ON BEING 'UNDIGITAL' WITH DIGITAL CAMERAS: EXTENDING DYNAMIC RANGE BY COMBINING DIFFERENTLY EXPOSED PICTURES

S. Mann and R. W. Picard

Massachusetts Institute of Technology,E15-383,20 Ames Street,Cambridge,MA02139
Corresponding author currently with University of Toronto
Contact info: Tel=(416) 946-3387; Fax=(416)971-2326
steve@media.mit.edu
http://n1nlf-1.eecg.toronto.edu
http://eecg.toronto.edu/~mann
http://www.wearcam.org

BibTeX info (original appearances):

```
@techreport{mann94b,
  author = "S. Mann and R.W. Picard",
  title = "Being `undigital' with digital cameras:
          Extending Dynamic Range by Combining
          Differently Exposed Pictures",
  number = "323",
  institution = "M.I.T. Media Lab Perceptual
                Computing Section",
  address = "Boston, Massachusetts",
  year = "1994",
  note = "Also appears, {IS\&T}'s 48th annual conference,
         Cambridge, Massachusetts, May 1995",
  pages = "422-428",
}
```


ABSTRACT

Most everyday scenes have a far greater dynamic range than can be recorded on a photographic film or electronic imaging apparatus (whether it be a digital still camera, video, etc.). However, a set of pictures, that are identical except for their exposure, collectively show us much more dynamic range than any single picture. The dark pictures show us highlight details of the scene that would be washed out in a "properly exposed" picture, while the light pictures show us some shadow detail that would also not appear in a "properly exposed" picture.

We propose a means of combining differently exposed pictures to obtain a single picture of extended dynamic range, and improved color fidelity. Given a set of digital pictures, we may produce a single picture which is, for all practical purposes, 'undigital', in the sense that it is a floating point image, with the kind of dynamic range we are accustomed to seeing in typical floating point representations, as opposed to the integer images from which it was generated.

The method is completely automatic; it requires no human intervention, and it requires no knowledge of the response function of the imaging device. It works reliably with images from a digital camera of unknown response, or from a scanner with unknown response, scanning an unknown film type.

1. INTRODUCTION

1.1. Advantage of being 'undigital'

Digital photography allows us to do many things we cannot do with traditional analog photography. However, being digital is not desirable in and of itself – it is desirable for what it facilitates (instant feedback, ability to rapidly transmit high quality anywhere in the world, ease of manipulation, etc).

Digital imaging imposes certain limitations on the ways we think about images. Ideally, what we want is not bits, but, rather, a mathematical or parametric representation of the continuous underlying intensity variations projected onto an image plane, represented in a form that allows for easy transmission, storage, and analysis.

As the spatial resolution of digital images has improved over the years, we are approaching a level where the image may be regarded as essentially continuous – it is essentially free of *pixels*. Thus high resolution digital images give us the spatial continuity of analog photography, together with the ability to view pictures right away, transmit them over wireless links, analyze them computationally, etc.

However, while there may be so many pixels that we can, for all practical purposes, assume the image is a function of two real coordinates, each of these pixels are still represented as an array of integers that can assume only 256 different values, for each color channel. So-called *24 bit color*, also known as *full color*, *true color direct visual*, etc., is not as "full" or "true" as these names imply. In particular, these images are also typically manipulated using 8-bit precision arithmetic. Any simple manipulations in an image editing program, such as Photoshop, quickly degrade the quality of the images, introducing gaps in the histograms that grow with each successive computation.

The purpose of this paper is to examine the recovery of the 'true image', a real-valued quantity of light projected onto a flat surface. We regard the 'true image' as a collection of **analog** photometric quantities that might have been measured with an array of linearized lightmeters having floating-point precision, and thus, being essentially, for all practical purposes, 'undigital'.

Of course, all images that are stored on a computer are digital. A floating point number is digital. But a double-precision (64 bit) floating point number is close to analog in spirit and intent.

With the growing word size of desktop computational hardware, floating point arithmetic is becoming more practical for large images. The new DEC 3000 (Alpha) computer has a word size of 64 bits, and can easily handle images as double precision arrays. Double precision is nothing new. For years, languages like FORTRAN have supported floating point arithmetic, used widely by the scientific community, but floating point calculations are not supported in any of the popular image manipulation software such as Photoshop or Live picture. Capturing images that are essentially unlimited in dynamic range, and, while digitally represented, behave as analog images, allows us to capture and surpass the benefits traditionally offered by truly analog image formats like film.

