

Increasing Intensity Resolution on a Single Display using Spatio-Temporal Mixing

Pawan Harish^{*}
Center for Visual Information
Technology
International Institute of
Information Technology
Hyderabad, India
harishpk@research.iiit.ac.in

Parikshit Sakurikar
Center for Visual Information
Technology
International Institute of
Information Technology
Hyderabad, India
parikshit.sakurikar@research.iiit.ac.in

P. J. Narayanan
Center for Visual Information
Technology
International Institute of
Information Technology
Hyderabad, India
pjn@iiit.ac.in

ABSTRACT

Displays have seen much improvements over the years, with enhancements in spatial resolution and vertical refresh, etc., to provide better and smoother visual experiences. Color intensity resolution, however, has not changed much over the past few decades. Most displays are still limited to 8-bits per channel. Simultaneously, much work has gone into capturing high dynamic range images. Mapping these directly to current displays loses information that may be critical to many applications. We present a way to enhance intensity resolution of a given display by mixing intensities over spatial or temporal domains. Our system sacrifices high vertical refresh and spatial resolution in order to gain intensity resolution. We present three ways to mix intensities: spatially, temporally and spatio-temporally. The systems produce in-between-intensities not present on the base display, which are clearly distinguishable by the naked eye. We evaluate our systems using both a camera and human subjects, evaluating whether they scale the intensity resolution and also ensuring that the newly generated intensities follow the display model.

Keywords

High Dynamic Range Display, Spatial and Temporal Color Integration, User Evaluation.

1. INTRODUCTION

High dynamic range images are commonplace today as cameras are capable of capturing 12 or more bits per channel. However, displays to view such images are not easily available; most displays are limited to a low color resolution, typically 8 bits per channel. Though much research is going on in high dynamic range (HDR) imaging, displays to show

such images are still far from being commonplace. Most techniques to render HDR images to low-bit displays involve image enhancement using tone mapping operators. Images rendered using these methods necessarily lose information in the process, though they are visually appealing. Retaining all intensity details calls for displaying in-between-intensity levels. Visualization of medical, scientific, and astronomical data are among the applications that can benefit from such displays. Regular images too can benefit from a mapping that shows more intensity levels as more detail is presented to the viewer.

Spatial resolution of a display has grown over the past few years even though the color resolution has not. Vertical refresh rate has also increased well beyond the limit dictated by the human visual system. In this paper, we present a method to add more perceivable intensities on everyday displays by exploiting their increased spatial and frequency characteristics. We mix pixels spatially and temporally on a single display to generate in-between-intensities not present on the base display. The additional facility sacrifices spatial resolution, refresh frequency, or both in order to gain intensity resolution. We can increase the intensity resolution without loss in human perception by switching between different intensities at each pixel. For example, on a display capable of 120 Hz vertical refresh, two intensities can be mixed to produce a new intensity, which is the average of both, by reducing the effective refresh rate to 60 Hz. A similar argument can be made for spatial mixing by treating a group of, say, 2×2 pixels as a high-resolution pixel. This can be done on ordinary displays without hardware modifications, providing a low-cost way to display more intensity/color levels.

Our method differs significantly from other methods that map HDR images to low-bit displays such as tone mapping, which aim to maintain the visual feel without changing the intensity scale. We increase the intensity resolution of the display at lower refresh rates or spatial resolution. This enables a viewer to see the in-between-intensities not present in the base display and make distinctions based on them. Our methods are built on the ideas of dithering and halftoning techniques which are typically used for the static print medium. We extend these techniques to an inherently dynamic computer display taking advantage of its ability to change across time.

We show results using high frequency CRT displays and evaluate our system both quantitatively and qualitatively.

^{*}Corresponding author

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICVGIP '12, December 16-19, 2012, Mumbai, India
Copyright 2012 ACM 978-1-4503-1660-6/12/12 ...\$15.00.

Quantitatively, we evaluate the method using a high-bit camera sensor which measures the actual intensities on the display to verify if more are seen. We also evaluate our system subjectively, using human subjects. A user study evaluates our technique over various intensity scales, the goal of which is to visibly distinguish in-between-intensities between the two adjacent levels of a basic 8-bit display.

2. RELATED WORK

Capturing and generating high dynamic range content is commonplace in today’s world. Debevec et al. presented a way to capture HDR images using conventional cameras by taking multiple exposures of the scene [5]. In many domains such as medical, astronomical and visualization, the data produced is of high dynamic range. Games simulate HDR using glow and bloom effects [7]. Real-time HDR texture mapping is also shown for computer graphics [4]. A high dynamic range display is thus needed to provide high color resolution required by these applications. Physical methods such as limiting the light scatter using per-pixel reflectors have been shown. These, however, are hard to scale with the number of intensities. High dynamic range can also be achieved using multiple lower-bit images overlaid on top of each other. Seetzen et al. prototype two HDR displays using this approach, one using a projector to back-project on an LCD and another using an array of LEDs behind an LCD [15]. Such a system can provide one additional bit to the display. Multiple projectors can also be used to produce a similar effect. By aligning four projectors pixel-by-pixel, one can add two additional bits to the color resolution of a single projector [14]. These methods, however, cannot scale beyond a few bits as the required number of physical devices keep on doubling.

