



Programmable Imaging: Towards a Flexible Camera

SHREE K. NAYAR AND VLAD BRANZOI

Department of Computer Science, 450 Mudd Hall, Columbia University, New York, N.Y. 10027

nayar@cs.columbia.edu

vlad@cs.columbia.edu

TERRY E. BOULT

Department of Computer Science, University of Colorado, Colorado Springs, Colorado 80933

tboult@cs.uccs.edu

Received February 23, 2005; Revised June 8, 2005; Accepted June 21, 2005

First online version published in June, 2006

Abstract. In this paper, we introduce the notion of a programmable imaging system. Such an imaging system provides a human user or a vision system significant control over the radiometric and geometric characteristics of the system. This flexibility is achieved using a programmable array of micro-mirrors. The orientations of the mirrors of the array can be controlled with high precision over space and time. This enables the system to select and modulate rays from the scene's light field based on the needs of the application at hand.

We have implemented a programmable imaging system that uses a digital micro-mirror device (DMD), which is used in digital light processing. Although the mirrors of this device can only be positioned in one of two states, we show that our system can be used to implement a wide variety of imaging functions, including, high dynamic range imaging, feature detection, and object recognition. We also describe how a micro-mirror array that allows full control over the orientations of its mirrors can be used to instantly change the field of view and resolution characteristics of the imaging system. We conclude with a discussion on the implications of programmable imaging for computer vision.

Keywords: programmable imaging, flexible imaging, micro-mirror array, digital micro-mirror device, MEMS, adaptive optics, high dynamic range imaging, optical processing, feature detection, object recognition, field of view, resolution, multi-viewpoint imaging, stereo, catadioptric imaging, wide-angle imaging, purposive camera.

1. A Flexible Approach to Imaging

In the past few decades, a wide variety of novel imaging systems have been proposed that have fundamentally changed the notion of a camera. These include high dynamic range, multispectral, omnidirectional, and multi-viewpoint imaging systems. The hardware and software of each of these devices are designed to accomplish a particular imaging function. This function cannot be altered without significant redesign of

the system. It would clearly be beneficial to have a single imaging system whose functionality can be varied using software without making any hardware alterations.

In this paper, we introduce the notion of a programmable imaging system. Such a system gives a human user or a computer vision system significant control over the radiometric and geometric properties of the system. This flexibility is achieved by using a programmable array of micro-mirrors. The orientations of

the mirrors of the array can be controlled with very high speed. This enables the system to select and modulate scene rays based on the needs of the application at hand.

The basic principle behind the proposed approach is illustrated in Fig. 1. The system observes the scene via a two-dimensional array of micro-mirrors, whose orientations can be controlled. The surface normal \mathbf{n}_i of the i^{th} mirror determines the direction of the scene ray it reflects into the imaging system. If the normals of the mirrors can be arbitrarily chosen, each mirror can be programmed to select from a continuous cone of scene rays. As a result, the field of view and resolution characteristics of the imaging system can be chosen from a very wide space of possibilities. Moreover, since the mirror orientations can be changed instantly, the field of view and resolution characteristics can be varied from one state to another without any delay. In short, we have an imaging system whose geometric properties are instantly controllable.

Now let us assume that each mirror can also be oriented with normal \mathbf{n}_b such that it reflects a black surface (with zero radiance). Let the integration time of the image detector be T . If the mirror is made to point in the directions \mathbf{n}_i and \mathbf{n}_b for durations t and $T - t$, respectively, the scene ray is attenuated by t/T . As a result, each imaged scene ray can also be modulated with high precision. This allows us to control the brightness modulation of each pixel with high precision. In other words, the radiometric properties of the imaging system are also controllable. As we shall show, the ability to instantly change the modulation of the light received by each pixel also enables us to perform simple but useful computations on scene radiance values prior to image capture.

Since the micro-mirror array is programmable, the above geometric and radiometric manipulations can all be done using software. The end result is a single imaging system that can emulate the functionalities of several existing specialized systems as well as new ones. Such a flexible camera has two major benefits. First, a user is free to change the role of the camera based on his/her need. Second, it allows us to explore the notion of a purposive camera that can, as time progresses, always produce the visual information that is most pertinent to the task.

We have developed a prototype of the programmable imaging system that uses a commercially available digital micro-mirror device. The mirror elements of this device can only be positioned in one of two ori-

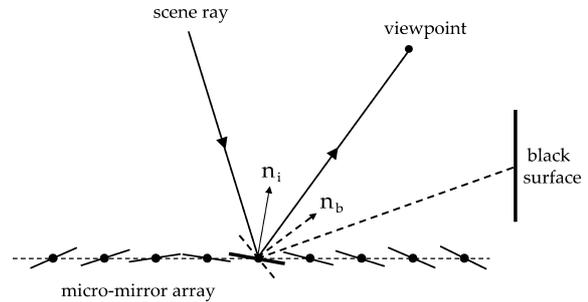


Figure 1. The principle underlying programmable imaging using a micro-mirror array. If the orientations of the individual mirrors can be controlled with high precision and speed, scene rays can be selected and modulated in a variety of ways, each leading to a different imaging system. The end result is a single imaging system that can perform the functions of a wide range of specialized cameras.

entations. Using this system, we demonstrate several functions including high dynamic range imaging, optical feature detection, and object recognition using appearance matching. We also show how this micro-mirror device can be used to instantly rotate the field of view of the imaging system. We believe that in the future micro-mirror devices may become available that provide greater control over mirror orientations. We show how such a device would enable us to change the field of view, emulate camera rotation, and create multiple views for stereo. We conclude with a discussion on the implications of programmable imaging for computer vision.

2. Imaging with a Micromirror Device

Ideally, we would like to have full control over the orientations of our micro-mirrors. Such devices are being developed for adaptive optical processing in astronomy Tyson (1998). However, at this point in time, they do not have the physical properties and programmability that we need for our purpose. To implement our ideas, we use the digital micro-mirror device (DMD) that was introduced in the 1980s by Hornbeck at Texas Instruments (Hornbeck, 1998; Hornbeck, 1989). The DMD is a micro-electromechanical system (MEMS) that has evolved rapidly over the last decade and has found many applications (Dudley et al., 2003). It is the key enabling technology in many of today's projection systems, Hornbeck (1995). The latest generation of DMDs have more than a million mirrors, each mirror roughly 14×14 microns in size (see Fig. 2). From our perspective, the main limitation of current DMDs is

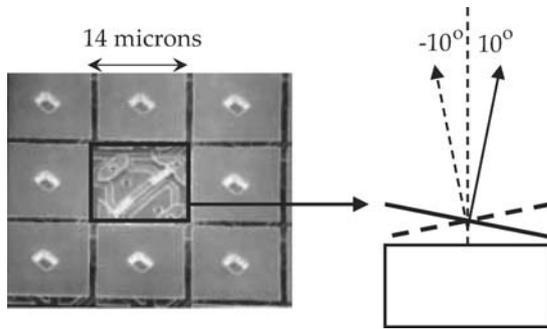


Figure 2. Our implementation of programmable imaging uses a digital micro-mirror device (DMD). The most recent DMDs have more than a million mirrors, each mirror roughly 14×14 microns in size. The mirrors can be oriented with high precision and speed at $+10$ or -10 degrees.

that the mirrors can be oriented in only two directions: -10° or $+10^\circ$ about one of the mirror's diagonal axes (see Fig. 2). However, the orientation of each mirror can be switched from one state to the other in a few *microseconds*, enabling modulation of incident light with very high precision.

Figure 3 shows the optical layout of the system we have developed using the DMD. The scene is first projected onto the DMD plane using an imaging lens. This means that the cone of light from each scene point received by the aperture of the imaging lens is focused onto a single micro-mirror. When all the mirrors are oriented at $+10^\circ$, the light cones are reflected in the

direction of a re-imaging lens which focuses the image received by the DMD onto a CCD image detector. Note that the DMD in this case behaves like a planar scene that is tilted by 20° with respect to the optical axis of the re-imaging lens. To produce a focused image of this tilted set of source points, one needs to tilt the image detector according to the well-known Scheimpflug condition (Smith, 1966).