2. WHAT IS A CAMERA

We regard an image as a collection of photometric measurements, and a camera as an array of light meters. However, in traditional imaging, each of these measurements (pixels) are made with a light meter (sensor element) that has some unknown nonlinearity followed by a quantization to a measurement having 8-bit precision.

2.1. Dynamic range and amplitude resolution

Many everyday scenes contain a tremendous dynamic range. For example, the scene might be a dimly lit room, with a window in the background; through the window we might observe a beautiful blue summer sky with puffy white clouds. Yet a picture that is exposed for the indoor scene will render the window as a white blob, blooming out into the room, where we can scarcely discern the shape of the window, let alone, see beyond it. Of course, if we exposed for the sky outside, the interior would appear completely black.

Cameras (whether analog or digital) tend to have a very limited dynamic range. It is possible to extend the dynamic range by various means. For example, in the case of photographic emulsion, the film can be made thicker, but there are tradeoffs (e.g. thicker emulsion results in increased scattering, which results in decreased spatial resolution). Nyquist showed how a signal can be reconstructed from a sampling of finite resolution in the domain (e.g. space or time), but assumed infinite dynamic range. On the other hand, if we have infinite spatial resolution, but limited dynamic range (even if we have only 1 bit of image depth), Curtis and Oppenheim [1] showed that we can also obtain perfect reconstruction. This tradeoff between image resolution, and image *depth* is also at work in a slightly different way in image *halftoning*.

Before the days of digital image processing, Charles Wyckoff formulated a multiple layer photographic emulsion [2][3]. The Wyckoff film had three layers that were identical in their spectral sensitivities (each was roughly equally sensitive to all wavelengths of light), and differed only in their overall sensitivities to light (e.g. the bottom layer was very *slow*, with an ISO rating of 2, while the top layer was very *fast* with an ISO rating of 600).

A picture taken on Wyckoff film can both record a dynamic range of one to a hundred million and capture very subtle differences in exposure. Furthermore, the Wyckoff picture has very good spatial resolution, and thus **appears** to overcome the resolution/depth tradeoff, by using different color dyes in each layer, which have a specular density as opposed the diffuse density of silver. Wyckoff printed his *greyscale* pictures on color paper, so the *fast* (yellow) layer would print blue, the medium (magenta) layer would print green, and the *slow* (cyan) layer would print red. His result was a *pseudocolor* image similar to those used now in data visualization systems to display floating point arrays on a computer screen of limited dynamic range.

Wyckoff's most well-known pictures are perhaps his motion pictures of nuclear explosions – one could clearly see the faint glow of a bomb just before it exploded (which would appear as blue, since it only exposed the fast top layer), as well as the details in the highlights of the explosion (which appeared white since they exposed all 3 layers – the details discernable primarily on account of the slow bottom layer).

2.2. Combining multiple pictures of the same scene

The idea of computationally combining differently exposed pictures of the same scene to obtain extended dynamic range has been recently proposed [4], where the images were assumed to have been taken from roughly the same position in space, with possibly different camera orientations (pan, tilt, rotation about optical axis), and different zoom settings. In this paper we describe, in further detail, the computational means of combining differently exposed pictures into a floating-point image array, and assume a simpler case, namely that all pictures are taken from a camera at a fixed location in space and a fixed orientation, with a fixed focal length lens. This simpler case corresponds to pictures that differ only in exposure.

We refer to a collection of pictures that differ only in exposure as a *Wyckoff set*, in honor of Charles Wyckoff, who was the first to exploit such a set of pictures collectively. Photographers, through a procedure called *exposure bracketing* (trying a variety of exposure settings and later selecting the one exposure that they most prefer) also produce Wyckoff sets but generally with the intent of later merely selecting the best image from the set, without exploiting the full potential value of using the images collectively.

3. EXPOSURE BRACKETING OF DIGITAL IMAGES

Whenever the dynamic range of the scene exceeds the range of the recording medium (which is almost always) photographers tend to expose for areas of interest in the scene. For example, a scene containing people is usually exposed to show the most detail in them (Fig. 1) at the expense of details elsewhere in the scene. Additionally, in our case, a picture was taken immediately afterward (Fig. 2), with four times the exposure time, so that the surrounding contextual details of the scene would show up nicely.