Directly viewing high dynamic range data on a low dynamic range display is cumbersome. A windowing method is usually employed to select the “best” high range intensities to map to lower bit displays. An alternate approach is to map the high color range to a lower range using local contrast adjustments or tone mapping [9, 8]. This produces pleasing images to a user as the images are adjusted according to the human visual system and are thus more appealing to a human observer. Though visually appealing, these methods do not add more intensities to the display. A comparison of this method to an HDR display is presented in [10].

Image enhancement algorithms such as dithering and half toning rely on human perception to produce intensities over spatial integration of nearby pixels. These are capable of producing better images by propagating quantization errors to neighboring pixels. The methods are typically used to produce visually acceptable images from low-bit inputs in print. Halftoning integrates pixels spatially in two ways using amplitude or frequency modulation [3]. In amplitude modulation, the location of error diffused pixels is static and the intensity is varied by changing the pixel size. Frequency modulation maintains the same pixel size but varies its frequency across the region to produce dense and sparse intensities. We extend these methods to a dynamic display using spatial and temporal integration to add more intensity levels on display than physically available on it.

3. BACKGROUND

In this section, we introduce the display and the human visual system models that our intensity expansion techniques are based upon. Both are elaborate subjects in themselves. Here we only focus on the aspects relevant to our system. Although our methods can be applied to any display, we focus on CRT displays since they provide the high refresh rates needed for our system.

Human Visual System:

The human visual system (HVS) is a complex machine. In the context of this article, our interest lies in the description of perceived intensities over spatial and temporal arrangement of pixels. From a distance, spatial arrangement of pixels produces an average intensity in the viewer’s eye as the eye cannot distinguish between the individual intensities [13]. The critical diameter defines the spatial area over which intensity summation occurs. Intensity summation occurs because a single intensity stimulates a point spread function (PSF) in the eye and multiple PSFs can overlap to create a single stimulus area, if the PSFs lie within the critical diameter, known as a receptive field [6]. This property is exploited by display manufacturers, stacking RGB pixels individually to produce a color pixel, and in imaging methods such as halftoning. We also use this property to increase intensity resolution by sacrificing spatial resolution. Temporally, the visual system can resolve 10 to 12 images per second, perceiving them as individual images [12]. Faster moving images produce the illusion of motion in the observer’s eye due to averaging of intensities over time. The critical flicker frequency (CFF) is defined as the frequency at which an intermittent intensity ceases to flicker and appears as a continuous intensity. The Talbot Plateau law states that above CFF, the subjective averaging of intermittent intensities will produce the same luminosity in the visual system as an objective intensity of the same color and brightness [16]. This is exploited by display techniques such as DLP which produce individual intensities by flipping micro-mirrors at extremely high speed to produce 8-bit intensity resolution. We also exploit this property to temporally mix intensity values to gain intensity resolution. It is also well established that spatial and temporal perception is non-separable in the human visual system [2]. A purely temporal or spatial integration of pixels will thus not suffice. We incorporate a spatial component to the temporal mixing and vice-versa to satisfy this condition. Perceivable intensity values for a given luminance is represented by just noticeable difference (JND) steps in the human visual system. On a display with 1000 : 1 luminance there are about 1000 JND steps [15]. A 10-bit image should thus be visibly distinguishable on such a display to a human observer.

The CRT model:

Many models have been proposed for explaining the behavior of CRT displays, based on the observable values, $F(d)$, and the predicted values $T(d)$ produced by the input digital frame buffer, d . The difference between these two curves evaluates the quality of the model. A well accepted model for CRT behavior is the gamma correction model given as [1]

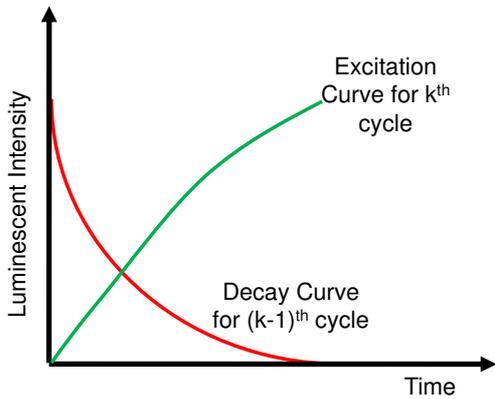


Figure 1: Phosphors excitation and decay curves.

$$T(d) = \left[k_g \left(\frac{d}{2^N - 1} \right) + k_o \right]^\gamma, \quad (1)$$

where N is the number of bits per channel. k_g and k_o are the gain and offset values along with γ as the tunable parameters. Though the model is accepted across all intensities, it may not produce correct results at very low intensities [1]. Our mixing methods should also produce intensities that follow this model for them to produce correct expansion of the color resolution for in-between-intensities. The behavior for decay and excitation of phosphors is also of interest, as it can affect the perceived averaging of pixel values. Phosphors response to excitation and decay is shown in Figure 3 for a single display cycle [11]. Adding the two curves for k and $(k - 1)$ cycles gives the intensity during the excitation of the k^{th} cycle. Clearly the transition holds an average of $(k - 1)$ and k^{th} intensities during excitation, this helps our mixing techniques to sustain a perceived average intensity across multiple display cycles.

