Painting with Looks: Photographic Images from Video Using Quantimetric Processing

Steve Mann*
University of Toronto

Corey Manders
University of Toronto

James Fung
University of Toronto

ABSTRACT

When we ask the fundamental question "What does a camera measure?", we arrive at the concept of quantimetric imaging, which uses a new quantimetric unit, q, characteristic of a particular camera (e.g. each kind of camera defines its own quantimetric unit q based on its spectral response, etc.). Fluctuations in interframe exposures, along a sequence of images, give rise to a *comparametric* relationship between successive pairs of images. This allows us to estimate the response function of the camera (to derive the quantimetric unit q) as well as the relative differences in exposure. A new method of video image processing that exploits multiple differently exposed pictures (frames of the video sequence) of overlapping subject matter is thus possible. The method may be used 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. Since almost all cameras have an automatic exposure feature, typically center weighted, 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. Such fluctuations in gain may be used to estimate the camera's response function, to estimate exposure differences, to do quantimetric processing, as well as to obtain images having both extended dynamic range and extended dynamic domain.

Keywords

video, image processing, comparametrics, comparametric equations, multiple exposures

1. INTRODUCTION

*Correspondence author: S. Mann, Dept. E.C.E., 10 King's College Road, Mailroom S.F. B540, mann@eecg.toronto.edu Tel. (416) 946-3387; Fax: (416) 971-2326

Multimedia'02, December 1-6, 2002, Juan-les-Pins, France.

Figure 1 illustrates how a camera with automatic exposure control takes in a typical scene. In this example, we have six differently exposed pictures depicting parts of the University College building and surroundings. 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. In general, automatic exposure cameras tend to capture the same subject matter at different exposures, depending on how the pictures are framed. 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. Such registration is among that which helps make the camera become a scientfic measuring instrument, together with correction of barrel distortion and correction for darkening at the corners of the image.

2. WHAT A CAMERA MEASURES

The question "what does a camera measure?" seems to at first have a simple answer. Certainly a camera measures light, but the light measured by a camera lies across a broad spectrum.

To demonstrate how a camera works, it is possible to consider a simple one pixel camera. Once we understand what one pixel measures, it will be easier to understand what is measured by many pixels working together.

2.1 Constructing a one pixel camera as a simple experiment

Perhaps the simplest way to construct a one pixel camera is to use a light sensor. The cheapest and most common kinds of light sensors are usually photoresistors, such as the cadmium sulphide photocells found in the small controllers that automatically turn streetlights on after dark.

Cameras measure electromagnetic energy, but unlike an ideal radio receiving antenna that one might connect to a spectrum analyzer, a camera measures electromagnetic radiation only within a certain region of the electromagnetic spectrum. Typically cameras are sensitive to that which we call light, e.g. to electromagnetic energy falling close to the visible, the infrared, or the ultraviolet portion of the spectrum. Moreover, the camera's sensitivity to light is not very flat across the spectrum over which it is sensitive. Thus a camera is certainly not a radiometer, and the measurements that it makes are certainly not radiometric.

Moreover, cameras have spectral sensitivities that are quite different from that of the human eye, and are therefore certainly not photometers.

Figure 1: Automatic exposure as the cause of differently exposed pictures of the same (overlapping) subject matter: **Leftmost frames:** 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. **Middle frames:** 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. *Rightmost frames:* As we look further up and to the right, the dimly lit interior 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 (over-exposed). However, the inscriptions on the wall (names of soldiers killed in the war) now become visible. *All six frames:* In this paper we will see how the differently exposed pictures of the same overlapping subject matter can be combined to provide a true and accurate photographic quantity for intelligent vision systems, or simply to extend dynamic range and tonal definition for a variety of multimedia applications.

2.1.1 Cameras as quantimetric sensing instruments

Different kinds of cameras have different spectral sensitivity profiles, that are neither flat (radiometric) nor match the human eye (photometric), and therefore cameras are referred to as quantimetric devices [Mann, 1998]. A quantimetric measurement refers to a measurement that can be quantified in some measurable units, where the camera itself can be used as the measurement instrument. The quantimetric unit, typically denoted by the letter q, is usually made relative to some reference value, q_0, so that it can be expressed as a ratio, or as a logarithmic ratio (often in decibels).

The quantity, q, has been referred to in image processing literature as a *photoquantigraphic* quantity [Mann, 1998], or just the photoquantity (or photoq) for short. This quantity is neither radiometric (i.e. neither *radiance* nor *irradiance*) nor photometric (i.e. neither *luminance* nor *illuminance*). Most notably, since the camera will not necessarily have the same spectral response as the human eye, or, in particular, that of the photopic spectral luminous efficiency function as determined by the CIE and standardized in 1924, q is neither brightness, lightness, luminance, nor illuminance. Instead, quantigraphic imaging measures the quantity of light integrated over the full range of wavelengths, λ in the spectral response of the particular camera system,

$$q = \int_0^\infty q_s(\lambda) s(\lambda)\, d\lambda, \qquad (1)$$

where $q_s(\lambda)$ is the actual light falling on the image sensor and $s(\lambda)$ is the spectral sensitivity of an element of the sensor array. It is assumed that the spectral sensitivity does not vary across the sensor array.

It is this measurement q which shall be dealt with in the following sections of the paper. When dealing with the output of a given camera, this quantity is easier to deal with than any other derived measurement.

2.1.2 Using a photoresistor as a single pixel

Photoresistors (photocells) are devices in which resistance is a function of incident light. Typically an increase in light results in a decrease in resistance. Most are known to obey an empirical law:

$$R = R_0 q^{-\gamma}, \qquad (2)$$

where γ is usually a positive constant less than one. The fact that γ is less than one indicates that a photocell tends to compress dynamic range. Cameras also compress dynamic range in a similar way.

If the assumption is made that a camera is an instrument for converting light into numbers, a one pixel camera may be constructed from a photocell by simply connecting it to an ohm meter.

