Quantigraphic Imaging:
Estimating the camera response and exposures from differently exposed images

Steve Mann
University of Toronto, Dept. E.C.E.
Toronto, Ontario, Canada, M5S 3G4

Richard Mann
Dept. C.S.
University of Waterloo

Abstract

Multiple differently exposed pictures of the same subject matter arise naturally whenever a video camera having automatic exposure captures multiple frames of video with the same subject matter appearing in regions of overlap between at least some of the successive video frames. Almost all cameras have some kind of automatic exposure feature. Generally automatic exposure is center weighted, so that when a light object falls in the center of the frame the exposure is automatically decreased, whereas the exposure is automatically increased when the camera swings around to point at a darker object. In this paper, it is assumed that the spatial (e.g. projective) coordinate transformation between successive frames of the sequence is known (or equivalently that it is the identity), and the contribution of the paper is an efficient way to estimate the tonal relationship between successive frames of the sequence. In particular methods are proposed to simultaneously estimate the unknown camera response function, as well as the set of unknown relative exposure changes among images, up to a single unknown scalar constant. The method comprises a succession of guesses each of which is a refinement of the previous. The first guess is often sufficient, so that no initial solution needs to be provided by the user. Each subsequent guess is a least squares solution so that no computationally expensive optimization is required. Since the method makes use of all the data, it is extremely immune to noise. The method is tested against state-of-the art laboratory measurement instruments to confirm the accuracy of the results.

1. Introduction: Variable gain image sequence processing

Many papers have been published on the problems of motion estimation and frame alignment; for review see [1]. Most of these assume fixed gain. In practice, however, camera gain varies to compensate for varying quantity of light, by way of Automatic Gain Control (AGC), automatic level control, or some similar form of automatic exposure.

In fact almost all modern cameras incorporate some form of automatic exposure control. Moreover next generation cameras, such as EyeTap devices (http://eyetap.org) that cause the eye itself to function, in effect, as if it were both a camera and display, also feature an automatic exposure control system to make possible a hands free gaze activated wearable system operable without conscious thought or effort. Indeed, the human eye itself incorporates many features akin to the automatic exposure or AGC of modern cameras.

(a) (b) (c)

Figure 1: Automatic exposure as the cause of differently exposed pictures of the same (overlapping) subject matter: (a) Looking from inside Hart House Soldier's Tower, out through an open doorway, when the sky is dominant in the picture, the exposure is automatically reduced, and we can see the texture (clouds, etc.) in the sky. We can also see University College and the CN Tower to the left. (b) As we look up and to the right, to take in subject matter not so well illuminated, the exposure automatically increases somewhat. We can no longer see detail in the sky, but new architectural details inside the doorway start to become visible. (c) As we look further up and to the right, the dimly lit interiour dominates the scene, and the exposure is automatically increased dramatically. We can no longer see any detail in the sky, and even the University College building, outside, is washed out (overexposed). However, the inscriptions on the wall (names of soldiers killed in the war) now become visible. (a,b,c) The differently exposed pictures of overlapping subject matter can be combined to extend dynamic range and tonal definition, or to provide a true photographic quantity "lightspace" for intelligent vision systems.

Figure 1 illustrates how such a camera takes in a typical scene. As we look straight ahead we see mostly sky, and the exposure is quite small. Looking to the right, at darker subject matter, the exposure is automatically increased. Since the differently exposed pictures depict overlapping subject matter, we have (once the images are registered, in regions of overlap) differently exposed pictures of identical subject matter. (Registration typically also includes correction of barrel distortion, correction for darkening at the corners of the image such as by $cos(\alpha^4)$, etc., to make the camera become a truly quantimetric instrument.) In this example, we have three very differently exposed pictures depicting parts of the University College building and surroundings.

1.1 Previous work on variable gain image sequences

Fully automated computational approaches to simultaneously estimate:

- the unknown nonlinear response function of the camera;
- the unknown overall change in gain caused by automatic exposure control or AGC; and
- the projective coordinate transformation relating the images to one another,

were presented (implemented and shown) in [2], and further explored in [3] using both parametric and nonparametric methods. A goal of this work was to combine variable gain image sequences into a single image of increased spatiotonal range and definition, as well as for a front end to a wearable vision system. More recently, Szeliski also considered the problem of estimating only the projective coordinate transformation between images [4], while Debevec and Malik [5] have considered the problem of estimating only the camera response function. Mitsunaga and Nayar have also considered the problem of estimating the response function using a low order polynomial [6]. Mann has also considered parametric estimates of the camera response function, by proposing a simple three parameter function that provides a very good fit to most camera response functions [7].

This paper concentrates on nonparametric determination of camera response functions. The problem of nonparametric reverse engineering a camera's response function, from differently exposed images of identical or overlapping subject matter, was first proposed and first solved in [2]. In this paper we present, in detail, such a computationally efficient maximum likelihood estimation based on least squares.