4. INTENSITY MIXING TO INCREASE RESOLUTION

We present our mixing techniques in this section. Our methods can be applied in three ways: temporal, spatial and spatio-temporal. Temporal and spatial methods both can add 1 or 2 bits to the displayed image due to the limited frequency and resolution of current displays. Spatio-temporal method is a combination of both approaches and can provide better scaling.

4.1 Temporal Mixing

We sacrifice vertical refresh rate to gain intensity resolution in this approach. Given a display with a vertical refresh rate r , we can divide r into a number of sub-images such that the average of these sub-images, when shown one after the other, produces an intensity not available in the actual display. To display a 0.5 intensity, for example, on a black and white display, the pixel is decomposed into black and white pixels. These are then flipped at high speeds to produce a grey value not present in the black and white display. The intensity is averaged out temporally as described by the Talbot Plateau law [16]. From the halftoning literature, such a mixing corresponds to an amplitude modulation, albeit in

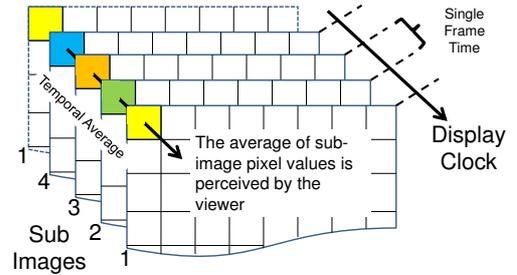


Figure 2: The temporal mixing, averaging over four sub-images, with each sub image being flipped at the vertical refresh of the display.

the temporal domain as opposed to the spatial mixing used in halftoning [3]. The flipping speed is critical and must not be lower than the acceptable frequency perceived by the human visual system; we use a lower limit of 30 Hz. This limits this scheme to a maximum of $r/30$ sub-images. At 120 Hz such a system can show four sub-images, adding two more bits to the already existing display. The flipping must also be synchronized to the display vertical refresh such that the sub-image is changed exactly with change in the display frame, as shown in Figure 2.

4.1.1 Intensity Decomposition

Decomposition of a high-bit pixel into lower-bit pixels is needed in order to produce the desired intensity by averaging the sub-pixels either spatially or temporally. We produce a mapping for each high bit intensity such that it maps to four lower bit intensity values. More sub-images can also be used if the display refresh is high. For a typical display with 120 Hz, displaying more than four frames violates the lowest allowable limit of 30 Hz. The mapping needs to be computed once and can be an off-line process. We decompose by mapping the minimum and maximum values of the input image scale (10-bit) to the minimum and maximum of the display scale (8-bit). The intensities in the high-bit scale can then map to a sum of lower bit intensities by incrementally increasing one intensity per sub-image as shown in Table 1. Intensities can be assigned in multiple ways, for

10-Bit Input Intensity	8-Bit Sub-Images				8-Bit Intensity
	F1	F2	F3	F4	
0	0	0	0	0	0
1	0	0	0	1	0.25
2	0	0	1	1	0.5
...
5	1	1	1	2	1.25
...
759	189	190	190	190	189.75
...

Table 1: Intensity mapping for a 10-bit pixel to four 8-bit pixels

example for intensity 0.5, the sequence [0 0 1 1] or [0 1 0 1] or any rotations of these may be used. For a display with high refresh rate we see no noticeable difference between these two permutations. However, for lower refresh [0 1 0 1] will produce better results as the intensities are more diffused.

The sub-images may be computed off-line if the input is an image but must be computed in real-time for videos. A GPU can be used for this as the process is embarrassingly parallel.

4.1.2 Spatial Component to Temporal Mixing

Multiple low-bit intensity combinations can be used to represent a single high bit intensity value as discussed. The redundancy produces better visual cues than a single combination shown temporally for the same intensity. Using a single mapping for an input high-bit pixel does not produce the best image for large regions of same intensity. Consider, for example, the entire screen filled with a single intensity value. Then if all pixels are allotted the same sub-image sequence, the entire screen will change simultaneously. This results in a noticeable flipping artifact and is particularly visible when the difference between the lowest and highest sub-image intensity results in a large number of JNDs. To avoid this we change the sub-image sequence given to a pixel based on its neighboring pixels. A pixel is allotted a rotation of the sequence that is different from the ones allotted to its direct north, north-west and west neighbors as shown in Algorithm 1, resulting in a spatial averaging of intensities and reducing the JND due to spatial mixing. This kind of spatial averaging relates to halftoned frequency modulation, since the pattern is dispersed over the entire region to create a better visual experience [3].

