Extreme-Dynamic-Range Sensing: Real-Time Adaptation to Extreme Signals

Ryan Janzen
University of Toronto

Steve Mann
Stanford University

Sensors—whether they sense light, vibration, temperature, or other phenomena—have a certain range over which they can sense reliably. The ratio of the largest magnitude signal that can be sensed accurately, to the magnitude of the smallest signal that can be sensed at all, is called dynamic range.

All sensors have a limited dynamic range. A physical phenomenon is sensed only as weakly as a certain minimum perceptible difference, up to a maximum amplitude or intensity. Digital sensors have a further limitation—quantization—where signals can only be expressed by a limited set of discrete numbers. Herein, again, signals can only be sensed over a limited dynamic range.

The dynamic range of a sensory process can be improved using high-dynamic-range (HDR) compositing, where one physical phenomenon is simultaneously sensed multiple times. In dynamic range compositing, multiple exposures (images, audio samplings, or so on) of a physical phenomenon can be taken at different exposure settings, either using a single sensor that switches between gain settings in a sequence, or using an array of differently configured sensors. Finally, the data is composited into an output that covers a wider dynamic range than that of one single capture of the signal (see Figure 1). In this way, it is possible to overcome the limited dynamic range of a camera, audio recorder, or other sensor.

Here, we present a system that can control the sensor itself—dynamically adjusting exposures or gains using negative feedback (see Figure 2), so that HDR compositing can respond in real time to changing conditions. Such dynamic dynamic-range (D2R) compositing can address extreme-dynamic-range situations, such as satellite imagery of the sun against stars, arc-welding diagnostics, or seismic waves capturing an insect’s footsteps during an earthquake.

Creating Composite-Dynamic-Range Signals

There are several well-known methods for combining exposures (outputs of sensors) into an HDR signal with ordinary HDR compositing. Typically, a camera’s nonlinear response function is determined, reversed for each exposure, and, based on the result, each pixel in each exposure can be weighted according to the precision (degree of certainty) each exposure’s pixel gives to the combined measurement. Pixels whose values are near the extrema (either cut off near zero or saturated near the maximum value) are given the lowest possible weighting.

For example, if an image exposure has any pixels that are completely white—saturated at the camera’s maximum possible value, as shown in Figure 1—then those pixels should be ignored, because they give insufficient information about exactly how bright the physical light really was. Similarly, completely black pixels are ambiguous in not revealing exactly how dim the physical light really was, below the camera’s minimum cut-off. As long as another exposure exists with well-exposed pixels having medium (gray) values, then those pixels contain more useful information that should be incorporated in the final composited image. A composite-dynamic-range (CDR) signal can then be reconstructed, after compensating for the different signal strengths entering the sensor for each exposure.

Audio-based HDR compositing generalizes HDR to other time-varying signals (see Figure 3).

Choosing Exposure Settings

Choosing exposure (input gain) settings wisely is critical to the HDR compositing process.
This research paper was published in the journal *IEEE MultiMedia* in 2017.

The initial paper won an award at IEEE ISM 2016, leading to the *IEEE MultiMedia* article. Additional details can be seen in the initial paper.

For copyright reasons the publisher has requested that the plain-formatting version be distributed, as seen below.
Extreme-Dynamic-Range Sensing:
Real-Time Adaptation to Extreme Signals

Ryan Janzen and Steve Mann
University of Toronto
http://eyetap.org/sensing

Abstract—The new concept of coupled dynamic dynamic-range (D2R) compositing operates by assembling sensor information, such as images or audio, from multiple “strong” and “weak” samplings or sensor snapshots, whose sensitivities drift and change over time, as lighting conditions or sound conditions change over time in their amplitude-domain properties. We have introduced a feedback-control method to automatically adjust multiple exposure settings for compositing, to increase the dynamic range of a sensory process such as video capture. The method uses a cost function to express measurement uncertainty, salience detection, and a “cross-uncertainty” metric between different exposures, fed into a dynamic control system. The system responds in real-time to changing ambient conditions and sensor motion, asymptotically tracking the exposures to minimize uncertainty, to capture an extremely high dynamic range for compositing.

Keywords—high dynamic range, HDR sensing, HDR video, HDR audio, HDR welding video, composite dynamic range

I. INTRODUCTION

Sensors — whether they sense light, vibration, temperature, or other phenomena — have a certain range over which they can sense reliably. The ratio of the largest magnitude signal that can be sensed accurately, to the magnitude of the smallest signal that can be sensed at all, is called dynamic range.