1.2 Quantimetric imaging

The quantity, denoted q, of light, to which an image sensor responds, is known as the **photoquantity** [7] which is neither **radiometric** nor **photometric**, but nevertheless provides quantifiable units of light. **Quantimetric imaging**, also known as (photo)quantigraphic imaging [7] measures the quantity of light in units that depend on the spectral response of the particular sensor used in a particular camera, rather than in units of flat spectral response (radiometry) or in units of the human eye's response (photometry).

Differently exposed images (e.g. individual frames of video) of the same subject matter are denoted as vectors: $f_0, f_1, \ldots, f_i, \ldots, f_{I-1}, \forall i, 0 \leq i < I$.

Each video frame is some unknown function, $f()$, of the actual quantity of light, $q(\mathbf{x})$ falling on the image sensor:

$$f_i = f\left(k_i q\left(\frac{\mathbf{A_i x} + \mathbf{b_i}}{\mathbf{c_i x} + d_i}\right)\right), \tag{1}$$

where $\mathbf{x} = (x, y)$ denotes the spatial coordinates of the image, k_i is a single unknown scalar exposure constant, and parameters $\mathbf{A}_i$, $\mathbf{b_i}$, $\mathbf{c}_i$, and d_i denote the projective coordinate transformation between successive pairs of images: $A \in \mathbf{R}^{2\times 2}$ is the linear coordinate transformation (e.g. accounts for magnification in each of the x and y directions and shear in each of the x and y directions), b is the translation in each of these two coordinate directions, and c is the projective chirp rate in each of these two coordinate directions[3]. The additional constant d makes the coordinate transformation into a group.

For simplicity, this coordinate transformation is assumed to be able to be independently recovered (e.g. using the methods of [3]). Therefore, without loss of generality, in this paper, it will be taken to be the identity coordinate transformation, which corresponds to the special case of images differing only in exposure.

Without loss of generality, k_0 will be called the reference exposure, and will be set to unity, and frame zero will be called the reference frame, so that $f_0 = f(q)$. Thus we have:

$$\frac{1}{k_i} f^{-1}(f_i) = f^{-1}(f_0), \forall i, 0 < i < I. \tag{2}$$

Taking the logarithm of both sides,

$$F^{-1}(f_i) - K_i = F^{-1}(f_0), \forall i, 0 < i < I, \tag{3}$$

where $K = log(k)$, and F^{-1} is the logarithmic inverse camera response function (e.g. a LookUp Table converting pixel values into exposure values).

Re-arranging, we have:

$$F^{-1}(f_i) - F^{-1}(f_0) = K_i, \forall i, 0 < i < I. \tag{4}$$

This relation suggests a way to estimate the camera response function, f, from a pair of differently exposed images of the same subject matter. Before estimating the camera response function, we consider how the noise will affect the estimation.

1.3 Estimation in the presence of noise

Consider first, a special case in which we know each of the relative exposure values K_i. For this special case, images are related to to the actual quantity q as well as to each other, by:

$$f_i(\mathbf{x}) = f\left(k_i q\left(\mathbf{x}\right) + n_{q_i}\right) + n_{f_i} \tag{5}$$

where n_{q_i} includes sensor noise, and n_{f_i} includes image noise due to quantization, compression, transmission. (For precise definitions of these two noise sources, see [7].)

In the presence of noise, each picture provides an estimate of the actual quantity of light falling on the image sensor:

$$\hat{q}_i(\mathbf{x}) = \frac{1}{\hat{k}_i}\hat{f}^{-1}\left(f_i(\mathbf{x})\right) \tag{6}$$

where $\hat{k}_i$ is an estimate of the actual exposure constant k_i, and $\hat{f}$ is an estimate of the true camera response function f, assuming $n_{q_i} >> n_{f_i}$ [7].

Multiple estimates of the actual quantity of light falling on the image sensor may be combined as follows:

$$\hat{q}(\mathbf{x}) = \frac{\sum_i \hat{c}_i \hat{q}_i(\mathbf{x})}{\sum_i \hat{c}_i} \tag{7}$$

Photographic film is traditionally characterized by the so-called "Density versus log Exposure" *characteristic curve* [8][9]. Similarly, in the case of electronic imaging, we may also use logarithmic exposure units, $Q = \log(q)$, so that one image will be $K = log(k)$ units darker than the other:

$$\log(f^{-1}(f_1(\mathbf{x}))) = Q = \log(f^{-1}(f_2(\mathbf{x}))) - K \tag{8}$$

The existence of an inverse for f follows from a semi-monotonicity assumption. Semimonotonicity follows from the fact that we expect pixel values to either increase or stay the same with increasing quantity of illumination, q[1]. Since the logarithm function is also monotonic, the problem comes down to estimating the semimonotonic function $F() = \log(f^{-1}())$ and the scalar constant K.

The unknowns (F and K) may be solved in a least squares sense[2].