3. Prototype System

It is only recently that developer kits have begun to appear that enable one to use DMDs in different applications. When we began implementing our system this option was not available. Hence, we chose to re-engineer an off-the-shelf DMD projector into an imaging system by reversing the path of light; the projector lens is used to form an image of the scene on the DMD rather than illuminate the scene via the DMD. Fig. 4(a) shows a partly disassembled Infocus LP 400 projector. This projector uses one of the early versions of the DMD with 800×600 mirrors, each 17×17 microns in size. The modulation function of the DMD is controlled by simply applying an 8-bit image (VGA signal) to the projector input. We had to make significant hardware changes to the projector. First, the projector lamp had to be blocked out of the optical path. Then, the chassis was modified so that the re-imaging lens and the camera could be attached to the system. Finally, a 2

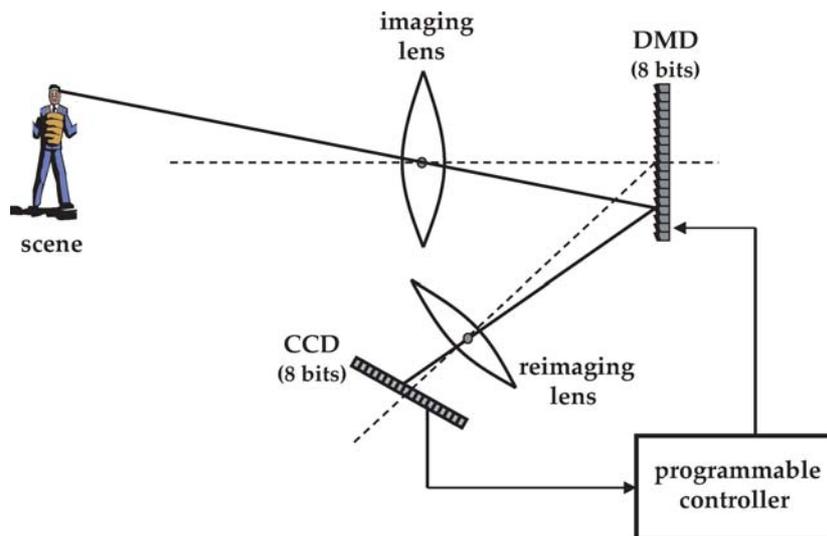
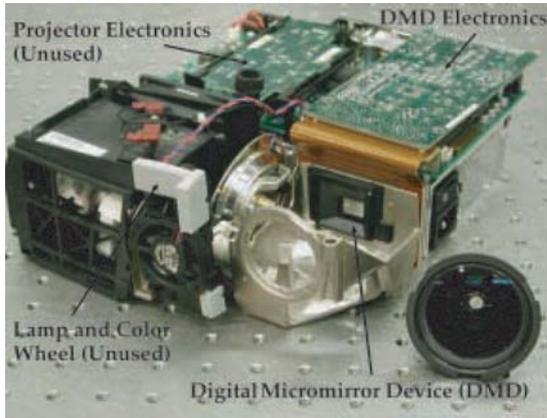


Figure 3. Imaging using a DMD. The scene image is focused onto the DMD plane. The image reflected by the DMD is re-imaged onto a CCD. The programmable controller captures CCD images and outputs DMD (modulation) images.



(a) Disassembled projector.



(b) Programmable imaging system.

Figure 4. (a) A disassembled Infocus LP 400 projector that shows the exposed DMD. (b) In this re-engineered system, the projector lens is used as an imaging lens that focuses the scene on the DMD. The image reflected by the DMD is re-imaged by a CCD camera. Camera images are processed and DMD modulation images are generated using a PC.

degree-of-freedom manual stage was used to orient the detector behind the re-imaging lens so as to satisfy the Scheimpflug condition.

The final system is shown in Fig. 4(b). It is important to note that this system is bulky only because we are using the electronics of the projector to drive the DMD. If instead we built the entire system from scratch, it would be very compact and not much larger than an off-the-shelf camera. The CCD camera used in the system is an 8-bit monochrome Sony XC-75 model with 640×480 pixels. The processing of the camera image and the control of DMD image is done using a Dell workstation with a 2.5 GHz Pentium 4 processor.

DMDs have previously been used in imaging applications, but for very specific tasks such as recording celestial objects in astronomy. For instance, in Malbet et al. (1995) the DMD is used to mask out bright sources of light (like the sun) so that dimmer regions (the corona of the sun) can be imaged with higher dynamic range. In Kearney and Ninkov (1998), a DMD is used to mask out everything except a small number of stars. Light from these unmasked stars are directed towards a spectroscope to measure the spectral characteristics of the stars. In Christensen et al. (2002) and Castracane and Gutin (1999), the DMD is used to modulate brightness values at a pixel level for high dynamic range multispectral imaging and removal of image blooming, respectively. These works address rather specific imaging needs. In contrast, we are interested in a flexible imaging system that can perform a wide range of functions.

An extensive off-line calibration of the geometric and radiometric properties of the system was conducted. The geometric calibration involves determining the mapping between DMD and CCD pixels. This mapping is critical to controlling the DMD and interpreting the images captured by the CCD. The geometric calibration was done by using a bright scene with more or less uniform brightness. Then, a large number of square patches were used as input to the DMD and recorded using the CCD, as shown in Fig. 5. Note that a dark patch in the DMD image produces a dark patch in the CCD image and a white patch in the DMD image produces a patch of the bright scene in the CCD image. In order to scan the entire set of patches efficiently, binary coding of the patches was done. The centroids of corresponding patches in the DMD and CCD images were fitted to a piecewise, first-order polynomial. The computed mapping was found to have an RMS error of 0.6 (CCD) pixels. This computed mapping as well as its inverse were resampled and stored as two look-up tables; one that maps CCD pixels to DMD pixels and another that maps DMD pixels to CCD pixels. Both of these are two-dimensional tables with two entries at each location and hence efficient to store and use.

The radiometric calibration was done in two parts. First, the CCD camera was calibrated using a Macbeth reflectance chart to obtain a one-dimensional look-up table that maps image brightness to scene radiance. This was done prior to mounting the camera on the system. Once the camera was attached to the system, the radiometric response of the DMD modulation system (including the DMD chip, the projector

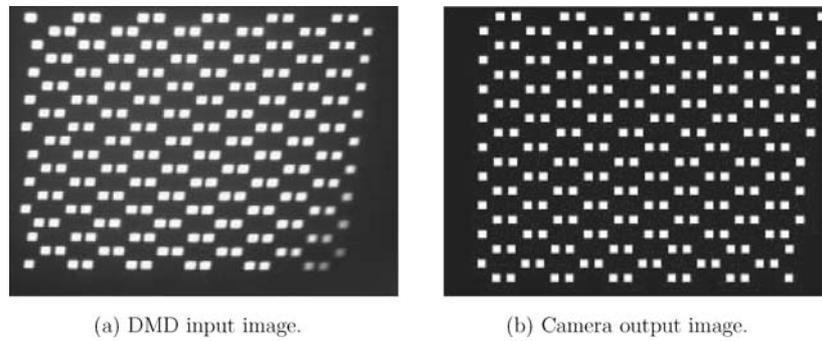


Figure 5. Geometric calibration of the imaging system. The geometric mapping between the DMD and the CCD images is determined by showing the system a scene of uniform brightness, applying patterned images to the DMD, and capturing the corresponding CCD images. The centroids of corresponding patches in the DMD and CCD images are then used to compute the forward and inverse transformations between the DMD and CCD planes.

electronics and the PC's video card) was estimated using a scene with uniform brightness. A large number of uniform modulation images of different brightnesses were applied to the DMD and the corresponding (linearized) camera images were captured. A few of the camera pixels (chosen around the center of the camera image) were used to compute a function that relates DMD input and camera output. This function was again stored as a one-dimensional look-up

table. One of the camera images was then used to compute the spatial brightness fall-off function (which includes vignetting and other effects) of the complete system.

