, the SLC is generally not flat, even if s_{object} is constant. Typically, it is curved [35] radially from the centre of the camera FOV. Since it no longer obeys Equation (8.1), we denote the SLC as $s_{\text{image}}^{\text{effective}}$. This effect is caused by lens aberrations, which have been considered as a hindering effect. Thus, optical engineering makes an effort to

[2]The transversal coordinates of the image (x, y) are a scaled version of the object coordinates. The scale is the magnification of the camera. Since the magnification is fixed for given camera settings, we do not make explicit use of this magnification. Thus for simplicity, we do not scale the coordinates, and thus (x, y) are used for both the image domain and the object.

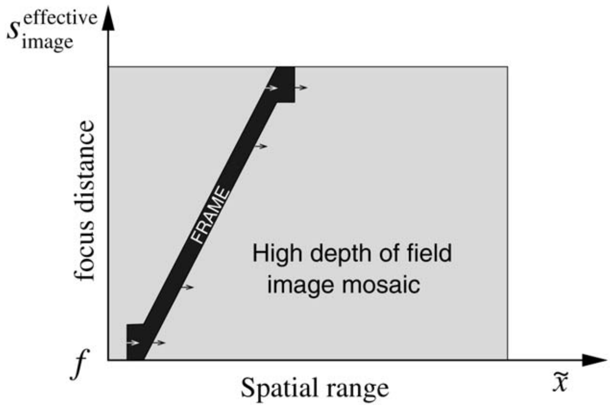


Figure 8.5 *Each frame has a spatially varying focus. Generalised mosaicing of such frames extends both the FOV and the DOF.*

flat-field optical systems [35], i.e. to minimise the departure of the SLC from a flat surface normal to the optical axis.

8.2.2 Intentional aberration

We now describe a unified way for expanding both the FOV and the DOF. The key is an intentionally aberrated imaging system, in which the distance between the SLC and the flat detector array spatially varies significantly. At a given frame, thus, different object points on a flat frontal plane are focused or defocused differently. This concept is depicted in Figure 8.5, by the support of the frame: across the FOV of a frame, the focus distance changes.

Now, the camera scans the scene *transversally*, in order to increase the FOV using mosaicing, as in Figure 8.1. However, due to the spatially varying focus of this system, during the transversal scan any object becomes focused at *some* frame k, as seen in Figure 8.5. By use of computational analysis of the acquired images, information about focus is extracted for each object.

A focused state may be obtained for all the pixels in a wide FOV image mosaic. In addition, information becomes available about the periphery of the central region of interest. The periphery has a gradually narrower DOF, but at least as wide as the inherent DOF of the camera (Figure 8.6). Such a gradual variation is analogous to foveated imaging systems, in which the acquisition quality improves from the periphery towards the centre of the FOV. The periphery is at most one frame wide, and is eliminated in 360° panoramic mosaics.

Optical implementation

Panoramic focusing based on this principle had been obtained by a system in which the CCD array was tilted relative to the optical axis [36]. We now describe an alternative

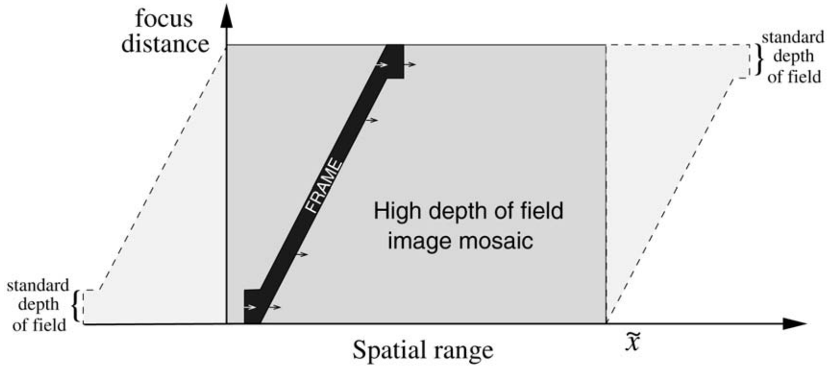


Figure 8.6 *Outside the main region of interest, the mosaic provides additional information about the scene periphery, whose quality gradually coincides with that of a single frame.*

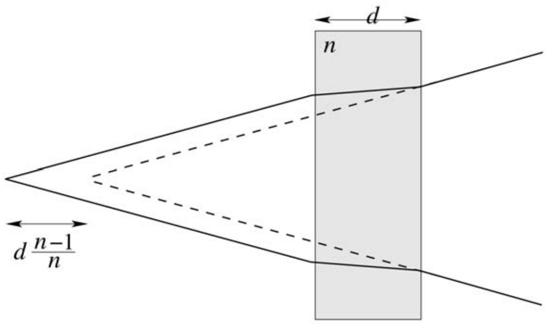


Figure 8.7 *A dielectric with refraction index n and thickness d. If it is placed before the detector, it shifts the focused image point away from the imaging optics.*

implementation, that can be more flexible. It can be based on mounting a transparent dielectric object on the detector array. By letting this transparent object have spatially varying characteristics, the SLC can be deformed to suit our needs.

First, we describe a simpler setup. Suppose a thin transparent slab is inserted between the imaging optics and the plane of best image focus. Now, let a light beam be focused through the slab. As seen in Figure 8.7, light refracts at the slab interfaces. Hence, the plane of best focus is shifted. The shift effectively increases s_{image} by approximately

$$\Delta s \approx d(n-1)/n \tag{8.3}$$

where d is the thickness of the slab and n is its refractive index relative to air.[3] The SLC is then

$$s_{\text{image}}^{\text{effective}} = s_{\text{image}} + \Delta s \tag{8.4}$$

[3]This is based on the paraxial approximation.

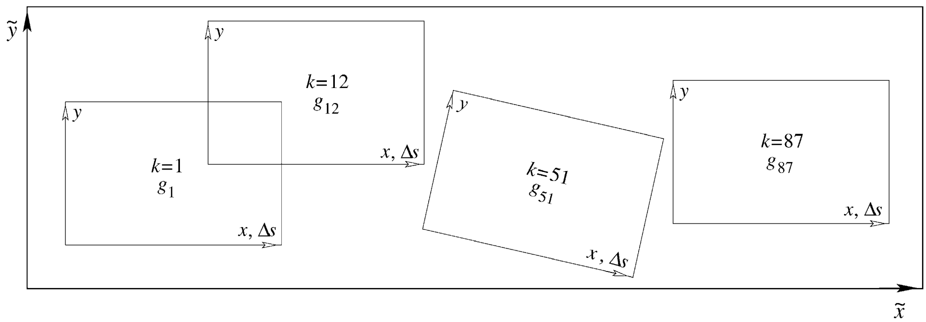


Figure 8.8 *The coordinates $(\tilde{x}, \tilde{y})$ denote the position of a scene point in the global coordinate system of the mosaic [32]. The mosaic is composed of raw frames indexed by k. The intensity in a raw frame is $g_k(x, y)$, where (x, y) are the internal coordinates in the frame. The axial shift Δs of the SLC is a function of (x, y).*

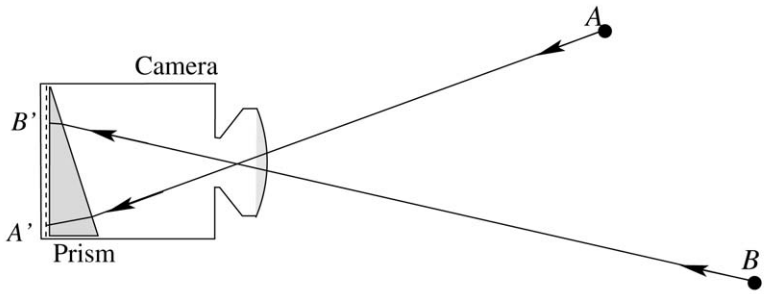


Figure 8.9 *The focus settings can change across the frame's FOV by placing a transparent object with variable thickness on the detector. Here, a wedge-prism is placed on the detector array, such that both objects A and B are focused even-though they are at different distances from the camera.*

where s_{image} is given by Equation (8.1). The distance between the focused image and the detector is thus $|s_{\text{image}}^{\text{effective}} - s_{\text{detector}}|$. It can be affected by setting d or n.