1.4 Estimating camera response function

From (4), we may find F by minimizing the sum of squared errors:

$$\epsilon = \sum_x \sum_y \left(F(f_1(\mathbf{x})) - F\left(f_2\left(\mathbf{x}\right)\right) + K\right)^2 \tag{9}$$

[1]Except in rare instances where the illumination is so intense as to damage the imaging apparatus, as, for example, when the sun burns through photographic negative film and appears black in the final print or scan.

[2]In a typical imaging situation with 480 by 640 images and 256 grey values, this amounts to solving 307200 equations in 257 unknowns: 256 for F and one for K.

Since F can only be found up to a single unknown scalar constant, $F(N)$ is fixed at zero[3], where N is the number of greyvalues (typically $N = 256$).

Equation 4 gives a set of equalities of the form: $\mathbf{AF} = -\mathbf{K}$, where $\mathbf{A} \in \mathbf{R}^{L+1 \times N}$, and L is the number of pixels in one of the images f_i. The first L rows of $\mathbf{A}$ are constructed by inserting 1 in the column index corresponding to the pixel value of f_1 and inserting -1 into the column index corresponding to the pixel value of f_2:

$$\begin{aligned} \mathbf{A}(x + wy, f_1(x, y)) &= 1 \\ \mathbf{A}(x + wy, f_2(x, y)) &= -1 \end{aligned} \tag{10}$$

where w is the width of one of the images, and the last row of $\mathbf{A}$ is all zeros except its last entry which is 1:

$$\mathbf{A}(L + 1, N) = 1 \tag{11}$$

All unspecified entries of matrix $\mathbf{A}$ are zero. Vector $\mathbf{K}$ is constructed by placing the value K in the first L entries and 0 in the last entry. This is an overdetermined system of equations.

The solution that minimizes the error $\varepsilon = \|AF + K\|^2$ in (9) is the maximum likelihood solution according to the noise model of (5), assuming $n_{q_i} >> n_{f_i}$ [7], and is given by:

$$\frac{d\epsilon}{d\mathbf{F}} = 2\mathbf{A}^T\mathbf{A}\mathbf{F} + 2\mathbf{A}^T\mathbf{K} = 0 \tag{12}$$

giving

$$\mathbf{F} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T(-\mathbf{K}), \tag{13}$$

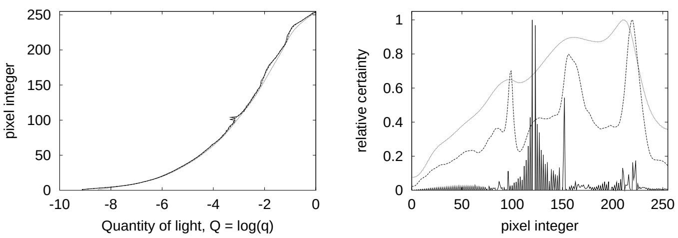
assuming additive white Gaussian noise. This solution gives us a way of estimating the camera response function from two or more differently exposed pictures of overlapping subject matter.

Although this system is massively overdetermined, the constraints follow a comparametric form that admits solutions having sinusoidal components (e.g. solutions tend to be "wavy") [7]. (Indeed, comparametric forms are very similar to difference equations as can be seen from the form of (4) which constrains the function over an interval step size of K_i allowing ripples in the solution.)

1.4.1 Smoothness and monotonicity constraints

We can make an inference that the true F is probably smooth, and, as mentioned previously, F **must** be semi-monotonic.

[3]This choice is arbitrary, as F could have been fixed at any point. However, it is common practice, when working with logarithmic units, to define the maximum quantity as zero (e.g. audio and video recorders typically have signal measurement meters calibrated so maximum signal input corresponds to 0dB), and since F is assumed to be monotonic, the constraint makes F entirely negative when images are of increasing exposure ($K > 0$).



(a) (b)

Figure 2: (a) Response functions estimated from the data for no smoothing (solid line), smoothing with $\lambda = 100$ (dotted line), and smoothing with $\lambda = 1000$ (dotted line). (b) Derivatives of the response functions make the effects of smoothing more evident. The derivative of the response function is called the certainty function, because it shows the sensitivity of pixel integer output as a function of changes in light [7]. Slight ripple in the response function becomes magnified and visible as periodicity components in these certainty functions. Note how the harmonic contributions have a period of K (normalized to stepsize $K = 1$ in this plot).