Figure 6 shows two simple examples that illustrate the modulation of scene images using the DMD. One can see that after modulation some of the scene regions that were previously saturated produce useful brightness values. Note that the captured CCD

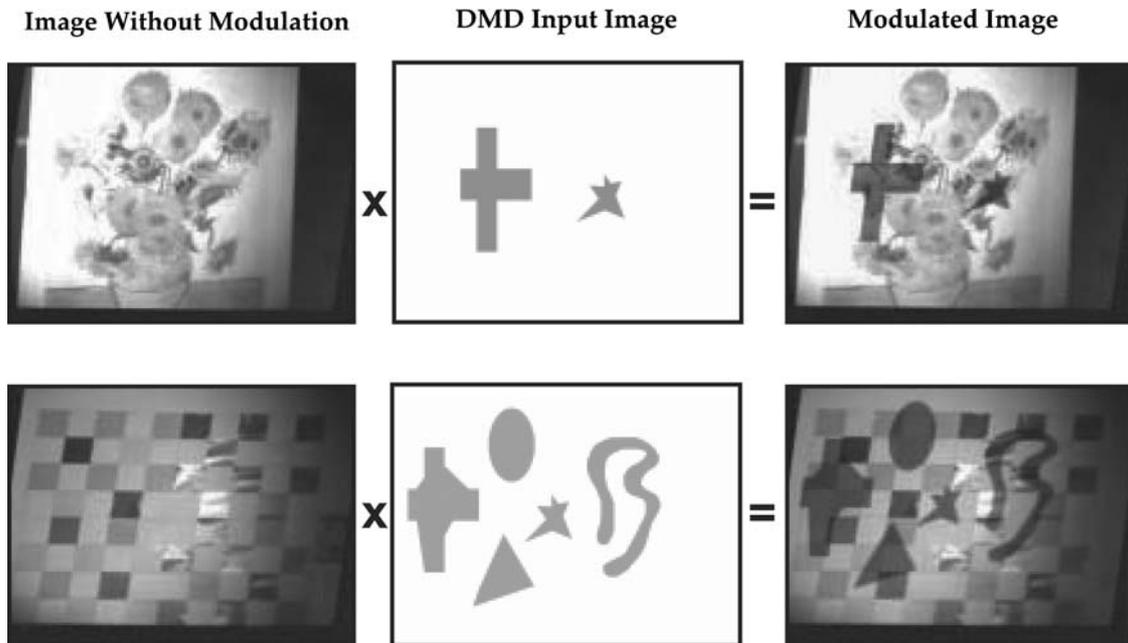


Figure 6. Examples that show how image irradiance is modulated with high resolution using the DMD.

image is skewed with respect to the DMD modulation image. This skewing is due to the required tilt of the CCD discussed above and is corrected using the calibration results. In our system, the modulation image can be controlled with 8 bits of precision and the captured CCD images have 8 bits of accuracy.

4. High Dynamic Range Imaging

The ability to program the modulation of the image at a pixel level provides us with a flexible means to implement several previously proposed methods for enhancing the dynamic range. In this section, we will describe three different implementations of high dynamic range imaging.

4.1. Temporal Exposure Variation

We begin with the simplest implementation, where the global exposure of the scene is varied as a function of time. In this case, the control image applied to the DMD is spatially constant but changes periodically with time. An example of a video sequence acquired in this manner is shown in Fig. 7, where 4 modulation levels are

cycled over time. It has been shown in previous work that an image sequence acquired in this manner can be used to compute high dynamic range video when the motions of scene points between subsequent frames is small (Ginosar et al., 1992; Kang et al., 2003). Alternatively, the captured video can be subsampled in time to produce multiple video streams with lower frame-rate, each with a different fixed exposure. Such data can improve the robustness of tasks such as face recognition, where a face missed at one exposure may be better visible and hence detected at another exposure.

Videos of the type shown in Fig. 7 can also be obtained by changing the integration time of the detector or the gain of the camera. Due to the various forms of camera noise, changing integration time or gain compromises the quality of the acquired data. In our case, since the DMD can be controlled with 8 bits of accuracy and the CCD camera produces 8-bit images, the captured sequence can be controlled with 16 bits of precision. However, it must be noted that this is not equivalent to using a 16-bit detector. The additional 8 bits of control provided by the DMD allows us to use 256 different exposure settings. The end result is 16 bits of control over the measure irradiance but the quantization levels are not uniformly spaced as in the case of a 16-bit detector.

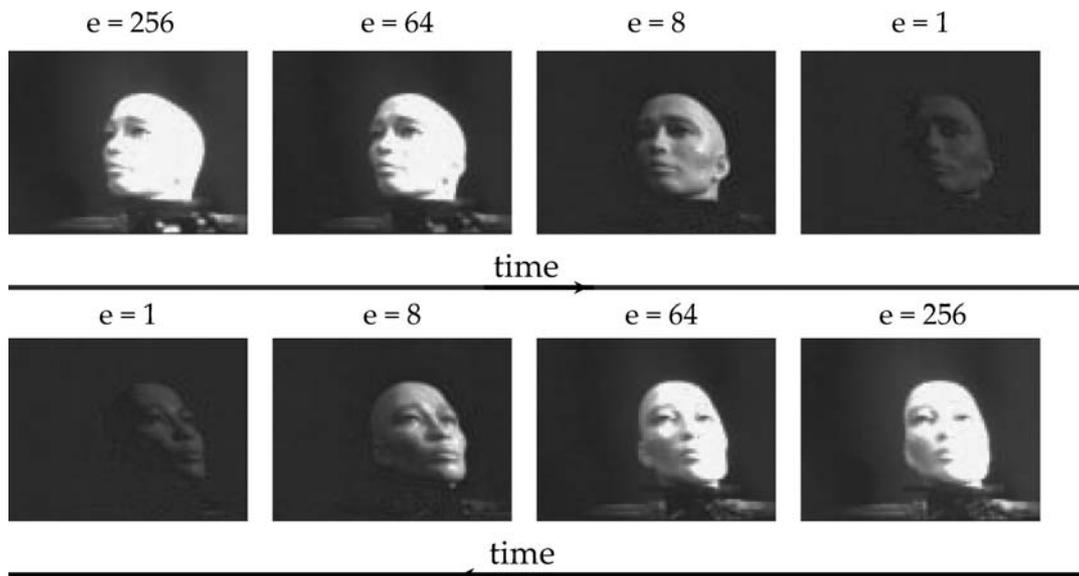


Figure 7. Spatially uniform but temporally varying DMD inputs can be used to generate a video with varying exposure (e). Using a DMD in this case produces high quality data compared to changing the exposure time or the camera gain.

4.2. Spatio-Temporal Exposure Variation

In Nayar and Mitsunaga (2000), the concept of spatially varying pixel exposures was proposed where an image is acquired with a detector with a mosaic of neutral density filters. The captured image can be reconstructed to obtain a high dynamic range image with a slight loss in spatial resolution. Our programmable system allows us to capture an image with spatially varying exposures by simply applying a fixed (checkerboard-like) pattern to the DMD. In Nayar and Narasimhan (2002), it was shown that a variety of exposure patterns can be used, each trading off dynamic range and spatial resolution in different ways. Such trade-offs are easy to explore using our system.

It turns out that spatially varying exposures can also be used to generate video streams that have higher dynamic range for a human observer, without post-processing each acquired image as was done in Nayar and Mitsunaga (2000). If one uses a fixed pattern, the pattern will produce a very visible modulation that would be distracting to the observer. However, if the pattern is varied with time, the eye becomes less sensitive to the pattern and a video with a larger range of brightnesses is perceived by the observer. Fig. 8(a) shows the image of a scene taken without modulation. It is clear that the scene has a wide dynamic range and an 8-bit camera cannot capture this range. Fig. 8(b) shows four consecutive frames captured with spatially varying exposures. The exposure pattern uses 4 different exposures (e_1, e_2, e_3, e_4) within each 2×2 neighborhood of pixels. The relative positions of the 4 exposures are changed over time using a cyclic permutation. In the images shown in Fig. 8(b), one sees the spatial patterns introduced by the exposures (see insets). However, when this sequence is viewed at 30 Hz, the pattern is more or less invisible (the eye integrates over the changes) and a wider range of brightnesses are visible.