Algorithm 1 Spatial Component to Temporal Mixing

- 1: Allot sub-image sequence to 1st pixel.
 - 2: **for** 2nd to total number of pixels **do**
 - 3: **if** pixel intensity is same as North, North-West Or West neighbor intensity **then**
 - 4: Rotate sub-image sequence by 1 until sequence for this pixel is different from North, North-West and West.
 - 5: **end if**
 - 6: Save sub-image sequence allotted to this pixel.
 - 7: **end for**
-

4.2 Spatial Mixing

Our spatial mixing technique borrows from halftoning to amplitude modulate intensities over a fixed spatial region [3]. The resolution of the input image is increased to map each pixel to a number of pixels on the screen. Each input pixel (super-pixel) may map to a window of 2×2 , 3×3 etc., such that the eye may not perceive an individual sub-pixel, but treat the group as a single pixel with average of all sub-pixel intensities, as shown in Figure 3(a). When viewed from a distance these sub-pixels average into a single new intensity not available on the base display.

4.2.1 Intensity Decomposition

In our implementation we allot four sub-pixels to a single pixel of the input image. A decomposition similar to decomposition given in Table 1 but applied spatially can be used for this. The mapping of the input intensity to low-bit intensities remains the same. Instead of arranging into sub-images, the intensities are now arranged as a group of four pixels, with each sub-pixel allotted a low-bit intensity.

4.2.2 Temporal Component to Spatial Mixing

Static spatial arrangement of pixels may not provide satisfactory visual cues to perceive an average intensity, even

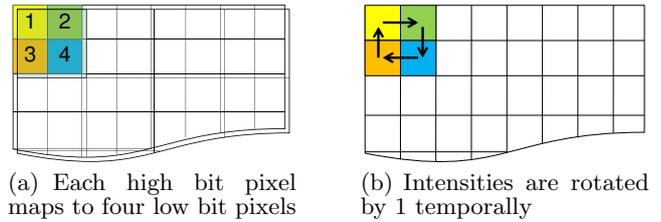


Figure 3: Spatial mixing, each high bit pixel is mapped to 4 sub-pixels that are rotated by one intensity each frame of the display.

when the sequence allotted to each super-pixel differs based on its neighboring super-pixels. This can be attributed to the fact that spatial and temporal perception in HVS cannot be decoupled. Temporal mixing of the spatial data helps to better perceive the average intensity per super-pixel and also to differentiate its identity from the neighboring super-pixels. This can be done by rotating the sub-pixel intensities by one for each display clock cycle as shown in Algorithm 2 and Figure 3(b). This results in better intensity perception per super-pixel and also defines the boundary between super-pixels in the input image. Current generation displays allow high spatial resolution, with up to 3M pixels. These can easily handle sufficiently large input images. However, due to the temporal dependence a high vertical refresh is also needed.

Algorithm 2 Temporal Component to Spatial Mixing

- 1: Allot sub-pixel sequence to 1st super-pixel.
 - 2: Allot sub-pixel sequences to rest of the super-pixels based on their neighbors according to Algorithm 1.
 - 3: **for** All super-pixels **do**
 - 4: Rotate sub-pixel sequence by 1 for each vertical refresh.
 - 5: **end for**
-

4.3 Spatio-Temporal Mixing

The above methods map an input high-bit image to lower bit displays and cannot scale beyond two bits more than the base display intensity resolution due to limitations both in vertical refresh and spatial resolution achievable by current displays. They can be combined into a single system which can scale to a greater intensity resolution. This can be done by decomposing the input intensity into spatial and temporal components. For better performance, there exists a trade-off while decomposing. More bits can be decomposed using temporal mixing if the display refresh rate is high, or alternatively more weightage can be given to spatial mixing if the display provides high resolution.

4.3.1 Intensity Decomposition

Spatial and temporal integration of sub-pixels can be combined to gain higher than 2-bit intensity resolution by decomposing the input image into a number of sub-images. Based on the weightage given to temporal method each of these will have lower intensity resolution than the input image. These sub-images can then be spatially decomposed individually to the even lower display resolution. The sub-images when shown temporally produce an integration of spatial intensity over time, as shown in Figure 4.

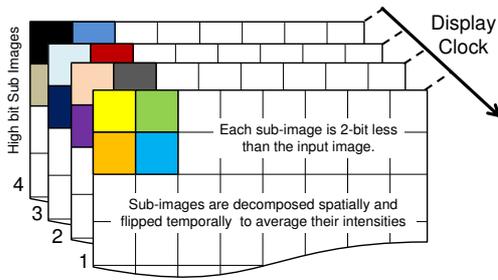


Figure 4: The spatio-temporal mixing, temporally averaging over four high-bit sub-images, with each sub-image decomposed spatially.