Thus the estimate of F can be constrained in both smoothness and monotonicity. A smoothness constraint may be formulated by appending $N - L_s$ extra rows to $\mathbf{A}$ and the same number of extra zeros to the end of $\mathbf{K}$, where L_s is the length of the appropriate smoothness *filter*, $\mathbf{s}$, and then solving (12). The extra rows appended to $\mathbf{A}$ are constructed as follows: Let $\mathbf{A_s} \in \mathbf{R}^{N-L_s \times N}$ denote the portion appended to $\mathbf{A}$, and create $\mathbf{A_s}$ as a toeplitz matrix in which the first L_s elements of the first row are the *filter* coefficients, and the remaining rows are appropriately shifted versions of the coefficients:

$$A_s = \begin{bmatrix} s_1 & s_2 & s_3 & \dots & 0 & 0 & 0 & \dots \\ 0 & s_1 & s_2 & s_3 & \dots & 0 & 0 & \dots \\ \dots & & & & & & & \\ 0 & 0 & 0 & \dots & s_1 & s_2 & s_3 & \dots \end{bmatrix} \quad (14)$$

In order to achieve smoothness, the *filter* needs to be a *highpass filter*. (The intuition for this comes from the fact that by "looking" at the function through a highpass filter, this makes it "expensive" for the curve to have high frequency (non-smoothness) content since the right hand side vector for this portion of the matrix equations is zero.)

The simplest filter is a three-tap filter $\lambda[1, -2, 1]$ for which the effect of appending the corresponding $\mathbf{A_s}$ to $\mathbf{A}$ is to impose a penalty for nonzero second derivatives (inflection) of the curve F. The *amplitude*, λ, of the filter, determines how heavily the smoothness constraint is weighted. Additionally, a monotonicity constraint may be imposed using Quadratic Programming (QP).

Examples of determining the response function, F, from two differently exposed images, are shown in Fig 2(a). The rippling is particularly evident when we consider the derivatives of the response functions as shown in Fig 2(b).

1.4.2 Computational efficiency

The least squares formulation of (13) can be solved using the standard approach for the normal equations. Debevec [5] presented a similar least squares formuation, but his was much more computationally intensive because he introduced extra unnecessary variables corresponding to our Q for each exposure. Due to his large number of unnecessary variables he was forced to use a small number of hand–selected points to calculate a computationally tractable solution.

In contrast, our method (and earlier methods such as that of [2]) is computationally tractable, such that it uses the information at all the pixels and hence could be expected to be much more immune to noise.

Our method is also fully automatic and does not require any hand selection of image portions in order to make it computationally tractable.

The method presented in this paper can be be made more efficient (typically hundreds of times more efficient than it already is) by the following:

- Recognizing that many rows of A are repeated, these rows can be grouped together into a single row and weighted by the square root of the number of repetitions. By weighting K also by the same factor, the same numerical solution results.
- Rows that are zero, which happen when the corresponding pixel location in both images has the same value, may be eliminated, further reducing the size of the matrix $\mathbf{A}$.

1.4.3 Robust statistics and yet more computationally efficient solutions: From " Lebesgue summation" to comparagram

Another way to interpret the error presented in (9) is to rewrite the summation as a sum over the range rather than over the domain of the images[4]:

$$\epsilon = \sum_{m=0}^{N} \sum_{n=0}^{N} J(m,n)(F(f_1(x,y)) - F(f_2(x,y)) + K)^2 \quad (15)$$

summing over the values of (x, y) for which $F(f_1(x, y)) = m$ and $F(f_1(x, y)) = n$.

The quantity J provides a weighting for the number of pixels in image f_1 that are equal to m and at the same coordinates (x, y) equal to n in image f_2.

It will be easier to understand the form of (15) by grouping (counting) all of the pixel values for which $f_1(x, y) =$

[4]This summation over level sets is reminiscent of Lebesgue integration (e.g. the form of summation in (15) is to the earlier form of the summation given in (9) as Lebesgue integration is to Riemann integration, so we will call it "Lebesgue summation".

m and, in the corresponding location $f_n(x, y) = n$. The result, denoted by the letter J, is called the *comparagram*[7] of the two images, and has dimensions N by N. It should be noted that J captures all of the information about the tonal relationship between the two pictures while discarding all of the spatial information in the images. This tonal relationship is all that is needed to solve (9) while discarding a great deal of irrelevant information. To solve (15), the matrix $\mathbf{A} \in \mathbf{R}^{N^2 \times N}$ and the vector $\mathbf{K}$ are constructed to have one row for each element of J, as follows:

$$\begin{aligned}
\mathbf{A}(mN + n, m) &= \sqrt{J(m, n)} \\
\mathbf{A}(mN + n, n) &= -\sqrt{J(m, n)} \qquad (16) \\
\mathbf{K}(mN + n) &= \sqrt{J(m, n)} \qquad (17)
\end{aligned}$$

with the additional constraints appended to A as before, resulting in the same method of solution as before (least-squares).

Looking at comparagrams of typical sets of differently exposed images, the comparagrams tend to be very sparse. Thus working with comparagrams provides us with a very tidy way of greatly increasing the computational efficiency, as well as performing robust statistics (pruning low entries).

