Generalized Mosaicing

Yoav Y. Schechner and Shree K. Nayar
Department of Computer Science, Columbia University, New York, NY 10027
{yoav,nayar}@cs.columbia.edu

Abstract

We present an approach that significantly enhances the capabilities of traditional image mosaicing. The key observation is that as a camera moves, it senses each scene point multiple times. We rigidly attach to the camera an optical filter with spatially varying properties, so that multiple measurements are obtained for each scene point under different optical settings. Fusing the data captured in the multiple images 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, or enhancements to other dimensions of imaging. We refer to this approach as generalized mosaicing. The approach was tested using a filter with spatially varying transmittance and a standard 8-bit black/white video camera, to achieve image mosaicing with dynamic range comparable to imaging with a 16-bit camera. In another experiment, we attached a spatially varying spectral filter to the same camera to obtain mosaics that represent the spectral distribution (rather than the usual RGB measurements) of each scene point. We also discuss how generalized mosaicing can be used to explore other imaging dimensions.

1 Multi-Dimensional Mosaics

Image mosaicing is a very popular way to obtain a wide field of view (FOV) image of a scene. The basic idea is to capture images as a camera moves and stitch these images together to obtain a larger image. Image mosaicing has found applications in consumer photography [3, 13, 18, 24, 25, 28] as well as uses in various scientific disciplines [11, 16, 32, 33]. It addresses the fundamental problem of increasing the FOV without sacrificing spatial resolution.

We show that image mosaicing can be generalized to extract much more information about the scene, given a similar amount of acquired data. We refer to this approach as *generalized mosaicing*. The basic observation is that a typical video sequence acquired during mosaicing has great redundancy in terms of the data it contains; as the camera moves, each scene point is observed multiple times. We wish to exploit this fact to explore additional aspects of imaging. Consider the setup shown in Fig. 1. A fixed filter with spatially varying properties is rigidly attached to the camera. Hence, as the camera moves (or simply rotates), each scene point is measured under different optical settings[1]. This simple optical filtering significantly reduces the redundancy in the captured video stream. In fact, the filtering embeds in the acquired data more information about each point in the mosaic FOV. Beside mounting the fixed filter, the image acquisition in generalized mosaicing is identical to traditional mosaicing.

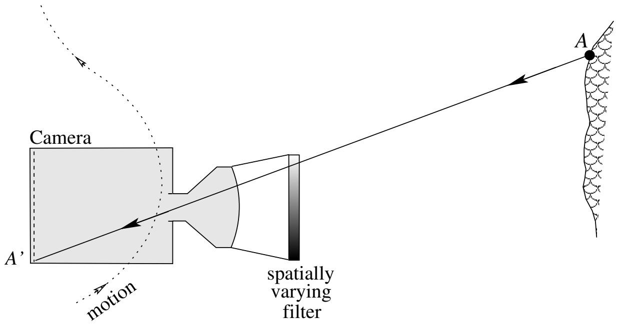

Figure 1. Scene point A is imaged on the detector at A' through a spatially varying filter attached to the camera. As the imaging system moves, each scene point is sensed through different portions of the filter, thus multiple measurements are obtained under different optical settings.

For example, if the filter has spatially varying transmittance, the scene is effectively measured with different exposures. These measurements can be combined to obtain a high dynamic range (HDR) mosaic. Alternatively, if the spectral band transmitted by the filter varies spatially, we obtain multispectral data for each scene point. If a different imaging dimension is of interest to the user, all he/she needs to do is change the optical filter. Note that in previous work, the enhancement of each imaging dimension (e.g., FOV, dynamic range, spectral resolution, polarization sensing, etc.) was considered separately from the others. In contrast, generalized mosaicing provides a single, unified framework to enhance all or some of these dimensions.

The only requirements for implementing generalized mosaicing are a simple optical filter and algorithms for im-

[1] The filter is not placed right next to the lens, as this would only alter the aperture properties [8] without producing spatially varying effects in the image.

Figure 2. Two generalized mosaicing systems. [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 camera and a cylindrical attachment that holds the filter. In both cases, the camera moves with the attached filter as a rigid system.

age registration and fusion that are compatible with the filter's spatially varying characteristics. The filter may be mounted inside the camera, though it is much simpler and more flexible to mount it externally. Fig. 2 shows two prototype systems we developed for our experiments. In one experiment, a standard 8-bit black-and-white video camera and a spatially varying neutral density filter were combined to form a mosaicing system[2] with dynamic range comparable to a 16-bit camera. In another experiment, we used a black/white video camera and a spatially varying spectral filter to obtain multispectral mosaics.

2 Dynamic Range from Motion

We can use generalized mosaicing to significantly extend the brightness dynamic range of imaging systems. Capturing both dim and bright objects with good quality is very important, and indeed there have been several approaches for extending the dynamic range. One approach to HDR imaging is to fuse differently exposed images [2, 6, 19, 21] acquired sequentially with a static camera. Other approaches are based on specialized hardware such as a mosaic of neutral density filters placed on the detector [23], or using detectors with logarithmic response [5].

