Spatio-Temporal Mixing to Increase Intensity Resolution on a Single Display

Pawan Harish, Parikshit Sakurikar and P. J. Narayanan

ABSTRACT

Mapping high dynamic range images 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. 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.

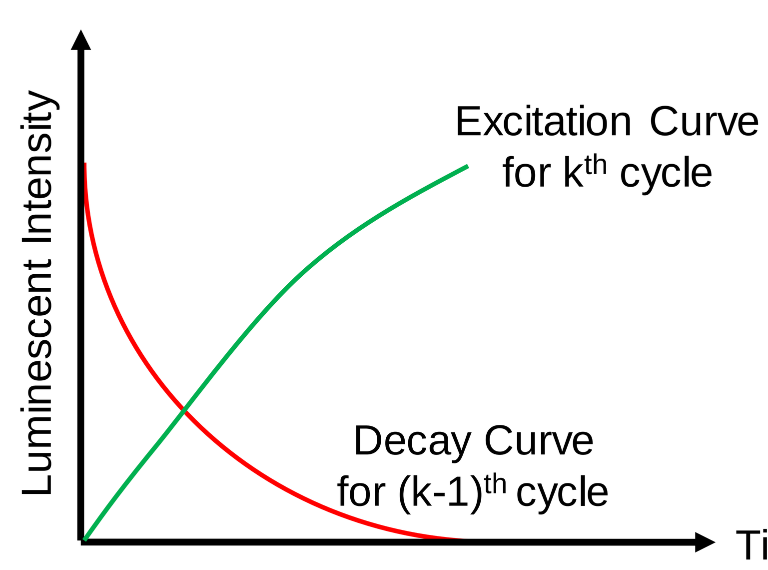
BACKGROUND

The Human Visual System (HVS)

- From a distance, spatial arrangement of pixels produces an average intensity in the viewer's eye.
- Temporally, the visual system can resolve 10 to 12 images per second. Faster moving images produce the illusion of motion in the observer's eye due to averaging over time.
- Spatial and temporal perception is non-separable.
- Perceivable intensity values for a given luminance is represented by just noticeable difference (JND) steps.

The CRT Display 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. A well accepted model for CRT behavior is the gamma correction model, as stated by the equation on the right.



CRT Excitation Curves

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

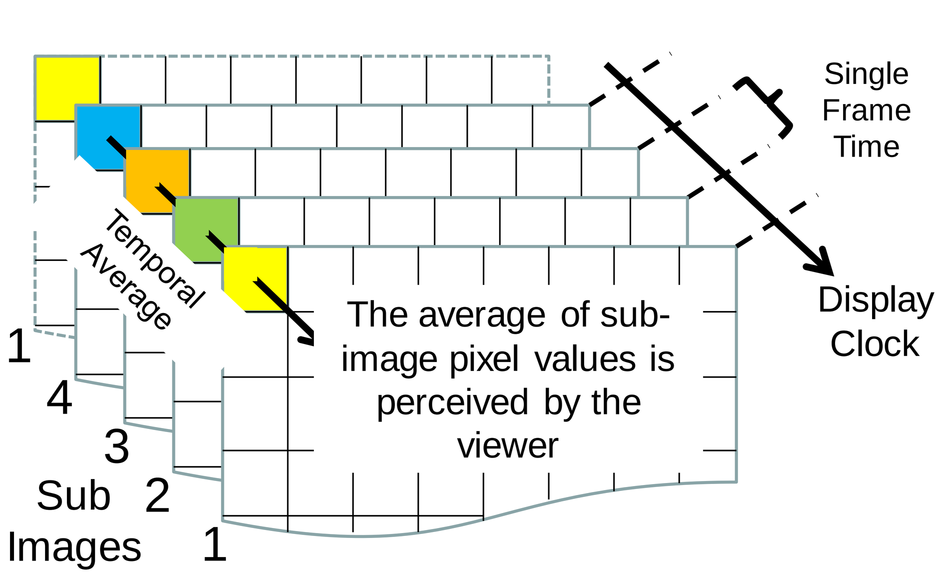
DECOMPOSING INPUT INTO LOW BIT PIXELS

10-Bit Input Intensity	8-Bit Sub-Images F1	F2	F3	F4	8 Bit Intensity
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
...	....				....

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

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. The minimum and maximum values of the input image scale (10-bit) maps to the minimum and maximum of the display scale (8-bit). The intensities in the high-bit scale map to a sum of lower bit intensities by incrementally increasing one intensity per sub-image as shown in the table.