It is more convenient to consider conductance, which is the reciprocal of resistance, and to define the camera's response function, f, as conductance of the photocell. Equation (2) now becomes:

$$f(q) = \beta q^{\gamma}, \qquad (3)$$

where $\beta = 1/R_0$. The quantimetric function, f, increases with increasing quantity of light, q. Thus the ohm meter will show $f = 0 = q$, and will show larger values as f and q increase.

What is required is to determine f as a function of q (e.g. suppose that we did not already know the relationship between f and q). For the purpose of the experiment, all that is necessary is a lamp and the camera (photocell plus ohm meter), together with a piece of black cardboard that can be used to cover half of the lamp.

It is therefore possible to cover half of the lamp, and then point the lamp at the camera (sensor), and take a reading $f(q_0)$, (see Fig 2(a)), and then uncover the lamp while leaving it exactly in the same place. Uncovering the lamp exactly doubles the quantity of light received by the camera (sensor), so that we then know what $f(2q_0)$ is. (See Fig 2(b).) Although the absolute quantity q_0 is not known, the relative quantity $2q_0/q_0 = 2$ is known. The fact that the quantity of light was exactly doubled, implies that it is now possible to record sets of ordered pairs $(f(q_0), f(2q_0))$. To obtain different points (sets of ordered pairs) the lamp may be placed at different distances from the camera (sen-

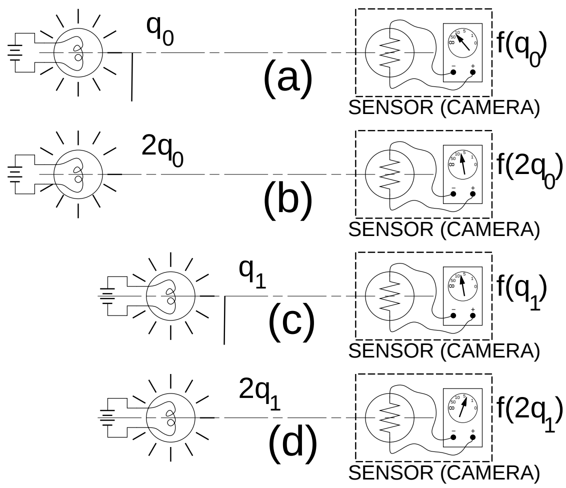


Figure 2: Photocell experiment: for each of various distances between the lamp and the camera, a reading is taken with the lamp exactly half covered, versus with the lamp not covered. Other than the camera, lamp, and black card, no other devices (such as rulers or other measurement instruments) are needed in order to characterize the response of the photocell to light.

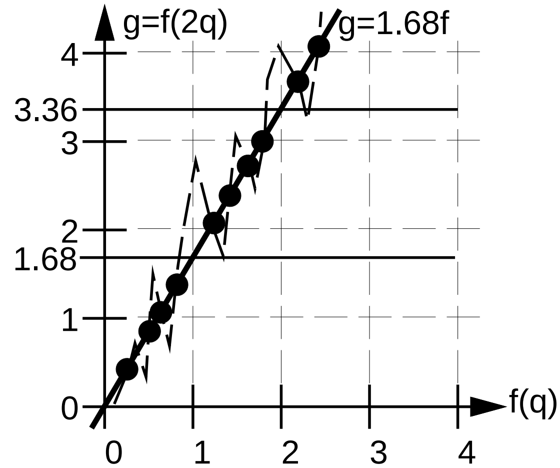


Figure 3: Datapoints from photocell experiment: ordered pairs of points taken by reading the camera with the lamp uncovered and exactly half covered.

sor) and for each such lamp position, another ordered pair will be produced.

Suppose that ten such ordered pairs of measurements are taken, e.g. continuing with $(f(q_1), f(2q_1))$ and so on, all the way up to $(f(q_9), f(2q_9))$. These ordered pairs may now be plotted on a graph, as shown in Fig 3. This is just like an $(x, f(x))$ plot, except that the axes are actually functions, $f(q)$, and $f(2q)$ instead of just scalar quantities. The first axis is f so it is convenient to use the next letter of the alphabet, g, after f in order to denote the other axis. Thus we have an (f, g) plot — a plot of a function against a plot of a function of a function, where $g = f(2q)$.

It is now possible to use a linear least squares algorithm to fit a line (or curve) through the points. The solid line shows one possible curve, namely a straight line of slope approximately 1.68. The dotted line shows another possible curve. Repeated measurements, however, lead us to believe that the relationship is simply $g = 1.68f$, as shown by the solid line. However, to prove this using this technique would imply the use of an uncountably infinite number of ordered pairs.

This (Fig 3) is simply a plot of the photocell's response function $f(q)$ against a contracted (squashed in) version of the same function $f(2q)$. Such a plot is called a *comparametric plot*.

The notion of a parametric plot is certainly a familiar concept. A parametric plot of a circle, for example, is simply a set of ordered pairs $(rcos(\theta), rsin(\theta))$. A comparametric plot is just a special kind of parametric plot, where both axes pertain to the same function at different rates.

2.1.3 Solving comparametric equations

From the comparametric plot shown in Fig 3 it has been determined that $g = f(2q) = 1.68f(q)$. This equation, $g = 1.68f$ is called a comparametric equation. In general, solving a comparametric equation $g(f(q)) = f(kq)$, for some comparametric ratio k (in this case $k = 2$) means analytically determining a family of possible functions $f(q)$ that satisfy this equation.

In the case above, therefore, what needs to be determined is what possible functions $f(q)$ give a straight line (of slope 1.68) when plotted against themselves contracted (by a factor of 2).

Fig 4 depicts two plots, one being a smooth function $f(q)$ whose comparametric plot is a straight line of slope 1.68, and the other being a quasi-periodic function whose comparametric plot is also a straight line of slope 1.68. Both of these functions have the same comparametric plot. Both are solutions to the comparametric equation $g = 1.68f$.