All sensors have a limited dynamic range. A physical phenomenon is sensed only as weakly as a certain minimum perceptible difference, up to a maximum amplitude or intensity. Digital sensors have a further limitation—quantization—where signals can only be expressed by a limited set of discrete numbers. Herein, again, signals can only be sensed over a limited dynamic range.

The dynamic range of a sensory process can be improved using high dynamic range (HDR) compositing, where one physical phenomenon is simultaneously sensed multiple times. In dynamic range compositing, multiple exposures (images [1], audio samplings [2], etc.) of a physical phenomenon can be taken at different exposure settings, either using a single sensor that switches between gain settings in a sequence, or using an array of differently-configured sensors. Finally the data is composited into an output that covers a wider dynamic range than that of one single capture of the signal (see Figure 1). In this way, it is possible to overcome the limited dynamic range of a camera, audio recorder, or other sensor.

Here, we present a system that can control the sensor itself—dynamically adjusting exposures or gains using negative feedback (see Figure 2), so that HDR compositing can respond in real time to changing conditions. Such dynamic dynamic-range (D2R) compositing can address extreme-dynamic-range situations, such as satellite imagery of the sun against stars, arc-welding diagnostics, or seismic waves capturing an insect’s footsteps during an earthquake.

II. CREATING COMPOSITE-DYNAMIC-RANGE SIGNALS

There are several well-known methods for combining exposures (outputs of sensors) into an HDR signal with ordinary HDR compositing [1], [3], [4], [5], [6], [7], [8], and [9]. Typically, a camera’s nonlinear response function is determined, reversed for each exposure, and, based on the result, each pixel in each exposure can be weighted according to the precision (degree of certainty) each exposure’s pixel gives to
the combined measurement [1] [4]. Pixels whose values are near the extrema (either cut off near zero or saturated near the maximum value) are given the lowest possible weighting.

For example, if an image exposure has any pixels that are completely white—saturated at the camera’s maximum possible value, as shown in Figure 1—then those pixels should be ignored since they give insufficient information about exactly how bright the physical light really was. Similarly, completely black pixels are ambiguous in not revealing exactly how dim the physical light really was, below the camera’s minimum cut-off. As long as another exposure exists with well-exposed pixels having medium (gray) values, then those pixels contain more useful information that should be incorporated in the final composited image. A composite-dynamic-range (CDR) signal can then be reconstructed, after compensating for the different signal strengths entering the sensor for each exposure.

Audio-based HDR compositing [2], generalizes HDR to other time-varying signals. See Figure 3.

A. Choosing Exposure Settings

Choosing exposure (e.g. input gain) settings wisely is critical to the HDR compositing process. Without carefully choosing the gain settings of the exposures, uncovered gaps might emerge in the dynamic range; in turn, unknown indeterminate samples or pixels can result due to a lack of information about the quantity of light at those locations. This can occur if there are insufficient exposures, or if the specific exposures are not chosen properly.

A common practice for setting exposures is rather unscientific: planning for a specific dynamic range capability and then trying a number of different exposure settings until satisfied by visual inspection of the output image. Visual inspection (or “eyeballing”), improvising, or guessing all have the limitations of human intervention and can be dependent on visual cues such as terraced lines in HDR output images.

We (and others) saw a need for an automated method, so in earlier work we developed a static exposure optimization method for HDR compositing of time-varying signals [2]. We found a set of constraints used to control exposure settings, based on the properties of a time-varying signal such as light or sound. Our method used an “exposure packing” dynamic range to compute the values of exposure gains, but at the time it did not dynamically control the exposures.

B. Understanding the Evolution of Exposure Control

Automatic gain control (AGC) is one method for dynamically adapting to different signal strengths (see Figure 4A). However, there are two flaws with ordinary AGC: the dynamic range is only matched to one amplitude regime at a time (very strong and very weak signals cannot be sensed simultaneously), and information is lost about the original signal strength.

We could call this “Generation-1 AGC”, and suggest a “Generation-2 AGC”, which records the exposure gain at all points in time, to be able to recover the original signal strength from this recording, after-the-fact, at each point in time. Still, this does not help in sensing both strong and weak signals simultaneously, such as direct sunlight and dim shadows in the same image.