In the realm of mosaicing, simultaneous extension of the FOV and dynamic range by exploiting the automatic gain control (AGC) feature of a camera was proposed in [18]. However, AGC has a global (or, at best regional) effect and does not guarantee the required measurements at all scene points. For instance, a bright scene point may remain saturated all through the image sequence if it happens to be surrounded by a large dark area. Similarly, a dim point may remain dark through the sequence when it is surrounded by a bright area.

In our approach of generalized mosaicing, we mount a filter on the camera whose intensity transmittance varies across the filter's extent. This causes an *intended vignetting*. Including vignetting effects originating from the lens, the overall effect is equivalent to spatially attenuating the image by a mask $M(x)$, where x is the axis along which the mask is changing[3]. 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. Thus, the system acquires both dark areas and bright ones with high quality while extending the FOV.

The method can be applied to almost any imaginable variable filter, such as a stepped filter [7]. We used an off-the-shelf *linear variable neutral density filter* [7], in which the transmittance changes exponentially across its extent. The mask $M(x)$ in this case is approximately exponential as well. The order of magnitude by which the transmittance of this mask varies over the camera FOV is linear.

2.1 Self Calibration of the Effective Mask

It is possible to calibrate the mask "on the fly" from the sequence itself. Let I be the intensity of light that falls on the detector when the transmittance is maximum ($M = 1$). The intensity readout[4] at pixel (x, y) in frame k is

$$g_k(x,y) = I(x,y)M(x) \ . \tag{1}$$

We approximate the mask by the average horizontal profile

$$\hat{M}(x) \propto \sum_{k=1}^{\text{frames}} \sum_{y} g_k(x,y) \ . \tag{2}$$

In the limit of an infinite number of independent frames the estimate converges to the true mask. This simple estimation does not require prior registration of the images, and actually improves the registration, as discussed in Sec. 6.

Once the images are registered, the estimation is refined in the following way. Let a scene point be seen in frame k at image point x_k, with unsaturated intensity readout

$$g_k = IM(x_k) \ . \tag{3}$$

Then, this same scene point is measured without saturation in frame p at image pixel x_p, with intensity readout

$$g_p = IM(x_p) \ . \tag{4}$$

Assuming the scene radiance is constant between frames, these points should satisfy

$$M(x_k)g_p - M(x_p)g_k = 0 \ . \tag{5}$$

Tracking some of the scene points in several images provides many such linear equations which the mask should satisfy at each image pixel x. This set of equations can be written as $\mathbf{F}M = 0$. We also impose smoothness by penalizing $|\nabla^2 M|^2$, the Laplacian of M. The result is an overconstrained system of equations and its least squares solution is

$$\hat{M} = \arg\min_{M} (M^{t}\mathbf{A}^{t}\mathbf{A}M) \ . \tag{6}$$

[2]To eliminate reflections from the filters we mounted a dark hood between the filter and the camera.

[3]For simplicity we assumed filter variations along one spatial dimension. The results can be applied to 2D filter variations [7] as well.

[4]The readout g and the intensity I are unit-less since they are normalized by the irradiance equivalent to detector noise level.

Figure 3. Images taken with the linear variable *density* filter. 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.

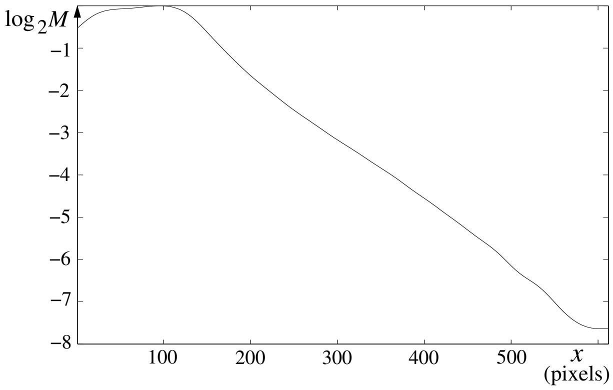

Figure 4. The self-calibrated log transmittance of the effective mask. The order of magnitude of the mask changes approximately linearly over most of the camera FOV. It extends the dynamic range of the camera by about 8 bits (factors of 2). The mask function drops on the left due to vignetting by system components other than the filter.

Here, $\mathbf{A} = [\mathbf{F}^{\mathrm{t}}\ \beta\mathbf{L}^{\mathrm{t}}]^{\mathrm{t}}$, $\mathbf{L}$ is the matrix form of an approximation to the Laplacian and β is a parameter that weights the penalty for unsmooth solutions relative to the penalty for disagreement with the data. Singular value decomposition yields the nontrivial solution up to a scale factor. The scale is set by letting $\max(\hat{M}) = 1$. These equations also enable the estimation of the covariance matrix of M, from which mask uncertainty estimates $\Delta\hat{M}(x)$ are derived [4].