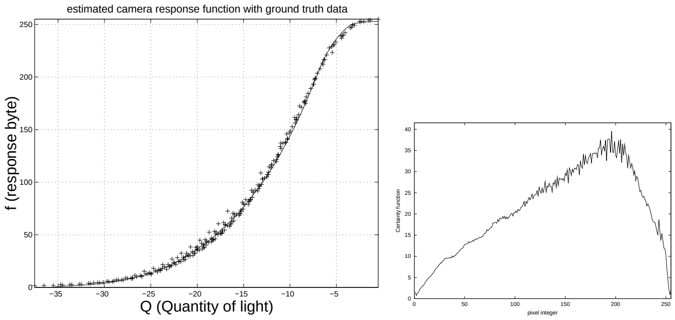
Thus the computation can be made still more efficient by removing rows of A and K that are zero, which is done by only including nonempty bins of J that are off the main diagonal (zero entries result from either empty bins of J or from diagonal entries of J). Jointly, the comparagram across two images seldom contains more than ten percent nonzero entries, so the resulting computational savings is very significant.

Robust statistics provide improved immunity to noise but this immunity is usually obtained at some computational expense. However, the formulation of (15) with solution (17) may be made both computationally more efficient as well as more robust by thresholding J. This is done by setting any entries in J that are less than a certain number equal to zero. A threshold of 10, for example, means that any bins containing less than 10 counts are set to zero prior to solving (17). Interestingly enough, this method of throwing away outliers results in further increases in computational efficiency.

1.5 Estimation with more than two input images

A simplified, though artificial situation, is when the camera is held still and the exposure is adjusted manually. Although not so realistic in today's world of mostly automatic cameras, this situation helps provide insight into the problem.

Here, a dataset of subject matter differing only in exposure, is used to calibrate the system. The sequence is from a dark interiour looking out into bright sunlight, with bright sky in the background, the dynamic range of the original subject matter being far in excess of what can be captured in any one of the constituent pictures. Such an image sequences is shown in Fig 3.



(a) (b)

Figure 4: (a) A least squares solution to the data shown in Fig 3, using a novel multiscale smoothing algorithm, is shown as a solid line. The plus signs denote known ground truth data measured from the camera using professional laboratory instruments (as described later in this paper). (b) The derivative of the computed response function is the certainty function. Note that despite the excellent fit to the known data, the certainty function magnifies slight roughness in the curve.

The comparagram (the square matrix J that arose from (15)) is a very powerful tool for understanding the relationship between differently exposed pictures of the same subject matter since it contains all that can be known about the response function of the imaging apparatus [7]. Fig 3 shows the variable image sequence together with comparagrams for various successive pairs of images as indicated.

In this case, rather than considering all possible pairs of images in a least squares unrolling of the comparagrams, it is only necessary to consider the three possible pairs of comparagrams to get a total least squares estimate.

Reverse engineering (e.g. discovering or determining) the response function from the comparagram may be achieved through a logarithmic unrolling of the comparagram, as if it were a logistic map [2]. Applying a least squares unrolling to this data provides the recovered response function shown in Fig 4(a). This solution recovers a lookup table, for converting an image into lightspace [7]. Although this estimate of the response function, $f(Q)$ looks reasonable, next to known points found with professional lab equipment, we can gather more insight by plotting the derivative (known as the *certainty function* [7]) of the response curve. This *certainty function* is shown in Fig 4(b).

The resulting recovered response function, $f(q)$, of Fig 4(a), found by unrolling the comparagrams, can be verified by regenerating ordered pairs $(f, g) = (f(q), f(kq))$ as shown in Fig 5.