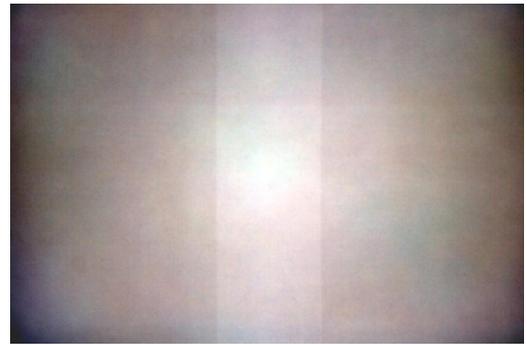
For our implementation, we decompose a high bit pixel into four sub-images, with each sub-image pixel further decomposed into four sub-pixels, corresponding to the display intensity resolution. This can result in a maximum of 4-bit higher intensity resolution than on the base display. A higher bit intensity image, however, will be indistinguishable by the HVS for most current displays due to the presence of more intensities than the allowable number of JND steps for a given luminance value. The method, however, works if the target is a photo sensor. For example a high-bit sensor camera looking at such an interpretation of a high dynamic range input image on a low-bit display can distinguish these intensities. Such a system can thus find application in computer vision and related domains.

5. EXPERIMENTAL EVALUATION

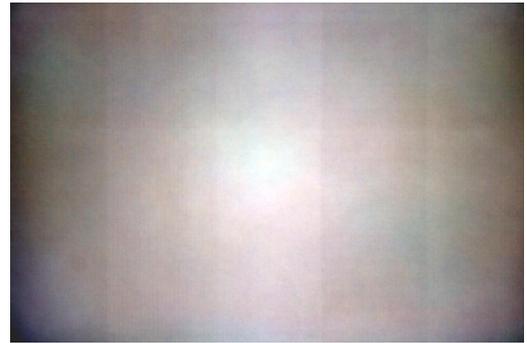
We evaluate our mixing techniques in two ways. A quantitative evaluation using a camera with a high-bit sensor is performed, the aim of which is to see the in-between-intensities and to make sure they follow the CRT gamma curve. A user evaluation is also presented in which we try to evaluate if users can perceive the generated intermediate intensities as expected based on JNDs.

5.1 Camera Based Evaluation

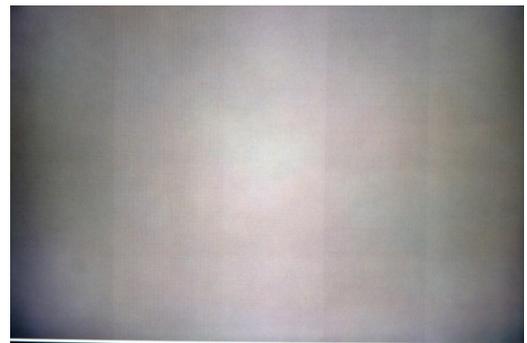
In the camera based evaluation, the intensities generated by our methods are captured using a 12-bit sensor of the Canon 350D DSLR camera. The images are taken in the RAW format to avoid color interpolation. In order to see the in-between-intensities we use a synthetic structured image displaying vertical bands of 10-bit intensities. The image is then mapped on to an 8-bit CRT display using our temporal and spatial mixing methods. The CRT is set at 640×480 pixels running at 160 Hz vertical refresh for the temporal method and 1280×960 spatial resolution running at 75 Hz for the spatial method. The spatio-temporal method uses the same display resolution and refresh as the temporal method. In order to compare perceived intensities we also display the sample 10-bit image directly on the CRT, which decimates it to an 8-bit version and makes the in-between-intensity bands indistinguishable. The image is captured using the camera which is calibrated to the intensity levels of the display. The exposure is set to twice the vertical refresh rate of the display in order for the camera to see the temporal averaging and to allow the intensities to mix on the camera sensor. The input test image has 5 bands of 10-bit values. Decimation to an 8-bit scale creates exactly



(a) Displaying directly using a 8-bit display



(b) Displaying using our temporal mixing



(c) Displaying using our spatial mixing

Figure 5: The 10-bit input image as seen using a 12-bit camera sensor on a CRT display. The difference between the lowest and highest intensity band is 1 level on a 8-bit scale. The images are enhanced to bring out the difference in print.

one intensity difference between the lowest and the highest intensity in the test image. In the 10-bit original image this corresponds to four levels of difference. The intensity bands are arranged out of order for the boundary between intensities to be prominent. Figure 5 shows the input image for the 8-bit, 10-bit temporal and 10-bit spatial versions. The images shown in Figure 5 are contrast enhanced in order for the bands to be distinguishable in print. As seen from these images, the camera can see the 10-bit in-between-intensity levels using our temporal and spatial schemes. The levels, however, are not visible on the display using the direct 8-bit mapping. Figure 6 shows an MRI slice of the knee region dis-