TEMPORAL MIXING

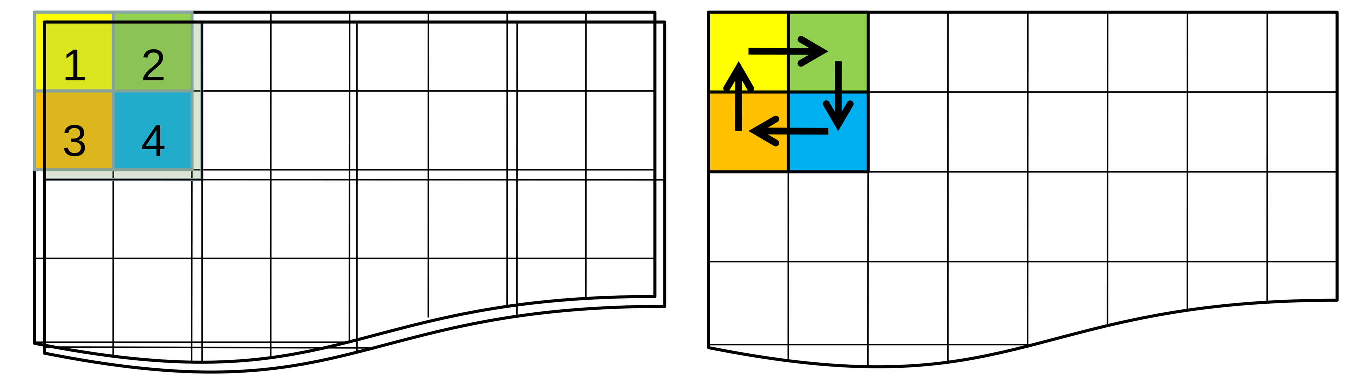


We sacrifice vertical refresh rate to gain intensity resolution in this approach. Given a display with a vertical refresh rate r, we can divide r into any number of sub-images such that the average of these sub-images produce an intensity not available in the actual display when flipped temporally.

Spatial Component to Temporal Mixing

If adjacent pixels in the input image have same intensity values, they map to the same sequence of sub-images, resulting in a noticeable flipping artifact. This 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 resulting in a spatial averaging of intensities.

SPATIAL MIXING

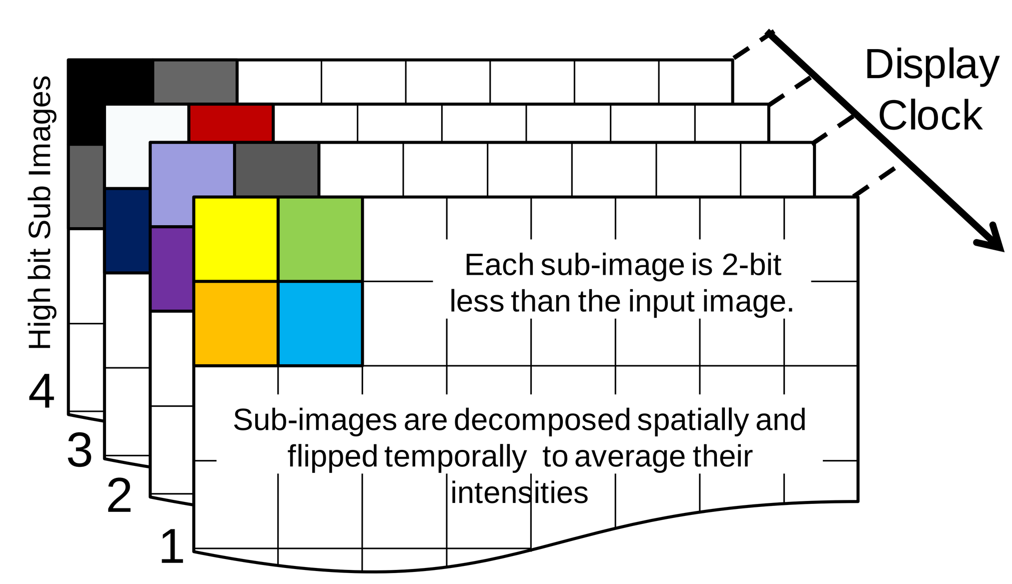


Each pixel may map to a window of 2x2, 3x3 etc. such that the eye may not perceive individual sub-pixels, but treat the group as a single pixel with average of all sub-pixel intensities. When viewed from a distance these sub-pixels average into a single new intensity not available on the base display.

Temporal Component to Spatial Mixing

Temporal mixing of the spatial data must also be present to 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.

SPATIO-TEMPORAL MIXING

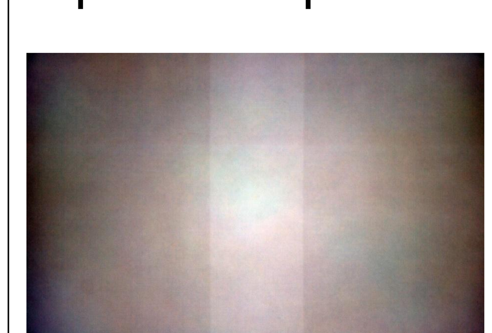


Spatial and temporal integration of sub-pixels can be combined to gain higher than 2-bit intensity resolution by first decomposing the input image into 4 sub-images, each of 2-bit less resolution. These sub-images can then be decomposed individually using the proposed spatial mixing method to a 2-bit lower display resolution. A higher bit intensity image, however, will be indistinguishable by the HVS on most current displays due to presence of more intensities than the allowable number of JND steps for a given luminance value. The method works if the target is a photo sensor.