4.3. Adaptive Dynamic Range

Recently, the method of adaptive dynamic range was introduced in Nayar and Branzoi (2003), where the exposure of each pixel is controlled based on the scene radiance measured at the pixel. A prototype device was implemented using an LCD attenuator attached to the front of the imaging lens of a camera. This implementation suffers from three limitations. First, since the LCD attenuator uses polarization filters, it allows only 50%

of the light from the scene to enter the imaging system even when the attenuation is set to zero. Second, the attenuation function is optically defocused by the imaging system and hence pixel-level attenuation could not be achieved. Finally, the LCD attenuator cells produce diffraction effects that cause the captured images to be slightly blurred.

The DMD-based system enables us to implement adaptive dynamic range imaging without any of the above limitations. Since the image of the scene is first focused on the DMD and then re-imaged onto the image detector, we are able to achieve pixel-level control. In addition, the fill-factor of the DMD is very high compared to an LCD array and hence the optical efficiency of the modulation is closer to 90%. Because of the high fill-factor, the blurring/diffraction effects are minimal.

In Christensen et al. (2002) and Castracane and Gutin (1999), a DMD has been used to implement adaptive dynamic range. However, these previous works do not adequately address the real-time spatio-temporal control issues that arise in the case of dynamic scenes. We have implemented a control algorithm very similar to the one in Nayar and Branzoi (2003) for computing the DMD modulation function based on each captured image. Results from this system are shown in Fig. 9. The first row shows a person under harsh lighting imaged without modulation (conventional camera). The second row shows the output of the programmable system and the third row shows the corresponding modulation (attenuation) images applied to the DMD. As described in Nayar and Branzoi (2003), the output and modulation images can be used to compute a video stream that has an effective dynamic range of 16 bits, although without uniform quantization.

5. Intra-Pixel Optical Feature Detection

The field of optical computing has developed very efficient and powerful ways to apply image processing algorithms (such as convolution and correlation) Goodman (1968). A major disadvantage of optical computing is that it requires the use of *coherent* light to represent the images. This has proven cumbersome, bulky, and expensive. It turns out that programmable modulation can be used to implement a limited class of image processing tasks directly to the *incoherent* optical image captured by the imaging lens, without the use of coherent sources. In particular, one can apply convolution at an *intra-pixel* level very efficiently. By

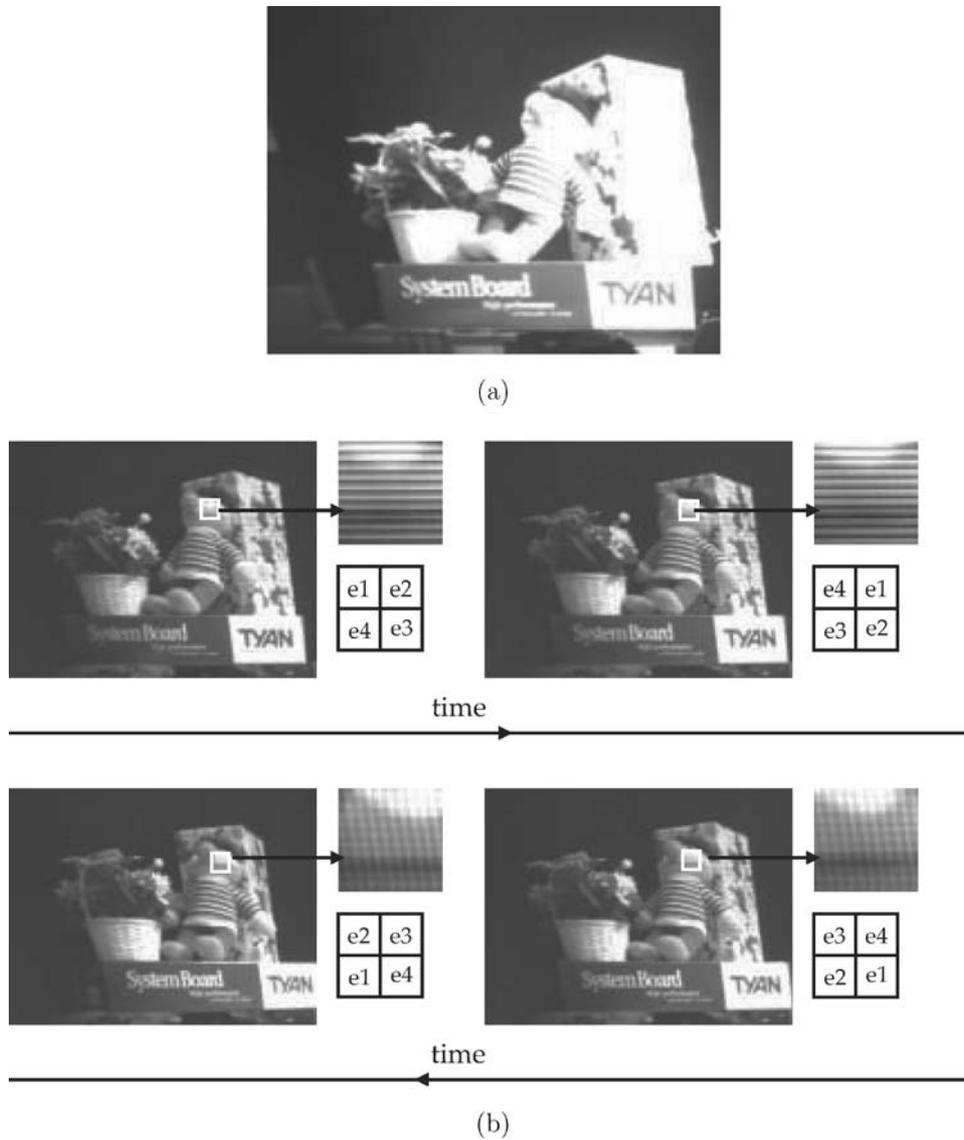


Figure 8. (a) A scene with a wide range of brightnesses captured using an 8-bit (low dynamic range) camera. (b) Four frames of the same scene (with moving objects) captured with spatio-temporal exposure modulation using the DMD. When such a video is viewed at frame-rate, the observer perceives a wider dynamic range without noticing the exposure changes.

intra-pixel we mean that the convolution mask is being applied to the distribution of light energy within a single pixel rather than a neighborhood of pixels. Intra-pixel optical processing leads to very efficient algorithms for finding features such as edges, lines, and corners.

Consider the convolution $f * g$ of a continuous optical image f with a kernel g whose span (width) is less than, or equal to, a pixel on the image detector. We can rewrite the convolution as $f * (g^+ - g^-)$ where g^+

is made up of only the positive elements of g and g^- has only the absolute of the negative elements of g . We use this decomposition since incoherent light cannot be negatively modulated (the modulation image cannot have negative values). An example of such a decomposition for the case of a first-derivative operator is shown in Fig. 10(a). As shown in the figure, let each CCD pixel correspond to 3×3 DMD pixels; i.e. the DMD has three times the linear resolution of the CCD.