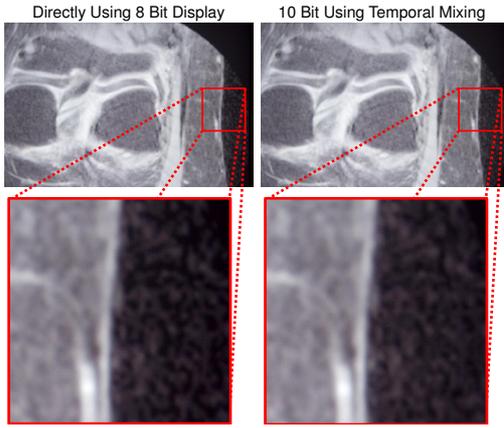


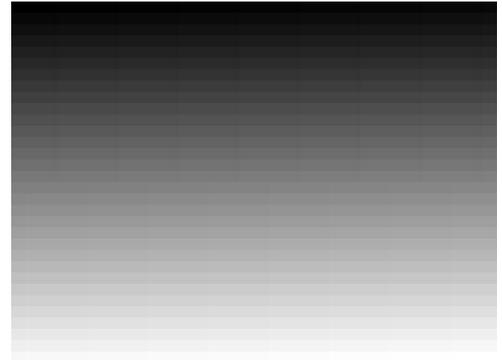
Figure 6: A 10-bit MRI image of the knee as viewed directly on an 8-bit display and using our Temporal method. Clearly more detail is present in the temporal method than on the base 8-bit display.

played using our temporal method in comparison to a direct mapping on the base display. Clearly our method produces more detail than that available on the base display.

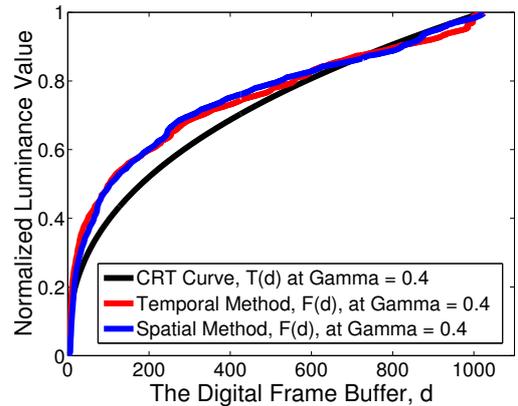
Though the intensities are distinguishable in these images using our methods, they should also follow the expected behavior of the CRT at 10-bit resolution. This is needed for our methods to successfully stretch the intensity resolution to 10-bits. In order to do this, we display a full 1024 intensity image using our methods. The intensities are incremented from left to right and from top to bottom as shown in Figure 7(a). This image is displayed using our methods and is captured by the camera. Since the location for each intensity is known in the image, we can map the observed intensity, $F(d)$, to the input digital frame buffer value, d . This provides us with the luminance for each digital frame buffer value. $F(d)$ can then be compared to the CRT model to see if our methods produce an expected behavior.

Figure 7(b) shows the resulting $F(d)$ for the input d for both temporal and spatial mixing methods. The CRT expected luminance is also plotted as $T(d)$ (Equation 1) for a 10-bit CRT. We see that both the spatial and temporal methods follow the expected behavior projected by $T(d)$. There are, however, some variations as the observed intensities do not account for the camera model, which may alter the intensity due to aperture, exposure, optical and viewing condition dependencies. Overall, however, we see that our methods follow the CRT gamma curve suggesting that both methods produce the in-between-intensities as expected.

We also evaluate the spatio-temporal method for an 11-bit input intensity image to gain more than 2-bit intensity resolution on the display. 11-bits are used here to make sure the 12-bit camera sensor picks up the differences clearly. A 12-bit image could also have been used, however, it would be harder to distinguish the intensities using our camera sensor. The input image is again made up of 5 bands of 11-bit intensities. On the 8-bit scale, there is no difference between the lowest and the highest intensity values. The 11-bit image is first decomposed into two 10-bit images, each of which is then decomposed into 8-bit images using our spatial mixing. The two images are then shown one after the other using temporal mixing. To compare our results we also show the 11-bit image directly on the CRT. The resulting images are



(a) The test image used to evaluate 10-bit intensity resolution



(b) The luminance curves for the 1024 level test image

Figure 7: The test image used and the observed luminance values using our Spatial and Temporal mixing methods for a digital frame buffer range 0–1024. Note the curves closely resemble the CRT gamma curve at the same gamma value. The offsets of the curves are due to the observing camera dependencies.

then captured using the camera. Figure 8 shows the images using 8-bit direct mapping and the 11-bit spatio-temporal mapping. It can be seen that in the 8-bit mapping all intensities merge into a single value, since there is no difference between them on a 8-bit scale. The 11-bit version, however, clearly shows 5 bands of different intensity values, validating our spatio-temporal mixing method. It should be stated that though the bands are visible after enhancement of the camera sensor image, they were very hard to distinguish by the naked eye, as the intensity difference is less than the consecutive JND steps.