The quasi-periodic function illustrates that any function can be specified on, for example, the interval from q to $2q$, and then merely replicated into the interval to the right scaling by $(2*, g\circ)$, e.g. by stretching to twice the width and composing to $g()$ of the height, in this case, merely multiplying by 1.68 times the height, since in this specific case, g is linear. This recipe can be applied recursively in both the left and right directions. Therefore, in general, solutions to comparametric equations are not unique. However, it may be possible to choose a function that is semimonotonic, with semimonotonic slope, semimonotonic curvature, (and so-on, including possibly further derivatives), of which $f(q) = \beta q^{\gamma}$ is the preferred general solution to $g = 2^{\gamma} f$.

2.2 Directly solving a comparametric equation by unrolling while collecting the data

In the previous section, the response function of the photocell was known, e.g. knowing the solution of the comparametric equation, and then confirming that this known function was in fact a solution. Now suppose the solution is not known, and one did not know how to solve the comparametric equation. In this case, there exists a method of simultaneously collecting and constructing comparametric data. This method arrives at a numerical solution to the comparametric equation, namely, to obtain samples of the function $f(q)$.

Refer back to Fig 2(a) where, to obtain $f(q_0)$, half the lamp is blocked with the black cardboard. Now suppose

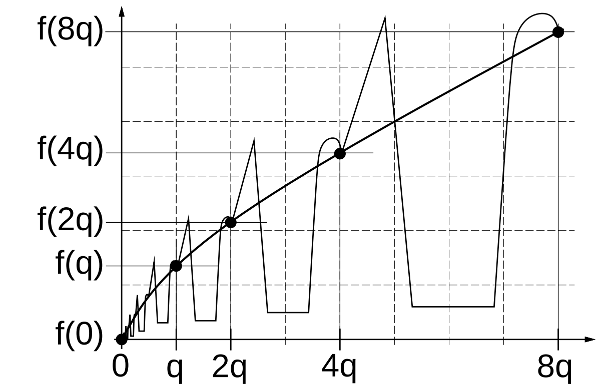


Figure 4: Two examples of possible solutions to the comparametric function $g = 1.68f$. One solution, plotted as the thick line, is given by $f(q) = q^{0}.75$. Another solution plotted as the thin line, consists of a triangular wave, square-shaped wave, and round shaped wave between q and $2q$ so that we can see how the waveform has comparametric periodicity. In each "period", the function is stretched by a factor of two as we go to the right (or contracted by a factor of two as we go to the left), and composed by a factor of $g()$ as we go to the right, or $g^{-1}()$ as we go to the left. This phenomenon is called comparametric periodicity. Thus there is a comparametric uncertainty in the solution to a comparametric equation.

we unblock the lamp, as shown in Fig 2(b) to obtain $f(2_{q_0})$. The next step is to now cover exactly half the lamp again and then move the lamp toward the photocell until the observed meter reading is exactly the same as it was when the lamp was not covered. Thus the situation as shown in Fig 2(c) will be that $q_1 = 2q_0$.

The procedure is then repeated. Uncover the lamp to obtain $f(2q_1) = f(4q_0)$. Then cover it again, and move it still closer to the sensor, until the reading is the same as it was in Fig 2(d). Then uncover the lamp again, to obtain $f(4q_1) = f(8q_0)$.

Plotting these data points will provide the eight points denoted as filled in black circles in Fig 4. Of course there still exists the comparametric uncertainty of what should be inserted between the points, but if a power law is suspected, we can assume the smooth monotonic function of the form $f = \beta q^\gamma$ and determine the value of γ from the data.

2.3 Doing the experiment in bulk

With an actual camera, there are multiple pixels, not just one. So rather than exactly doubling the exposure by using a black card to cover half the lamp, suppose that we take two pictures of the same subject matter, the two pictures differing only in exposure. Suppose that one picture is exactly twice the exposure of the other.

Pictures usually consist of a two dimensional lattice of pixels, upon which falls a two dimensional distribution of light $q(x, y)$ for the first picture, and $2q(x, y)$ for the second picture (because the exposure is twice as much for the second picture. Thus the two pictures may be written as:

$$v_f = f(q(x, y)) \quad (4)$$

$$v_g = f(2q(x, y)) = g(q(x, y)). \quad (5)$$

A typical size for a picture is an array that is 480 pixels high and 640 pixels wide (same aspect ratio as standard NTSC television aspect ratio, namely 4:3). Let us consider a simpler example, namely two pictures that are each 3 pixels high and 4 pixels wide:

$$v_f = \begin{bmatrix} 1 & 3 & 2 & 3 \\ 3 & 2 & 1 & 2 \\ 0 & 0 & 2 & 0 \end{bmatrix}; v_g = \begin{bmatrix} 2 & 3 & 3 & 3 \\ 2 & 2 & 2 & 2 \\ 0 & 1 & 3 & 0 \end{bmatrix}. \quad (6)$$

In this example, assume that we have a 2 bit camera, such that it has grey values that go from 0 to 3.