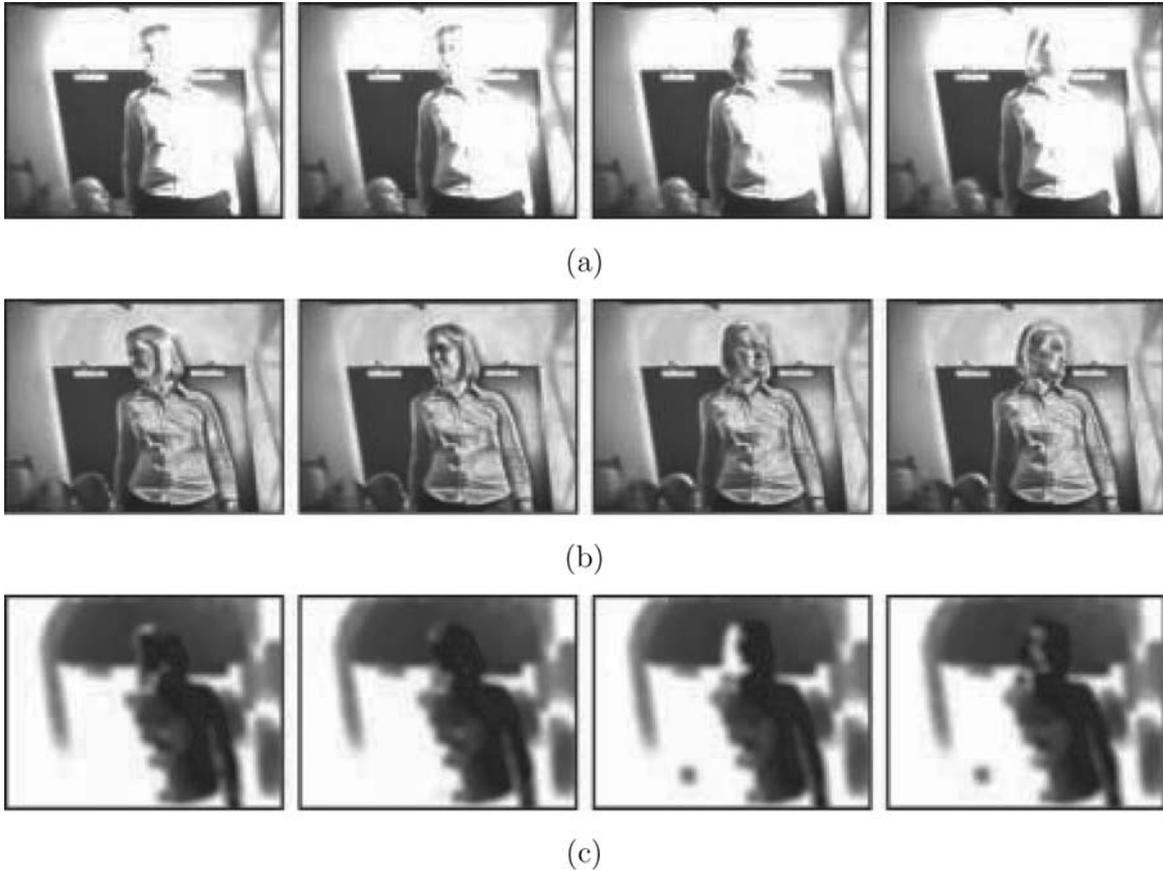


Figure 9. (a) Video of a person taken under harsh lighting using a conventional (8-bit) camera. (b) The raw output of the programmable system when the DMD is used to achieve adaptive dynamic range. (c) The modulation images applied to the DMD. The raw camera output and the DMD modulation can be used to compute a video with very high dynamic range.

Then, the two components of the convolution (due to g^+ and g^-) are directly obtained by capturing two images with the modulation images shown in Fig. 10(b). The difference between these images gives the final result ($f * g$).

Figure 10(c) shows the four *optically processed* images of a scene obtained for the case of the Sobel edge operator. The computed edge map is shown in Fig. 10(d). Since our DMD has only 800×600 elements, the edge map is of lower resolution with about 200×150 pixels. Although four images are needed in this case, it can be applied to a scene with slowly moving objects where each new image is only used to update one of the four component filter outputs in the edge computation. Note that all the multiplications involved in the convolutions are done in optical domain (at the speed of light).

6. Optical Appearance Matching

In the past decade, appearance matching using subspace methods has become a popular approach to object recognition (Turk and Pentland, 1991; Murase and Nayar, 1995). Most of these algorithms are based on projecting input images to a precomputed linear subspace and then finding the closest database point that lies in the subspace. The projection of an input image requires finding its dot product with a number of vectors. In the case of principal component analysis, the vectors are the eigenvectors of a correlation or covariance matrix computed using images in the training set.

It turns out that optical modulation can be used to perform all the required multiplications in optical domain, leaving only additions to be done computa-

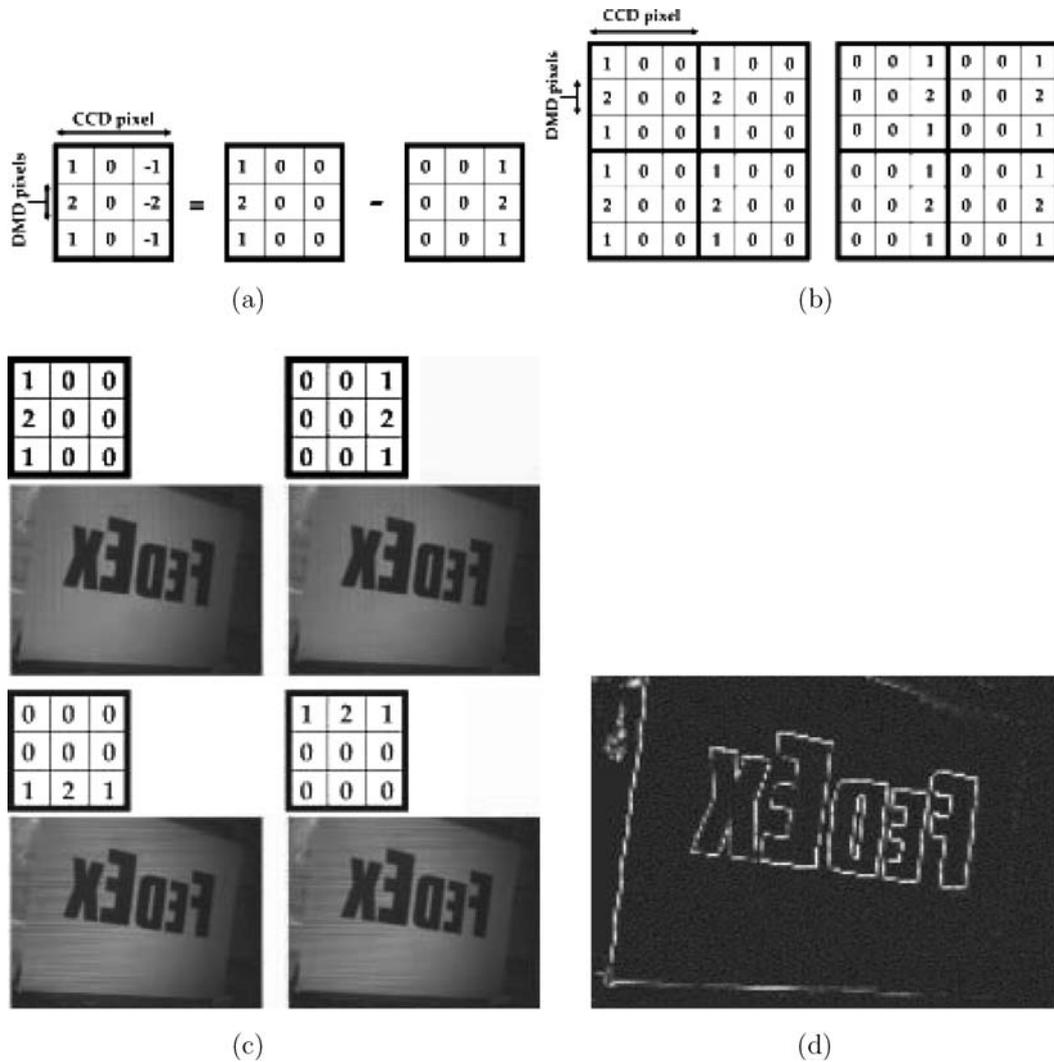


Figure 10. (a) Decomposition of a convolution kernel into two positive component kernels. (b) When the resolution of the DMD is higher than that of the CCD, intra-pixel convolution is done by using just two modulation images and subtracting the resulting CCD images. (c) Four images that result from applying the four component kernels of a Sobel edge operator. (d) The edge map computed from the four images in (c).

tionally. Let the input image be m and the eigenvectors of the subspace be e_1, e_2, \dots, e_k . The eigenvectors are concatenated to obtain a larger (tiled) vector $B = [e_1, e_2, \dots, e_k]$ and k copies of the input image are concatenated to obtain the (tiled) vector $A = [m, m, \dots, m]$. If the vector A is “shown” as the scene to our imaging system and the vector B is used as the modulation image, the image captured by the camera is a vector $C = A .* B$, where $.*$ denotes an element-by-element product of the two vectors. Then, the image C is raster scanned to sum up its k tiles to obtain the k coefficients that correspond to the subspace

projection of the input image. This coefficient vector is compared with stored vectors and the closest match reveals the identity of the object in the image.