To see the behavior of our spatio-temporal mixing on an 11-bit image, we compare the observed intensities, $F(d)$ to the CRT gamma curve, $T(d)$, at 11-bits. A sample image similar to the one shown in Figure 7(a) but with 2048 intensity values is displayed using our method and is captured by the camera. We plot the digital frame buffer vs the luminance in Figure 9. The method clearly follows the overall shape of the gamma curve, however, even with the



(a) Displaying directly using a 8-bit display



(b) Displaying using our spatio-temporal method

Figure 8: The 11-bit band image image as seen using a 12-bit camera sensor on a CRT display using spatio-temporal method. The difference between the lowest and highest intensity band is not visible in the 8-bit scale. The images are enhanced to bring out the difference in print.

unknown camera dependencies, it should be explained why $F(d)$ slightly diverges from $T(d)$. This can be explained by the mixing method: mixing is taking place over eight intensity values in the spatio-temporal system, each of which is gamma related to the previous intensity, thus the difference between intensities becomes less substantial than on a true 11-bit CRT. To display an average intensity, we use the discussed linear decomposition of intensity values. However, on displaying each of the sub-images, all the intensity values undergo a non-linear transformation in accordance with the gamma curve of the display. Thus the sensor now records an average of these transformed intensities, contributing to the slight observed divergence. It should also be noted that this may change on a different display model such as an LCD.

5.2 Human Evaluation

Our experiments for human evaluation are limited to spatial and temporal methods only, as more than 1024 levels cannot be distinguished by a human on current displays due to limited number of JNDs for the given luminance. The purpose of this evaluation is to establish that in-between-intensities can be perceived by a human observer. In order to do this, we show 10-bit images with varying number of bands (5 – 7) with 1 – 2 levels of difference between the lowest and the highest intensity bands on an 8-bit scale. The

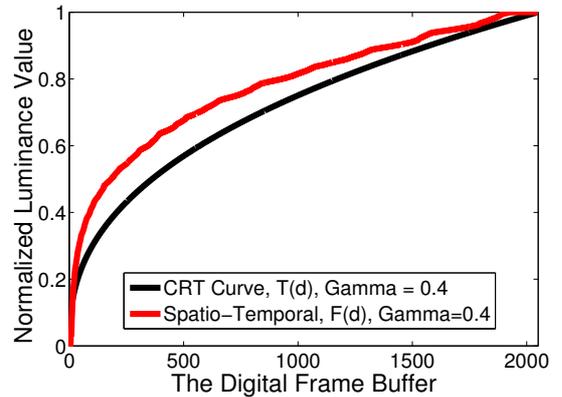


Figure 9: Spatio-Temporal method for a digital frame buffer range 0 – 2048. The curve closely resembles the CRT gamma curve at the same gamma value.

test consists of images simultaneously shown spatially and temporally on two adjacent CRTs. 12 selected images are randomly shown using direct 8-bit mapping to none, either or both the displays. Table 2 defines the experimental input used for human evaluation. The user is asked to report the number of bands visible in both displays. The images span the entire intensity range and random RGB combinations.

Task No.	Intensity Range	No. Of 10-bit bands	Color Channel	Display Bits	
				Temporal/No mixing	Spatial/No mixing
1	508-512	5	RGB	10	8
2	340-345	6	G	8	10
3	676-682	7	RGB	8	10
4	340-344	5	GB	8	8
5	844-849	6	RGB	10	8
6	508-513	6	RG	8	10
7	676-680	5	RGB	8	8
8	340-346	7	RGB	10	8
9	844-848	5	RGB	8	10
10	676-681	6	RB	10	10
11	844-850	7	RGB	10	8
12	508-514	7	G	10	8

Table 2: Experimental images used for human based evaluation.

Thirty test subjects were tested for 12 images each. The results for 8 and 10 bit spatial and temporal methods are summarized in Figure 10. The mean and maximum deviation from actual number of bands in the 8-bit and 10-bit images are reported for each task, with **T** specifying temporal method and **S** the spatial method. There are more errors in distinguishing 10-bit images, however, errors were also seen in the 8-bit versions. The mean error is a good indicator for the human perception of 10-bit images. We see that the mean error never exceeds two bands with the exception of task number 10. This was a particularly hard task with only red and blue channels making up the image. We found the distinction of bands in red and blue to be the hardest due to fewer transmitters and receptors for these channels in the CRT display and the human eye respectively. Green channel images were the easiest to distinguish, as seen in task T12. Over the entire experiment we found that spatial mixing produced slightly more errors than the temporal method.