To obtain a spatially varying SLC, we may vary n or d across the camera FOV. Let us better define the spatial coordinates we use. The coordinates $(\tilde{x}, \tilde{y})$ denote a scene point, in the global coordinate system of the mosaic, as depicted in Figure 8.8. In each frame, the internal coordinates of a pixel are (x, y), equivalent to a position on the detector array. As an example for spatial variation of d, let the transparent object be a wedge prism, as illustrated in Figure 8.9. In the paraxial approximation (small angles), d and Δs change linearly along the detector's x-axis.

$$d = \gamma x \tag{8.5}$$

where γ encapsulates the wedge-prism slope. Following Equations (8.3) and (8.5),

$$\Delta s = \gamma x(n-1)/n \tag{8.6}$$

The defocus blur changes accordingly.

Figure 8.10 *The frame in which an object is focused depends on its position in the frame. The teddy bear is focused only when it appears on the right-hand side of the frame's FOV. The words 'don't panic' are focused only when they appear in the left-hand side of the frame's FOV.*

To demonstrate this, we placed a wedge-prism on the CCD array of a Sony monochrome machine vision camera, and then mounted a C-mount lens over it. The system was then positioned on a tripod, and panned to obtain video data for a wide FOV mosaic. Four sample raw frames extracted from the video sequence are shown in Figure 8.10. Consider two objects in the scene: a teddy bear and the words 'don't panic' written on a newspaper. The two objects have a significantly different distance from the camera. The teddy bear is defocus blurred when it appears on the left-hand side of an image frame, while focused on the right-hand side of another frame. In contrast, the words 'don't panic' focus at the opposite frame part.

A wedge-prism is a special case. More generally, the transparent object placed on the detector array can have other shapes, leading to more general SLCs. Actually, even in the absence of any transparent object attachment, the focal distance is somewhat spatially varying (curved radially) by default, unless an effort is made to flat field the system [35]. Thus, one can use simpler lenses in which no effort is made to minimise the Petzval curvature [35]. In any case, the raw images need to have spatially varying focusing characteristics. While this makes the raw images look strange, the spatially varying focus is compensated for and even exploited using fusion algorithms, as described in the next sections.

8.2.3 Data fusion

Thanks to algorithms, a wide FOV and a wide DOF can be obtained from the raw images. The algorithms combine the principles of image mosaicing (registration; image fusion that reduces artifacts) with the principles of wide DOF imaging (focus search and image fusion). Therefore, mosaicing includes three consecutive stages: image registration, focus sensing for the registered images, and image fusion that reduces the appearance of artifacts.

First of all, the frames should be registered, assuming the motion between frames is unknown. A scene point has different coordinates in each frame. The measurements corresponding to this point should be identified before they can be fused. Registration algorithms for image mosaicing are well developed, and we refer the readers to [1,4,9,27,29, 32,33] for details. We registered the images, samples of which are shown in Figure 8.10, by minimising the mean squared difference between the frames.

Recall that $(\tilde{x}, \tilde{y})$ are the coordinates of a scene point in the global coordinate system of the mosaic (see Figure 8.8). Let k be the index of the individual frames that compose the mosaic. Since the raw frames are registered, we have for each scene point in the mosaic FOV a set of intensity measurements $\{g_k(\tilde{x}, \tilde{y})\}_{k=1}^{K}$. Now we may analyse the focus or defocus blur.

8.2.3.1 The best focus

Once the frames are registered, we may find which one of them is focused at each mosaic coordinate $(\tilde{x}, \tilde{y})$. The focused state of an object point is detected by maximising some focus criterion [37–45]. Since focus is associated with high sharpness, we maximise a sharpness criterion, say the image Laplacian

$$\hat{k}(\tilde{x}, \tilde{y}) = \arg\max_k \left|\nabla^2_{\tilde{x},\tilde{y}}\, g_k(\tilde{x}, \tilde{y})\right| \tag{8.7}$$

where $\nabla^2_{\tilde{x},\tilde{y}}$ is the Laplacian over the mosaic's spatial domain.

Based on $\hat{k}$, a depth map of the scene can be estimated, as in standard methods for *depth from focus* [38–41,43,46–49]. As illustrated in Figure 8.8, once we know $\hat{k}$ at $(\tilde{x}, \tilde{y})$, we can retrieve the corresponding (x, y) coordinates in the $\hat{k}$th frame. Denote them as $(x_{\hat{k}}, y_{\hat{k}})$. The SLC $s^{\text{effective}}_{\text{image}}(x, y)$ of the imaging system can be known, e.g., by pre-calibration of the system. Hence the retrieved position $(x_{\hat{k}}, y_{\hat{k}})$ immediately indicates $s_{\text{image}}(x_{\hat{k}}, y_{\hat{k}})$. As explained in Section 8.2.1, s_{image} is equivalent to the object distance, i.e., the depth map.

For example, consider again our experiment using the wedge-prism. At focus, $s^{\text{effective}}_{\text{image}} \approx s_{\text{detector}}$. Thus, following Equations (8.4) and (8.6),

$$s_{\text{image}} = s_{\text{detector}} - \gamma x_{\hat{k}} \frac{n-1}{n} \tag{8.8}$$

Figure 8.11 *A map equivalent to a rough object distance, estimated in an experiment. Brighter pixels correspond to objects at larger object distances.*

This was applied to our image data, samples of which were shown in Figure 8.10. The resulting depth-equivalent map s_{image} is shown in Figure 8.11, where brighter pixels correspond to objects at larger object distances. This demonstrates the feasibility of obtaining rough depth information. The depth map was smoothed by median filtering. As mentioned in Section 8.2.2, the peripheral scene parts are measured in fewer frames, thus with less focus states. Therefore, the dynamic range of depths gradually decreases in the mosaic periphery, leaving partial knowledge of depth from focus.[4]

8.2.3.2 Wide DOF mosaic by fusion

In Section 8.2.3.1, Equation (8.7) led to a rough estimate of the object depth. In this section, Equation (8.7) is the basis for forming a sharp image of the objects in view (as if they are all in focus), no matter their depth. The output image is $\hat{I}$. Let us assign an output pixel $(\tilde{x}, \tilde{y})$ the value

$$\hat{I}(\tilde{x}, \tilde{y}) = g_{\hat{k}(\tilde{x}, \tilde{y})}(\tilde{x}, \tilde{y}) \tag{8.9}$$

Here, each object point is extracted at its sharpest state. This principle is common in techniques that create high DOF images from a stack of differently focused frames, in a stationary camera [37,44], as described in Section 8.2.1.

While Equations (8.7) and (8.9) are a simple recipe, the result may not be visually pleasing. Two kinds of artifacts are created. One of them is a noisy appearance, stemming from a noisy estimation of the state of best-focus. The other artifact is seam-lines in the mosaic, corresponding to the boundaries of raw frames. It stems from the slight inconsistencies of the exposure and illumination settings between frames. Both artifacts are inherent to image fusion and image mosaicing, and have been observed consistently in past work about either mosaicing or focusing. In the following, we detail the reasons for these two visual artifacts, and describe a way to overcome both. This is not just a cosmetic technicality: overcoming visual artifacts in elegant ways had been a topic for research in the mosaicing and fusion communities.

---

[4] Even if no frame measures the focus state at an area, as occurs in the periphery, depth may still be estimated there. A method that enables this is *depth from defocus* [50–53], which requires as few as two differently defocused measurements in order to estimate depth.

Focus artifacts Focus measurement as in Equation (8.7) is not an ideal process. Image noise may cause a random shift in the maximum focus measure. In other words, there is randomness in the value of $\hat{k}(\tilde{x}, \tilde{y})$ about the value that would have been obtained in the absence of noise. This results in a noisy looking mosaic. This randomness can be attenuated if the focus measure (e.g., Equation (8.7)) is calculated over a wide patch. This would have been fine if the object was equidistant from the camera. However, $s_{\text{object}}(\tilde{x}, \tilde{y})$ may have significant spatial variations (depth edges). In this case, large patches blur the estimated $\hat{k}$ map, degrading the resulting fused image.