We have used our system to implement this idea and develop a real-time face recognition system. Fig. 11(a) shows the 6 people in our database; 30 poses (images) of each person were captured to obtain a total of 180 training images. PCA was applied and the 6 most prominent eigenvectors are tiled as shown in Fig. 11(b) and used as the DMD modulation image. During recognition, the output of the video camera is also tiled in the same way as the eigenvectors and displayed on a

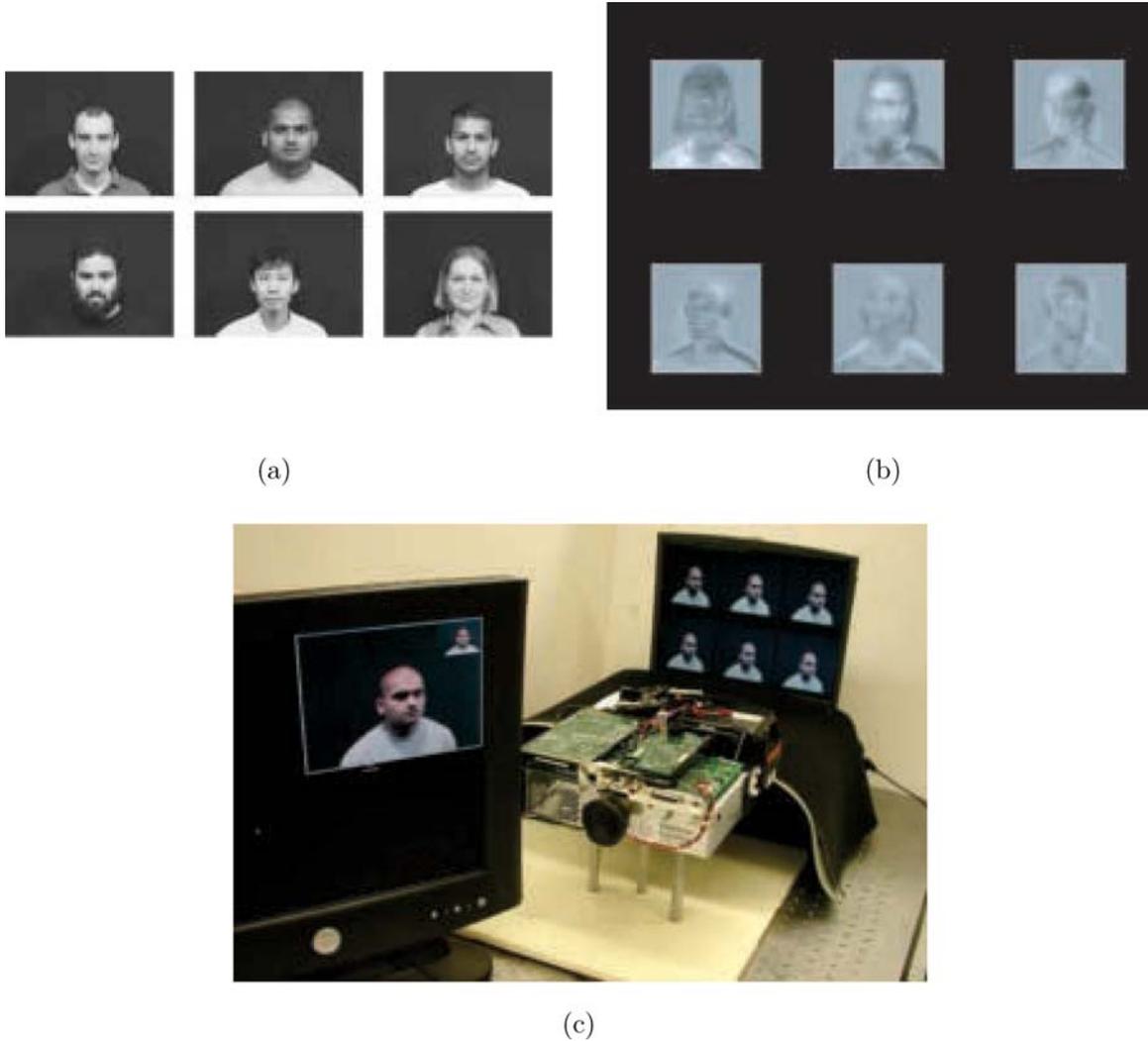


Figure 11. (a) People used in the database of the recognition system (30 different poses of each person are included). (b) The 6 most prominent eigenvectors computed from the training set, tiled to form the modulation image. (c) A tiling of the input (novel) image is “shown” to the imaging system by using an LCD display. Simple summation of brightness values in the captured image yields the coefficients needed for recognition.

screen that sits in front of the imaging system, as shown in Fig. 11(c). The 6 parts of the captured image are summed to obtain the 6 coefficients. A simple nearest-neighbor algorithm is applied to these coefficients to recognize the person in the input image.

7. Programmable Imaging Geometry

Thus far, we have mainly exploited the radiometric flexibility made possible by the use of a programmable micro-mirror array. Such an array also allows us to very quickly alter the field of view and resolution characteristics of an imaging system¹. Quite simply, a planar

array of mirrors can be used to emulate a *deformable mirror* whose shape can be changed almost instantaneously.

To illustrate this idea, we do not use the imaging system in Fig. 4 as its optics would have to be substantially altered to facilitate field of view manipulation. Instead, we consider the case where the micro-mirror array does not have an imaging lens that focuses the scene onto it but instead directly reflects the scene into the camera optics. This scenario is illustrated in Fig. 12, where the array is aligned with the horizontal axis and the view-point of the camera is located at the point P at height h from the array.

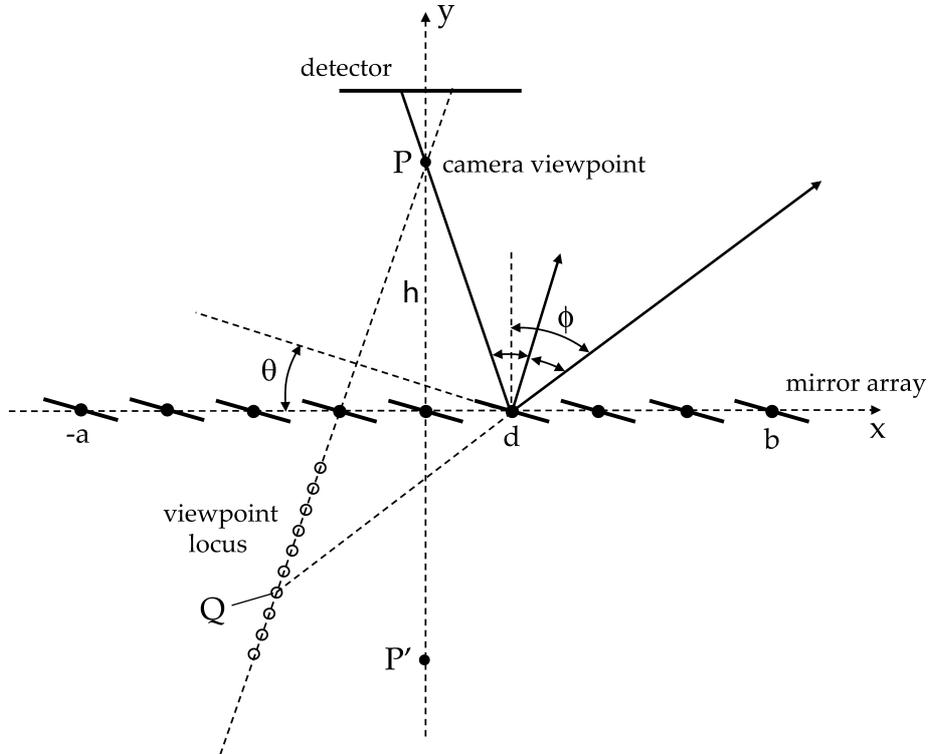


Figure 12. The field of view of an imaging system can be controlled almost instantly by using a micro-mirror array. The scene is being reflected directly by the array into the viewpoint P of the camera. When all the mirrors are tilted by the same angle (θ) , the effective field of view of the system is the same as that of the camera but is rotated by 2θ . In this case of parallel micro-mirrors, the system has a locus of viewpoints (one for each mirror) that lies on a straight line.