The images shown in Fig. 3 are samples of a sequence taken through a commercial filter, rigidly attached to an 8-bit video camera (shown in Fig. 2). As the camera was rotated about its center of projection, almost all scene points were measured 14 times across the camera's FOV.

We registered the sequence of frames using the method described in Sec. 6. Then, we obtained about 50,000 equations as Eq. (5) based on randomly picked pairs of corresponding image points, to determine the mask. For stability, each image point used for this estimation was unsaturated and also non-dark. The log of the self-calibrated mask is shown in Fig. 4. This mask enables the extension of dynamic range by about 8 bits beyond the intrinsic dynamic range of the detector. Therefore, using an ordinary 8 bit camera we can obtain image mosaics with dynamic range close to that produced by a 16 bit camera.

2.2 Fusing the Measurements

We now describe the method we used to estimate the intensity at each mosaic point, given its multiple corresponding measurements. 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 $\hat{M}$ with uncertainty $\Delta\hat{M}$. Compensating the readout for the mask, the scene point intensity is $I_k = g_k/\hat{M}$ with uncertainty

$$\Delta I_k = \sqrt{\left(\frac{\partial I_k}{\partial g_k}\Delta g_k\right)^2 + \left(\frac{\partial I_k}{\partial \hat{M}}\Delta\hat{M}\right)^2} \ . \qquad (7)$$

We assumed 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 high uncertainty, thus its corresponding Δg_k is set to be a very large number[5].

If the measurements I_k are Gaussian and independent, the log-likelihood for a value I behaves like $-E^2$, where

$$E^2 \equiv \sum_k \left(\frac{I - I_k}{\Delta I_k}\right)^2 \ . \qquad (8)$$

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

$$\hat{I} = \widehat{\Delta I}^2 \sum_k \frac{I_k}{\Delta I_k^2} \ , \qquad (9)$$

where

$$\widehat{\Delta I} = \left(0.5 * \frac{\mathrm{d}^2 E^2}{\mathrm{d}I^2}\right)^{-1/2} = \left(\sqrt{\sum_k \frac{1}{\Delta I_k^2}}\right)^{-1} \ . \qquad (10)$$

At the boundaries of the frames that compose the mosaic there is a transition between points that have been estimated using somewhat different sources of data, causing seams to appear in the mosaic. We used feathering [28] to remove these seams, i.e., weighting each pixel according to its distance from the frame boundary. This weighting fits easily into our ML estimation; the uncertainty ΔI_k is multiplied by a factor that smoothly increases to ∞ towards the image boundaries. More details are given in [26].

[5] This estimate for ΔI_k assumes that the estimated mask is independent of the signal. Even if the mask is estimated using many points from the sequence, its dependence on a single intensity measurement is small.