This artifact is bypassed by an elegant approach which yields fusion results that are more perceptually pleasing, as has been shown in various fusion studies [54–57]. It is based on calculations performed in *multiple scales* [54–57], as described next. For each frame $g_k(\tilde{x}, \tilde{y})$, derive Laplacian and Gaussian pyramids [58] having M pyramid levels. Denote the image at level $m \in [0, M-1]$ of the Laplacian pyramid as $\mathbf{l}_k^{(m)}$, with a corresponding Gaussian pyramid image $\mathbf{g}_k^{(m)}$. As m increases, the image $\mathbf{l}_k^{(m)}$ expresses lower spatial frequencies. Moreover, a pixel of $\mathbf{l}_k^{(m)}$, denoted as $l_k^{(m)}(\tilde{x}^{(m)}, \tilde{y}^{(m)})$ represents an equivalent image area in the full image domain $(\tilde{x}, \tilde{y})$ that increases with m. The lowest spatial frequencies are represented by another image in the Gaussian pyramid, denoted as $\mathbf{g}_k^{(M)}$. Based on $\mathbf{g}_k^{(M)}$ and $\{\mathbf{l}_k^{(m)}\}_{m=0}^{M-1}$, the raw image $g_k(\tilde{x}, \tilde{y})$ can be reconstructed [58].

Now, instead of Equation (8.7), the relation

$$\hat{k}\left(\tilde{x}^{(m)}, \tilde{y}^{(m)}\right) = \arg\max_k \left| l_k^{(m)}\left(\tilde{x}^{(m)}, \tilde{y}^{(m)}\right)\right| \tag{8.10}$$

determines which is the sharpest frame $\hat{k}$, in each level m of the pyramid, and in each pixel $(\tilde{x}^{(m)}, \tilde{y}^{(m)})$. Define an image

$$l_{\text{best}}^{(m)}\left(\tilde{x}^{(m)}, \tilde{y}^{(m)}\right) = l_{\hat{k}}^{(m)}\left(\tilde{x}^{(m)}, \tilde{y}^{(m)}\right) \tag{8.11}$$

where $\hat{k}$ is given by Equation (8.10). The image $\mathbf{l}_{\text{best}}^{(m)}$ defined in Equation (8.11) is analogous to Equation (8.9), but it fuses the images in each scale of the Laplacian pyramid, as in [55,56]. Effectively, Equation (8.11) can be interpreted as using large areas of the full image domain $(\tilde{x}, \tilde{y})$ to fuse rough image components (low frequencies), while using small effective areas to fuse small features (high frequencies).

The images $\{\mathbf{g}_k^{(M)}\}_{k=1}^{K}$ represent the lowest spatial frequencies. These components are least affected by defocus blur, as defocus is essentially a low-pass filter. Therefore, there is not much use in trying to select the sharpest $\mathbf{g}_k^{(M)}$ among all K frames. Hence, this component is fused by a linear superposition of $\{\mathbf{g}_k^{(M)}\}_{k=1}^{K}$, yielding a new image, which we term $\mathbf{g}_{\text{best}}^{(M)}$. Finally, using all the frequency components (pyramid levels) $\mathbf{g}_{\text{best}}^{(M)}$ and $\{\mathbf{l}_{\text{best}}^{(m)}\}_{m=0}^{M-1}$, we reconstruct the full size, complete fused image, $g_{\text{best}}(\tilde{x}, \tilde{y})$. This is in analogy to the operation mentioned above, of reconstructing $g_k(\tilde{x}, \tilde{y})$ from its pyramid components.

Figure 8.12 *A wide DOF mosaic. It is composed of a sequence of frames, each having a narrow FOV and a spatially varying focus. Samples of the sequence are shown in Figure 8.10.*

Mosaicing artifacts Image mosaicing creates artifacts in the form of *seams* in lines that correspond to the boundaries of the raw frames [4,9,16,17,25,28,54]. One reason for seams is spatial variability of exposure, created when an object is seen through different parts of the camera's FOV. For example, it may be caused by vignetting. Other reasons include slight temporal variations of illumination or camera gain between frames. This problem has been easily solved in traditional image mosaicing using image feathering. There, fusion of frames $\{\mathbf{g}_k\}_{k=1}^{K}$ is obtained by a weighted linear superposition of the raw pixels. The superposition weight of a pixel (x, y) in frame k decreases the closer the pixel is to the boundary of this frame [28,54]. Seam removal by the feathering operation is particularly effective if done in low-frequency components [17], since exposure variations (which cause seams) change very smoothly across the camera FOV.

We easily adapted this principle to our problem. Recall that we create the low frequency component $\mathbf{g}_{\text{best}}^{(M)}$ of the fused image by a linear superposition of $\{\mathbf{g}_k^{(M)}\}_{k=1}^{K}$. Hence, we set the weights of the superposition of $\{\mathbf{g}_k^{(M)}\}_{k=1}^{K}$ according to the described feathering principle. To conclude, the multi-scale fusion approach handles both kinds of artifacts: those associated with non-ideal focusing and those stemming from non-ideal mosaicing.

As an example, we analysed a sequence, samples of which are shown in Figure 8.10. The resulting wide DOF mosaic appears in Figure 8.12. The leftmost part of the image contains defocused objects: as explained in Figure 8.6, the peripheral parts are seen in fewer frames, thus with less focus states. As said, the periphery is at most one frame wide.

8.3 Panorama with intensity high dynamic range

8.3.1 Image acquisition

In many scenarios, object radiance changes by orders of magnitude across the FOV. For this reason, there has recently been an upsurge of interest in obtaining HDR image data and in their representation [59–63]. On the other hand, raw images have a limited optical dynamic range [64], set by the limited dynamic range of the camera detector. Above a certain detector irradiance, the images become saturated and their content at the saturated region is lost. Attenuating the irradiance by a shorter exposure time, a smaller aperture, or a neutral (space invariant) density filter can ensure that saturation is avoided. However, at

Figure 8.13 *Two generalised mosaicing systems [31–33]. (Left) A system composed of a Sony black/white video camera and an extended arm which holds the filter. (Right) A system that includes a Canon Optura digital video camera and a cylindrical attachment that holds the filter. In both cases, the camera moves with the attached filter as a rigid system.*

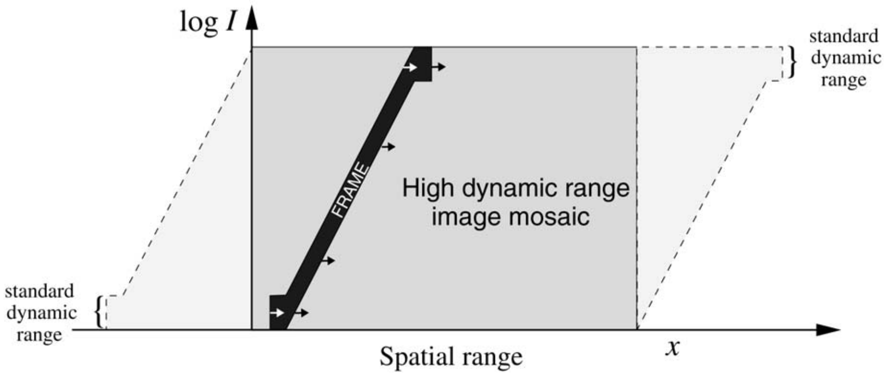


Figure 8.14 *Image mosaicing coupled with vignetting effects yields HDR image mosaics [33]. Besides the FOV, it also extends the intensity dynamic range of the system. Outside the main region of interest, the mosaic provides additional information about the scene periphery, whose quality gradually coincides with that of a single frame.*

the same time other information is lost since light may be below the detection threshold in regions of low image irradiance.

Using generalised mosaicing, extension of both the dynamic range and the FOV are done in a unified framework [33]. We mount a fixed filter on the camera, as in Figures 8.2 and 8.13. The intensity transmittance varies across the filter's extent. This causes an *intentional vignetting*. Including vignetting effects originating from the lens, the overall effect is equivalent to spatially attenuating the image by a mask $A(x, y)$.