We now introduce the so-called comparagram. The comparagram is a two dimensional array of size M by N where M is the number of grey values in the first image, and N is the number of grey values in the second image, where entry $J[m, n]$ is a count of how many times a pixel in image 1 has greyvalue m and the corresponding pixel in image 2 has greyvalue n. In this case both images have 4 grey values, so the comparagram is a 4 by 4 matrix:

$$\xrightarrow{\quad g \quad}$$

$$f \big\downarrow \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 2 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

Summing across rows gives the histogram of the first image:

$$h_f = \begin{bmatrix} 3 & 2 & 4 & 3 \end{bmatrix}.$$

Summing down columns gives the histogram of the second image,

$$h_g = \begin{bmatrix} 2 & 1 & 5 & 4 \end{bmatrix}.$$

3. TYPICAL CAMERAS

Some video cameras function much like the photocell, but with a value of $\gamma = 0.54$ instead of the γ values of 0.6 to 0.9 that are typical of photocells. However, most cameras follow a more complicated law than the simple power law. In particular, various laws such as a simple 2 parameter law $(\frac{q^a}{q^a+1})^c$ have been proposed [Mann, 2001] and used in various research applications such as wearable imaging systems [Mann, 2001], computer vision, and robotics [Candocia, 2002]

Initially, no specific assumptions are made about the response function, f, and a comparagram is constructed from two differently exposed pictures of the same subject matter. (See Fig 1.) Typically, a difference in exposure will result in any two successive frames due to some amount of camera motion in the sense that most cameras have some kind of automatic exposure mechanism. Typically, due to noise in images (sensor noise, as well as the inter-frame motion artifacts, etc.), comparagrams do not provide points that only lie on a slender curve. More typically, comparagrams define a cloud of points clustered along an underlying curve. Thus slenderization of the comparagram [Mann, 2001] is a typical first step in recovery of the comparametric function, $g(f)$. (See Fig 5.)

4. CONSTRUCTING A COMPARAGRAPH FROM A COMPARAGRAM

Ideally, a comparagram is a function which plots $f(q)$ against $f(2q)$ which appears much like the solid line in Fig 3. Unfortunately, the comparagram does not usually appear as this ideal comparagraphic function. Rather, the comparagram appears as a cloud of points around this comparagraph. Three methods are commonly used for recovering the underlying comparametric function from the comparagram:

- first moments may be calculated down each of the columns of the comparagram or across each row of the comparagram. This would correspond to Bayes' Least Error if the comparagram were to be regarded as a joint probability distribution;
- a maximum likelyhood estimator can be used to pick the indices of the comparagram which have the highest values as entries. Alternatively a combination of row and column calculations could be taken;
- the comparagraph may be determined using only the marginals of the comparagram. This process is called marginalization. The comparagram is slenderized into a comparagraph by getting rid of the joint information in the distribution. This is equivalent to estimation of a joint PDF knowing only the marginals. It turns out that throwing away the joint information leads to a good estimate of the comparagraph from the comparagram. Prior to marginalizing the comparagram, it should first be processed, by thresholding and tone scale adjustment. This initial processing is important, and dramatically improves the results. Without this initial processing, marginalization is the same as histogram specification. But with this initial processing, the marginals are no longer mere histograms. We refer to the marginals of a processed comparagram as *histographs* to distinguish them from mere histograms. Typically comparagrams are displayed on a thresholded logarithmic scale. Preferably also, any entry in the comparagram that has an entry of 1, is set to 0. We have found this to be the minimum amount of thresholding acceptable.

An interactive computer program, called gunroll, has been developed (by Mann and Manders) to explore different values of the threshold and see the results in real time. This program is available for free (freesource under GNU General Public License) from sourceforge.net.

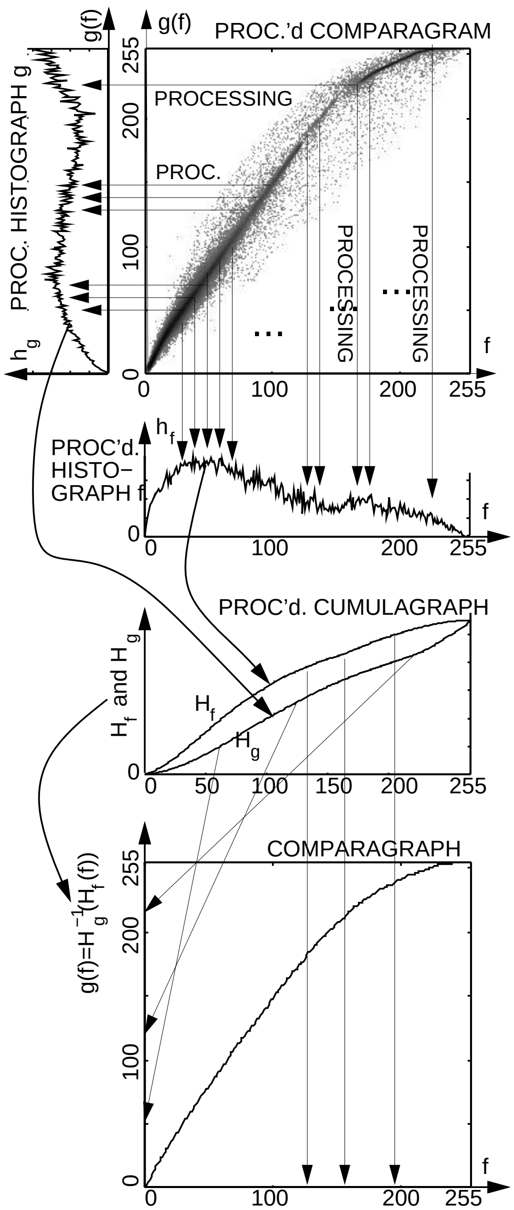


Figure 5: **Slenderizing a comparagram by marginalization:** A comparagram is constructed from two differently exposed pictures of the same subject matter. The dynamic range of comparagrams tends to be quite high, so here it is displayed on a logarithmic scale, $log(J+\epsilon) - log(\epsilon)$, $\epsilon > 0$ preventing calculation of $log(0)$ (and thresholding noise below -60dB or so). Next summing down columns produces the marginal h_f along the f axis. Summing along rows produces the marginal h_g along the g axis. Whatever tone scale is selected for the optimal display of the comparagram (typically logarithmic or power law) will affect h_f and h_g, so h_f and h_g are typically not histograms. We refer to them as *histographs* to make this important distinction. Cumulative sums are taken, giving "cumulagraphs". The comparagram is then re-constructed from only its marginals, such that only a slender ridge remains. By also constraining this reconstruction to be monotonic, the result is a graph, and is referred to as a *comparagraph* of the underlying *comparamateric function*[Mann, 2001].