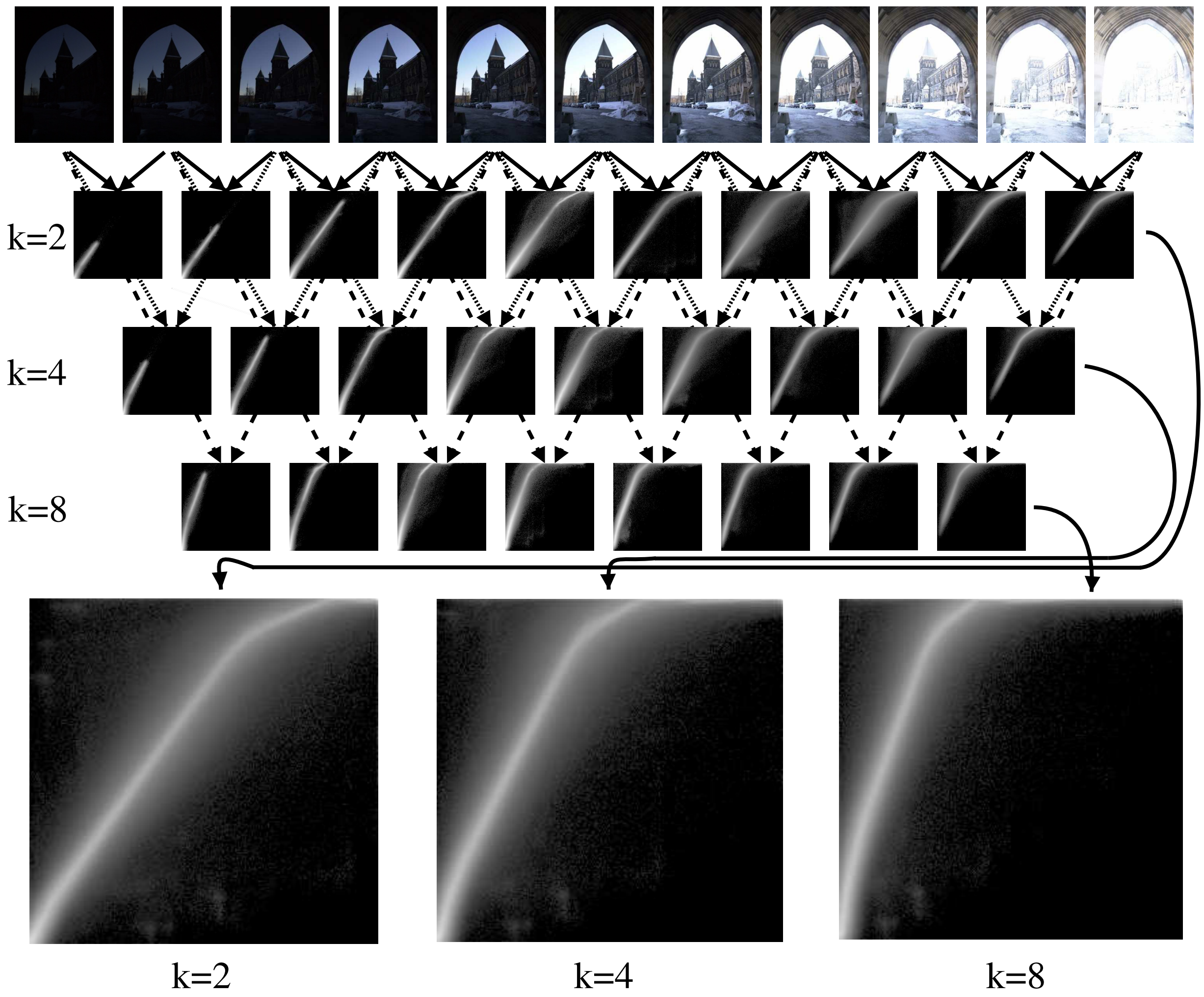


Figure 3: A sequence of differently exposed pictures of overlapping subject matter. Such variable gain sequences give rise to a family of comparagrams. In this sequence the gain happens to have increased from left to right. The square matrix J (called a comparagram) that arose from (15) is shown for each pair of images under the image pairs themselves, for $k = 2^1 = 2$. The next row shows pairwise comparagrams for skip=2 e.g. $k = 2^2 = 4$, and then for skip=3, e.g. $k = 2^3 = 8$. Various skip values give rise to families of comparagrams that capture all the necessary exposure difference information. Each skip value provides a family of comparagrams.

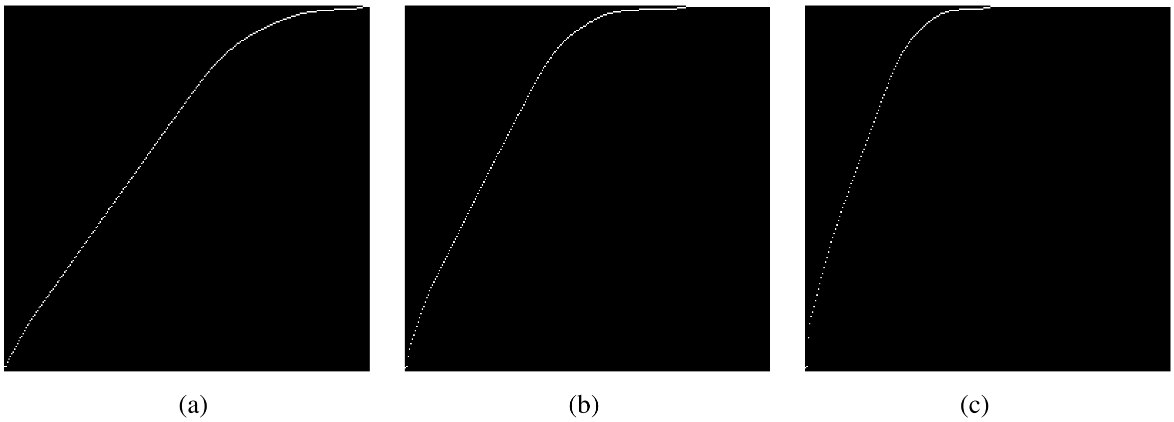


Figure 5: Verification of compargram unrolling: The response function $f(q)$ is unrolled from the comparagrams. A comparametric plot of the data for each response function is made, as ordered pairs $(f, g) = (f(q), f(kq))$ as a comparagram image with bin counts equal to zero where the comparametric equation is not true. If we compare these images with the original comparagrams, we can easily verify that the estimates of the response function must have been very close to correct. Note that this method of verification does not require any special laboratory instruments.