Now, as in Section 8.2.2, the scene is scanned by the motion of the camera. The moving system attenuates the light from any scene point differently in each frame. Effectively, the camera captures each point with different exposures during the sequence. Therefore, the system acquires both dark and bright areas with high quality while extending the FOV. It may be viewed as introducing a new dimension to the mosaicing process (Figure 8.14). This dimension leads to the introduction of the concept of the *spatio-intensity space*. In

Figure 8.14, the spatio-intensity support of a single frame occupies a diagonal region in the spatio-intensity space. This occurs if $\log[A(x, y)]$ varies linearly with x. The spatial frame motion then covers the intensity dynamic range, as a by product. High definition intensity may be obtained for all the pixels in a wide FOV image mosaic. In addition, information becomes available about the periphery of the central region of interest: the periphery has a smaller dynamic range, but at least the standard dynamic range of the detector (Figure 8.14).

To demonstrate the appearance of frames taken this way, we used [33] a linear variable density filter, 3 cm long, rigidly attached to an 8-bit CCD camera system, ≈ 30 cm in front of its 25 mm lens. The filter has a maximum attenuation of 1:100. The camera was rotated about its centre of projection so that each point was imaged 14 times across the camera FOV. Some images of this sequence of 36 frames are presented in Figure 8.15.

8.3.2 Data fusion

We now describe the method we used [33] to estimate the intensity at each mosaic point, given its multiple corresponding measurements. As in Section 8.2.3, this is done after the images have been registered. Let a measured intensity readout at a point be g_k with uncertainty Δg_k, and the estimated mask be $\hat{A}$ with uncertainty $\Delta \hat{A}$. Compensating the readout for the mask, the scene point's intensity is

$$I_k = \frac{g_k}{\hat{A}} \tag{8.12}$$

and its uncertainty is

$$\Delta I_k = \sqrt{\left(\frac{\partial I_k}{\partial g_k}\Delta g_k\right)^2 + \left(\frac{\partial I_k}{\partial \hat{A}}\Delta \hat{A}\right)^2} \tag{8.13}$$

For instance, we may set the readout uncertainty to be $\Delta g_k = 0.5$, since the intensity readout values are integers. Any image pixel considered to be saturated (g_k close to 255 for an 8-bit detector) is treated as having a high uncertainty. Thus, its corresponding Δg_k is set to be a very large number.

Assuming the measurements I_k to be Gaussian and independent, the log-likelihood for a value I behaves as $-E^2$, where

$$E^2 \equiv \sum_k \left(\frac{I - I_k}{\Delta I_k}\right)^2 \tag{8.14}$$

The maximum likelihood (ML) solution for the intensity I in this scene point is the one that minimises E^2:

Figure 8.15 *Frames 4, 9, 11, 15, 17, 23, and 31 from a sequence taken with a linear variable density filter [33]. Scene features become brighter as they move leftwards in the frame. Bright scene points gradually reach saturation. Dim scene points, which are not visible in the right-hand side of the frames, become visible when they appear on the left.*

$$\hat{I} = \widehat{\Delta I}^2 \sum_k \frac{I_k}{\Delta I_k^2} \tag{8.15}$$

where

$$\widehat{\Delta I} = \left(0.5 \cdot \frac{\mathrm{d}^2 E^2}{\mathrm{d}I^2}\right)^{-1/2} = \left(\sqrt{\sum_k \frac{1}{\Delta I_k^2}}\right)^{-1} \tag{8.16}$$

Although Equation (8.15) suffices to determine the value at each point, annoying seams may appear at the boundaries of the frames that compose the mosaic. At these boundaries there is a transition between points that have been estimated using somewhat different sources of data. Seams appear also at the boundaries of saturated areas, where there is an abrupt change in the uncertainty Δg, while the change in g is usually small. These seams are removed by feathering techniques: see [33].

Images from a sequence, of which samples are shown in Figure 8.15, were fused into a mosaic using this method. The histogram equalised version of $\log \hat{I}$ is shown in Figure 8.16. Contrast stretching of $\hat{I}$ in selected regions shows that the mosaic is not saturated anywhere, and details are seen wherever $I \geqslant 1$. The HDR range of this mosaic is evident. The periphery parts of the mosaic are left of the $\mathcal{L}$ mark and right of the $\mathcal{R}$ mark, having a width of a single frame. These parts were not exposed at the full range of attenuation. HDR is observed there as well, but the range gradually decreases to that of the native detector.

8.4 Multispectral wide field of view imaging

Multispectral imaging is very useful in numerous imaging applications, including object and material recognition [65], colour analysis and constancy [66–68], remote sensing [65,69–71], and astronomy [72]. The applications for multispectral imaging are expanding [73], and include for example, medical imaging, agriculture, archaeology and art.

In this section we describe the use of a spatially varying interference filter in the generalised mosaicing framework [32]. In particular, a linear interference filter passes a narrow wavelength band, and its central wavelength λ varies linearly with x. Such a filter can be used in a system as depicted in Figures 8.2 and 8.13. In this case, the spectral information in each raw frame is multiplexed with the spatial features which appear in ordinary images. This is seen, for example, in frames shown in Figure 8.17, acquired using a monochrome camera. This scene was acquired under incandescent lighting. The spatial details of the scene are clearly recognisable (e.g., the computer monitor). The spatial features are clear because, as depicted in Figure 8.2, the system is an imaging device, thus each of its frames captures an area of the scene.

Once the raw frames are registered, we have for each scene point in the mosaic FOV a set of wavelength samples $\{\lambda_k(\tilde{x}, \tilde{y})\}$, and corresponding intensity measurements $\{g_k(\tilde{x}, \tilde{y})\}$, where k is the index of the individual frames that compose the mosaic. This raw data structure is converted [32] to a multispectral *image cube*, denoted $g(\tilde{x}, \tilde{y}, \lambda)$. The multispectral data is now available at each scene point, and can be used in multispectral