5. A REAL WORLD EXAMPLE

We test our theory of comparametric imaging, by estimation of the exposures and response function from a sequence of differently exposed pictures of the same (in regions of overlap) subject matter.

Multiple differently exposed images of the same subject matter occur naturally in video sequences, where Automatic Gain Control (AGC) is present, or in cameras having some form of automatic exposure.

This naturally occuring exposure fluctuation allows us to estimate the camera's response function, as well as the exposure differences, as shown in Fig 6. The results of the exposure estimates appear in Table 1.

5.1 Fitting Comparametric Data to a Known Function

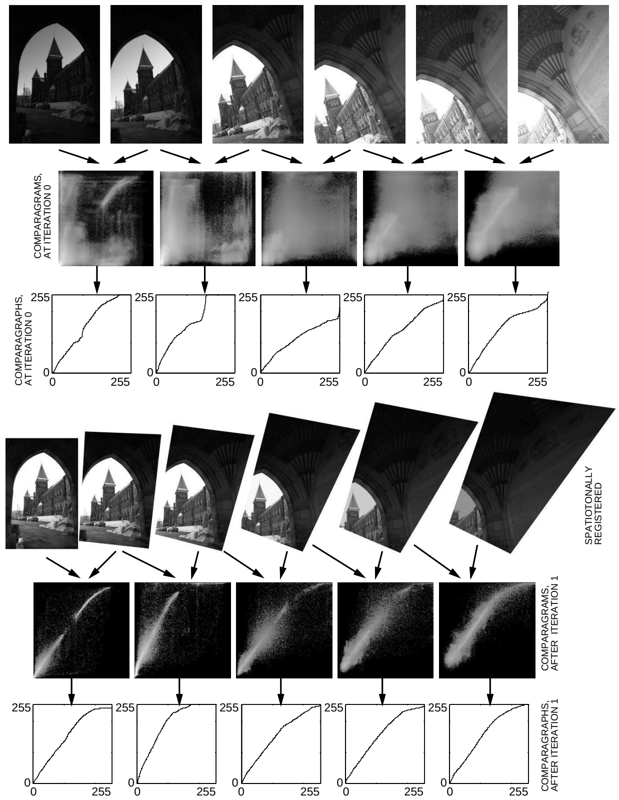


Figure 6: Variation in frame-to-frame exposure allows us to estimate the camera's response function, as well as the exposure differences. At zero iterations the estimates of the comparagraphs are quite crude, but after one iteration, they converge to satisfactory plots. Note that the second comparagraph reaches higher, because the exposure difference jumped two steps, rather than just one as was the case in the other four comparagraphs.

Iteration 0	Iteration 1	True value
2.530293	2.011216	2
3.444645	4.051448	4
2.785258	1.984015	2
2.771012	1.996743	2
2.226825	1.986862	2

Table 1: Results of comparametric interframe exposure estimator applied to the six frames in the image sequence of Fig 6. Although the results are initially crude, after one iteration, the error falls to well under one percent for all five interframe estimates.

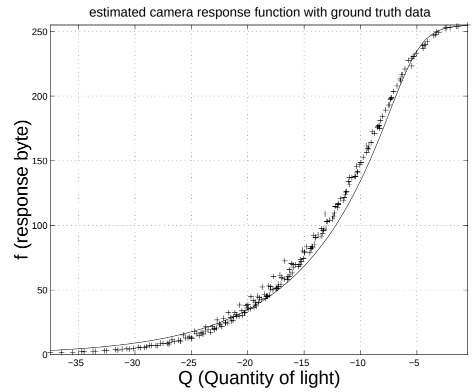


Figure 7: Using a non-linear solver to attain values for a and c to the model function $f(q) = \left(\frac{e^b q^a}{e^b q^a+1}\right)^c$

It has been shown [Mann, 2001] that the function

$$f(q) = \left(\frac{e^b q^a}{e^b q^a + 1}\right)^c \tag{10}$$

is able to model response functions from a variety of cameras. Not only does the function have the general shape of camera response functions, it has a shape which given reasonable parameters is able to fit data collected in sensor experiments from several studies.

Using 10, we may use it and it's inverse

$$f^{-1}(p) = \left(\frac{\sqrt[c]{p}}{e^b(1-\sqrt[c]{p})}\right)^{\frac{1}{a}} \tag{11}$$

to find optimum values of a and c using a non-linear solver. This may be accomplished by first asking "what quantimetric value q will produce a pixel value of x", where x is a pixel value normalized to the range $[0, 1]$. This may be done for pixel values $[0, 255]$. Knowing q for each value (by use of f^{-1}), it is now trivial to multiply the values of q by k and produce the comparagram which results from a and c as parameters. Using a slimmed version of the comparagram, the derived comparagram may be subtracted from the true comparagram by treating each element as a vector in $\mathbb{R}^{255}$. Taking any p norm of the resulting vector gives a value which may be minimized by a non-linear solver (on the values a, c, and possibly k).

This procedure used on data collected from a Kodak DC260 camera produced the approximation to the response function shown in Fig 7 where the method is plotted against ground truth data.

5.2 Estimation using Log Unrolling of a Comparagram with Spline Interpolation

An alternative method of obtaining a closed-form solution to the response function of a sensor is demonstrated.

If a comparasum (sum of registered comparagrams over various exposure ranges) is considered, an approximating function is constructed. The function ideally returns the mean $f(kq)$ for any $f(q)$ at a discrete point on $[0, 255]$ and interpolates sensibly for any intermediate value. A close approximation to an arbitrary comparagram is found as follows. First, bin counts lower than a given value are disregarded (to remove outliers resulting from sensor noise, JPEG compression noise and the like). Then first moments are taken for each column of the comparagram. Finally, not-a-knot splines [de Boor, 1978] are used across each column. For the process which follows, it is helpful to have a continuous function. Splines give the required continuity and are a sensible method for finding intermediate values of $f(kq)$. For simplicity, the continuous function which approximates the comparagram shall be $\phi(q)$ where q is a value from $[0, 255]$.