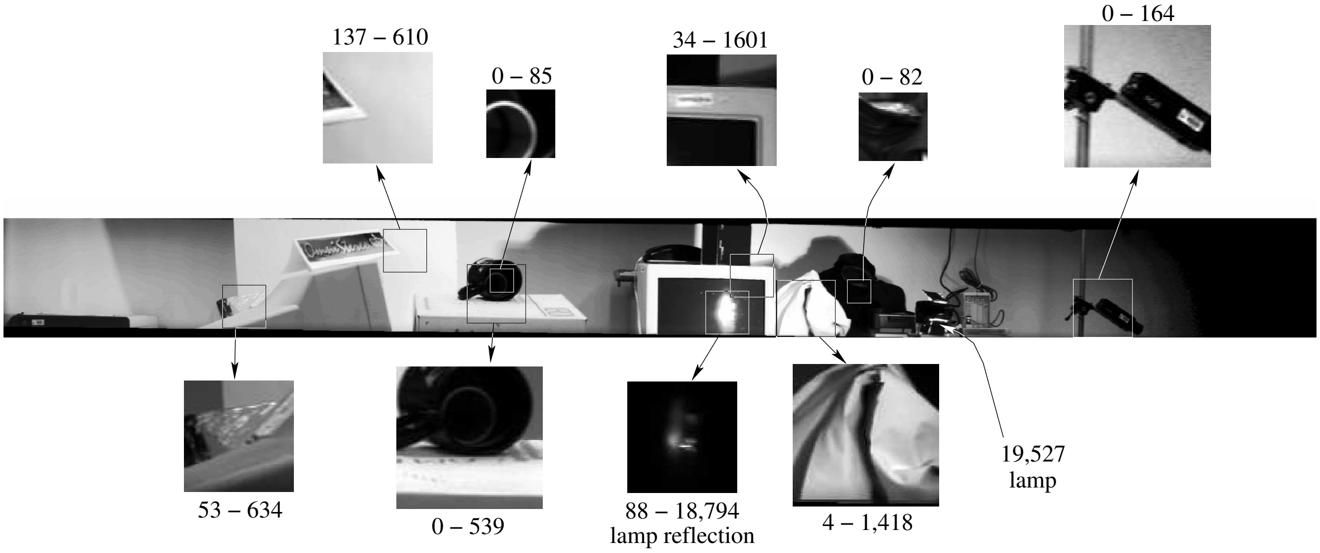

Figure 5. An image mosaic of 51^o horizontal FOV, created with a generalized mosaicing system having dynamic range of 16 bits. It is based on a single rotation about the center of projection of an 8 bit video camera. Contrast stretching in the selected squares reveals the details that reside within the computed mosaic. The numbers beside 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 images from the sequence, of which samples are shown in Fig. 3, were fused into a mosaic using this method. The histogram equalized version of $\log \hat{I}$ is shown in Fig. 5. Contrast stretching of $\hat{I}$ in selected regions shows that the mosaic is not saturated anywhere, and details are seen wherever $I \geq 1$. The HDR of this mosaic far exceeds that of the 8-bit camera we used, due to the dynamic range of the generalized mosaicing system.

3 Spectrum from Motion

Multispectral imaging has proven extremely useful in numerous imaging applications [9, 12, 17, 20, 29, 30, 34] including, object and material recognition, color analysis and constancy, remote sensing and medical imaging. However, the wide variety of methods used to capture multispectral images require the use of specialized and expensive hardware. In contrast, multispectral imaging nicely fits into the generalized mosaicing paradigm. We may use the same method that was used for HDR mosaicing. However, now the filter transmits a spectral band that varies across it. For example, if the left, central and right parts of the filter were red, green and blue, respectively, then RGB values could be obtained for each image pixel. This is in contrast to interpolated values based on subsampled measurements obtained using CCDs coated with RGB filters. Such a configuration yields three broad band samples of the spectrum. With generalized mosaicing it is possible to obtain much finer spectral information than RGB.

For enhanced spectral resolution, we mounted a commercial *linear variable interference (spectral) filter* [7] on a

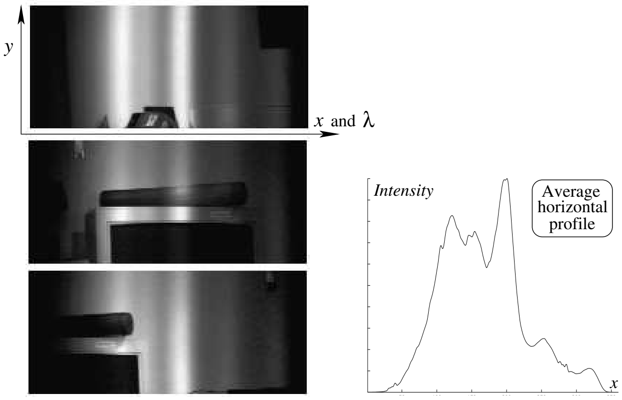

Figure 6. Images taken with the linear variable *interference* filter. The left of the FOV senses the energy density at 0.7μ, while the right senses it at 0.4μ. Besides the spatial features of the scene, its illuminant may be characterized easily by using a few frames. Here the *fluorescent illumination* is revealed by the typical spectral peaks [31], seen as bright columns appearing in the same places in all the images. The peaks appear also in the horizontal profile averaged over 5 frames in the sequence.

black/white video camera. The filter passes a narrow band around a central wavelength which changes linearly across the filter (horizontally). Its spectral range spans the visible light wavelengths[6]. Therefore, mosaicing images taken with such a filter enables multispectral measurements of each scene point. Fig. 6 shows black-and white images acquired through the filter.

[6]In other filters the spectral range may include the infra-red or UV.