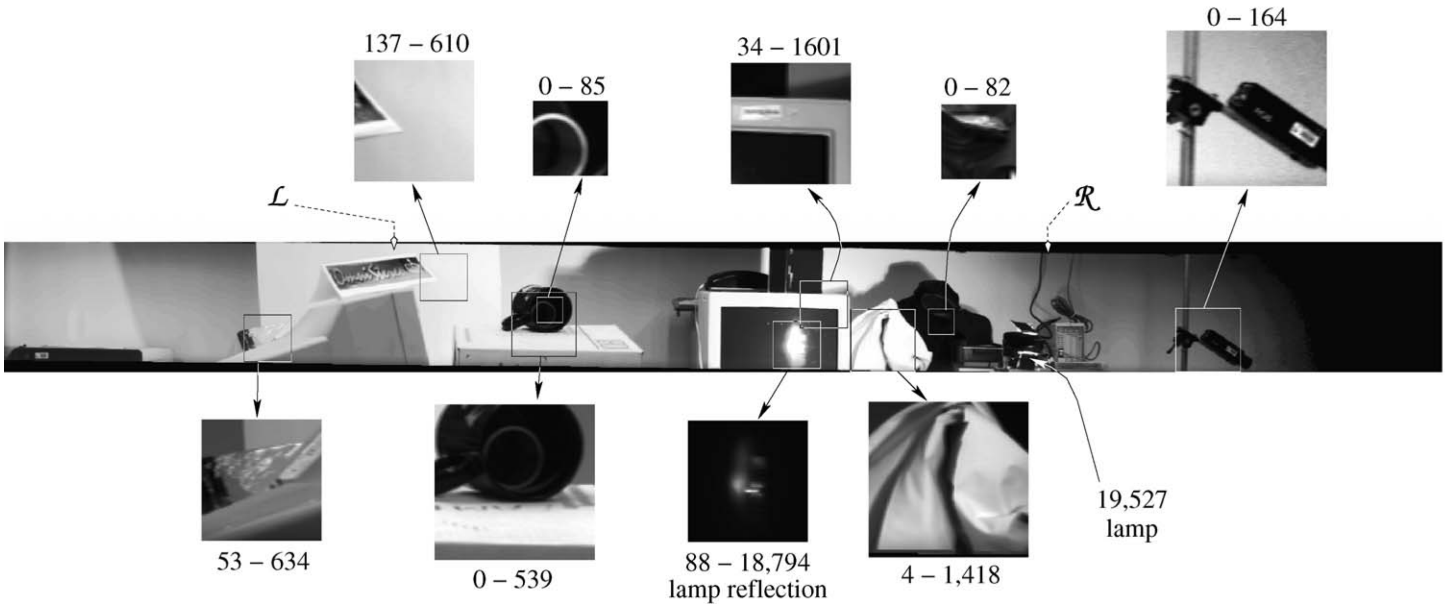
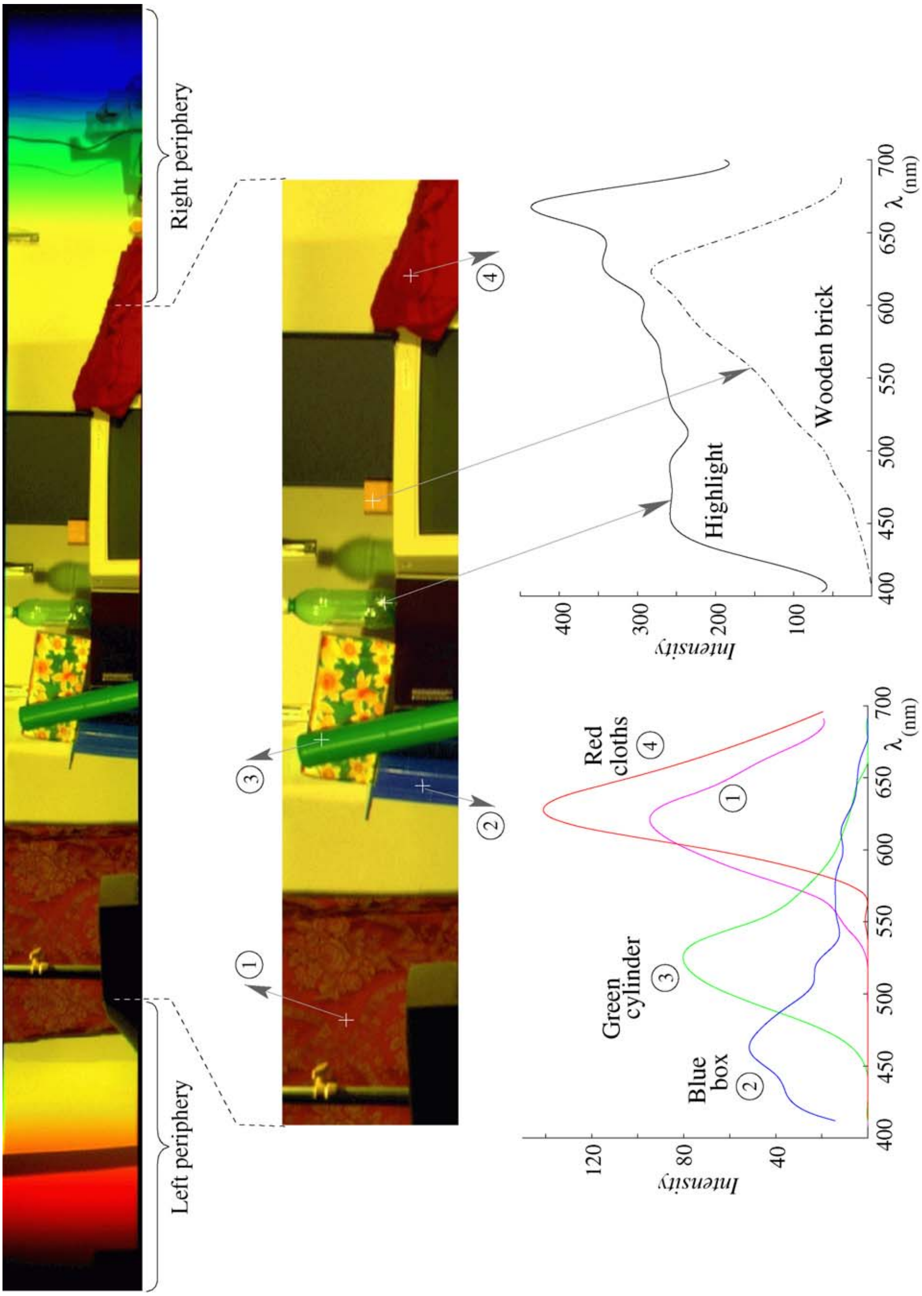


Figure 8.16 *An image created using a generalised mosaicing system [33]. It is based on a single rotation about the centre of projection of an 8-bit video camera. Contrast stretching in the selected squares reveals the details that reside within the computed mosaic. The numbers near the squares are the actual (unstretched) brightness ranges within the squares. Note the shape of the filament of the lamp in its reflection from the computer monitor. The periphery regions are left of the $\mathcal{L}$ mark and right of the $\mathcal{R}$ mark.*

Figure 8.17 *Frames 29, 32, 38, 40, 60, and 64 of a sequence [32] taken through the linear variable interference filter. The left of the FOV senses the energy density at 700 nm, while the right senses it at 400 nm. The spatial features of the scene are clearly seen.*

imaging applications. For display, it is possible to convert the spectrum in each point to Red–Green–Blue values, as described in [32].

In an experiment corresponding to the images in Figure 8.17, about 21 spectral samples were acquired for each scene point. The grabbed images were compensated for cam-



Right periphery
Left periphery
Highlight
Wooden brick
Intensity
400
300
200
100
400
450
500
550
600
650
700
λ (nm)
Red cloths
Green cylinder
Blue box
120
80
40

Figure 8.18 *(Top) A colour image mosaic [32] rendered using the spectral data acquired at each point in its FOV, based on a single pass (rotation about the centre of projection) of an ordinary black/white camera with a single fixed filter. The scene was illuminated by incandescent lamps. In the mosaic periphery the spectral range becomes narrower towards the outer boundaries, thus gradually deteriorating the colour rendering. (Middle) The mosaic's central region of interest contains the full spectral range the filter can scan. (Bottom) The spectrum is plotted for selected points.*

era vignetting effects that were calibrated beforehand. We registered the images using a method discussed in [32]. The registration yielded a wide multispectral image mosaic, where the spectrum can be computed for each point. The multispectral mosaic was then converted to a colour mosaic, shown in Figure 8.18. Similarly to Sections 8.2 and 8.3, a full range spectrum is obtained in the central region of interest. This region seems yellowish in Figure 8.18 because the light coming from the incandescent lamps is rather yellow.[5] Other than that, the estimated colours were consistent with the appearance of the objects.

Information is obtained also about the periphery, though with decreasing spectral range. The top of Figure 8.18 indeed shows the periphery regions, one frame wide, on both sides of the mosaic central part. Since the spectral range in these regions changes gradually from the central part, there is no abrupt decrease of quality in the periphery. However, towards the outer boundaries of the left and right periphery, the image becomes red and blue respectively. This is due to the absence of data on the complimentary wavelengths in these regions [32]. Even there, substantial information can still be available for algorithms that make do with partial spectral data, or that do not rely on colour but on spatial features. For example, the objects in the right periphery clearly appear in the raw frame shown at the last photograph in Figure 8.17. It shows [32] loose dark cables hanging down through the frame, and their shadows on the wall behind. This can indicate the number and spatial distribution of the light sources in the scene, as in [74]. In addition, other objects (shaped bricks) can be recognised in this region. Therefore, the peripheral regions are not wasted data, but can be useful for computer vision.

8.5 Polarisation as well

Polarimetric imaging has been used in numerous imaging applications [75,76], including object and material recognition [77,78], shape recovery [79–81], and removal and analysis of specular reflection in photography and computer vision [82–85]. It has also been used for removal of scattering effects [86,87], e.g. in haze [88–90], underwater [78,88, 91–94] and tissue [95].

The polarisation state has several parameters. In linear polarisation these parameters are the intensity, the degree of polarisation, and the orientation of the plane of polarisation. To recover the polarisation parameters corresponding to each scene point, it is usually sufficient to measure the scene several times, each time with different polarisation settings [34]. Typically, man-made systems achieve this by filtering the light through a linear

[5]The human visual system adapts when it is embedded in such a coloured illumination (colour constancy).