Automatic exposure bracketing (AEB), or auto-bracketing, uses a similar principle to ordinary AGC, but chooses a set of exposures with fixed ratios between them. Cameras currently on the market typically implement auto-bracketing using AGC to find one middle exposure, and using a pre-set ratio to choose two or more other exposures at a fixed ratio away from the middle exposure [10] [11] [12].

Generation-3 AGC could be defined as multiple separate AGC units, each working separately from each other as if they were different cameras, audio recorders, or other unconnected sensor exposures [13] [6] [14].

In a paper presented at the 2016 IEEE International Symposium on Multimedia (ISM 2016) [15] we proposed a Generation-4 AGC, where a dynamic coupled control system adapts to signal conditions, and jointly controls a vector of exposure values. This coupled dynamic-range (CDR²) system built on previous work by not only merging dynamic-range data in real-time, but also dynamically and jointly controlling a vector of exposure settings themselves in real-time. This article, an extended version of our ISM paper [15], includes new developments in salience metrics for exposure array control to reduce the required computation on each data frame.

III. DYNAMIC²-RANGE COMPOSITING

A key limitation of previous work has been its lack of adaptation to time-varying conditions. For example, an HDR video processor might use two exposures that happen to capture a complete dynamic range in an office setting, but may be woefully inadequate when brought to a welding facility with low light levels plus extremely bright light at small points where the welding is taking place. As noted in Figure 1, the result can be portions of an image that are all black, while the welding work is seen clearly (the case of dark exposures), or can be a completely white, saturated, indeterminate region where the welding takes place, while the rest of the room is visible (the case of bright exposures). This static nature of HDR exposures is illustrated in Fig. 4B.

Instead, we wish to make a dynamic (Figure 4C) and coupled (Figure 4D) control system to control dynamic ranges of each exposure, through the exposure settings. Adding a feedback loop allows the system to asymptotically approach an optimum and also adapt to time-varying lighting conditions (in
the case of a camera) in concert with the particular nonlinear response of the camera. Furthermore, we perform joint control of the exposures, based on joint metrics of their real-time response to light.

A. Exposure Parameters, and Re-Alignment of Scales

The exposures are controlled with a time-varying parameter, $\chi_e(t)$, individually for each of the exposures (numbered $e = 1\ldots E$). Exposure variation in a camera, for example, involves changing one or more of the following settings:

- exposure speed;
- aperture size (F-stop);
- sensor sensitivity (commonly called “ISO”);
- spatial light modulator (SLM) attenuation; or
- any other control input that affects the dynamic range or dynamic level of sensitivity.

Audio-based exposure control involves varying one or more of the following:

- gain value of an amplifier before the sensor;
- gain value of the sensor itself; or
- any other control input that affects the dynamic range or dynamic level of sensitivity.

Exposures are captured as signals, and these signals are realigned according to their relative original physical scales, to prepare them for compositing.

For camera exposures, for example, scale realignment requires us to first overcome the nonlinear response of a typical camera. The tonal range of the image should be converted into an equivalent tonal range of physical light levels, for each pixel. The available tonal range given from an image is referred to as imagespace [16], typically having pixel values between 0 to 255 for an 8-bit image, across red, green, and blue channels. A nonlinear model of the camera’s response function (determined offline by calibrating initially) is used to convert each pixel value into an estimated true quantity of light. This true, physical range of values is referred to as lightspace [16].

Audio systems, on the other hand, are designed to be as linear as possible in their response, unlike a camera. In this case, the exposure setting is a linear gain value, $\chi_e = \alpha_e$, and the dynamic ranges of the exposures can be re-aligned by a factor of $1/\alpha_e$. However, audio HDR requires exposure weighting functions that are designed for the harmonic nature of audio signals [2]—unlike image-based HDR, where the harmonics of each pixel typically do not need to be considered. (Audio HDR compositing requires additional exposures to overcome a phenomenon we call eclipse, where one frequency range at one dynamic level masks another frequency range at a lower dynamic level—that is, cross-effects in the amplitude-frequency plane.)

B. Optimization Cost Functions

Imagine if HDR exposures were chosen such that each of the dynamic ranges neatly abutted against each other, so that one exposure saturates exactly at the point when the next exposure barely senses a signal. Is that sufficient?

If the sensor’s response function, $f(q)$, has low-precision anywhere in its dynamic range $[q_{\text{min}}, q_{\text{max}}]$, then a greater amount of overlap with another exposure is desirable at that point in the dynamic range.