EXPERIMENTAL EVALUATION

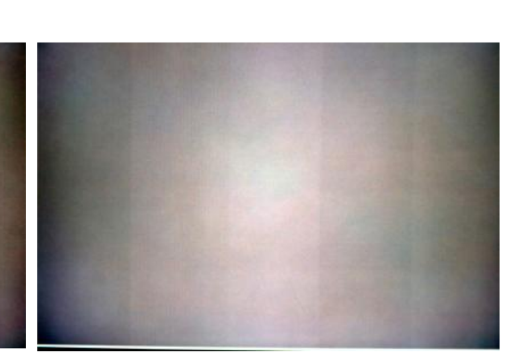
Camera Based Evaluation

The intensities generated by our methods are captured using a 12-bit sensor of the Canon 350D DSLR camera. In order to see the in-between-intensities we use a synthetic structured images displaying vertical bands of 10 and 11 bit consecutive intensities. The bands are arranged out order for the boundary between intensities to be prominent. Images are then mapped to an 8-bit CRT display using our mixing methods and also shown directly using the same. The CRT is set at 640x480@160Hz for the temporal and spatio-temporal methods and at 1280x960@75Hz for the spatial method.

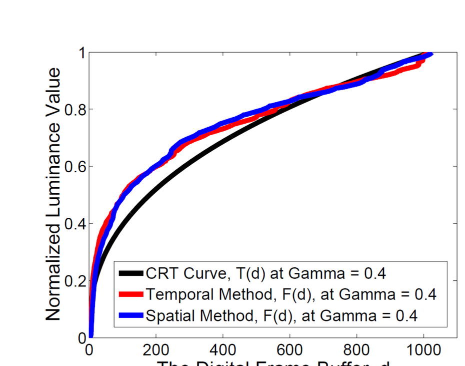


10 bit On 8 bit Directly

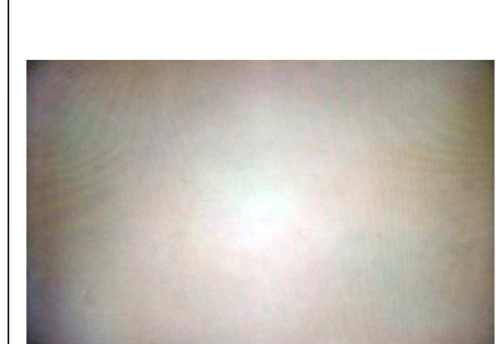
10 bit On 8 bit Temporally



10 bit On 8 bit Spatially

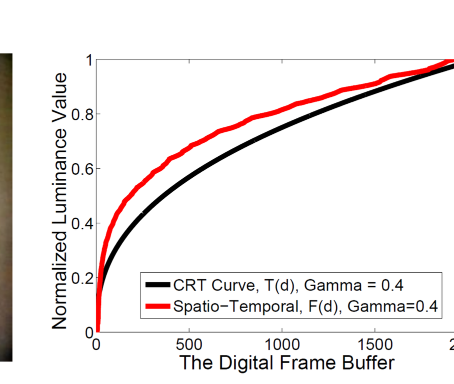


Luminance Curves for Spatial Temporal Methods

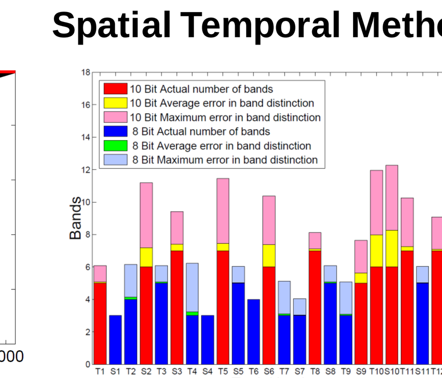


11 bit On 8 bit Directly

11 bit On 8 bit Spatio-Temporally



Luminance Curve for Spatio-Temporal Method



Mean and Max Error for Human Evaluation

Human Based Evaluation

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 a 8-bit scale. Twelve random RGB combination images are simultaneously shown spatially and temporally on 2 adjacent CRTs. The images are randomly shown using direct 8-bit mapping to none, either or both the displays. Thirty users are asked to report the number of bands visible in both displays.

Center for Visual Information Technology. International Institute of Information Technology-Hyderabad, India