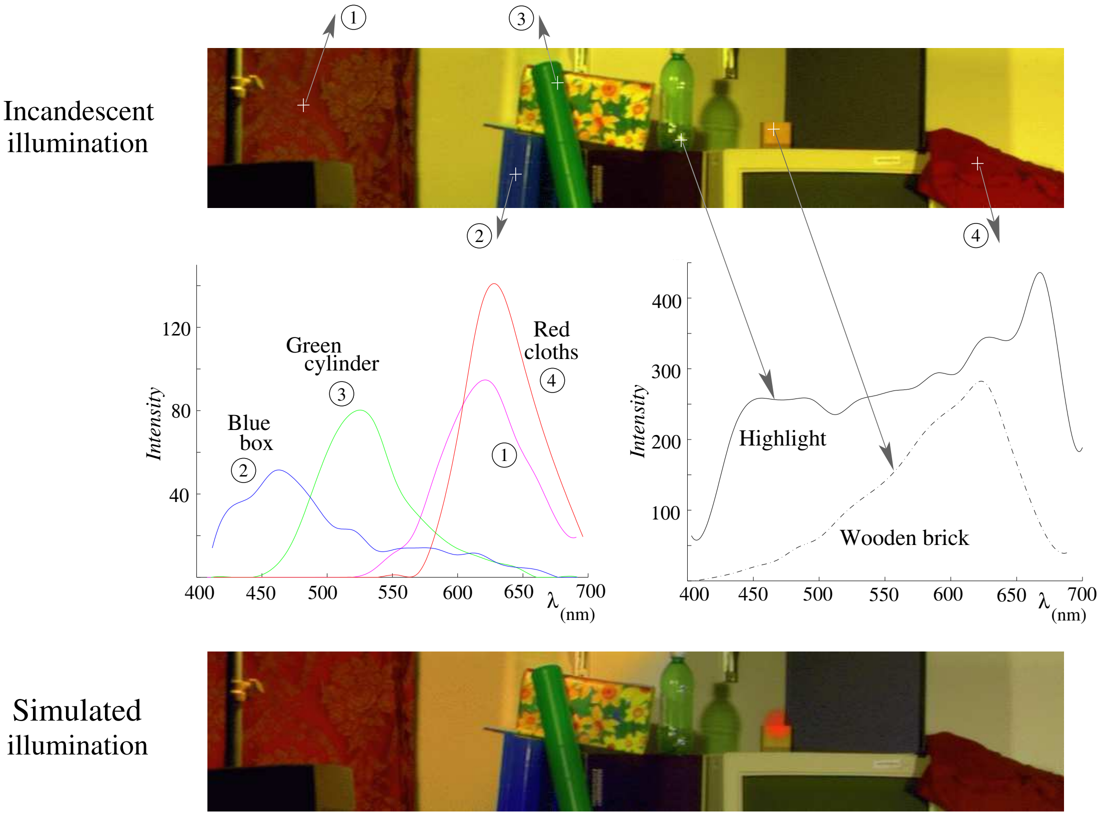

Figure 7. [Top] A color image mosaic rendered using the spectral data acquired at each point in its FOV, based on a single pass (rotation about the center of projection) of an ordinary black/white camera with a single fixed filter. The scene was illuminated by incandescent lamps. [Middle] The spectrum is plotted for selected points. [Bottom] Using known spectra, the scene was rendered *as if* its left part is illuminated by a setting sun [22], while its right part is illuminated by 'cool-white' fluorescent lamps [31]. In addition, a simulated red HeNe laser beam spot ($\lambda = 633nm$) is "shone" at the top edge of the wooden brick. **High resolution *color* images can be found at www.cs.columbia.edu/CAVE/ .**

In an experiment, the same black/white camera that was used to yield the HDR mosaic was used with this spatially variable spectral filter[7]. The grabbed images were compensate for vignetting effects in the camera which were computed before hand. We registered the images using the method discussed in Sec. 6. The registration yielded a wide FOV multispectral image mosaic, that represents the spectrum at each point. The multispectral mosaic was then converted to the RGB mosaic shown at the top of Fig. 7, via calculation of the CIE tristimulus values [10]. The yellowish appearance of the scene should not come as a surprise, because the illumination was by incandescent lamps (whose light is rather yellow). Using a simple interactive tool we created, one can get the spectrum at any point simply by clicking on the point with the computer's mouse[8].

As mentioned above, the availability of multispectral images already plays a very important role in many imaging applications. As a demonstration in the area of image rendering, we may render the scene under any illumination spectra. For example, given models of the spectra of fluorescent lamps [31], sunset [22] and a HeNe laser, we rendered the image shown on the bottom of Fig. 7. To do this, we first estimated the illumination spectrum of the acquired images, by measuring the spectrum of a patch in the scene (the monitor's label) that appeared to be white.

In another experiment, we created a multispectral mosaic of the Macbeth ColorChecker, for which the reflectances $R(\lambda_b)$ in narrow bands λ_b are known for each of the chart's 24 patches [10]. For this verification experiment we measured hundreds of wavelength samples per patch. We estimated the expected spectrum of each patch, based on the known reflectances and the chart's average spectrum. The plots of the expected spectrum and the measured one were in strong agreement for all the patches. The results for two patches are shown in Fig. 8. The correlation between the known reflectances and the measured intensities is very high: the correlation coefficient is ≈ 0.98 on average. We

[7]The characteristics of the interference filter somewhat depend on the angle at which light is incident on it. However, in the experiment the FOV angle of the camera was narrow enough to neglect this effect.

[8]The compensation of vignetting enabled measurements of intensities higher than 255 with the 8-bit camera, as seen in Fig. 7.

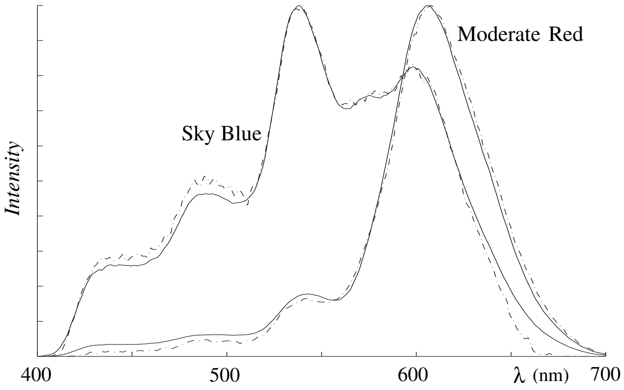

Figure 8. [Solid] Expected spectra (normalized) for two patches of the Macbeth ColorChecker, based on the chart's average spectra and known reflectances. Each patch spectrum is the multiplication of its reflectance by the estimated illumination spectra. [Dashed] The measured spectra.

note that it was significantly larger than the average correlation coefficient, 0.65, between randomly picked bands[9]. According to this test, our measurements are consistent with the known reflectances.