First, we define an output value observed from the sensor:

$$\rho = f(q)$$

The original physical quantity, $q$, might be at a variety of different levels (bright v.s. dark, or loud v.s. soft), but the only observable from the sensor is its output signal, $\rho$. (Indirectly, then, a sensor’s nonlinear response function, $f$, can be determined indirectly by testing its output using comparametric equations [3]. Without needing a precisely-controlled test signal for $q$, this indirect method lets us use available data from different exposures, compare nonlinear responses between the scaled-$q$ inputs as $g(f(q)) = f(kq)$, and use comparametric equations to transform to the original $f(q)$ function. This can be used to recover the unknown physical quantity $q$.)

Precision of the sensor within its dynamic range can be expressed by an uncertainty function, $u(\rho)$, for each possible sensor output $\rho$. For example, if the sensor’s nonlinear response function is smooth and continuous, the uncertainty function at each quantity in imagespace could be written as:

$$u(\rho) = \left( \frac{\partial f}{\partial \log(q)} \right)^{-1} f^{-1}(\rho)$$
Automatic Gain Control (AGC)

- Loss of information of original signal strength
- Failure to capture outlier elements in array sensing (e.g. saturated image pixels)

Typical D.R. Compositing

- Dynamic Range is covered by multiple exposures
- Predefined exposure settings
- Risk of information loss if signal becomes larger or smaller than the anticipated dynamic range.
- Risk of inefficient D.R. coverage.
- Excessive sacrifices to sample/frame rate due to inflexible exposure settings.

Isolated Dynamic Dynamic Range (D²R) Compositing

- Dynamic response to varying input signal conditions.
- Independent AGC units, each working separately on unconnected sensor exposures.
- Inefficient coverage of the entire dynamic range, due to uncontrolled exposure overlap.

Coupled Dynamic Dynamic Range (D²R) Compositing

- Dynamic response to varying input signal conditions.
- Continuous optimization of dynamic range coverage.
- Efficient coverage of the entire dynamic range, by joint, coupled control of exposure overlap.
- Each exposure balances its own uncertainty forcing functions against those of its neighbours.

Fig. 4. Combining exposure control with HDR compositing: (A) A typical autoexposure camera or sound recorder with automatic gain control (AGC) uses a feedback control to vary the exposure or gain setting. (B) Typical HDR compositing uses a fixed set of gains, or exposures, to try to capture an entire dynamic range. However, these settings cannot anticipate the particular signal that will be sensed—for example, a photo of the sky might produce many pixels which lie in a suboptimal portion of the dynamic range. (C) We add dynamic feedback control of exposures. (D) By cross-linking the exposure control, we dynamically control each exposure according to a cost function based on the response of neighbouring exposures.
Fig. 5. A lattice network of uncertainty calculators forms the basis of the exposure control system, in one embodiment of CD²R. Sensor exposures are controlled in a mutually-coupled manner, to reduce the uncertainty in the resulting HDR composited signal.

In this case, the more unchanging the sensor’s response at a particular signal level, the less precision is available to express the original phenomenon, and thus the less weighting should be placed on that sample/pixel from that exposure, as opposed to the same pixel from other exposures [3]. Uncertainty is thus a measure of the degree of usefulness of each sample/pixel in an HDR composite, based on the particular exposure settings.

Joint, coupled control of exposures requires us to further distinguish between optimization cost functions, to form specialized types of couplings between the exposures. We define: **Supra-uncertainty**, a cost function expressing a penalty on near-saturation and strong samples/pixels in each exposure’s image, $I_e(x,y) \in [\rho_{\min}, \rho_{\max}]$. It will influence the exposure control, forcing it down to a lower amplification:

$$u_{H,e} = \sum_{x,y} I_e(x,y) - \rho_{\min} \left( \frac{\partial f}{\partial \log(q)} \right)^{-1} \left( f^{-1}(I_e(x,y)) \right)$$  \hspace{1cm} (3)

**Infra-uncertainty**, a penalty on weak signals near a sensor’s minimum cut-off level of $q$:

$$u_{L,e} = \sum_{x,y} \rho_{\max} - I_e(x,y) \left( \frac{\partial f}{\partial \log(q)} \right)^{-1} \left( f^{-1}(I_e(x,y)) \right)$$  \hspace{1cm} (4)