Having a function which approximates the comparagram, some assumptions are made about the underlying response function which gave rise to the comparagram. Recall that we generally assume that the response function is such that $f(0) = 0$. We also assume that $f(1) = 255$, such that if we normalize the function, we may assume $f(1) = 1$ (the maximum pixel value of an image). However, to simplify computations, this normalization will not be applied until the response function has been determined. Since the quantity of light q may be rescaled later, an assumption shall be made that $\hat{f}_1(1) = \epsilon$ (this is the first approximation to the response function f). In practice, ϵ will be some small number greater than machine epsilon. Using the approximation to the comparagram, ϕ, the comparagram may be unrolled in the forward direction. Remembering that the comparagram is a plot of $f(q)$ against $f(2q)$, it is now possible to find $\hat{f}_1(2)$ by evaluating $\phi(\epsilon) = y_1$. The function may continue to be unrolled in the forward direction. To find $\hat{f}_1(4) = y_2$, ϕ may be evaluated at y_1. If the comparagram is monotonic (which it should be given the physical applications), then we may continue in this manner until $\phi(y_n) \geq 255$. This will imply $\hat{f}_1(2^{(n-1)}) = 255$ for some $n \in \mathbb{N}$. Since the function ϕ is itself an approximation to the comparagram which has included sensor noise, compression noise, etc., $\hat{f}_1$ shall be the notation for the function which approximates the true camera response function f, using the forward unroll procedure.

Since this unrolling only gives discrete points, to achieve a continuous and differentiable approximation, cubic splines may be used to interpolate between the unrolled points.

One downfall of the method is that if we have a small error in the approximation of the comparagram due to noise, compression, etc., then this error will propagate through the calculation. One method to counter this error is to find the point in the comparagram which first attains 255. The value for which this first occurs shall be 2^n for some γ_1 such that $\phi(\gamma_1) = \phi(2^n) = 255$ where n is to be determined later. If the assumption is made that $\hat{f}_2(2^n) = 255$, then $\hat{f}_2(2^{n-1}) = \gamma_1$. $\hat{f}_2(2^{n-1})$ may be recovered by locating γ_2 such that $\phi(\gamma_2) = \gamma_1$. Corresponding to what has been done in the forward unroll, we shall end the procedure at some point when $\phi(\gamma_m) \approx 0$. We may then assign the value of n appropriately (where $\hat{f}_2(1)$ is the smallest non-zero value achieved during the reverse unroll and $\hat{f}_2(2^n) = 255$).

It is natural to expect that the forward unroll will be accurate for values close to 0, but will lose accuracy as $\hat{f}_1(x)$ approaches 255. Similarly, it is expected that the reverse unroll will be accurate for $\hat{f}_2(x)$ which yields values close to 255, but will degrade for values of x such that $\hat{f}_2(x)$ is close to 0.

At this point, both approximations were normalized so that the domain of both f_1 and f_2 is $[0, 1]$. Both the forward and reverse results were put into a logarithmic scale, and once again splines were used to achieve a continuous and

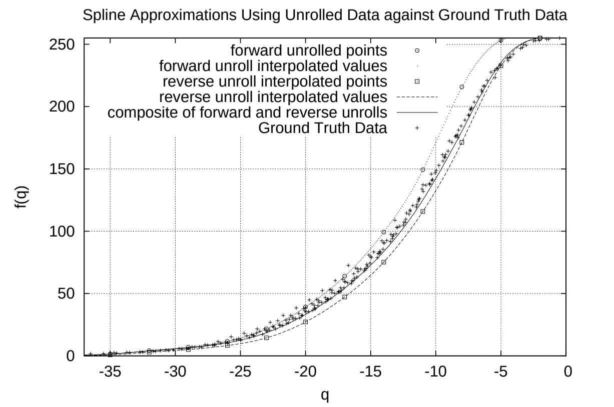


Figure 8: The results of the forward and reverse unrolled data compared to known ground truth data. The ground truth data was recovered by images taken of patterns of known reflectivity (DCS Labs test pattern) by the same camera from which the comparagram was derived.

differentiable approximation to the true camera response function. To achieve a compromise between the two approximations which benefits from the strength of both solutions, both approximations were combined into a single approximation which was weighted to transition from $\hat{f}_1$ for values close to 0, shifting to $\hat{f}_2$ as the function approaches 255.

The results of this procedure are shown in 8, where they are compared to the camera function as recovered from use of a test chart containing a pattern of known reflectivity (ground truth).

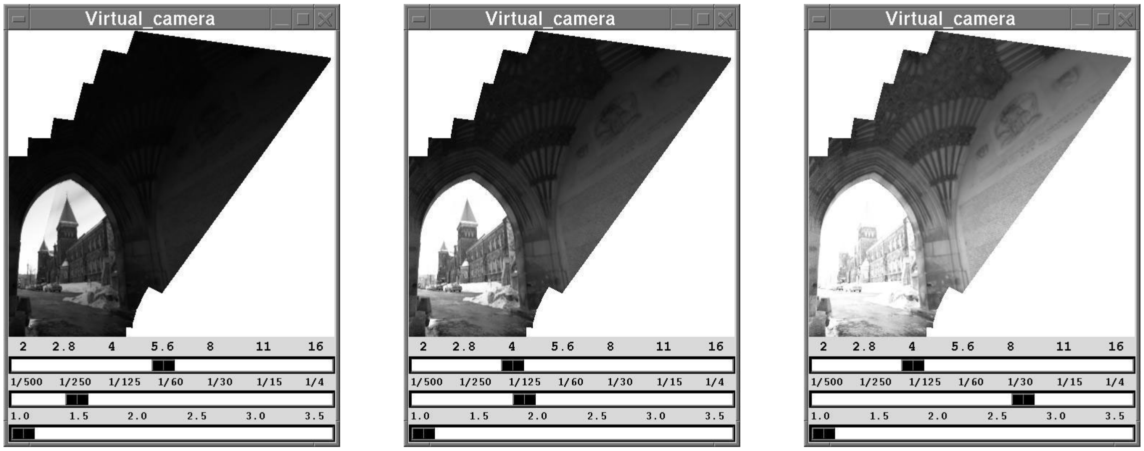
6. MULTIMEDIA IMAGING

6.1 Composing Images with Certainty Functions

The inverse camera response function f^{-1} can be used to convert pixel values into photoquantities, and vice-vera (with f).