4 Illumination at a Glance

One of the aspects of generalized mosaicing is the ease of obtaining spectral information. This information exists in each raw frame along with the spatial features which appear in ordinary images. Consider the images shown in Fig. 6. The spatial details of the scene are clearly seen (e.g, the computer monitor). This is due to the fact that the system is an imaging device which captures an area of the scene. At the same time, the frames are dominated by vertical bands which appear at the same places in all the images.

To understand this, note that the illumination spectrum is similar in all the frames. Since the measured band changes horizontally across the frame, wavelengths at which the illumination energy is low manifest as dim vertical lines. Wavelengths at which the illumination is strong typically manifest as bright vertical lines (unless an object greatly absorbs these wavelengths). The bright bands in Fig. 6 are thus due to the characteristic spectral peaks of the fluorescent lamp which illuminated the scene.

Therefore, just a few frames may suffice to obtain cues about the illumination type and spectral distribution, while each frame supplies information about the 2D FOV. This is in contrast to many traditional methods of spectral imaging. Some of these methods capture, in each frame, the spectrum for a single 1D line of the 2D FOV [17, 30, 34]. These methods are based on prisms and gratings. There are other methods, such as those based on dichroic mirrors [34] and tunable filters, that capture the 2D FOV, but each frame is measured with a single band.

[9]This is the correlation between $R(\lambda_b)$ and the measured intensity $I(\lambda_q)$ for randomly picked bands b, q.

5 More Imaging Dimensions

It is possible to measure other imaging dimensions with the same system that was used to obtain HDR mosaics and multispectral mosaics. We are currently exploring estimation of the polarization of light, depth from defocus and a focused image of the entire scene (as in [15]) using different spatially varying filters.

Further, it may be possible to measure *multiple* scene properties *simultaneously*, using a "mega-filter" which can be formed by stacking individual filters, each of which modulates a single dimension (brightness, spectrum, polarization, etc.). Each dimension may be modulated spatially in different octaves. Alternatively, different portions of the camera FOV may correspond to different types of filtering.

6 Image Registration

A scene point has different coordinates in each image of the sequence. The measurements corresponding to this point should be identified before they can be fused. Image registration is based on the redundancy that exists between frames. However, the essence of generalized mosaicing is the removal of this redundancy in the raw images. Therefore, image registration becomes more challenging the more information we try to encode via filtering. For example, in multispectral mosaicing the scene may look quite different in the "red" wavelengths than in the "blue" wavelengths.

If the motion is small enough, as in most video streams, the readout changes between consecutive frames will be small and traditional registration methods are adequate. However, when images are taken with large displacements, the spatially varying but temporally static effects of the filter become significant. Consider the images shown in Fig. 3. Although features appear to be moving through the camera's FOV, the static mask clearly dominates the images. This would bias traditional algorithms towards estimating a motion slower than the true one. The mask varies gradually across the image, thus highpass filtering the raw images as in [14, 27] reduces the biasing effect. Nevertheless, it is not removed completely, from our experience. Also the images shown in Fig. 6 appear *as if* taken through a static intensity mask. Indeed, the bias occurred also in sequences of images captured with a spectral filter.

Measurements of scene points that become darker due to strong attenuation are relatively noisy. The same applies to scene points in the spectrally filtered images that are dim due to low illumination intensity in some spectral bands. Therefore, instead of matching the original image readouts, we use a transformed version of them that takes into account the attenuation-dependent uncertainties. We adapted some aspects of traditional techniques to enable the registration to cope with the spatially varying filtering effects.

Appendix A.1 describes the principles underlying our algorithm (details are given in [26]). The algorithm maximizes the likelihood of the matched data, and is robust to the biasing problem. We note that registering ordinary (not filtered) images by minimizing their sum of squared difference is obtained as a special case of this algorithm.

7 Conclusions

We have proposed generalized mosaicing as a framework for capturing information along multiple imaging dimensions (or dimensions of the plenoptic function [1]). This is achieved by acquiring a similar amount of data as in the case of traditional mosaicing. Thus far, we have used this framework to compute high dynamic range and multispectral mosaics. However, generalized mosaicing is not limited to these dimensions; it is a general concept that can be applied to other valuable dimensions (e.g., polarization, focus, depth). In addition, it permits the simultaneous enhancement of multiple dimensions. Generalized mosaicing can be applied in the presence of general motion, and other scenarios encountered in traditional mosaicing (e.g., lens distortions). More involved algorithms may be needed in such cases. However, as long as the spatially varying filtering removes part of the inter-frame redundancy, additional information can be extracted. In summary, generalized mosaicing has implications for several aspects of computer vision as well as other scientific disciplines such as astronomy and remote sensing. The simplicity of the technique suggests that it can also add significant value to digital photography.