We compute a set of uncertainty images, $U_{H,e}(x,y)$ and $U_{L,e}(x,y)$ for each exposure’s sensed image $I_e(x,y)$, formed by separate uncertainty functions computed at each sample or pixel using the previous two equations without summations.
These uncertainty images can be seen in Fig. 6:

\[ U_{H,e}(I_e) = \text{supra-uncertainty image (penalty array expressing pixels whose uncertainty is caused by strong or saturated signals)} \]
\[ U_{L,e}(I_e) = \text{infra-uncertainty image (penalty array expressing pixels whose uncertainty is caused by weak or cut-off signals)} \]

We can then define:

**Cross-uncertainty**, a penalty or cost function, expressing a joint uncertainty caused by gaps in the dynamic range between two adjacent exposures, or poor-quality information in the overlap between two adjacent exposures due to an unchanging sensor response function at those amplitudes:

\[ u_{C,e}(I_e, I_{e+1}) = \sum_{x,y} \min(U_{L,e}(x,y), U_{H,e+1}(x,y)) \] (5)

**Delta-uncertainty**, accounting for the difference in an exposure’s uncertainties caused by its relationships with its different neighbours. Boundary conditions apply to the two end exposures, because they are each solely responsible for the dynamic range at one of their extrema:

\[ \Delta u_e = \begin{cases} 
    u_{C,e}(I_e, I_{e+1}) - u_{H,e}(I_e), & \text{for } e = 1 \\
    u_{C,e}(I_e, I_{e+1}) - u_{C,e-1}(I_{e-1}, I_e), & \text{for } 1 < e < E \\
    u_{L,e}(I_e) - u_{C,e-1}(I_{e-1}, I_e), & \text{for } e = E
\] (6)

We separately compute \( \Delta u_e \) for each exposure, to control each of the \( E \) exposure settings, \( \chi_e \), for the \( e^{th} \) exposure, through a proportional-integral-derivative (PID) controller as in Fig. 2. By accounting for the amount of missing or imprecise information at either end of each exposure’s dynamic range, we can use \( \Delta u_e \) to provide an assessment of whether an exposure needs to shift upward or downward in the overall dynamic range.

The control system is governed by two countervailing influences. Each exposure setting is influenced by a push-pull mechanism from both ends of the dynamic range of each exposure. (This will be compared to a coupled-mass system in a following section.) The feedback loop is then completed with the sensor’s actual signal-capture, giving \( I_e(\chi_e, t) \), resulting from those exposure settings for each exposure.

The basic idea is to compare uncertainty caused by high-valued pixels and low-valued pixels. In two adjacent exposures, this involves the cross-uncertainty of one and the cross-uncertainty of the next, except in the case of the first and last exposures, where we must use an absolute uncertainty caused by saturation in that given exposure alone. See Fig. 5.

These cross-linkages implement the cross-effects illustrated in Fig. 4D. A smooth and continuous control response is made possible by accounting for the floating-point difference in uncertainties, rather than simply incrementing and decrementing integers for the exposure setting.

C. Proportional-Integral-Derivative Control

To complete the exposure control system, we used a PID control, as shown in Fig. 2, to govern the exposure change velocity, as follows:

\[ \nu_e(t) = K_P \cdot \Delta u_e + K_I \cdot \int_0^t \Delta u_e(\tau)d\tau + K_D \cdot \frac{d}{dt}\Delta u_e \] (7)

Finally, the exposure setting \( \chi_e(t) \) is composed of the exposure control velocity \( \nu_e \), as:

\[ \chi_e(t) = \int_0^t \nu_e(\tau)d\tau \] (8)

IV. PHYSICAL ANALOGY

This camera control system can be compared to a coupled-mass system in physics. See Fig. 8. The exposure settings are analogous to the positions of each mass, and the uncertainty cost functions are analogous to forces on each mass.

To see this, first we imagine one force controlling one mass: This is analogous to a single exposure, with only one factor controlling it, according to uncertainty in the image. If we used purely the “I” term of the PID controller, the velocity control function \( \nu \) becomes

\[ \nu_e(t) = K_I \cdot \int_0^t \Delta u_e(\tau)d\tau \] (9)
and therefore, the exposure setting is

$$\chi_c(t) = K_f \cdot \int_0^t \left[ \int_0^t \Delta u_c(t') dt' \right] dt$$

(10)