Ideally, only one picture would be needed to capture the entire dynamic range of the scene, and we wouldn't even need to worry about whether the picture was overexposed or underexposed because we could lighten or darken it later on, by simply using the appropriate 'lookup operator'. By 'lookup operator', we mean any spatially invariant nonlinearity: $g(x, y) = g(f(x, y))$. A 'lookup operator' is the continuous analog of a *lookup table*. Gamma correction is an example of a 'lookup operator'.

However, due to various noise sources, such as quantization noise, a 'lookup operator' will only be able to compensate

Figure 1: The Mann family standing outside an old building with the camera inside. Here the exposure was selected so that the people would show up nicely.

Figure 2: The exposure was increased by a factor of $k = 4$, compared to Fig. 1; as a result, the interior of the building is nicely visible.

for a very limited amount of overexposure or underexposure. For example, we will never recover the detail in the faces of the people from Fig. 2. The increased exposure has caused this information to be lost by the combined effect of saturation and noise. Similarly, nothing can be done to recover the shadow details in the darker portions of Fig. 1, because these areas have pixel values that are uniformly zero. Even in slightly brighter areas, where there is variation in the pixels, this variation is subject to extreme quantization noise. For example, in dark areas where the pixel values fluctuate between zero and one, there is only one bit of precision. A camera with a small number of bits of depth (such as a one-bit camera), but which has very high spatial resolution, may be used to capture a continuous tone image [1]. Indeed, a stat camera, used in a photo mechanical transfer (PMT) machine, is able to capture images that appear to be continuous-tone (due to the *halftoning screen*), even though the film can only record two distinct levels. This is possible because the film has essentially unlimited spatial resolution, and is recording through a screen of much lower (e.g. 85dpi) spatial resolution.

However, in most digital photography and video applications, spatial resolution is much lower than we would like. We do not have the luxury of essentially infinite spatial resolution that PMT systems have, and so we are not at liberty to trade spatial resolution for improved dynamic range.

Therefore, we propose the use of exposure bracketing as an alternative, whereby we make the tradeoff along the time axis, exchanging reduced frame-rate for improved dynamic range, rather than reduced spatial resolution for improved dynamic range. In particular, often a still image is all that is desired from a video camera, and in many other digital video applications, all that is needed is a few frames per second, from a camera capable of producing 30 frames per second or more.

4. SELF-CALIBRATING CAMERA

The numerical quantity appearing at a pixel in the image is seldom linearly related[1] to the quantity of light falling on the corresponding sensor element. In the case of an image scanned from film, the density of the film varies nonlinearly with the quantity of light to which it is exposed. Furthermore, the scanner will most likely introduce a further unknown nonlinearity.

We propose a simple algorithm for finding the pointwise nonlinearity of the entire process, f, that maps the light q projected on a point in the image plane to the pointwise value in the picture, $f(q)$, up to a constant scale factor. We ignore, until Section 4.1, the fact that each pixel can only assume a finite number of values, the fact that there are a finite number of pixels in the image, and the effects of image noise:

1. Select a relatively dark pixel from image a, and observe both its location, (x_0, y_0), and its numerical value, f_0. We do not know the actual quantity of light that gave rise to f_0, but we will call this unknown quantity q_0. Since f_0 is the result of some unknown mapping, f, applied to the unknown quantity of light, q_0, we denote $a(x_0, y_0)$ by $f(q_0)$.
2. Locate the corresponding pixel in image b, namely $b(x_0, y_0)$. We know that k times as much light gave rise to $b(x_0, y_0)$ as to $a(x_0, y_0)$. Therefore $b(x_0, y_0) = f(kq_0)$. For convenience, we denote $b(x_0, y_0)$ by $f(q_1)$, so that $f(q_1) = f(kq_0)$. Now search around in image a for a pixel that has the numerical value $f(q_1)$, and make a note of the coordinates of the found pixel. Call these coordinates (x_1, y_1), so that we have $a(x_1, y_1) = f(q_1)$.
3. Look at the same coordinates in image b and observe the numerical quantity $b(x_1, y_1)$. We know that k times as much light fell on $b(x_1, y_1)$ as did on $a(x_1, y_1)$. Therefore $b(x_1, y_1) = f(kq_1)$. For convenience, we denote $b(x_1, y_1)$ by $f(q_2)$. So far we have that $f(q_2) = f(kq_1) = f(k^2 q_0)$. Now search around in image a for a pixel that has the numerical value $f(q_2)$ and note these coordinates (x_2, y_2).
4. Continuing in this fashion, we obtain the nonlinearity of the image sensor at the points $f(q_0)$, $f(kq_0)$, $f(k^2 q_0)$, ..., $f(k^n q_0)$.