Acknowledgments

This work was supported in parts by a National Science Foundation ITR Award, IIS-00-85864, a DARPA/ONR MURI Grant, N00014-95-1-0601, and the Morin Foundation.

A Appendix

A.1 Registration Algorithm

The following is a brief description of the aspects of our registration algorithm, that handle the spatially varying effects of the filter. Traditional registration by minimizing the sum of squared difference between frames is obtained as a special case.

- Each frame $g(x,y)$ is roughly flat fielded using $1/M(x)$, to yield an estimate of $I(x,y)$. This is based on an estimate of the mask M by Eq. (1). The intensity uncertainty ΔI is estimated using Eq. (7), given the readout uncertainty Δg and an estimated mask uncertainty ΔM (which may be initially set to zero). In case a spatially varying filter is not present, $M \equiv 1$, thus ΔI is constant.

- Let I_1 and I_2 be the intensity measurements at candidate corresponding pixels in two images, with respective uncertainties ΔI_1 and ΔI_2. As in Eq. (8), the squared distance between this pair of pixel measurements is

$$\hat{E}^2_{\text{pixel pair}} = [(\hat{I} - I_1)/\Delta I_1]^2 + [(\hat{I} - I_2)/\Delta I_2]^2 \ . \quad (11)$$

where $\hat{I}$ is given by Eqs. (9,10). The distance measure for the entire images is

$$\hat{E}^2_{\text{total}} = \sum_{\text{all pixels}} \hat{E}^2_{\text{each corresponding pair}} \ . \quad (12)$$

The best registration between two frames (or between a new frame and an existing mosaic) according to this objective function is the one that minimizes $\hat{E}^2_{\text{total}}$. If the measurements are Gaussian and independent, this is the *most likely match*. When the spatially varying filter is not present, $\Delta I_1 = \Delta I_2$, hence Eq. (12) is proportional to the sum of square difference between the images.

- The registration is done hierarchically, from coarse to fine resolution similar to [14, 25, 27]. We create a *Maximum Likelihood pyramid*, where not only the image value is stored at each scale, but also its uncertainty. The weights used in the construction of the pyramid structure depend also on the uncertainties of the pixels in each neighborhood, so that more reliable pixels contribute more to their coarse representation. If all the uncertainties are the same, the result is the same as in a traditional image pyramid. Details on this structure are given in [26]. The representation of $I_1, I_2, \Delta I_1, \Delta I_2$ at each scale enables efficient image registration by maximizing the likelihood of the match.

- To reduce the accumulation of matching errors, each new sequence frame is registered to the current mosaic [13, 25], and then fused into it (see Sec. 2.2).

- In HDR data, we prefer to penalize for relative errors rather than absolute ones. We thus calculated $s(x,y) = \log I(x,y)$ and $\Delta s = |ds/dI|\,\Delta I$. Then, we applied the above algorithm to the $s(x,y)$ images rather than the intensity images $I(x,y)$.

A.2 Sampling Criteria for Still Images

If a set of still images are acquired rather than a video stream, we cannot assume that frame displacements are small. The question then is what should the frame displacements be, or, how many times should each scene point be seen? In HDR mosaicing we required a transmittance change (due to motion) by at most a factor of 2 between consecutive frames, for each scene point. Each scene point is imaged at least $1 - \log_2(\min M)$ times, if $\max(M) = 1$. This increment yields, for each scene point, one measurement relatively close to saturation. This bright measurement has the highest quality since its representation contains a maximal number of significant bits.

In multispectral mosaicing, it is possible to minimize aliasing in the derived spectrum while using a relatively small number of images. This is due to the finite width of the transmitted wavelength bands. We have developed a criterion analogous to Nyquist sampling for the frame displacement. For rotations the angular change in the viewing direction between image acquisitions is $\approx D/2A$ radians, where A is the distance of the filter from the lens and D is the lens aperture diameter. Detailed proofs of these results as well as other sampling criteria are given in [26].

References

[1] E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision," in *Computational models of visual processing*, Eds: M. Landy and J. A. Movshon, pp. 3-20 (MIT Press, 1991)

[2] P. J. Burt and R. J. Kolczynski, "Enhanced image capture through fusion," Proc. ICCV, pp. 173-182 (1993).

[3] D. Capel and A. Zisserman, "Automated Mosaicing with super-resolution zoom," Proc. CVPR, pp. 885-891 (1998).

[4] S. C. Chapra and R. P. Canale, *Numerical methods for engineers with programming and software applications*, 3rd Ed., pp. 443,444,463-467 (WCB/McGraw-Hill, 1998).