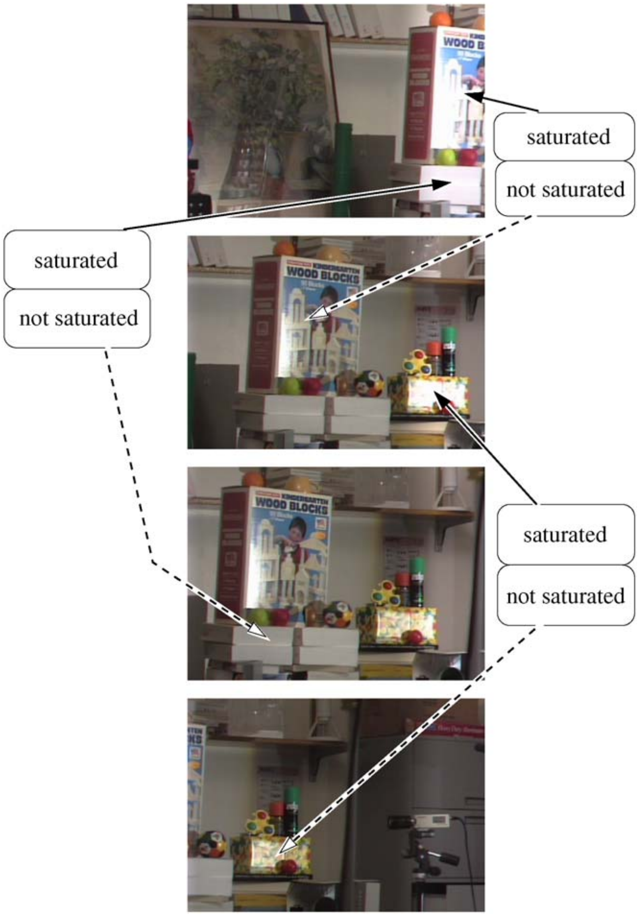


Figure 8.19 *A few frames from a sequence [34] taken through a mounted filter. The filter has spatially varying attenuation and polarisation characteristics. The effective filtering variations are gradual and not easily visible, due to defocus blur of the filter. Nevertheless, the spatial variations make objects appear brighter on the right-hand side. Thus saturated pixels on the right appear unsaturated in other frames, in corresponding left-hand side pixels.*

polariser, oriented differently in different images. Biological systems, however, use other mechanisms to capture polarisation images. Specifically, the retina of the mantis shrimp [96,97] has several distinct regions, each having different optical properties. In order to capture all the information in high quality, the shrimp moves its eyes to scan the FOV, thereby sequentially measuring each scene point using different optical settings [96]. This gives the shrimp vision high quality colour, polarisation and spatial information in a wide FOV. This biological evidence motivates the application of generalised mosaicing to polarisation sensing.

This principle can be demonstrated by attaching a spatially varying polarisation filter in front of the camera [34]. As the camera moves, a given scene point is measured multiple times, each through a filter part with different polarisation characteristics and/or orientation. Computational algorithms tailored to this kind of a system extract the required polarisation parameters, in a wide FOV. To demonstrate this, a camera was mounted, as in Figure 8.13, with a filter that varies the polarisation filtering across it. Furthermore, the system was built such that its transmittance varied across the camera FOV as well (see details in Ref. [34]). This enhances the dynamic range of the mosaic, as in Section 8.3. For instance, let scene points appear saturated in a frame, when viewed through a filter part having high transmittance. These points may become unsaturated in other frames, when the points are viewed through darker portions of the filter. Furthermore, the polarisation sensitivity of the mounted spatially varying filter can significantly reduce specular highlights, thus aiding in the dynamic range extension. This is seen in sample frames taken by the system [34], shown in Figure 8.19. Note that objects are brighter (and even saturated) when they appear on the right-hand side of the frame.

Defocus blur affects the filtering properties of the mounted filter. Thus Ref. [34] describes a single framework that handles analysis of data acquired by non-ideal polarisation filters (partial polarisers), variable exposures and saturation. As the images are automatically registered, they can be analysed by algorithms described in Ref. [34], to yield a panoramic, polarimetric seamless image mosaic [34].

8.6 Conclusions

Generalised mosaicing is a framework for capturing information in multiple imaging dimensions. In Section 8.2 we described how the DOF can be extended while enlarging the FOV. Contrary to common optical designs, an SLC that is not flat or not normal to the optical axis can be beneficial, as it enables the extraction of depth information when the scene is scanned. This has implications for several aspects of computer vision, such as image-based rendering [98]. It may also be applied to machine vision systems (e.g., microscopic ones) used for industrialised inspection.

In addition to focus, we demonstrated this framework by deriving mosaics having HDR, multispectral or polarimetric outputs. Nevertheless, generalised mosaicing permits simultaneous enhancement of multiple dimensions. One example for this was described in Section 8.5, where both an extended intensity dynamic range and polarisation information were extracted using a single spatially varying filter. Other possibilities may include simultaneous extraction of polarisation and spectral information in a wide field, as the mantis shrimp [96,97]; simultaneous expansion of focus and intensity range, or other combinations. Furthermore, generalised mosaicing can be used for self-calibration of simultaneous radiometric effects that occur in the camera, including vignetting, automatic-gain control (AGC) and radiometric nonlinearity [99,100]. This is achieved by exploiting the redundancy encapsulated in multiple overlapping frames to retrieve these radiometric parameters.

Acknowledgements

The studies described in this work were supported in parts by a National Science Foundation ITR Award, IIS-00-85864, a David and Lucile Packard Fellowship, the Louis Morin

Fellowship, the US–Israel Binational Science Foundation (BSF), and the Ollendorff Minerva Center in the Electrical Engineering Department at the Technion. Minerva is funded through the BMBF. Yoav Schechner is a Landau Fellow – supported by the Taub Foundation, and an Alon Fellow.

References

[1] D. Capel and A. Zisserman, 'Automated mosaicing with super-resolution zoom', in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1998, pp. 885–891.

[2] S. Peleg, M. Ben-Ezra and Y. Pritch, 'Omnistereo: Panoramic stereo imaging', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, 2001, pp. 279–290.

[3] R.M. Batson, 'Digital cartography of the planets: New methods, its status, and its future', *Photogrammetric Engineering and Remote Sensing*, Vol. 53, 1987, pp. 1211–1218.

[4] M.L. Duplaquet, 'Building large image mosaics with invisible seam lines', in *Visual Information Processing VII*, in *Proc. of the SPIE*, Vol. 3387, 1998, pp. 369–377.

[5] R. Eustice, O. Pizarro, H. Singh and J. Howland, 'UWIT: Underwater Image Toolbox for optical image processing and mosaicking in MATLAB', in *Proc. IEEE International Symposium on Underwater Technology*, 2002, pp. 141–145.

[6] R. Garcia, J. Batlle, X. Cufi and J. Amat, 'Positioning an underwater vehicle through image mosaicking', in *Proc. IEEE International Conference on Robotics and Automation*, Part 3, 2001, pp. 2779–2784.

[7] R. Kwok, J.C. Curlander and S. Pang, 'An automated system for mosaicking spaceborne SAR imagery', *International Journal of Remote Sensing*, Vol. 11, 1990, pp. 209–223.

[8] S. Hsu, H.S. Sawhney and R. Kumar, 'Automated mosaics via topology inference', *IEEE Computer Graphics and Application*, Vol. 22, No. 2, 2002, pp. 44–54.

[9] M. Irani, P. Anandan, J. Bergen, R. Kumar and S. Hsu, 'Efficient representations of video sequences and their application', *Signal Processing: Image Communication*, Vol. 8, 1996, pp. 327–351.

[10] A. Smolić and T. Wiegand, 'High-resolution image mosaicing', in *Proc. International Conference on Image Processing*, Vol. 3, 2001, pp. 872–875.

[11] C.J. Lada, D.L. DePoy, K.M. Merrill and I. Gatley, 'Infrared images of M17', *The Astronomical Journal*, Vol. 374, 1991, pp. 533–539.

[12] J.M. Uson, S.P. Boughn and J.R. Kuhn, 'The central galaxy in Abel 2029: An old supergiant', *Science*, Vol. 250, 1990, pp. 539–540.

[13] E.M. Reynoso, G.M. Dubner, W.M. Goss and E.M. Arnal, 'VLA observations of neutral hydrogen in the direction of Puppis A', *The Astronomical Journal*, Vol. 110, 1995, pp. 318–324.