Now we can construct points on a plot of $f(q)$ as a function of q, where q is the quantity of light measured in arbitrary (reference) units. We illustrate this process diagrammatically (Fig 3(a)), where we have introduced a plot of the numerical values in the first image, $a = f(q)$ against the numerical values in the second image, $b = f(kq)$, which we call the 'range-range' plot, as the axes are both the range of f, with a constant domain ratio, k. Once the camera is calibrated, we may use the calibration curve to combine sets of pictures like the ones in Fig. 1 and 2. The pictures that are used to calibrate the camera need not be the same ones used to make the composite. In fact, had we used a smaller value for k to calibrate the camera (e.g. 1.4 or 2 instead of 4),

[1] In fact, quite often, photographers **desire** a nonlinear relationship: the nonlinearities tend to make the image look better when printed on media that have limited dynamic range.

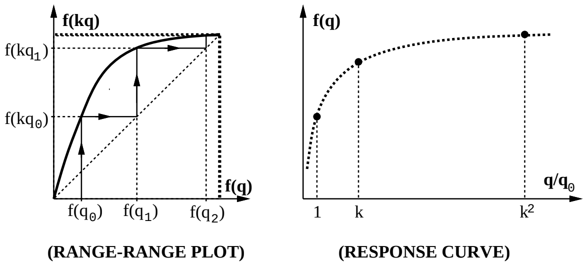

Figure 3: Procedure for finding the pointwise nonlinearity of an image sensor from two pictures differing only in their exposures. (RANGE-RANGE PLOT) Plot of pixel values in one image against corresponding pixel values in the other, which we call the 'range-range' plot. (RESPONSE CURVE) Points on the response curve, found from only the two pictures, without any knowledge about the characteristics of the image sensor. If we use a logarithmic exposure scale (as most photographers do) then the samples fall uniformly on the $log(q/q_0)$ axis.

we would have obtained more sample points on the response curve (Fig 3(b)).

In general, estimating a function, $f(q)$, from a graph of $f(q)$ versus $f(kq)$, is a difficult problem. However, we can place certain restrictions on f. For example, we suppose that f is semi-monotonic[2] (increases or remains constant with increasing q). Since the response curve is semi-monotonic, so is the plot depicted in Fig 3(a). We can also impose that $f(0) = 0$ by taking a picture with the lens cap on, and subtracting the resulting pixel value from each of the two (or more) images. This step will insure that the plot of Fig 3(a) passes through the origin.

We may be willing to place even stronger restrictions on the response curve. For example, a commonly used empirical law for film is $f(q) = \alpha + \beta q^\gamma$. This gives rise to the canonical *D log E* (density versus log exposure) curve much of which is linear. The *Dmin* (minimum density), α, would be subtracted off as suggested, using a picture with the lens cap on, and the (a, b) plot would take the form $b = k^\gamma a$, from which we could find the film's contrast parameter γ by applying regression to the points known on the range-range plot.

4.1. Quantization and other noise

In practice, the pixel values are quantized, so that the range-range plot is really a *staircase function*. It is still semi-monotonic, since it is a quantized version of a continuous semi-monotonic function.

In addition to quantization effects, we also have noise, which may be due to a variety of causes, such as thermal noise in the image sensor, grain in the film, slight misregistration of the images, or slight changes in camera position, scene content, and lighting. We consider a 'joint histogram' of the two images (Fig. 4(a)), which is the discrete equivalent of the

[2] The only practical situation that would likely violate this assumption, is where a negative film is being used, the sun is in the picture, and the sun's rays are concentrated on the film for a sufficiently long time to burn a hole through a negative film. The result is a print where the brightest object in the scene (the sun) appears black.