[5] A. Davis, "Logarithmic pixel compression," *TechOnLine Review* **2**/3 (1998).

[6] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," Proc. SIGGRAPH pp. 369-378 (1997).

[7] Edmund Industrial Optics, *2001 Optics and optical instruments catalog*, pp. 71,85. Stock #'s L45-645, L41-960, L32-599, L32-700.

[8] H. Farid and E. P. Simoncelli, "Range estimation by optical differentiation," *JOSA A* **15**, pp. 1777-1786 (1998).

[9] N. Gat, U.S. Pat. 5166755 (1992).

[10] A. S. Glassner, *Principles of digital image synthesis*, Appendix G.4 (Morgan-Kaufmann, 1995).

[11] R. Gore "Ancient Ashkelon," *National Geographic* **199**/1 pp. 66-93 (2001).

[12] M. Hauta-Kasari, K. Miyazawa, S. Toyooka and J. Parkkinen, "A prototype of the spectral vision system," Proc. SCIA vol. 1, pp. 79-86 (1999).

[13] M. Irani, P. Anandan, J. Bergen, R. Kumar and S. Hsu, "Efficient representations of video sequences and their application," *Signal Processing: Image Communication* **8**, pp. 327-351 (1996).

[14] M. Irani and P. Anandan, "Robust multi-sensor image alignment," Proc. ICCV, pp. 959-966 (1998).

[15] A. Krishnan and N. Ahuja, "Panoramic image acquisition," Proc. CVPR, pp. 379-384 (1996).

[16] R. Kwok, J. C. Curlander, and S. Pang, "An automated system for mosaicking spaceborne SAR imagery," *Int. J. Remote Sensing* **11**, pp. 209-223 (1990).

[17] Y. Manabe, S. Kurosaka, O. Oshiro and K. Chihara, "Simultaneous measurement of spectral distribution and shape," Proc. ICPR, vol. 3, pp. 811-814 (2000).

[18] S. Mann, "Joint parameter estimation in both domain and range of functions in same orbit of the projective-Wyckoff group," Proc. ICIP, pp. 193-196 (1996).

[19] S. Mann and R. W. Picard, "On being 'Undigital' with digital cameras: extending dynamic range by combining differently exposed pictures," IS&T 48th Annual Conference, pp. 422-428 (1995).

[20] A. M. Mika, "Linear-wedge spectrometer," Proc. SPIE **1298**, *Imaging Spectroscopy of the Terrestrial Environment*, pp. 127-131 (1990).

[21] T. Mitsunaga and S. K. Nayar, "Radiometric self calibration," Proc. CVPR Vol. I, pp. 374-380 (1999).

[22] M. R. Nagel, H. Quenzel, W.Kweta and R. Wendling, *Daylight illumination - color-contrast tables,* pp. 78-87 (Academic Press, 1978).

[23] S. K. Nayar and T. Mitsunaga "High dynamic range imaging: Spatially varying pixel exposures," Proc. CVPR, Vol. I, pp. 472-479 (2000).

[24] S. Peleg, M. Ben-Ezra and Y. Pritch "Omnistereo: Panoramic stereo imaging," *IEEE Trans. PAMI* **23**, pp. 279-290 (2001).

[25] H. S. Sawhney, R, Kumar, G. Gendel, J. Bergen, D. Dixon, V. Paragano, "$\mathrm{VideoBrush}^{TM}$: Experiences with consumer video mosaicing," Proc. WACV, pp. 52-62 (1998).

[26] Y. Y. Schechner and S. K. Nayar, "Generalized Mosaicing: High dynamic range in a wide field of view," Submitted to *IJCV*. And "Generalized Mosaicing: Wide field of view multispectral imaging," Submitted to *IEEE PAMI* (2001).

[27] R. K. Sharma and M. Pavel, "Multisensor image registration," Society for Information Display, Vol. XXVIII, pp. 951-954 (1997).

[28] H. Y. Shum and R. Szeliski, "Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment, " *IJCV* **36**, pp. 101-130 (2000).

[29] D. Slater and G. Healey, "Material classification for 3D objects in aerial hyperspectral images," Proc. CVPR vol. 2, pp. 268-273 (1999).

[30] H. M. G. Stokman, T. Gevers and J. J. Koenderink, "Color Measurement by Imaging Spectrometry," *Comp. Vis. Imag. Under.* **79**/2 pp. 236-249 (2000).

[31] Osram Sylvania Corp., "Fluorescent lamps," *Engineering Bulletin* 0-341 (Danvers, MA).

[32] J. M. Uson, S. P. Boughn, J. R. Kuhn, "The central galaxy in Abell 2029: an old supergiant," *Science* **250**, pp. 539-540 (1990).

[33] 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* **135**, pp. 265-275 (1998).

[34] J. B. Wellman, "Multispectral mapper: imaging spectroscopy as applied to the mapping of earth resources," Proc. SPIE **268** *Imaging Spectroscopy*, pp. 64-73 (1981).