Once the camera response function is estimated, along with the projective coordinate transformations between successive frames of video, the images may be brought together into a common coordinate space, as shown in the fourth row of the six rows in Fig 6. Each image may be brought into the same photoquantimetric range by multiplicative scaling by an appropriate scalar exposure constant k. Each image provides an estimate of q in areas of overlap. Looking at Fig 8, we see that the camera is most accurate as a photoquantimetric tool in the middle of its range where the derivative of the response function is relatively high. The derivative of the response function is known as the *certainty function* [Mann, 2001]. Thus to combine two or more images, a logical method is to multiplicatively scale each image to that of one reference frame, then assign a quantimetric value given appropriate weightings of certainty. This procedure will have the effect of favoring photoquantimetric values where the camera was most accurate at collecting data. Consequently this gives lower significance to data which was collected in less responsive areas. This gives us the following equation for a particular photoquantimetric value:

$$X = \frac{\sum_{s=1}^{d} \omega_{s_X} q_{s_X}}{\sum_{s=1}^{d} \omega_{s_X}} \tag{12}$$



(a) (b) (c)

Figure 9: The virtual camera shows an image composite having all the information from the images combined together, so that the exposure settings of the camera can be set retroactively (e.g. after the picture is taken). Thus at picture taking time, we merely measure the quantities of light, and then render these into an image at the desired exposure later. (a) If we desire to see outside through the archway, we can adjust the exposure settings on the virtual camera appropriately. These settings are roughly equivalent to what the automatic camera exposure would select, when pointed out through the open doorway. Here we can see the buildings outside, together with the needle-like object to the left of the buildings (Toronto's CN Tower). (b) An intermediate exposure shows some detail in the bright exterior, and some of detail in the darker interior, but neither is ideal. (c) The exposure may be adjusted appropriately to show the textual inscription on the wall inside. This exposure is roughly equivalent to that of the automatic camera when pointed primarily at the text on the wall.

Where q_{s_X} is the photoquantimetric value of lightspace image s at location $X = (x, y)$ given d lightspace images. And ω_{s_X} is the certainty of the photoquantimetric value at that location.

6.2 Viewing High-Dynamic Range Images

After composing the images, we have a single array of immense dynamic range (as well as range resolution) and spatial extent (as well as spatial resolution). To display such an image on a printed page, we may wish to reduce the dynamic range down to that of typical print media, which we can do by presenting the data on a $Q = log(q)$ scale, as shown in Fig 10. Similarly, a "virtual camera" program may be used to display photoquantimetric maps at different camera f-stops retroactively as shown in Fig 9.

Now we seek to produce an image that shows the entire dynamic range on a high contrast, visually appealing tone scale. To achieve such a dynamic range compression, we relax the *monotonicity constraint*[Mann, 2001], globally, while maintaining it locally.

6.3 Mean Regression

The first step in producing a tonally appealing image is to shift the mean quantimetric value of local regions towards a desired global mean. This is done in order to reduce the maximum dynamic range so that the image lies within some displayable range of values, while still preserving local detail. It is also desired that brighter regions maintain a higher mean quantimetric value than darker regions, so that the image retains the *impression* of global monotonicity (even though global monotonicity may be violated).

Thus, regions of high quantimetric mean have their mean value reduced, while regions of low quantimetric mean have their means increased. In this fashion, local means are moderated toward the desired global mean. This shift must also be dependent upon on the original regional mean. For instance, if all means are moved to a single midrange value, the image will have an undesirable flat grey quality. Thus it is desired that regions which are extremely light or dark have their means shifted proportionally less to keep them proportionally lighter or darker.

Figure 10: Cemented video frames of Fig 1 after being spatiotonally aligned, are now displayed on a logarithmic Quantimetric scale, Q. Here the entire dynamic range is visible, but the image lacks the pleasing effect of high contrast, which we often desire in images.

To achieve this, the local means are *regressed* towards some midrange value. This is regression in the true sense of the word since it moderates local means by regressing them towards a global mean[Galton, 1886]. The formulation for the regression used was:

$$\alpha = (\frac{\mu_g}{\mu_l})^a \tag{13}$$

where μ_g is the global mean, μ_l is the local mean, and $0 < a < 1$ is a scaling constant which affects the severity of the regression. α is then used as a constant multiplier on the quantimetric each of the values in the local region, so α regresses the local mean μ_l towards the global mean μ_g.

The fraction μ_g/μ_l determines the amount of shift by comparing the local mean with global mean. When $\mu_l > \mu_g$, the fraction $\mu_g/\mu_l < 1$ and so $\alpha < 1$ which means the local mean will be shifted downwards. Similarly, when $\mu_l < \mu_g$, the fraction $\mu_g/\mu_l > 1$ and the local mean is shifted upwards. The constant a affects the severity of the shift. A typical value for a is $a = 1/2$, which takes the square root of μ_g/μ_l. This moderates the amount of shift by making the extreme light and dark regions shift proportionally less. Local means which are further from the global mean are thus shifted proportionally less. It was found that this type of mean regression was able to bring all regions of the image into a displayable dynamic range while preserving the impression of global monotonicity.