If all the mirrors are parallel to the horizontal axis, the array behaves like a planar mirror and the viewpoint of the system is simply the reflection P' of the camera's viewpoint P . The field of view in this case is the field of view of the camera itself (only reflected) as long as the mirror array fills the field of view of the camera. Now consider the mirror located at distance d from the origin to have tilt θ with the horizontal axis, as shown in Fig. 12. Then, the angle of the scene ray imaged by this mirror is $\phi = 2\theta + \alpha$, where $\alpha = \tan^{-1}(d/h)$. It is also easy to show that the viewpoint of the system corresponding to this particular mirror element is the point Q with coordinates $Q_x(d) = d - \sqrt{(h^2 + d^2)} \cos \beta$ and $Q_y(d) = d - \sqrt{(h^2 + d^2)} \sin \beta$ where $\beta = (\pi/2) - \phi$. If all the micro-mirrors have the same tilt angle θ , then the field of view of the system is rotated by 2θ . In this case the system has a *locus of viewpoints* (caustic) that is a segment of the line that passes through P and Q .

Figure 13 shows how the imaging system can be used to capture multiple views of the same scene to

implement stereo. In this example, the mirrors on the left half and the right half of the array are oriented at angles θ and $-\theta$, respectively. A single image captured by the camera includes two views of the scene that are rotated with respect to each other by 4θ . Note that stereo does not require the two views to be captured from exactly two centers of projection. The only requirement is that each point in the scene is captured from two different viewpoints, which is satisfied in our case.

If the mirrors of the array can be controlled to have any orientation within a continuous range, one can see that the field of view of the imaging system can be varied over a wide range. In Fig. 14 the mirrors at the two end-points of the array have orientations $-\theta$ and θ , and the orientations of mirrors in between vary smoothly between these two values. In this case, the field of view of the camera is enhanced by 4θ .

As we mentioned, the DMDs that are currently available can have only one of two mirror orientations ($+10$ or -10 degrees) in their active (powered) state. Therefore, if all the mirrors are initially inac-

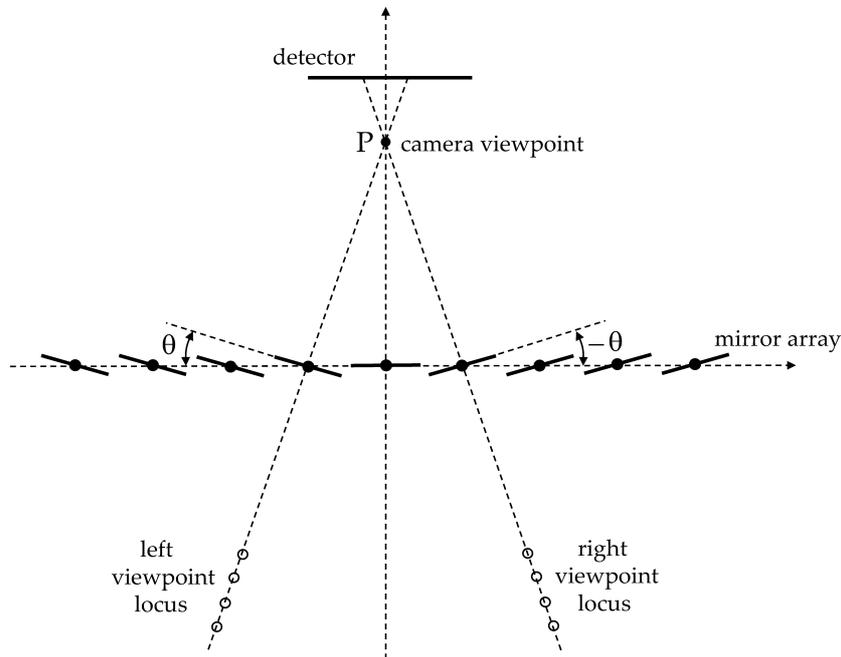


Figure 13. Here, the mirrors on the left half and the right half of the array are oriented at angles θ and $-\theta$, respectively. The result is a pair of stereo views of the same scene captured within a single camera image. Each view has a linear viewpoint locus and the two views are rotated with respect to each other by 4θ .

tive (0 degrees) and then powered and oriented at 10 degrees, the field of view remains the same but its orientation changes by 20 degrees as described earlier. This very case is shown in Fig. 15, where the left image shows one view of a printed sheet of paper and the right one shows the other (rotated) view of the same.

One can see that both the images are blurred. This is because we are imaging the scene directly through a DMD without using a re-imaging lens and hence many mirrors lie within the light cone that is imaged by a single pixel. Since the mirrors are tilted, the surface discontinuities at the edges of the mirrors cause diffraction effects. These effects become negligible when the individual micro-mirrors are larger.

8. Discussion

We have shown that programmable imaging using a micro-mirror array is a general and flexible approach to imaging. It enables one to significantly alter the geometric and radiometric characteristics of an imaging system using software. We now conclude with a few observations related to the proposed approach.

- Programmable Raxels:** Recently, a general imaging model was proposed in Grossberg and Nayar (2005) which allows one to represent an imaging system as a discrete set of “raxels”. A raxel is a combination of a ray and a pixel. It was shown in Grossberg and Nayar (2005) that virtually any imaging system can be represented as a three-dimensional distribution of raxels. The imaging system we have described in this paper may be viewed as a distribution of programmable raxels, where the geometric and radiometric properties of each raxel can be controlled via software independent of all other raxels. This does not imply, however, that *any* imaging system can be emulated using our approach. For instance, while we have significant control over the orientations of the raxels (ray directions), it is not possible to use a single mirror-array to control the positions of the raxels independent of their orientations. Even so, the space of raxel distributions one can emulate is large.
- Optical Computations for Vision:** As we have shown, our approach can be used to perform some simple image processing tasks, such as image multiplications and intra-pixel convolutions, in optical domain. While optical image processing was not the goal of our work, the ability to perform these com-

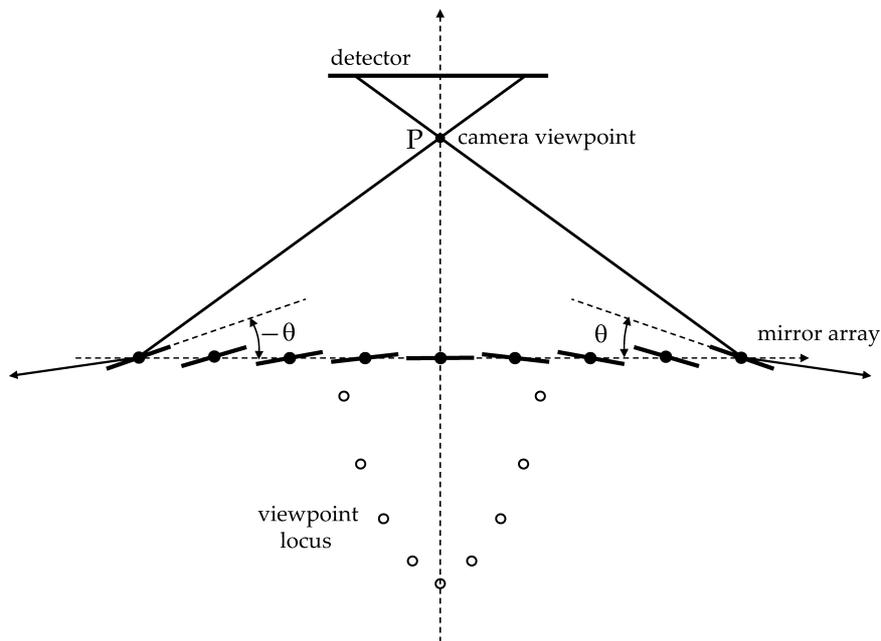


Figure 14. A planar array of planar mirrors can be used to emulate curved mirrors. Here the orientations of the mirrors vary gradually from $-\theta$ to θ . The image captured by the camera has a field of view that is 4θ greater than the field of view of the camera itself. Again, the system has a locus of viewpoints which in this case lies on a curve. Such a system can be used to instantly switch between a wide range of image projection models.