[14] R. Bernstein, 'Digital image processing of earth observation sensor data', *IBM Journal of Research and Development*, Vol. 20, No. 1, 1976, pp. 40–57.

[15] M. Hansen, P. Anandan, K. Dana, G. van der Wal and P. Burt, 'Real-time scene stabilization and mosaic construction', in *Proc. IEEE Workshop on Applications of Computer Vision*, 1994, pp. 54–62.

[16] E. Fernandez, R. Garfinkel and R. Arbiol, 'Mosaicking of aerial photographic maps via seams defined by bottleneck shortest paths', *Operations Research*, Vol. 46, No. 3, 1998, pp. 293–304.

[17] L.A. Soderblom, K. Edwards, E.M. Eliason, E.M. Sanchez and M.P. Charette, 'Global color variations on the Martian surface', *Icarus*, Vol. 34, 1978, pp. 446–464.

[18] A.R. Vasavada, A.P. Ingersoll, D. Banfield, M. Bell, P.J. Gierasch and M.J.S. Belton, 'Galileo imaging of Jupiter's atmosphere: The great red spot, equatorial region, and white ovals', *Icarus*, Vol. 135, 1998, pp. 265–275.

[19] R.D. Ballard, 'Ancient Ashkelon', *National Geographic Magazine*, Vol. 199, No. 1, 2001, pp. 61–93.

[20] D. Drollette, 'Photonics defies the depths', *Photonics Spectra*, Vol. 34, No. 11, 2000, pp. 80–90.

[21] S. Negahdaripour, X. Xu, A. Khemene and Z. Awan, '3-D motion and depth estimation from sea-floor images for mosaic-based station-keeping and navigation of ROV's/AUV's and high-resolution sea-floor mapping', in *Proc. IEEE Workshop on Autonomous Underwater Vehicles*, 1998, pp. 191–200.

[22] S. Coorg, N. Master and S. Teller, 'Acquisition of a large pose-mosaic dataset', in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1998, pp. 872–878.

[23] S. Gumustekin and R.W. Hall, 'Mosaic image generation on a flattened sphere', in *Proc. IEEE Workshop on Applications of Computer Vision*, 1996, pp. 50–55.

[24] S. Mann, 'Joint parameter estimation in both domain and range of functions in same orbit of the projective-Wyckoff group', in *Proc. International Conference on Image Processing*, 1996, pp. 193–196.

[25] S. Peleg and J. Herman, 'Panoramic mosaics by manifold projection', in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 338–343.

[26] H.S. Sawhney and S. Ayer, 'Compact representations of videos through dominant and multiple motion estimation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, 1996, pp. 814–830.

[27] H.S. Sawhney, R. Kumar, G. Gendel, J. Bergen, D. Dixon and V. Paragano, 'VideoBrush™: Experiences with consumer video mosaicing', in *Proc. IEEE Workshop on Applications of Computer Vision*, 1998, pp. 52–62.

[28] H.Y. Shum and R. Szeliski, 'Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment', *International Journal of Computer Vision*, Vol. 36, 2000, pp. 101–130.

[29] R. Szeliski, 'Image mosaicing for telereality applications', in *Proc. IEEE Workshop on Applications of Computer Vision*, 1994, pp. 44–53.

[30] M. Aggarwal and N. Ahuja, 'High dynamic range panoramic imaging', in *Proc. IEEE International Conference on Computer Vision*, Vol. I, 2001, pp. 2–9.

[31] Y.Y. Schechner and S.K. Nayar, 'Generalized mosaicing', in *Proc. IEEE International Conference on Computer Vision*, Vol. I, 2001, pp. 17–24.

[32] Y.Y. Schechner and S.K. Nayar, 'Generalized mosaicing: Wide field of view multispectral imaging', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, 2002, pp. 1334–1348.

[33] Y.Y. Schechner and S.K. Nayar, ‘Generalized mosaicing: High dynamic range in a wide field of view’, *International Journal of Computer Vision*, Vol. 53, 2003, pp. 245–267.

[34] Y.Y. Schechner and S.K. Nayar, ‘Generalized mosaicing: Polarization panorama’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, 2005, pp. 631–636.

[35] E. Hecht, *Optics*, fourth ed., Addison–Wesley, New York, 2002, pp. 264–266.

[36] A. Krishnan and N. Ahuja, ‘Panoramic image acquisition’, in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1996, pp. 379–384.

[37] K. Itoh, A. Hayashi and Y. Ichioka, ‘Digitized optical microscopy with extended depth of field’, *Applied Optics*, Vol. 28, No. 16, 1989, pp. 3487–3493.

[38] R.A. Jarvis, ‘A perspective on range-finding techniques for computer vision’, *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol. 5, No. 2, 1983, pp. 122–139.

[39] H.N. Nair and C.V. Stewart, ‘Robust focus ranging’, in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1992, pp. 309–314.

[40] S.K. Nayar, ‘Shape from focus system’, in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1992, pp. 302–308.

[41] M. Noguchi and S.K. Nayar, ‘Microscopic shape from focus using active illumination’, in *Proc. International Conference on Image Processing*, Vol. 1, 1994, pp. 147–152.

[42] Y.Y. Schechner, N. Kiryati and R. Basri, ‘Separation of transparent layers using focus’, *International Journal of Computer Vision*, Vol. 89, 2000, pp. 25–39.

[43] M. Subbarao and J.K. Tyan, ‘The optimal focus measure for passive autofocusing’, in *Videometrics VI*, in *Proc. of the SPIE*, Vol. 2598, 1995, pp. 88–99.

[44] P. Torroba, N. Cap and H. Rabal, ‘Defocus detection using a visibility criterion’, *Journal of Modern Optics*, Vol. 41, No. 1, 1994, pp. 111–117.

[45] T.T.E. Yeo, S.H. Ong, Jayasooriah and R. Sinniah, ‘Autofocusing for tissue microscopy’, *Image and Vision Computing*, Vol. 11, No. 10, 1993, pp. 629–639.

[46] T. Darrell and K. Wohn, ‘Pyramid based depth from focus’, in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1988, pp. 504–509.

[47] K. Engelhardt and G. Hausler, ‘Acquisition of 3-D data by focus sensing’, *Applied Optics*, Vol. 27, No. 22, 1988, pp. 4684–4689.

[48] S.A. Sugimoto and Y. Ichioka, ‘Digital composition of images with increased depth of focus considering depth information’, *Applied Optics*, Vol. 24, No. 14, 1985, pp. 2076–2080.

[49] Y. Xiong and S.A. Shafer, ‘Depth from focusing and defocusing’, in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1993, pp. 68–73.

[50] P. Favaro and S. Soatto, ‘A geometric approach to shape from defocus’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, 2006, pp. 406–417.

[51] A. Pentland, S. Scherock, T. Darrell and B. Girod, ‘Simple range camera based on focal error’, *Journal of the Optical Society of America A*, Vol. 11, 1994, pp. 2925–2934.