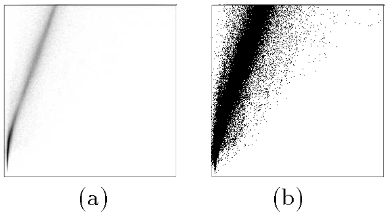

Figure 4: Cross histogram of the images in Figs. 1 and 2. The cross-histogram of two images is itself an image. Since the two images have a depth of 8 bits, the cross histogram is a 256×256 image regardless of the sizes of the two images from which it is obtained. The bin count at the origin (lower left corner) indicates how many pixels were black (had a value of zero) at the same location in both images. (a) Cross histogram displayed as an image. Darker areas correspond to greater bin counts. (b) All non-empty bins are shown as black. Ideally, there should only be a slender "staircased" curve of non empty bins, but due to noise in the images, the curve fattens.

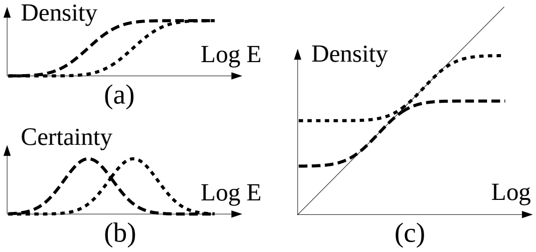

Figure 5: Response curves of the Wyckoff set (note the log scale as opposed to the scale of Fig 3 (RESPONSE CURVE) which was linear). (a) Response curves corresponding to two different exposures, depicted as though they were taken on two different films. The dashed line may be thought of as either a longer exposure or the faster layer of a 2-layer Wyckoff film, while the dotted line may be regarded as a shorter exposure or the slow layer of a 2-layer Wyckoff film. (b) "Certainty functions" calculated by differentiating the two response curves. (c) Hypothetical curves re-aligned as they will be when the images are later combined. The response of the ideal composite is indicated by the thin solid line. Using more exposure bracketing (or more layers on a Wyckoff film), we can extend this response indefinitely.

'range-range' plot of Fig 3. It is a 256 by 256 array since each pixel of the two images can assume 256 distinct values. Due to noise, we see a fat ridge, rather than a slender "staircase". Ideally there should be no points off of the staircase, defined by quantizing the range-range plot, but in practice we find a considerable number of such non-empty bins (Fig. 4(b)).

5. COMBINING IMAGES OF DIFFERENT EXPOSURE

At this point we have found the response curve (by fitting to the data in the range-range plot, as in Fig 4), and can shift the response curve to the left or right to get the curves of the two or more exposures (Fig. 5(a)). In the shadow areas (areas of low exposure, E) the same quantity of light in the scene has had a more pronounced effect on the dashed-exposure, so that the shadow detail in the scene will still be on a portion of the dashed line that is relatively steep. The highlight detail will saturate this exposure, but not the dotted-exposure.

In general, for parts of the film that are exposed in the

Figure 6: 'Crossover image' corresponding to the two pictures in Figs. 1 and 2. Black denotes pixel locations where Fig. 1 is the more "certain" of the two images, and thus where Fig. 1 should contribute to the composite. White denotes pixel locations where Fig. 2 is the more "certain" of the two images, and thus where it should contribute to the composite. In practice, we take a weighted sum of the images rather than the abrupt switchover depicted in this figure.

Figure 7: Wyckoff composite, derived from Fig. 1 and Fig. 2, reduced in contrast and then quantized to 8 bit image depth.

extremes (greatly overexposed or greatly underexposed), detail is lost – we can no longer distinguish small changes in light level since the resulting changes in film density are so small that they fall below the noise floor (e.g. we are operating on the flat parts of Fig 5). On the other hand, steep portions of the response curves correspond to detail that can be more accurately recovered, and are thus desirable operating points. In these regions, small changes in light will cause large changes in the measured value of the response function, and even if the measurements are highly quantized (e.g. only made with 8 bit precision), small differences in the measured quantities will remain discernable.