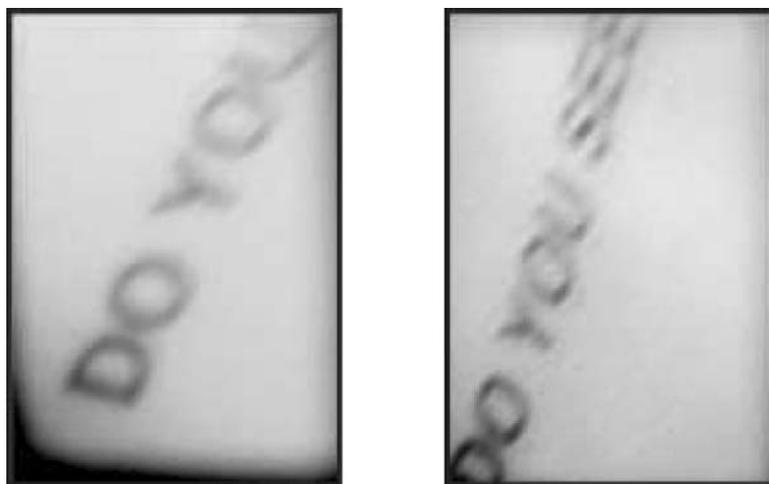


Figure 15. Two images of the same scene taken by pointing a camera directly at a DMD. The image on the left was taken with all the mirrors at 0 degrees (inactive DMD) and the image on the right will all the mirrors at 10 degrees. The fields of view corresponding to the two images are the same in terms of their solid angles but they are rotated by 20 degrees with respect to each other. Both images are blurred as the DMD is a very small and dense array of mirrors that is not appropriate for capturing direct reflections of a scene.

putations in optical domain happens to be an inherent feature of programmable imaging using a micro-mirror array. There are two major advantages to processing visual data in optical domain. The first is that

any image processing or computer vision system is resource limited. Therefore, if any computations can be moved to the optical domain and can be done at the speed of light, it naturally frees up resources

for other (perhaps higher levels) of processing. The second benefit is that the optical processing is being done while the image is formed. That is, the computations are applied to the signal when it is in its purest form—light. This has the advantage that the signal has not yet been corrupted by the various forms of noise that occur between image detection and image digitization.

- **Towards a Purposive Camera:** Any imaging system is limited in terms of its resources. At a broad level, one may view these resources as being the number of discrete pixels and the number of brightness levels (bits) each of these pixels can measure. Different specialized cameras (omnidirectional, high dynamic range, multispectral, etc.) can each be viewed as a specific assignment of pixels and bits to the scene of interest. From this viewpoint, programmable imaging provides a means for dynamically changing the assignment of pixels and bits to the scene. We can now begin to explore the notion of a purposive camera—one that has intelligence to automatically control the assignment of pixels and bits so that it always produces the visual information that is most pertinent to the task.
- **Future Implementation:** We are currently pursuing the implementation of the next prototype of the programmable imaging system. There are two limitations to the existing system that we are interested in addressing. The first is its physical packaging. We intend to redesign the hardware of the system from scratch rather than re-engineer a projector. Recently, DMD kits have become available and by using only the required components we believe we can make the system very compact. The second problem we wish to address relates to optical performance. In the current system, we have used the projector lens as the imaging lens and an inexpensive off-the-shelf lens for the re-imaging lens. Higher optical resolution can be achieved by using lenses that match the properties of the DMD and the CCD.

Finally, the full flexibility of programmable imaging will become possible only when mirror arrays provide greater control over mirror orientation. Significant advances are being made in MEMS technology as well as adaptive optics that we hope will address this limitation. When micro-mirror arrays allow greater control over the orientations of their mirrors, programmable imaging will have the potential to impact

imaging applications in several fields of science and engineering.

Acknowledgments

This work was done at the Columbia Center for Vision and Graphics. It was supported by an ONR contract (N00014-03-1-0023).

Note

1. This approach to controlling field of view using a mirror array is also being explored by Andrew Hicks at Drexel University, Hicks (2003).

References

- Castracane, J. and Gutin, M. A. 1999. DMD-based bloom control for intensified imaging systems. In *Diffractive and Holographic Tech., Syst., and Spatial Light Modulators VI*, vol. 3633, pp. 234–242. SPIE.
- Christensen, M.P., Euliss, G.W., McFadden, M.J., Coyle, K.M., Milojkovic, P., Haney, M.W., Gracht, J. van der, and Athale, R. A. October 2002. Active-eyes: An adaptive pixel-by-pixel image-segmentation sensor architecture for high-dynamic-range hyperspectral imaging. *Applied Optics*, 41(29):6093–6103.
- Dudley, D., Duncan, W., and Slaughter, J. February 2003. Emerging digital micromirror device (DMD) applications. White paper, Texas Instruments.
- Ginosar, R., Hilsenrath, O., and Zeevi, Y. September 1992. Wide dynamic range camera. U.S. Patent 5, 144,442.
- Goodman, J. W. 1968. *Introduction to Fourier Optics*. McGraw-Hill, New York.
- Grossberg, M.D. and Nayar, S.K. 2005. The raxel imaging model and ray-based calibration. *IJCV*, 61(2):119–137.
- Hicks, R. A. 2003. *Personal Communication*.
- Hornbeck, L.J. 1988. Bistable deformable mirror device. In *Spat. Light Mod. and Apps.*, vol. 8. OSA.
- Hornbeck, L.J. August 1989. Deformable-mirror spatial light modulators. In *Projection Displays III*, vol. 1150, pp. 86–102. SPIE.
- Hornbeck, L.J. 1995. Projection displays and MEMS: Timely convergence for a bright future. In *Micromachined Devices and Components*, vol. 2642. SPIE.
- Kang, S.B., Uyttendaele, M., Winder, S., and Szeliski, R. 2003. High dynamic range video. *ACM Trans. on Graph. (Proc. of SIGGRAPH 2003)*, 22(3):319–325.
- Kearney, K.J. and Ninkov, Z. 1998. Characterization of digital micromirror device for use as an optical mask in imaging and spectroscopy. *Spatial Light Modulators*, 3292:81–92. SPIE.
- Malbet, F., Yu, J., and Shao, M. 1995. High dynamic range imaging using a deformable mirror for space coronagraphy. *Public. of the Astro. Soc. of the Pacific*, 107:386.
- Murase, H. and Nayar, S.K. 1995. Visual learning and recognition of 3d objects from appearance. *IJCV*, 14(1):5–24.

- Nayar, S.K. and Branzoi, V. 2003. Adaptive Dynamic Range Imaging: Optical Control of Pixel Exposures over Space and Time. *Proc. of Inter. Conf. on Computer Vision (ICCV)*, pp. 1168–1175.
- Nayar, S.K. and Mitsunaga, T. 2000. High dynamic range imaging: Spatially varying pixel exposures. *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1:472–479.
- Nayar, S.K. and Narasimhan, S.G. 2002. Assorted pixels: Multi-sampled imaging with structural models. *Proc. of Euro. Conf. on Comp. Vis. (ECCV)*, 4:636–152.
- Smith, W.J. 1966. *Modern Optical Engineering*. McGraw-Hill.
- Turk, M. and Pentland, A.P. 1991. Face recognition using eigenfaces. In *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 586–591.
- Tyson, R.K. 1998. *Principles of Adaptive Optics*. Academic Press.