6.4 Variance Scaling

After the local means have been regressed to the global means, all the light and dark regions are visible and have detail present. However, the local detail is not very pronounced nor does it exhibit much local dynamic range. To enhance local details after the mean regression, the local image variance is increased. The variance is increased about the regressed mean so that the impression of global monotonicity is preserved even though local detail will likely violate global monotonicity.

A dilation factor to adjust local variance was calculated as:

$$\beta = \beta_m \log(\frac{\sigma_g}{\sigma_l}) + \sigma_a \qquad (14)$$

where σ_l is the local variance, σ_g is the maximum variance found in any of the local regions of the image, σ_a and β_m are constants. β is then used as a constant multiplier on the pixel variance (about the shifted mean) to increase the local variance.

The logarithm is used to properly scale the large range of image variances. In some areas the variance is sufficient before processing and so no dilation factor is used. σ_a is a constant, below which no dilation is applied ($\beta = 1$), and above which, the local variance is multiplied by $\beta > 1$. Thus, σ_a defines a smooth threshold above which variances are increased. β_m is a constant which determines the maximum amount of variance dilation which can occur in any region.

The result is shown in Fig 11

6.5 Determining and processing of neighborhoods of large dynamic range

Having regressed the means and expanded the variances, there are cases where adjustment may need to be moderated. In particular, local regions which contain large differences in photoquantity will result in neighborhoods of high variance. The methods used to scale the means and variances in these local regions tend to be particularly sensitive to local variance. Specifically, even small changes in variance in such an area produces an "over-processed" look. The first challenge in dealing with this problem is to determine local regions which possess this characteristic. In the images presented up to this point in the paper (such as Fig 6), any local region around the archway will demonstrate this phenomena. Similarly, neighborhoods around the border of the sky and the top of building to the right will possess this troublesome property.

To detect a high variance region which contains sharp differences in lighting, certainly many techniques may be used. Among these are harmonic analysis, Fourier analysis, etc., but are typically computationally expensive. This fact is made worse by the high resolution needed to produce superior quality images as well as the floating point precision needed to store the high dynamic range lightspace images. For this reason, one logical approach is to calculate a *general non-uniformity factor* for each local neighborhood. To be more specific, a local neighborhood shall be an $n \times n$ block, calculated at intervals of one pixel in both the vertical and horizontal directions for the entire image. For each $n \times n$ region, first order spatial gradients are computed, and the largest spatial gradient in the block (i.e. the largest pixel difference in either the horizontal or vertical direction) is compared to a threshold value. If the largest spatial gradient is greater than the given threshold, the block is accordingly characterized as non-uniform. The factor by which the gradient is greater than the threshold is similarly characterized thus accounting for its description.

A slightly more computationally expensive method would be to find subsections of the lightspace image which differ greatly in photoquantity. Given that such subsections may be determined, the question still remains as to what factor this should effect the corresponding calculations and similarly how the size of the the subsections should factor in.

Given that non-uniform regions can be consistently determined, we may then deal with the region in a variety of manners. One possible method of dealing with the region is to vary the block size (or a generalized local neighborhood) upon which compressive adjustments are made, thereby reducing the local variance. Any method chosen to act on a region in lightspace will thereby be tempered in a logical manner. Alternatively, we may calculate a general non-uniformity factor and directly apply it to either a mean-shifting method, a variance shifting method or both.

7. CONCLUSION

Asking the fundamental question "what does a camera measure", provided us with a new conceptual framework upon which to understand the processing of image sequences. In particular, fluctuations in interframe exposures, among multiple pictures of the same subject matter, allowed us to estimate the camera response function f as well as the relative differences in exposure, k.

A key concept is that of comparagram slenderization to recover the comparametric function (plotted as a compargraph through the cloud of data points in the comparagram).

Two methods were described for estimation of the the unknown camera response function and the spatiotonal transformation between images in a sequence of differently exposed overlapping subject matter. Either of these two methods may be applied directly to the comparagram, or to the comparagraph. It was found that results were improved by slenderization (e.g. applying the methods to the comparagraph rather than the comparagram).

The first method provided the simplest description, namely a two parameter closed-form response function. Although providing a good fit to the ground-truth data, that can be

Figure 11: Cemented video frames of Fig 1 after being spatiotonally aligned, are now dynamic-range compressed in a way that maintains high contrast by relaxing the *monotonicity constraint*[Mann, 2001] so that the entire dynamic range is visible and the contrast is high, resulting in a visually pleasing image, typical of what might be desired in many multimedia applications.

described with just two constants that fully characterize a given camera, there may be situations where an extremely accurate fit is required. The second method therefore uses a more sophisticated model than the first method, and is thus more accurate.

These methods were also shown to be useful for the generation of a virtual camera in a multimedia application. The recovered camera response function may be used to bring the images into a common quantimetric range. The methods presented were able to spatiotonally align images to produce a composite image which captured the favorable properties from each image in a logical manner. This final image is in fact as good as the original images in regions of non-overlap, and superior in areas where overlap does occur.

Acknowledgments

Thanks to NSERC, CITO, Altera, Xilinx, DSC labs, and Canada Council for the Arts.

8. REFERENCES

Candocia, F. M. (2002). Synthesizing a panoramic scene with a common exposure via the simultaneous registration of images. *FCRAR*. avail. at http://iul.eng.fiu.edu/candocia/Publications/Publications.htm.

de Boor, C. (1978). *A Practical Guide to Splines*, volume 27 of *Applied Mathematical Series*. Springer-Verlag. ISBN 0-387-90356-9.

Galton, F. (1886). Regression towards mediocrity in hereditary stature. http://www.jmpdiscovery.com/news/jmpercable/ summer98/regression.html.

Mann, S. (1998). Humanistic intelligence/humanistic computing: 'wearcomp' as a new framework for intelligent signal processing. *Proceedings of the IEEE*, 86(11):2123–2151+cover. http://wearcam.org/procieee.htm.

Mann, S. (2001). *Intelligent Image Processing*. John Wiley and Sons. ISBN: 0-471-40637-6.