Thus we are tempted to plot the derivatives of these hypothetical response curves (Fig. 5(c)), which we call the *certainty functions.*

At first glance, one might be tempted to make a composite from two or more differently exposed pictures by manually combining the light regions from the darker pictures and the dark regions from the lighter pictures (e.g. manually selecting the middle of Fig. 1 and pasting on top of Fig. 2). However, we wish to have the algorithm automatically combine the images. Furthermore, the boundary (Fig. 6) between light regions and dark regions is, in general, not a smooth shape, and would be difficult to trace out by hand. Pasting this irregular region of Fig. 1 into Fig. 2, amounts to choosing, at each point of the composite, the source image that has the higher *certainty* of the two. However, abrupt changes resulting from suddenly switching from one image to another ocasionally introduce unpleasant artifacts, so instead, we compute a weighted average. Every pixel of the composite, whether shadow or highlight, or in the transition region, is drawn from all of the input images, by weighting based on the certainty functions. This provides a gradual transition between the images, where the shadow detail comes **primarily** from the lighter image, and the highlight detail comes **primarily** from the darker image.

The extended-response image array from the two pictures of Figs. 1 and 2 is a floating point array which has more than 256 distinct values, and therefore cannot be displayed on a conventional 8-bit display device.

6. DYNAMIC RANGE; 'DYNAMIC DOMAIN'

Tekalp, Ozkan, and Sezan [5], Irani and Peleg [6], and Mann and Picard [7] have proposed methods of combining multiple pictures that are identical in exposure, but differ in camera position. The result is increased *spatial resolution.* When one of these images is too big to fit on the screen, we look at it through a small movable viewport, scrolling around and exploring one part of the 'image domain' at a time.

In this paper, the composite image is a *floating point* array, and is therefore too *deep* for conventional screen depths of 24 bits (8 bits for each color channel), so we constructed a slider control to allow the user to interactively look at only part of the 'image range' at a time. The user slides the control back and forth depending on the area of interest in the composite image. This control is to screen range as the scrolling window is to screen domain – showing the vast tonal range one piece at a time. Of course we were able to obtain the underexposed view much like Fig 1, by sliding the control left, and the overexposed view much like Fig 2 by sliding the control right.

When an image is too big to fit on the screen, one can also subsample its domain to make it fit on the screen. Analogously, we applied the appropriate range-subsampling (quantization to 8 bits) to our floating-point composite image for screen display, or print (Fig 7). Before quantization, we applied a nonlinearity which restored the appearance of the image to the familiar tonal scale to which photographers are accustomed, and we added the appropriate amount of noise (*dither*)[3]. It is worth mentioning that the final nonlinearity before quantization selects the tonal range of interest. We can regard its derivative (the 'certainty function') as depicting the 'Wyckoff spectrum' (which regions of greyvalue are emphasized and by how much) analogous to a conventional bandpass filter which selects the frequencies of interest. The elements of a Wyckoff set, having equally spaced certainty functions of identical shape, are analogous to a bank of *constant Q* filters.

If all that is desired is a single print, why not just try to formulate a super-low-contrast film or image sensor? The superiority of the Wyckoff composite lies in the ability to control the process of going to the low contrast medium. For example, we might apply a homomorphic [8] filtering operation to

[3] The dither did not have a perceivable effect on an 8 bit image, but when reducing a Wyckoff composite to 5 bits or less, the dither made a noticable improvement.

the final composite, which would bring out improved details at high spatial frequencies, while reducing the unimportant overall changes in density at low spatial frequencies.

7. WYCKOFF ANALYSIS AND SYNTHESIS FILTERBANKS

We can regard the Wyckoff film (or exposure bracketing) as performing an *analysis* by decomposing the light falling on the sensor into its 'Wyckoff layers'. The proposed algorithm provides the *synthesis* to *reconstruct* a floating point image array with the dynamic range of the original light falling on the image plane. This *analysis-synthesis* concept is illustrated in Fig 8.

The analysis-synthesis concept suggests the possibility of using the Wyckoff layer decomposition as a "Wyckoff filter" that could, treat the shadows, midtones, and highlights of an image differently. For example, we might wish to sharpen the highlights of an image without affecting the midtones and shadows.

The Wyckoff filter provides a new kind of filtering – 'amplitude domain' filtering – as opposed to the classic *Fourier domain*, *spatial domain*, *temporal domain*, and *spatiotemporal* filters. We envision a generalized Nyquist-like theory for reconstruction from 'amplitude samples', to augment classic sampling theory.

8. 'LIGHTSPACE'

The concept presented in this paper is part of a larger framework called 'lightspace'[9], which is a description of the way a scene responds to light. 'Lightspace' is the space of all possible photometric measurements taken for each possible photometric excitation.