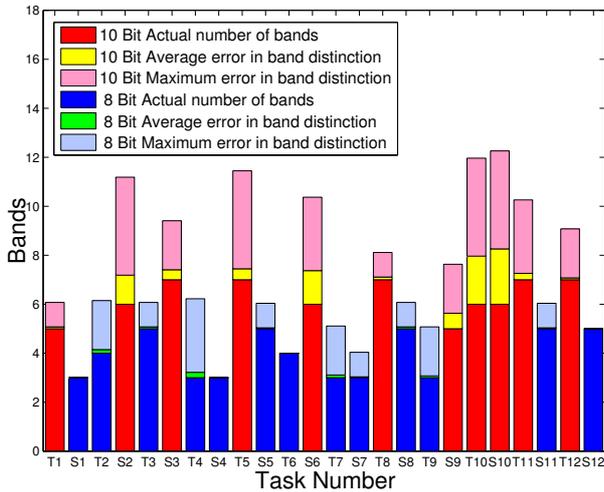


Figure 10: Results of the human experiment. The left column for each task represents the temporal (T) mixing while right represents spatial (S) mixing. The graph shows actual bands in the image, the average error per task and the maximum error per task. For example, for task no S6, it shows it is a 10-bit task with 6 bands, average error of 1.4 over all subjects and the maximum reported error of 3 bands.

This may be due to the high resolution and low refresh rate used in this method. For the 10-bit images, we see a mean error of 0.8 bands with standard deviation of 0.76 over all subjects. Overall, our methods produced a low error for both spatial and temporal approaches and were capable of displaying more intensities to a human observer than those available on the base display.

5.3 Limitations of Our Methods

The main limitation of our methods for HDR displays is that they do not increase the upper intensity limit. Only the display intensity resolution is increased keeping the same brightness range. This can result in our method producing more intensity levels than the allowable JND steps. Since the mixing occurs in the eye and not on the display, an increase of more than two bits of intensity cannot be verified by HVS. Our methods are also limited by the vertical refresh and the spatial resolution of the target display. Sacrificing the vertical refresh can produce stuttering artifacts. Lowering the spatial resolution can also cause loss in image detail.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we presented three methods to map HDR images to a display of low-bit resolution. Our methods produce in-between-intensities that are capable of visual distinction as shown in our experiments both by human observers and by a camera sensor. Our methods can also be used for HDR videos if the video is of less than 30 frames per second. We have experimented with these by decomposing each frame into sub-images on the fly using GPUs. Our methods can visually enhance images as required by many visualization applications. The methods are technology independent and can be applied even on a high-bit per

channel display to increase intensity resolution further. We would like to explore our methods for regular images and compare them with other mapping methods such as tone mapping to see if they produce better visual image quality. Advent in display contrast, resolution and vertical refresh has an increasing trend. Our methods when combined with such high contrast, resolution and vertical refresh displays can truly reproduce HDR images at enhanced intensity resolutions.

7. REFERENCES

- [1] P. Bodrogi, K. Muray, and J. Schanda. Accurate colorimetric calibration of CRT monitors. In *SID*, page 455, 1995.
- [2] K. R. Boff, L. Kaufman, and J. P. Thomas. *Handbook of perception and human performance*, volume 1, pages 13–1 to 13–62. 1986.
- [3] K. K. Charles Hains, Shen-Ge Wang. Chapter 6, digital color halftones. In *Digital Color Imaging Handbook*. Xerox Corporation, 2003.
- [4] J. Cohen, C. Tchou, T. Hawkins, and P. E. Debevec. Real-time high dynamic range texture mapping. In *Eurographics Workshop on Rendering Techniques*, pages 313–320. Springer-Verlag, 2001.
- [5] P. E. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH 1997*, pages 369–378, 1997.
- [6] E. F. Helga Kolb, Ralph Nelson and B. Jones, editors. *WebVision: The Organization of the Retina and Visual System*. University of Utah, 1995.
- [7] G. James and J. O’Rorke. Real-time glow. In *GPU Gems*, pages 343–362. 2004.
- [8] A. G. Kirk and J. F. O’Brien. Perceptually based tone mapping for low-light conditions. *ACM Trans. Graph.*, 30(4):42:1–42:10, Aug. 2011.
- [9] G. Krawczyk, K. Myszkowski, and H. P. Seidel. Lightness perception in tone reproduction for high dynamic range images. In *Eurographics 2005*, volume 24, pages 635–645, 2005.
- [10] P. Ledda, A. Chalmers, T. Troscianko, and H. Seetzen. Evaluation of tone mapping operators using a high dynamic range display. In *ACM Trans. Graph.*, pages 640–648, 2005.
- [11] A. Pfahnl. Properties of fast-decay cathode ray tube phosphorus. In *Bell System Technical Journal*, v42, pages 181–201, 1963.
- [12] P. Read, M. Meyer, and G. Group. *Restoration of Motion Picture Film*. Butterworth-Heinemann Series in Conservation and Museology. Butterworth-Heinemann, 2000.
- [13] D. Salomon. *Data Compression: The Complete Reference*. Second edition, 2000.
- [14] M. Sato, M. Inoue, M. Kasuga, and N. Hashimoto. Hdr display with a composite response function. In *Joint Virtual Reality Conference of EuroVR - EGVE posters*, JVRC 2011 posters, pages 86–88, 2011.
- [15] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcovs. High dynamic range display systems. In *ACM Trans. Graph.*, pages 760–768, 2004.
- [16] N. J. Wade. *A Natural History of Vision*. MIT Press, 1998.