[52] A.N. Rajan and S. Chaudhuri, 'Simultaneous estimation of super-resolved scene and depth map from low resolution defocused observations', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, 2003, pp. 1102–1117.
[53] Y.Y. Schechner and N. Kiryati, 'Depth from defocus vs. stereo: How different really are they?', *International Journal of Computer Vision*, Vol. 89, 2000, pp. 141–162.
[54] P.J. Burt and E.H. Adelson, 'A multiresolution spline with application to image mosaics', *ACM Transactions on Graphics*, Vol. 2, 1983, pp. 217–236.
[55] P.J. Burt and R.J. Kolczynski, 'Enhanced image capture through fusion', in *Proc. IEEE International Conference on Computer Vision*, 1993, pp. 173–182.
[56] J.M. Odgen, E.H. Adelson, J.R. Bergen and P.J. Burt, 'Pyramid-based computer graphics', *RCA Engineer*, Vol. 30, No. 5, 1985, pp. 4–15.
[57] A. Toet, 'Hierarchical image fusion', *Machine Vision and Applications*, Vol. 3, 1990, pp. 1–11.
[58] P.J. Burt and E.H. Adelson, 'The Laplacian pyramid as a compact image code', *IEEE Transactions on Communications*, Vol. 31, 1983, pp. 532–540.
[59] F. Durand and J. Dorsey, 'Fast bilateral filtering for the display of high-dynamic-range images', *ACM Transactions on Graphics*, Vol. 21, No. 3, 2002, pp. 257–266.
[60] R. Fattal, D. Lischinski and M. Werman, 'Gradient domain high dynamic range compression', *ACM Transactions on Graphics*, Vol. 21, No. 3, 2002, pp. 249–256.
[61] G.W. Larson, H. Rushmeier and C. Piatko, 'A visibility matching tone reproduction operator for high dynamic range scenes', *IEEE Transactions on Visualization and Computer Graphics*, Vol. 3, 1997, pp. 291–306.
[62] A. Pardo and G. Sapiro, 'Visualization of high dynamic range images', in *Proc. International Conference on Image Processing*, Vol. 1, 2002, pp. 633–636.
[63] D.A. Socolinsky, 'Dynamic range constraints in image fusion and realization', in *Proc. IASTED International Conference on Signal and Image Processing*, 2000, pp. 349–354.
[64] W. Ogiers, 'Survey of CMOS imagers', IMEC Report P60280-MS-RP-002, Issue 1.1, Part 1, 1997.
[65] D. Slater and G. Healey, 'Material classification for 3D objects in aerial hyperspectral images', in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, 1999, pp. 268–273.
[66] A. Abrardo, L. Alparone, V. Cappellini and A. Prosperi, 'Color constancy from multispectral images', in *Proc. International Conference on Image Processing*, Vol. 3, 1999, pp. 570–574.
[67] M. Hauta-Kasari, K. Miyazawa, S. Toyooka and J. Parkkinen, 'A prototype of the spectral vision system,' in *Proc. Scandinavian Conference on Image Analysis*, Vol. 1, 1999, pp. 79–86.
[68] H.M.G. Stokman, T. Gevers and J.J. Koenderink, 'Color measurement by imaging spectrometry', *Computer Vision and Image Understanding*, Vol. 79, No. 2, 2000, pp. 236–249.
[69] A.M. Mika, 'Linear-wedge spectrometer', in *Imaging Spectroscopy of the Terrestrial Environment*, in *Proc. of the SPIE*, Vol. 1298, 1990, pp. 127–131.
[70] X. Sun and J.M. Anderson, 'A spatially variable light-frequency-selective component-based, airborne pushbroom imaging spectrometer for the water environment', *Photogrammetric Engineering and Remote Sensing*, Vol. 59, No. 3, 1993, pp. 399–406.

[71] J.B. Wellman, ‘Multispectral mapper: Imaging spectroscopy as applied to the mapping of earth resources’, in *Imaging Spectroscopy*, in *Proc. of the SPIE*, Vol. 268, 1981, pp. 64–73.
[72] G. Monnet, ‘3D spectroscopy with large telescopes: Past, present and prospects’, in *Tridimensional Optical Spectroscopic Methods in Astronomy*, in *ASP Conference Series*, Vol. 71, 1995, pp. 2–17.
[73] N. Gat, ‘Imaging spectroscopy using tunable filters: A review’, in *Wavelet Applications VII*, in *Proc. of the SPIE*, Vol. 4056, 2000, pp. 50–64.
[74] I. Sato, Y. Sato and K. Ikeuchi, ‘Illumination distribution from brightness in shadows: Adaptive estimation of illumination distribution with unknown reflectance properties in shadow regions’, in *Proc. IEEE International Conference on Computer Vision*, 1999, pp. 875–883.
[75] G.P. Können, *Polarized Light in Nature*, Cambridge University Press, Cambridge, 1985.
[76] W.A. Shurcliff and S.S. Ballard, *Polarized Light*, Van Nostrand, Princeton, 1964.
[77] L.B. Wolff, ‘Polarization camera for computer vision with a beam splitter’, *Journal of the Optical Society of America A*, Vol. 11, 1994, pp. 2935–2945.
[78] L.B. Wolff, ‘Polarization vision: A new sensory approach to image understanding’, *Image and Vision Computing*, Vol. 15, 1997, pp. 81–93.
[79] D. Miyazaki, M. Masataka and K. Ikeuchi, ‘Transparent surface modeling from a pair of polarization images’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26, 2004, pp. 73–82.
[80] S. Rahmann and N. Canterakis, ‘Reconstruction of specular surfaces using polarization imaging’, in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 1, 2001, pp. 149–155.
[81] A.M. Wallace, B. Laing, E. Trucco and J. Clark, ‘Improving depth image acquisition using polarized light’, *International Journal of Computer Vision*, Vol. 32, 1999, pp. 87–109.
[82] H. Farid and E.H. Adelson, ‘Separating reflections from images by use of independent component analysis’, *Journal of the Optical Society of America A*, Vol. 16, 1999, pp. 2136–2145.
[83] Hermanto, A.K. Barros, T. Yamamura and N. Ohnishi, ‘Separating virtual and real objects using independent component analysis’, *IEICE Transactions on Information & Systems*, Vol. E84-D, No. 9, 2001, pp. 1241–1248.
[84] S. Lin and S.W. Lee, ‘Detection of specularity using stereo in color and polarization space’, *Computer Vision and Image Understanding*, Vol. 65, 1997, pp. 336–346.
[85] Y.Y. Schechner, J. Shamir and N. Kiryati, ‘Polarization statistical analysis of scenes containing a semireflector’, *Journal of the Optical Society of America A*, Vol. 17, 2000, pp. 276–284.
[86] L.J. Denes, M. Gottlieb, B. Kaminsky and P. Metes, ‘AOTF polarization difference imaging’, *Proc. of the SPIE*, Vol. 3584, 1999, pp. 106–115.
[87] J.S. Tyo, M.P. Rowe, E.N. Pugh Jr. and N. Engheta, ‘Target detection in optically scattering media by polarization-difference imaging’, *Applied Optics*, Vol. 35, 1996, pp. 1855–1870.
[88] Y.Y. Schechner and Y. Averbuch, ‘Regularized image recovery in scattering media’, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, in press.
[89] Y.Y. Schechner, S.G. Narasimhan and S.K. Nayar, ‘Polarization-based vision through haze’, *Applied Optics*, Vol. 42, No. 3, 2003, pp. 511–525.

[90] S. Shwartz, E. Namer and Y.Y. Schechner, 'Blind haze separation', in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, 2006, pp. 1984–1991.

[91] S. Harsdorf, R. Reuter and S. Tönebön, 'Contrast-enhanced optical imaging of submersible targets', *Proc. of the SPIE*, Vol. 3821, 1999, pp. 378–383.

[92] Y.Y. Schechner and N. Karpel, 'Recovery of underwater visibility and structure by polarization analysis', *IEEE Journal of Oceanic Engineering*, Vol. 30, 2005, pp. 570–587.

[93] J.S. Talyor Jr. and L.B. Wolff, 'Partial polarization signature results from the field testing of the shallow water real-time imaging polarimeter (SHRIMP)', in *Proc. MTS/IEEE Oceans*, Vol. 1, 2001, pp. 107–116.

[94] T. Treibitz and Y.Y. Schechner, 'Instant 3Descatter', in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, 2006, pp. 1861–1868.

[95] S.G. Demos and R.R. Alfano, 'Optical polarization imaging', *Applied Optics*, Vol. 36, 1997, pp. 150–155.

[96] T.W. Cronin and J. Marshall, 'Parallel processing and image analysis in the eyes of the mantis shrimp', *The Biological Bulletin*, Vol. 200, 2001, pp. 177–183.

[97] T.W. Cronin, N. Shashar, R. Caldwell, J. Marshall, A.G. Cheroske and T.H. Chiou, 'Polarization vision and its role in biological signaling', *Integrative and Comparative Biology*, Vol. 43, 2003, pp. 549–558.

[98] K. Aizawa, K. Kodama and Kubota, 'Producing object-based special effects by fusing multiple differently focused images', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 10, 2000, pp. 323–330.

[99] A. Litvinov and Y.Y. Schechner, 'Radiometric framework for image mosaicking', *Journal of the Optical Society of America A*, Vol. 22, 2005, pp. 839–848.

[100] A. Litvinov and Y.Y. Schechner, 'Addressing radiometric nonidealities: A unified framework', in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, 2005, pp. 52–59.