Differentiating reveals the acceleration of this exposure, which can be compared to a force, $F$, acting on a mass, $m$: $F/m = K_f \cdot \Delta u_c$. Typically, $\Delta u_c(t)$ responds to the negative of $I_x(t)$, which in turn roughly scales with $\chi$ to first order. Therefore, this forcing function behaves similarly to Hooke’s law, representing the physics of a spring: $F = -k\chi$. This would produce simple harmonic motion (see Fig. 9a) in an ideal model, with amplitude $A$, angular frequency $\omega$, and phase $\phi$:

$$\chi(t) = A \sin(\omega t + \phi)$$

(11)

For two exposures, the equivalent mass-spring dynamics are:

$$m \frac{d^2 \chi_1}{dt^2} = -k\chi_1 + k(\chi_2 - \chi_1)$$

(12)

$$m \frac{d^2 \chi_2}{dt^2} = -k\chi_2 + k(\chi_1 - \chi_2)$$

(13)

which leads to two independent normal modes of oscillation [17], in the absence of any control or damping. Additional interactions and normal modes are created with three or more exposures (as in Fig. 9b). However, in this system, by using damping, the “P” term, and the compressive nonlinearity of an image sensor, we can prevent sustained oscillation of the system.

Interestingly, though, we observed wave-like motion of the exposure settings, as they settled within a fraction of a second, which is analogous to the mathematics of high-order mass-spring systems (see Fig. 9c). In the theory of coupled resonators, it can be mathematically shown that as the number of masses and springs approaches infinity, the solution to the equations of motion approaches that of a travelling wave [17]. (This yields the underlying physics behind physical waves such as sound, light, and water waves!)
V. SALIENCE DETECTION TO EFFICIENTLY CONTROL EXPOSURES

Rather than performing the above calculations on every single sensor reading (for example, on every pixel in an image), we search for and selectively compute the most salient sensor readings (pixels) for dynamic range control. When much of the demand for dynamic range is concentrated in a small area of the image, such as with bright lights or arc welding (as in Fig. 1), selective computation of the exposures is especially relevant.

We initially selected pixels in a sparse grid pattern, but unfortunately, missing relevant pixels while jumping by fixed intervals caused problems with stability and accuracy in the control system. Instead, we implemented a salience search algorithm, which progressively expands and contracts regions-of-interest in the image that are most relevant to dynamic range control. These "salience zones" are controlled to gradually change in size and shape according to a real-time gradient descent search. The objective is to maximize the distribution of a histogram of available data, with a penalty on pixels that do not exercise full coverage of the dynamic range. That way, as scenery changes, the system attempts to isolate increasingly relevant and representative data about each exposure’s dynamic range.

Fig. 10 illustrates this system. At each salience zone in the image, a trial run searching dispatcher oscillates between larger and smaller zones at each boundary, testing the improvement or worsening of a dynamic-range histogram cost function. This cost function penalizes extra pixels which are redundant in covering the dynamic range, and penalizes uncovered areas inside the dynamic range. In other words, if we lose relevant information in the salience zone, we gradually shift the salience zone to absorb regions where the composited image has an increasingly broad and diverse histogram.

By implementing this secondary control system, we enable real-time video-based HDR compositing, where the camera may move around over time, and the system tracks relevant features in the image—specifically relevant to measuring the dynamic range and controlling exposures. Fig. 11 shows the time-evolution of a salience zone, automatically selecting a region that includes the widest-possible dynamic range. The salience zone searches and constricts itself automatically, to reduce computational load for the CD²R processing. (This same red rectangle, in the 5th frame, can be seen in Fig. 6. This image is viewed in reverse in Fig. 11.)

The ability to focus on highly salient regions of a video is especially relevant to HDR images of welding. Arc welding consists of a small volume of extremely bright light, surrounded by a large volume of relatively darker light. Traditionally welders use a mask that uniformly darkens the entire spatial field, protecting the eyes but unfortunately preventing the welder from seeing the darker areas around the torch, making it difficult to precisely position the torch. HDR imaging and HDR vision assistance can help address this problem.

Salience zones are relevant in HDR arc welding, thanks to the grouping together of extremely bright pixels and extremely dark pixels. This situation lends itself well to isolation of a salient zone of the image to allow computational speed-up.

VI. IMPLEMENTATION AND RESULTS

The coupled dynamic dynamic-range (CD²R) compositing system was implemented and evaluated using MATLAB code interfaced to a single sensor, cycled between exposure settings for each exposure. This method simplifies the exposure alignment problems when using a multi-sensor configuration.