1.6 Estimation with unknown exposure values

Generally the relative exposure values K are not known a-priori. Thus simultaneously solving for K_i and F^{-1} is an optimization problem:

- Initially, pairwise comparagrams are constructed and unrolled, each assuming an arbitrary value of $K = 1$ (e.g. $k = 2$). The resulting F_i^{-1} function estimates for each pair of images will have the same overall shape, but each be scaled differently, apart from noise.

- Next these estimates of F_i^{-1} are used to estimate the *relative* k_1 values. Without loss of generality, $K_1 = 1$ may be assumed, and $K_i \forall 1 < i < I$ may be found.

- Next the estimates of F_i^{-1} are consolidated (averaged together) by using the above estimates of $K_i \forall 1 < i < I$ to register them. This registered and averaged $\hat{F}^{-1}$ is now used to re-estimate the K_i values, by comparing ratios:
$$\hat{K}_i = \hat{F}^{-1} - \hat{Q} \tag{18}$$
where $Q = log(q)$ is the logarithm of the reference quantity of light. (At this point, this first guess is often good enough to stop, but may undergo successive refinements by continuing as follows...)

- Finally, if desired, images are sorted in increasing order of exposure (least to greatest), and a final estimate of $\hat{F}^{-1}$ is made across all pairs of images, using the estimates of K_i.

- If desired, this process may again be repeated, e.g. using that estimate $\hat{F}^{-1}$ to again estimate K_i and so on.

1.6.1 The general case: multiple images with unknown K

It can be seen, from (12) that changing the value of K changes only the amplitude (range scaling) of $\mathbf{F}$. Therefore, $\mathbf{F}$ may first be solved using a "generic" value of $K = 1$, so that the shape of F can be found up to another single unknown scalar constant. Subsequently, when working with more than two images, a function F may be found for each image pair, and since all the estimates of F should vary only in overall amplitude, the curves may be scaled to the same height and averaged to obtain an average estimate of the general shape of F for the entire image sequence. Then the individual K values may be found by comparing each F to the averaged F.

After a crude solution found by averaging, as a first of a succession of guesses, the result may be refined by a least squares fit across all possible pairs of images. Alternatively, a combination of averaging and least squares across all pairs

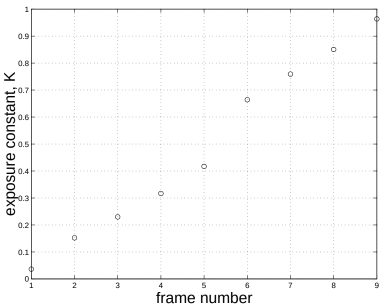


Figure 6: Estimated exposure values, K_i from nine of the eleven test images. We can clearly see, in the estimated values, where two of the images were left out of the sequence.

may be made, especially if some pairs turned out to be related by the same K values (in which case comparagrams with like K values may be averaged, to reduce the number of all possible image pairs).

Since K can only be determined up to an unknown offset and gain constant, without loss of generality, the K values are assumed to be on the interval from 0 to 1.

As an example, nine of the eleven test images were used (two were left out), and the K values returned after one iteration appear in Fig 6.

1.6.2 Weighting the solution by the *certainty* function

It is expected that pixel values near the middle of their range in the image will change more rapidly for a given change in exposure than pixel values close to the extremes. For example, when a pixel is saturated (e.g. 255 in the case of an eight bit camera) no additional light will change the pixel value. Therefore, we expect that the midtones in the image will give greater *certainty* in estimating f than will the shadows and highlights. This intuition is made formal by the so-called *certainty functions* [10][3], which are the derivatives of the response functions.

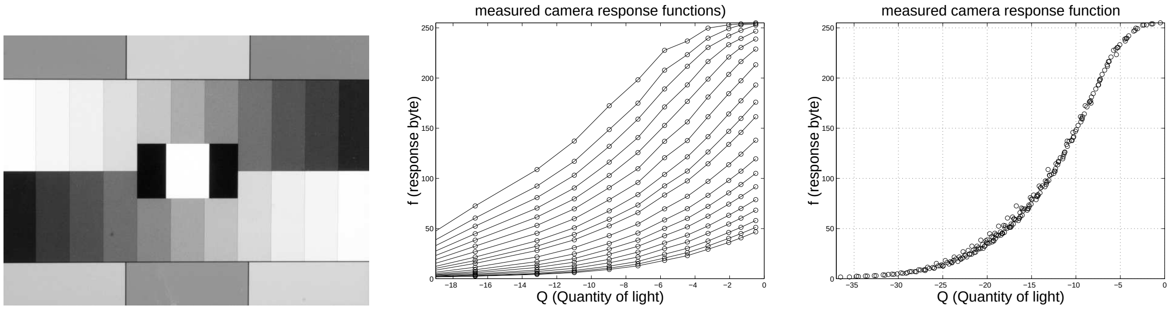
The certainty functions may therefore be used to weight the columns of $\mathbf{A}$, and the corresponding entries of the vector $\mathbf{K}$, when solving for $\mathbf{F}$:

$$\begin{aligned} \mathbf{A_w} &= \mathbf{A} \circ (\mathbf{w1}) \\ \mathbf{A_w}^T \mathbf{A_w} \mathbf{F} - \mathbf{A_w}^T \mathbf{K_w} &= 0 \end{aligned} \tag{19}$$

where $\circ$ denotes Hadamard multiplication[5], $\mathbf{w}$ is a column vector in which each entry is made up of $c(n)$ where c is the certainty function, and n is the column index in which the quantity -1 appears for that row.

In practice, we do not know the response function (this is the very entity we are trying to estimate) so we also do not know a-priori, its derivative, the certainty function, which

[5]Also known as *element-by-element multiplication* and denoted by ".*" in Octave, or nonstandard proprietary work-alikes such as Matlab.



(a) (b) (c)

Figure 7: (a) One of 19 differently exposed pictures of test pattern. (b) Each of the 19 exposures produced 11 ordered pairs in a plot of $f(Q)$ as a function of Q. (c) Shifting these 19 plots left or right by the appropriate K_i allowed them all to line up to produce the ground truth known-resonse function $f(Q)$.

we need to use in the weighting. However, we can use a certainty function that has the general shape of the derivative of a typical response function. It has been found that the solution F is not particularly sensitive to the shape of the certainty function. A Gaussian weighting is generally used. If desired, once F is found, F^{-1} can be differentiated and used as the certainty function to weight the columns of $\mathbf{A}$ to re-estimate F. This procedure can be repeated again (e.g. giving rise to an iterative estimation of F).

2 Ground truth: verifying the result

A CamAlign-CGH test chart from DSC Laboratories, Toronto, Canada, (Serial No. S009494), as shown in Fig 7(a) was used to verify the response function recovered using the method presented in this paper.

The individual bars were segmented automatically by differentiation to find the transition regions, and then robust statistics used to determine an estimate of $f(q)$ for each of the 11 steps, as well as the black regions of the test pattern. Using the known reflectivity of each of these 12 regions, a set of 12 ordered pairs $(q, f(q))$ was determined for each of the 19 exposures, as shown in Fig 7(b). Shifting these results appropriately (by the K_i values, to line them up, gives the ground-truth, known response function, f, shown in Fig 7(c).

Acknowledgments

The first author would like to thank David and Sue Corley of DSC labs for donation of the test chart. The Comparunroll package and related computer programs were written by Steve Mann, Corey Manders, Adnan Ali, and James Fung. The authors would also like to thank NSERC, CITO, and Nortel Networks.

References

[1] S.S. Beauchemin J.L. Barron, D.J. Fleet. Systems and experiment performance of optical flow techniques. *International journal of computer vision*, 12(1):43–77, 1994.

[2] S. Mann. Compositing multiple pictures of the same scene. In *Proceedings of the 46th Annual IS&T Conference*, pages 50–52, Cambridge, Massachusetts, May 9-14 1993. The Society of Imaging Science and Technology. ISBN: 0-89208-171-6.

[3] Steve Mann. Joint parameter estimation in both domain and range of functions in same orbit of the projective-Wyckoff group. pages 193–196, Lausanne, Switzerland, December 1996. Also appears in: M.I.T. M.L. T.R. 384, 1994.

[4] R. Szeliski. Video mosaics for virtual environments. *Computer Graphics and Applications*, 16(2):22–30, March 1996.

[5] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. *SIGGRAPH*, 1997.

[6] T. Mitsunaga and S. K. Nayar. Radiometric self calibration. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June 1999.

[7] S. Mann. Comparametric equations with practical applications in quantigraphic image processing. *IEEE Trans. Image Proc.*, 9(8):1389–1406, August 2000. ISSN 1057-7149.

[8] Charles W. Wyckoff. An experimental extended response film. *S.P.I.E. NEWSLETTER*, pages 16–20, JUNE-JULY 1962.

[9] Charles W. Wyckoff. An experimental extended response film. Technical Report NO. B-321, Edgerton, Germeshausen & Grier, Inc., Boston, Massachusetts, MARCH 1961.

[10] S. Mann and R.W. Picard. Being ‘undigital’ with digital cameras: Extending dynamic range by combining differently exposed pictures. In *Proc. IS&T’s 48th annual conference*, pages 422–428, Washington, D.C., May 7–11 1995. Also appears, M.I.T. M.L. T.R. 323, 1994, http://wearcam.org/ist95.htm.