Regarding an image of size $M \times N$ pixels as a point or vector in $\mathbb{R}^{MN}$, allows us to consider each of a set of differently exposed images, prior to nonlinearities and quantization, as colinear vectors in $\mathbb{R}^{MN}$.

Furthermore, if we obtain multiple pictures of the same scene differing only in lighting, they span a subspace of $\mathbb{R}^{MN}$, which we call the 'lightvector subspace'. From any set of 'lightvectors' (pictures of a scene taken with particular lighting) that span a particular 'lightvector subspace' we can synthesize pictures taken with any combination of the light sources.

To the extent that a multichannel image (such as color, having three channels: R,G,B), having L channels is a collection of L vectors, then for each of a set of multiple channel pictures differing only in lighting, we can associate L vectors. We call the set of L vectors a 'lightmodule'.

It has been shown[10] that a set of 'lightmodules' (which we call a 'lightmodule subspace') also spans a useful space. For example, a set of color pictures of a scene differing only in lighting, taken with white lights at various places in the scene, was used to synthesize the result of having taken a picture with colored lights at these same locations.

9. SUMMARY

We have presented a means of combining multiple digital images that differ only in their exposure, to arrive at an extended-response floating point image array. The method proceeds as follows:

1. From the set of pictures (or from another set of pictures taken with the same camera) determine the camera's pointwise response function using the "self-calibration" method of Section 4.
2. Linearize the images (undo the nonlinear response of each), if desired, or map the response curves onto one desired final response curve.
3. Compute the *certainty function* by differentiating the response function. The certainty function of each image is found by appropriately shifting this one certainty function along the exposure axis.
4. Compute the weighted sum of these images, weighting by the *certainty functions*.

The composite may be explored interactively or contrast-reduced and quantized, for a conventional display device. Furthermore, we can regard the Wyckoff film (or exposure bracketing) as performing an *analysis* by decomposing the light falling on the sensor into its 'Wyckoff layers'. The proposed algorithm provides the *synthesis* to *reconstruct* a floating point image array with the dynamic range of the original light falling on the image plane. This suggests the possibility of a 'Wyckoff filter' that could, for example, blur the highlights of an image while sharpening the midtones and shadows. Wyckoff filters work in the 'amplitude domain', in contrast to Fourier filters which work in the frequency domain, or spatiotemporal filters which work in the space and time domains.

10. REFERENCES

[1] S. R. Curtis and A. V. Oppenheim, "Signal reconstruction from Fourier transform sign information," Technical Report No. 500, MIT Research Laboratory of Electronics, May 1984.

[2] C. W. Wyckoff, "An experimental extended response film," Tech. Rep. NO. B-321, Edgerton, Germeshausen & Grier, Inc., Boston, Massachusetts, MARCH 1961.

[3] C. W. Wyckoff, "An experimental extended response film," *S.P.I.E. NEWSLETTER*, JUNE-JULY 1962.

[4] S. Mann, "Compositing multiple pictures of the same scene," in *Proceedings of the 46th Annual IS&T Conference*, (Cambridge, Massachusetts), The Society of Imaging Science and Technology, May 9-14 1993.

[5] A. Tekalp, M. Ozkan, and M. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," in *Proc. of the Int. Conf. on Acoust., Speech and Sig. Proc.*, (San Francisco, CA), pp. III–169, IEEE, Mar. 23-26, 1992.

[6] M. Irani and S. Peleg, "Improving Resolution by Image Registration," *CVGIP*, vol. 53, pp. 231–239, May 1991.

[7] S. Mann and R. W. Picard, "Virtual bellows: constructing high-quality images from video," in *Proceedings of the IEEE first international conference on image processing*, (Austin, Texas), Nov. 13-16 1994.

[8] T. G. Stockham, Jr., "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828–842, July 1972.

[9] C. Ryals, "Lightspace: A new language of imaging," *PHOTO Electronic Imaging*, vol. 38, no. 2, pp. 14–16, 1995. http://www.novalink.com/pei/mann2.html.

[10] S. Mann, "Lightspace." Unpublished report (Paper available from author). Submitted to SIGGRAPH 92. Also see example images in http://wearcam.org/lightspace, July 1992.