To provide a worst-case test of the control system due to a very limited dynamic range, we used an off-the-shelf USB camera with 8-bit RGB output images, ranging from 0 to 255 in each of the red, green, and blue channels. For computational efficiency we approximated $u(\rho)$ functions with piecewise linear functions.

Examples of the system’s operation are illustrated in Figs. 6, 7, and 11. Fig. 6 shows examples of exposures in the network of interconnected uncertainty metrics. Fig. 7 shows the time evolution of three exposures (dark, medium, light) under the control of the algorithm. The control system attempts to minimize the mutual uncertainty between the exposures over the combined dynamic range. It attempts to ensure that sufficient information is known about every pixel—for example, that no pixel is saturated in all three input exposures.

Coupled, wave-like motion of multiple exposures, as they adjust themselves over a fraction of a second, can be further seen in our initial paper. [15].

The final result is a composited image. In Fig. 11, we see a progressive improvement from high-noise output to low-noise output as the control system adjusts itself during startup. In the composite image, each pixel is composed of tonal information from at least one of the three corresponding input exposure pixels. Each triplet of input exposure pixels were combined by compensating for the known gain of each exposure, and the nonlinear response function of the camera for each exposure.

The algorithm was applied to a battery of different input scenarios, including high dynamic range visual scenes, using
incandescent lighting and RGB-controllable LED lighting. In some cases, the lighting was pointed directly at the camera, and in others it was directed at elements of the visual scene to create a mix of very bright and dark areas.

We evaluated the system performance of the CD²R system alongside the performance of a traditional static-exposure HDR system. The evaluation method compared the uncertainty metrics as time evolved, comparing a static exposure scenario with broadly-distributed exposure settings, against dynamically-controlled exposure settings (see Table I). The analysis method is presented more fully in our ISM paper [15].

The observations in Table I are under conservative test conditions, where the control system was given a short, insufficient time to settle. The system was repeatedly given only 5 frames of video in which to demonstrate its accuracy, which replicates extremely non-ideal operating conditions, such as rapid camera motion.

Therefore, even under harsh conditions where the dynamic control system is not given sufficient time to fully adapt to the input signal, we still observe an average improvement in measurement uncertainty. The standard deviation indicates variation from one instance to the next.

In a more typical situation (in which the video camera does not move significantly within a few milliseconds), the system takes about 20 video frames (less than a second) to adapt to changing conditions as the input signal (such as ambient lighting) changes or the sensor (such as a camera) moves — as was observed in Fig. 7.

As a camera is moved, the control system might initially have poorer performance than a static exposure system, but it then approaches a more optimal set of exposures which dynamically adapt to the changing conditions.

<table>
<thead>
<tr>
<th>Exposure Control Method</th>
<th>Uncertainty (norm. histogram cost fn.)</th>
<th>Control sig. activity (RMS normalized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STATIC</td>
<td>0.25</td>
<td>0</td>
</tr>
<tr>
<td>DYNAMIC</td>
<td>0.19</td>
<td>0.114</td>
</tr>
<tr>
<td>Improvement</td>
<td>22%</td>
<td>42% var.</td>
</tr>
</tbody>
</table>

TABLE I. CONSERVATIVE PERFORMANCE TEST OF THE CD²R SYSTEM ALONGSIDE THE PERFORMANCE OF A TRADITIONAL STATIC-EXPOSURE HDR SYSTEM.
VII. CONCLUSION

Exposure array control systems, such as the coupled method presented here, have great potential to improve the dynamic range of cameras, audio recorders, and other sensors. Imagine an audio recorder that can be taken out of one’s pocket and used to record an earthquake or a ballistics test, and then the whispers of a mouse in a quiet room, all without adjusting any volume or gain adjustments, and using ordinary sensors and ordinary analog-to-digital converters (ADCs) with a limited inherent dynamic range.

Welding-vision systems, autonomous robot and spacecraft vision systems, acoustic recorders in geology/mining, and scientific cameras and signal recorders, are all applications of CD²R, requiring extreme dynamic ranges with unpredictable, nonstationary signals. See http://eyetap.org/sensing for more information on extreme-dynamic-range and augmented-reality sensing, along with sample code and examples.

ACKNOWLEDGMENTS

This special issue is a collaboration between the 2016 IEEE International Symposium on Multimedia (ISM 2016) and IEEE MultiMedia. This article is an extended version of “Feedback Control System for Exposure Optimization in High-Dynamic-Range Multimedia Sensing,” presented at ISM 2016.

REFERENCES