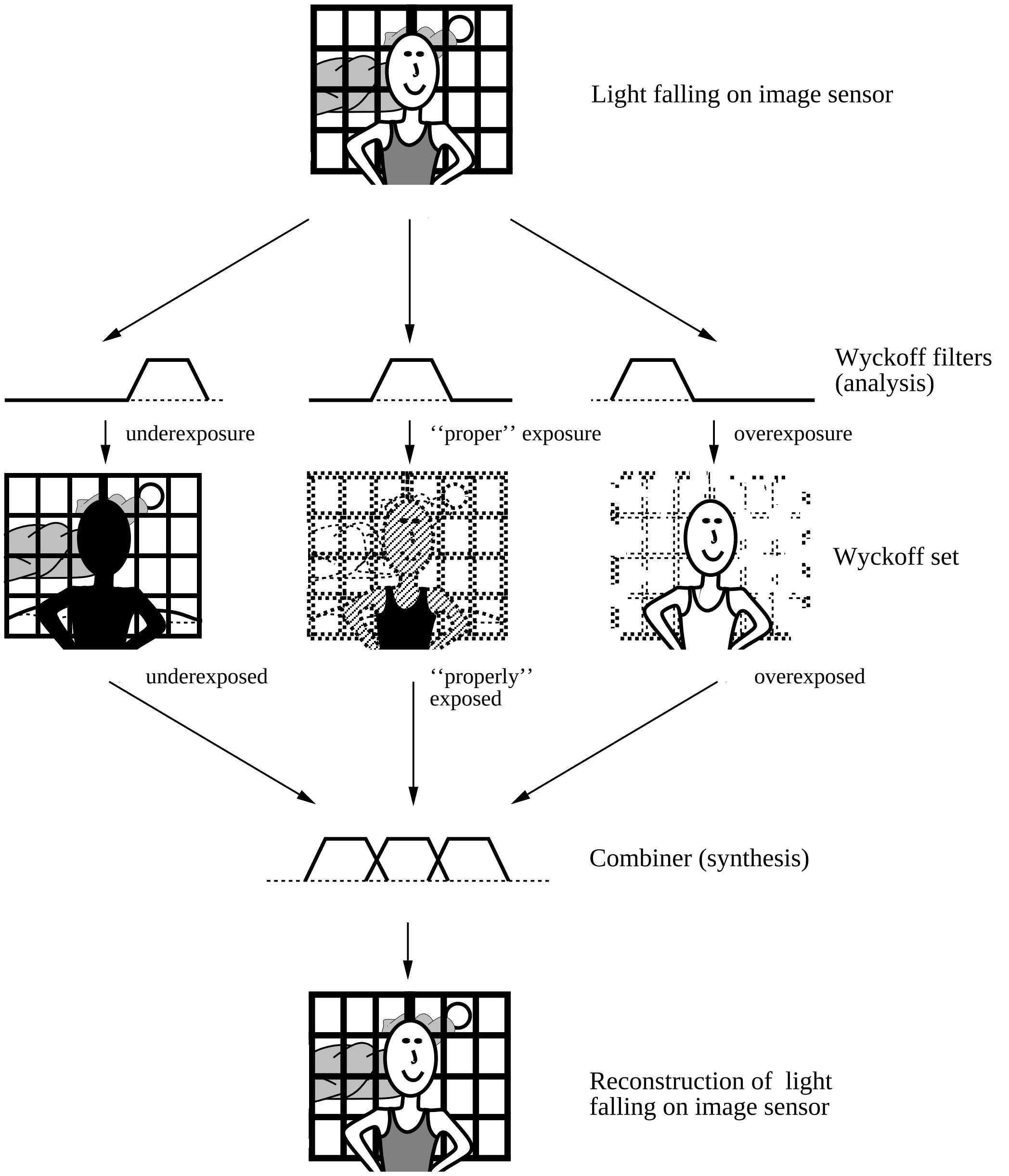

Figure 8: The layers of a Wyckoff film decompose the light falling on the film into differently exposed images. Each of these images may be regarded as a filtered version of the light falling on the image sensor. These 'Wyckoff filters' act as a filterbank to capture overlapping portions of the exposure "spectrum", and perform an analysis of the light falling on the image sensor. The set of pictures can then be used to obtain perfect reconstruction of the original light intensity falling on the image sensor.